Ransomware Detection with Honey Credentials

Kent Yamada
NYU Tandon School of Engineering
Cyberfellows, Information Security and Privacy, Fall 2021
New York, NY
ky2303@nyu.edu

***Abstract*— In this paper I propose the use of honey credentials on networked machines and IoT devices to detect network attacks in the account discovery/lateral movement phase [MITRE ATT&CK TA0007, TA0008]. After an initial foothold is established in a target network, the most likely next step is a lateral movement or elevation of privileges. Honey credentials can be an effective and useful tool for warning network defenders against compromised devices early in the attack chain.**

Keywords—honey credentials, honey tokens, honey pot, MITRE ATT&CK, ransomware, lateral movement, privilege escalation

I. Introduction

Samples of ransomware variants such as Ryuk and Conti change their code and obfuscation techniques with each new attack, leveraging different initial vectors of intrusion, but once initial access is gained attackers utilize similar lateral movement and privilege escalation techniques. A playbook leaked by a lower-level Conti operator, translated and published by Cisco Talos in September 2021 [4][4] showed how attackers use the same techniques elevate privileges once they have gained an initial foothold in a target network. The playbook also showed security researchers that there are hierarchies in the Ransomware-as-a-Service groups and newer operators with less experience can be given directions from such a playbook to compromise large enterprise networks and cause damage. A fake user account and password combination that has no access to valuable resources but trigger alerts for defenders when used in a login attempt could catch lower-skill threat actors early in the attack chain and save organizations from further exploitation.

Some attack vectors can bypass authentication altogether and move to later stages such as contacting command and control (C2) servers for data exfiltration and encryption. The Conti playbook showed how PowerShell is used by attackers to dump shares, disable defensive monitoring tools, and run various scripts to escalate privileges and contact C2 servers. Tools that inject PowerShell scripts into existing processes such as PSinject can run arbitrary code as an existing privileged user. Other authentication bypass techniques such as ProxyShell or PrintNightmare allow adversaries to create new administrator accounts for themselves. These types of attacks would unfortunately not be caught with the use of honey credentials.

However, other techniques that search for existing privileged credentials and/or brute force passwords would theoretically be thwarted by honey credentials. Mimikatz, a widely used post-exploit credential dumper is included in popular security frameworks such as Metasploit and Cobalt Strike. A potential usage of Mimikatz, outlined in detail in the leaked Conti playbook, is to acquire local passwords and domain hashes for offline brute-forcing. Dumping the LSASS.exe process, which can be done with Mimikatz or manually gives the attacker plaintext passwords and NTLM hashes which the playbook explains can be brute-forced or used in pass-the-hash attacks later in the attack chain.

In the following sections, the possibility for honey credentials to be used in ransomware defense will be studied. First, previous research related to honey credentials and ransomware will be discussed and different approaches are considered. Then, my hypothesis and potential empirical evidence to study will be presented. Finally, potential future work and the extensions to the experiment to be explored are evaluated.

II. Related Research

Researchers and security industry vendors have extensively studied, developed, and use honey pots in network implementations with honey credentials usually being treated as a subset of this defense tactic. For example, Rapid7 includes honey folder and honey credential modules in their InsightIDR security product [5][5]. However, since most companies and organizations handle user and password controls internally, not much research or open-source material on honey credentials is available. This paper focuses on using honey credentials in catching lateral movement in the attack chain. Related studies exist in finance and healthcare fields, but this specific use does not appear to be thoroughly researched.

Herley and Florencio [1] propose the use of thousands to millions of honeypot credentials to protect bank accounts and observe attacker's behaviors when trying to cash out using compromised accounts. This research was published in 2008, thus the focus is on thwarting attackers using brute-force password spraying techniques to compromise many user accounts at once. This strategy is similar to this paper's because the idea is to catch adversaries who have already broken in and are past the initial network defenses but have not yet compromised anything of value. Fake bank accounts are created with existing user data anonymized with fake personal information. These accounts are designed to be (nearly) impossible to access by real bank customers but appear completely real to an attacker who has found a fake user/password login combination. When the attacker tries to "cash-out" of the account, the bank is alerted to the attempt and also gains knowledge of the cashier or real account they have tried to use.

Sibi Chakkaravarthy, et al. [2] experiment using a testbed network of Raspberry Pi's, honeyfolders and a novel Intrusion Detection Honeypot (IDH) to detect 1000 samples of ransomware. Although only one part of the IDH, the proposed honey folder is used to detect suspicious file activity. Though conceptually similar to honey credentials, honey folders are not useful for detecting lateral movement or privilege escalation and mainly focus on the encryption phase of the attack chain. The researchers created honey folders with random dummy documents and media that an attacker would consider valuable and placed on each host in the testbed network. Then an agent modeled with the researcher's proposed algorithm monitors the folder's contents for state transitions to determine if actions are ransomware or normal user activity. When the algorithm determines ransomware activity, a firewall is alerted, and the offending process is killed.

Similarly, S. Sheen and A. Yadav [3] use a large malware sample set (16243) along with a set of benign executables (3620) to identify malware. They first search for API calls within samples then run the set through a classification model to determine if the sample is malicious or benign. Their experiment was able to produce a best performance of 0.9653 true positive classification of malware executables. Though not exactly honey credential or honeypot related, this study is interesting because the use of machine learning classifiers to detect ransomware. As the never-ending battle against threat adversaries evolves, machine learning and data mining will become a more useful tool for detecting and possibly even predicting attacks.

III. HYPOTHESIS AND EMPIRICAL EVIDENCE

The hypothesis for this paper is that creating honey credentials is a reliable method to alert defense teams to an initial breach and update firewalls before a ransomware attack can occur. The Cybersecurity and Infrastructure Security Agency's (CISA) September 2021 Alert AA21-265A [6] warns of rising reports of successful Conti ransomware attacks and includes a table of MITRE ATT&CK techniques observed in the attacks. Of interest to this paper is the "Credential Access" section shown in Table 1.

The methodology for testing was to create a test network with a honeypot server machine and a vulnerable host machine. The host machine is connected to, but cannot access, the target server. The honey credentials are stored in memory on the vulnerable machine using the Windows `runas` command with the `/netonly` flag [7]. This exposes the honey credential to popular credential stealing techniques such as procdump and Mimikatz. In this case, a dump file is created using Windows Task Manager (a living-off-the-land attack technique) and analyzed using Mimikatz. The ssh port (22) was opened on the server and a logger, the opensource credentials catching honeypot, heralding [8], was used to view and record login attempts on the server.

TABLE I. MITRE ATT&CK TECHNIQUES – CONTI

Credential Access		
Technique Title	**ID**	**Use**
Brute Force	T1110	Conti actors use legitimate tools to maliciously scan for and brute force routers, cameras, and network-attached storage devices with web interfaces.
Steal or Forge Kerberos Tickets: Kerberoasting	T1558.003	Conti actors use Kerberos attacks to attempt to get the Admin hash.
System Network Configuration Discovery	T1016	Conti ransomware can retrieve the ARP cache from the local system by using the `GetIpNetTable()` API call and check to ensure IP addresses it connects to are for local, non-internet systems.
System Network Connections Discovery	T1049	Conti ransomware can enumerate routine network connections from a compromised host.
Process Discovery	T1057	Conti ransomware can enumerate through all open processes to search for any that have the string `sql` in their process name.
File and Directory Discovery	T1083	Conti ransomware can discover files on a local system.
Network Share Discovery	T1135	Conti ransomware can enumerate remote open server message block (SMB) network shares using `NetShareEnum()`.

IV. CONCLUSION AND FUTURE WORK

This paper explored the potential for honey credentials to be used as a defense technique against ransomware. Although it will not catch all attacks, a honey credential setup can be a good layer of defense and act like a canary in certain attack chains. Also, it is a relatively easy defense system to setup and can add value to logging and monitoring systems. Finally, a honey credential system can act like a canary for on-going attacks and help defenders and the community gain more insight into attack vectors and attackers' behaviors.

Further research in this domain can go in many directions. The next experiment will be carried out with the network accessible to the internet to see if real attackers access it. This test network will be left open over a period and could possibly yield a few new metrics to examine. The evidence of honey credentials effectiveness would be reinforced if the fake credential appears in the server login attempts. Over the experiment period, if there are signs of access on the vulnerable machine and no login attempts on the server, the hypothesis is rejected.

Other metrics that could be useful to observe are the time between the vulnerable machine being exploited and the authentication attempt on the server, and a list of IP addresses from which the servers were exploited. A measure of time between initial exploit and login attempt using honey credentials could be studied further to examine whether initial exploits are automated or human operated. This could also be studied over a longer period to study attackers' most active operation times. The other interesting metric that could be gathered from the experiment is attacker IP addresses. These could be useful for automation, triggering a firewall block for incoming traffic for any address that attempts to authenticate using the honey credential on the honeypot server or anywhere on the network in a production enviornment. Finally, relating this work to the machine learning classification research by Sheen and Yadev, a honey credential authentication attempt event could be used in AI-based network defense systems. For example, classification of ransomware, identification of threat actors, or even potential intrusion vectors could be calculated in a machine learning-based network defense system by giving a failed authentication attempt a class weighting in a classification model, thus improving the detection system over time.

Acknowledgment

Thank you to the Professors and TAs this semester for the instruction and guidance. Also, thanks to Tomer Gross for the insightful discussion on ransomware and attack techniques.

References

[1] C. Herley and D. Florˆencio, "Protecting Financial Institutions from Brute-Force Attacks," Proceedings of The Ifip Tc 11 23rd International Information Security Conference, 2008, pp 681-685, doi: 10.1007/978-0-387-09699-5_45

[2] S. Sibi Chakkaravarthy, D. Sangeetha, M. V. Cruz, V. Vaidehi and B. Raman, "Design of Intrusion Detection Honeypot Using Social Leopard Algorithm to Detect IoT Ransomware Attacks," in IEEE Access, vol. 8, pp. 169944-169956, 2020, doi: 10.1109/ACCESS.2020.3023764.

[3] S. Sheen and A. Yadav, "Ransomware detection by mining API call usage," 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2018, pp. 983-987, doi: 10.1109/ICACCI.2018.8554938.

Other Resources

[4] https://blog.talosintelligence.com/2021/09/Conti-leak-translation.html

[5] https://docs.rapid7.com/insightidr/deception-technology

[6] https://us-cert.cisa.gov/ncas/alerts/aa20-302a

[7] https://logrhythm.com/blog/using-honeywords-to-make-password-cracking-detectable/

[8] https://github.com/johnnykv/heralding