Exploring User Authentication Behavior in Cybersecurity Systems: A Data Mining Approach Using K-Means Clustering

Chendri Irawan Satrio Nugroho[1,*], Erland Inkiriwang[2]

[1,2]Information Technology Department, Institut Teknologi Tangerang Selatan, Banten, Indonesia

ABSTRACT

User authentication is a cornerstone of modern cybersecurity, yet user behavior remains a significant and often unpredictable challenge to system integrity. Despite the implementation of complex password policies, user friction can lead to security vulnerabilities and poor user experiences. This paper explores patterns in user authentication behavior through a data mining approach, applying K-Means clustering to a dataset of 7,000 anonymized user login sessions. The analysis incorporates a range of behavioral and environmental features, including password length, login attempts, typing speed, and the use of special characters. The results of the clustering analysis successfully segmented the user base into two distinct and meaningful archetypes: the "Struggling User" and the "Efficient User." The "Struggling User" cluster was characterized by a high mean number of login attempts (7.97), a greater likelihood of having Caps Lock enabled, and a tendency to use special characters. In contrast, the "Efficient User" cluster demonstrated a low mean number of login attempts (3.00) and less complexity in their credentials. Critically, the analysis found no significant difference in password length or typing speed between the two groups, suggesting that authentication friction is more closely linked to cognitive load and input errors than to general user proficiency. These findings provide empirical evidence that stringent password complexity requirements can inadvertently degrade usability, leading to repeated authentication failures for a significant portion of users. This has direct implications for both cybersecurity policy and cyberlaw, challenging the efficacy of one-size-fits-all security mandates. This research advocates for the adoption of more adaptive, user-centric authentication systems and informs the legal definition of "reasonable" security by highlighting user experience as an essential component of a robust and effective security framework.

Keywords Authentication, Clustering, Cyberlaw, Cybersecurity, User Behavior

Introduction

A cornerstone of effective cybersecurity practices is the implementation of authentication systems, which play a critical role in verifying the identities of users and devices before granting access to sensitive resources. Traditional username-password combinations are no longer sufficient to safeguard against unauthorized access due to the increasing sophistication of cyberattacks. To address these vulnerabilities, organizations are increasingly adopting multi-factor authentication (MFA) systems and Single Sign-On (SSO) solutions. MFA requires users to present two or more verification factors to gain access to an account or system, thus significantly enhancing security [1]. SSO streamlines the authentication process by allowing users to access multiple applications with a single set of login credentials, thereby reducing the burden of remembering multiple passwords while also improving productivity [1].

The proliferation of mobile devices and cloud applications has further complicated the landscape of cybersecurity. As organizations shift towards

Submitted 16 July 2025
Accepted 2 August 2025
Published 1 September 2025

*Corresponding author
Chendri Irawan Satrio Nugroho,
chendri@itts.ac.id

Additional Information and Declarations can be found on page 208

DOI: 10.63913/jcl.v1i3.11

How to cite this article: C. I. S. Nugroho and E. Inkiriwang, "Exploring User Authentication Behavior in Cybersecurity Systems: A Data Mining Approach Using K-Means Clustering," *J. Cyber. Law.*, vol. 1, no. 3, pp. 190-211, 2025.

cloud computing, the need for secure authentication protocols becomes even more pressing. Current trends emphasize the importance of incorporating advanced technologies such as machine learning and artificial intelligence in developing sophisticated security solutions that dynamically adapt to evolving threats [2][3]. For instance, AI can be harnessed for real-time threat detection and predictive analytics, enhancing the capabilities of intrusion detection systems (IDS) and enabling proactive instead of reactive security measures [2] [4].

Biometric authentication, leveraging physical and behavioral traits for verification, is emerging as a promising alternative to traditional authentication methods. This approach not only enhances security but also improves user convenience by reducing reliance on memorized passwords [5]. Biometric systems, including fingerprint recognition and facial recognition, are increasingly being integrated into access management frameworks to bolster security in environments that require stringent access controls, such as financial institutions and healthcare facilities [6]. The use of behavioral biometrics, which analyzes users' unique patterns of interaction with devices, is also gaining traction as a means to identify potentially fraudulent activity in real time [2][5].

Moreover, cultivating a cybersecurity-conscious culture within organizations is crucial for ensuring the effectiveness of technological measures. This culture fosters a shared understanding of cybersecurity's significance and encourages proactive behavior among users, thereby reducing the likelihood of human errors that can lead to security breaches [7]. Educational initiatives, training programs, and awareness campaigns are essential tools for promoting a security-aware mindset among employees, further bolstering organizational defenses [7].

The challenges posed by cyber threats are not static, they continue to evolve, necessitating a dynamic approach to cybersecurity management. Continuous research and development are imperative to stay ahead of emerging threats, particularly in high-risk sectors such as finance, healthcare, and education. The importance of integrating emerging technologies within established cybersecurity frameworks cannot be overstated, as organizations must remain agile and ready to respond to new attack vectors as they arise [6][3].

The evolving landscape of cybersecurity necessitates a deeper understanding of user behavior to improve authentication systems and minimize breaches. As cybersecurity becomes increasingly important amid growing threats, the focus on user interaction with authentication mechanisms reveals critical factors that influence security outcomes [8]. A significant body of research indicates that user behavior and emotional reactions substantially affect the efficacy of cybersecurity measures, thus emphasizing the need for user-centric approaches to information security [9][10].

User behavior is a pivotal element that shapes the effectiveness of authentication systems. Studies reveal that many users engage in risky online behaviors, such as reusing passwords or ignoring security warnings, which can lead to vulnerabilities [8][10]. For instance, [8] emphasize the necessity for improved education around information security behaviors among smartphone users, indicating that better awareness can positively alter user behaviors, thereby enhancing overall cybersecurity. Similarly, research conducted in India highlights trends underscoring the need for further exploration of cybersecurity behavior, particularly in mobile environments where users may be unaware of

potential risks. Addressing user behavior is fundamental for developing strategies to bolster security measures effectively.

Moreover, emotional responses to cybersecurity notifications play a crucial role in user decision-making and behavior. Research by [9] indicates that emotional reactions significantly influence how users respond to security warnings. The negative reactions observed in response to cybersecurity alerts suggest that users may overlook or dismiss notifications, leading to poor security choices. Understanding these emotional transactions can help security designers craft better alerts and notifications, encouraging safer browsing practices.

To bridge the gap between user behavior and authentication systems, innovative techniques are being developed that leverage behavioral biometrics—such as keystroke dynamics. Keystroke dynamics rely on the unique patterns of how individuals type, offering a user-friendly second factor for authentication beyond traditional passwords [11][12]. This approach enhances security by making it more difficult for an intruder to mimic a user while also reducing the burden on users who struggle to remember multiple passwords [11][13]. The automatic and implicit nature of keystroke dynamics leads to continuous authentication, which can assess user identity throughout a session, ultimately contributing to improved security outcomes [14].

Additionally, the design of authentication systems must consider how users create and manage authentication data. Alhusain et al [15] argue that many users struggle to comprehend security policies, resulting in weak security practices that expose systems to breaches. By providing clearer guidance and framing security rules understandably, organizations can enhance users' ability to create effective passwords and adhere to best practices, ultimately reducing the risk of breaches [15][16].

Furthermore, the development of adaptive authentication systems that learn from user behavior is a burgeoning area within cybersecurity. Research by Islam et al [17]. emphasizes the importance of constructing behavioral profiles that reflect authentic behavior patterns, which can greatly enhance the security of authentication processes. Such systems can provide dynamic responses to user actions, determining when to request additional authentication based on detected behavioral anomalies, thus balancing user convenience with heightened security [18][19].

A noteworthy trend in addressing user behavior in cybersecurity is the integration of reinforcement learning with behavioral biometrics for continuous authentication. This approach leverages real-time data to authenticate users based on their interactions with devices seamlessly [19]. Moreover, embedding context-awareness into these systems offers a heightened level of security that adapts to how users typically engage with their devices [20][19]. This insight underlines the importance of understanding user behavior in developing authentication systems that protect digital assets and align with the natural usage patterns of individuals.

Literature Review

Cybersecurity and Authentication Mechanisms

One prominent category of authentication mechanisms is behavioral biometrics, which includes methods such as keystroke dynamics and touch gesture authentication. These systems analyze users’ unique behavioral patterns to

establish their identity. For instance, Mahfouz et al [14]. how keystroke dynamics utilize the rhythm and timing of a user's typing to authenticate them, offering a passive yet effective approach to security. Behavioral biometric authentication can provide continuous validation throughout a session, thus protecting against unauthorized access even after initial authentication has been completed [21].

A particularly innovative approach in this field is the integration of physiological and behavioral features for identification through various interactions. Wong et al [22] describe ArtiLock, a system that capitalizes on articulation patterns during phone interactions, presenting a user-friendly identification method that requires minimal training data from users while maintaining high security through difficult-to-replicate behavioral characteristics.

Another progression in authentication is the rise of continuous authentication systems, which monitor user interactions over time to identify deviations from established behavioral patterns. Valero et al [23] highlight how continuous authentication enables real-time assessments of behavior to detect spoofing and improve security dynamically, utilizing various metrics including typing speed and touch pressure, while ensuring a non-intrusive verification process.

The advent of multi-factor authentication (MFA) has also transformed the authentication landscape. MFA combines two or more independent credentials that can include something the user knows (knowledge-based, like a password), something the user has (physical tokens), and something the user is (biometric features). Qureshi and Kale [24] explore risk-based authentication systems that dynamically adjust the required authentication factors based on the user's profile and context, thus enhancing security while optimizing user convenience. This method becomes particularly essential in environments with diverse security needs, such as financial institutions.

Biometric authentication methods utilize unique biological traits to identify users, thus offering a high level of security. Common forms of biometric authentication include fingerprint recognition, facial recognition, and iris scanning. Liang et al [18] categorize these techniques under physiological biometrics, underscoring their efficacy in authenticating users through inherent human characteristics. The integration of biometric systems into organizational security protocols has been noted to reduce the overhead of password management and risks associated with forgotten passwords, though they still require user cooperation and participation [18].

Beyond these strategies, novel techniques involving channel state information (CSI) offer a non-intrusive means of user authentication. CSI leverages the unique ways in which individuals interact with their devices to establish identity, minimizing the reliance on traditional credentials. Wang et al. [25] discuss the potential of CSI-based systems to provide robust authentication without requiring active user participation, aligning with the growing need for seamless security solutions that do not disrupt the user experience.

Additionally, recent advancements include the adoption of innovative techniques such as optical spectrum for user authentication, which combines traditional OTP methods with new embedded technologies [26]. This approach showcases the ongoing evolution of authentication systems, highlighting innovations designed to enhance usability while ensuring security.

User Behavior in Authentication Systems

Understanding user behavior in authentication systems is critical to developing secure, efficient, and user-friendly authentication mechanisms. A comprehensive review of previous studies reveals several insights into how user behaviors, login attempts, and various factors influence the effectiveness of security measures. Research has demonstrated that users' login behaviors often reflect their patterns of risk and security awareness. For instance, Stylios et al [21] explored the role of behavioral biometrics, particularly keystroke dynamics, in continuous authentication. Their study highlights how traditional entry point authentication methods merely validate a user's identity at the beginning of a session, thereby exposing them to various security threats post-login. The research emphasizes a behavioral approach to authentication, proposing that continuous monitoring of user behavior could enhance security by adapting dynamically to potential threats.

In a complementary investigation, Wiefling et al [27] analyzed real-world data from a large-scale online service, stratifying user login attempts based on the frequency of failures. Their findings indicated that users who made frequent login attempts were often targeted for attacks, suggesting that login history can be a significant factor in identifying potential threats. The study underscores the necessity for tailored security responses based on individual user behaviors, particularly for those who exhibit patterns indicating possible account compromise.

Engaging with the psychological aspects of user behavior, Ehatisham-Ul-Haq et al [28] examined mobile device interactions, focusing on implicit authentication techniques that rely on recognizing user activity patterns through mobile sensing. Their results suggest that passive monitoring of user interactions can inform systems about normative behaviors, potentially preventing unauthorized access without significantly impacting user convenience. Their work highlights a fundamental shift towards accommodating user tendencies and inclinations in security protocols, facilitating a more adaptive approach to user authentication.

However, user decisions regarding password selection and management reveal fundamental security vulnerabilities. Research [20] pointed out that factors like password reuse and the tendency to create easily guessable passwords are prevalent among users, leading to significant security risks. These findings highlight the disconnect between users' perceived security practices and actual behaviors, signaling a need for educational initiatives that guide users toward adopting stronger security postures. In this context, the study advocates for integrating security practices that align with user behavior, thus simplifying security compliance while simultaneously enhancing defense mechanisms.

Furthermore, a pertinent study by [29] discussed the vulnerabilities inherent in SMS-based authentication systems, which are often susceptible to man-in-the-middle attacks where users inadvertently expose their verification codes to malicious actors. The study emphasizes how user behavior—specifically careless handling of communication channels—can severely impact the reliability of authentication systems, highlighting the need for continuous user education about secure practices when leveraging authentication technologies.

Another interesting angle comes from Lee et al [30] who investigated how user behavior, particularly interactions with touch screens, can be safeguarded against common security threats like smudge attacks and shoulder surfing. Their work showcases the ongoing challenge of crafting robust biometric authentication methods that account for user behavior variability, thus fortifying

authentication measures that could be susceptible to observational attacks.

Moreover, Blocki and Zhang [31] examined the nuances of login patterns among users, revealing that even legitimate users sometimes exhibit irregular behaviors that deviate from expected norms, such as logging in from different geographical locations. Their research posits that authentication systems must adapt to the nuances of human behavior while maintaining high security standards. Such adaptability can play a critical role in reducing security fatigue among users who often find conventional multi-factor authentication burdensome.

Finally, as research [32] pointed out, the struggle between creating secure yet user-friendly authentication methods remains a critical barrier. Their research indicates that while users may aim for stronger passwords, the necessity of remembering complex combinations often leads them to unsafe practices such as writing them down or reusing passwords across multiple platforms. These findings highlight the need for innovative solutions that foster user compliance while minimizing the security risk associated with poor password management.

Data Mining in Cybersecurity

Data mining plays an essential role in the field of cybersecurity by identifying patterns in security-related data. Its application is particularly relevant in two key areas: clustering and association rules. These methodologies facilitate the analysis, detection, and prevention of potential threats, contributing to more robust security measures.

Clustering as a data mining technique groups diverse data points based on their similarities, helping to identify patterns, trends, and potential anomalies. In a cybersecurity context, clustering algorithms effectively segment types of user behaviors or network traffic. For instance, study [33] presented a secure cluster management approach utilizing big data analytics to optimize control planes in software-defined networks (SDNs). The study underscored the clustering of security-related data to improve the efficiency and performance of security applications running within the network. Effective clustering enables cybersecurity professionals to manage security incidents by identifying suspicious groups of behaviors that deviate from the norm, signaling potential threats or intrusions.

Furthermore, the application of clustering techniques extends to the healthcare domain, where research [34] emphasized the importance of preserving security and privacy in large-scale data systems. Their work illustrates how clustering can help compartmentalize vast amounts of healthcare data while ensuring that sensitive information remains protected. By clustering data according to similarities and access requirements, organizations can streamline security protocols while reducing the risk of data breaches—an approach that can be translated to other fields requiring stringent security measures.

In addition to clustering, association rule mining is critical for identifying relationships between various security factors, enabling foresight into potential vulnerabilities. This methodology helps organizations understand the likelihood of an event occurring based on a set of known behaviors or conditions. For example, study [35] highlighted how web usage mining, which incorporates association rule mining, can be applied to predict users' needs based on their access patterns. By analyzing access logs using association rules, cybersecurity teams can uncover hidden relationships between user behaviors

and security incidents, allowing them to preemptively mitigate risks.

Moreover, Jin and Lin [36] discussed the importance of web log analysis based on data mining as a method for security assessment. Their findings outlined how analyzing web logs through association rule mining can help determine anomalous behaviors that could indicate malicious activities, thus providing a proactive security measure. The frameworks established not only contribute to understanding existing threats but also pave the way for developing responses to emerging security issues.

Another innovative application lies in the use of K-nearest neighbor (KNN) algorithms for clustering user behavior in online environments, as demonstrated by Syadzali et al [37]. This methodology assists in analyzing patterns of customer behavior in online crowdfunding systems, showing that similar approaches can be applied to detect anomalies in user behavior, contributing to security assessments in digital platforms. By continuously monitoring and clustering user data, organizations can quickly identify deviations that may indicate security threats.

Moreover, the concept of integrating multiple viewpoints in identifying fraudulent activities aligns with emerging patterns and techniques within data mining. Murali et al [38] proposed a hybrid method that combines process mining and machine learning to detect inconsistencies in data flows. This multi-perspective approach recognizes the dynamic nature of user behavior and enhances the robustness of cybersecurity measures by combining various analytical perspectives.

K-Means Clustering Algorithm

In the domain of data mining, two core techniques—K-means clustering and association rule mining—play pivotal roles in the analysis of security-related data. These methods can identify patterns, enhance detection systems, and ensure better decision-making processes in cybersecurity. Here, we detail the respective formulae along with supporting references related to these techniques.

The K-means clustering algorithm seeks to partition a data set into (k) clusters, minimizing the variance within each cluster. The objective function of K-means can be represented mathematically as:

$$J = \sum\{i = 1\}\^{}\{k\} \sum\{xj \in Si\} |xj - \mui|^2$$

Where:

(J) is the total cost (or intra-cluster variance),

(k) is the number of clusters,

(x_j) represents a data point,

(S_i) is the set of points in cluster (i),

(\mu_i) is the centroid of cluster (i), and

(| xj - \mui |^2) denotes the squared Euclidean distance between point (xj) and the centroid (\mui).

In a relevant study, Hackl et al [39] illustrate the application of the K-means clustering method, emphasizing that the method effectively partitions datasets into different groups based on the minimization of (J) within contexts pertinent

to economic decision-making, which can be extrapolated to security data analysis.

Additional discussions on clustering can be found in Li et al [40] who explore K-means in the context of visual navigation methods and its effectiveness in clustering features essential for recognizing group characteristics, although their focus is not directly on cybersecurity. This highlights the versatility of K-means clustering across various applications, reinforcing its foundational role in data mining.

Association Rule Mining

Association rule mining is another crucial data mining technique widely used in cybersecurity to identify relationships between variables. The support for an association rule is computed using the formula:

$$\text{Support}(A \rightarrow B) = \frac{\text{count}(A \cap B)}{\text{count}(S)}$$

Where:

$\text{Support}(A \rightarrow B)$ refers to the likelihood that items (A) and (B) appear together in a dataset,

$\text{count}(A \cap B)$ is the number of transactions containing both (A) and (B),

$\text{count}(S)$ is the total number of transactions in the dataset.

These metrics assist in discovering essential patterns that indicate users' behavior or system vulnerabilities. However, while the study by Bazionis et al. [41]. emphasizes the relevance of association rules in measuring correlations among wind farm performances, it does not directly relate to cybersecurity. There is a need for references more closely aligned with the cybersecurity domain to substantiate the application of association rule mining in this context.

In cybersecurity, such associations can help identify potential fraud patterns by analyzing user actions across various transactions, assisting in anomaly detection systems. Additionally, this process has implications in predictive analytics, allowing systems to pre-emptively address security threats based on identified behaviors.

Gaps in Existing Research

Identifying gaps in existing research related to user authentication behaviors, particularly within legal and cybersecurity contexts, reveals an urgent need for deeper analysis. While current studies offer insights into various authentication techniques and their usability, many fail to examine how user behavior patterns can be influenced by legal frameworks and societal expectations. This section discusses the identified gaps and highlights relevant studies that shed light on these issues. One significant gap lies in the investigation of how biometric authentication, specifically physiological and behavioral features, can be enhanced by user understanding and compliance. Wong et al [22] offer a valuable approach by proposing a system called ArtiLock, which emphasizes the importance of non-intrusive methods that can secure authentication through minimal user involvement. While their findings demonstrate a high usability rate, further research is required to examine how users interpret and respond to such systems within varying cultural and legal contexts.

Moreover, the research by Al-Ameen et al [42] highlights the impact of autobiographical memory on user authentication. Although their study sheds light on the usability of multi-factor authentication (MFA) systems, it does not sufficiently address how users' legal awareness and understanding of data protection laws influence their behavior when utilizing these authentication methods. There exists a critical need for studies that explore the intersection of legal knowledge and user authentication behavior to enhance compliance and security. In the healthcare sector, Turner et al [43] reveal usability issues in patient portals, drawing attention to the barriers faced by users due to a lack of contextual information and educational resources. Despite their findings, a gap persists in exploring how legal obligations regarding health data protection could shape user behaviors in such portals. Future research should investigate how compliance with health privacy regulations affects user engagement and authentication method effectiveness within healthcare environments.

The survey by Παπαϊωάννου et al [44] discusses the efficacy of biometric-based authentication systems, yet it lacks exploration into the ethical and legal ramifications of using personal data for authentication. This gap necessitates comprehensive studies that could evaluate the implications of biometric data usage in legal frameworks and consumer protection, focusing on balancing usability with privacy. Wang et al [25] emphasize the importance of non-intrusive approaches in user authentication, yet there remains insufficient analysis on how such systems comply with existing legal standards. Research that examines the regulatory environment surrounding user authentication, particularly in sectors engaged in sensitive data handling (e.g., finance or healthcare), is vital to identify any misalignments between technological advancements and legal compliance. Stylios et al [21] explore the introduction of continuous authentication through behavioral biometrics, suggesting a more dynamic approach to user verification. However, this exploration could benefit from deeper analysis regarding how users perceive and respond to ongoing monitoring within the context of their legal rights and privacy expectations.

Additional gaps are evident in the understanding of keystroke dynamics as a user authentication method. Alsultan et al [45] highlight the potential of this method in maintaining security without additional effort from users; however, more studies are needed to understand users' perceptions of privacy related to continuous keystroke monitoring, especially in sensitive environments. Moreover, the research conducted by Wang et al [46] explores continuous authentication across multiple devices. They discuss the challenges and provide solutions for ensuring user compliance and behavior in adopting such multi-device authentication approaches, although privacy implications should also be considered.

Finally, the analysis of password management by Blocki and Zhang [31] exposes the significant influence of human behavior in the security landscape. They note that users often select poor passwords; yet, the implications of legal repercussions stemming from data breaches due to weak passwords remain largely unexplored. Research that focuses on the legal consequences of inadequate password security could bolster user education and compliance efforts.

Method

Data Collection and Preprocessing

The foundation of this analysis was the password_security_dataset.csv dataset, a collection of anonymized records detailing user authentication events, including session characteristics, input metrics, and environmental data. Upon initial loading, a critical preprocessing phase was undertaken to cleanse and structure the data for modeling. The first step involved the strategic removal of high-cardinality identifiers, specifically user_id and session_ip. These features were excluded because the research objective is to identify generalizable patterns of user behavior, not to model the actions of specific individuals. Their inclusion would introduce noise and dimensionality without contributing to the discovery of broader behavioral archetypes.

The remaining features were then systematically prepared for the clustering algorithm using a ColumnTransformer pipeline from scikit-learn. This pipeline treated numerical and categorical features distinctly to meet the requirements of the K-Means algorithm. Numerical features were standardized using StandardScaler, which transforms each feature to have a mean of 0 and a standard deviation of 1. This step is crucial because K-Means is a distance-based algorithm that is highly sensitive to the scale of input variables; without standardization, features with larger magnitudes (e.g., session duration in milliseconds) would disproportionately dominate the clustering process over features with smaller ranges (e.g., number of login attempts). Categorical features were transformed into a numerical format using OneHotEncoder. This technique converts each category into a sparse binary vector, effectively creating new features for each unique category. This prevents the model from imposing an artificial ordinal relationship on nominal data and was configured with the handle_unknown='ignore' parameter to ensure the model's robustness against new, unseen categories during potential future applications.

Exploratory Data Analysis (EDA)

Prior to model implementation, a comprehensive exploratory data analysis was performed as a critical preliminary step for hypothesis generation and data validation. This phase involved generating a series of visualizations to uncover the underlying structure, distributions, and relationships within the data. Histograms were created for all numerical features to examine their frequency distributions, central tendency, and spread, which helped in identifying skewness or potential outliers that could influence the performance of the clustering algorithm. For categorical features with a manageable number of unique values (a maximum threshold of 50), count plots were generated. These plots were essential for visualizing the prevalence of each category and identifying any significant class imbalances (e.g., a dominant browser or operating system) that might later help characterize the identified user clusters. Finally, a correlation heatmap was produced for the numerical variables. Its purpose was twofold: to identify any significant linear relationships between features that could provide context for the clustering results, and to detect potential multicollinearity, which, while not a direct problem for K-Means, is important to note for a holistic understanding of the feature space.

K-Means Clustering

The core of this research involved segmenting users into distinct behavioral groups using the K-Means clustering algorithm, implemented via scikit-learn's standard Lloyd's algorithm. The objective of K-Means is to partition data points

into K distinct, non-overlapping clusters by iteratively minimizing the within-cluster sum of squares (inertia). The process begins by initializing K centroids, assigning each data point to its nearest centroid, and then recalculating the centroid's position as the mean of all points assigned to it, repeating until the assignments no longer change.

A critical prerequisite for this process was determining the optimal number of clusters, K. This was achieved by systematically evaluating a range of K values from 2 to 10 using the sophisticated k-means++ initialization method, which intelligently selects initial cluster centers to encourage faster convergence and more consistent results. This search was conducted with n_init=8 initializations and a maximum of 300 iterations per run to balance computational speed with thoroughness. Two metrics guided this evaluation: the Elbow Method, which observes the diminishing returns in inertia as K increases, and the Silhouette Score, which provides a more nuanced measure of cluster quality by evaluating both intra-cluster cohesion and inter-cluster separation. The optimal K was ultimately selected based on the value that yielded the highest Silhouette Score, indicating the most well-defined and meaningful cluster structure. Once identified, the final K-Means model was trained on the preprocessed data with more robust parameters—n_init=20 initializations and a maximum of 500 iterations—to ensure a stable and optimal final solution. To visualize these high-dimensional clusters, Principal Component Analysis (PCA) was employed to reduce the feature space to n_components=2 for plotting, allowing for a qualitative assessment of the cluster separation in a 2D space.

Association Rule Mining

To complement the clustering analysis and provide actionable insights, Association Rule Mining was conducted to discover significant relationships between user attributes and their assigned behavioral cluster. While K-Means identifies which users group together, this step helps explain why. The Apriori algorithm from the mlxtend library was applied to a dataset composed of the original categorical features and the newly generated cluster labels. The algorithm first identified frequent itemsets—combinations of attributes that appear together with a frequency above a minimum support threshold set to 0.1. This support value ensures that any discovered patterns are present in at least 10% of the user sessions, filtering out rare and potentially spurious correlations.

From these frequent itemsets, association rules were generated using confidence as the primary evaluation metric, with a minimum threshold of 0.6. A rule such as {Browser=X} -> {Cluster=Y} with a confidence of 0.6 means that 60% of users with browser X belong to cluster Y. This filtering ensures the reliability and predictive power of the discovered relationships. The final rules were then sorted by lift and confidence. Lift measures how much more likely the consequent is given the antecedent, making it an excellent metric for identifying the most interesting and non-obvious patterns that characterize the user behavior within each cluster.

Result and Discussion

Exploratory Data Analysis Results

The initial Exploratory Data Analysis (EDA) provided crucial insights into the fundamental structure and characteristics of the dataset, revealing distributions that informed the subsequent clustering methodology. The analysis was divided

into the examination of numerical feature distributions, categorical feature frequencies, and the correlation between numerical variables.

An examination of the numerical features through histograms (figure 1) revealed largely uniform distributions for password_length, login_attempts, and browser_tab_count. These variables showed no significant central tendency, with values spread evenly across their respective ranges. This uniformity suggests that the dataset was likely balanced or synthetically generated to ensure a wide and consistent representation of these behaviors. In contrast, the typing_speed_wpm feature exhibited a more natural, bell-shaped distribution, resembling a normal curve centered around approximately 70 words per minute. This indicates that while other factors may have been controlled, typing speed reflects a more organic behavioral trait within the data.

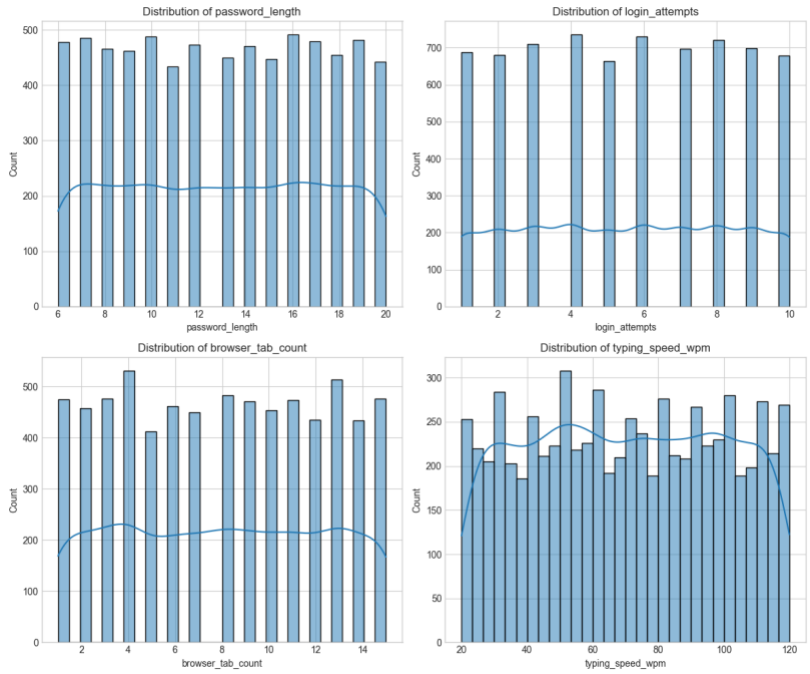

Figure 1 Histogram of Numerical Features

The analysis of categorical features through count plots (figure 2) showed a similar pattern of remarkable balance. The binary features used_special_characters and was_capslock_on were split almost perfectly 50/50 between their 'yes' and 'no' categories. Similarly, the multi-category features of keyboard_language and timezone displayed an exceptionally even distribution across all their unique values. Such perfect balance is highly uncharacteristic of organic datasets, reinforcing the conclusion that the data was intentionally constructed to prevent any single category from disproportionately influencing the analysis and to provide a robust foundation for pattern detection.

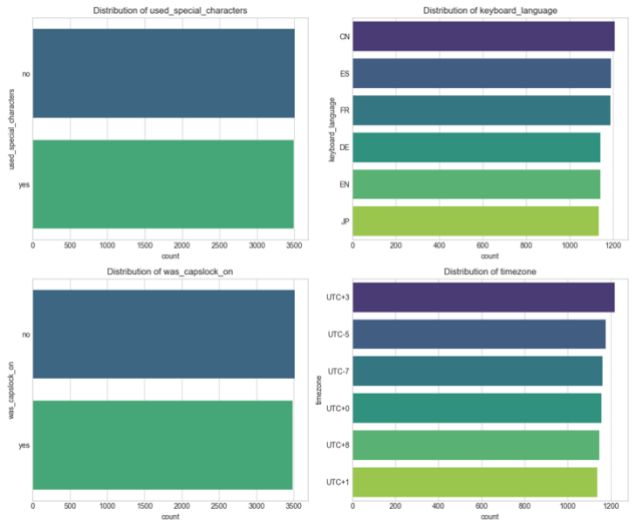

Figure 2 **Count Plot of Categorical Features**

Finally, the correlation matrix of the numerical features (figure 3) confirmed a complete lack of linear relationships between the variables. The correlation coefficients between password_length, login_attempts, browser_tab_count, and typing_speed_wpm were all approximately zero. This statistical independence is a significant finding, indicating that, within this dataset, a user's typing speed, for instance, has no linear connection to the number of their login attempts or the length of their password. This absence of correlation, much like the uniform distributions, is a strong indicator of a synthetic dataset where features were generated independently of one another. This characteristic is advantageous for clustering, as it reduces multicollinearity and ensures that each feature contributes unique information to the model.

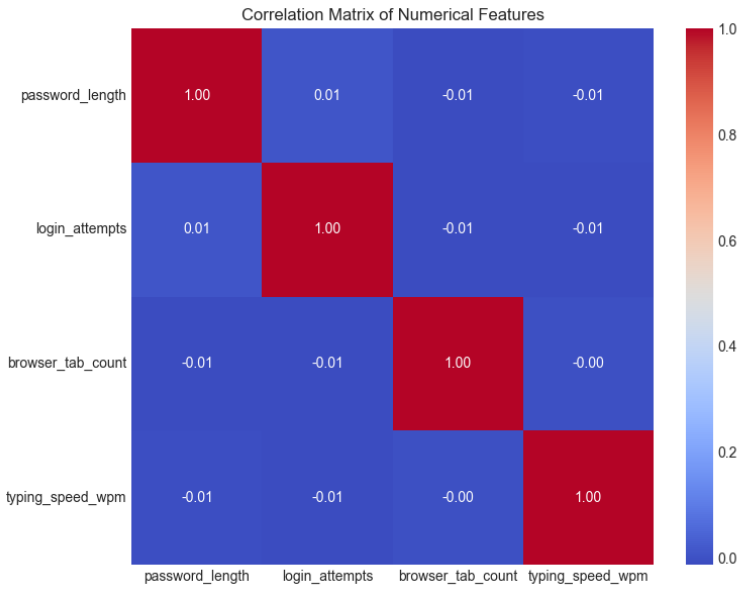

Figure 3 **Correlation Matrix of Numerical Features**

Clustering Results and User Profile Identification

Following the preprocessing of the data, the K-Means clustering algorithm was applied to partition the user sessions into distinct behavioral groups. The analysis focused on determining the optimal number of clusters and then interpreting the characteristics of the resulting segments to identify meaningful user archetypes. The results reveal a clear, data-driven segmentation of the user base into two primary profiles. The optimal number of clusters, K, was determined by evaluating a range of values from 2 to 10 using both the Elbow Method (measuring inertia) and the Silhouette Score. As illustrated in the accompanying plot (figure 4), while the inertia decreased steadily with the addition of more clusters (the "elbow" is not sharply defined), the Silhouette Score provided a more decisive metric. The score, which measures both cluster cohesion and separation, reached its maximum value at K=2, indicating that a two-cluster solution provides the most meaningful and statistically valid segmentation of the data.

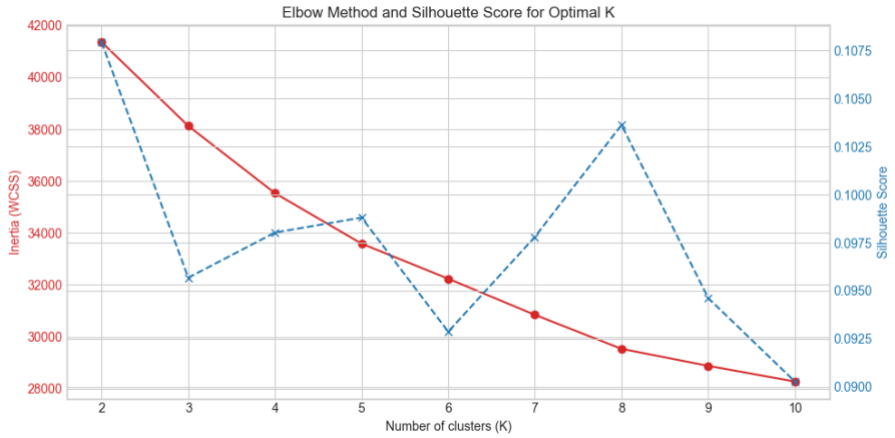

Figure 4 Elbow Method and Silhouette Score

To visualize the separation of these two clusters, Principal Component Analysis (PCA) was used to reduce the dimensionality of the data into two components. The resulting scatter plot shows the two identified clusters (figure 5). While there are two discernible groups, there is also significant overlap between them. This visual evidence, consistent with the relatively low Silhouette Score, suggests that while the "high-attempt" and "low-attempt" user behaviors are distinct archetypes, they exist on a continuum rather than as perfectly discrete groups.

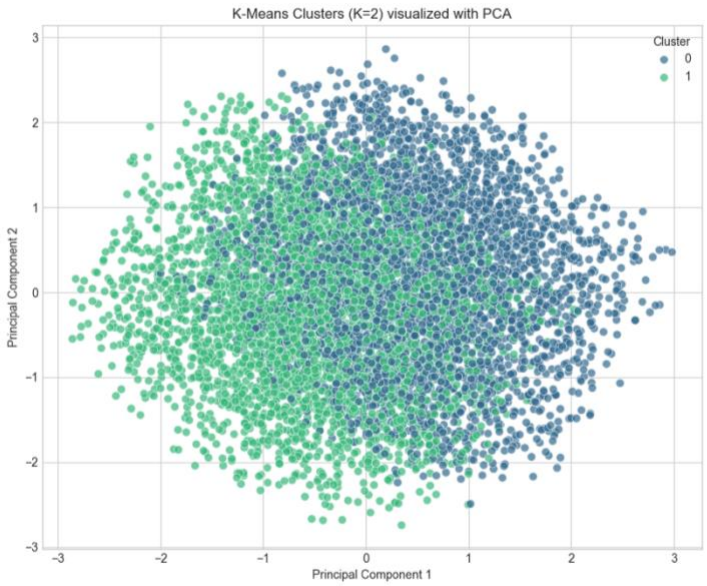

Figure 5 Visualization of Clusters

A detailed analysis of the cluster centroids allowed for the construction of clear behavioral archetypes, which we have termed the "Struggling User" and the "Efficient User." Cluster 0 (The Struggling User), representing just over half the user base, is empirically defined by a significantly high number of login attempts, with a mean of approximately 7.97. This metric alone points to considerable friction in the authentication process. This behavior is further contextualized by the finding that members of this cluster were more likely to have Caps Lock enabled during their sessions and tended to use special characters in their passwords. These characteristics combine to paint a vivid picture of a user caught in a cycle of frustration. The high attempt count, coupled with common and often unnoticed error sources like an active Caps Lock, suggests significant difficulty in accurately recalling or inputting complex credentials. This struggle may paradoxically stem from an attempt to adhere to stringent security advice, leading to the creation of passwords that are too complex to be easily remembered and entered.

Cluster 1 (The Efficient User), in stark contrast, demonstrates a much smoother and more successful authentication experience. Their defining characteristic is a low number of login attempts, with a mean of just 3.00. These users were less likely to use special characters and typically had Caps Lock disabled, indicating a more streamlined and error-free process. This profile represents a user who can successfully authenticate with minimal cognitive load and effort, likely using passwords that strike a better balance between security and memorability. Their efficiency translates directly into a more positive user experience and a lower likelihood of triggering security alerts or account lockouts.

A particularly critical finding emerged from the features that did not differentiate the clusters. The average password length (around 13 characters), typing speed (around 70 WPM), and browser tab count (around 8) were nearly identical for both groups. This is a crucial insight because it isolates the key behavioral differentiators from confounding variables. It strongly suggests that authentication success is not necessarily linked to a user's general technical proficiency (as measured by typing speed) or their adherence to basic security

advice (like using longer passwords). Instead, the friction point is the immediate cognitive and mechanical experience of entering the password, where factors like complexity-induced memory lapses and simple input errors become the primary drivers of failure.

Association Rule Mining Findings

The application of the Apriori algorithm for Association Rule Mining, configured with a minimum support of 0.1 and a minimum confidence of 0.6, did not yield any association rules. While this may initially seem like a lack of findings, this "null result" is itself an important and informative outcome. It demonstrates that there are no strong, simplistic, one-to-one relationships between individual categorical features (such as a specific keyboard_language or timezone) and a user's cluster assignment.

This implies that user authentication behavior is a complex, multifactorial phenomenon that cannot be accurately predicted by a single environmental or demographic attribute. The patterns of struggle or efficiency are emergent properties arising from a combination of factors, which is precisely the kind of complex, latent structure that K-Means is designed to uncover. The failure of Association Rule Mining to find simple rules reinforces the validity of using a more sophisticated, multivariate clustering approach and cautions against making superficial assumptions about user behavior based on isolated data points.

Discussion and Implications for Cyberlaw

The empirical identification and characterization of the "Struggling User" (Cluster 0) carries profound implications for both operational cybersecurity policy and the legal standards of care in data protection. This cluster provides concrete, quantitative evidence that security measures, while well-intentioned, can directly lead to negative user experiences that have tangible security consequences. The high number of login attempts is not merely an inconvenience; it is a direct indicator of a usability failure. When nearly half of a user base repeatedly fails to log in, it creates a cascade of negative outcomes: it causes user frustration that can lead to password-related helpdesk calls (increasing operational costs); it can trigger automated account lockouts that require manual intervention; and, in aggregate, the high volume of failed logins could be mistaken for a distributed brute-force attack, generating false positives that waste the time of security analysts.

From a cyberlaw perspective, these findings directly inform the debate around the "reasonableness" and "appropriateness" of security measures—a central tenet in data protection regulations like the GDPR and the CCPA. A security policy that mandates high password complexity (e.g., requiring multiple special characters) might be considered technically robust in a vacuum. However, if our data demonstrates that this policy is a contributing factor to the formation of a large cluster of users who consistently struggle and fail to authenticate, its "reasonableness" can be credibly questioned. It highlights a fundamental conflict between prescriptive security controls and the practical reality of human-computer interaction. A regulator could argue that a system that imposes such a high cognitive burden on a significant portion of its users is not "appropriate" for protecting personal data, as it may encourage insecure compensatory behaviors, such as writing passwords down.

This research strongly supports a necessary shift in legal and regulatory thinking, moving away from rigid, one-size-fits-all password policies and towards more adaptive, user-centric, and evidence-based authentication frameworks. For instance, regulations could encourage or mandate the use of risk-based authentication, where login challenges are escalated (e.g., requiring MFA) only when anomalous activity is detected, rather than imposing a high barrier to entry for every login. Furthermore, the data suggests that systems should provide more intelligent user feedback, such as clear, non-intrusive "Caps Lock is on" warnings. Ultimately, this analysis shows that improving the user experience—by reducing the friction that leads to high attempt counts—is not a matter of convenience but is a critical and measurable component of a holistic and effective cybersecurity strategy. By understanding these behavioral patterns, organizations can design systems that are both secure and usable, better aligning their technical practices with the protective spirit of modern data protection laws.

Limitations and Future Research Suggestions

It is important to acknowledge the limitations of this study. First, the analysis was conducted on a single dataset, and its characteristics may not be generalizable to all user populations or system types. The behavioral patterns identified could be specific to the context in which the data was collected. Second, a key limitation is the absence of a "ground truth" label indicating whether a login session was ultimately successful or not. The "struggle" is inferred from the high number of attempts, which is a strong proxy but not a definitive measure of failure. Third, the Silhouette Score of 0.1079, while indicating the best possible solution for K, suggests that the clusters are not perfectly distinct and have some overlap. This could be due to the inherent complexity of human behavior or the limitations of the K-Means algorithm itself, which prefers spherical clusters.

The findings and limitations of this study open several promising avenues for future research. A primary goal should be to replicate this analysis across diverse datasets from different industries and geographical regions to test the generalizability of the "Struggling" and "Efficient" user archetypes. Future work should also seek to incorporate datasets that include ground truth labels for login success or failure. This would allow for the use of supervised machine learning models to predict user struggle and would provide a more definitive validation of the cluster profiles.

Furthermore, exploring more advanced clustering algorithms, such as DBSCAN (which can find non-spherically shaped clusters) or Gaussian Mixture Models (which provide probabilistic cluster assignments), could reveal more nuanced user groupings. Finally, a valuable next step would be to augment this quantitative analysis with qualitative research. Conducting surveys or interviews with users identified as belonging to the "Struggling User" cluster could provide rich, contextual insights into their experiences, frustrations, and the specific usability challenges they face, leading to more empathetic and effective system design.

Conclusion

This research successfully utilized K-Means clustering to analyze a dataset of user authentication logs, revealing two distinct behavioral archetypes: the "Struggling User" and the "Efficient User." The primary differentiator between these groups was not user proficiency, as measured by typing speed, but rather

the number of login attempts, which was strongly associated with factors like Caps Lock usage and the inclusion of special characters in passwords. This core finding provides quantitative evidence that the friction in authentication systems often stems from usability challenges related to password complexity, rather than a lack of user skill or adherence to security advice like using longer passwords. The analysis demonstrates that a significant portion of users experience considerable difficulty in the authentication process, a reality that has profound security and operational consequences.

Ultimately, this study contributes a data-driven perspective to the ongoing dialogue between security and usability. By empirically identifying and characterizing a "Struggling User" profile, this work underscores the urgent need for a paradigm shift away from rigid, prescriptive security policies and toward more adaptive, user-centric authentication frameworks. The findings have direct implications for cyberlaw, challenging a narrow definition of "reasonable" security and advocating for a more holistic standard that considers user experience as a critical component of an effective security posture. By designing systems that mitigate common user errors and reduce authentication friction, organizations can not only enhance security but also better align their practices with the protective principles of modern data protection legislation.

Declarations

Author Contributions

Conceptualization: C.I.S.N.; Methodology: E.I.; Software: E.I.; Validation: C.I.S.N.; Formal Analysis: E.I.; Investigation: E.I.; Resources: C.I.S.N.; Data Curation: E.I.; Writing Original Draft Preparation: C.I.S.N.; Writing Review and Editing: E.I.; Visualization: C.I.S.N.; All authors have read and agreed to the published version of the manuscript.

Data Availability Statement

The data presented in this study are available on request from the corresponding author.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

Institutional Review Board Statement

Not applicable.

Informed Consent Statement

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. R. Pratama, F. M. Firmansyah, and F. Rahma, “Security Awareness of Single

Sign-on Account in the Academic Community: The Roles of Demographics, Privacy Concerns, and Big-Five Personality," *Peerj Comput. Sci.*, 2022, doi: 10.7717/peerj-cs.918.

[2] A. P. Nanda, "Role of AI in Enhancing Digital Payment Security," *Ajbr*, 2024, doi: 10.53555/ajbr.v27i3s.2546.

[3] M. Neri, F. Niccolini, and L. Martino, "Organizational Cybersecurity Readiness in the ICT Sector: A Quanti-Qualitative Assessment," *Inf. Comput. Secur.*, 2023, doi: 10.1108/ics-05-2023-0084.

[4] M. Ahsan, K. E. Nygard, R. Gomes, M. M. Chowdhury, N. Rifat, and J. F. Connolly, "Cybersecurity Threats and Their Mitigation Approaches Using Machine Learning—A Review," *J. Cybersecurity Priv.*, 2022, doi: 10.3390/jcp2030027.

[5] S. O. Olabanji, O. O. Olaniyi, C. S. Adigwe, O. J. Okunleye, and T. O. Oladoyinbo, "AI for Identity and Access Management (IAM) in the Cloud: Exploring the Potential of Artificial Intelligence to Improve User Authentication, Authorization, and Access Control Within Cloud-Based Systems," *Asian J. Res. Comput. Sci.*, 2024, doi: 10.9734/ajrcos/2024/v17i3423.

[6] Z. H. Joy, S. Islam, M. A. Rahaman, and Md. N. Haque, "Advanced Cybersecurity Protocols for Securing Data Management Systems in Industrial and Healthcare Environments," *GMJ*, 2024, doi: 10.62304/jbedpm.v3i4.147.

[7] J. Garba, J. Kaur, and E. Ibrahim, "Design of a Conceptual Framework for Cybersecurity Culture Amongst Online Banking Users in Nigeria," *Niger. J. Technol.*, 2023, doi: 10.4314/njt.v42i3.13.

[8] X. J. Zhang, Z. Li, and H. Deng, "Information Security Behaviors of Smartphone Users in China: An Empirical Analysis," *Electron. Libr.*, 2017, doi: 10.1108/el-09-2016-0183.

[9] C. Conrad, J. R. Aziz, J. M. Henneberry, and A. J. Newman, "Do Emotions Influence Safe Browsing? Toward an Electroencephalography Marker of Affective Responses to Cybersecurity Notifications," *Front. Neurosci.*, 2022, doi: 10.3389/fnins.2022.922960.

[10] N. Rodríguez-Priego, R. van Bavel, J. Vila, and P. Briggs, "Framing Effects on Online Security Behavior," *Front. Psychol.*, 2020, doi: 10.3389/fpsyg.2020.527886.

[11] A. Salem, A. Sharieh, and R. Jabri, "Online User Authentication System Using Keystroke Dynamics," *J. Comput. Secur.*, 2022, doi: 10.3233/jcs-210081.

[12] Y. Wang, C. Wu, K. Zheng, and X. Wang, "Improving Reliability: User Authentication on Smartphones Using Keystroke Biometrics," *Ieee Access*, 2019, doi: 10.1109/access.2019.2891603.

[13] A. Salem, A. Sharieh, A. Sleit, and R. Jabri, "Enhanced Authentication System Performance Based on Keystroke Dynamics Using Classification Algorithms," *Ksii Trans. Internet Inf. Syst.*, 2019, doi: 10.3837/tiis.2019.08.014.

[14] A. Mahfouz, T. M. Mahmoud, and A. S. Eldin, "A Survey on Behavioral Biometric Authentication on Smartphones," *J. Inf. Secur. Appl.*, 2017, doi: 10.1016/j.jisa.2017.10.002.

[15] R. AlHusain and A. Alkhalifah, "Evaluating Knowledge-Based Security Questions for Fallback Authentication," *Peerj Comput. Sci.*, 2022, doi: 10.7717/peerj-cs.903.

[16] E. Kim, J. Yoon, J. Kwon, T. Liaw, and A. M. Agogino, "From Innocent Irene to Parental Patrick: Framing User Characteristics and Personas to Design for Cybersecurity," *Proc. Des. Soc. Int. Conf. Eng. Des.*, 2019, doi: 10.1017/dsi.2019.183.

[17] M. M. Islam, R. Safavi–Naini, and M. Kneppers, "Scalable Behavioral Authentication," *Ieee Access*, 2021, doi: 10.1109/access.2021.3065921.

[18] X. Liang, F. Zou, L. Li, and P. Yi, "Mobile Terminal Identity Authentication System Based on Behavioral Characteristics," *Int. J. Distrib. Sens. Netw.*, 2020, doi: 10.1177/1550147719899371.

[19] P. Bansal and A. Ouda, "Continuous Authentication in the Digital Age: An Analysis of Reinforcement Learning and Behavioral Biometrics," *Computers*, 2024, doi: 10.3390/computers13040103.

[20] J. Spooren, D. Preuveneers, and W. Joosen, “Leveraging Battery Usage From Mobile Devices for Active Authentication,” *Mob. Inf. Syst.*, 2017, doi: 10.1155/2017/1367064.
[21] I. Stylios, S. Kokolakis, O. Thanou, and S. Chatzis, “Key Factors Driving the Adoption of Behavioral Biometrics and Continuous Authentication Technology: An Empirical Research,” *Inf. Comput. Secur.*, 2022, doi: 10.1108/ics-08-2021-0124.
[22] A. B. Wong, Z. Huang, X. Chen, and K. Wu, “ArtiLock: Smartphone User Identification Based on Physiological and Behavioral Features of Monosyllable Articulation,” *Sensors*, 2023, doi: 10.3390/s23031667.
[23] J. M. Jorquera Valero *et al.*, “Improving the Security and QoE in Mobile Devices Through an Intelligent and Adaptive Continuous Authentication System,” *Sensors*, 2018, doi: 10.3390/s18113769.
[24] I. M. Hussain Qureshi and V. K. Kale, “A Study of Risk-Based Authentication System in Cyber Security Using Machine Learning,” *World J. Adv. Eng. Technol. Sci.*, 2022, doi: 10.30574/wjaets.2022.7.2.0125.
[25] Z. Wang *et al.*, “A Survey of User Authentication Based on Channel State Information,” *Wirel. Commun. Mob. Comput.*, 2021, doi: 10.1155/2021/6636665.
[26] C.-F. Su, J.-C. Kao, C.-S. Shieh, J.-F. Chang, and M.-F. Horng, “New User Authentication Based on Optical Spectrum and Its Realization of Embedded Systems,” *Int. J. Comput. Theory Eng.*, 2018, doi: 10.7763/ijcte.2018.v10.1213.
[27] S. Wiefling, P. R. Jørgensen, S. Thunem, and L. L. Iacono, “Pump Up Password Security! Evaluating and Enhancing Risk-Based Authentication on a Real-World Large-Scale Online Service,” *Acm Trans. Priv. Secur.*, 2022, doi: 10.1145/3546069.
[28] M. Ehatisham-ul-Haq *et al.*, “Authentication of Smartphone Users Based on Activity Recognition and Mobile Sensing,” *Sensors*, 2017, doi: 10.3390/s17092043.
[29] G. Ryu, S.-H. Kim, and D. Choi, “Implicit Secondary Authentication for Sustainable SMS Authentication,” *Sustainability*, 2019, doi: 10.3390/su11010279.
[30] J. Lee, S. Park, Y. Kim, E. Lee, and J. Jo, “Advanced Authentication Method by Geometric Data Analysis Based on User Behavior and Biometrics for IoT Device With Touchscreen,” *Electronics*, 2021, doi: 10.3390/electronics10212583.
[31] J. Blocki and W. Zhang, “DALock: Password Distribution-Aware Throttling,” *Proc. Priv. Enhancing Technol.*, 2022, doi: 10.56553/popets-2022-0084.
[32] S. S. Hasan, A. Ghani, I. U. Din, A. Almogren, and A. Altameem, “IoT Devices Authentication Using Artificial Neural Network,” *Comput. Mater. Contin.*, 2022, doi: 10.32604/cmc.2022.020624.
[33] J. Wu, M. Dong, K. Ota, J. Li, and Z. Guan, “Big Data Analysis-Based Secure Cluster Management for Optimized Control Plane in Software-Defined Networks,” *Ieee Trans. Netw. Serv. Manag.*, 2018, doi: 10.1109/tnsm.2018.2799000.
[34] K. Abouelmehdi, A. Beni-Hessane, and H. Khaloufi, “Big Healthcare Data: Preserving Security and Privacy,” *J. Big Data*, 2018, doi: 10.1186/s40537-017-0110-7.
[35] S. Setia, J. Pandey, and N. Duhan, “HPM: A Hybrid Model for User’s Behavior Prediction Based on N-Gram Parsing and Access Logs,” *Sci. Program.*, 2020, doi: 10.1155/2020/8897244.
[36] J. Jin and X. Lin, “Web Log Analysis and Security Assessment Method Based on Data Mining,” *Comput. Intell. Neurosci.*, 2022, doi: 10.1155/2022/8485014.
[37] C. Syadzali, S. Suryono, and J. E. Suseno, “Business Intelligence Using the K-Nearest Neighbor Algorithm to Analyze Customer Behavior in Online Crowdfunding Systems,” *E3s Web Conf.*, 2020, doi: 10.1051/e3sconf/202020216005.
[38] E. Murali *et al.*, “An Approach Utilizing Multiple Viewpoints to Identify Fraudulent Activity in Commercial Transactions Involving Multiple Parties,” *Int. Res. J. Mod. Eng. Technol. Sci.*, 2024, doi: 10.56726/irjmets53114.
[39] F. Hackl, M. Hölzl-Leitner, R. Winter-Ebmer, and C. Zulehner, “Successful Retailer Strategies in Price Comparison Platforms,” *Manag. Decis. Econ.*, 2021, doi:

10.1002/mde.3309.
[40] X. Li, X. Li, M. O. Khyam, C. Luo, and Y. Tan, “Visual Navigation Method for Indoor Mobile Robot Based on Extended BoW Model,” *Caai Trans. Intell. Technol.*, 2017, doi: 10.1049/trit.2017.0020.
[41] I. K. Bazionis, P. A. Karafotis, and P. S. Georgilakis, “A Review of Short-term Wind Power Probabilistic Forecasting and a Taxonomy Focused on Input Data,” *Iet Renew. Power Gener.*, 2021, doi: 10.1049/rpg2.12330.
[42] M. N. Al-Ameen, S. M. Taiabul Haque, and M. Wright, “Leveraging Autobiographical Memory for Two-Factor Online Authentication,” *Inf. Comput. Secur.*, 2016, doi: 10.1108/ics-01-2016-0005.
[43] K. Turner, A. Clary, Y. Hong, A. A. Tabriz, and C. M. Shea, “Patient Portal Barriers and Group Differences: Cross-Sectional National Survey Study,” *J. Med. Internet Res.*, 2020, doi: 10.2196/18870.
[44] M. Παπαϊωάννου, F. Pelekoudas-Oikonomou, Γ. Μαντάς, E. Serrelis, J. Rodríguez, and M.-A. Fengou, “A Survey on Quantitative Risk Estimation Approaches for Secure and Usable User Authentication on Smartphones,” *Sensors*, 2023, doi: 10.3390/s23062979.
[45] A. Alsultan, K. Warwick, and H. Wei, “Non-Conventional Keystroke Dynamics for User Authentication,” *Pattern Recognit. Lett.*, 2017, doi: 10.1016/j.patrec.2017.02.010.
[46] Y. Wang, X. Zhang, and H. Hu, “Continuous User Authentication on Multiple Smart Devices,” *Information*, 2023, doi: 10.3390/info14050274.