AI for Security and Security for AI

Elisa Bertino
bertino@purdue.edu
Purdue University

Murat Kantarcioglu
muratk@utdallas.edu
The University of Texas at Dallas

Cuneyt Gurcan Akcora
Cuneyt.Akcora@umanitoba.ca
University of Manitoba, Canada

Sagar Samtani
ssamtani@iu.edu
Indiana University

Sudip Mittal
(Moderator)
mittals@uncw.edu
University of North Carolina Wilmington

Maanak Gupta
(Moderator)
mgupta@tntech.edu
Tennessee Technological University

ABSTRACT

On one side, the security industry has successfully adopted some AI-based techniques. Use varies from mitigating denial of service attacks, forensics, intrusion detection systems, homeland security, critical infrastructures protection, sensitive information leakage, access control, and malware detection. On the other side, we see the rise of Adversarial AI. Here the core idea is to subvert AI systems for fun and profit. The methods utilized for the production of AI systems are systematically vulnerable to a new class of vulnerabilities. Adversaries are exploiting these vulnerabilities to alter AI system behavior to serve a malicious end goal. This panel discusses some of these aspects.

ACM Reference Format:
Elisa Bertino, Murat Kantarcioglu, Cuneyt Gurcan Akcora, Sagar Samtani, Sudip Mittal, and Maanak Gupta. 2021. AI for Security and Security for AI. In *Proceedings of the Eleventh ACM Conference on Data and Application Security and Privacy (CODASPY '21), April 26–28, 2021, Virtual Event, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3422337.3450357

Statement of Elisa Bertino

Cyberattacks with different goals, such as data ransoms, denial of service, sabotage, data theft, are on a dramatic increase. To make defenses more effective, recent solutions leverage AI techniques. However, a major problem is that the application of AI techniques to cybersecurity is not trivial. For example, if one would like to apply reinforcement learning, one has to understand how to properly design reward functions. In addition different security tasks may need different AI techniques. Thus an effective AI-enhanced defense in depth must be based on a comprehensive cybersecurity lifecycle and an analysis of adversarial tactics and techniques. Also, AI-based cybersecurity approaches have been intensely scrutinized with respect to ethics. Thus, for AI to be effectively and quickly adopted in cybersecurity, AI security and ethics need also to be assured. AI security requires assurance processes for data used in AI training as well as systematic approaches to AI security testing. With respect to ethics, we need to develop technical approaches to embed ethics principles in intelligent systems.

Elisa Bertino is the Samuel D. Conte Professor of Computer Science at Purdue University. She serves as Director of the Purdue Cyberspace Security Lab (Cyber2Slab). In her role as Director of Cyber2SLab she leads multi-disciplinary research in data security and privacy. Prior to joining Purdue, she was a professor and department head at the Department of Computer Science and Communication of the University of Milan. She has been a visiting researcher at the IBM Research Laboratory (now Almaden) in San Jose, at the Microelectronics and Computer Technology Corporation, at Telcordia Technologies, and visiting professor at the Singapore Management University and the National University of Singapore. Her recent research focuses on cybersecurity and privacy of cellular networks and IoT systems, and edge analytics and machine learning for cybersecurity. Elisa Bertino is a Fellow member of IEEE, ACM, and AAAS. She received the 2002 IEEE Computer Society Technical Achievement Award for "For outstanding contributions to database systems and database security and advanced data management systems", the 2005 IEEE Computer Society Tsutomu Kanai Award for "Pioneering and innovative research contributions to secure distributed systems", the 2014 ACM SIGSAC Outstanding Contributions Award with citation "For her seminal research contributions and outstanding leadership to data security and privacy for the past 25 years", and the 2019-2020 ACM Athena Lecturer Award.

Statement of Murat Kantarcioglu

Direct application of AI techniques to cyber security domain may be misguided. Unlike most other application domains, cyber security applications often face adversaries who actively modify their strategies to launch new and unexpected attacks. Thus AI techniques for cyber security need to be resilient against the adaptive behaviors of the adversaries, and are able to quickly detect previously unknown new attack instances. Recently, various adversarial AI techniques (including our proposed techniques developed using Army Research Office funding) have been developed to counter adversaries' adaptive behaviors. For example, in our earlier work, we developed a game theoretic framework to discover an optimal set of attributes to build machine learning models against active adversaries. In another work, we modified an existing, and popular machine learning tool named Support Vector Machine to be more resistant against adversarial attacks. The attack models are defined in terms of the adversaries' capabilities of modifying data. Our solutions minimize the worst-case loss corresponding to the attack

CODASPY '21, April 26–28, 2021, Virtual Event, USA

ACM ISBN 978-1-4503-8143-7/21/04.
https://doi.org/10.1145/3422337.3450357

models, and show that such tailored tools could be more resistant to adversarial behavior compared to existing SVM alternatives.

Murat Kantarcioglu is a Professor in the Computer Science Department and Director of the Data Security and Privacy Lab at The University of Texas at Dallas (UTD). He received a PhD in Computer Science from Purdue University in 2005 where he received the Purdue CERIAS Diamond Award for Academic excellence. He is also a visiting scholar at Harvard Data Privacy Lab. Dr. Kantarcioglu's research focuses on the integration of cyber security, machine learning, data science and blockchains for creating technologies that can efficiently and securely store, analyze and share data and machine learning results. His research has been supported by grants including from NSF, AFOSR, ARO, ONR, NSA, and NIH. He has published over 170 peer reviewed papers in top tier venues such as ACM KDD, SIGMOD, ICDM, ICDE, PVLDB, NDSS, USENIX Security and several IEEE/ACM Transactions as well as served as program co-chair for conferences such as IEEE ICDE, ACM SACMAT, IEEE Cloud, ACM CODASPY. He is the recipient of various awards including NSF CAREER award, the AMIA (American Medical Informatics Association) 2014 Homer R Warner Award and the IEEE ISI (Intelligence and Security Informatics) 2017 Technical Achievement Award presented jointly by IEEE SMC and IEEE ITS societies for his research in data security and privacy.

Statement of Cuneyt Gurcan Akcora

Recent years have seen increased user participation in digital technologies. In the past, social networks had been the primary venue for a user to reach a global audience. Nowadays a user has many devices that collect data and share on the web. From self-driving cars to mobile and smart home devices, users connect to a wider audience through popular applications. This increased connection has not always been a boon. Data collected from other users without proper auditing is used to train ML models and make decisions with them. Companies seem to be most interested in increasing their revenues by using the data - any data. Suddenly, we have found ourselves governed by decisions that are conditioned on other people's preferences and actions. Well-documented cases of bias and racism seeping into machine learning models have caused outrage. With these harmful effects, it is imperative to learn how Machine Learning models use data to make decisions that affect users' lives. Governments have joined efforts as well and passed legislation that forces companies to explain their ML models' decisions. AI and ML researchers have responded to this growing issue with novel directions called Interpretable (IAI) and Explainable (XAI) AI. Interpretability attempts to explain the cause and effect observed within a system, whereas Explainability deconstructs a machine learning system to explain it in human terms. Both areas are developing tools that wrap around existing ML models and explain algorithm decisions and predictions. ML research must note the growing fields of IAI and XAI, and attempt to explain their ML models before model deployment. The benefits can be far-reaching as we believe that this practice will speed up ML adoption in society.

Cuneyt Gurcan Akcora is an Assistant Professor of Computer Science and Statistics at the University of Manitoba, Canada. He received his Ph.D. from the University of Insubria, Italy. His primary research interests are Data Science on complex networks and large-scale graph analysis, with applications in social, biological, IoT and Blockchain networks. He has worked at and collaborated with Qatar Computing Research Institute, Yahoo Research Barcelona, and Huawei Research in Istanbul. He is a Fulbright Scholarship recipient, and his research works have been published in leading conferences and journals such as IEEEtran, VLDB, ICDM, SDM, IJCAI, and ICDE.

Statement of Sagar Samtani

Modern society's irreversible dependence on information technology has placed a significant impetus on cybersecurity analysts to enhance the confidentiality, integrity, and availability of their ever-increasing asset-bases from a rapidly evolving threat landscape. In particular, methodologies such as deep learning, machine learning, network science, text analytics, and others can help human analysts sift through large quantities of heterogeneous cybersecurity data with unprecedented efficiency and effectiveness to detect patterns missed by conventional approaches. To date, AI for cybersecurity has been leveraged in four major cybersecurity applications areas: (1) cyber threat intelligence (CTI) to create, manage, and leverage information about emerging threats and key threat actors to enable effective cybersecurity decision making, (2) security operations centers (SOCs) to assist human analysts in tactical tasks such as alert management, vulnerability management, security orchestration, and others, (3) disinformation and computational propaganda to identify how fake content can sway public actions and perceptions, and (4) adversarial machine learning (AML) that relies on techniques such as generative adversarial networks (GANs) to generate fake or synthesized content to enhance offensive and defensive cybersecurity postures. Despite significant advancements in these areas from both industry and academia alike, key challenges remain that require significant attention. First, many practitioners and academics often work in siloes. Second, there is currently a dearth of publicly accessible datasets that accurately depict the nuances and complexities of production environments. Third, many AI-based models deployed in cybersecurity contexts often lack interpretability. Finally, many students and faculty may lack the resources to start and/or execute their AI for cybersecurity research.

Sagar Samtani is an Assistant Professor and Grant Thornton Scholar in the Department of Operations and Decision Technologies at the Kelley School of Business at Indiana University (IU). He is also a Fellow within the Center for Applied Cybersecurity Research at IU. Samtani graduated with his Ph.D. in May 2018 from the Artificial Intelligence Lab in University of Arizona's Management Information Systems (MIS) department at the University of Arizona. From 2014 – 2017, Samtani served as a National Science Foundation (NSF) Scholarship-for-Service (SFS) Fellow. Samtani's research centers around Explainable Artificial Intelligence for Cybersecurity and cyber threat intelligence. Samtani has published over three dozen journal and conference papers on these topics in leading venues such as MIS Quarterly, Journal of Management Information Systems, ACM Transactions on Privacy and Security, IEEE Intelligent Systems, Computers and Security, IEEE S&P, IEEE ICDM, and others. His research has received nearly $1.8M (in PI and Co-PI roles) from the NSF CICI, CRII, and SaTC-EDU programs. Samtani has won several awards for his research, including the ACM SIGMIS Doctoral Dissertation award in 2019.