Research Statement

Anika Tabassum

Which power grids or substations are the most critical and in need of immediate action to recover during a natural disaster? Which regions are highly affected during a power failure? What are the similarities and differences in the power failures occurring in a certain region? How can we model the energy-usage pattern to plan for building microgrid systems (electric generators, solar plants, etc.) in the neighborhood? I aim to develop explainable and robust frameworks to solve the challenges that emerge from Cyber-Physical Systems (CPS) and energy-grid with extensive applications for emergency management and Power System Domain Experts (DEs). My interests lie in utilizing Explainable Machine Learning (ML), Domain-Informed ML, and optimization techniques to solve computational challenges for emerging applications from CPS. My strengths are in developing novel problems, models, and algorithms for solving the computational challenges that have arisen from our increasing knowledge of Cyber-Physical Systems and energy-grid data.

1. Research Philosophies.

a. **Developing domain-independent explainability frameworks.** Explainability in CPS refers to systematically providing answers to the DEs to questions concerning data and events relevant to the subject at hand, which can assist DEs in improving grid resiliency as well as minimize the impact of a disaster. The rule-based techniques used as explanations may vary for different domains (e.g., rules in power systems are different from those in healthcare). Thus, it is necessary to formulate a problem independent of the domain that can provide explanations for the respective domain. For example, to explain which power grids are most critical during a disaster we developed a model [1] to identify different phases during a disaster and pinpoint critical regions (grids) at different phases. To make the problem domain-independent (in this case, the domain is the power system), we formulate the problem in a manner such that it takes multivariate sequences and graph structures among the respective sequences as inputs. Consequently, our model can be modified to identify disaster phases not only for power failures, but also failures concerning other CPS (where multivariate sequences can be a set of sensors and graphs can be connections among the respective sensors). Our model, however, cannot be utilized for answering a number of other relevant questions, as developing explainability may vary depending on the type of answers to provide, and, as such, mapping domain characteristics to adapt for generalization can be rather challenging and diverse. *Formulating domain-independent explainability problems* to *answer questions like 'when', 'what', 'why', and 'where' can offer an in-depth understanding of any CPS vulnerabilities for domain experts.*

b. **Modeling uncertainty to build robust frameworks.** Developing robust models using ML for any domain in CPS can be quite challenging due to the presence of uncertainties. For example, in [3], we developed a "what-if" simulation tool that analyzes national critical infrastructures (CIs)[1] by mapping the data into a heterogeneous network. Our goal was to identify the critical components that can cause significant damage (vulnerable) to disasters and exploitations (e.g., cyber-attacks). Due to limited domain knowledge and lack of exogenous variables (e.g., the

[1] Critical infrastructures (CIs) (a sector of CPS), refer to systems, technologies, facilities, and networks that are vital to security, public health, and socio-economic well being of people, e.g., power, transportation, communication, etc.

weather), there exist multiple uncertainties in graph construction, parameter inputs, and examples of the like, which makes it difficult to attain the final objective. The incorporation of these uncertainties into the network, however, can lead to the development of more robust models to predict and measure critical and vulnerable components in the network. *My aim is to identify and model uncertainties to build robust frameworks for analyzing and understanding data.*

c. **Leveraging high-performance computing (HPC) for large scale data.** Leveraging HPC for large-scale data can accelerate high complexity optimization processes and solve ML problems on supercomputers considerably faster. In [3] to analyze criticalities and vulnerabilities among different CI components, we developed analytic modules using graph-theoretic approaches. To efficiently handle these modules for graph-theoretic analysis, we leverage ORNL's supercomputers. In [2] we implemented a shared-memory parallel processing implementation to analyze the patterns on high dimensional time-series data faster. *My objective is to build methodologies for real-world large-scale data by leveraging supercomputers*.

d. **Interdisciplinary collaboration.** Building explainable and robust models and making them adaptable for and available to domain experts requires heavy interdisciplinary collaboration. Based on my experience, I can confidently state that my works have been greatly benefited from strong collaboration among multidisciplinary teams that included emergency management operatives, power engineers, and epidemiologists. My works [1, 3, 4] were utilized by multidisciplinary teams at ORNL, including ML and visualization experts, HPC, IoT, and power engineers. *I intend to make my future projects adaptable for DEs and hence aim to collaborate closely with interdisciplinary teams.*

2. Recent relevant projects

a. **Explainability as identifying disaster phases and pinpointing vulnerabilities by characterizing failures.** We characterize the severity of power failure by measuring the number of outages in a hurricane-affected region (a county, in our case) over the entire period of the hurricane. Three critical questions need to be answered for the characterization of this process. (i) How can we identify different phases of a hurricane as a function of the severity of the damage (sudden changes) to Critical Infrastructures (CIs) (such as power grids) using sparse power outage data? (ii) Which counties are the culprits for characterizing each phase? (iii) How can the counties be grouped based on their overall failure dynamics during the hurricane? **The goal of this project** is to help emergency management DEs answering the aforementioned questions.

Our algorithm, Cut-n-Reveal [1], contains two parts: detecting a good segmentation of the power loss data to capture the sudden changes, and finding the corresponding explanations (a subset of culprit counties) per segment. For segmentation, we assume a known underlying graph structure that captures the relationship among these power failure time-series. E.g., the number of electricity outages in all counties affected by disaster can be used to form a set of time series. The relationship among the counties can be based on geographical proximity. *With this knowledge of the segmentation and the explanations for each segment, the expert has a holistic picture of the different phases of the failure process and the specific time series that contributed significantly to each phase change.* Fig. 1(a) depicts the overall temporal segmentation or phases of Hurricane

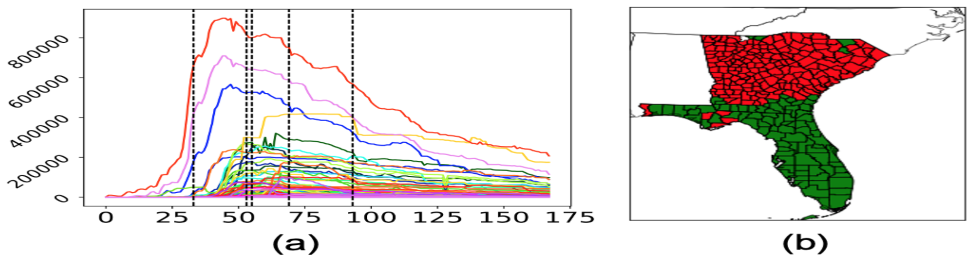

Fig 1: An example of the holistic spatial and temporal analysis results from our novel CnR model to characterize the damage to the power grid during Hurricane Irma.

Irma over the power failure data, where each time series indicates the total number of households that lost power in a single county throughout the hurricane. Fig. 1(b) represents the spatial clustering of all counties that experienced significant damage (in green) to power infrastructures throughout the course of the hurricane.

Existing works in temporal segmentation are limited by the fact that only a few of them are able to easily incorporate spatial information. That none of the existing models provides an explanation framework wherein the 'culprit' counties (at each segment of a power failure) can be identified in space and time merely aggravates the situation and further exposes the limitations of these existing works.

b. **Identifying critical components in what-if simulation tools through failure dynamics.** The aim is to model uncertainty levels over the inter-dependency and failure dynamics present in the power systems by leveraging ORNL's "what-if" simulation tool Urban-Net, which is built on the idea of viewing the heterogenous power systems as large heterogeneous networks [2]. The intent behind the "what-if" simulation tool is to measure the degree to which the criticality of a network (i.e., power system components) is affected by different exogenous variables such as extreme weather or cyber-attacks. However, measuring this criticality can be challenging due to the presence of multiple uncertainties, which may take different forms of failure rates among different components, impacts of exogenous events, or protection systems to prevent damage to a component. We propose a problem to measure the criticality in presence of hidden failures, where the criticality is defined as "the nodes whose failure causes the maximum damage over the network [5]". *We introduce randomness over existing cascade behavior in networks and develop optimization over randomness to find the critical nodes.* The existing power simulation tools are generally restricted to only one type of component within a smaller region, and are quite slow. *Where even parallel computation using a supercomputer takes approximately 10 days to identify the critical components, our model can adapt to all types of power systems over the US and can measure the criticality over far-reaching regions within approximately 10 mins.*

3. Future research.

My future research plans pertain to solving explainability, prediction, and uncertainty problems in the cyber-physical domain using deep learning. Considering that my recent projects are closely related to power systems, the motivation behind the formulations of these problems have been induced from the energy domain, whereas my goal for future projects lies in formulating domain-independent problems and tractability for generalization.

a. **Developing explainability with human-in-the-loop.** The motivation behind this project stems from the need to make explainability more easily understandable to DEs by providing simple explanations instead of intricate models or a group of features for understanding the behaviors of the models. DEs are more interested in knowing what violations in power grids cause such failures instead of model behaviors. There are, of course, challenges to be faced in the process of developing such explanations from power failure sequences during a disaster: (i) There can be hundreds of conditions whose violations may cause power failure, thus making it difficult to provide rule-based explanations. (ii) The conditions may vary depending on the scenario. One idea to develop such explainability is to get a set of scores from DEs for ideal conditions of power grids. Along with the number of power failures for every region (e.g., as the data described in Sec. 2a), we can map the condition of the power grids for each region at each timestamp as a set of values. For each disaster phase, then, we can characterize the culprit time-series based on the set of values, which can be solved using inverse Reinforcement Learning (RL) ideas where expert rewards are based on a set of condition values of power grids and characterization of the time-series (using inverse RL) in terms of different conditions. From these different conditional values generated by inverse RL, DEs can discover the condition that has been violated in a particular scenario.

b. **Prediction model incorporating domain knowledge.** The power generation capacities of power plants depend on several exogenous variables (e.g., renewable or non-renewable energy sources, wind, water, natural gas, etc.). If the generation capacities can be predicted in time, it will be far easier for the DEs to strategically determine how they will meet the demands for utilities. However, building a data-driven predictive model from these exogenous variables is not enough, as multiple unseen variables (such as climate change, pollution, rust, etc.) can have a slower impact on exogenous variables as well as generation capacity. I plan to design a prediction model by incorporating the impact of these unseen sources with a data-driven model, possibly by incorporating climate change physics therein.

c. **Uncertainty estimation with ML in CPS.** Finding domain knowledge to model the uncertainties that pose challenges to understanding energy-grid failures (using a "what-if" simulation tool) may itself be difficult due to the multitude of both seen and unseen variables present, along with the fact that DEs cannot always have clear and detailed answers to all questions. My project, thus, aims to model such uncertainties based on observed data, and thereafter incorporate these uncertainties into the model to analyze failures. This project can be beneficial not only for energy grids or CPS but also can be used for computational purposes by different fields.

4. Impact of my research.

Industrial Impact. My research has a direct impact on improving awareness concerning emergency management situations and augmenting the power-flow simulation tools. Our projects to identify the different phases of a natural disaster by characterizing power failure sequences [1, 2] were motivated by the need to integrate the Oak Ridge National Laboratory **(ORNL) energy monitoring tool VERDE** with the **situation-awareness tool EAGLE-I**. Our efforts to pinpoint domain-based critical components from time-series data can capture the states in which non-pharmaceutical interventions (NPI) were declared exactly 2 weeks prior due to COVID-19. Epidemiologists, the CDC, as well as the FEMA will find this model especially useful, as it will allow for the incorporation of data concerning the NPIs, for enabling forecasting models to forecast COVID-19 cases and deaths. My recent work**,** inspired by the ORNL's DOE-based project, namely, the **North American Energy Resiliency Model (NAERM)**, was aimed towards identifying and evaluating the threat mitigation strategies in energy grids.

Academic Impact. My work is equally influential in academia. My research has been funded by **NSF NRT Urban Computing Fellowship** which is approximately **1M USD** funding. Besides, as a team, our model DEEP OUTBREAK for forecasting COVID-19 hospitalizations and mortality won *1st prize* of worth **50k USD** in the COVID-Facebook Symptom Data Challenge 2020.

My research goals are not only associated with emergency management or modeling uncertainty levels for energy grids, but also with adaptability for the systems discussed, for adaptability is especially pertinent to other emerging scientific applications including, but not limited to, Cyber-Physical Systems, health-care, and scientific applications.

References

[1] Nikhil Muralidhar, **Anika Tabassum**, Liangzhe Chen, Supriya Chinthavali, Naren Ramakrishnan, and B. Aditya Prakash. Cut-n-Reveal: Timeseries segmentations with explanations. ACM Transactions on Intelligent Systems and Technology. May 2020.

[2] **Anika Tabassum**, Supriya Chinthavali, Tansakul Varisara, B. Aditya Prakash. Rationalizing Time-series Segmentations. [in submission].

[3] **Anika Tabassum**, Supriya Chinthavali, Sangkeun Lee, Liangzhe Chen, B. Aditya Prakash. Urban-Net: A System to Understand and Analyze Critical Infrastructure Networks for Emergency Management. ACM SIGKDD 2019.

[4] Supriya Chinthavali, Varisara Tansakul, Sangkeun Lee, **Anika Tabassum**, Jeff Munk, Jan Jakowski, Michael Starke, Teja Kuruganti, Heather Buckberry, JimLeverette. Quantification of Energy Cost Savings through Optimization and Control of Appliances within Smart Neighborhood Homes. ACM International Workshop on Urban Building Energy Sensing, (UrbSys), 2019.

[5] **Anika Tabassum**, Supriya Chinthavali, Sangkeun Lee, Bill Kay, Stenvig Nils, B. Aditya Prakash. Efficient Contingency Analysis in Power System via Network Trigger Nodes. [in submission].