Explainable AI for Cybersecurity

We are living in the age of smart computing systems powered by artificial intelligence as well as hardware (electronics) and software (applications). Recent attacks have demonstrated that an attacker can exploit hardware and software vulnerabilities in a wide variety of systems. In order to design trustworthy systems, it is crucial to identify and mitigate both hardware and software vulnerabilities. There are recent efforts in utilizing artificial intelligence for defending against cybersecurity attacks. Explainable AI has received significant attention due to its ability to provide insights into the decision making process. This enables efficient detection and localization of malicious software attacks (e.g., malware and ransomware), hardware attacks (e.g., hardware Trojans), as well as combinational attacks (e.g., Spectre and Meltdown attacks).

This course will provide a comprehensive overview of security vulnerabilities and state-of-the-art countermeasures using explainable AI. Specifically, the participants will understand how explainable AI can be effectively used to design secure and trustworthy systems. The course will also cover hardware acceleration of explainable AI using field programmable gate array (FPGA), graphics processing units (GPU), and tensor processing units (TPU). Finally, the course will provide insights into the security threats towards machine learning models and presents effective countermeasures. The participants will be able to get a complete picture of cybersecurity challenges and how to detect them using machine learning techniques. This course will serve as a stepping stone for researchers for designing secure and trustworthy systems.

Duration: 10 – 21 March 2025

Modules	• **Week 1 (10th March 2025 to 14th March 2025)**: Cybersecurity Landscape for Computer Systems, Explainable AI, Malicious Software (Malware), Malware Detection using Explainable AI, Security-aware Compilation, Hardware Trojans, Hardware Trojan Detection using Reinforcement Learning, Physical Side-Channel Analysis • **Week 2 (17th March 2025 to 21st march 2025)**: IP Piracy and its Protection, Microarchitectural Attacks, Speculative Load Forwarding Attack on Modern Processors, Spectre and Meltdown Detection using Explainable AI, Cryptographic Libraries, Adversarial Machine Learning, Spectral Normalization for Enhancing Robustness of Deep Neural Networks, AI Trojan Attacks and Countermeasures, Hardware Acceleration of Explainable AI, The Future of AI-Enabled Cybersecurity. **Number of participants for the course will be limited to hundred.**
Who should Attend?	▪ Faculty from academic and technical Institutions. ▪ Executives, engineers and researchers from manufacturing, service and government organizations including R&D laboratories. ▪ Student students at all levels (BTech/MSc/MTech/PhD) or Faculty from reputed academic institutions and technical institutions.
Fees	The participation fees for taking the course is as follows: • **Participants from abroad: US $500** • **Industry/ Research Organizations: INR 10000/- +18% GST** • **Faculty members from Academic Institutions: INR 3000/- +18% GST** • **Student: INR 1000/- +18% GST** The above fee includes all instructional materials, computer use for tutorials and assignments, laboratory equipment usage charges, 24 hr free internet facility. The participants will be provided with accommodation on payment basis.

The Faculty

Prabhat Mishra is a Professor in the Department of Computer and Information Science and Engineering and a UF Research Foundation Professor at the University of Florida, where he leads the CISE Embedded Systems Lab. His research interests include embedded systems, design automation, hardware security, energy-aware computing, formal verification, system-on-chip validation, machine learning, and quantum computing. He received his Ph.D. in Computer Science from the University of California at Irvine in 2004. Prof. Mishra has published 9 books, 35 book chapters, 23 patents/copyrights, and more than 200 research articles in premier international journals and conferences. His research has been recognized by several awards including the NSF CAREER Award from the National Science Foundation, IBM Faculty Award, three Best Paper Awards (ISQED'16, VLSID'11 and CODES+ISSS'03) as well as eight Best Paper Nominations (ASPDAC'23, DATE'19, ASPDAC'17, NANOARCH'13, VLSI'13, DATE'12, DAC'09, VLSI'09), and EDAA Outstanding Dissertation Award from the European Design Automation Association. He has also received several awards from UF College of Engineering including Doctoral Dissertation Mentoring Award and International Educator of the Year Award. His research projects are sponsored by both federal agencies (NSF, AFRL, ARO, AFOSR, and DARPA) and industry (SRC, Raytheon, Intel, Cisco, Harris, IBM, Edaptive, Synopsys and Nimbis).

Dr. Chandan Karfa is currently an Associate Professor in the Department of Computer Science and Engineering, IIT Guwahati where he is working since August 2016. Prior to that, he has worked for five years as Sr. R&D Engineer in Synopsys (India) Pvt. Ltd. He has visited New York University during the summer of 2019. His research interests include formal verification, high-level synthesis, hardware security and formal methods. He has published more than senenty research papers in reputed international journals and conferences. He has received Qualcomm Faculty Award from Qualcomm in 2021, the TechnoInventor Award by India Electronics and Semiconductor Association in 2014, Innovative Student Projects Award from the Indian National Academy of Engineers in 2008 and 2013, Best Paper Awards in the ADCOM 2007 and in I-CARE 2013, Microsoft Research India PhD Fellowship in 2008. He is a senior member of IEEE.

For Registration:

https://iitg.ac.in/cet/gian/2024/ckarfa

Course Co-ordinator

Dr. Chandan Karfa
Phone: 0361-2582375
E-mail: ckarfa@iitg.ac.in

http://www.gian.iith.ac.in