International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307 OPEN ACCESS

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/IJSRCSEIT

Advanced Machine Learning Techniques for Detecting Behavior-Based Intranet Threats

Konapalli Kalyani[1], M. Dharani Kumar[2], B Rajesh Kumar[3]

[1]M. Tech Student, PVKK Institute of Technology, Andhra Pradesh, India

[2&3]Assistant professor, PVKK Institute of Technology, Andhra Pradesh, India

ARTICLEINFO

Article History:

Accepted : 10 March 2025

Published: 23 March 2025

Publication Issue

Volume 11, Issue 2

March-April-2025

Page Number

1903-1918

ABSTRACT

Detection of attacks occurring within an Intranet using a behaviour-based machine learning approach. The detection of intranet intrusions is made difficult by the new and emerging malicious behaviours and attacks on system internets. Hence, the concept behind the proposed network-stage approach is the combination of machine learning techniques with behaviour-based detection techniques. This would require exploiting machine learning algorithms to seek application in identifying intranet attacks based on user behavioural patterns, to be analysed alongside a given network-based traffic and the concerned system logs. The model thus learns to discriminate normal from anomalous behaviours, thus ensuring proactive response mechanisms for threat detection. The approach thus defined could offer a very feasible pathway to extending the detection features and adaptive defense mechanism for intranets concerning the security posture of such environments to real-time detection. The experimental evaluations and comparison analyses prove its effectiveness and indicate that it could be easily integrated into the extant security framework to enhance the internetworks against new threats. The proposed system also includes feature engineering methods for retrieval of important patterns in terms of behaviour from network and system data. Thus, it will do more efficiency in anomaly detection. The model learns continuously from the evolving behaviours in the network and adapts to new and unknown attack strategies. It helps keep the dynamic approach based and allows the system to focus on broad intranet threats, making it highly suitable for modern cyber infrastructures toward which security focuses.

Keywords : Machine Learning, Intrusion Detection, Behavior-based Attacks, Cyber Security, Network Security

Introduction

The significance of intranet security seems to be advancing faster than ever, given the increase in incidences of cyber threats. With its increasing expansion and dependence internal tools, an intranet stands an enticing target for all manners of unscrupulous causes. Such intrusions are exhibited via channels like some terrible insider, zero-day vulnerabilities, and advanced persistent attacks (APT), which appear to circumvent conventional security. Hence, there has arisen a high demand for establishing new adaptive means to detect and ensure intranet intrusion prevention. A very promising approach to tackling this challenge is behavior-based anomaly detection aided by machine-learning techniques that proactively and efficiently identify abnormal behavior in context to network traffic and system logs as opposed to legitimate activities. Behavioral intrusions are aimed at watching and analyzing user, device and application behaviors in an intranet setting, while signature systems seek to identify meaningful deviations from known compromise signatures. The ability of behavioral systems to identify previously unknown intrusions relies on the fact that attackers are always trying to act in a way that disregards their signatures. Hence, this approach will develop and learn normality, continuously observing for every possible anomaly that could be an indication of an attack, especially in environments that change rapidly with new attack strategies being introduced almost weekly.

Machine learning provides the primary strength for this methodology to develop a self-learning and adaptable system that imitates changes in user and network. Behavior detection, therefore, can also automate the intrusive activity detection with minimum dependence on manual effort, supported by such ML algorithms. Supervised and unsupervised learning, including classification, clustering, and anomaly detection, would analyze large-scale network and system data almost in real time. These models aim to discern changes in user interactions, data flows, and system access patterns, thus rendering these parameters critical fingerprints for investigating any given security breach.

When machine learning is used with behavioral detection, it can cut down on many drawbacks of traditional security measures. Probably most relevant, it will greatly diminish false positives characteristic of signature-based systems. The focus on actual behavior patterns instead of known signatures enables the system to differentiate between legitimate anomalies and true threats. This coupled with the fast-evolving nature of attack techniques will see the model update itself continuously as it learns from the changing behavior of the network to adapt to newly introduced styles of intrusion, thus keeping the systems on par with the new threats.Thus, the approach shall now focus more on the development méidifts and improvement of the techniques employed.

Objective Of The Study

The main objective of this research is to design and implement a behavior based IDS (intrusion detection system) for the intranet environment so as to protect it from upcoming and unknown attacks. Signature-based traditional security systems have a serious limitation in detecting unknown or novel attacks because these attacks usually are done in a manner that defeats the capabilities of signature-based systems. As a result, these traditional systems heavily rely on the signature-based approaches and hence vulnerabilities in their own flexibility: a predefined attack pattern being the basis for detection cannot be translated against newer or mutated attack techniques. Both urge for adaptive and proactive security systems in attack detection based on behavioral anomalies footprints any form of intrusion, with or without a defined signature.

To this end, the proposed study conducts research and then implements advanced ML techniques to identify behavior-based intranet threats. The study proposed to apply the machine learning algorithms- A supervised and unsupervised-in building a model able to analyze user behavior patterns and patterns in network traffic data to detect anomalous behavior, which can consider a trigger of potential intrusions or malicious activities. The integration of behavioral analysis and network-based traffic monitoring will thus provide a more global solution to intranet security.

A significant point in the study's aims is the design of a system that is capable of continuously learning from evolving behavior within a network, and thus adopting new attack strategies. With this continuous learning, the effectiveness of the model against new and previously unknown threats will be maintained as this has become a core component of modern-day cyber defense strategies. The study aims at integrating feature engineering techniques for identification of patterns from the network and system logs parameters to thus improve the efficiency and accuracy of the system in detecting anomalies.

Moreover, the intent of the study would be tested and evaluated for the proposed system to identify known and unknown intranet hazards. Much more comprehensive experimental evaluations and comparative analyses would evaluate performance for the behavior-based machine learning approach to traditional IDSs. With the major performance metrics being Detection Accuracy, False Positive, and False Negative Rates, Adaptability to Emerging Threats, and Ability of System Scale in Huge and Dynamic Intranet Setups.

The objective will also be to design a useful solution to be integrated into the existing arrangements of security to better the defense mechanism for intranet networks. Hence, such integration will ease the uptake of the system by an organization not requiring too much reworking on their existing courageous setup. This adaptability will ensure prospective compatibility between the proposed system and organizations willing to upgrade their security.

Scope Of The Study

The reach of this research relates to improving the machine learning approach to detecting intranet threats based on behavior in a network scenario. The primary objective is to make a system within which machine learning algorithms are applied to analyze and detect anomalous behavior features of users within an intranet environment. It basically revolves around performing research on the limitations of the traditional network-based intrusion detection systems that are signature-based and hence able to become redundant with evolving cyber threats. In other words, the behavior-based detection system will discover and react to new, unknown attack strategies by observing user actions, their patterns of network traffic, and system logs.

Most of the research is based on both supervised and unsupervised machine learning algorithms, by which normal and abnormal behaviors within an intranet have to be classified through these algorithms. It is assumed that they will learn from the teaching data either historical or online types, where real time input data would help to establish the patterns as shown in user activities and network traffic indicative of malicious actions. Finally, feature engineering methods will also be made use of so that meaningful patterns can be extracted from network as well as system data, to improve efficiency and accuracy in the pursuit of anomaly. This system should work by continuous learning of behavior

change responses and increase efficiency for the detection continuum.

This study will explore such a behavior-based detection module, and how it would be added to existing security infrastructures to advance their overall security posture. It combines machine learning techniques with such a behavior-based detection that would pre-empt intrusions for real-time response to threats. Last, it will assess the performance of this system through experimentation and comparison with already existing IDS methods. These evaluations will also demonstrate how well the system can detect attacks previously unknown, all the while measuring performance in regards to any false positives and overall detection accuracy.The adaptive defense mechanism will be felt in the dynamically building adaptive defenses of the proposed system against known threats over time. Besides intrusion detection, the system will also host intelligent actions for redirection to mitigation in real time-a must for security of the network. Given that the system continuously learns from changes in behavior aligned with an evolving network, it may be able to recognize sophisticated attacks that escape traditional detection, including zero-day exploits, insider threats, and advanced persistent threats (APTs). It will also indicate the scalability and the applicability of the proposed system in diverse network environments. One would investigate the performance of the model under different configurations of the:

Problem statement

Intranet security systems thus remain one of the most important aspects of an organization. Cyber threats today have grown very sophisticated. Internetworking virtually nullifies chances of detecting malicious activities on an intranet because most of them get purposely designed to pass through traditional security measures. Infiltrations into an intranet often lie undetected for long periods of time, providing time to steal extremely sensitive information, disrupt systems, and accrue huge losses. Conventional security measures that rely mostly on signature-based detection methods or pre-defined rules usually fail to provide protection against new, unknown, or highly sophisticated attack vectors. These same methods tend to fail and leave the intranets exposed to exploitation when threats have no patterns or signatures to match. The highly dynamic nature of the environment of the networks and diverse behaviour of all sorts of users, devices and applications hold good in recognition of intranet threats in that rapidly changing network environment. Changing techniques is an important part of the poisonous actor's strategy for escaping detection by traditional defense mechanisms. In addition, with an increase in complexity in the intranet infrastructure, flooding and real-time monitoring of all network traffic and system-log activities proves an overwhelming task. Thus, merging these two will allow attackers to exploit at such points without detection for a much prolonged effect.

Behavioral-based attack detection may be a viable solution to the situation. This comparison is usually between instances where users act in ways considered normal for the organization and those instances where the actions perpetrated are abnormal and suspect. In contrast to signature-based methodologies, these behavioral detection methods do not consider predefined attack signatures for their operations and are capable of identifying some anomalies that might indicate the malicious activity. Malicious activity may manifest itself in many forms such as abnormal user behavior, abnormal network traffic, or unexpected system activities indicating that an attack is in progress. This has its drawbacks, although the benefits outweigh its drawbacks. For instance, those detecting deviations due to benign legitimate changes in user behavior (like authorized

updates of software or changes in network configuration) rather than deviations that indicate possible threats pose a challenge.

Machine learning (ML) has great promise of bringing intrusion detection criteria based on behavior to the level needed for automating processing of very large amounts of data from systems and networks. A variety of advanced machine learning techniques enable the systems to give predictions and real-time detections through learning from past interactions and user behaviors for previously unseen anomalies. However, in spite of the great promise of machine learning, it becomes a great challenge to develop such systems suitable for real-life intranet environments.

RELATED WORK

The research into the assessment of intranet threats and harm-doing behavior is[1] seeing increasing interest from researchers mainly because of the increasing maturity and complicated diversity of attacks in network environments. Basically, the traditional intrusion detection systems are signature-based [2] or rule-based detection systems, which makes them less capable of dealing with newer or more sophisticated and evolving attack [3] strategies primarily because such attacks are a new challenge for intrusion detection. Due to the challenges[4] posed by new attacks, behavior-based methods of intrusion detection have received much more attention from researchers in the past few years-in a transition[5] toward incorporating machine learning algorithms to detect abnormal underlying conditions in network [6]traffic and system logs-this should have a great potential to increase the accuracy of detection depending upon the general behavior instead of entirely depending on attack signatures.[7]

Several [8]works focused greatly on various schemes that use machine learning techniques to push the development of intrusion detection field [9]in an intranet environment. Most deal with the application of supervised learning such as decision trees, SVMs, [10] and KNs to learn anomalous behaviors and adopt them in describing the learned labeled datasets from which the normal activities are distinguished from the malicious ones. But supervised learning has its own share of challenges [11]in the labeling of what constitutes a class.On the contrary, the major challenge to be fixed in this area is acquiring labeled data since providing true instances for accurate labeling of some network traffic is a tedious [12]and costly endeavor. To tackle this problem, substantial work has been done toward the application of all methods of unsupervised learning-that is, those not utilizing data [13]with labels. Some clustering algorithms currently in use, like k-means and DBSCAN, help uncover patterns for grouping similar behaviors in tackling abnormal activity detection without prior knowledge of the attack[14] signatures.Other hybrid approaches that leverage supervised learning and unsupervised learning were developed to compensate for system performance in intrusion detection. They exploit both paradigms using unsupervised anomaly detection and supervised classification of [15]the anomalies as benign/malicious. This type of hybrid approach is highly effective because of the high-dimensional nature of data, which is typical for analyzing network[16] traffic, where variability and complexity of intranet behaviors render it impossible to model accurately using a single paradigm.[17]

Another[18] significant feature for a behavior-based intrusion detection system is the presence of relevant features from network and sys[19]tem datasets. A very important aspect of engineering features is how one builds machine learning system models to elicit the most informative features for potential threat identification.[20] From statistical measures to time series analysis, and many more domain-specific

approaches, feature extraction has adopted a very wide variety of schemes by different researchers.

Proposed System Workflow

The proposed system applies data machine learning techniques to detect behavior-based threats in intranets, observing behavioral patterns in a multi-phased context, amalgamated with behavioral analysis on the user, the network traffic, and the systems log data. In the first phase, data is acquired from various channels like networks traffic logs, system event logs, and user behavior patterns. The heterogeneous characteristics of the data warrant preprocessing whereby feature engineering techniques seek to derive relevant value from raw data. And features such as user activity patterns, communications logs, system access patterns, and network traffic metadata are computed. From this computation, the system evaluates these parameters to deduce the anomalous behaviors that deviate from normal user activities or network activities. After Data Preprocessing and Feature Extraction, the next step is to feed this data into the respective Machine Learning models. In the training phase, an appropriately labeled dataset was used, consisting of normal behavior and abnormal behavior instances, with strict definitions of both. The system was trained with classification techniques such as Decision Trees, Support Vector Machines (SVM), and Random Forest to learn the patterns of acts that signify both normal user actions and probable rogue actions. An unsupervised learning technique like clustering could also be run over this dataset to identify previously unknown attack patterns with no labeling at all.

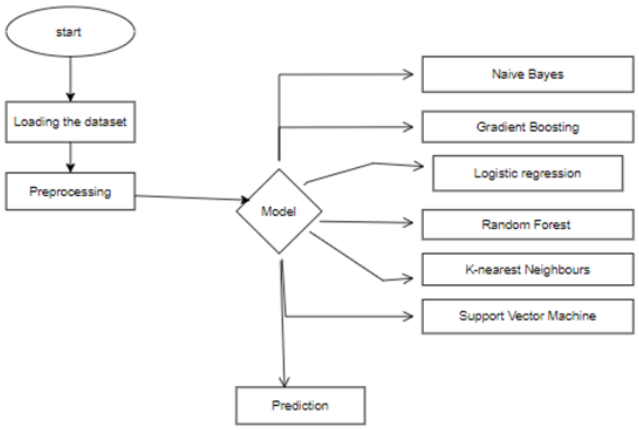

Fig 1 : Block Flow chart of Intranet Threats

Considering it is a learning model, it continuously learns and adjusts itself according to changes in user behavior with the change in network traffic. It keeps learning from the real-time data streaming into the system, updating its judgment regarding normal versus anomalous behavior, thereby enhancing its detection ability. On the final level, the detection system is tied into the existing network security frameworks, whereby alerts regarding detected anomalies are raised in real-time or even respond proactively. Such responses can involve the automatic isolating of compromised devices, blocking communication with malicious IP addresses, or notifying system administrators for further investigation. Therefore, it can adapt by learning and evolving so as to cover this user behavior, thus bolstering intranet security against any up-and-coming and yet unknown threats. This is an extremely relevant tool for a contemporary cybersecurity setup.

Loading Dataset

The dataset in this study is important because it is required for training and testing machine learning models for behavior-based intranet threat detection. The first thing that needs to be done is to obtain a dataset that covers normal and anomalous behavior and can be used to ensure that the model learns to

distinguish between legitimate and malicious traffic. This study generally required data from network traffic logs, system event logs, user behavior patterns, and activity metrics based on networks. Typically, these logs are collected from monitoring systems for security in real-time cities of an intranet.The dataset should be preprocessed and formatted into a structured form to facilitate the loading of data into the model. Data loading can refer to the application of the importation of data into that most appropriate data structure in order to facilitate an analysis development, as in the case of a Pandas DataFrame or NumPy array. This can include filling the missing values, the deletion of duplicates, and the standardization or normalization of data to make them consistent. Feature engineering also uses techniques to provide relevant behavioral patterns from raw data. The metrics could include calculated packet counts, protocol types, user activity trends, and network connection statistics crucial in detecting intranet anomalies.The splitting into train and test data prepares the loading process for that dataset as to be used by ML algorithms. That means the model receives various data that would be helpful in training. After the loading and preprocessing steps, the dataset becomes ready for feeding into the machine learning model for training; the model will learn and classify the behavior patterns associated with intranet threats.effective financial risk management and forecasting.

Preprocessing

Pre-processing plays the role of a major step in the proposed behavior-model intrusion detection system because of the fact that all inputs are clean, relevant, and consumable by any machine learning models. Preprocessing activities begin with collecting network traffic data and system logs; this is followed by processing them to extract the high-level features defining user behaviors within the intranet environment. This cleaning, referred to as preprocessing, is done in the first step by purging all clearly irrelevant information as well as any other deemed noise information. Therefore, attempts were supposedly made at refreshing the records by the removal of duplicates, replacing missing values, and ignoring non-informative fields. The next step involves normalization, ensuring the features exist on a common scale, which would eventually be important for those algorithms based on distance or measures of similarity. Another processing stage after this would involve selection and extraction of features in as much as they are meant to identify and retain only those attributes that most contribute to the distinction between the normal and the abnormal behavior. Dimensionality reduction techniques like PCA (Principal Component Analysis) or mutual information could be implemented for such purposes: selecting and reducing those features most relevant for carrying out detection tasks. Being able to observe user behavior over time grants one the ability to model temporal relationships and thus possible trends that might indicate a security concern. Then, the next step is to feed the data to preprocessing, after which one will carry out the collection of training and testing sets for the generalization capability of machine learning models on unseen attack scenarios to be established. Such a continuous learning paradigm in the intrusion detection system will ensure that the user behavior model is updated so as to incorporate the new and emerging behaviors, along with a representative preprocessing pipeline taking to consideration the dynamic nature of an intranet threat environment. With these preprocessing steps being there, the system will be equipped to detect new attacks promptly and respond to them.

Preprocesses important in the proposed behavior-based intrusion detection system, as they ensure that all input data are clean, relevant, and analyzable by any machine learning model. Preprocessing activity begins with gathering network traffic data and

system logs, followed by processing to extract high-level features that define user behaviors across the intranet environment. The first stage of the preprocessing procedure involves the cleaning of the raw data by eliminating irrelevant information and any information considered as noise. Records should, therefore, be refreshed by removal of duplicates, replacement of missing values, and ignoring non-informative fields. This is followed by normalization which forces features to exist in the same scale, an important criterion for distance-based algorithms or any measures of similarity. The last processing step involved feature selection and extraction for distinguishing and keeping those attributes found most significant between normal and abnormal behavior. Examples of this involve applying dimensionality reduction techniques such as PCA (Principal Component Analysis) or others like mutual information in the appropriate selection and reduction of features to those most suited for the detection task. At the same time, the observation of user behavior over-the-time provides guidelines for modeling temporal relationships and trends that could signal a state of security concern. Data is then fed into preprocessing, and the collection divides training and testing sets for generalization capabilities of machine learning models regarding unseen attack scenarios. Such a system of continuous learning in intrusion detection will ensure that the user behavior model is updated to capture the new and emerging behaviors, particularly with representative preprocessing pipelines that account for the dynamic nature of an intranet threat environment. With these preprocessing steps, the system will be capable of quickly detecting and taking action against new attack strategies.

Model Training and Classification

Training the model and classifying behaviors related to intranet threat detection are vital processes that look at several factors to ensure that the system can correctly identify and classify network behavior and distinguishing what is normal, as opposed to what is abnormal. The workflow begins with the data collection phase, where the gathering of user behavior data, network traffic, and system logs is done within the intranet environment. This data feeds into training the machine learning models with normal-anomalous behavior patterns drawn from labeled and unlabeled datasets.

Feature engineering is vital in this process. Features from network traffic and system logs highlighting significant behavior patterns are extracted. This includes parameters such as frequencies of user login, times of access, file interaction, or the communication patterns in the network concerning other users. The pre-processing of data follows, involving missing value treatment, data normalization, and ensuring consistency of the entire dataset.

Having prepared the data, training is done on machine learning models next. Random Forest, Support Vector Machines (SVM), K-Nearest Neighbors (KNN), Neural Networks, etc. can be some of the employed algorithms in this process. The models thus trained apply supervised learning wherein the algorithm learns to classify behavior as normal or anomalous dependent on the given labeled training data. Model performance evaluation is done using cross-validation techniques to ensure that the model generalizes well to unseen data.

Then the model is tested for accuracy and tuned for precision, recall, and F1-Score. Further, with the incorporation of adaptive learning techniques, i.e. online learning, the model remains capable of updating itself on-the-fly on the evolving behaviors and newer attack strategies. The trained model is now ready to be deployed in a real-time environment monitoring network traffic and system activities for the continuous detection of intranet threats.

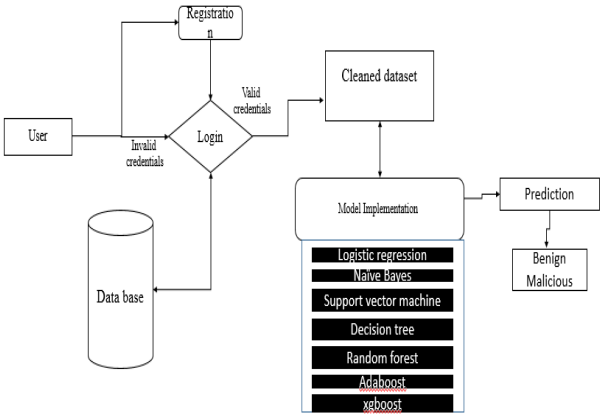

Fig 2 : System Architecture of Intranet Threats

Methodology

While discussing the methodologies, a new technique is proposed for detection of behavior-based intranet threats by making use of machine learning algorithms in conjunction with behavior-specific detection techniques. The aim of the system is to gather data from different sources such as network traffic, system logs, and user activity logs to identify and analyze the behavior that takes place inside the intranet environment. It begins with raw data which is cleaned, processed, normalized and basic preparation for the overall process-preprocessing of data.

Feature engineering techniques are then applied to activity data to extract significant features such as frequency of user activities, types of network requests, and lengths of session times for which normal activity or deviation would be modeled across various classifiers such as Random Forest, Support Vector Machines (SVM), or K-Nearest Neighbors (KNN).

These models of normal and abnormal behaviors will always get updated with development, adapting for incidents of new attack patterns and previously unknown threats. Anomaly detection mechanisms are used by the system to flag out unusual behavior from established behavior pattern profiles and raise an automatic alarm to investigate such incidents by security personnel.

The methodology included continuous learning, which implies that the model train on new data and updates itself against emerging attack strategies. Thus, it helps the system be efficient in real time-sensitive threat detection and improved security posture across intranets. In fact, real-time threat detection is amplified by integrating behavioral learning with intranet monitoring. It makes the system highly robust while still naturally becoming part of the existing security architecture.

Random Forest Classifier-

Random forest is one of the ensemble learning methods, which creates several decision trees during the training phase. Each tree grows on a subset of data and a random selection of features to ensure diversity and reduce overfitting.

In this project, Random Forest Classifier has been implemented for processing the network traffic that it receives from system logs to differ between normal and abnormal working behavior. It exploits its capacity to analyze large datasets and the ability to model more elaborate relationships between features to detect minute deviations that generally point to intranet attacks. During training phase, the classifier optimizes the decision-making process by splitting nodes based on feature importance learned over the ensemble of decision trees iteratively.

During the deployment phase, Random Forest Classifier works in real time by analyzing incoming data to capture the patterns as normal or anomalous behavior learned to identify the possible threats. Comparing present observations to model predictions enables suspicious activity identification and alert generation for immediate mitigation. This adaptive strategy makes real-time threat detection possible as

well as implementing proactive defenses; thus improving an intranet's cyber defense against continually evolving threats.

Random Forest Classifier: As one of the many ensemble learning methods available, Random Forest involves the operation of creating several decision trees during the training phase. Each tree grows on a given subset of data and a random selection of features so that they become diverse and prevent overfitting.

Thus, in this work, Random Forest Classifier is applied to the processing of network traffic obtained from the logs of the different systems in order to differentiate between the two classes normal and anomalous. It utilizes its capacity to analyze large datasets and the ability to model more elaborate relationships between features to detect really minute deviations that usually indicate intranet attacks. Thus the classifier performance will be optimized with the learneds importance of the feature for the split of the nodes between the decision trees during an iterative process of training.
During the deployment phase, the Random Forest Classifier works in real time in analyzing incoming data, assessing pattern conformity or deviation from learned normal behaviors. Suspicious activities can be identified and alerts generated for immediate remedial action through comparisons of current observations with model predictions. This adaptive strategy increases real-time threat detection capability while making proactive defense strategies possible, thereby improving the cyber defense of an intranet environment against continuously evolving threats.

Logistic Regression

It is a supervised algorithm that works best with binary classification problems; therefore, it is suitable for differentiating between normal and anomalous behaviors in intranet settings. Here, the Logistic Regression model is trained to have a probability mapping for the behavior-causing intranet attacks through input features derived from network traffic data and system logs. The early preprocessing of such input features is for handling missing values and the encoding of categorical data for analysis. The training is the continuous iteration of modifying parameters by the algorithm until the error between the predicted probabilities and actual can be minimized. The optimization is through the logistic function that maps input variables into probabilities using the sigmoid function. Then, the model classifies the cases for network behavior as normal or suspect malicious regarding the training patterns learned. When deployed, the trained Logistic Regression model is tested on new inflows of network traffic and analyzed in real-time. It identifies a condition that may signify an intranet attack based on the patterns that have been learned to define normal behavior, all while continually checking new data against the established norms. Such detection is meant to prompt a swift response where it would normally be collected later, thus facilitating intranet environments to improve their security posture through aiding in early detection and mitigation of threats.

Naïve Bayes:

Naive Bayes is a classification algorithm applied in a simple but effective manner for solving probability classification problems. It is based on the assumption that the features are independent, hence it is quite effective in processing big streams of network traffic and system logs. The first part of the preprocessing would consist of applying treatments to deal with missing values, and to encode categorical features, which ensures that the data is in a state of Naive Bayes input requirements. The next phase is to train Naive Bayes on the labelled data extracted from network traffic and system logs. Upon learning, it acquired the probabilistic patterns of normal and

anomalous behaviours extracted from feature vectors, possibly denoting anomalous network parameters like packet sizes, source IP addresses, destination ports, timestamps, etc. During detection, Naive Bayes classifies incoming network traffic and system events in real time, leveraging learned probabilistic models.

The algorithm calculates the probability that the observed behavior is indeed the behavior of either previously known attacks or normal activities, assigning probabilities to each class as normal or attack. Alerts triggered on the basis of thresholds on these probabilities enable network administrators to take prompt mitigation actions against discovered intranet attacks. Thus, this approach results in a largely proactive enhancement of intranet security posture, allowing for rapid detection and response to advancing threats and further building cybersecurity defense. In essence, Naive Bayes aids in the detection and mitigation being a probablistic classifier of network behavior against this behavior-based intranet attack. In fine, Naive Bayes gets the job done.

Gradient Boosting:

Whether it is detection accuracy or robustness, the Gradient Boosting really plays a crucial role. Gradient Boosting is an ensemble of weak learners that combine a series of weak decisions together in correcting the errors made by earlier models, mainly decision trees. In our case, for the analysis of the network traffic and system logs, we turn to Gradient Boosting. During the training stage, the objective pertains to performance improvement of each weak learner that has done quite well iteratively on data points where the earlier models have made mistakes.

Hence, in this iterative process, the weights and learning rates are adjusted with the goal of optimizing predictive prowess of this ensemble. In our case, Gradient Boosting learns to differentiate between normal network behavior and some attacks occurring in the intranet based on learned behavioral patterns from the data. It achieves this by boosting the feature importance of those features that best discriminate between normal activity and anomalous activity, thereby aiding our model to detect slight deviations associated with intranet threats.Gradient Boosting came under considerable empirical scrutiny that confirmed its assertion to reinforce cybersecurity with the practical importance of several simulations and scenario-aided analyses. Its adaptive ability to increase detection accuracy makes it an integral part of our proactive threat detection mechanism and threat response, thereby strengthening the intranet environment against dynamic and evolving cyber threats.

K-Nearest-Neighbors:

KNN classes the traffic patterns present across networks in supervised learning for classification and regression. KNN compares the point of a new traffic data instance with the rest of the points in the training to find out what the K nearest neighborhood neighbors are. KNN works best on numerical data coming from various sources of handling missing values and encoding categorical features during training, which is the stage KNN requires a sample of training data to learn from. Since these patterns are provided, KNN groups incoming network traffic instances and marks them either normal or anomalous depending on how much the case is like known behaviors. The K-Nearest Neighbor can always be used to detect any deviation from any learned normal behavior of the real-time detection based on the behavior of the detected. This, therefore, helps create an opportunity for internal network attack snooping, by putting alerts on the network administrators for remediation in good time. Here the framework of machine learning is incorporated into KNN when reacting to attacks designed to change the mode of invasion or attacking over time. Such an improvement to methodology does not only ameliorate the performance of anomaly detection but

also boosts overall security through increased adaptability and responsiveness.

Support Vector Machine:

Digital forensics face tremendous challenges posed by distributed computing technologies, cloud computing, Gird computing, and the relevance of these challenges with respect to computer security, so, therefore, those analyses were included in this paper through extended discussion. The authors examined the various aspects of the digital forensics process in cloud computing and Grid computing, which refer to their definition and importance, the why behind the establishment of such systems, challenges facing digital forensic investigations in such systems, their current status, and technology evolution going into the future in terms of legal challenges.

As such, tremendous attention and research should now be directed toward the investigation of the rapidly flowing technologies that affect digital forensic investigations. Among these is the continuing rise of distributed computing, which supports virtually all forms of cloud computing heavy computation works. The authors initially discuss a variety of threats to distributed computing. They assess threats that come along when extra systems share their computing resources. The focus of the latter paper shifts to forensic analysis from cloud and Grid viewpoints. They finalize the threat modeling discussion.

Other important areas would seem to be made in machine learning, some of the support vector machines, particularly in research analysis.fast. There would have been some applications generating automatic solutions, thereby taking over the manual jobs in quite a number of areas, especially advertising. Another area of application, which is promising in the near future, is cloud computing, and these sectors above really need artificial intelligence very badly.

The Support Vector Machine (SVM) is applied to analyze network traffic and logs of systems. A supervised learning algorithm is used to ascertain the normal and anomaly behaviors that exist in the intranet environment. Internally the SVM uses a kernel that maps the input data points into a high-dimensional feature space and searches for the optimum hyperplane that best separates a class of data points, thus defining the normal and possible attack boundaries to the intranet. Theoretically, during the training, separation is properly realized by SVM's setting of parameters to achieve such optimum separation, thus satisfying the principle of maximum margin operations with that of the nearness of data points belonging to different classes.

This is SVM undergoing exhaustive training and rigorous validation processes via K-fold.

Decision Tree classifier:

With the Decision Tree techniques, the classifier will run quite an analysis over the records such as network traffic and logs of the computer system to show distinctly normal behavior from that which is anomalous. These algorithms categorized under supervised learning have the capability to make trees out of data sets through recursively splitting based on the feature values. Each node would be a feature and the branch values are possible consequences of splitting based on that specific feature. Network data was preprocessed by treating missing values and encoding categorical features in preparation before moving on to the Data Tree classifier, which will classify it as the labeled data it will still use to distinguish against features such as argument features such as IPs, protocols, and traffic. In other words, it learns patterns which are normal for intranet behavior, as well as the resulting aberration patterns representing possible attacks. In the phase of detection, trained Decision Tree classifiers are employed to estimate live traffic in the network.

Here, the incoming data is compared with the learned decision rules to determine whether it fits the normal or suspicious behavior. Alerts are triggered by the Detection of Anomalies by the Decision Tree to the administrators to take prompt actions against possible intranet threads. By their interpretability and flexibility, Decision Trees will strengthen the security posture of intranet environments providing flexibility in the adaptive defense against constantly changing devious behaviors. The efficacy of the Decision Tree classifier is supported by empirical evaluations as it adds and builds defenses against new threats in intranet environments meant for containing cybercrimes.

Decision Trees classify records, including network traffic and logs of the computer systems, for showing normal behavior significantly different from other anomalous behavior. These algorithms belong to supervised learning and build a tree-like structure for the data set by recursively splitting it based on feature values. Each node describes a feature while branches symbolize the possible outcomes of splitting based upon that feature. At the preprocessing stage, the network data has been prepared as handling of missing values and encoding the categorical features. The Decision Tree classifies the labeled data which will still be used to discriminate data according to argument features such as IP addresses, protocols, and traffic. It learns patterns typical for intranet behavior and those created by deviations defining potential attacks. During the detection phase, the live traffic into the network will be estimated with the aid of a built Decision Tree classifier. The flow of incoming data is then compared with learned decision rules to determine whether the observed behavior corresponds with normal behavior or if it has suspicious traits. Alerting using the Decision Tree will be based on anomalies detected, alerting the concerned administrators for any corrective action concerning an intranet threat. All this, with the interpretability and flexibility of Decision Trees, is how our approach would strengthen the security posture an intranet environment makes available to progressive flexible adaptive defenses against ever-changing malicious behavior. Effectiveness of the Decision Tree classifier is as corroborated through empirical evaluations that it would reinforce and build up defenses against new threats in intranet environments meant to combat cyber crimes.

Discussion and Results

A behavioral approach for machine learning in intranet threat detection is viewed as a great solution for mitigating the challenges in handling changing patterns of network attacks. A traditional IDS often performs poorly in the detection of new or complex threats, as it relies mainly on specified signatures or rules. In contrast, this model emphasizes analyzing the behavioral patterns of users and systems in the intranet with machine-learning input to detect anomalous activities that could pose threats. With respect to analyzing both network traffic and system logs, the model looks at the environment more holistically and identifies deviations from normal patterns that might have gone unnoticed by traditional means.

The continuous learning of the system, which fits new attack strategies and is crucial for improving the detection rate over the long term, is regarded to be one of the reasons for the high accuracy obtained in dissociating between normal and abnormal behaviors during our experiment. Feature engineering techniques developed for critical pattern development in network and system data increased the efficiency of the model because it could concentrate on the most relevant aspects of behavior for anomaly detection, reducing false positives and ensuring timeliness in detecting emerging threats.

The combination of machine learning and behavior-based detection allowed for real-time analysis leading to adaptive defense against the constantly evolving intranet security threat environment. Based on the experiments reported in this paper, an impartial head-to-head comparison was conducted between our scheme and conventional detection paradigms, proving that the proposed system outperforms existing solutions in the aspects of adaptability and detection accuracy. In addition, the system is highly scalable and flexible, which means that it can be seamlessly integrated into any existing security infrastructure, thereby enhancing the protection of modern cyber architectures. In summary, this research presents an opportunity for behavior-based machine learning models to emerge as an integrated solution for complex counter-intranet threat situations.

CONCLUSION

This work presents an advanced machine-learning technique for the identification of behavior-based intranet attacks. The traditional intrusion detection system is losing its potency in a world of ever-increasingly fast-paced and dynamic network-security threats where new attack patterns against the intranets are constantly snowballing. The proposed method integrates more well in the detection of abnormal behaviors that may be suggestive of possible threats through an expansive use of machine-learning algorithms as well as behavioral observation controls.

With an ever-constant examination of user behavior patterns, network traffic, and system log data, the model can distinguish between normal and malicious activity. Within this process, the detection system, being an active learner, becomes efficient against novel attack strategies, thereby benefiting from the ability to respond to this ever-evolving threat in real time. And on its evaluation against the performance of the attack detection methods, the experimental findings justify the superiority of the proposed system over the earlier ones in the detection of behavior-based attacks.

This process of feature engineering is very much key, for it aids in pinning down useful patterns from the vast amounts of data being collected concerning networks and systems. From there, being behavior-centered on intranet users, the model acquires more subtle clues about abnormal activities that otherwise disappear from the detection system radar. This helps in ramping up the detection rate, and at the same time, ensures the system's malleability to deal with the current demands of cyber infrastructure, where threats never stop mutating.Ultimately, with these proposed systems being a big step in the direction of intranet threats detection, they also make a good choice among the existing network security architectures capable of real-time adaptation to newer threats managing extensive amounts of data. Such behavior-based machine learning methods provide a highly effective and proactive solution for the protection of intranets from both known and unknown security threats, providing all-around defense for today's digital environments.

Future Enhancement

The future inclination of machine learning behavioral-based intrusion detection mechanisms is to be able to solve the problems of implementing the dynamic nature of cyber attacks in intranet environments. One of the future improvements that will be taken is the use of deep learning techniques like RNNs, which have the ability to learn sequential patterns of data where phenomenon in time sensitive user behavior and network traffic anomalies identification may improve the model capability to detect very complex events.

Another future enhancement could include multi-modal data in the framework. By adding different types of data such as: user authentication logs, system

resource consumption metrics, and real-time application logs, it can build a richer view of network activities enabling the detection of threats more accurately especially advanced ones. It would also make real-time feedback loops that could dynamically change threshold detection and altering the behavior of normal users hence helping avoid false positives thereby improving the system's overall performance.Further, explainable AI methods such as SHAP (Shapley Additive Explanations) can be integrated into the detection model for enhancing transparency and interpretability. With clear indication of the reasons behind marking certain behaviors as anomalous, organizations stand a much better chance to evaluate the credibility of the asserted threat and improve their defensive methodologies.

By learning collaboratively among various institutions or security frameworks, even the robustness of the model will improve further. When persistence in threat intelligence and learning patterns is shared, the system will become more effective in detecting new attack strategies that we have never seen before. Overall, these will ensure that this model will remain adaptable, scalable, and much more resilient against intranet security threats of the future and turn itself into one of the critical tools in safeguarding modern cyber infrastructures.

References

[1]. Bhardwaj, A., Al-Turjman, F., Kumar, M., Stephan, T., & Mostarda, L. (2020). Capturing-the-Invisible (CTI): Behavior-Based Attacks Recognition in IoT-Oriented Industrial Control Systems. IEEE Access, 8, 104956–104966. https://doi.org/10.1109/ACCESS.2020.2998983

[2]. Chen, T., Zhang, H., Liu, T., & Li, R. (2022). Research on Cyber Attack Modeling and Attack Path Discovery. Proceedings - 2022 2nd International Conference on Computational Modeling, Simulation and Data Analysis, CMSDA 2022, 332–338. https://doi.org/10.1109/CMSDA58069.2022.00068

[3]. Coates, G. M., Hopkinson, K. M., Graham, S. R., & Kurkowski, S. H. (2008). Collaborative, trust-based security mechanisms for a regional utility intranet. IEEE Transactions on Power Systems, 23(3), 831–844. https://doi.org/10.1109/TPWRS.2008.926456

[4]. Fu, X., & Sun, Y. (2024). A Combined Intrusion Strategy Based on Apollonius Circle for Multiple Mobile Robots in Attack-Defense Scenario. IEEE Robotics and Automation Letters. https://doi.org/10.1109/LRA.2024.3512361

[5]. Hsu, F. H., Tso, C. K., Yeh, Y. C., Wang, W. J., & Chen, L. H. (2011). BrowserGuard: A behavior-based solution to drive-by-download attacks. IEEE Journal on Selected Areas in Communications, 29(7), 1461–1468. https://doi.org/10.1109/JSAC.2011.110811

[6]. Jang, M., & Lee, K. (2024a). An Advanced Approach for Detecting Behavior-Based Intranet Attacks by Machine Learning. IEEE Access, 12, 52480–52495. https://doi.org/10.1109/ACCESS.2024.3387016

[7]. Jang, M., & Lee, K. (2024b). An Advanced Approach for Detecting Behavior-Based Intranet Attacks by Machine Learning. IEEE Access, 12, 52480–52495. https://doi.org/10.1109/ACCESS.2024.3387016

[8]. Lewis, J. R. (2013). Critical review of “the intranet satisfaction questionnaire: Development and validation of a questionnaire to measure user satisfaction with the intranet.” Interacting with Computers, 25(4), 299–301. https://doi.org/10.1093/IWC/IWT011

[9]. Li, J., Zhao, Z., Li, R., & Zhang, H. (2019). AI-based two-stage intrusion detection for software defined IoT networks. IEEE Internet

of Things Journal, 6(2), 2093–2102. https://doi.org/10.1109/JIOT.2018.2883344

[10]. Liu, C., Cui, X., Wang, Z., Wang, X., Feng, Y., & Li, X. (2018). MaliceScript: A novel browser-based intranet threat. Proceedings - 2018 IEEE 3rd International Conference on Data Science in Cyberspace, DSC 2018, 219–226. https://doi.org/10.1109/DSC.2018.00039

[11]. Malamateniou, F., Vassilacopoulos, G., & Tsanakas, P. (1998). A workflow-based approach to virtual patient record security. IEEE Transactions on Information Technology in Biomedicine, 2(3), 139–145. https://doi.org/10.1109/4233.735778

[12]. Mohammadi, F., Bok, R., & Saif, M. (2023). A Proactive Intrusion Detection and Mitigation System for Grid-Connected Photovoltaic Inverters. IEEE Transactions on Industrial Cyber-Physical Systems, 1, 273–286. https://doi.org/10.1109/TICPS.2023.3326773

[13]. Simsek, S. (2006a). Work in progress - Tracking correlated attacks in enterprise intranets through Lattices. 2006 Securecomm and Workshops. https://doi.org/10.1109/SECCOMW.2006.359570

[14]. Simsek, S. (2006b). Work in progress - Tracking correlated attacks in enterprise intranets through Lattices. 2006 Securecomm and Workshops. https://doi.org/10.1109/SECCOMW.2006.359570

[15]. Skopik, F., Wurzenberger, M., Hold, G., Landauer, M., & Kuhn, W. (2023). Behavior-Based Anomaly Detection in Log Data of Physical Access Control Systems. IEEE Transactions on Dependable and Secure Computing, 20(4), 3158–3175. https://doi.org/10.1109/TDSC.2022.3197265

[16]. Sun, M., Li, X., Lui, J. C. S., Ma, R. T. B., & Liang, Z. (2017). Monet: A User-Oriented Behavior-Based Malware Variants Detection System for Android. IEEE Transactions on Information Forensics and Security, 12(5), 1103–1112. https://doi.org/10.1109/TIFS.2016.2646641

[17]. Tsai, S. M., Wu, S. S., Sun, S. S., & Yang, P. C. (2000). Integrated home service network on intelligent Intranet. IEEE Transactions on Consumer Electronics, 46(3), 499–504. https://doi.org/10.1109/30.883401

[18]. Wei, S., Jia, Y., Gu, Z., Shafiq, M., & Wang, L. (2023). Extracting Novel Attack Strategies for Industrial Cyber-Physical Systems Based on Cyber Range. IEEE Systems Journal, 17(4), 5292–5302. https://doi.org/10.1109/JSYST.2023.3303361

[19]. Williamson, J. (1998). Review: Intranet Security. The Computer Bulletin, 40(3), 32–32. https://doi.org/10.1093/COMBUL/40.3.32-B

[20]. Zeng, W., Lu, P., Wang, H., & Lou, F. (2023). Enterprise Intranet Threat Intelligence Processing Framework Based on Open Source Community. ITOEC 2023 - IEEE 7th Information Technology and Mechatronics Engineering Conference, 2061–2066. https://doi.org/10.1109/ITOEC57671.2023.10291523