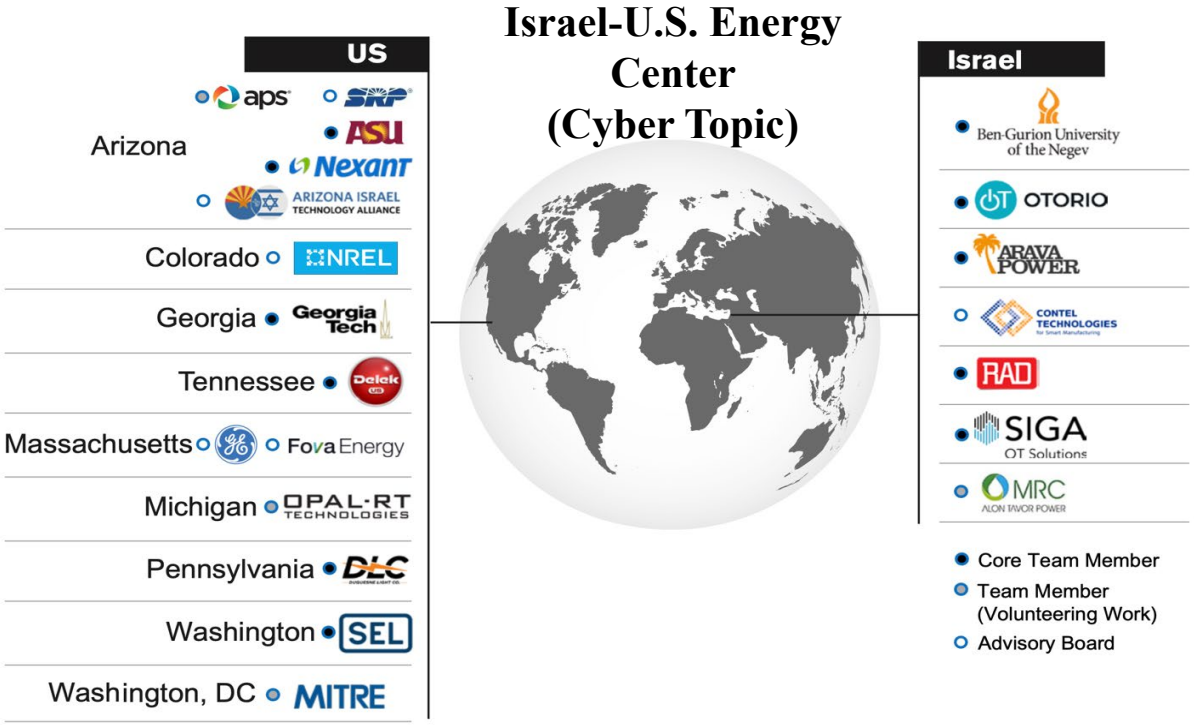
Comprehensive Cybersecurity Technology for Critical Power Infrastructure AI-Based Centralized Defense and Edge Resilience



Prepared for

Itai Ganzer and **Ofer Goldhirsh**

Israel Innovation Authority

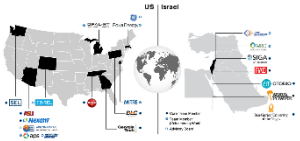
Avi Shavit and **Eynan Lichterman**

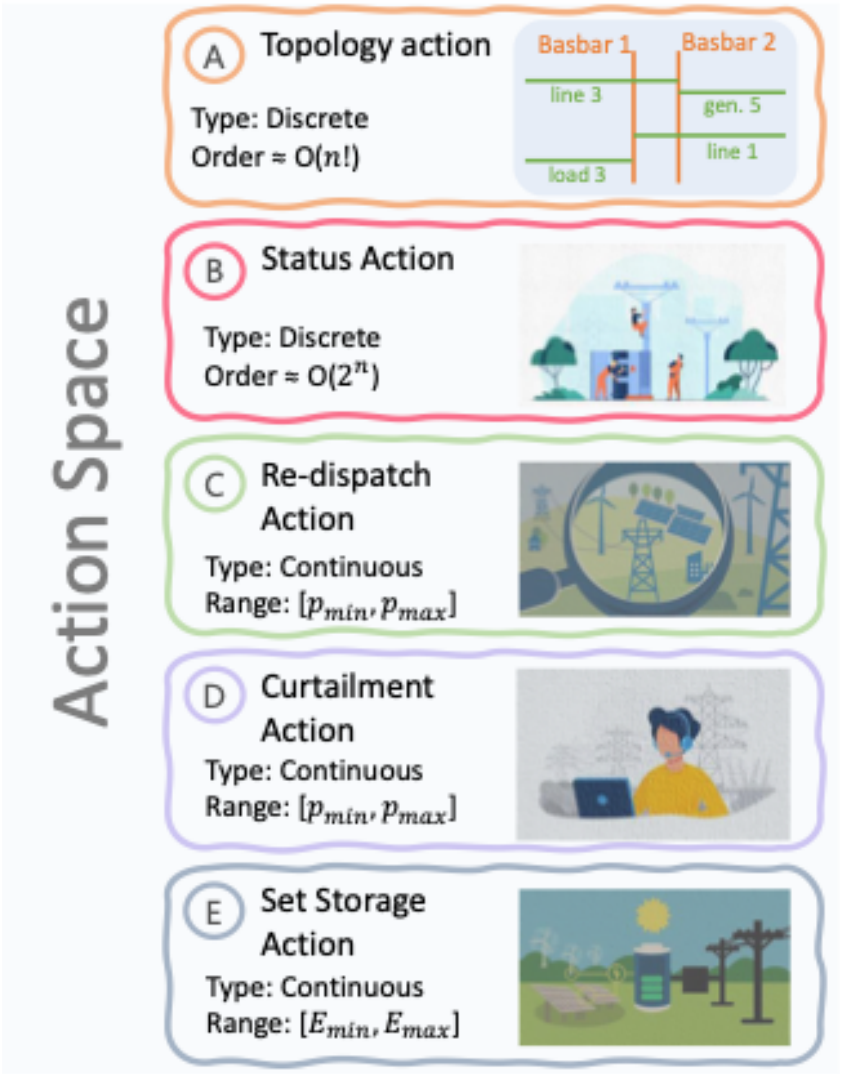
Israel Ministry of Energy

Task 16: Reinforcement Learning Control for Cyber Physical Systems

- Students: Mohammadamin Moradi, Zheng-Meng Zhai (new)
- Task Lead: Dr. Y.-C. Lai

Large Power Grids: Large and Diverse Action Space





Discrete actions:

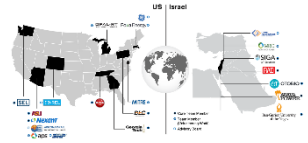
- *Topology actions*: changing the topology of certain substations (TG)
- *Status actions*: transmission or power line switching (PLS)

Continuous actions:

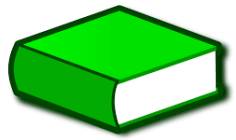
- *Redispatch actions*: changing the operating schedule of power plants
- *Curtailment actions*: limiting the production of renewable generators
- *Set-storage actions*: changing the role of some storage units from loads to generators or vice versa

Example: IEEE 118-Bus system: about 12 million possible actions

Reinforcement Learning CPS Control Analogy

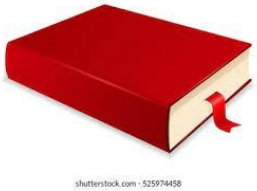


Agent #1



Topology actions

Agent #2



Status actions

Agent #3

Redispatch actions

Agent #4

Curtailment actions

Agent #5

Set-storage actions

Different subspaces of action

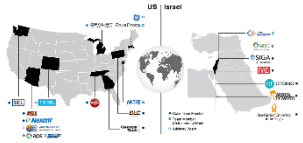
Test includes questions from the 5 books
(Grid under attack)

RL Environment

Goal: maximizing reward or grid survival

Coordination: Temporal Graph Convolutional Network

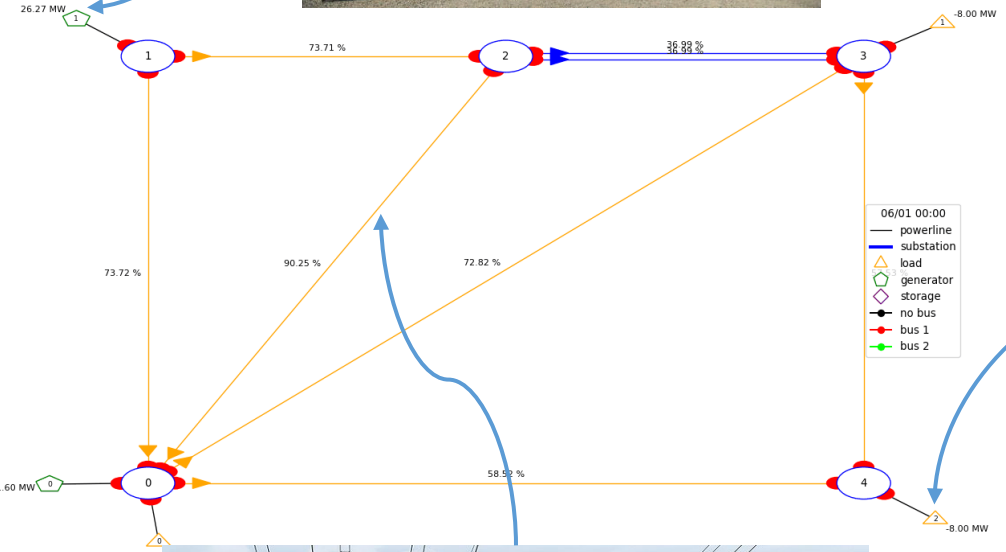
Power Grid on Grid2op Platform

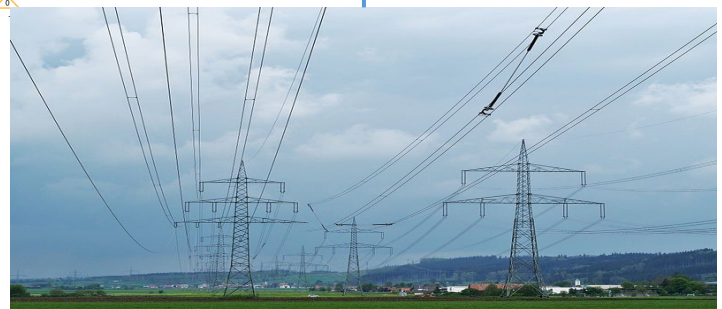


Generator station

Load: a Small Town

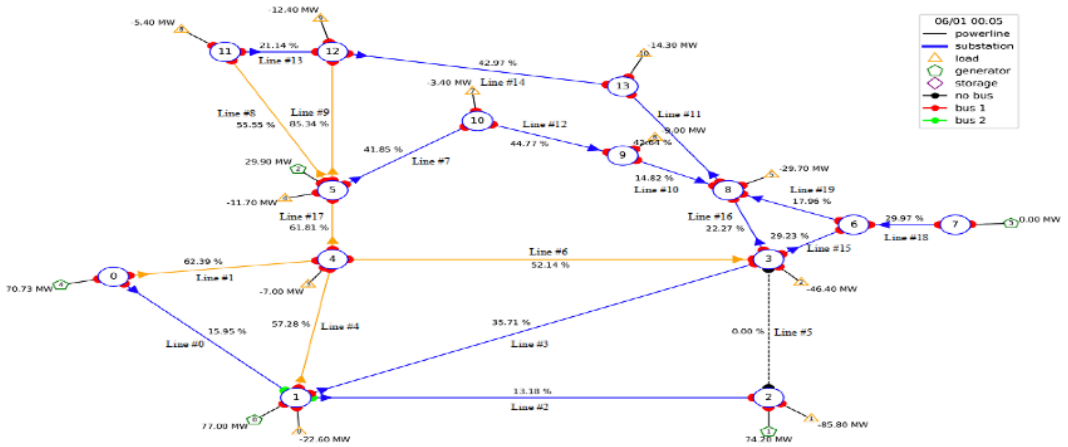
Substation (Bus)

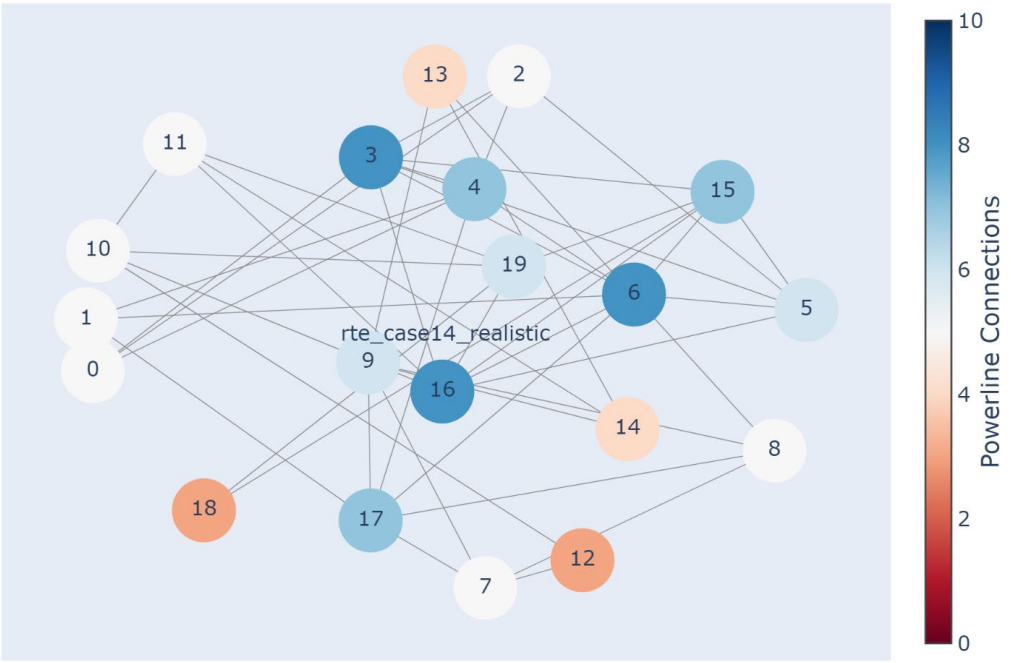


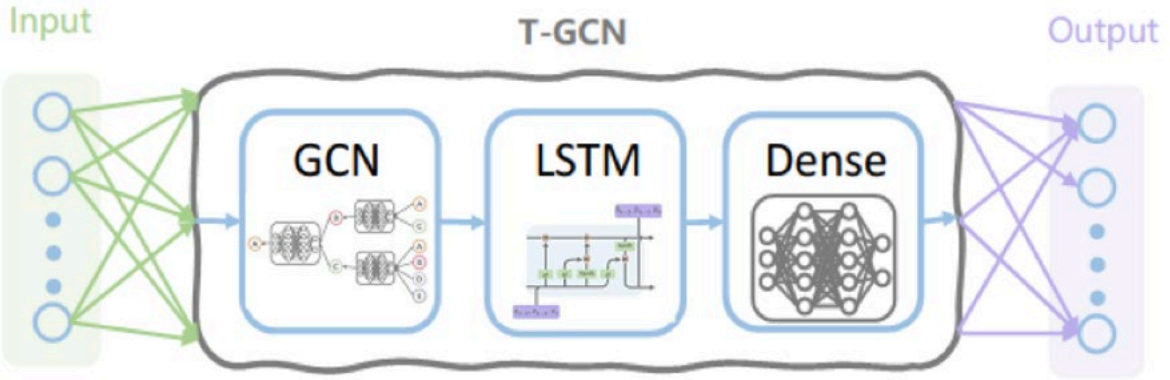


Transmission lines

Line Graph of a Power Grid Network and TGCN

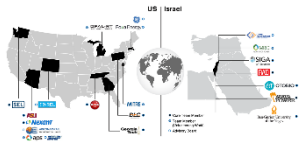




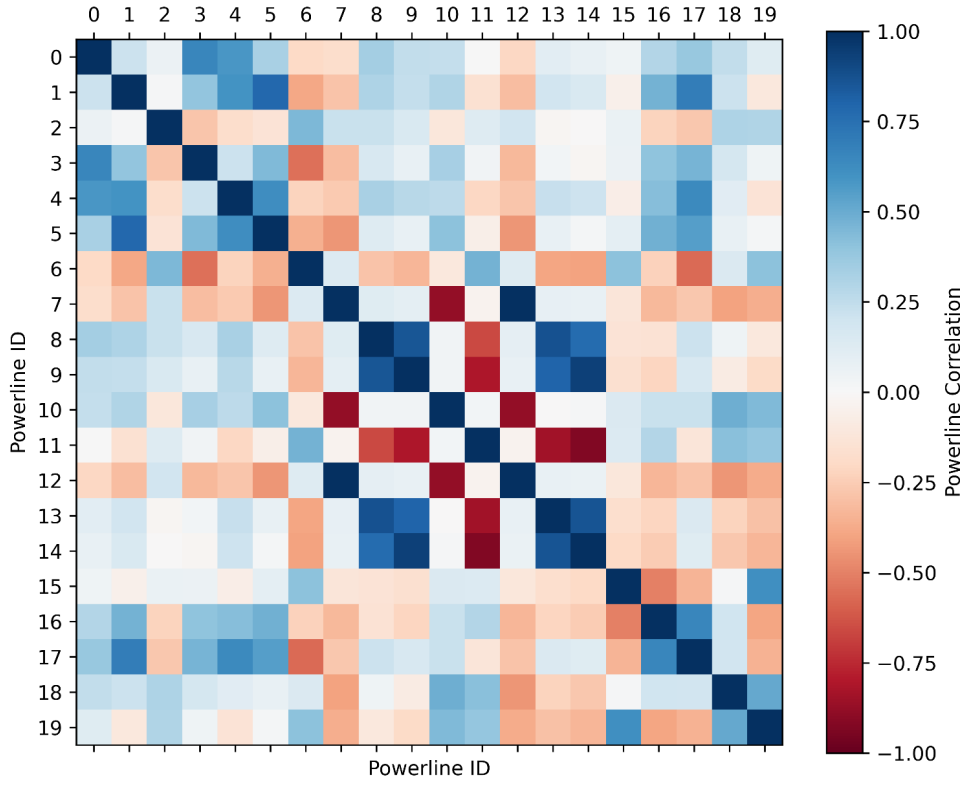


- TGCN: action specific (e.g., five different TGCNs depending on the action types)
- Input: currents from all nodes in the line graph
- Output: currents from all nodes in the line graph
- Training data: Grid2op simulations

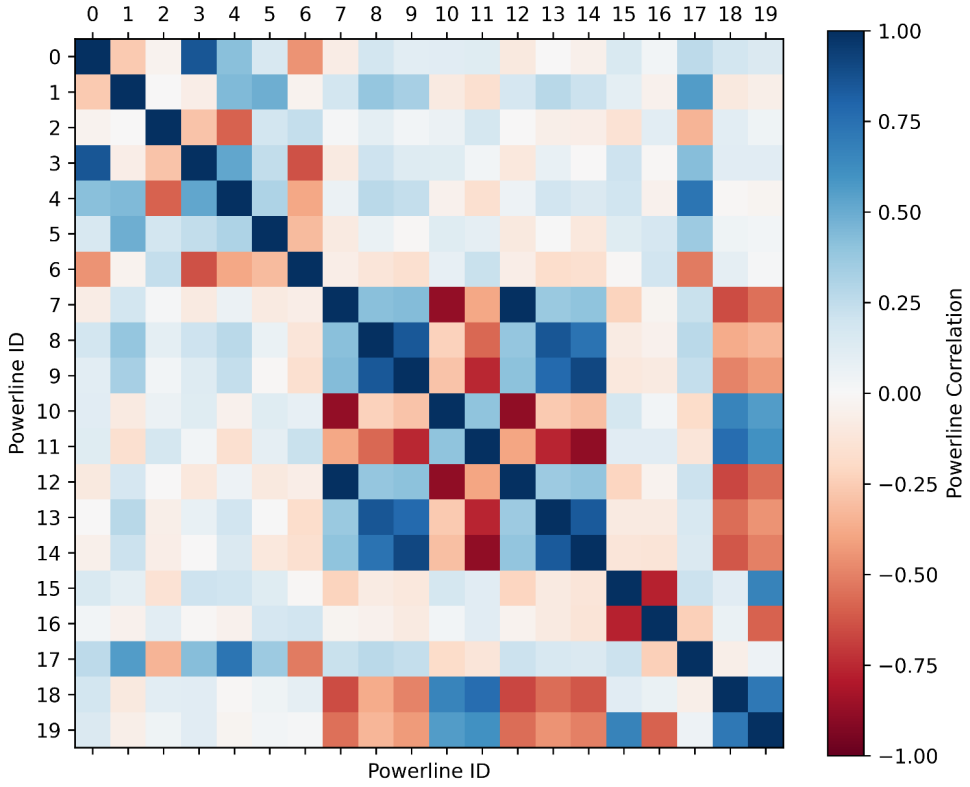
Correlation of Line Current Flow under Attack



PLS Agent

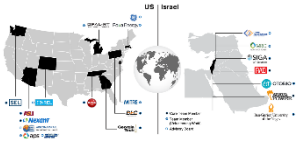


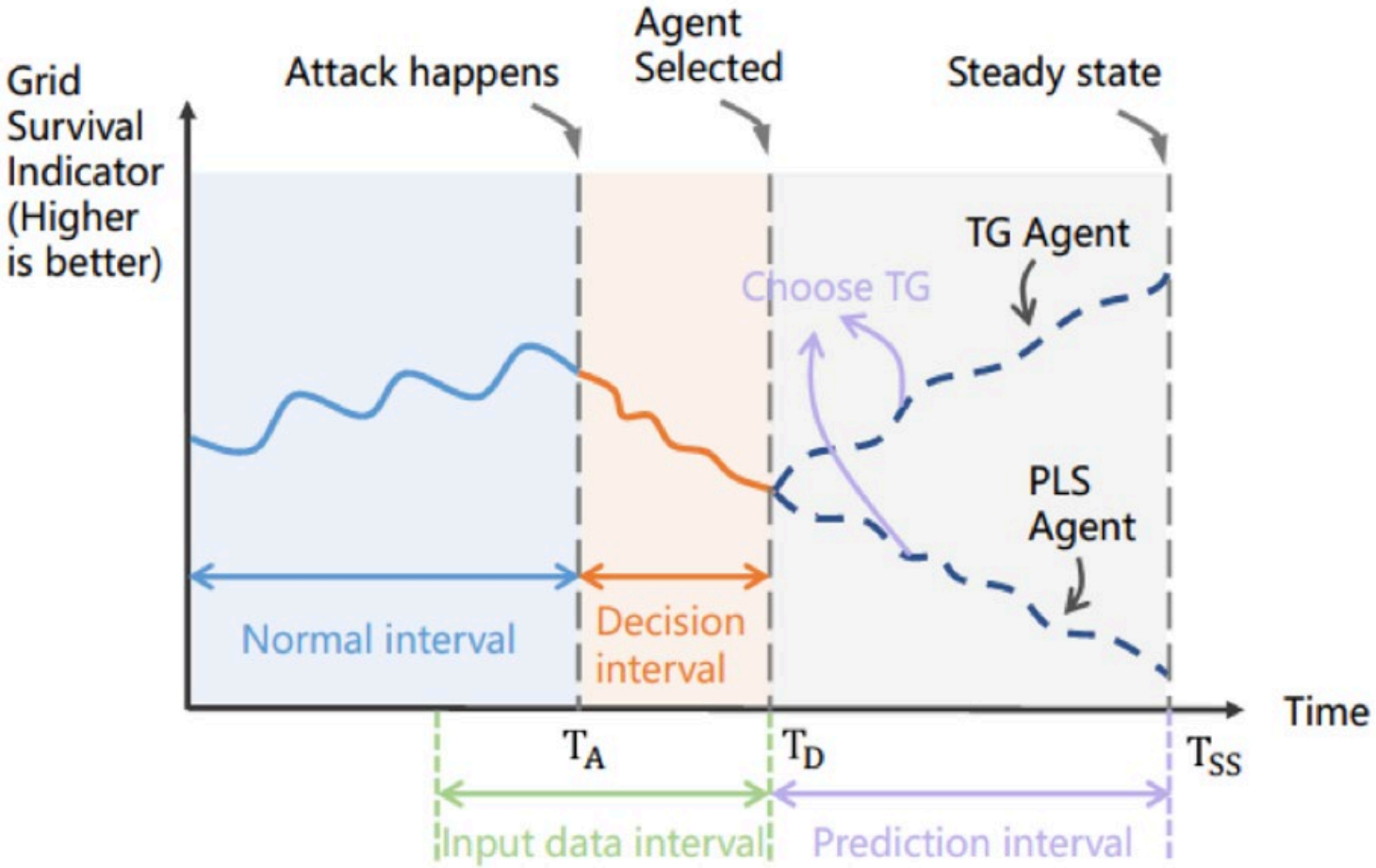
TG Agent



Correlations are neither too small nor too large, justifying TGCN

TGCN Aided RL for Optimal Agent Selection Under Attack



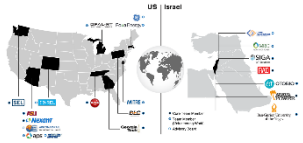


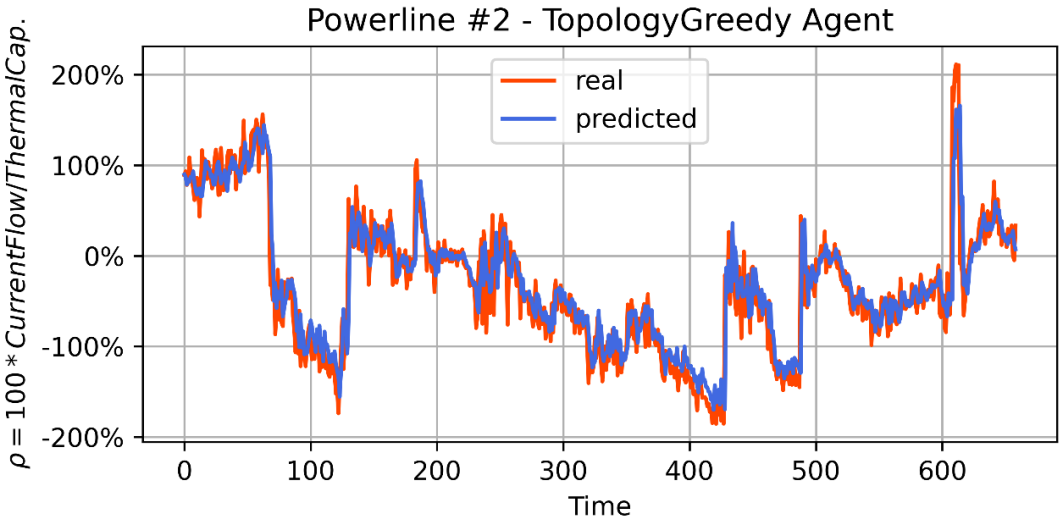
$$\varrho \equiv \frac{Line\ Current}{Line\ Thermal\ Capacity}$$

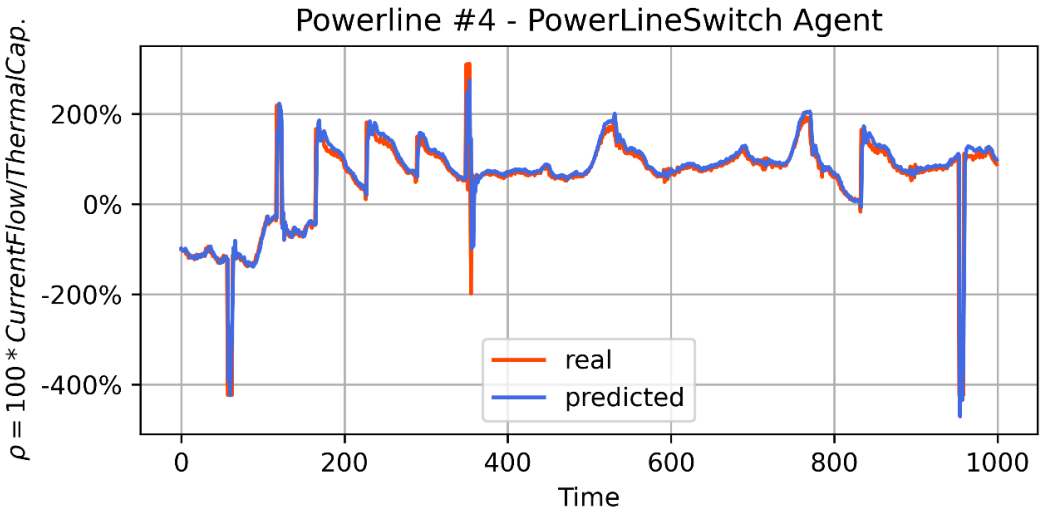
$N_>$ - number of lines in the network with ϱ greater than a threshold

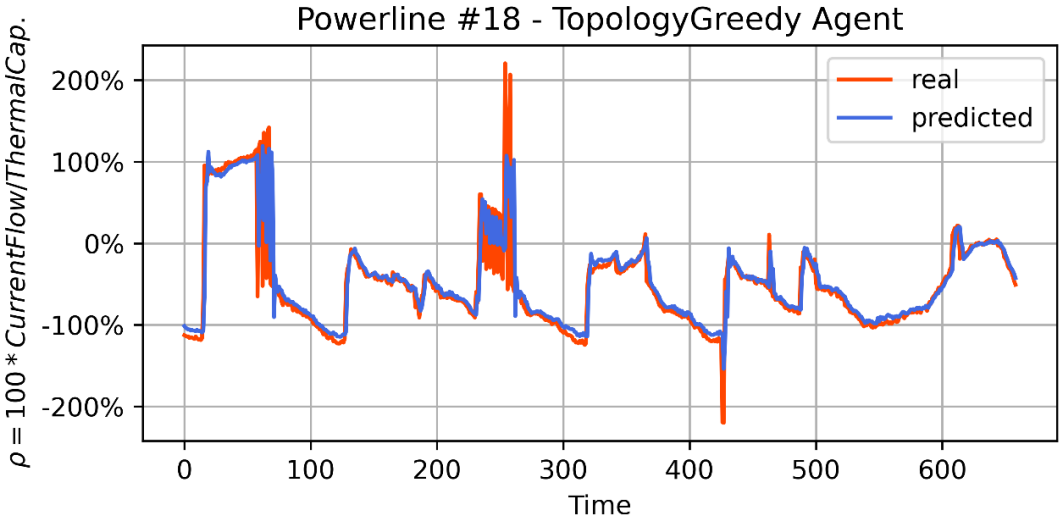
$$\text{Reward} \equiv \frac{1}{1+N_>}$$

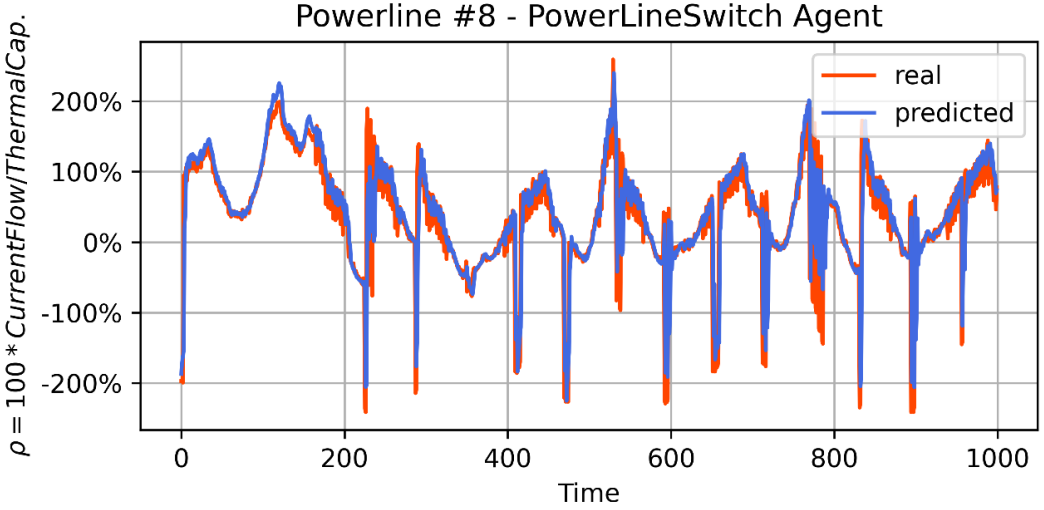
Examples of TGCN Prediction



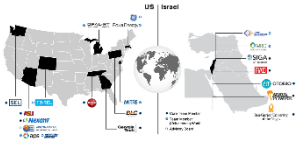


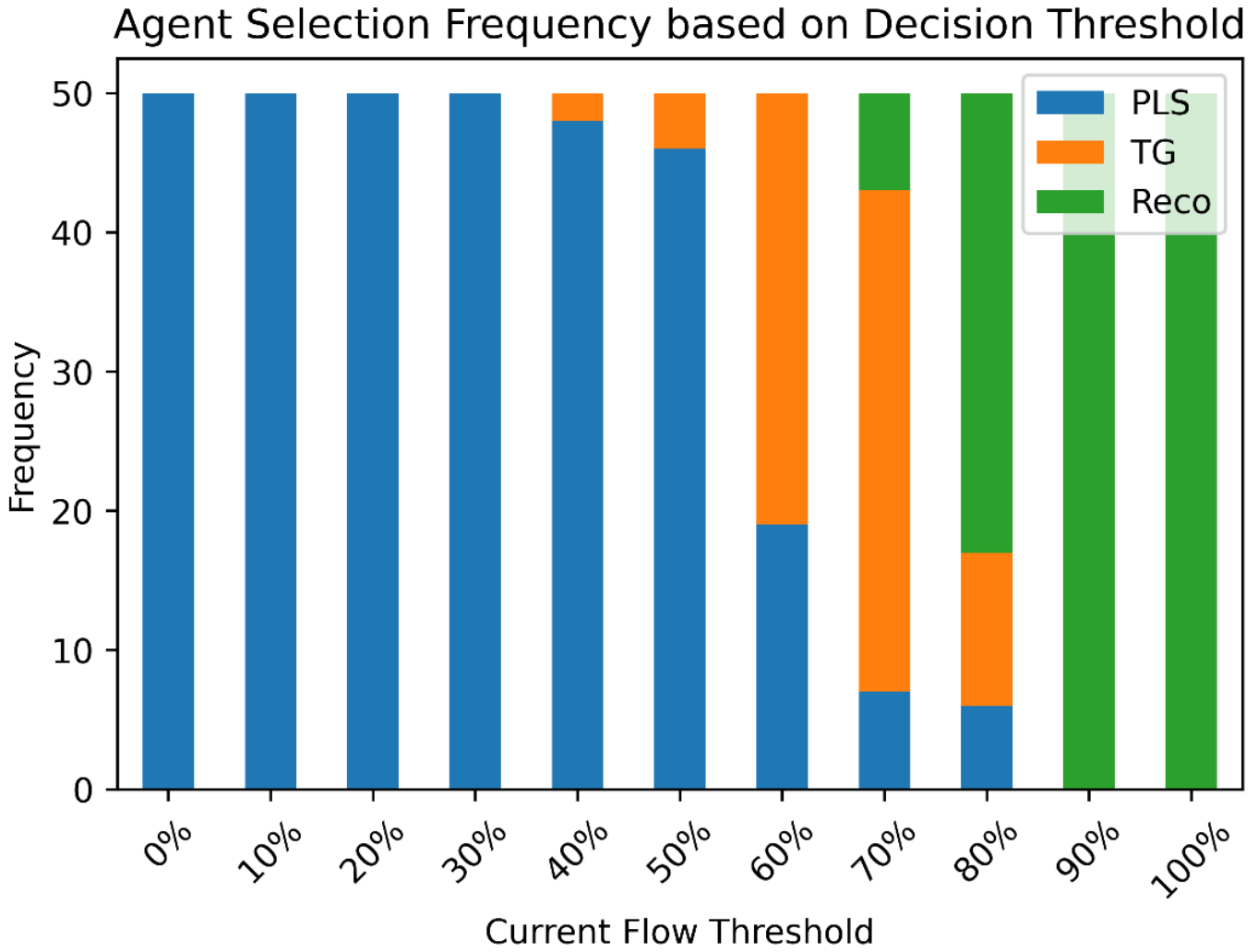




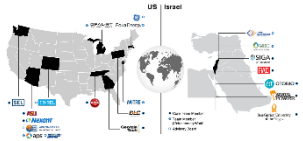


Selection of Controller Based on TGCN Prediction





Industrial Partner and Commercialization



The ASU Task 16 team is working with John Dirkman's team at *Resource Innovations Nexant* to implement the principle and methodologies of reinforcement learning control of cyber physical systems into the existing industrial software tools.

Parallel heterogeneous reinforcement learning for defending large power grids

Mohammadamin Moradi,[1] Shirin Panahi,[1] Zheng-Meng Zhai,[1] Yang Weng,[1] John Dirkman,[2] and Ying-Cheng Lai[1,3,*]

[1] *School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287, USA*

[2] *Resource Innovations, Nexant Inc., 6620 Southpoint Drive South, Jacksonville, FL 32216-8098*

[3] *Department of Physics, Arizona State University, Tempe, Arizona 85287, USA*

(Dated: March 2, 2023)

Reinforcement learning (RL) has been employed to devise the best course of actions in defending the critical infrastructures such as large power networks against cyberattacks. However, as the size of the power grid system grows, the RL action space increases exponentially, making it practically infesible for the RL agent to efficiently explore. The current RL algorithms tailored to power grids are generally not suited when the system size becomes large, in spite of trade-offs. We exploit temporal graph convolutional neural networks (TGCNs) to develop a parallel but heterogeneous RL framework to meet this challenge. In particular, we divide the action space into smaller subspaces, each explored by a RL agent. How to efficiently organize the spatiotemporal action sequences then becomes a great challenge. We invoke TGCN to meet this challenge through accurately predicting the performance of each individual RL agent in the event of an attack. The top performing agent is selected, resulting in the optimal sequence of actions. We demonstrate the effectiveness of our TGCN method to capture both the temporal and spatial dependencies of the graph structured data. The framework is validated by simulations of the RTE 14-Bus system using the Grid2OP platform. Our TGCN framework provides a computationally efficient framework for generating the best course of actions to defend large cyberphysical systems against attacks.

John on 3/1/2023:
"Regarding the manuscript for RL (Task 16) I would be happy to co-author this work and I will also consider the potential for commercialization. I will also evaluate if we could also use Grid360 to further valuate and test the developed code."

3/10/2023: A commercialization meeting with John, who deemed. the TGCN framework interesting and potentially commercializable.