How can user behavior logs and machine learning models help in improving the Incident response strategies, a conceptual walk through

Pranith Shetty
Senior Information security and Risk lead
Cisco, New Jersey, USA
Pranith.shetty14@gmail.com

Abstract:
This study aims to dive into the usage of Machine Learning (ML) along with User behavior logs in Incident response-related measures, it introduces a cutting-edge methodology for reinforcing cybersecurity threat detection processes by harnessing the power of machine learning technologies. It delves into a detailed examination of user behavior patterns to establish a dependable framework for the preemptive detection of security vulnerabilities. Utilizing a blend of both supervised and unsupervised ML algorithms, this paper investigates the algorithms' ability to navigate through extensive datasets to pinpoint anomalies that diverge from the expected behavioral patterns. A critical aspect of this research is the rigorous process of extracting and choosing behavioral features that play a pivotal role in the precision of predicting security incidents. The document elaborates on an extensive experimentation phase carried out on heterogeneous datasets, covering a broad spectrum of real-life conditions, to assess the efficacy of the suggested ML models. The outcomes highlight the transformative potential of ML in bolstering threat detection systems, providing a forward-looking, preventative strategy for cybersecurity. Beyond contributing to scholarly discussions around the application of ML in safeguarding digital environments, this research offers actionable strategies for entities keen on amplifying their security postures in response to the dynamic nature of cyber threats.

Introduction

Incident response is a crucial process under (Security Operations Center (SOC). User behavior logs provide a detailed record of user activities, allowing organizations to spot potential security incidents before they escalate. By analyzing these logs, security teams can identify unusual behavior that may indicate a threat. For example, sudden access to unusual files or systems can indicate a compromised account. Therefore, user behavior logs are essential in preventing security threats.

[1] [2] [3] User behavior logs are used to detect and investigate security incidents. They provide a detailed record of user actions that help determine the root cause of a breach, identify the accessed data, and assemble a timeline of events.

To manage incidents effectively, organizations should review user behavior logs. These logs provide insight into the extent of the damage caused, help prioritize the response process, and guide mitigation efforts. Accurate and up-to-date logs are essential for improving incident response planning. Models when trained on pre-processed data sets relying on user behavior logs can help the firms be more proactive and predict frauds and attacks. These self-learning models can enhance threat detection and incident response capabilities. The following sections dive into the benefits, and challenges of each of the prerequisites for this strategy.

Benefits of User Behavior Logs

User behavior logs can provide valuable insights into user activity that can help detect potential security incidents, investigate breaches, and understand the impact of an incident.

1) Pre-incident analysis: Teams and systems can learn from user behavior logs, over some time. The log will contain out-of-band and suspicious activities. These logs will have enough information about potential threats and vulnerabilities that could emerge from anywhere in the network. By being on a constant lookout for suspicious user activities, firms can take preventive actions to prevent incidents before escalations
2) Post-incident or root cause analysis: User behavior logs are essential for investigating security incidents, holding users accountable, and monitoring compliance. These logs track user activity and access, enabling analysts to identify the root cause of incidents and prevent future problems. By communicating clearly about their importance, organizations can help keep data safe.
3) Forensic deep dive: To respond to an incident effectively, it is crucial to analyze user behavior logs thoroughly. These logs contain essential information about user activities, such as login times, file access, system modifications, and application usage. By examining this data carefully, investigators can piece together the sequence of the incident, assess its scope and impact, identify the root cause, and develop strategies to events leading prevent similar occurring in the future. This information is invaluable in enabling incidents from organizations to respond to incidents in a timely and effective manner.
4) Policy Compliance to IR controls: By monitoring user behavior records, security policies, laws, and industry standards can be more closely adhered to. Organizations can discover instances of non-compliance and implement corrective measures to limit risks by keeping an eye on user activities in comparison to set policies and guidelines.

The analysis of user behavior records aids refinement of incident response plans and techniques. Through the identification of common attack routes, gaps in monitoring capabilities, and flaws in security controls, companies can enhance their readiness for incident response and fortify their defenses.

All things considered, user behavior logs are critical for efficient incident response since they offer insightful information about user behavior, help identify security events, and reduce risks to protect company property and information.

Challenges with User behavior logs

[4] [5] User behavior logs present a number of challenges because of the volume of information that must be gathered, which necessitates big data databases and logistics. Additionally, because these logs are crucial for cybersecurity incident detection, there may be issues with their efficient use and storage. Among these difficulties are:

1) Extensive and intricate data: Logs of user behavior produce enormous volumes of data, frequently spanning several different systems and applications. It can be difficult to analyze this amount of data, especially in light of the complexity of contemporary IT infrastructures with their many interrelated systems and endpoints.
2) Normalization: Correlation and normalization of user behavior logs from various sources can be challenging due to their disparate forms and structures. Although it can be difficult, aligning timestamps, user IDs, and event descriptions across diverse logs is essential for efficient analysis.
3) Contextual Understanding: A thorough grasp of the environment in which users behave is necessary to interpret user behavior logs. When actions that seem suspicious are not properly contextualized, they can seem harmless, which can cause misunderstandings and poor incident detection. False positives, in which perfectly normal activity is reported as suspicious, or false negatives, in which real security incidents are missed, can result from the analysis of user behavior records. Fine-tuning detection algorithms and thresholds is necessary to maximize detection accuracy while reducing false warnings.
4) Privacy implications: User activity records may include private information about people's activities, which is cause for concern. Businesses must strike a balance between the requirement for security monitoring, user privacy, and data protection law compliance.
5) Quality compromise: Accurate incident detection depends on the quality and integrity of user behavior logs. Logs that are tampered with, incomplete, or erroneous can cause security incidents to go unnoticed or misconstrued, undermining the efficacy of security monitoring operations.
6) Financial and resource limitations: It takes a lot of time, money, and computing power to analyze user behavior logs. Many organizations might not have the tools or knowledge needed to use user behavior logs for incident detection in an efficient manner, which could leave gaps in their security posture.
7) Enterprise tool integration: Because of interoperability limitations, lack of standard formats, and compatibility difficulties, integrating user behavior logs with current security products and platforms can be difficult. Streamlining incident detection and response operations requires seamless interaction.

To tackle these obstacles, a comprehensive strategy incorporating automation, advanced analytics, cooperation between security teams, and continuous improvement of detection tactics is needed. Organizations can experiment and implement the IR tactic using user behavior logs by being mindful of the challenges and ensuring the appropriate regulations and controls are taken care of. By surmounting these challenges, entities can proficiently employ user activity logs to augment their cybersecurity stance and expeditiously identify and address security breaches.

Discussion and Challenges with using ML

[10] [11] If machine learning (ML) is to improve cybersecurity incident management, a number of obstacles must be overcome. These include:

1) Amount of data: For ML algorithms to train efficiently, they need plenty of high-quality data. But, for new or uncommon security risks, it could be difficult to get labelled data for training ML models. Data can also be noisy or missing certain pieces, which can cause models to be erroneous or biased.
2) Complex layout: The complexity of many ML algorithms–especially deep learning models–makes it difficult to understand how they arrived at their predictions. Security analysts may find it challenging to comprehend and verify model results if they are not interpretable, which might erode confidence in incident management systems powered by ML. Protecting ML-powered event management systems from malicious attacks is crucial for their dependability and security.
3) Outdated Models: ML models can get outdated real quick, Rapid change occurs in cybersecurity threat environments as adversaries tweak their strategies, tactics, and procedures (TTPs) all the time. Continuous updating and retraining of ML models is essential to keep their effectiveness in detecting new threats and changing assault trends.
4) Inherent Bias: The possibility of bias and unfairness in ML models raises concerns about the potential for unequal treatment or representation of different groups. If there is underrepresentation or a skewed distribution in the training data, it can influence the ML models and produce discriminating results. If we want ML-based event management systems that detect and respond to threats fairly and effectively, we must ensure that these systems are fair and eliminate biases.
5) Resourcing requirements: Significant computational resources, specialists, and infrastructure are typically necessary for training and deploying ML models for incident management.

There are gaps across sectors and industries when it comes to building and managing security solutions powered by machine learning.

Concerns about the processing of sensitive data, such as personally identifiable information (PII) and private company information, may arise when using ML-based incident management

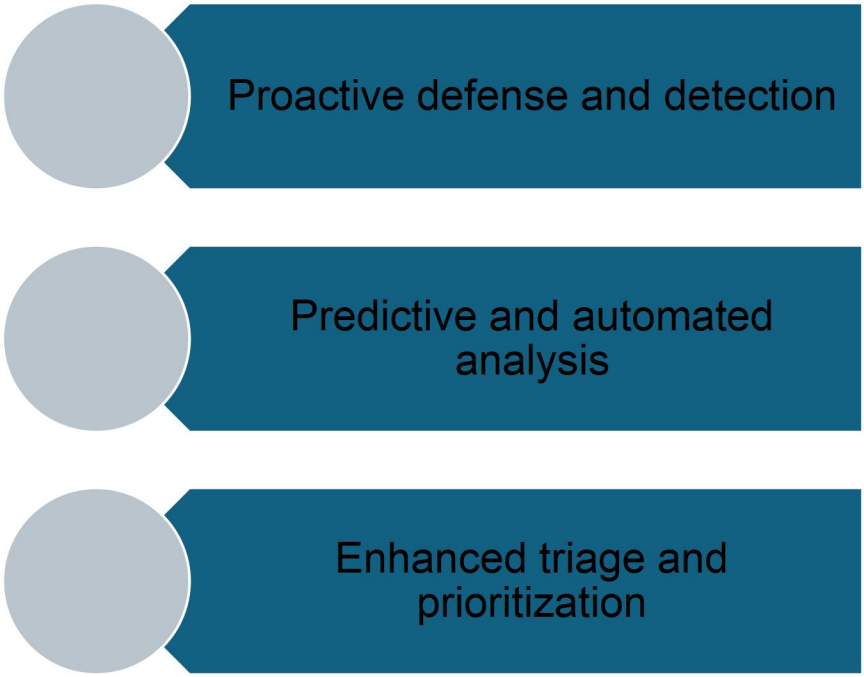
systems. 11. For ML models to be used in cybersecurity, privacy must be protected, and data protection laws must be followed. Protecting users' privacy when using machine learning for incident management can be achieved using techniques like federated learning and differential privacy.

It may be difficult to integrate incident management solutions driven by ML with preexisting security infrastructure and processes. When it comes to security, ML models aren't complete without integrating with SIEM systems, EDR platforms, and other technologies to help identify and respond to threats efficiently. Machine learning models also need to be scalable so they can deal with different contexts and large-scale deployments.

Strong, open, and ethical ML-based solutions for cybersecurity incident management are necessary to tackle these difficulties. This can only be achieved if researchers, practitioners, and lawmakers work together.

Effectiveness of using ML

[6] [7] [8] [9] Because machine learning (ML) has increased capabilities for detection, response, and mitigation, it can significantly improve cybersecurity incident management. The following are some ways that machine learning might enhance incident handling:



1) Proactive defense and detection: Machine learning algorithms are able to identify abnormalities suggestive of possible security problems by analyzing vast amounts of data from several sources, such as system logs, network traffic, and user activity. ML models are able to recognize abnormalities from typical patterns of behavior that could indicate dangers like malware infections, insider threats, or unauthorized access attempts. Behavioral analytics algorithms that can spot minute clues of harmful activity in huge

datasets can be developed thanks to machine learning. Machine learning algorithms are able to identify suspicious activity and unusual behavior that could point to security incidents by examining user behavior, system interactions, and application usage patterns. Incoming data can be analyzed and threat intelligence streams integrated using machine learning (ML) techniques to find trends or indicators of compromise (IOCs). Organizations may automatically correlate threat intelligence with their own security telemetry by utilizing ML algorithms, which will improve their ability to identify and address new attacks.

2) Automated and predictive analysis: Machine learning algorithms are capable of analyzing incident data in the past to spot patterns and forecast possible security incidents in the future. Organizations can proactively allocate resources, adopt preventative measures, and minimize risks before they become big security events by utilizing predictive analytics. With the help of machine learning (ML), adaptive security controls can be created that can dynamically modify security setups in response to shifting environmental factors and growing threats. Systems with machine learning capabilities can continuously analyze and modify security measures to reduce new threats and weaknesses. Response actions can be automated by ML-powered incident management systems on the basis of preset rules, regulations, or discovered patterns. Blocking erroneous network traffic, isolating infected endpoints, or sending alarms to security experts for more research are examples of automated reactions.
3) Enhanced Triage and Prioritization: By evaluating each security incident's severity, impact, and likelihood, machine learning algorithms can help with the triage and prioritization process. Machine learning (ML)-driven technologies assist security teams in allocating their limited resources to addressing the most pressing risks by automatically classifying and ranking incidents. Security monitoring systems can produce fewer false positive warnings with the aid of machine learning algorithms. Machine learning (ML)-driven incident management solutions can decrease the amount of time spent on non-threatening incident investigations and increase accuracy by continuously improving detection models through learning from past data.

All things considered, ML has the ability to completely transform cybersecurity event management by offering cutting-edge capabilities for detection, response, and mitigation. Organizations can improve their entire cybersecurity posture by using ML algorithms and approaches to improve their ability to detect and respond to security problems in real-time.

Convergence of ML and User Behavior logs

The utilization of machine learning (ML) in conjunction with user behavior logs can yield significant value across a range of applications, including personalization, recommendation systems, fraud detection, anomaly detection, and others. Below is a comprehensive methodology for effectively employing machine learning techniques with user behavior logs:

1. Acquire user behavior logs from many sources like websites, mobile applications, IoT devices, and others. The logs may contain data regarding user interactions, clicks, purchases, searches, timestamps, device details, location data, and any other pertinent contextual information.
2. Derive significant features from the preprocessed data that encapsulate crucial elements of user behavior. These attributes may encompass metrics such as the frequency of activities, duration of time spent on various sites, session length, device category, geographic location, and so on.
3. Develop, construct, and implement the trained models into operational systems to generate predictions or offer insights using real-time user behavior logs. This may entail incorporating the models into pre-existing applications, APIs, or data pipelines.
4. Consistently oversee the performance of the trained models, ensuring thorough examination, analysis, and preprocessing of the data sets.

Using machine learning with a training data collection of user behavior records offers several significant advantages:

Fraud Detection and Security: User behavior logs provide significant indicators that can be utilized to identify fraudulent actions and security risks. Machine learning models have the ability to acquire knowledge about typical patterns of behavior and detect any deviations or potentially suspicious behaviors in real-time. This capability is valuable for preventing fraud and strengthening cybersecurity measures.

Predictive Analytics: ML models can utilize historical user activity data to forecast future user activities, including purchasing behavior, likelihood of churn, and product demand. These forecasts facilitate proactive decision-making and focused actions to achieve desired results.

Conclusion

Incident response is crucial in multiple industries as it enables firms to not only prosper but also endure. Numerous instances exist where firms have experienced failure due to inadequate incident response protocols. The utilization and combination of User behavior logs alongside Machine learning methods in Incident Response measures will effectively enhance defensive strategies. The benefits of using this strategy outweigh the challenges, furthermore, the majority of those challenges can be conquered while being mindful of the laws, regulations, and overarching policies. This technique aims to implement a comprehensive and holistic risk response strategy for incidents. It will dismantle obsolete techniques and prove highly efficient due to its predictive analysis grounded in ongoing learning.

References

[1] Writer, R. L. C. (2023, December 6). *Enhancing Incident Response Playbooks With Machine Learning*. [Online]. Available: https://www.darkreading.com/cybersecurity-operations/automation-via-machine-learning-makes-cybersecurity-playbooks-better

[2] *The role of User Entity Behavior Analytics to detect network attacks in real time*. (2018, November 1). IEEE Conference Publication | IEEE Xplore. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8855782

[3] *Role of User and Entity Behavior Analytics in Detecting Insider Attacks*. (2020, October 20). IEEE Conference Publication | IEEE Xplore. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9292394

[4] *ACM: Digital Library: Communications of the ACM*. (n.d.). [Online]. Available: https://dl.acm.org/doi/fullHtml/10.1145/2076450.2076466

[5] *Advances and Challenges in Log Analysis - ACM Queue*. (n.d.). [Online]. Available: https://queue.acm.org/detail.cfm?id=2082137

[6] *AI-powered incident response: Harnessing the potential of self-healing endpoints - Atos*. (2023, September 4). Atos. [Online]. Available: https://atos.net/en/lp/detect-early-respond-swiftly/ai-powered-incident-response-harnessing-the-potential-of-self-healing-endpoints

[7] Mktg, S., & Mktg, S. (2022, May 19). *AI in Cybersecurity: Incident Response Automation Opportunities*. SISA. [Online]. Available: https://www.sisainfosec.com/blogs/ai-in-cybersecurity-incident-response-automation-opportunities/

[8] Doerrfeld, B. (2023, October 2). *Using ML to Accelerate Incident Management - Security Boulevard*. Security Boulevard. [Online]. Available: https://securityboulevard.com/2023/10/using-ml-to-accelerate-incident-management/

[9] *Machine learning approach to quick incident response*. (2020, June 1). IEEE Conference Publication | IEEE Xplore. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9141989

[10]Bailey, G. (2022, August 31). *The challenges of developing Machine Learning and statistical models - Riskaware*. Riskaware. [Online]. Available: https://www.riskaware.co.uk/insight/the-challenges-of-developing-machine-learning-and-statistical-models/

[11] Review, M. (2023, November 1). *7 Common Challenges In Adopting Machine Learning For Business*. Mirror Review. [Online]. Available: https://www.mirrorreview.com/7-common-challenges-adopting-machine-learning-business/