Pacific Northwest
NATIONAL LABORATORY

PNNL-33401

ROTOR: Research to Operations and Operations to Research

September 2022

Aaron R Phillips
Chance R Younkin
Glenn A Fink
Matthew R Oster
Edgar W Thomas
Theora R Rice
Angie E Chastain

U.S. DEPARTMENT OF
ENERGY

Prepared for the U.S. Department of Energy
under Contract DE-AC05-76RL01830

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, **makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights**. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

**Available to DOE and DOE contractors from
the Office of Scientific and Technical
Information,
P.O. Box 62, Oak Ridge, TN 37831-0062**
www.osti.gov
**ph: (865) 576-8401
fox: (865) 576-5728
email: reports@osti.gov**

**Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
or (703) 605-6000
email: info@ntis.gov
Online ordering: http://www.ntis.gov**

PNNL-33401

ROTOR: Research to Operations and Operations to Research

September 2022

Aaron R Phillips
Chance R Younkin
Glenn A Fink
Matthew R Oster
Edgar W Thomas
Theora R Rice
Angie E Chastain

Prepared for
the U.S. Department of Energy
under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

Acknowledgments

This research was supported by the **National Security Mission Seed**, under the Laboratory Directed Research and Development (LDRD) Program at Pacific Northwest National Laboratory (PNNL). The computational work was performed using PNNL Computing at Pacific Northwest National Laboratory. PNNL is a multi-program national laboratory operated for the U.S. Department of Energy (DOE) by Battelle Memorial Institute under Contract No. DE-AC05-76RL01830.

Contents

Figures

No table of figures entries found.

Tables

No table of figures entries found.

1.0 Introduction

PNNL's Research to Operations/Operations to Research (ROTOR) Program fosters a cybersecurity operations to researcher collaboration, identifying hard problems and challenges for cyber defenders and bringing researcher science disciplines to these challenges. ROTOR allows PNNL's cyber defenders to better protect the laboratory but also fosters innovation and incubation of solutions for our sponsor missions across DOE, DOD, DHS and the intelligence community. Our sponsors and their missions are faced with many of the same cyber defense challenges that PNNL faces. By using our own laboratory security operations environment as an innovation generator and testing ground, allowing research to be tested and tried, our sponsors are directly benefitted.

PNNL is in a unique position to combine our world class research organization with enterprise cybersecurity operations. ROTOR is taking advantage of this to bring PNNL cybersecurity research projects into an operational environment within PNNL's Cyber Security Operations Center (CSOC).

With ROTOR, PNNL researchers gain the advantage of operational experience and expertise and have an avenue for showcasing research in an operational environment. This helps to advance PNNL research, prove operability of PNNL projects to its sponsors, and provide real-world insight to real-world problems for our cybersecurity research agenda.

ROTOR is the conduit for PNNL cybersecurity research to find its way to operational use. PNNL researchers and engineers are informed by real-world operations and operations has access to current PNNL research and engineering capabilities.

All of this leads to improved reputation for PNNL as a provider of national security solutions that are tried and tested. To date, ROTOR has collaborated with six projects, enabling each to conduct operational work within the CSOC.

1.1 SerialTap

PNNL and other federal facilities utilize OT systems to maintain and operate facilities and laboratory capabilities. A large percentage of these OT systems still leverage legacy point-to-point and shared bus communication infrastructure, such as RS232/485. Common IT security monitoring tools are designed around Ethernet/IP based communication infrastructure and cannot monitor legacy communication systems. To achieve holistic monitoring of these environments a solution is needed that allows collecting data from legacy OT systems and transmitting actionable information to IT security's triage and incident response systems.

SerialTap is a passive monitoring device that sits on legacy serial communication buses to monitor the data and messages between components of the building control systems. Through the ROTOR Pilot the SerialTap team deployed an initial proof of concept data collection and analysis system in a PNNL building, and coordinated with CSOC analysts and engineers to route SerialTap data into existing CSOC monitoring solutions. By working with the CSOC the team was able to improve the SerialTap project thanks to the invaluable feedback from the CSOC team about how to integrate the SerialTap into real world workflows. As a result, the

CSOC now has a method to monitor a previously unmonitorable set of PNNL systems and continue to work toward addressing a DOE notable.

1.2 MLSTONES

Malware detection is a complex and dynamic research field. Many malware detection systems rely on binary signatures or rules that can be easily fooled by making small changes to the malware. MLSTONES is a bio-inspired malware detection system that can detect similarity between malware samples, even in the presence of modifications. The MLSTONES ROTOR Pilot involved retraining the system using an updated dataset of malware, and testing using CSOC artifacts. Additionally, the MLSTONES team designed a new container-based deployment package designed to be used in a CSOC environment. This effort provided extremely valuable user interaction and feedback by providing direct exposure to CSOC analysts, as well as necessary funding to improve and expand both the trained malware model, as well as the container-based method for deploying MLSTONES into CSOC environments. This MLSTONES instance is still in use by CSOC analysts and has been used to identify malware samples that were not caught by other systems.

1.3 BastionML

Bastion hosts are a key element in the cyber security strategy of PNNL. These machines regulate access control to the mission critical servers in PNNL’s infrastructure. Thus, securing these machines is critically important. The BastionML project was intended to help CSOC analysts quickly identify anomalous behavior and identify non-privileged users who are behaving as administrators on bastion nodes.

Log file analysis is a time-consuming, human intensive process. Currently, human analysts must trust the analysis of log aggregation tools and may spend significant time investigating whether detections are indicators of compromise (IOCs) or false positives. The BastionML project worked to help human users establish the meaning of various types of activity and provide a baseline of normality for behavior. This will save analyst time and allows a baseline of behavior change over time to be established. Additionally, this also allows for visual analytics tools to be created that would ease the reporting burden of analysts and enable better explainability of classifications.

1.4 Cymbiote

Operational Technology often depends upon specialized embedded computing devices to monitor and control critical processes. However, these embedded devices often lack the traditional security controls we have come to expect and depend upon in IT to monitor and manage the security of the device. There is a need for a solution to provide the security oversight around these devices, which are often operating on isolated and not easily monitored communication networks, and controls to mitigate and respond to any potential threat activity.

The Cymbiote provides a means to add on the common host-based logging and alerting functionality in addition to some mitigative control to specialized embedded computing systems. With the Cymbiote data feeds, cyber analysts can gather a better situational awareness of the health and potential threat activity of OT system edge nodes. With this gained knowledge they can detect and respond more efficiently and effectively to threats targeting OT systems.

1.5 Conditional Context Analyzer: COCOA

CSOC receives many logs and alerts of varying types a day, ranging from VPN traffic logs ingested in Splunk to malicious attachment alerts from O365. Fortunately, the number of alerts in need of manual attention is many orders of magnitude lower than incoming data volumes, yet false positives are inevitable. The need for reducing these false positive alerts while acquiring quick contextual awareness of which user generated such alerts/logs becomes more critical as data volumes remain high and cyber threat surfaces increase. The Conditional Context Analyzer (COCOA) addresses these needs by providing a mechanism for conditioning on alert context, to either filter out false positives, or highlight them so that a more informed manual filtering can be done efficiently. CSOC can leverage these detection and filtering mechanisms in order to build new alerting mechanisms and dashboards that enhance their workflows.

The team collected cyber log data provided by CSOC, and utilized both supervised and unsupervised machine learning methods to build a classifier that could distinguish between users, as well as classes of users. Preliminary results indicate that distinguishing individual user activity from others is more performant than distinguishing user types (for example determining whether a user is acting like a “data scientist”). Furthermore, the team discovered that that finer granularity data is more performant (i.e., binning records into 5-minute intervals is better than 15min, which in turn is better than 1hour bins). These findings show promise for future research and development in this area, and can be leveraged by CSOC analysts to build novel alerting mechanisms and dashboards that should easily integrate into existing Splunk-based workflows.

1.6 Cyber Log Embeddings for VPN: CLEM4VPN

The increase in remote work has made VPN a major target for bad actors to gain unauthorized access to internal resources. Malware on the endpoints could subvert VPN functionality and attempt to make illegitimate connections outside the enterprise or to compromise enterprise assets. Given the large number of VPN connections, this attack traffic may be difficult to separate from normal traffic. To combat these threats, CSOC teams utilize VPN log capture and analysis systems to monitor these connections and discover issues. While these systems can help discover security issues with VPN systems, they also create large amounts of data.

To assist CSOC analysts in analyzing this data, the CLEM4VPN project created a prototype machine-learning tool that ingests and trains on VPN log data. Using this trained model, the team was able to plot IP address locations and found that they cluster consistently. They were also able to identify IP addresses that move between clusters, indicating a change in function. This research confirmed that these machine learning methods can be used to help establish a baseline for normal VPN behavior, and in the future could be used to develop alerting mechanisms and dashboards to facilitate CSOC monitoring of VPN data.

2.0 Conclusion

While we have seen success with these projects, there have been challenges in sustaining ongoing project engagements. The primary challenge has been identifying projects that could be applied to the CSOC domain with a relatively small amount of research and development. Our experience has been that most projects are working with sponsors that have an operational space that looks very different from the CSOC environment. This means that significant research and development would need to be done to develop a version of the technology that could be utilized by CSOC analysts. Future ROTOR efforts would benefit from a broader set of operational partners that could provide operational collaboration beyond the CSOC. This would allow for engagement with projects that would otherwise be unable to benefit from ROTOR.

Additionally, future efforts should continue the “Trench Walks” and “Trench Talks” functions of ROTOR. Trench Walks are comprised of a series of short meetings between researchers and CSOC analysts where the researcher can observe CSOC workflows and ask questions of analysts in order to better understand their operational context. Trench Talks are a series of presentations given by CSOC staff that provide valuable insight into the problems faced by the CSOC. These presentations have been very well received, but in the future ROTOR team should strive to strike a balance between researcher and CSOC presentations, in order to give both sides insight into the work being done by their colleagues.

Pacific Northwest
National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354

1-888-375-PNNL (7665)

www.pnnl.gov