ACM KDD AI4Cyber: The 1st Workshop on Artificial Intelligence-enabled Cybersecurity Analytics

Sagar Samtani†
Department of Operations and Decision Technologies
Indiana University
Bloomington, IN, USA
ssamtani@iu.edu

Shanchieh Yang
Department of Computer Engineering
Rochester Institute of Technology
Henrietta, New York, USA
Jay.Yang@rit.edu

Hsinchun Chen
Department of Management Information Systems
University of Arizona
Tucson, Arizona, USA
hsinchun@arizona.edu

ABSTRACT

Despite significant contributions to various aspects of cybersecurity, cyber-attacks remain on the unfortunate rise. Increasingly, internationally recognized entities such as the National Science Foundation and National Science & Technology Council have noted Artificial Intelligence can help analyze billions of log files, Dark Web data, malware, and other data sources to help execute fundamental cybersecurity tasks. Our objective for the 1st Workshop on Artificial Intelligence-enabled Cybersecurity Analytics (half-day; co-located with ACM KDD) was to gather academic and practitioners to contribute recent work pertaining to AI-enabled cybersecurity analytics. We composed an outstanding, inter-disciplinary Program Committee with significant expertise in various aspects of AI-enabled Cybersecurity Analytics to evaluate the submitted work. Significant contributions to the half-day workshop were made in the areas of CTI, vulnerability assessment, and malware analysis.

CCS CONCEPTS

• Security and Privacy • Computing methodologies ~Artificial intelligence ~Knowledge representation and reasoning • Computing methodologies ~Machine learning ~ Machine Learning Approaches

KEYWORDS

Cybersecurity; artificial intelligence; analytics; machine learning

ACM Reference format:
Sagar Samtani, Shanchieh Yang, and Hsinchun Chen. 2021. ACM KDD AI4Cyber: The 1st Workshop on Artificial Intelligence-enabled Cybersecurity Analytics. In *Proceedings of 2021 ACM Conference Knowledge Discovery and Data Mining (KDD'21), August 14– 18, Virtual Event.* ACM, New York, NY, USA. 2 pages. https://doi.org/10.1145/3447548.3469450

KDD '21, August 14–18, 2021, Virtual Event, Singapore

ACM ISBN 978-1-4503-8332-5/21/08. https://doi.org/10.1145/3447548.3469450

1 Introduction and Workshop Objective

Modern society's irreversible dependence on computing technology has helped cybersecurity emerge as a grand societal concern. While significant contributions have been made to numerous aspects of cybersecurity, cyber-attacks remain on the rise. A key reason for this unfortunate trend is the ever-growing volume, heterogeneity, and velocity of cybersecurity data such as malware, log files, Dark Web data, social media data, and other data types that make conducting key cybersecurity tasks non-trivial [1]. Increasingly, internationally recognized entities such as the National Science Foundation (NSF) and National Science & Technology Council (NSTC) have noted Artificial Intelligence (AI) and Machine Learning (ML) can analyze billions of cybersecurity records to execute fundamental cybersecurity tasks such as asset management, vulnerability prioritization, threat forecasting, and control allocations [2, 3].

Despite initial promising results in this space, substantial work is needed to carefully develop novel AI-enabled algorithms, methods, and systems for the unique volume, variety, velocity, and veracity of cybersecurity data. Moreover, industry and academic AI-enabled cybersecurity analytics initiatives are often siloed. Our objective for the 1st Workshop on Artificial Intelligence-enabled Cybersecurity Analytics (co-located with ACM KDD) was to gather academic and practitioners to share, disseminate, and communicate completed research papers, work in progress, and review articles pertaining to AI-enabled cybersecurity analytics. Significant contributions to the half-day workshop were made in the areas of CTI, vulnerability assessment, and malware analysis.

2 Topics of Interest for the Workshop

Four major themes of AI for cybersecurity analytics research exist [4]: malware analysis and evasion, Cyber Threat Intelligence (CTI), Security Operations Centers (SOCs), and disinformation and computational propaganda. Therefore, areas of interest for this workshop include, but are not limited to:

- Static and/or dynamic malware analysis and evasion
- IP reputation services (e.g., blacklisting)
- Anomaly and outlier detection
- Phishing detection (e.g., email, website, etc.)
- Dark Web analytics (e.g., multi-lingual threat detection)
- Spam detection
- Large-scale and smart vulnerability assessment
- Real-time threat detection and categorization

- Real-time alert correlation for usable security
- Weakly supervised and continual learning for intrusion detection
- Adversarial attacks to automated cyber defense
- Automated vulnerability remediation
- IoT analysis (e.g., fingerprinting, measurements)
- Misinformation and disinformation
- Deep packet inspection
- Automated mapping of threats to cybersecurity risk management frameworks

Authors were encouraged to clearly articulate their data (e.g., key metadata, statistical properties), analytical procedures (e.g., algorithm details), and evaluation set up (e.g., performance metrics, statistical tests, case studies) in their submissions. Making data, code, and processes publicly available to facilitate scientific reproducibility were strongly encouraged to help facilitate a culture of data/code sharing in this quickly developing discipline.

3 Summary of Program Committee Members

We composed an outstanding, inter-disciplinary Program Committee (PC) with significant expertise in various aspects of AI-enabled Cybersecurity Analytics, including adversarial malware evasion, Dark Web Analytics, CTI, vulnerability assessment, disinformation, social cybersecurity, and more. The PC spans both the academic and practitioner landscapes. The PC members are listed below in alphabetical order based on last name:

- **Mr. Benjamin Ampel,** University of Arizona
- **Dr. Hyrum Anderson,** Microsoft
- **Dr. Victor Benjamin,** Arizona State University
- **Dr. Elias Bou-Harb,** University of Texas, San Antonio
- **Dr. Yidong Chai,** Hefei University of Technology
- **Dr. Sriram Chelleppan,** University of South Florida
- **Dr. Sven Krasser,** CrowdStrike
- **Dr. Weifeng Li,** University of Georgia
- **Dr. Yunji Liang,** Northwestern Polytechnical University
- **Dr. Xiaojing Liao,** Indiana University, Bloomington
- **Dr. Sudip Mittal,** Mississippi State University
- **Dr. Edward Raff,** Booz Allen Hamilton
- **Dr. Ethan Rudd,** FireEye
- **Dr. Ankit Shah,** University of South Florida
- **Mr. Steven Ullman,** University of Arizona
- **Dr. Ziming Zhao,** University of Buffalo
- **Dr. Lina Zhou,** University of North Carolina, Charlotte
- **Dr. Hongyi Zhu,** University of Texas, San Antonio

4 Background of the Workshop Organizers

The workshop organizers have extensive expertise in numerous AI for Cybersecurity analytics related topics and also serve in key leadership roles within the broader AI for Cybersecurity discipline. A brief biography of each member is summarized below:

- **Dr. Sagar Samtani** is an Assistant Professor and Grant Thornton Scholar of Operations and Decision Technologies at Indiana University. Dr. Samtani's research on CTI for Dark Web analytics and scientific cyberinfrastructure security have been funded by the NSF SaTC, CICI, and CRII programs. Dr. Samtani has published 40+ articles at *MIS Quarterly, Journal of MIS*, *ACM TOPS*, IEEE S&P, IEEE ICDM, and others. He has served as Program Chair at IEEE ISI 2020, and PC member at ACM CCS, IEEE S&P and other AI for Cybersecurity venues. He is a member of ACM and IEEE.
- **Dr. Shanchieh (Jay) Yang** is a Professor in Computer Engineering and the Director of Global Outreach for Global Cybersecurity Institute at Rochester Institute of Technology. His research focuses on advancing machine learning, modeling, and simulation for predictive cyber intelligence and anticipatory cyber defense. He has worked over 20 sponsored research projects supported by NSF, IARPA, DARPA, NSA, AFRL, ONR, and ARO. His team has developed several prototypes, including ASSERT to continuously learn and generate emerging statistical attack models, CASCADES to simulate synthetic attack scenarios, and CAPTURE to forecast cyberattacks using unconventional signals in the public domain. He has published more than 70 peer-reviewed papers.
- **Dr. Hsinchun Chen** is a Regents' Professor of Management Information Systems at the University of Arizona. Dr. Chen is the founder and director of the Artificial Intelligence Lab, an internationally recognized research lab renowned for its research on AI cybersecurity. Dr. Chen has received over $50M of federal funding from funding agencies such as the NSF, DoJ, DHS, and others. As director of the AZSecure Cybersecurity program at UArizona, Dr. Chen has received over $10M from the NSF SFS, SaTC, and CICI programs since 2013. Dr. Chen has published over 900 papers in highly visible IEEE, ACM, and information systems journals and conferences. He is also the founding conference chair for several leading security informatics conferences and workshops. He is a Fellow of the IEEE, ACM, and AAAS.

ACKNOWLEDGMENTS

This workshop is based upon work funded by DGE-2038483 (SaTC-EDU), OAC-1917117 (CICI), and CNS-1850362 (CRII SaTC). We would like to thank all of the authors for their interest and contributions to this workshop. We would also like to thank each Program Committee Member for their tireless efforts reviewing and providing thoughtful comments for the submitted papers. Finally, we would like to thank the ACM KDD 2021 Workshop Chairs, Dr. Beibei Li (Heinz College at Carnegie Mellon University) and Dr. Lauren Rhue (Robert H. Smith School of Business at University of Maryland) for their advice.

REFERENCES

[1] Elisa Bertino, Murat Kantarcioglu, Cuneyt Gurcan Akcora, Sagar Samtani, Sudip Mittal, Maanak Gupta. 2021. AI for Security and Security for AI. In *Proceedings of the 11th. ACM Conference on Data and Application Security and Privacy (CODASPY)*. ACM Press, New York, NY, 226-236. DOI: https://doi.org/10.1145/3422337.3450357

[2] National Science and Technology Council. 2019. The National Artificial Intelligence Research and Development Strategic Plan: 2019 Update. Washington, DC. Retrieved from https://www.nitrd.gov/pubs/National-AI-RD-Strategy-2019.pdf

[3] National Science Foundation. 2019. National Artificial Intelligence (AI) Research Institutes (2019). nsf20503 | NSF – National Science Foundation. Retrieved from https://www.nsf.gov/pubs/2020/nsf20503/nsf20503.pdf

[4] Sagar Samtani, Murat Kantarcioglu, and Hsinchun Chen. 2020. Trailblazing the Artificial Intelligence for Cybersecurity Discipline. ACM Trans. Manag. Inf. Syst. 11, 4 (December 2020), 1–19. DOI:https://doi.org/10.1145/3430360