# GRIDSHIELD: Detecting Patient Zero Threats in Distributed Energy Ecosystem

**Amin Kharraz, Florida International University**
**mkharraz@fiu.edu**

**Overview.** Identifying previously unseen adversarial operations, Patient Zero Threats, has never been a trivial task. These threats are continuously getting more efficient and scalable, and their destructive impact are increasingly getting more consequential. This issue has become even more challenging in distributed computing environments such as power grids where millions of independent smart devices should be integrated into critical infrastructure. While adversaries have significant freedom to scan potential attack surfaces and identify weaknesses, the domain knowledge about these operations is still very abstract. This lack of knowledge places security defenders at a distinct disadvantage, making it very difficult to verify the trustworthiness and security posture of mission-critical energy systems in a real-time fashion. This proposal aims to close this gap by developing an augmented defense layer to the distributed energy systems, called GRIDSHIELD, for real-time attack inferencing and low-latency response. GRIDSHIELD provides three crucial enhancements over the state-of-the-art: (1) GRIDSHIELD offers fine-grained forensics data about run-time behavior of cyber-physical systems integrated into the energy grids without introducing any human intervention or changing the underlying logics of involved energy systems, (2) GRIDSHIELD provides an unsupervised approach for this problem space. Supervised learning methods fall short to address zero patient threats in this rapidly evolving ecosystem because modern attacks can deviate easily from those seen before and new forms of attacks constantly emerge, (3) understanding the nature and veracity of new attacks is largely manual and often requires significant human intervention. This introduces additional overhead that further increases incident response latency and reduces defense agility. GRIDSHIELD's project offers AI-based techniques to quickly analyze temporal artifacts and locate entities once diverging from the expected behaviors.

**Intellectual Merits:** The goal of this proposal is to develop foundational principles that will guide the development of techniques to improve our defense agility against unknown attack on energy systems.To achieve this goal, PI Kharraz's team will invest efforts along the following components to fill the research gaps in four interconnected areas:

- **Component I: Run-time Forensics Engine for Distributed Energy Systems (§2.3.1)** development of a novel forensics engine at the kernel level to generate spatio-temporal artifacts about cyber-physical systems integrated into the power grids to contextualize attacks without imposing any changes in the underlying semantics of the underlying infrastructure;
- **Component II: Energy System Integration, Trace Collection, and Cataloging (§2.3.2)** design of novel frameworks to deploy GRIDSHIELD's defense service on energy systems using a customized OS library that exposes GRIDSHIELD's features. GRIDSHIELD also proposes a novel encoding mechanism to build a generalizable representation for the collected artifacts for automatic behavioral cataloging.
- **Component III: Unsupervised Detection and Human-Directed Response (§2.3.3)** propose new algorithms to advance the state-of-the-art in event aggregation and correlation and synthesize knowledge from disparate sources of dirty information to quickly assess behavioral intents of the connected devices, predict divergence from expected behavior, and generate automated responses;
- **Component IV: Case studies (§2.3.4)** evaluate the utility of GRIDSHIELD, the abstraction method, the integrated interfaces at various software layers. The project will also develop empirical methods to analyze the output of GRIDSHIELD and build a rich suite of software-chain to streamline the deployment workflow.

**Broader Impacts.** The overall impact of the GRIDSHIELD project is potentially substantial. Implementing integrated defensive capabilities in energy systems similar to the GRIDSHIELD project will develop important research underpinnings for making more resilient and trustworthy cyberspace. It can also preempt adversarial operations before successful attacks damage mission-critical cyber-infrastructure. Furthermore, the project will provide a natural mechanism to strengthen diversity in cybersecurity research. The PI works with one female African-American Ph.D. student and two Hispanic students. If funded, the project will support the student who has just started her Ph.D. in spring 2022.