Mitigating the Nexus: Dark Triad Traits and Insider Threats in Cloud Territories

Mohammad Hafiz Hersyah
mohammad.hafiz_hersyah.mc4@is.naist.jp
Nara Institute of Science and Technology
Nara, Japan
Universitas Andalas
West Sumatera, Indonesia

Md Delwar Hossain
delwar.hossain@is.naist.jp
Nara Institute of Science and Technology
Nara, Japan

Hideya Ochiai
ochiai@elab.ic.i.u-tokyo.ac.jp
The University of Tokyo
Tokyo, Japan

Yuzo Taenaka
yuzo@is.naist.jp
Nara Institute of Science and Technology
Nara, Japan

Youki Kadobayashi
youki-k@is.naist.jp
Nara Institute of Science and Technology
Nara, Japan

Abstract

Ensuring information security in the cloud is crucial because insiders with authorized access can pose risks to organizations by engaging in behaviors that compromise information security. Such sinister acts come from behaviors described in the dark triad traits that comprise destructive behaviors into Narcissism, Psychopathy, and Machiavellianism and its 5 additional composite traits, that require to be mitigated appropriately. This study employs diverse algorithms like Random Forest, eXtreme Gradient Boosting (XGBoost), and Support Vector Machine (SVM) to scrutinize insider threats via the dark triad traits framework that highlights the prominence of dark triad traits among men relative to women, yielding substantial experimental precision: 98.09% for males and 97.34% for females, and proposing theoretical mitigation towards insider threats.

CCS Concepts: • **Security and privacy → Social aspects of security and privacy**.

Keywords: Insider Threats, Dark Triad Traits, Malicious Behaviors, Profile Ranking, Cybersecurity

ACM Reference Format:
Mohammad Hafiz Hersyah, Md Delwar Hossain, Hideya Ochiai, Yuzo Taenaka, and Youki Kadobayashi. 2023. Mitigating the Nexus: Dark Triad Traits and Insider Threats in Cloud Territories. In *Proceedings of The Annual Computer Security Applications Conference (ACSAC), Workshop on Research for Insider Threat (WRIT), December 4, 2023, Austin, TX USA. 7 pages. (ACSAC '23, December 4–8, 2023, Austin, TX, USA)*. ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 Introduction

The aftermath of the COVID-19 pandemic has impacted businesses across the globe for the past two years, thereby providing further impetus to the shift towards cloud-based intelligent asset management, which is called cloud digital transformation, where it would reduce reliance on on-premise data centers at a constant pace. Companies have increasingly embraced the new normal by introducing remote work and granting employees access to sensitive, confidential data. As a result, insiders can access company assets in the cloud with excessive freedom, making it more difficult for IT teams to observe the security boundary. Unlike external attackers, insider threats have emerged as one of the most prominent cybersecurity challenges and can engage in malicious activities without actual technical skills and/or experience [1] - [4]. Even worse with those who possess technical abilities, given the adversarial nature of the insider threat domain, malicious insiders employ various techniques to evade detection and establish persistence to conceal their activities, which can range from the theft of confidential data to the sabotage of critical systems and can cause severe harm to an organization's reputation, financial stability, and operational continuity [5] - [7].

Organizations are progressively allocating a growing share of their annual budgets to address insider threats, emphasizing the demand for intelligible systems defense to detect and prevent insider attacks [8]. One of the most recognizable forms of workplace bias is unconscious gender bias due

ACSAC '23, December 4–8, 2023, Austin, TX, USA,
,
© 2023 Association for Computing Machinery.
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

to the gender traits and stereotypes associated with malicious activities, as a study in paper [21] suggests. Existing publications such as [22], [23], and [24] utilize inaccurate psychometrics traits to describe insider threats as research indicated that malicious insiders suffer from personality disorders such as narcissistic disorder and psychopathy and tend to have an appropriate sense of self-esteem such as Machiavellianism as describes in paper [25] are regarded feasible trigger precursor to antisocial and criminal outcomes. On top of that, the matters related to gender and mitigation efforts in these publications remain unsolved.

In this work, we attempt to solve the shortcomings in detecting insider threats by contemplating a machine learning-based scheme. In the course of our investigation, we made deliberate use of an up-to-date dataset sourced from paper [19] that is uniquely tailored to address the Dark Triad traits as they pertain to gender-related dynamics. This dataset was chosen to ensure the relevance and accuracy of our research findings in understanding the manifestations of dark features and personalities within different gender groups and attempting to draw dynamic user profiles in a theoretical approach to mitigate the risk of insider threat. Codified on these facts, this study offered approaches for utilizing machine learning in the following contributions :

- Contribute to the proactive addressing of gender-based differences in malicious insider attributes, particularly focusing on Dark Triad traits, using machine learning.
- Propose machine learning algorithms approach that outperformed previous state-of-the-art findings publications.
- Propose theoretic insight regarding dynamic user profiles in a cloud computing environment that build around dark triad traits and their mitigation concept.

The organization of this paper is as follows. Section 2 describes the relevant background and related works. Section 3 proposed the risk assessment activities. Experimentation in section 4. Discussion regarding limitations and opportunities in Section 5 and the Conclusion in Section 6.

2 Background and Related Works

This section discusses underlying theories and synonymous works regarding Insider Threats.

2.1 Insider Threat Definition, Progression, and Taxonomy

Elmrabit in paper [9] defined a clear rationale of insider threat as "malicious or unintentional activities on the part of an employee, who has, or has had, authorized access to the organization's IT assets, that trigger damage to the assets and/or has a considerable unfavored impact on information security aspects on confidentiality, integrity and availability.

Several publications such as paper [10] proposed a hierarchical taxonomy in five (5) dimensions, namely Cloud Deployment (D), Source (S), Impact (I), Attack Behavior(B), and Cloud Services (Cs). Greitzer et al. in Paper [11] developed a SOFIT taxonomy of insider threat risk that consist of sociotechnical and organizational factor. Paper [12] has brought forward a literature review of insider threat where multiple dimension of Insider and Insider Threat Detection is classified as a set of taxonomy. Paper [13] proposed an approach to gathering organic narratives of an insider threat incident that then uses a computational approach to map the descriptions to an existing insider threat framework. Paper [14] points out a behavioral consultant's assistance towards investigators in several aspects, precisely the deductive profiling approach and case management decision-making with illustrative case studies.

2.2 Behavioral and Psychological Theories

Schultz's paper [16] defines five behavioral indicators that are predictive of an insider seeking to conduct malicious activities toward a system:

- *Deliberate markers and Errors* refer to the fact that various users may engage in deviant behavior online.
- *Preparatory and Verbal behavior* in which a user might use hateful language and a range of system-level commands to perform surveillance on a system.
- *Correlated usage patterns* in which a user might exhibit a particular behavior on multiple loosely coupled subnetworks.
- *Personality traits* such as introversion are assumed to be correlated to the likelihood of a user posing an insider threat.

2.3 Pathological Personality Traits - Dark Triad

The trio personality traits of psychopathy, namely: Machiavellianism, Psychopathy, and Narcissism, are suitable observed traits when dealing with antisocial behavior [20]. The description of each malevolent personality is as follows :

- *Machiavellianism*: Individuals that possess this trait are accurate in their self-knowledge and dishonest in a cautious way. Flexible in comparing the costs and benefits of a specific course of action and do not deviate when the consequences outweigh the benefits.
- *Psychopathy*: Individuals with this trait, affiliated with superficial charm and persistent manipulation, may be prone to antisocial activities and exhibit a deficiency in impulsive control. They are often trained to emulate emotions for a short period, which contributes to these tendencies.
- *Narcissism*: Individuals exhibiting this trait often display unwarranted overconfidence in their decision-making abilities and tend to act impulsively. They frequently overestimate their knowledge and experience a sense of entitlement, which serves as a primary catalyst for their engagement in antisocial behaviors.

2.4 Related Works

Paper [17] aims to identify specific semantic preferences related to dark triad traits by addressing the relationship between users' text and their psychological characteristics. However, the accuracy result is not satisfactory. Achieving 56.2% for narcissism, 55.8% for psychopathy, and 51.6% for Machiavellian. Multiple publications such as [22], [23], and [24] utilize diverse psychometric traits to describe insider threats. In the paper [27], the authors predict the dark-triad(psychopath) personality trait from online content sources from social media. They proposed a deep neural network technique, namely BILSTM. However, the dark triad traits consist of three (3) different personality traits, where each other can be merged, resulting in a compound trait that is out of the boundary paper discussed in this paper.

In publication [28], the authors proposed a novel comprehensive security psychological model that integrates the big-5 and Dark Triad personality traits. The dataset is sourced from email data that contains approximately 150 users' email text. However, the samples from four malicious insiders are not entirely taken from emails; therefore, the authors synthetically constructed positive emails by extracting specific sentences via TF/IDF (Term Frequency - Inverse Document Frequency), achieving a 10.4% false positive rate. In this paper, we intend to observe all possible variations of the dark triad traits in males and females, validated by the paper [26]

3 Methodology

This section aims to identify insider threats for cloud employee screening using Random Forest, XGBoost, and SVM algorithms. Which is inspired by paper [15] and [18].

3.1 Dataset Information and Preprocessing

The dataset was gathered from a paper [19] conducted in 2021. The final result was obtained by conducting a report from three studies, two of which were preregistered and worked on both WEIRD (the US American) and non-WEIRD (Turkish) with 800 total samples (N = 880). The content is obtained by a simulation where participants were informed that the study was about to hypothesize personality by observing male and female faces and directed to complete an online questionnaire. There were male and female versions of comparisons on a total of 8 personalities, and the survey had 16 sets of images. In each set, two faces were presented side by side. The participants were provided with a brief definition of the related personality traits and were asked to guess which person scored higher dark triad score.

The result was that the US-American and Turkish sample participants successfully predicted all Dark Triad personality traits by examining the survey object. A total of 87 features is being reduced to 24 features and transformed into a complete numeric form. The narcissism traits in males and females are represented in NDM_LH_Q, NDM_RH_Q, NDF_LH_Q, and NDF_RH_Q, the similar approach for the rest of the Dark Triad traits representation. The measured treatments taken towards this approach combine all six data columns into three inputs containing each dark triad trait and conduct a label encoding of person 1 and person 2 into 1 and 2 to anticipate NaN values.

The details of the granular dark males and dark females are shown in Table 1, and the workflow in Figure 1. In dark males, we identify the dark triad traits based on their average values of 1 and 0 with records of narcissism 0.63% and 0.36%. The psychopath average value is 0.71% and 0.29%, and Machiavellianism is 0.61% and 0.39%.

While in dark females, we found the average values of 1 and 0 values with records of 0.68% and 0.32% for narcissism, 0.69% and 0.31% for psychopathy, and 0.66% and 0.34% for Machiavellianism.

Table 1. Detected dark Triad Traits on Males and Females

Dark Triad Counts	Narcissim	Psychopathy	Machiavellianism
Males	553	625	535
Females	603	610	583

3.2 Model Architecture

We define a multiclass classification based on paper [26], combining all possible dark triad trait outputs. The adopted methodology, validated through a systematic fusion of conceivable dark traits, emanates from the intrinsic characteristics inherent within the dataset. This involves the formulation of composite representations generated through combinations of binary values (1 and 0) across three distinct dark triad traits: Narcissistic, Machiavellian, and Psychopathic. The integration extends to incorporate permutations of these traits, such as Psychopathic Narcissistic, Manipulative Narcissistic, Anti-Social, and Maleficent attributes. These distinct configurations encapsulate a unique synthesis of complex psychological dimensions from dark traits, which is the efficient approach to describe the dataset.

The process of representation entails encoding each of the aforementioned traits with a coding schema spanning from 0 to 7. This encoding mechanism delineates a comprehensive and structured framework wherein every numeric value corresponds to a distinct integration of dark traits. Utilizing a coding representation ranging from 0 to 7 not only facilitates the unambiguous identification of specific trait configurations but also enables the establishment of a potent quantitative basis for comparison and analysis.

After deciding on eight (8) possible combinations of dark triad traits, we define the features and classes of each male and female aspect into dark males and dark females separation by considering the correlation matrix in Figure 2.

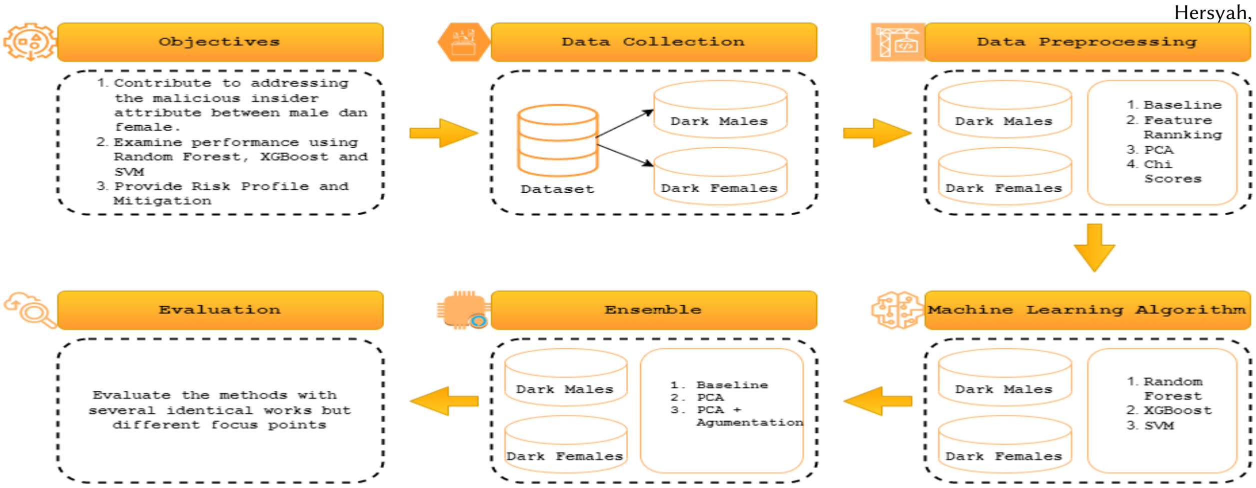
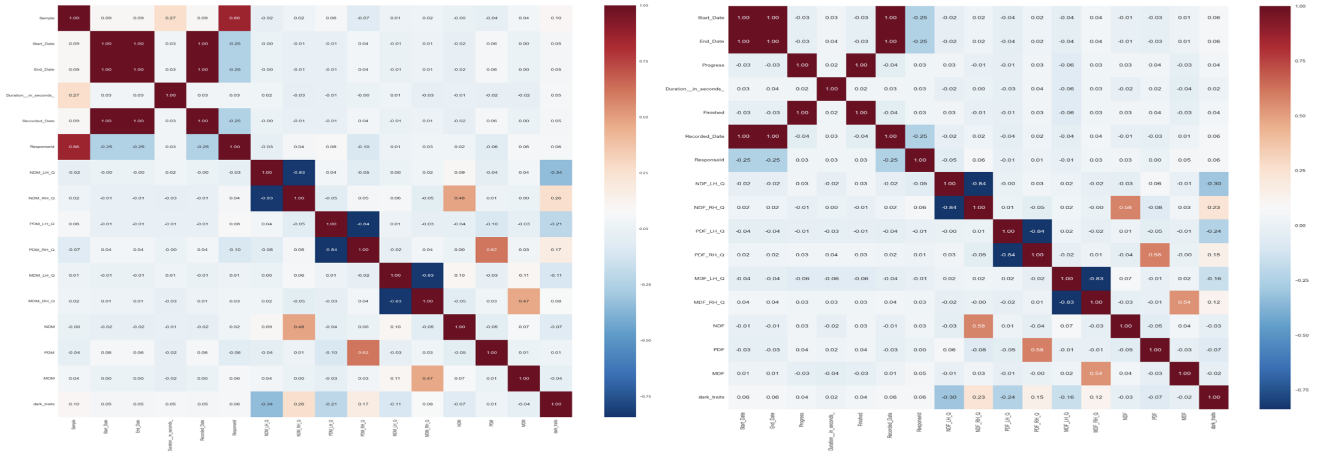

Objectives
1. Contribute to addressing the malicious insider attribute between male dan female.
2. Examine performance using Random Forest, XGBoost and SVM
3. Provide Risk Profile and Mitigation
Data Collection
Dataset
Dark Males
Dark Females
Data Preprocessing
Dark Males
Dark Females
1. Baseline
2. Feature Rannking
3. PCA
4. Chi Scores
Machine Learning Algorithm
Dark Males
Dark Females
1. Random Forest
2. XGBoost
3. SVM
Ensemble
Dark Males
Dark Females
1. Baseline
2. PCA
3. PCA + Agumentation
Evaluation
Evaluate the methods with several identical works but different focus points

Figure 1. The High-level workflow diagram of the proposed work.



(a) Dark Males Correlation Matrix

(b) Dark Females Correlation Matrix

Figure 2. Dark Males and Dark Females Correlation Matrix.

4 Experiments Results and Performance Evaluation

4.1 Experiment Environment

For our experiments, we applied the particular environment used to evaluate this scheme, Anaconda 23.3.1, and built Python 3.10.9 using Jupyter Notebook 6.5.2. We experimented on an AMD Ryzen 7 3750H CPU 2.3Ghz, 32 GB RAM, Windows 11 (64-bit), and NVIDIA GeForce RTX 2060.

4.2 Evaluation Metrics

In assessing the effectiveness of distinct constituents within our devised framework, we employ a comprehensive set of performance metrics, including accuracy, precision, recall, and the F1 score. These metrics hold prominence within the academic domain and have been frequently implemented to evaluate various methodologies' efficacy. These quantifiable measures are characterized in Equations 1 to 4 sequentially:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$F_1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \tag{4}$$

Where TP (True Positives) represents the instances where the model correctly predicted the positive class when the actual class was positive, TN (True Negatives) represents the instances where the model correctly predicted the negative class when the actual class was negative. FP (False Positives) occurs when the model predicts the positive class; however, the actual class is negative. FN (False Negative) occurs when the model predicts the negative class, but the actual class is positive.

4.3 Dark Males and Dark Females Experiments Results using Machine Learning

In this section, we verify the capability of machine learning for identifying dark triad traits. The integration of machine learning and employee screening using dark triad traits is demonstrated in Figure 5.

4.3.1 Experimental Result: Dataset Preprocessing.

we conduct dataset preprocessing to examine the feature selection on each dataset feature to grasp holistic comprehension as described in the following Table 2.

Table 2. Dataset Preprocessing Accuracy

Data Preprocessing	Types	Accuracy
Baseline	Dark Males	90.49%
	Dark Females	90.49%
Feature Ranking	Dark Males	93.15%
	Dark Females	93.15%
PCA (Principal Component Analysis)	Dark Males	95.81%
	Dark Females	95.81%
Chi Scores	Dark Males	90.49%
	Dark Females	90.45%

As presented in Table 2, the analytical assessment engenders an inference that the unity in accuracy between dark male and dark female subsets. This coherence is particularly pronounced across multiple metrics, except for the domain of chi scores. The differential number of features inherent in the dark males and dark females subsets can highlight the divergence observed in the chi scores, which introduces an inherent variability in the computed chi scores.

A supplementary experimental undertaking is conducted within the ensemble experiments to illuminate the analysis’s robustness further and enhance the depth of insights. This additional investigation entails the rigorous examination of accuracy and recall outcomes, employing the polar extremities of the evaluation spectrum represented by the baseline and PCA (Principal Component Analysis) methodologies.

4.3.2 Experimental Result: Algorithm Accuracy.

The research experiment endeavor involves the practical application of several algorithms, namely the Random Forest, Gradient Boosting employing the XGBoost framework, and the Support Vector Machine (SVM). These algorithms, acknowledged for their robustness and adaptability, are sequentially used on distinct subsets of the dataset encompassing dark males and dark females. The tabulated summary in Table 3 provides a comprehensive delineation of the algorithmic configuration and parameters specific to each method.

Table 3. Several Model Classification Results

Dataset Preprocessing	Types	Accuracy	Precision	Recall	F1 Socre
Random Forest	Dark Males	94.31%	94.07%	91.15%	91.94%
	Dark Females	97.72%	98.38%	96.77%	97.43%
XGBOOST	Dark Males	98.09%	97.01%	96.65%	96.74%
	Dark Females	97.34%	97.81%	95.83%	96.30%
Support Vector Machine	Dark Males	98.09%	97.01%	96.65%	96.74%
	Dark Females	97.34%	97.81%	95.83%	96.30%

4.3.3 Experimental Result: Ensemble Techniques.

Incorporating extensive observations, this study employs ensemble techniques to scrutinize the precision and reliability of the dataset meticulously. Furthermore, Principal Component Analysis (PCA) is leveraged to produce graphical representations for each distinct dataset subset, categorized as "Dark males" and "Dark females." The visualizations generated through PCA elucidate notable consistencies; particularly, a steadfast pattern emerges within the context of the 9th feature among the dataset referenced to Dark Males, while an analogous pattern surfaces concerning the 10th feature within the Dark Females.

These identified patterns warrant in-depth examination through ensemble techniques, as elucidated and expounded upon in Figures 3 and 4. Deploying these analytical strategies enhances understanding of the dataset’s complexity and sets the stage for comprehensive insights into the underlying factors driving the observed patterns within the dataset subsets.

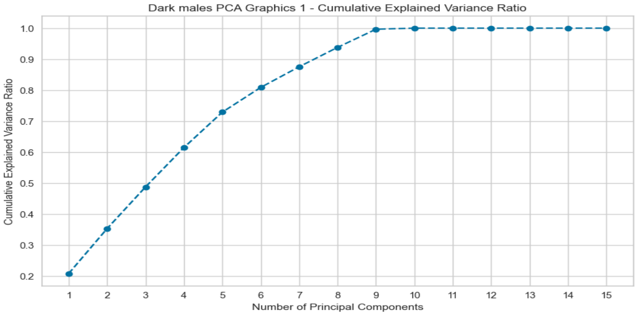


Figure 3. Dark Males PCA Graphic.

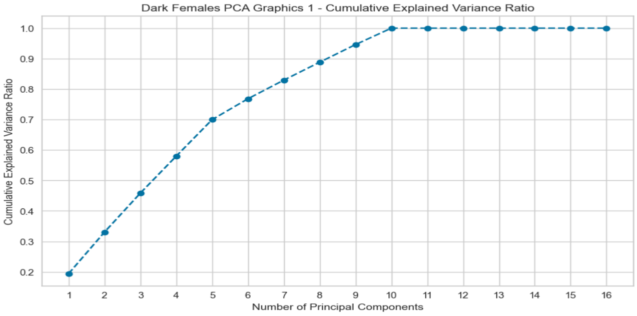


Figure 4. Dark Females PCA Graphic.

Table 4. Ensemble Techniques

Dataset Preprocessing	Types	Baseline		PCA		Augmentation + PCA	
		Accuracy	**Recall**	**Accuracy**	**Recall**	**Accuracy**	**Recall**
Ensemble Boosting and Bagging	Dark Males	**96.59%**	**94.49%**	**97.15%**	**95.69%**	**97.27%**	**96.09%**
	Dark Females	**93.73%**	**89.78%**	**94.88%**	**92.5%**	**95.17%**	**90.82%**
Ensemble Boosting and Bagging using Decision Tree	Dark Males	**92.01%**	**88.53%**	**97.15%**	**95.37%**	**97.57%**	**96.21%**
	Dark Females	**95.07%**	**92.21%**	**96.59%**	**94.94%**	**96.92%**	**94.11%**

The empirical investigations in Table 4 reveal a pattern of consistent and exponential enhancements within the accuracy and recall metrics across both the Dark male and Dark female datasets. The observed phenomenon of steady improvement indicates a systematic and gradual refinement process characterized by a pronounced exponential growth trajectory.

4.4 Comparison with Existing Findings

Several recent works on the dark triad have been investigated. The performance comparison of the methods with the proposed work is described in Table 5. Previous results focused on different focus points with diverse dataset acquisition and utilization techniques; most datasets were acquired through social media such as Facebook, Twitter, and email conversation. It is worth noticing that the methods of this work address only the dark triad traits, achieve the best possible results, and outperform the other existing models in Accuracy and recall by complimenting several machine learning algorithms.

5 Discussion

The organization should consider each of its potential insider risks related to cloud services and determine whether and to what extent cloud components such as service level agreements (SLAs) would cover the risk from insiders in case of a provider's business failure. Early identification of insider threats through the implementation of the dark triad traits is a breakthrough innovation that can be administered in the pre-screening activities of cloud administrator recruitment. This section presents a theoretical insight that assimilates dark triad traits and dynamic user profiles that adhere to risk principles in ISO 31000.

Employee Screening By leveraging the insights provided by Table 4 in describing a possible combination of dark triad traits, we can harness the findings from Figure 2 to construct a thorough employee screening protocol. A meticulous set of preventive measures can be structured to span the pre-recruitment stages, during recruitment, and post-recruitment. Initiatives, such as crafting precise job descriptions, establishing early-stage criteria for identifying insider threat profiles and technical prerequisites, and conducting candidate evaluations that encompass both technical and soft skills bolstered by cloud certifications, have the potential to enhance these endeavors with heightened efficacy.

Within the recruitment phase, the process consists of scrutinizing candidates' applications and resumes, conducting comprehensive interviews to unveil their behaviors, and subjecting them to psychological assessments. These assessments are strategically applied, testing candidates against various pressure points relevant to the specific position they are seeking. Proceeding into the post-recruitment phase, diligent background checks and, where applicable, the rigorous pursuit of security clearances come into play. This sequential approach ensures that the organization remains steadfast in its commitment to security and integrity throughout the hiring process.

Dynamic User Profile and Risk Mitigation With reference to ISO 31000 [29] and elaborative source [31], the identification of malicious dark triad traits combined with records of violations and obvious behaviors that were obtained from the employee screening phase would be resourceful in taking insider threat detection to remarkable progress. It is also worth noting that insider risk could be triggered by maladaptive organizational behaviors in response to human resources that would cause an imbalance in employee treatment. We proposed stepwise creating a dynamic user profile and risk mitigation by considering dark triad traits. The objective is to minimize the possibilities of insider threat occurrences as effectively as possible in an organization.

- **Define the Traits** Each description from dark triad traits D_t is assigned a customary weight to reflect its significance and contribution to the overall risk. In this publication, we already defined 8 different personalities ranging from benevolent *Be*, narcissistic *Na*, psychopathic *Ps*, Machiavellian *Mc*, psychopathic narcissistic *Pn*, manipulative narcissistic *Mn*, anti-social *As*, and maleficent *Ml*, as described in Section 3.1.

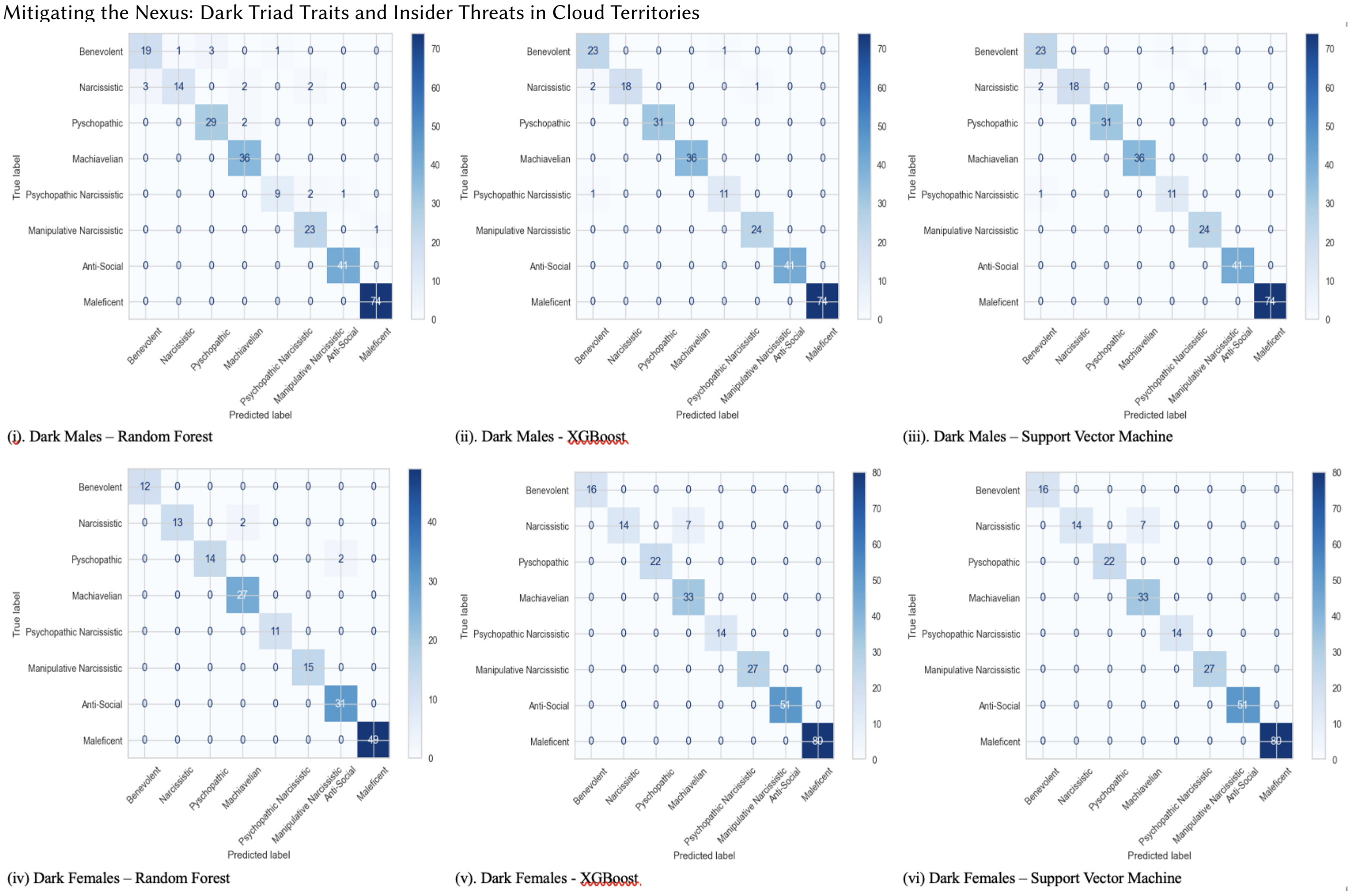
$$D_t = Be \cup Na \cup Ps \cup Mc \cup Pn \cup Mn \cup As \cup Ml \tag{5}$$

- **Establish the context** Determine the regular and irregular working times related to the cloud computing environment, as it can impact the organization's risk landscape, as shown in Equations 6 and 7.

$$\mathrm{Reg_{Time}} = \mathrm{B_{hr}} \cap \mathrm{Reg_{work}} \tag{6}$$

$$Irreg_{Time} = Shf \cup Ovr \cup ER \cup RW \cup Pj_w \tag{7}$$

The notation Reg_{Time} stands for regular time and is a product of the intersection of regular business hours B_{hr} and regular workday Reg_{work}. Irregular time, represented in $Ireg_{Time}$, is a product of a combination of Shift time *Sht*, Overtime *Ovr*, Emergency Response *ER*, Remote Work *RW*, and Project-Based Work Pj_w. along with the definition of the following activities that accumulated in Equation 8, validated from the paper [32] and paper[34]:



(i). Dark Males – Random Forest

(ii). Dark Males - XGBoost

(iii). Dark Males – Support Vector Machine

(iv) Dark Females – Random Forest

(v). Dark Females - XGBoost

(vi) Dark Females – Support Vector Machine

Figure 5. Dark Males and Dark Females Confusion Matrix for Machine Learning.

Table 5. Performance Comparison with Existing Publication on Dark Triad

Number	References	Year	Method Used	Gender	Accuracy (%)	Recall (%)	Precision (%)	F1-Scores (%)
1	Moskvichev et al [17]	2018	Random Forest	N/A	78.33%	78%	79%	78%
2	Yang et al [27]	2018	OCSVM	N/A	91.8%	N/A	N/A	N/A
3	Ahmad et al [28]	2020	BILSTM	N/A	85%	85%	85%	85%
4	Proposed Model	2023	ML with Random Forest	Males	93.15%	91.47%	88.77%	89.66%
				Females	97.72%	98.38%	96.77%	97.43%
			ML with XGBoost	Males	98.09%	97.01%	96.65%	96.74%
				Females	97.34%	97.81%	95.83%	96.30%
			ML with SVM	Males	98.09%	97.01%	96.65%	96.74%
				Females	95.81%	97.81%	95.83%	96.30%
			ML with Ensemble Boosting and Bagging	Males	99.23%	99.16%	98.80%	98.09%
				Females	96.96%	96.05%	95.05%	97.34%

- **Provisioning(P)**: Creating and configuring cloud resources, such as virtual machines, storage, and database
- **Deploying(D)**: uploading and launching applications or services on cloud infrastructure.
- **Managing(M)**: Handling and maintaining cloud resources, including scaling, monitoring, and optimizing performance.
- **Storing(S)**: Saving and retrieving data and files in cloud storage solutions, ensuring data availability and durability.
- **Analyzing(An)**: Processing and deriving insights from data using cloud-based analytics and machine learning tools and services.

$$N_aj = \frac{\alpha1|P_a^j| * \alpha2|D_a^j| * \alpha3|M_a^j| * \alpha4|S_a^j| * \alpha5|An_a^j|}{|A_u|} \qquad (8)$$

The number of activities in a given period for user a in a cloud computing environment j is denoted with N_{aj}. Where $|P_a^j|$ is provisioning activities from user a, $|D_a^j|$ is deploying activities, $|M_a^j|$ is managing activities, $|S_a^j|$ is storing activities, and $|An_a^j|$ is analyzing activities. $\alpha 1$ up to $\alpha 5$ are each respected weight calculated experimentally, such that the accumulation of the overall α equals 1. $|A_u|$ is the total number of user's activities.

- **Risk Identification** In this phase, we identify the possibilities of insider threat by defining the time span (short, middle, and long) in Figure 6 to observe the evolution of the employee's traits and behaviors over time, which would immediately mitigate the risk of insider threats, using scenarios and situations where alteration in insider behavior may pose a risk.

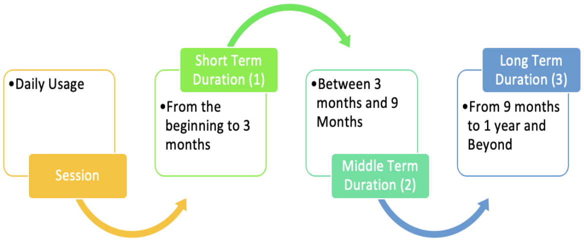

•Daily Usage
Session
Short Term Duration (1)
•From the beginning to 3 months
•Between 3 months and 9 Months
Middle Term Duration (2)
Long Term Duration (3)
•From 9 months to 1 year and Beyond

Figure 6. Time Span of Working Duration.

We represent various time spans spent by employees with notation t. For example, employees who have begun working for a short duration are expressed in $0 \leq t \leq 3$. Within the time spent in activities related to cloud computing, The user activity level engagement in a cloud computing environment PC_a^j can be obtained by formulating accumulated actions amount N_{aj} and a combination of regular Reg_{Time} and irregular time $Irreg_{Time}$ which incorporates Equations 7 and 8 into 9, adapted from a paper [33]:

$$PC_a^j = N_{aj} * (Reg_{Time} + Irreg_{Time}) \qquad (9)$$

- **Risk Analysis** In this phase, The ultimate objective is to obtain a formula that would separate expert users from regular users, assisting in insider threat evaluation of the difference of authority value possessed by both users. The sequences begin with determining several thresholds regarding the technical aspect of cloud computing and its users, which are described to explain risk analysis comprehensively, as follows:
  - **Authentication (a)**: Define how often specific actions or behaviors are expected such as user logins, access files, or perform particular tasks.
  - **Data Volume (v)**: Volume of data or transactions considered normal.
  - **Geographics or IP-based (g)**: Monitor user access from/to different geographic locations or IP Addresses.
  - **Sensitive Data Access (d)**: Define the threshold for accessing sensitive or critical systems, including uploading, transferring, and downloading.

  We can generate dynamic user profile DUP information after deciding on specific user recent activity information PC_a^j. The dynamic user profile in Equation 10 reflects recent preferences related to his / her engagement in a cloud computing environment related to the defined technical threshold by utilizing recent activities in Equation 9 and the description of the cloud computing threshold inspired by the paper [33].

$$DUP = PC_a^j >= a, v, g, d \qquad (10)$$

After determining the dynamic user profiles of insiders, we need to vectorize the target user and another user. We take advantage of cosine similarity SM to calculate the similarity of dynamic profile information of the two users DUP_a and DUP_b. The users with higher cosine values than the specified threshold are regarded as similar to the target user, as shown in Equation 11.

$$SM = \frac{\sum_{k=1}^{n}(DUP_a \cdot DUP_b)}{\sqrt{\sum_{k=1}^{n}(DUP_a)^2} \cdot \sqrt{\sum_{k=1}^{n}(DUP_b)^2}} \qquad (11)$$

To identify expert users whose activity exceeds the specified threshold and has an upper authority value compared with regular users, we describe several authority values to signify the significance of the roles sequentially as follows:

- **Administrators(Weight:4)**: Broad authority over cloud resources.
- **Power-users(Weight:3)**: Intermediate-level permissions, such as performing many actions, but still have restrictions, such as limited access to manage IAM users and resource utilization.
- **Regular User(Weight:2)**: Limited normal users can typically perform only limited and specific tasks or actions on given cloud resources.
- **Read-Only User (Weight:1)**: These users can only view AWS resources but cannot make any modifications.

We can formulate the authority value $Auth$ by multiplying weight a_q and value of authority value R_q, described in Equation 12.

$$Auth = a_q \cdot R_q \qquad (12)$$

Utilizing previous authority values, we can determine the formula for the expert user EU set which consists of the intersection of similar user SM and Authority value $Auth$ that exceed threshold d, explained in Equation 13.

$$EU = SM \cap Auth > d \qquad (13)$$

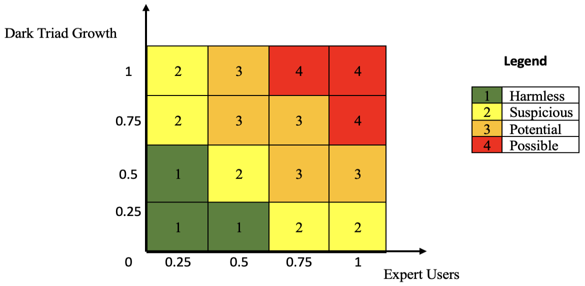
- **Risk Evaluation** Conducting several comparisons between risk analysis focused on authority value possessed by expert users in the cloud computing environment and the growth of dark triad traits in employee behavior. We defined an exponential growth or decay of Dark Triad traits ED_t by relating initial value D_t from Equation 5, where e is an Euler's constant that determines the percentage of decay or growth. Growth or Decay constant k and time in t notation is an extract from the time span spent by an employee in an organization, and every employee could experience a different kind of trait growth. Explained in Equation 14.

$$ED_t = D_t \cdot e^{(kt)} \tag{14}$$

Risk evaluation *Eval* could expressed as the intersection of expert users *EU* with exponential growth or decay of dark triad ED_t. Explained in Equation 15.

$$Eval = EU \cap ED_t \tag{15}$$

- **Risk Mitigation and Treatment** Conducting final mapping between expert users and Dark Triad traits growth. This paper uses several terms to describe the level of insider threats such as Possible (4), Potential (3), Suspicious (2), and Harmless (1) as referenced from an article [35]. Figure 7 describes the risk matrix between threat level, Expert user score *EU*, and Exponential growth of Dark Triad traits D_t ranging from 0 to 1.



Dark Triad Growth				
1	2	3	4	4
0.75	2	3	3	4
0.5	1	2	3	3
0.25	1	1	2	2
0	0.25	0.5	0.75	1

Legend	
🟩 1	Harmless
🟨 2	Suspicious
🟧 3	Potential
🟥 4	Possible

Figure 7. Risk Matrix of Expert Users and Dark Triad traits growth.

For each possible level of insider threat, the following mitigation suggestions are applied.

- **Possible**: Immediate suspension of user access, forensic investigation, and legal action.
- **Potential**: Temporary access limitation, internal investigation, Human Resource assessment.
- **Suspisious**: Continous log monitoring, access review, Human resource involvement.
- **Harmless**:Regular log monitoring.

Limitations The limitation of this publication is mainly limited to a single dataset with only 880 rows. To be comprehensively utilized, the dataset should contain even more samples of face templates from diverse human races worldwide and be supported with diversified features. Moreover, the real-world dataset would present multiple dimensions not limited to physical attributes but evidence such as historical logs and daily recorded conversations. To construct a complete employee screening, multiple means of approaches should be elaborated to create a precise result. We may be considering presenting improvements works of aforementioned aspects in the future.

The discussion section offers a theoretical approach by describing dynamic user profiles as elaboration aspects due to dataset feature limitations. We encouraged future works to intensively address the interconnection between dark triad traits and technical user profile features.

Future Works This work is a sequential activity that initializes from a risk assessment activity to address the multi-human-system perspectives of cloud computing. Extension to this work by considering N-shot learning due to the rare occurrence of insider threats in organizations and scrutinizing features of cloud continuity planning and incident handling in cloud disaster recovery. Ultimately, digital cloud capability and maturity dimension model formulation are in the scheduled timeframe.

6 Conclusion

This paper extensively explores insider threats in cloud employee screening, underlining that Dark Triad traits should not be seen as the exclusive indicators of insider risk. Our main goal is to introduce dynamic user-profiles and effective mitigation strategies for this complex issue. We specifically look at how machine learning can be used to screen male and female employees for dark triad traits, and we apply various methods to analyze the data. Our results show that a straightforward machine-learning approach can accurately identify dark triad traits in males and females. For males, the accuracy is an impressive 98.09%, and for females, it reaches a high of 97.72% using ensemble techniques, aligning with the traits outlined in the paper [30].

In addition, we integrate dynamic user profiles and mitigation strategies following ISO 31000 guidelines by addressing multi-dimension approaches. This framework divides mitigation activities into four stages: Possible, Potential, Suspicious, and Harmless, based on expert user skills and developing dark triad traits among employees. This approach enhances our understanding of and response to insider threats in cloud environments, ultimately improving security and risk management, in line with academic standards in the field.

Acknowledgments

This study was funded by the ICSCoE Core Human Resources Development Program and supported by the Laboratory of Cyber Resilience, Information Science Division, Nara Institute of Science and Technology, Japan.

References

[1] T. E. Senator et. al., "Detecting Insider Threats in a Real Corporate Database of Computer Usage Activity," in Proceedings of the ACM-SIGKDD Conference on Knowledge Discovery and Data Mining, page 1393-1401, ACM (2013)

[2] W. R. Claycomb, C. L. Huth, B. Phillips, L. Flynn, and D. McIntire, "Identifying indicators of insider threats: Insider IT sabotage," 2013 47th International Carnahan Conference on Security Technology (ICCST), Medellin, Colombia, 2013, pp. 1-5, doi: 10.1109/CCST.2013.6922038.

[3] Marc Dupuis and Samreen Khadeer. 2016. Curiosity Killed the Organization: A Psychological Comparison between Malicious and Non-Malicious Insiders and the Insider Threat. In Proceedings of the 5th Annual Conference on Research in Information Technology (RIIT '16). Association for Computing Machinery, New York, NY, USA, 35–40. https://doi.org/10.1145/2978178.2978185.

[4] Asaf Shabtai, Maya Bercovitch, Lior Rokach, Ya'akov (Kobi) Gal, Yuval Elovici, and Erez Shmueli. 2016. Behavioral Study of Users When Interacting with Active Honeytokens. ACM Trans. Inf. Syst. Secur. 18, 3, Article 9 (April 2016), 21 pages. https://doi.org/10.1145/2854152.

[5] M. Collins, "Common sense guide to mitigating insider threats," Carnegie Mellon Univ., Pittsburgh, PA, USA, Tech. Rep. CMU/SEI-2016-TR-015,2016.

[6] E. Ted, H. G. Goldberg, A. Memory, W. T. Young, B. Rees, R. Pierce, D. Huang, M. Reardon, D. A. Bader, and E. Chow, "Detecting insider threats in a real corporate database of computer usage activity," in Proc. 19th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2013, pp. 1393 1401

[7] L. Liu, C. Chen, J. Zhang, O. De Vel and Y. Xiang, "Insider Threat Identification Using the Simultaneous Neural Learning of Multi-Source Logs," in IEEE Access, vol. 7, pp. 183162-183176, 2019, doi: 10.1109/ACCESS.2019.2957055

[8] IBM Security (2020). Cost of a data breach report. Ponemon Institute, IBM

[9] N. Elmrabit, S. -H. Yang and L. Yang, "Insider threats in information security categories and approaches," 2015 21st International Conference on Automation and Computing (ICAC), Glasgow, UK, 2015, pp. 1-6, doi: 10.1109/IConAC.2015.7313979.

[10] M. J. Alhanahnah, A. Jhumka and S. Alouneh, "A Multidimensional Taxonomy of Insider Threats in Cloud Computing," in The Computer Journal, vol. 59, no. 11, pp. 1612-1622, Nov. 2016, doi: 10.1093/comjnl/bxw020.

[11] Greitzer, F., Purl, J., Leong, Y. M., and Becker, D. S. (2018, May). Sofit: Sociotechnical and organizational factors for insider threat. In 2018 IEEE Security and Privacy Workshops (SPW) (pp. 197-206). IEEE.

[12] Alsowail, R. A., and Al-Shehari, T. (2020). Empirical detection techniques of insider threat incidents. IEEE Access, 8, 78385-78402.

[13] Paxton-Fear, K., Hodges, D., and Buckley, O. (2020). Understanding Insider Threat Attacks using Natural Language Processing: Automatically mapping organic narrative reports to existing insider threat frameworks. In HCI for Cybersecurity, Privacy and Trust: Second International Conference, HCI-CPT 2020, Held as Part of the 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24, 2020, Proceedings 22 (pp. 619-636). Springer International Publishing.

[14] Shaw, E. D. (2006). The role of behavioral research and profiling in malicious cyber insider investigations. Digital investigation, 3(1), 20-31.

[15] Nurse, J. R., Buckley, O., Legg, P. A., Goldsmith, M., Creese, S., Wright, G. R., and Whitty, M. (2014, May). Understanding insider threat: A framework for characterizing attacks. In 2014 IEEE security and privacy workshops (pp. 214-228). IEEE.

[16] Schultz, E. E. (2002). A framework for understanding and predicting insider attacks. Computers and security, 21(6), 526-531.

[17] Moskvichev, A., Dubova, M., Menshov, S., Filchenkov, A. (2018). Using linguistic activity in social networks to predict and interpret dark psychological traits. In Artificial Intelligence and Natural Language: 6th Conference, AINL 2017, St. Petersburg, Russia, September 20–23, 2017, Revised Selected Papers 6 (pp. 16-26). Springer International Publishing.

[18] Ackerman, D., and Mehrpouyan, H. (2016, April). Modeling human behavior to anticipate insider attacks via system dynamics. In 2016 Symposium on Theory of Modeling and Simulation (TMS-DEVS) (pp. 1-6). IEEE.

[19] Alper, S., Bayrak, F., and Yilmaz, O. (2021). All the Dark Triad and some of the Big Five traits are visible in the face. Personality and Individual Differences, 168, 110350.

[20] Moshagen, M., Hilbig, B. E., and Zettler, I. (2018). The dark core of personality. Psychological Review, 125(5), 656–688. https://doi.org/10.1037/rev0000111

[21] Giddens, L., Amo, L. C., and Cichocki, D. (2020). Gender bias and the impact on managerial evaluation of insider security threats. Computers and Security, 99, 102066.

[22] Le, D. C., Zincir-Heywood, N., and Heywood, M. I. (2020). Analyzing data granularity levels for insider threat detection using machine learning. IEEE Transactions on Network and Service Management, 17(1), 30-44.

[23] Liu, L., Chen, C., Zhang, J., De Vel, O., and Xiang, Y. (2019). Insider threat identification using the simultaneous neural learning of multi-source logs. IEEE Access, 7, 183162-183176.

[24] AlSlaiman, M., Salman, M. I., Saleh, M. M., and Wang, B. (2023). Enhancing false negative and positive rates for efficient insider threat detection. Computers and Security, 126, 103066.

[25] Liang, N., Biros, D. P., Luse, A. (2016). An empirical validation of malicious insider characteristics. Journal of Management Information Systems, 33(2), 361-392.

[26] Garcia, D., Moraga, F. R. G. (2017). The Dark Cube: dark character profiles and OCEAN. PeerJ, 5, e3845.

[27] Yang, G., Cai, L., Yu, A., Ma, J., Meng, D., Wu, Y. (2018, March). Potential malicious insiders detection based on a comprehensive security psychological model. In 2018 IEEE Fourth International Conference on Big Data Computing Service and Applications (BigDataService) (pp. 9-16). IEEE.

[28] Ahmad, H., Arif, A., Khattak, A. M., Habib, A., Asghar, M. Z., Shah, B. (2020, January). Applying deep neural networks for predicting dark triad personality trait of online users. In 2020 International Conference on Information Networking (ICOIN) (pp. 102-105). IEEE.

[29] ISO, I. 31000: 2018 Risk Management—Guidelines. 2018. International Organization for Standardization: Vernier, Geneva, SUI.

[30] Paulhus, D. L., Williams, K. M. (2002). The dark triad of personality: Narcissism, Machiavellianism, and psychopathy. Journal of research in personality, 36(6), 556-563.

[31] Shaw, E. (2023). The Psychology of Insider Risk: Detection, Investigation and Case Management. CRC Press.

[32] Racherache, Badis, Paria Shirani, Andrei Soeanu, and Mourad Debbabi. "CPID: Insider Threat Detection using Profiling and Cyber-Persona Identification." Computers Security (2023): 103350.

[33] Jun, Yang, and Li Peilin. "A new method of group information recommendation based on the user dynamic profile information optimization." In 2021 7th International Conference on Information Management (ICIM), pp. 57-61. IEEE, 2021.

[34] Azzam, Fatima, Mohammed Kayed, and Abdelmgied Ali. "A model for generating a user dynamic profile on social media." Journal of King Saud University-Computer and Information Sciences 34, no. 10 (2022): 9132-9145.

[35] Abulencia, J. (2021). Insider attacks: Human-factors attacks and mitigation. Computer Fraud Security, 2021(5), 14-17.