International Journal of Core Engineering & Management
Volume-6, Issue-12, 2021 ISSN No: 2348-9510

ADAPTIVE ML-DRIVEN IDENTITY VERIFICATION FOR PREVENTING ACCOUNT TAKEOVERS ON HIGH-VALUE ACCOUNTS

Hariprasad Sivaraman
Shiv.hariprasad@gmail.com

Abstract

Criminals target high-value accounts for Account Takeover (ATO) attacks, because the rewards (i.e., financial and proprietary assets) combined with sensitive personal data putting reputations of institutions at risk. Sophisticated attack techniques, such as credential stuffing attacks, phishing campaigns or automated bots to steal identities are rendering traditional static identity verification systems less and less effective. This paper proposes an Machine Learning (ML) driven adaptive identity verification framework to help target ATO risk associated with high-value accounts. It applies a multi-layered model to Behavioral biometrics, device fingerprinting, and contextual risk assessment with evolving thresholds that baselines user-specific behaviors and dynamically assessing elevating threat levels. The introduced system improves both accuracy and detection while reducing user friction with the help of ML-driven adaptive learning. This flexible method enhances detection and decreases false positives for suspicious account behavior, allowing organizations to secure their most important accounts without undermining user experience

Keywords: ATO, Identity Verification, High-Value Accounts, Machine Learning, Behavioural Analysis, Device Fingerprinting, Adaptive Security, User Experience, Authentication, Real-Time Monitoring

I. INTRODUCTION

Reliance on online accounts for products and services in different industries from finance to government, healthcare and e-commerce has grown due to the digital transformation of business. As, ATO compromises the accounts of high-value assets (financial, health, personal data & enterprise) grows to be increasingly prevalent and harmful. Using automated credential stuffing, phishing and social engineering attacks, attackers are bypassing traditional static means of authentication leaving organizations with little they can do to stop it.

To overcome these problems, an adaptive ML-driven identity verification framework may be employed for high value accounts. Combining real-time behavioural analysis, device fingerprinting, and contextual risk factors, this framework adjusts verification requirements dynamically based on user behaviour and the characteristics of a session. Adaptive models enable bespoke security responses to ATO threats which may not only increase success rates in preventing ATO attempts from succeeding but also facilitate minimal legitimate user interruption, thereby maximizing both effective security and an optimal customer experience.

This paper elaborates upon this adaptive ML-driven framework and its architecture and components as a system capable of re-envisioning identity verification on high-value accounts.

II. PROBLEM STATEMENT

Accounts with high value are preferred victims for ATO attacks because they provide a potential financial gain or access to sensitive information. Password, two-factor authentication (2FA), and knowledge-based authentication (KBA); are static, inflexible, and can be bypassed via a range of attack vectors. High-value accounts have to take both high security and user experience into consideration; Account takeovers need more strict security.

Moreover, these conventional methods find it difficult to mirror the evolving and revolutionary style of attack techniques. Intrusion Detection Systems (IDS) play a crucial role in safeguarding our computers against unauthorized intrusions by malicious hackers. To evade static defences, attackers are increasingly utilizing machine learning, automated bots and stolen credentials. All this leads to an immediate requirement for a verification mechanism that is fluid and adaptive in nature which needs to be able to continuously evolve with the behaviour of the user as well as new adversarial activities.

III. PROPOSED SOLUTION: ADAPTIVE ML-DRIVEN IDENTITY VERIFICATION FRAMEWORK

In response to these challenges, a multiple-layer identity verification framework is proposed that is driven by adaptive ML wherein each layer can potentially uncover specific risk associated with the identity.

1. **Behavioural Biometrics Layer**: Behavioural Biometrics Behavioural biometrics is about observing and capturing how a user interacts in terms of specific patterns, from the speed they type or mouse movement to interaction with the screen itself, to gestures on a mobile device. This type of continuous authentication is hard for hackers to imitate. The trained ML model creates a baseline profile of users based on their historical behavioural data and uses it for anomaly detection. For instance, if a user typically logs in by typing with specific intervals between keystrokes, an abrupt change in their typing behaviour, might initiate extra verification steps like 2FA prompt or prompted security question. Hence, Behavioural biometrics offers a continuous verification layer in a non-intrusive manner and sensitive to the context, assessing the authenticity of user behaviour during an entire session.

2. **Device Fingerprinting Layer:** Device fingerprinting is the identification of the unique characteristics of a user device like IP address, browser configuration, operating system and other hardware-specific identifiers. Every user session is assigned with a device fingerprint, through which the system can identify violations such as logging in from abnormal devices and geographical locations. If one tries to access from a new device, it will be detected by the ML model, and additional verification steps will be needed. The fingerprinting layer identifies the characteristics of devices and appends them continuously, resulting in every session being verified to be consistent with past sessions; providing an additional security against ATOs.

3. **Contextual Risk Assessment Layer:** If content risk assessment assesses an issue in a session, the contextual layer evaluates external factors that could be the underlying risk of a session related to location accessibility, timing of access and type transaction. For example, an access

attempt during abnormal hours or from a suspected territory can activate dynamic response according to risk assessment. With the additional context from Behavioural biometrics and device fingerprinting, the ML model is able to make much better predictions on whether an ATO attempt is actually occurring thereby resulting in dynamic evolvement for efficient defences causing less friction to the legitimate user. Further, it can dynamically change verification rules based on account activity and offer a robust review of activities associated with higher transaction values.

4. **ML-Driven Feedback Loop:** The feedback loop results seamlessly feeding the ML model to help train it at each instance with new data reducing risk of being unable to adapt if the user behaviour evolves or evolved threat vectors along with malicious networks. Every interaction is a part of a self-optimizing model that evolves operational threshold limits on risk based upon both real-time and historical data. The same user device does not have to exhibit bad behaviour, however three times within a ten-minute range should be considered suspicious for the model to increase the risk threshold. This ML model adapts when there are legitimate changes in a user's habits (such as if they opened the account on a new device post-upgrade) thereby false positives and customer frictions, are kept to a minimum.

IV. ARCHITECTURE OF THE ML-DRIVEN IDENTITY VERIFICATION FRAMEWORK

The architecture of ML-driven identity verification framework divides the entire procedure into modules where each layer could work independently as well as a part of adaptive security model.

Key components include:

- Data Ingestion Layer: It fetches the real-time data from each session, such as behavioural, device, and contextual input for the respective ML models. The pre-processing layer pre-processes inputs using data pipelines, such as normalization, encoding, and noise reduction.
- ML Processing Layer: Contains the ML models (behavioural biometrics, device fingerprinting, and contextual risk assessment) and applies individual data inputs to each model in real time. This layer is parallel processing enabled to reduce latency and provide instant feedback for high-volume systems.
- Risk Scoring Engine: This is where the output from each ML model gets aggregated using a weighted algorithm to derive a composite risk score. The weighted score is the sum of contributions from each of the models based on their importance level, which can be modified based on risk level and user preference.
- Adaptive Verification Layer: This layer decides what kind of verification measures to take based on the risk score. If the score goes above a certain threshold, secondary authentication methods are triggered for 2FA, or security questions must be answered, or biometric verification is performed.
- Feedback Loop and Model Updating Module: Iteratively improves every individual ML model with data from newer sessions, user feedback, and threat intelligence. This practical approach uses this module to provide real-time adjustment of model parameters i.e., the framework tends to adapt with respect to real-time changes in behaviour or attack pattern.

A. Operational Flow within the Adaptive Model

- Session Initiation: When a user starts a session, the framework starts tracking information

regarding the users such as the behaviour, devices, contextual data (location), etc. This data is then passed post pre-processing to the ML models.

- Risk Scoring Process: Each model works on its data and generates different risk scores. To illustrate, the behavioural model examines factors like keystrokes and device model investigates verifying the accessing device. A composite of these scores produces a final risk score for the session.
- Dynamic Response Generation: Based on the risk score, the appropriate security response is given. As long as the score does not exceed the set thresholds, the session continues uninterrupted. In the case of high-risk scores, an adaptive verification layer based on risk level encourages extra authentication such as requesting a code sent via SMS or email.
- Model Feedback and Optimization: After each session, the feedback loop records whether the processing was successful and tunes the model accordingly. This makes way for adaptive learning where the system is improving, at distinguishing between authentic users and potential attackers.

B. Advantages of ML-Driven Identity Verification

- Enhanced Security: The multi-layered approach backed by individual ML models can be highly sensitive to abusive actions, which would help provide a solid barrier against ATOs.
- User Experience Optimization: Adaptive verification reduces the disruption when a real user is attempting to access a service.
- Real-Time Adaptability: Due to the results of the feedback loop and threat intelligence integration, it enables real time changes in the model so that it can be adaptive against new user behaviours & threat vectors.
- Scalability: This style of architecture can easily be deployed in a scalable manner across cloud environments and is flexible enough to accommodate high-traffic systems and applications with fluctuating security demands.

V. USES

The ML-based dynamic identity verification framework proposed in this paper can find use cases across disciplines for high-value accounts that need extra layers of security. With Amazon Cognito, both finance, healthcare, government and retail, require a security layer that is responsive to individual user actions as well as robust protection against unauthorized access. For example, banks may leverage this system to safeguard financial data related accounts and healthcare providers may use it for securing patient information. This architecture specifically lends itself to an ML-driven approach that makes it ideally suited for organizations handling vast quantities of sensitive data requiring strict compliance with security regulations.

VI. IMPACT

The effectiveness of adaptive ML-based identity verification enhances both security and the user experience. Introducing ML-based dynamic thresholds and verification methods improves the ATO detection rates and also minimizes user friction for genuine users. Such a multi-layered approach creates stronger security, while ensuring the user journey is optimized; with verification measures tailored in real-time based on risk. Additionally, organizations have the opportunity to decrease the economic impact of ATOs with this methodology as it prevents unauthorized access

attempts that could result in data breaches and other responsibilities.

This framework not only meets data security regulations (such as GDPR and PCI DSS) from a compliance perspective, but also reduces the risk of sensitive user information being compromised. Implementing adaptive security measures, reduces the risk of data breaches and unauthorized access, which enhances compliance posture.

VII. SCOPE AND FUTURE DEVELOPMENT

Future improvements will add to the strength of the framework increasingly with the development of complex AI and ML algorithms such as deep learning models for capturing analogs of real-world behaviour sequences, and federated learning capabilities that allow organizations to share information on malware signatures they have been previously observed. Lowering the response time and improving model efficiencies backed by biometric verification such as facial or voice recognition along with edge computing allows for processing of requests in a fast and efficient manner. Making the model adaptable to function over various platforms such as web and mobile applications will be important in meeting changing security requirements.

VIII. CONCLUSION

This work proposes an adaptive ML-driven identity verification framework designed to meet user-centric high-assurance ATO prevention in high-value accounts. Along with integrating behavioural biometrics, device fingerprinting and contextual risk analysis; this framework is addressing identity verification at scale while evolving against the threat landscape and user patterns. Adaptive learning continuously refines the model to ensure accuracy with false positives being very rare. The future work must extend the framework with intelligent AI techniques and cross-platform compatibility for more effective and robust detection to protect high-value accounts from ATO assaults.

REFERENCES

1. Verma, A., & Gupta, K. (2020). Machine Learning in Cybersecurity: A Comprehensive Review. ACM Computing Surveys.
2. Robinson, M., & Tan, L. (2020). Federated Learning for Threat Intelligence in Financial Security. Proceedings of the 27th Annual Network and Distributed System Security Symposium.
3. Chen, L., Xu, X., & Wang, Y. (2020). Real-Time Adaptive Authentication Based on Behavioural Biometrics and Contextual Data. IEEE Access.
4. Zhang, S., & Li, F. (2020). Device Fingerprinting Techniques for Secure User Authentication. Journal of Cybersecurity Technology.
5. Kumar, R., & Patel, T. (2020). Reinforcement Learning for Continuous Optimization in Cybersecurity Applications. IEEE Transactions on Cybernetics.
6. Park, H., & Lee, S. (2019). Adaptive Machine Learning Techniques for Authentication in High-Security Environments. Computers & Security.
7. Martinez, R., & Chang, H. (2019). Behavioural Biometrics for Continuous Authentication Using Machine Learning Approaches. IEEE Access.
8. [8] Johnson, T., & Wu, P. (2018). Using Machine Learning for Dynamic Risk Scoring in

Financial Security Applications. Journal of Financial Crime.
9. Smith, M., & Torres, E. (2018). An Overview of Machine Learning Techniques in Real-Time Fraud Detection Systems. Journal of Information Security and Applications.
10. Feng, L., & Yang, Z. (2017). Adaptive Security Models in Cyber-Physical Systems Using Behavioural Analytics. International Journal of Cyber Security and Digital Forensics.