**CYBER THREAT HUNTING**

JOSHUA ASIACHI MURUNGA

FACULTY OF SCIENCE EGERTON UNIVERSITY NJORO

COMPUTER SCIENCE DEPARTMENT

REG. NO.: SP13/21799/14

EMAIL: [joshmurush@gmail.com](mailto:joshmurush@gmail.com)

**ABSTRACT**

**It is no secret that expecting security controls to block every infection vector is unrealistic. For most organizations, the chances are very high that threats have already penetrated their defenses and are lurking in their network. Pinpointing such threats quickly is essential, but traditional approaches to finding these needles in the haystack often fall short. This is so due to the fact that most organizations still focus more of their resources on prevention rather than detection by deploying tools such as firewalls, anti-spam, sandboxing, intrusion prevention among others. Now, there is a unique opportunity for more feasible, more effective threat hunting capabilities, and it stems from rethinking the approach to wide area networking. This can be achieved through Software-Defined Wide Area Networking, which is a cloud computing service suited for the threat hunting process. The primary notion of threat hunting is to proactively search the network for threats that have evaded the existing security measures. Security experts around the globe have acknowledged that organizations can get infected no matter how good their security controls are.**

***KEYWORDS -*** Threat Hunting, Indicators of Compromise (IoC), Software-Defined Wide Area Networking (SD-WAN), Malware, Ransomware, Cloud, Cloud Service Provider

1. **INTRODUCTION**

Infection vectors change rapidly and continuously. Cyber criminals are constantly evolving and their tactics and techniques are increasingly changing with an alerting rate. The attackers use new delivery methods from social engineering to zero day exploits. With an almost equal rate of technological advancements and with everyone rushing to “upgrade” to the latest technology that exists, cyber criminals are somewhat being offered a bigger playing field. Cybercrime has even found its way to third world countries whose government have been rolling out its services through online portals. These portals have been gateways for cyber criminals who have siphoned cash from the national government without being realized. When the crimes were noticed, it was far too late. This is a classic example of how threats go unnoticed when attackers break through the primary security tools that only create a perimeter around the system and does not account for what happens when attackers slip through the defense.

To prevent such cases from constantly occurring, we delve into threat hunting, a post infection measure from the cyber kill chain, that aggressively intercepts, tracks and eliminates cyber adversaries. Having cyber threat covering a wide scope, security teams may seem to be staring down at a huge disadvantage. The dark web, being a haven for cyber criminals to share bypassing and hacking tools and scripts, there has become an increase in the number of malicious programs. To fish out these programs from an already infected system requires the task of analyzing system data, looking for anomalies. Most security professionals lack this particular expertise in data analytics. Likewise, data analysts lack the necessary security expertise. A combination of the two is required to effectively carry out the threat hunting process, which can be achieved by monitoring and analyzing system logs promptly, both by automation and human analysis, which helps identify the indication and threshold that can be used in threat hunting (Rasheed, Hadi, & Khader, 2017) . Malicious programs such as ransomware, which has been dominating the threat landscape, require a timely detection, which absolutely depends on how fast and accurately system logs can be analyzed and mined to hunt for anomalies and stop the intrusion. Ransomwares utilize different infection vectors ranging from social engineering and spam emails to botnets for distribution. A great technique in the detection of ransomware is to use sequential pattern mining to find the maximal frequent patterns of activities within ransomwares (Homayoun, Dehghantanha, Ahmadzadeh, Hashemi, & Khayami, 2017).However, it is not only ransomware threats that can be hunted using sequential pattern mining but it may also be effective against hidden threats, which are considered advanced targeted threats or Advanced Persistent Threats (APT). These kind of threats have the ability to hide themselves for instance in memory where they cannot be easily detected. They also have some unique feature of being able to work around security controls. (Rasheed, Hadi, & Khader, 2017)

1. **RELATED WORK**



*Fig. 1. Threat Hunting Process* (Miazi, Pritom, Shehab, Chu, & Wei, 2017)

The threat hunting process as a whole, as shown in the diagram above, requires inputs in the form of cyber intelligence, which requires an expertise in cyber security, different system logs and alerts from IDS, Firewalls and other security tools, for the purpose of scrutinizing, that requires an expertise in analytical skills (Miazi, Pritom, Shehab, Chu, & Wei, 2017). The threat hunting process itself is divided into stages. The first stage primarily is planning. Here, the assets of importance are identified, meaning, the most important, say information in the data centers held by enterprises, are identified and close security measures are deployed in these areas. Different security teams deployed by the organizations are responsible for the different threat hunting tasks, looking out for the enormous variety of attacks and intrusions that may threaten the integrity of the system. One team could be performing host hunting, another network hunting and others could be performing the threat intelligence analysis. The collective work of each team, under a supervisory role that serves as the primary command and control node responsible for planning and execution of the threat hunting operations, make up the threat hunting process a success. But to facilitate this, these security teams require data. Data is collected from system logs where user activities are logged at the appropriate level of detail, avoiding too much details leading to processing overhead and too little details causing lack of visibility.

The second stage is pursuing and detection. Here threats are pursued using two approaches, one is inspecting known threats whereby it relies much on the Indicators of Compromise (IoC) such as known signature attacks. The other approach is detecting unknown threats by setting the baseline of what is considered normal activities, thereby enabling the ability to detect deviations and, with high possibility, unknown threats. The last stage in the hunting process is responding. This is the active stage of threat hunting process and that includes using the needed and efficient means to detect and stop the attack at the earliest stage possible. (Rasheed, Hadi, & Khader, 2017)

Threat hunting has been on the rise and requires quiet some amount of technical skills and the right mindset i.e. a shift to a post-infection mentality. Some of the skills required include threat intelligence, where intelligence reports are gathered from various sources, which will help security teams in incident detection and analysis. Besides that, a set of threat detection tools have been developed to help in the massive workload of threat hunting. These tools include SIEM (Security Incident and Event Management), EDR (Endpoint Detection and Response) and NDR (Network Detection and Response). These tools are useful in detecting malware programs such as ransomwares, which are reportedly becoming a dominant tool for cybercriminals and a growing threat to the IT infrastructure. Cases have even arisen where cybercriminals have encrypted users’ data as part of a Denial of Service (DoS) attack, and with recent adoption of eCurrencies such as Bitcoin, these cybercriminals demand for ransom payment in order to decrypt the user data. Such cases have led to the development of an educational framework that is tailored to ransomware threats as well as tools which mimicked ransomware attacks and has proved to be useful in reducing ransomware infections (Homayoun, Dehghantanha, Ahmadzadeh, Hashemi, & Khayami, 2017). For any cyber threat, prevention is ideal but detection is a necessity.

1. **METHODOLOGY**

Despite all the techniques, tools and skills of cyber security, cases about hackers bypassing systems have barely gone down. Naturally, threat hunting relies on available substantive data at a single point in time in order to provide the appropriate solutions. But the questions asked include: what kind of data are we looking for? Do we necessarily need to get infected or bypassed in order to collect these data? It should be noted that, even with the necessary threat hunting tools, the threat hunting process is somewhat ‘heavy’. The data to be collected involves installing agents on endpoints and/or hardware placed on networks which can actually be quite expensive even for large enterprises. Another problem is that this data lacks a broader context and historical perspective for security analysts to pinpoint an infection.

Few organizations have the necessary skill set and resources to analyze data and identify persistent threats. It may pose a financial burden even with the implementation of data mining through machine learning. The main objective would then be to perform the threat hunting process and come up with solutions in a cost effective manner. This can be attained by rethinking the approach to wide area networking along with the sharing in cybersecurity threat intelligence and threat information, and also enabling the sharing of Indicators of Compromise (IoC) (Ginn & Ionescu, 2017).

Software-Defined Wide Area Networking (SD-WAN), is a cloud based service which offers an alternative means to conducting threat hunting that addresses the shortcomings of the existing approaches. Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources such as servers, storage, networks, applications and services, that can be rapidly provisioned and realized with minimal management effort or service provider interaction. Cloud based SD-WAN is a new architecture, whereby all the entities of any typical enterprise network are all connected into a network in the cloud. These entities include all the enterprise’s branches, the data centers, cloud infrastructure and even mobile device users (Marek, Hardesty, Meyer, Hardcastle, & Longwell, 2018). All these elements connect to the cloud network backbone through a global series of Points of Presence (PoPs). This creates a single unified network, that carries all traffic of the various enterprise entities that are connected, including corporate internet plus WAN traffic. Having all this traffic flow on one network forms a valuable dataset for threat hunting. Cloud providers could use this single, unified source of data flowing across their cloud network as input to a new threat hunting service, that would be a great addition to the already implemented security offerings such as firewall as a service, secure web gateways and advanced threat protection. Traditional network security solutions are built at the level of a single branch network. All the traffic they inspect is isolated and limited to a specific location, such as a branch or a geographic location. Because the cloud service providers have their own network backbone into which they have full visibility, they can see all the network traffic, from all consumers, all over the world. This visibility into so many network flows and so much data, are unique and it allows the cloud service providers to build the models that will enable full threat hunting based on unlimited raw data. These models evolve three aspects of data context: client classification, the target and communication over time.

The client classification context element, when other security solutions inspect the source client with network flow, entities such as source IP (Internet Protocol), username and device name are considered. This information is usually used to distinguish different devices over the network but it is rarely used in the actual decision making of whether the traffic is malicious or not. Client classification can then be expanded into a broader scheme using elements such as whether HTTP (Hyper Text Transfer Protocol) or TLS (Transport Layer Security protocol) is part of the main communications, the unique fingerprints of various web browsers and the various types of libraries they use. These items provide much more detail and by analyzing this data with machine learning, cloud service providers can classify different clients on their networks quite accurately.

The other context element is the target, i.e. the IP (Internet Protocol) or domain address that a client is connecting to. The target is commonly part of the network flow that is used in the decision making process of whether something is malicious or not. Most security solutions simply compare the target against a list of security feeds. This model goes further by creating a popularity score to each target seen in the network. The score is calculated based on the number of times client communicate with the targets. Scores of all targets are then collected and typically, the lowest scored target are indicators of malicious or command and control websites.

The last context parameter is communication over time. Active malware keeps communicating over time, for instance, to exfiltrate data. The time or repetitiveness is an important data element. The more the external communication is repeated uniformly, the more likely it is a machine or bot that is generating this traffic and thus more likely to be malicious traffic. (Greenfield, 2018) (Kumar, 2018)

1. **CONCLUSION**

The cloud infrastructure is built, managed and accessed in different types of models, mainly the public cloud, the private cloud, the community cloud and the hybrid cloud. Normally, when compared to public cloud, the private and community cloud offers enterprises a greater degree of security. It is in these cloud models that we expect to get the SD-WAN services. Since not all enterprises can sustain the expense of a private cloud, they may look into a community cloud, where the cloud services are provisioned for use to a specific community of consumers from enterprises that share the same concerns such as security. With the sharing of Indicators of Compromise (IoC) and cybersecurity threat intelligence across different community clouds and among cloud service providers, threat hunting becomes a timely and successful process. Responses would even be more intense if an attack has already succeeded or is in progress (Rasheed, Hadi, & Khader, 2017).

Threat hunting is an effective and critical security measure that is needed to minimize business impact of attacks, improving visibility of enterprises’ environments and weaknesses, and achieve proactive, accurate and early detection of threats. Proper and effective threat hunting requires a thorough study of the monitored systems and segments to define the normal behavior that will help filter out unexpected and malicious activities. This will help define IoC for the systems, which can be monitored to trigger the needed action once met. The threat hunting process requires the continuous monitoring of systems to help keep IoC and thresholds of normal behavior updated and matching the changes in the monitored systems (Rasheed, Hadi, & Khader, 2017). This is best achieved through SD-WAN, a cloud computing architecture where all the entities of any typical enterprise network are all connected into a network in the cloud. This approach of threat hunting is particularly unique since the cloud providers not only monitor the system logs but also the raw network data that carries a lot of information.

# **REFERENCES**

Ginn, R. J., & Ionescu, I. (2017). Cyber Threat Analysis:. *Cyber Threat Analysis:*, 1.

Greenfield, D. (2018, May 27). *Cato Adds Threat Hunting Capabilities to Cato Cloud*. Retrieved from Cato Networks: https://www.catonetworks.com/blog/cato-adds-threat-hunting-capabilities-cato-cloud/

Homayoun, S., Dehghantanha, A., Ahmadzadeh, M., Hashemi, S., & Khayami, R. (2017). Know Abnormal, Find Evil: Frequent Pattern. *Know Abnormal, Find Evil: Frequent Pattern*, 1.

Kumar, M. (2018, June 11). *A New Paradigm For Cyber Threat Hunting*. Retrieved from The Hacker News: https://thehackernews.com/2018/06/cyber-threat-hunting.html

Marek, S., Hardesty, L., Meyer, D., Hardcastle, J. L., & Longwell, A. (2018, March 26). *What is Software-Defined WAN (or SD-WAN or SDWAN)?* Retrieved from SDxCentral, LLC: https://www.sdxcentral.com/sd-wan/definitions/software-defined-sdn-wan/

Miazi, M. N., Pritom, M. M., Shehab, M., Chu, B., & Wei, I. (2017). The Design of Cyber Threat Hunting Games:. *The Design of Cyber Threat Hunting Games:*, 1.

Rasheed, H., Hadi, A., & Khader, M. (2017). Threat Hunting using GRR Rapid Response. *Threat Hunting using GRR Rapid Response*, 1.