Statement of Research Interests

Aydin Aysu

I develop secure AI systems and AI-driven security solutions that prevent advanced cybersecurity threats targeting hardware vulnerabilities. To that end, my research interests lie at the intersection of AI, cryptography, computer architecture, and digital hardware design.

Trusted computing in hardware is fundamental to ensuring information security practices, particularly in the era of artificial intelligence (AI)-driven systems. The foundation of security guarantees in digital infrastructure lies within a hardware root of trust. However, as AI becomes integral to critical systems, it introduces both novel vulnerabilities and new opportunities for securing hardware. Advanced cyberattacks increasingly target the hardware layer, exploiting vulnerabilities that are especially challenging to detect and mitigate from higher abstraction levels. This is particularly acute in security-critical domains such as Cyber-Physical Systems (CPS) and Internet-of-Things (IoT) applications, where the integration of AI heightens both the attack surface and the need for robust defenses.

My research addresses these challenges by analyzing the vulnerabilities of hardware implementations in AI-integrated cyberinfrastructure and designing innovative solutions at the intersection of **AI for security** and **security for AI**. On AI for security, I leverage advanced learning techniques to identify and mitigate hardware threats, such as side-channel attacks and fault injections, enhancing the resilience of security-critical systems. On security for AI, I focus on protecting the integrity and trustworthiness of AI models and their hardware accelerators, ensuring they resist to adversarial attacks and unauthorized access.

Proposed Research Plan

To provide practical and scalable security solutions, my systems emphasize implementation security, hardware/software efficiency, and end-to-end system integration. By combining AI-based analytics with cryptographic principles, I aim to design tools that quantify a provable security level for a given threat model while enabling developers to make automated trade-offs between security, performance, and cost. These efforts ultimately contribute to securing AI systems and leveraging AI to fortify the next generation of trusted computing. To that end, I aim to pursue the following research directions.

1) **Trusted Hardware for AI/ML:** Artificial Intelligence (AI) and Machine Learning (ML) have seen widespread adoptions including safety- and security-critical applications in military, consumer electronics, and healthcare sectors, among others. Enabling trusted execution of AI/ML in hardware is essential for such use-cases. While there is a significant focus on algorithmic and software-level issues, e.g., through adversarial learning [1] and data poisoning [2], hardware aspects of trusted AI/ML are largely unexplored. My research focuses on the trusted hardware design and security enforcement for AI/ML. The goal of this effort is to bring the security concepts of cryptographic hardware to platforms running AI/ML applications. This will effectively broaden the scope of hardware-security research which so far has been limited to cryptographic applications. This is especially important for edge/IoT hardware running AI/ML because adversaries can have physical access to these devices. Transforming lessons from cryptographic attacks/defenses is, however, a non-trivial task as AI/ML has unique compute requirements and building blocks in hardware.

   My CAREER award and SRC-funded projects are on the side-channel analysis of AI/ML hardware. The goals of these projects are to evaluate the impact of side-channel analysis on AI/ML applications and to develop effective defenses. Side-channel attacks can steal trained AI/ML models that are valuable intellectual property (IP). My research has shown that such attacks are even possible on highly-parallelized hardware and are much more effective than mathematical/theoretical attacks using input-output queries [3]. We have also built side-channel countermeasures by transforming solutions used in cryptographic hardware such as masking and hiding to the AI/ML workloads [3–6]. My long-term vision in this thrust is to build provably-secure and automated side-channel mitigation techniques that allow push-button security and seamless integration to common libraries such as TensorFlow.

Side-channel analysis is just one example of research in this thrust. As I have articulated in position papers [7, 8], there are many other attack vectors such as fault injection attacks, cold boot attacks, hardware Trojans, logic locking, logic encryption/obfuscation, probing and bus snooping attacks, and (micro-)architectural attacks which have been thoroughly studied on cryptographic workloads but are largely unknown for hardware running AI/ML applications. For example, our recent work has shown that while scan-chain attacks are straightforward on cryptographic hardware, they only scale towards AI/ML hardware when coupled with a novel algebraic analysis [9]. There are numerous opportunities in extending and protecting against all these different types of attacks on AI/ML hardware.

I have a particular interest in exploring fault injection attacks in the future. These attacks can disrupt the behavior of the device and cause faulty computations. Such attacks have been shown on cryptographic applications to extract secret keys or to evade access control mechanisms. They can cause critical misclassifications in AI/ML and have the potential to be much more effective than adversarial attacks. Defenses built for adversarial attacks thus cannot protect fault injection attacks by default. Interestingly, recent works have shown that fault injection attacks can execute remotely with software through dynamic-voltage frequency scaling (DVFS) interfaces [10, 11], alleviating the need to have physical access. I recently won an Office of Naval Research (ONR) award and another industry award on this topic that aims exposing and mitigating software-induced fault injection attacks on critical cyberinfrastructure and chiplets running AI/ML algorithms. The project will characterize the effects of these attacks, build a fault injection simulator for modeling them, and develop fault-injection-aware training to generate resilient neural networks by taking the attacks' effects into account.

2) **AI/ML for Security: Efficient and Secure Post-Quantum Cryptosystems.** Large-scale communication protocols in use today base their cryptographic security on the difficulty of solving mathematical problems such as integer factorization. Quantum algorithms, however, are proven to solve these problems quickly (in polynomial time)—quantum computers can thus break current cryptographic systems. Recent developments in quantum computing technologies have therefore spurred significant interest in post-quantum (PQ) cryptography alternatives basing security on other mathematical problems such as the shortest vector problem. Recent events showed the growing importance of this field. National Institute of Standards and Technology (NIST) has been standardizing PQ algorithms [12] for a large-scale transition from existing to quantum-secure protocols, which is now underway with significant industry support.

There are two major problems with current PQ cryptosystems. And both can be addressed with AI/ML. First, the algorithms are rather complex and their optimized implementations, especially for constrained/real-time systems like edge/IoT devices, are challenging. My PhD research in this field has resulted in seminal papers. I proposed the first hardware optimization techniques on the fundamental compute unit in lattice-based post-quantum cryptography—the number theoretic transform (NTT)—that resulted in an improved memory organization and datapath area-performance trade-offs, which are now common practice in hardware and software designs. I later showed pre-computation techniques in software that can achieve over 10× efficiency in energy for energy harvesting systems [13], and over 100× reduction in latency for real-time applications using hardware/software co-design [14]. Although these works made PQ more practical, they represent ad-hoc, point solutions.

The second major challenge of PQ systems is implementation security. Although these algorithms provide theoretical guarantees, their practical implementations can be vulnerable to side-channel attacks. Such attacks exploit the correlation of secret keys to implementation characteristics like execution time, power consumption, or memory access patterns. Physical (hardware-based) side-channels are especially important and difficult to mitigate in embedded settings since the adversary can have physical access to the device. Even after decades of intense study, side-channels on traditional cryptosystems are still an active area of research. Extending side-channel attacks and countermeasures to PQ algorithms is a non-trivial task, as the majority of these new proposals use fundamentally different arithmetic constructions. My post-doctoral research has demonstrated the *first side-channel attacks*

on post-quantum key exchange protocols [15], breaking algorithms used by Germany and Google, and evaluated at NIST. I have also built low-cost countermeasures through algorithm-specific features [16].

As an assistant professor, I was awarded an NSF CRII and a CAEML NSF IUCRC project as a sole-PI on the use of AI/ML for hardware efficiency and security for PQ cryptosystems. I recently won two new NSF SaTC awards on these themes. I am leading a team of graduate students who demonstrated the first side-channel vulnerabilities in 7 post-quantum protocols evaluated at NIST [17–22]. I am also designing a range of flexible, efficient, and side-channel resilient hardware, software, and architectural support solutions [23–26]. This research showed that AI/ML can be leveraged to optimize such solutions and to secure them. My ongoing research aims furthering this thrust in the context of end-to-end applications like the PQ Transport Layer Security (TLS) protocol and on emerging PQ algorithms. At the same time, I work on extending these techniques to other lattice-based cryptosystems such as homomorphic encryption.

To pursue my future research plans in this topic, I seek collaborations within or across the department on VLSI, circuits, and architecture design to tapeout the PQ encryption chips along with energy optimization and side-channel security. To that end, I see approximate computing as a special enabler for low-energy designs that have not yet explored. Traditional cryptosystems cannot use approximate computing as a single-bit difference within cryptographic computations would reveal completely different results. However, lattice-based and coding-based PQ constructions such as Learning-with-Errors (LWE) or Medium-Dense-Parity-Check (MDPC) codes work by introducing errors into computations and recovering them later on, providing an opportunity for approximately computed cryptographic systems. More broadly, I also seek collaborations with theoretical cryptographers to develop efficient and secure implementations of their cryptographic constructions. Research on this thrust would require EDA tools for chip tape-out and measurement equipment like high-end oscilloscopes, electromagnetic probes, and microscopes to investigate various side-channels such as power consumption, electromagnetic radiation, and photonic emissions.

3) **Secure Architectures for Heterogeneous/FPGA Cloud Servers.** Multi-tenant use in the cloud servers has well-known security issues and much research has been conducted to mitigate them. But the FPGA usage in cloud is relatively new and cloud providers are starting to experiment with multi-tenant use in cloud FPGA, which means two applications can share the same FPGA fabric at the same time (spatial tenancy), or applications can be paused, moved in and out, and resumed in a time multiplex manner (temporal tenancy). FPGAs enable configuring hardware and thus have superior flexibility and performance compared to multi-core or GPU-based design. This configurability, however, can introduce vulnerabilities on cloud FPGA that doesn't exist for other systems.

My currently funded ONR project focuses on secure and safe virtualization of cloud FPGA-based heterogeneous servers. The goal is to provide memory isolation, reconfiguration determinism, and denial-of-service protection for multi-tenant cloud FPGAs. Since multi-tenancy is not yet supported by commercial cloud FPGA tools, such security and safety aspects currently do not exist in practice. For example, a (temporal) tenant can read the earlier tenants' residual data without an explicit isolation mechanism, or a tenant can tamper with the bus infrastructure to avoid or delay reconfiguration needs of incoming applications, or even use excessive power to cause device shutdown. My research will explore such unique attack vectors stemming from cloud FPGA-based heterogeneous applications and related defenses for the detection or mitigation of such threats.

Another research direction I pursue is on "remote" physical side-channel attacks on the cloud FPGAs. The configurability of FPGA gives adversaries the capability to program a time-to-digital converter hardware and to infer the power consumption of the entire platform. This in turn can leak sensitive information about the spatial tenant running on the FPGA [27] or even other attached GPU/CPU components [28]. Unfortunately, academic and industrial secure computer architecture solutions like Aegis, Sanctum, Intel SGX, or Arm TrustZone do not consider physical side-channel attacks in their threat model [29]. Therefore, there are no off-the-shelf architectural defenses and I aim

to explore a combination of such defenses with RTL- and netlist-level checks for malicious designs.

Although physical side-channels are typically considered as an attack vector, for applications that use no secret information, analog behavior such as power consumption can also be an integrity check mechanism—an orthogonal method to evaluate if the device indeed follows the desired set of operations with the desired data. Therefore, I envisage that the use of hardware behavior fingerprinting combined with advanced AI/ML classifiers to be an efficient and *architecture-agnostic* identifier of otherwise difficult to detect zero-day attacks. My future research is to extend this approach to applications in smart grids, automotive, aerial, advanced manufacturing, and wearable/bio-implementable circuits. To realize this vision, I seek collaboration with researchers in or across the departments with domain-specific knowledge about these applications or about system security. This research would require setting up an infrastructure/environment to test the related devices or systems in such applications and carrying out the attacks.

4) **AI-Guided Automation for Hardware Security.** Unfortunately, security evaluation and countermeasures for AI and cryptographic systems are carried out manually and in an ad-hoc manner for each setting. For example, research on AI system side-channels requires a domain expert to fully understand new algorithms, to know how to implement them on specific platforms, to figure out the associated side-channel vulnerabilities, to propose new countermeasures for effectively mitigating vulnerabilities, and to finally evaluate the proposed solution thoroughly on the target platform with respect to some metric/method. Given that there are N algorithms, M possible implementations, and P side-channel attacks, there is a space of N×M×P configurations to evaluate. Performing an entire side-channel evaluation for a single configuration is typically sufficient today to publish a paper at premier security conferences. Even for that single setting, the evaluation process is error-prone hence each year there is yet another analysis/improvement on prior work. While this procedure may be possible in the short-term, e.g. for PQ cryptosystems or AI applications, we must develop new tools to automate security analysis. For hardware implementations, I envisage the use of high-level synthesis (HLS) tools and AI/ML based reasoning to produce secure hardware. These tools generate a hardware design from a high-level description like a C program. Recent work showed that existing HLS tools can provide a reasonable design compared to hand-coded hardware for cryptographic applications [30] and academic tools show a similar success for limited use cases [31]. No prior work, however, considered hardware security aspects in their analysis. The main challenge is to express hardware security properties into the tools in such a way that the resulting hardware will have formal guarantees. This research, therefore, requires the collaboration of a domain expert like me with researchers working on electronic design automation (EDA), and test and verification.

Funding Opportunities

Funding opportunities for cybersecurity research and AI are plentiful. I have thus far obtained about $5.8M in funding ($4.4M in personal share) from a range of sources from Department of Defense (DoD) agencies, National Science Foundation, and industry. I have pending proposals totaling about $2M in personal share submitted to such sponsors.

Security against quantum cryptanalysis is a *national security* issue because quantum computers are likely to be developed by motivated nation states to break into military-grade encryptions. Likewise, hardware security attacks in general and hardware supply-chain problems are likely to be orchestrated by advanced, government-funded organizations. Therefore, DoD is likely to fund this line of research. My engagements in the DoD sphere include AFOSR, DARPA, ONR, and HSARPA. Hardware security flaws in wearable/bio-implementable devices have a direct impact on healthcare, hence this research has potential for National Institute for Health (*NIH*) proposals. My research has also been sponsored by Semiconductor Research Corporation (*SRC*) industry liaisons and I have collaborated with researchers in charge of other cybersecurity funding programs at *Intel*, *Google*, *CISCO*, *NXP*, *Lockheed Martin* and *NIST*.

References

[1] D. Lowd and C. Meek, "Adversarial learning," in *Proceedings of the Eleventh ACM SIGKDD International Conference on Knowledge Discovery in Data Mining*, ser. KDD '05. New York, NY, USA: Association for Computing Machinery, 2005, p. 641–647. [Online]. Available: https://doi.org/10.1145/1081870.1081950

[2] B. Biggio, B. Nelson, and P. Laskov, "Support vector machines under adversarial label noise," in *Asian conference on machine learning*. PMLR, 2011, pp. 97–112.

[3] A. Dubey, R. Cammarota, and **A. Aysu**, "Maskednet: The first hardware inference engine aiming power side-channel protection," in *2020 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, 2020, pp. 197–208.

[4] A. Dubey, R. Cammarota, and **A. Aysu**, "Bomanet: Boolean masking of an entire neural network," in *2020 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*, 2020, pp. 1–9.

[5] A. Dubey, A. Ahmad, M. A. Pasha, R. Cammarota, and **A. Aysu**, "Modulonet: Neural networks meet modular arithmetic for efficient hardware masking," *IACR Transactions on Cryptographic Hardware and Embedded Systems*, vol. 2022, no. 1, p. 506–556, Nov. 2021. [Online]. Available: https://tches.iacr.org/index.php/TCHES/article/view/9306

[6] A. Dubey, R. Cammarota, V. Suresh, and **A. Aysu**, "Guarding machine learning hardware against physical side-channel attacks," 2021, accepted to ACM JETC.

[7] R. Cammarota and **A. Aysu** et al., "Trustworthy ai inference systems: An industry research view," 2020.

[8] Regazzoni, F., **A. Aysu** et al., "Machine learning and hardware security: Challenges and opportunities -invited talk-," in *2020 IEEE/ACM International Conference On Computer Aided Design (ICCAD)*, 2020, pp. 1–6.

[9] S. Potluri and **A. Aysu**, "Stealing neural network models through the scan chain: A new threat for ml hardware." 2021.

[10] K. Murdock, D. Oswald, F. D. Garcia, J. Van Bulck, D. Gruss, and F. Piessens, "Plundervolt: Software-based fault injection attacks against intel sgx," in *Proceedings of the 41st IEEE Symposium on Security and Privacy (S&P'20)*, 2020.

[11] A. Tang, S. Sethumadhavan, and S. Stolfo, "CLKSCREW: Exposing the perils of security-oblivious energy management," in *26th USENIX Security Symposium (USENIX Security 17)*. Vancouver, BC: USENIX Association, Aug. 2017, pp. 1057–1074.

[12] National Institute of Standards and Technology, "Post-Quantum cryptography," https://csrc.nist.gov/Projects/Post-Quantum-Cryptography.

[13] **A. Aysu** and P. Schaumont, "Precomputation methods for hash-based signatures on energy-harvesting platforms," *IEEE Transactions on Computers*, vol. 65, no. 9, pp. 2925–2931, 2016.

[14] **A. Aysu**, B. Yuce, and P. Schaumont, "The future of real-time security: Latency-optimized lattice-based digital signatures," *ACM Trans. Embed. Comput. Syst.*, vol. 14, no. 3, pp. 43:1–43:18, Apr. 2015.

[15] **A. Aysu**, Y. Tobah, M. Tiwari, A. Gerstlauer, and M. Orshansky, "Horizontal side-channel vulnerabilities of post-quantum key exchange protocols," in *2018 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, 2018, pp. 81–88.

[16] **A. Aysu**, M. Tiwari, and M. Orshansky, "A novel hardware design for binary Ring-LWE with power side-channel countermeasures," in *2018 Design, Automation Test in Europe Conference Exhibition (DATE)*, Accepted 2018.

[17] P. Kashyap, F. Aydin, S. Potluri, P. D. Franzon, and **A. Aysu**, "2Deep: Enhancing side-channel attacks on lattice-based key-exchange via 2d deep learning," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 2020.

[18] F. Aydin, P. Kashyap, S. Potluri, P. Franzon, and **A. Aysu**, "DeePar-SCA: Breaking parallel architectures of lattice cryptography via learning based side-channel attacks," in *Embedded Computer Systems: Architectures, Modeling, and Simulation*, A. Orailoglu, M. Jung, and M. Reichenbach, Eds., 2020, pp. 262–280.

[19] E. Karabulut and **A. Aysu**, "Falcon down: Breaking falcon post-quantum signature scheme through side-channel attacks," 2021, accepted.

[20] E. Karabulut, E. Alkim, and **A. Aysu**, "Single-trace side-channel attacks on w-small polynomial sampling," in *Accepted to the 2021 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, 2021, accepted.

[21] F. Aydin, **A. Aysu**, M. Tiwari, A. Gerstlauer, and M. Orshansky, "Horizontal side-channel vulnerabilities of post-quantum key exchange and encapsulation protocols," *ACM Trans. Embed. Comput. Syst.*, vol. 20, no. 6, oct 2021. [Online]. Available: https://doi.org/10.1145/3476799

[22] Z. Chen, E. Karabulut, **A. Aysu**, Y. Ma, and J. Jing, "An efficient non-profiled side-channel attack on the crystals-dilithium post-quantum signature," 2021.

[23] E. Karabulut, E. Alkim, and **A. Aysu**, "Efficient,flexible,and constant-time gaussian sampling hardware for lattice cryptography," *IEEE Transactions on Computers*, pp. 1–1, 2021.

[24] A. C. Mert, E. Karabulut, E. Ozturk, E. Savas, and **A. Aysu**, "An extensive study of flexible design methods for the number theoretic transform," *IEEE Transactions on Computers*, pp. 1–1, 2020.

[25] E. Karabulut and **A. Aysu**, "Rantt: A risc-v architecture extension for the number theoretic transform," in *2020 30th International Conference on Field-Programmable Logic and Applications (FPL)*, 2020, pp. 26–32.

[26] A. C. Mert, E. Karabulut, E. Öztürk, E. Savaş, M. Becchi, and **A. Aysu**, "A flexible and scalable ntt hardware : Applications from homomorphically encrypted deep learning to post-quantum cryptography," in *2020 Design, Automation Test in Europe Conference Exhibition (DATE)*, 2020, pp. 346–351.

[27] F. Schellenberg, D. R. Gnad, A. Moradi, and M. B. Tahoori, "An inside job: Remote power analysis attacks on fpgas," in *2018 Design, Automation Test in Europe Conference Exhibition (DATE)*, 2018, pp. 1111–1116.

[28] I. Giechaskiel, K. B. Rasmussen, and J. Szefer, "C¡sup¿3¡/sup¿apsule: Cross-fpga covert-channel attacks through power supply unit leakage," in *2020 IEEE Symposium on Security and Privacy (SP)*, 2020, pp. 1728–1741.

[29] V. Costan and S. Devadas, "Intel SGX explained." *IACR Cryptology ePrint Archive*, vol. 2016, p. 86, 2016.

[30] E. Homsirikamol and K. Gaj, "Can high-level synthesis compete against a hand-written code in the cryptographic domain? A case study," in *ReConFigurable Computing and FPGAs (ReConFig), 2014*. IEEE, 2014, pp. 1–8.

[31] A. Khalid, M. Hassan, G. Paul, and A. Chattopadhyay, "Runfein: a rapid prototyping framework for feistel and spn-based block ciphers," *Journal of Cryptographic Engineering*, vol. 6, no. 4, pp. 299–323, Nov 2016.