Implementing Artificial Intelligence (AI) in the Security Operation Center (SOC)

Le Minh Nguyen
Riaa Sehgal

ABSTRACT

With the rapid advancement of technology, there is also an advancement of the threats that lurk in cyberspace, hence creating a demand for advanced security measures to safeguard digital infrastructure. SOC (Security Operation Center) stands as the frontline of the defense in detecting and combating those threats. The research focuses on tools and technical framework that assists SOC analysts during triage analysis, enhancing their detection capability while also reducing alert fatigue, an issue every security engineer faces. The paper focuses on practical application of the framework in SOC's workflow like reducing log monitoring workload, a large part of tier 1 SOC responsibility.

Techniques such as Holt-Winters Forecasting and Long Short-Term Memory (LSTM) networks are employed to detect complex threats, including malware with long-term dependencies and Denial of Service (DoS) attacks, outperforming traditional signature-based intrusion detection systems (IDS). Framework like A^2C that utilizes DL (Deep Learning) algorithm that allows customizable automation level, with each level suiting specific stages of triage analysis.

By combining cutting-edge AI methodologies with collaborative approaches, this work proposes a transformative SOC framework capable of proactively mitigating risks and adapting to evolving cyber threats. The findings underscore the importance of continued innovation to address current limitations and ensure robust digital ecosystem security.

INTRODUCTION

The swift advancement of cyber threats has made the creation of sophisticated safeguards for digital ecosystems necessary. Security operation centers (SOCs) must use cutting-edge tools to preserve resilience and reduce risks as threat actors adopt more complex tactics. A critical focus of this research lies in optimizing human-AI collaboration within SOC operations. By enabling dynamic adjustments to AI automation levels, SOCs can tailor AI support to specific tasks, ranging from the categorization of vast volumes of log data to proactive threat hunting by Tier 3 analysts. Emphasizing self-awareness (SA) and shared situational awareness (shared SA) further enhances this collaboration, enabling both human analysts and AI systems to operate more effectively. Techniques such as inverse reinforcement learning (IRL) are explored to refine the AI's understanding of SOC analysts' decision-making processes, ensuring improved recommendations during triage analysis.

This research proposes a technical framework that enhances traditional intrusion detection systems (IDS) by integrating AI-based threat detection, leveraging anomaly-based approaches that surpass conventional signature-based methods. The models are trained using extensive network traffic data and past security incidents, significantly improving detection capabilities. The deployment of deep learning (DL) models, including recurrent neural networks (RNNs) and autoencoders (AE), is examined for their ability to detect malware characteristics within systems. Additionally, frameworks such as A^2C and IRL are proposed to empower SOC analysts, facilitating efficient threat monitoring and response through accurate data and actionable recommendations. By advancing AI-driven methodologies and fostering human-AI collaboration, this research aims to revolutionize the operational efficiency of SOCs, enabling them to stay ahead of adversaries in the evolving landscape of cybersecurity threats.

RELATED WORK

Anomaly-based detection implements DL (Deep Learning) algorithms like Holt-Winters Forecasting stated in [3] which successfully defends against network-based attacks. The algorithm uses statistical analysis to calculate traffic patterns and keeps track of network processes. The model is capable of adjusting the standard threshold for incoming traffic based on seasonality, meaning it adapts to a certain volume of data depending on a business's operation time, significantly reducing false-positive cases that could lead to alert fatigue in SOC's environment. Holt-Winters Forecasting technique uses exponential-smoothing to predict traffic during a period, it includes the calculation of three components: Baseline, Linear Trend, and Seasonal Trend.

$$y_{t+1} = a_t + b_t + c_{t+1-m}.$$

The variable Baseline (a***t***) shows the average volume of data at any given time of the day. The Linear Trend (b***t***) represents changes in data volume resulting from organizational changes like network growth or policy changes. Finally, Seasonal Trend (c***t***) shows the fluctuations of data during the day, and (c***t***+1-m) is the prediction of the seasonal coefficient based on the previously computed coefficient within the same period of the cycle. The future forecast of the data will be (y***t***+1) which is the addition of all three components. With the data standard threshold set up, Holt-Winter Forecasting begins detecting anomalies within the system. Multiple variations of DoS (Denial of Service) attacks like UDP flooding, and ICMP flooding where attackers repeatedly send packets to the host machine on a specific port, forcing the host to endlessly generate replies for these requests, consuming the target machine's resources and bandwidth to the point of system shut down. The algorithm uses a range to determine if the specific deviation in network traffic should raise an alert, this is achieved by constructing a traffic range for decision-making.

$$\hat{y}_t \pm \delta \times d_{t-m}.$$

The interval is calculated with the addition and subtraction of predicted traffic with the multiplication of scaling factor value (δ), meaning the confidence level of data traffic falling within the normal threshold with seasonal deviation (d***t***-m). After constructing the confidence interval, the residual is computed to determine if anomalies are detected in the traffic at a certain time (***t***) using the following formula.

$$r_t = | \ y_t - \widehat{y}_t \ |$$

Suppose the traffic falls within either LOW or HIGH categories, meaning it lies outside the confidence interval. In that case, the DL algorithm will raise flags and escalate this alert to the administrator for further investigation. Overall, Holt-Winters forecasting demonstrates its effectiveness in detecting unknown malicious traffic due to its adaptability to scenarios that might challenge conventional signature-based detection methods. With the algorithm itself not requiring a database of information about predefined attack signatures presents itself as a more cost-effective solution in an environment that values scalability and real-time performance.

PROPOSED SOLUTIONS

A challenge that every machine learning development encounters is its practical implication within the context of business daily operations. To successfully utilize machine learning capability in processing large volumes of data, appropriate training is essential to both the AI model and the SOC analysts to create a dynamic security environment. In [4], the framework A^{2}C offers SOC analysts flexibility in adjusting the AI's level of automation like *Full Automation, Selective Deferral, and Collaborative Exploration*, each can be applied to the triage analysis depending on the process stages. In the process of analyzing network traffic, a combination of anomaly-based detection and an A^{2}C framework could reduce the mental toll on the analyst's side. The *Full Automation* mode automatically validates any traffic that falls within the expected pattern, any difference within the data flow will generate an alert and get redirected to tier 1 or 2 analysts for further inspection.

In case the AI system encounters uncertainty in the decision-making process, the alert gets deferred to human experts for a deep analysis. This is possible due to Shared SA (situational awareness) being implemented within the model, recognizing its capability to understand the SOC's limitations and collaborating with humans to address each side's weaknesses. This feature also comes into play for 3 SOCs during threat-hunting activities, instead of performing automation, the AI system collaborates by providing accurate recommendations and understanding the information needed throughout the process. Enhancing threat hunters' capability in spotting vulnerability through collaborative exploration, supports an active security environment where analysts seek and mitigate threats before a disaster happens.

Anomaly-based detection mechanism presents a significant advantage in comparison to conventional anti-virus software advertised by companies like Microsoft, Mandiant, and Mcafee that employ signature-based detection, which reveals its shortfall against unknown threats. In [1], Deep learning algorithms like RNNs (Recurrent neural networks) were implemented to detect sequential data and store them in a hidden state, allowing it to identify differences in patterns between the computed data and information in its storage. However, RNN suffers from issues like *vanishing gradients,* the model struggles to create complex temporal patterns over an extended period.

The solution to this is integrating LTSM (Long Short-Term Memory) networks within the anomaly malware detection process. The model addresses RNN's weakness by incorporating memory cells and a gating mechanism to store relevant information in the long term, any changes that happen over an extended period would be detected. LTSM uses a forget gate, input gate, and output gate to determine which information should be stored or discarded, utilizing limited storage efficiently. The input gate decides which data in the current input should be written into the memory cell. Additionally, data in the previous hidden state and current data can be used by the input gate to calculate potential values for pattern prediction. Forget gate discards unnecessary information in the current input based on the previous cell state, keeping the memory accessible for the next input. The output gate selects which information to be exposed for calculation of the next memory cell state.

Overall, LTSM is an improved variation of RNN that effectively deals with long-term dependency issues using memory cells and gating mechanisms. While RNN can be useful for simple tasks, LTSM surpasses it when managing complex sequential patterns. This improvement in the security infrastructure enhances the intrusion detection system's capability and significantly reduces the number of threats to critical infrastructures.

LIMITATIONS

While anomaly-based detection systems, especially those employing deep learning (DL) algorithms like Holt-Winters Forecasting and LSTM networks, represent an important advancement in cybersecurity, they also have limitations. These limitations affect their usefulness and practical application in real-world situations, particularly in dynamic and complicated network systems.

One of the primary limitations of anomaly-based detection technologies is the risk of model bias and false positives. These systems rely heavily on training data to identify patterns and detect anomalies. The model may acquire biases that lead to incorrect identification if the training data is not extensive or varied. For instance, a false positives would result if unimportant actions that slightly migrate from the usual pattern could be reported as threats. In addition to wasting resources, this makes SOC analysts more prone to alert fatigue. Overwhelming false positives can erode system credibility and increase the likelihood that these analysts might overlook real threats.

Another difficulty is integrating AI-powered anomaly detection technologies with current SOC procedures. A lot of AI models are created separately, emphasizing computational efficiency above real-world operational usefulness. This lack of connectivity between the SOC analysts and the AI systems, can cause problems and would end up forcing SOC analysts to modify their methods to accommodate this technology rather than vice versa. Another significant issue is the scalability and flexibility of these models. As networks grow and evolve, the detection system must adapt to increased traffic and new attack methods. However, many AI models struggle to scale effectively, particularly when faced with high volumes of data or rapidly changing attack vectors.

Despite their promise, anomaly-based detection technologies face limitations that must be addressed for widespread, effective adoption. These limitations furthermore prove the necessity for continued innovation and optimization. Future advancements focusing on improving adaptability, reducing resource requirements, and incorporating seamless integration with SOC processes would amplify their practical utility in defending against evolving cybersecurity threats.

CONCLUSIONS

The incorporation of advanced deep learning (DL) techniques and generative AI into Security Operations Centers (SOCs) represents a revolutionary step in strengthening cybersecurity defenses. Anomaly-based detection techniques, such as Holt-Winters Forecasting and LSTM networks, can identify complex threats like malware with persistent dependencies and DoS attacks, helping SOCs to surpass traditional signature-based systems. These methods enable SOCs to spot anomalies in network activity and precisely handle new attack avenues. This research highlights the importance of human-AI collaboration within SOC frameworks.

By employing tools like A^2C and shared situational awareness (SA), AI systems can dynamically adjust automation levels and further support analysts in diverse tasks such as log categorization and proactive threat hunting. This reduces alert fatigue, enhances efficiency, and ensures a proactive approach to identifying vulnerabilities. Despite their promise, these systems face challenges, including false positives, scalability issues, and integration with existing SOC workflows.

In conclusion, using these AI-driven processes combined with collaborative frameworks can revolutionize SOC operations. Addressing current limitations through ongoing innovation will enable SOCs to safeguard against evolving cyber threats, and protect digital ecosystems with robust defenses.