Smart Grids, Cyber Security, and the Big Spillover

What Connects Fridges to Smart Grid Cyber Security and What can we do About it?

Eric MSP Veith <eric.veith@uol.de> , 2026-03-11

Carl von Ossietzky Universität Oldenburg

% whoami

- Eric MSP Veith `<eric.veith@uol.de>`
- Scientific director at OFFIS, Oldenburg, Germany
- Computer scientist by heart: First ICT, then distributed heuristics, then Multi-Agent Systems, now advanced Deep Reinforcement Learning
- PhD in 2017: “Universal Smart Grid Agent for Distributed Power Generation Management.”
- Creator of the Adversarial Resilience Learning methodology (advanced DRL in CNIs)

Carl von Ossietzky
Universität
Oldenburg

IARIA

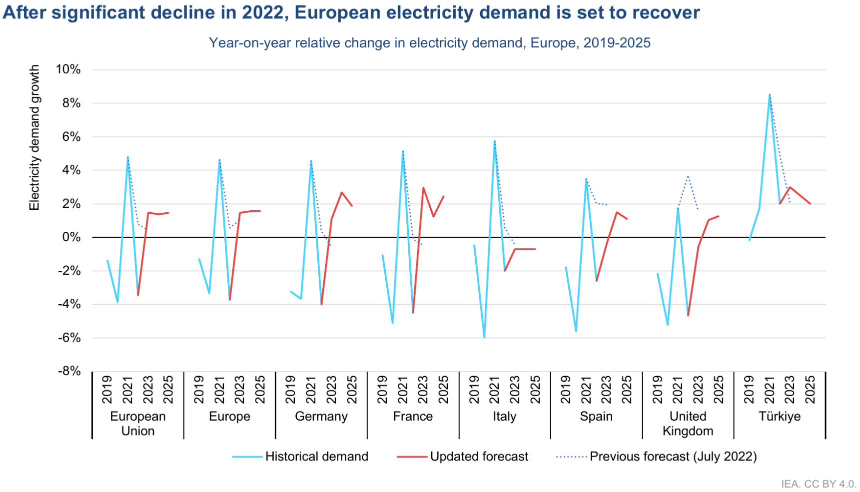
% whereami

A Tale About the Good Ol' Days

Carl von Ossietzky Universität Oldenburg

IARIA

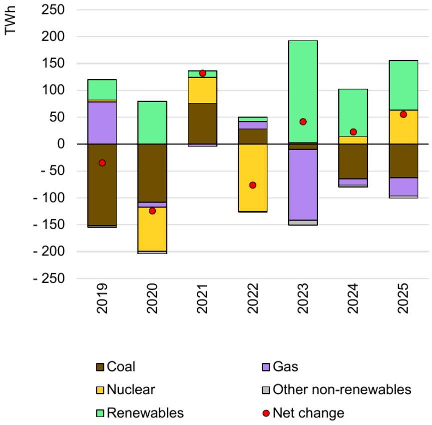
Electricity Demand Rising



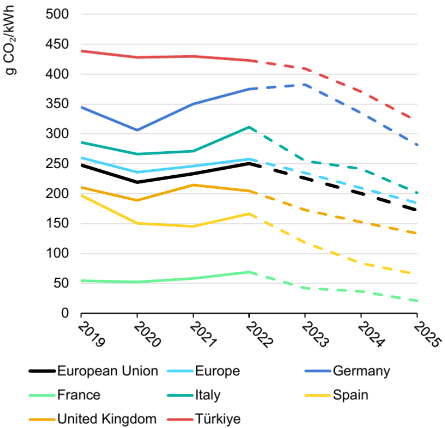
Renewables Are Replacing Fossil Fuels

Following two years of increases, CO_2 intensity starts to decline again from 2023 onward

Year-on-year change in electricity generation, European Union, 2019-2025

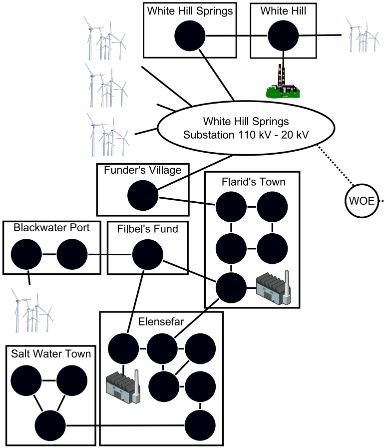


Development of average CO_2 intensity, Europe, 2019-2025



AI as Promise of an Alternative

- Multi-Agent Systems promise local, more more efficient grid operation
- Each node (subgrid, ...) an agent
- Nodes (agents) forecast local power generation/consumption
- On disequilibrium, match forecasts to achieve equilibrium
- An example, based on the literature [6, 3]



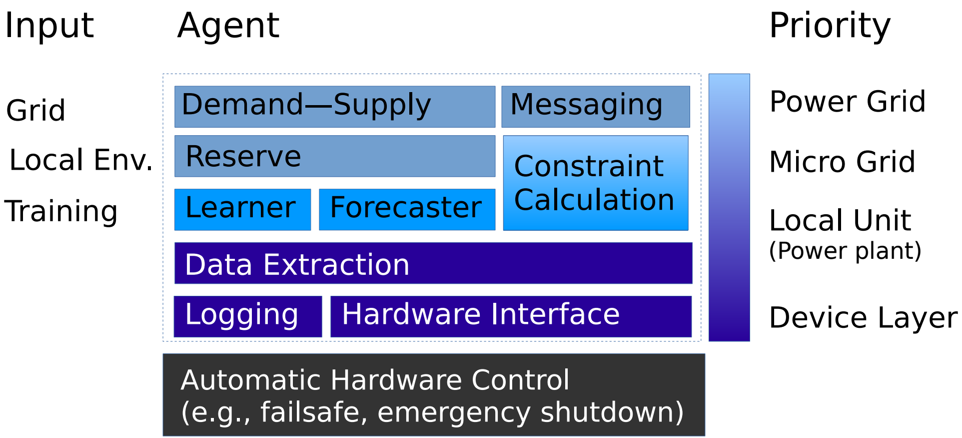
Carl von Ossietzky Universität Oldenburg

Pillars of CPES-MAS

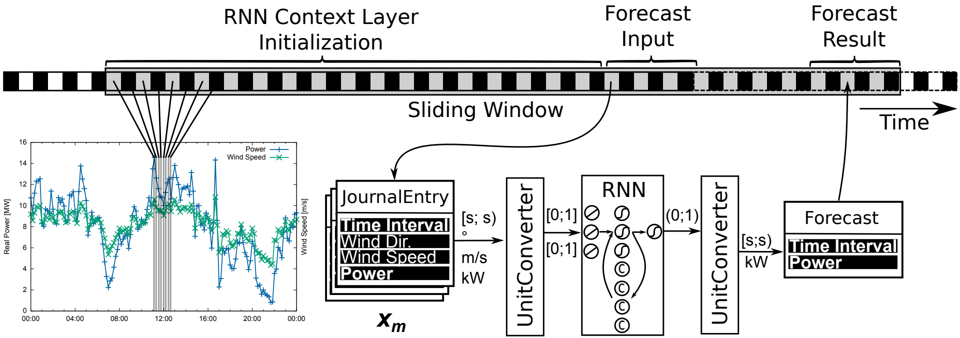
An approach for a Multi-Agent System (MAS) that manages high shares of volatile generators and consumers in an energy system is based on three pillars:

1. Forecasting of local generation or demand
2. Communicating demand and generation (distributed snapshotting with the power grid in mind)
3. Solving the combinatorical problem of demand and supply

Agent Design

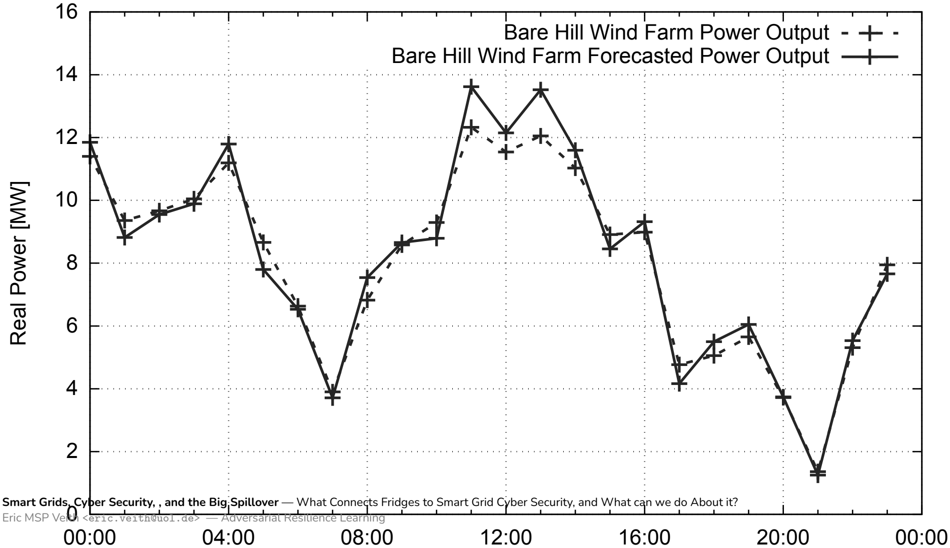


Forecasting



Forecasting is industry state of the art now.

Forecasting



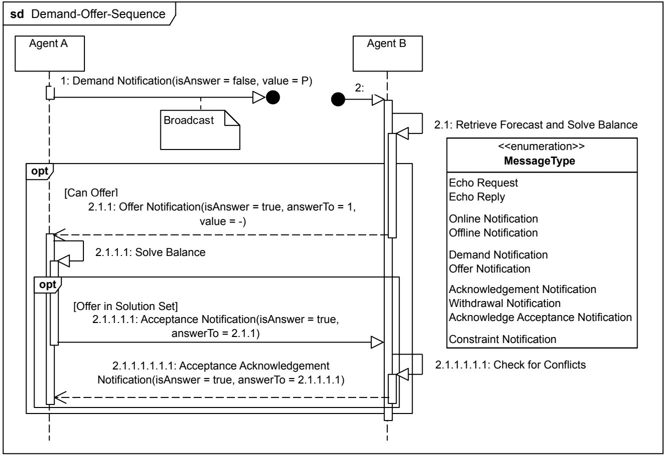
Communication

After forecasting works (i. e., imbalances are known in advance), each agent (= node) must ask its neighbors to help attain the equilibrium.

- Approach: Use overlay network – (virtual) communication lines between agents based on actual grid lines
- Requests (demand for power, or increase in feed-in that should be consumed) travel via selective broadcast
- Each node along the way must try to contribute!

Carl von Ossietzky Universität Oldenburg

Four-Way Handshake



Carl von Ossietzky Universität Oldenburg

LPEP Forwarding

- L_i: *Links* of the i-th agent
- $l_{i,k}$: k-th link of i-th agent
- distance($l_{i,k}$): Distance metric
- m_j: j-th message
- M_i: Message Journal of the i-th agent

$$M_i = \{m_1 \mapsto \{(l_{i,1}, m_{1,\text{distance}(l_{i,1})}), \ldots, (l_{i,n}, m'_{1,\text{distance}(l_{i,n})})\},$$
$$\ldots,$$
$$m_n \mapsto \{(l_{i,1}, m_{n,\text{distance}(l_{i,1})}), \ldots, (l_{i,n}, m'_{n,\text{distance}(l_{i,n})})\}\}$$

$$l_{i,1}(t) \leq l_{i,2}(t) \quad \Leftrightarrow \quad l_{i,1,\text{distance}}(t) \leq l_{i,2,\text{distance}}(t)$$

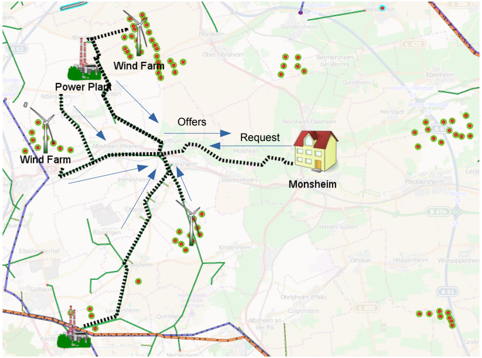
Forwarding

1. Respect *Constraint Notifications*:
   1.1 No answer if $\min(M(m))$ a constraint notification to m, additionally
   1.2 send *Withdrawal Notification* iff already answered
2. $m_{isAnswer}$: forward on best connect ($\min(M(m_{answerTo})))$
3. *Selective Broadcast* for requests:
   3.1 Replace request with *Constraint Notification*, if necessary
   3.2 $M(m) = \emptyset$: forward on $|L| - 1$ links
   3.3 $m' = \min(M(m'))$: Update by fowarding
   3.4 Otherwise: no forwarding

Carl von Ossietzky Universität Oldenburg

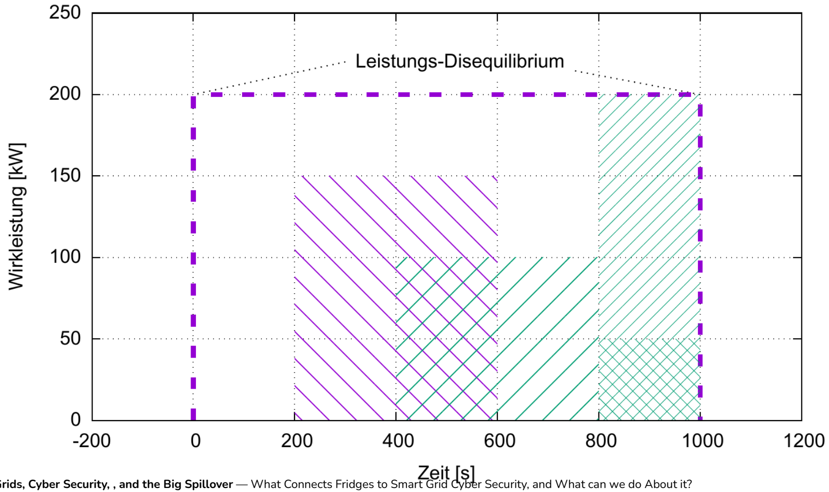
IARIA

How to Decide...?



1. Local forecasting shows demand or oversupply of energy
2. Requests are sent
3. Other nodes make offers
4. Offers reach requestor
5. **Decision about offers?**

Power Balance Concept



Set of Mappings $[t_1; t_2) \mapsto P$.

Carl von Ossietzky Universität Oldenburg

Problem Statement

'Power Balance Algebra':

$$\{[t_1; t_3) \mapsto P_1\} \cup \{[t_2; t_4) \mapsto P_2\} = \{[t_1; t_2) \mapsto P_1, [t_2; t_3) \mapsto P_1 + P_2, [t_3; t_4) \mapsto P_2\}\ , \quad (1)$$

$$[t_1; t_2) \mapsto P_1 \ \subseteq\ [t_3; t_4) \mapsto P_2 \Leftrightarrow \quad t_1 \geq t_3 \ \wedge\ t_2 \leq t_4 \ \wedge\ P_1 \leq P_2\ ; \quad (2)$$

Distance Function:

$$\mathrm{d}(r_i) : r_i \mapsto \mathbb{R} \quad (3)$$

Problem Statement:

$$\sum_i b_i r_i \subseteq r_0\ ,\ \ i \neq 0, b_i \in \{0, 1\}\ , \quad (4)$$

$$\text{Subject to: } \min \sum_i b_i \mathrm{d}(r_i),\ \ i \neq 0, b_i \in \{0, 1\}\ . \quad (5)$$

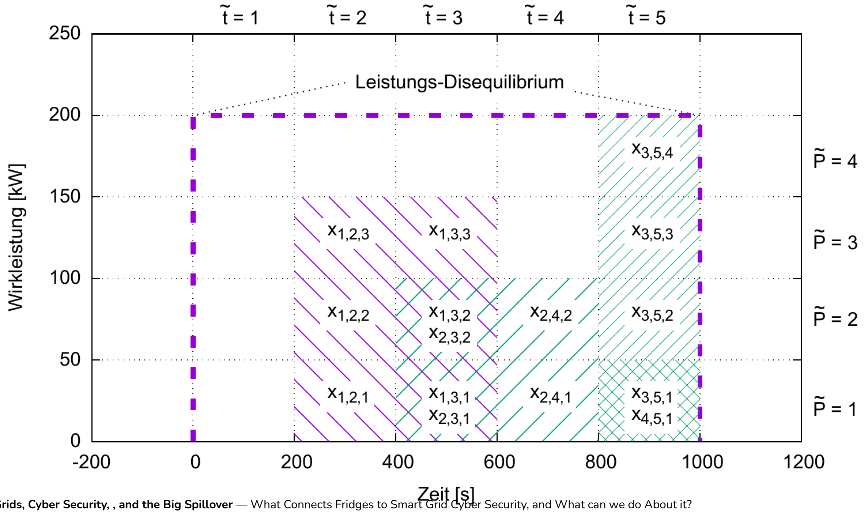
Carl von Ossietzky Universität Oldenburg

Atomization

$$
\begin{aligned}
\boldsymbol{P} &= (|P_0|, |P_1|, \ldots, |P_i|, |P_C|) \;, \\
\boldsymbol{t} &= (t_{2,0} - t_{1,0}, t_{2,1} - t_{1,1}, \ldots, t_{2,i} - t_{1,i}) \;, \\
\Delta P &= \mathrm{ggT}(\boldsymbol{P}) \;, \\
\Delta t &= \mathrm{ggT}(\boldsymbol{t}) \;,
\end{aligned}
$$

$$
x_{i,\tilde{t},\tilde{P}} = \begin{cases} 1 & \text{if agent } i \text{ influences the grid in time-subinterval } \tilde{t} \text{ with} \\ & \text{power from the power-subinterval } \tilde{P}, \\ 0 & \text{else.} \end{cases}
$$

Atomization Illustrated



Model of the Disequilibrium

A symmetric function for each time-subinterval:

$$S_k^n(\boldsymbol{x}_{\boldsymbol{i},\tilde{\boldsymbol{t}}=\boldsymbol{k},\tilde{\boldsymbol{P}}}) = \begin{cases} 1 & \text{if } n \text{ variables in } \boldsymbol{x}_{\boldsymbol{i},\tilde{\boldsymbol{t}}=\boldsymbol{k},\tilde{\boldsymbol{P}}} \text{ equal } 1, \\ 0 & \text{else;} \end{cases}$$

Full Disequilibrium:

$$S = \bigcap_{k=1}^{m} S_k^n(\boldsymbol{x}_{\boldsymbol{i},\tilde{\boldsymbol{t}}=\boldsymbol{k},\tilde{\boldsymbol{P}}})$$

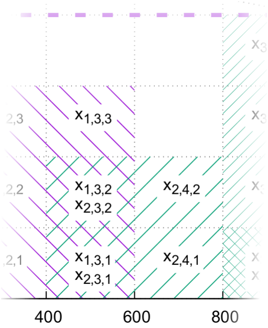
Carl von Ossietzky Universität Oldenburg

Modelling Responses

Acceptance Function:

$$r_i(\boldsymbol{x}_{\boldsymbol{i},\tilde{\boldsymbol{t}},\tilde{\boldsymbol{P}}}) = \begin{cases} 1 & \text{if } \boldsymbol{x}_{\boldsymbol{i},\tilde{\boldsymbol{t}},\tilde{\boldsymbol{P}}} \text{ describes a valid interval for accepting the response of } i, \\ 0 & \text{else.} \end{cases}$$

$$r_2(\boldsymbol{x}_{\boldsymbol{i},\tilde{\boldsymbol{t}},\tilde{\boldsymbol{P}}}) = \bar{x}_{2,3,1} \wedge \bar{x}_{2,3,2} \wedge \bar{x}_{2,4,1} \wedge \bar{x}_{2,4,2} \vee x_{2,3,1} \wedge x_{2,3,2} \wedge \bar{x}_{2,4,1} \wedge \bar{x}_{2,4,2} \vee x_{2,3,1} \wedge x_{2,3,2} \wedge x_{2,4,1} \wedge x_{2,4,2}$$



Equilibrium

$$S = \bigcap_{k=1}^{m} S_k^n(\boldsymbol{x}_{\boldsymbol{i},\tilde{\boldsymbol{t}}=\boldsymbol{k},\tilde{\boldsymbol{P}}})$$

$$R = \bigcap_{i \in I', \tilde{t}, \tilde{P}} r_i(\boldsymbol{x}_{\boldsymbol{i},\tilde{\boldsymbol{t}},\tilde{\boldsymbol{P}}}) \,,$$

$$C = S \cap R \,.$$

Equilibrium

$$S = \bigcap_{k=1}^{m} \mathrm{S}_k^n(\boldsymbol{x}_{\boldsymbol{i},\tilde{\boldsymbol{t}}=\boldsymbol{k},\tilde{\boldsymbol{P}}})$$

$$R = \bigcap_{i \in I', \tilde{t}, \tilde{P}} r_i(\boldsymbol{x}_{\boldsymbol{i},\tilde{\boldsymbol{t}},\tilde{\boldsymbol{P}}}) ,$$

$$C = S \cap R .$$

- Best solution through ordering: $r_i \leq r_{i'} \quad \Leftrightarrow \quad \mathrm{d}(r_i) \leq \mathrm{d}(r_{i'})$
- Generating next vector in S through permutation
- Exploiting the commutative property of the intersection operator: $R_n \cap (\ldots \cap (R_2 \cap (R_1 \cap S)))$

Efficiency

Data Effect $\kappa = \frac{W}{D} \left[\frac{\text{kWh}}{\text{kB}}\right]$

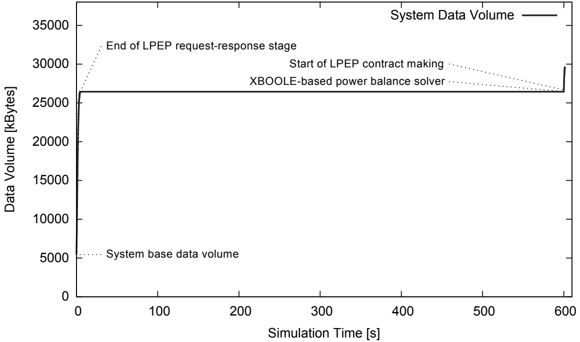
Data Efficiency $\xi = \frac{\Delta P}{D} \left[\frac{\text{kW}}{\text{kB}}\right]$

Comparison

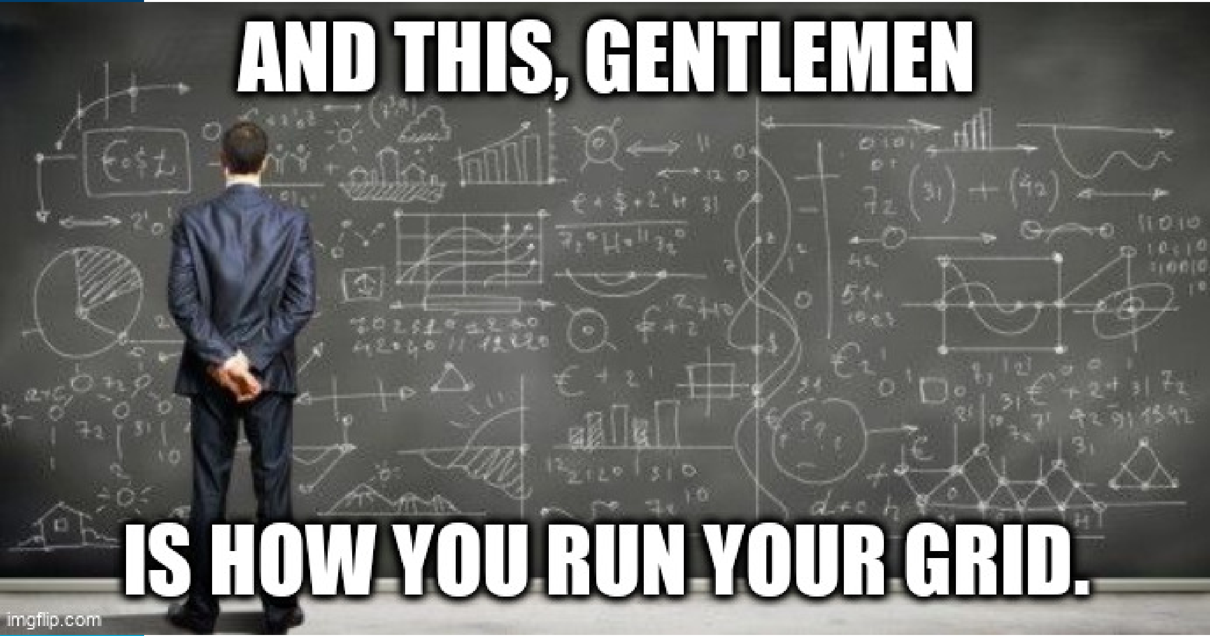
Comparison with BDD approach by Inoue *et al.* (2014):

	BDD	Universal Agent
Loss Avoided (ΔP)	17 208 kW	17 208 kW
Runtime	> 16 min	< 11 min (simulated)
D	100 MB	28.9 MB
ξ	0.168 kW/kB	0.581 kW/kB

Universal Agent Efficiency



- BDD approach in low-load situation: 100 kB
- *Universal Agent* concept especially useful in complex load situations

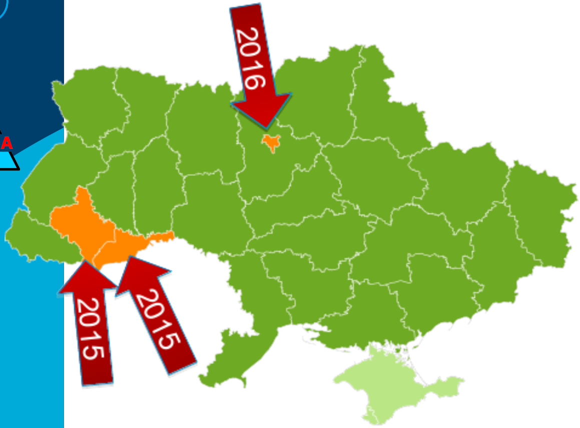

AND THIS, GENTLEMEN
IS HOW YOU RUN YOUR GRID.
imgflip.com

Carl von Ossietzky Universität Oldenburg

Threats No Longer Originate From Within the System


"There are only two types of companies:
those who have been hacked,
and those who don't yet know
they have been hacked."
— John T. Chambers

Energy Systems Fit The Bill Just As Well



Dec 23rd, 2005

- **Cyber attack** causes blackout in the Ukraine
- **3 DSOs** targeted
- **High level of automation** helps attackers
- Operative intrusion in **OT**; disconnection of **several substations**
- Several months in preparation

KA-SAT 2022

- Feb 24th, 2022: Russia invades Ukraine
- At the same time: cyber attack on Viasat
- Attackers exploit a mis-configured VPN appliance at an administration office in Turino
- Further, they gain access to the Viasat management servers and deploy "AcidRain"
- AcidRain deletes critical config data in the flash memory of approx. 40,000 to 45,000 satellite modems

MATT BURGESS SECURITY MAR 23, 2022 7:00 AM

A Mysterious Satellite Hack Has Victims Far Beyond Ukraine

The biggest hack since Russia's war began knocked thousands of people offline. The spillover extends deep into Europe.

PHOTOGRAPH: BJOLZX/GETTY IMAGES

Matt Burges/WIRED, 2022

KA-SAT 2022: The Cascade

A suprising cascade of collateral damage:

- 5,800 Enercon wind turbines in Germany and accross central Europe loose control
- Total installed power: 11 GW
- Loss of over 30,000 satellite terminal accross Europe
- Over 9,000 internet PoA in France down
- Several thousand customers in Poland, Hungary, Greece, and Italy

Nobody knew in advance that an attack on the Ukrainian satellite network would take out wind turbines in Europe. Enercon considers itself collateral damage.

IARIA

Cloud-Controlled PV Inverters

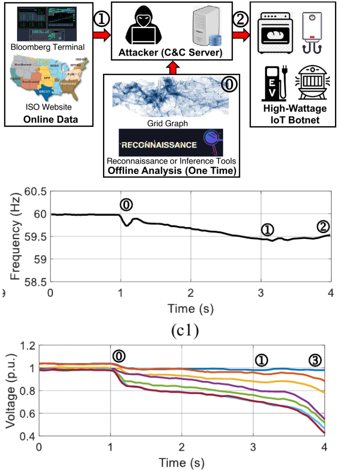
- 168 GW European PV via Chinese cloud platforms
- Projected to exceed 400 GW by 2030
- Vendors: Huawei and Sungrow

Sungrow Advertising

MadIoT: IoT Devices as Botnets

MadIoT [4] uses IoT devices in a concerted fashion to trip power grids.

- High-wattage IoT devices coordinated via botnets
- 90,000 ACs or 18,000 heaters sufficient for blackout
- Invisible to traditional grid monitoring



Common Structural Pattern

- External control plane: foreign or unregulated platforms
- Massive aggregation: individually small, collectively critical
- Regulatory blind spot: outside traditional grid regulations
- Loss of visibility and control for TSOs/DSOs

Carl von Ossietzky Universität Oldenburg

Attacks from the Cyber Space

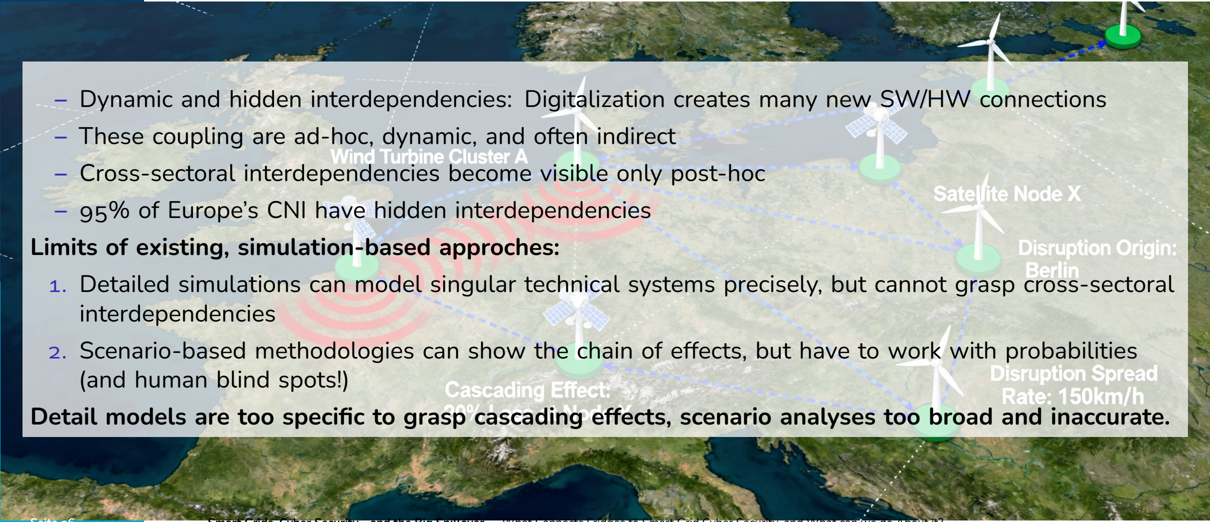
Cyber attacks are a unique threat:

- Global reach: Attacking one location can have global repercussions
- No time limits: A malware can be activated any time
- Low ressources: Much more cost-efficient compared to physical attacks
- Exploiting the supply chain: Legit update mechanisms can be used as attack vectors
- Unforseeable spillover effects: Cascade well beyond sector boundaries

Increased complexity:

- ICT is being deployed in breadth and depth (IoT, battery swarms, fused IIoT sensor-actuator meshes, cloud-controlled inverters, . . .)
- The corresponding infrastructure is all but under grid operators' control!
- Cyber threats go well beyond any CNI operator's ability to prevent or manage
- An increase in fragmentation and decentralization increases the attack surface

Prognosis? Not an Option



- Dynamic and hidden interdependencies: Digitalization creates many new SW/HW connections
- These coupling are ad-hoc, dynamic, and often indirect
- Cross-sectoral interdependencies become visible only post-hoc
- 95% of Europe's CNI have hidden interdependencies

Limits of existing, simulation-based approches:

1. Detailed simulations can model singular technical systems precisely, but cannot grasp cross-sectoral interdependencies
2. Scenario-based methodologies can show the chain of effects, but have to work with probabilities (and human blind spots!)

Detail models are too specific to grasp cascading effects, scenario analyses too broad and inaccurate.

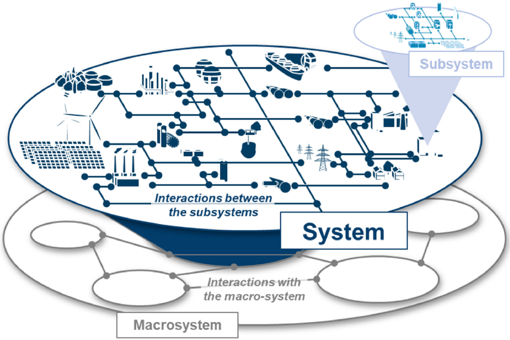
Carl von Ossietzky Universität Oldenburg

Technical Foundations of a Possible Solution

Carl von Ossietzky Universität Oldenburg

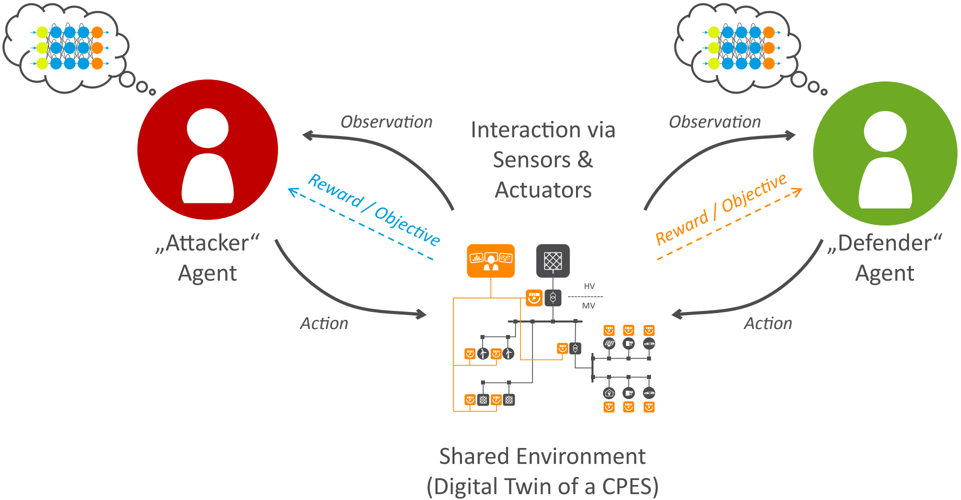
Learning Resilient Control

- **Interconnected CPS have always attack surface due to their inherent complexity**
- Low latency of ICT and OT
- High interdependence
- Complexity in breadth and depth
- Cricital Services as SPOF (DNS, BGP, SCADA, SDL)
- **Learning Stratgies for automatic issue mangement**
- “Adversarial Resilience Learning”

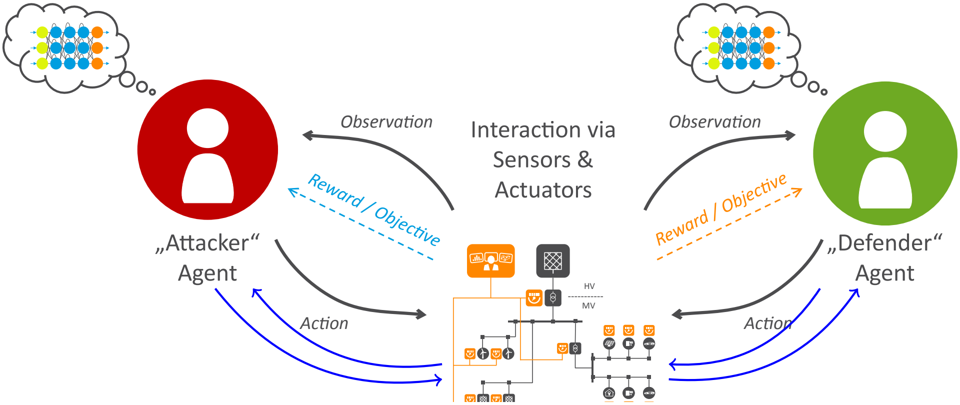


Kotzur, Leander, et al. “A modeler’s guide to handle complexity in energy systems optimization.” Advances in Applied Energy 4 (2021): 100063.

Adversarial Resilience Learning

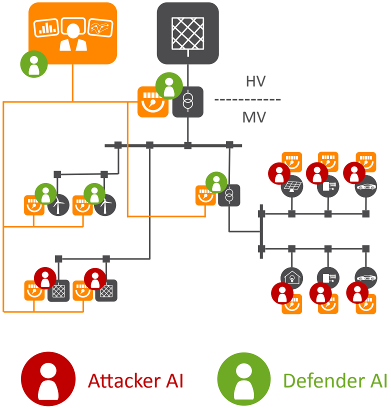


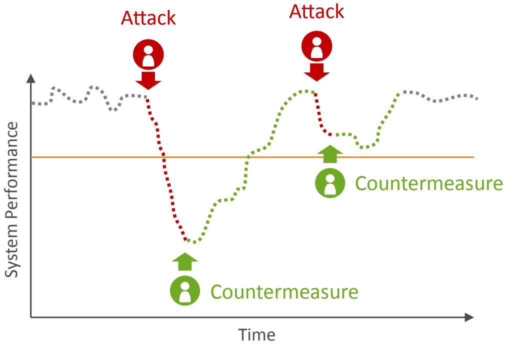
Adversarial Resilience Learning

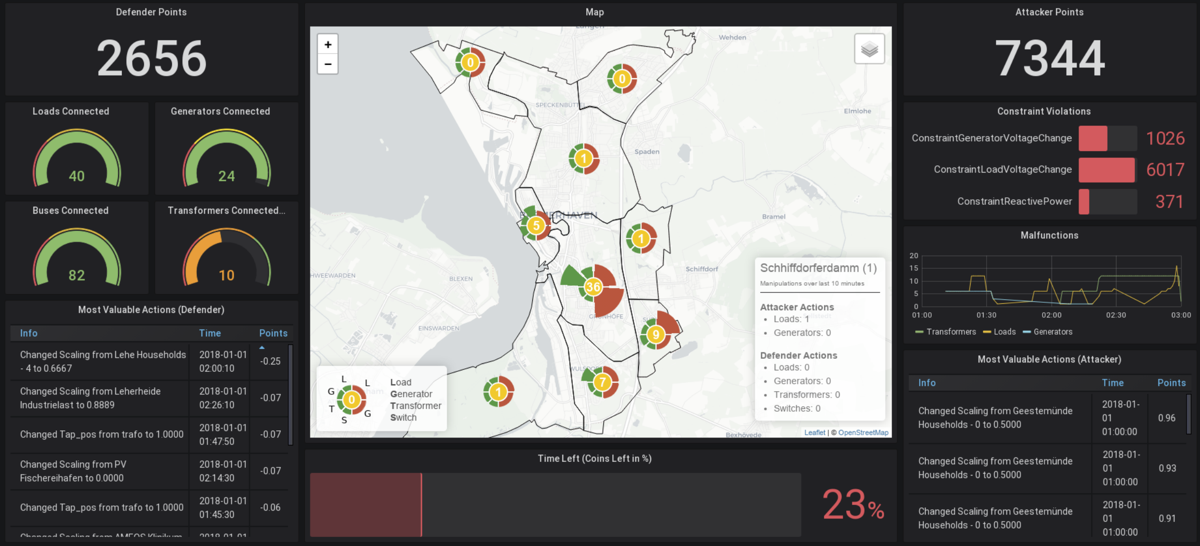


Adversarial
System-of-Systems Reinforcement Learning

ARL Agent Interaction






Defender Points
2656
Map
Attacker Points
7344
Loads Connected
40
Generators Connected
24
Buses Connected
82
Transformers Connected...
10
Constraint Violations
ConstraintGeneratorVoltageChange 1026
ConstraintLoadVoltageChange 6017
ConstraintReactivePower 371
Malfunctions
20 15 10 5 0
01:00 01:30 02:00 02:30 03:00
Transformers Loads Generators
Wehden
SPECKENBÜTTEL
Elmlohe
Spaden
Bramel
Schiffdorf
BLEXEN
EINSWARDEN
GRÜNHÖFE
Bexhövede
Schhiffdorferdamm (1)
Manipulations over last 10 minutes
Attacker Actions
Loads: 1
Generators: 0
Defender Actions
Loads: 0
Generators: 0
Transformers: 0
Switches: 0
Load
Generator
Transformer
Switch
Leaflet | © OpenStreetMap
Most Valuable Actions (Defender)
Info Time Points
Changed Scaling from Lehe Households - 4 to 0.6667 2018-01-01 02:00:10 -0.25
Changed Scaling from Leherheide Industrielast to 0.8889 2018-01-01 02:26:10 -0.07
Changed Tap_pos from trafo to 1.0000 2018-01-01 01:47:50 -0.07
Changed Scaling from PV Fischereihafen to 0.0000 2018-01-01 02:14:30 -0.07
Changed Tap_pos from trafo to 1.0000 2018-01-01 01:45:30 -0.06
Time Left (Coins Left in %)
23%
Most Valuable Actions (Attacker)
Info Time Points
Changed Scaling from Geestemünde Households - 0 to 0.5000 2018-01-01 01:00:00 0.96
Changed Scaling from Geestemünde Households - 0 to 0.5000 2018-01-01 01:00:00 0.93
Changed Scaling from Geestemünde Households - 0 to 0.5000 2018-01-01 01:00:00 0.91

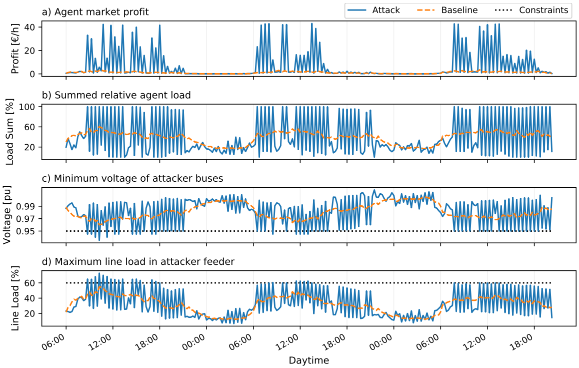
ARL Agent Can Discover Attacks

- Attack on voltage level
- Attacker controls Q feed-in
- Known attack: Oscillating behavior
- ARL agent indepently disovers attack, but also finds variant

Veith, Wellßow, and Uslar [5]

Transactive Energy Can Be Gamed

- Economic and control techniques, based on market standard values
- There is no “sound” market design yet than cannot be gamed
- Worse yet: Agents can find weaknesses & gain market dominance without system knowledge



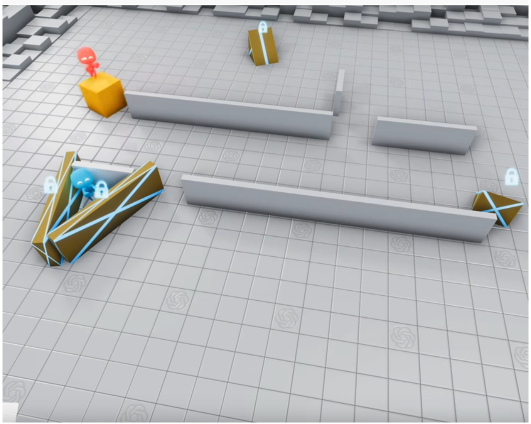
Agents learn to “game” local energy markets

Wolgast, Veith, and Nieße [8]

Carl von Ossietzky Universität Oldenburg

Multi-Agent Autocurricula

- ARL is an autocurriculum setup
- Independently known & verified to work
- Example Setup: Two groups of agents play hide and seek
- No domain information; agents learn strategies and tool use independently
- Result: Agents learn to exploit bugs in the underlying game engine
  - Holes in walls
  - Sliding boxes
  - Edge/corner jumps



Carl von Ossietzky Universität Oldenburg

Autocurricula Helpful in Theory

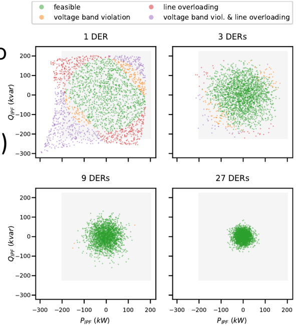
- DRL agents collect initial samples from random actions
- However, random actions over correlated actuators lead to convolution problem, i. e., if $X,\ Y \sim \mathcal{U}$, then

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\, dx\ , \qquad (6)$$

  which is a triangle distribution
- Equally, consider SAC's entropy maximization,

$$\pi^* = \arg\max_{\pi} \mathop{\mathbb{E}}_{\tau\sim\pi}\left[\sum_{t=0}^{\infty} \gamma^t \left(R(s_t, a_t, s_{t+1}) + \alpha \underbrace{H\left(\pi(\cdot|s_t)\right)}_{\text{Entropy term}} \right)\right] \qquad (7)$$

- … obviously, a “push” is required



Autocurricula Helpful in Theory II

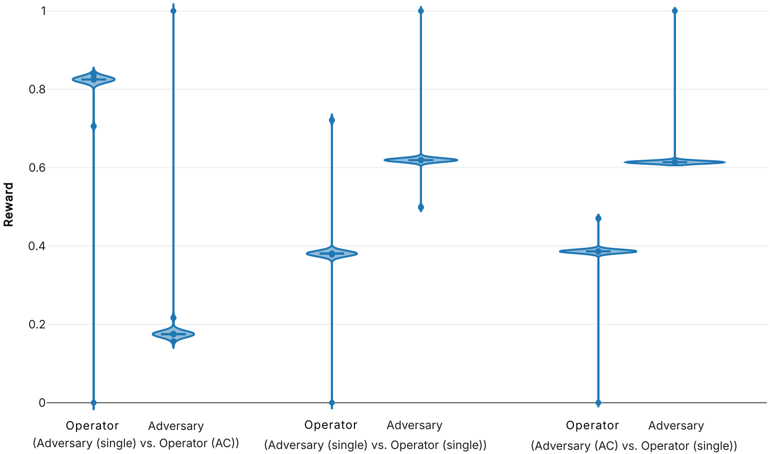
1. Formally, DRL approximates the unknown environment distribution p with q, i. e.,

$$\begin{array}{ll} \text{minimize} & KL(p, q) \\ \text{subject to} & q \end{array} \tag{8}$$

2. Learn a policy to exploit q, π_Ω
3. (Single agent: get stuck in local optimum because p is mostly unknown because of missing sample data)
4. Adversary agent: Observe p as influenced by π_Ω
5. $R_A(\boldsymbol{s}_t \sim p) = -R_\Omega(\boldsymbol{s}_t \sim p)$, therefore $\pi_A \widehat{=} -\pi_\Omega$
6. Result: agents observe adversarial sampels from the “other end” of p’s spectrum
7. Agents try to counter adverse effects: efficent state/action space exploration

Carl von Ossietzky Universität Oldenburg

… and in Practice



Carl von Ossietzky Universität Oldenburg

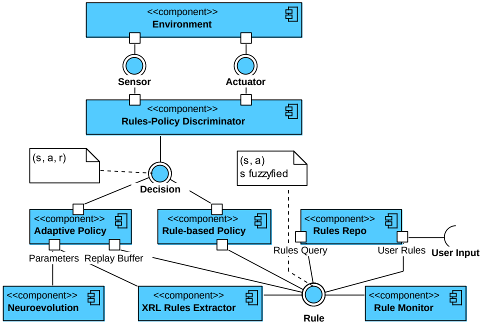
ARL Works

To summarize…

- ARL works for finding attack vectors (“easy”)
- ARL defender learn resilient control (“not quite so easy, but still…”)
- ARL agents learn faster & more robust strategies through the autocurriculum setup (“proove me, I’m only circumstantial evidence!”)
- ARL defender agents can control modern power grids (“ha-ha, as if that would be acceptable…”)
- **There is still a lot missing:**
  - Behavior guarantees
  - Adhere to constraints (rulesets)
  - Learn from existing domain knowledge
  - Adapt during production use (not just retraining)
  - …

ARL Agent Architecture

- Learn from sensor inputs (policy: DRL)
- Deploy & forget, don't design policy networks: Neuroevolution
- Explainability
- Learn from domain knowledge
- Follow rules, if given



Carl von Ossietzky Universität Oldenburg

Learning from Domain Knowledge

Example: Misuse Cases

Stage 1 — Expert Knowledge
Unstructured, not archived domain knowledge

Stage 2 — (Mis)Use Case Template
Structured, arciheved domain knowledge; not machine readable

Stage 3 — UML Diagram
Structured, archieved domain knowledge, limited to the depicted information

Stage 4 — XML file
Structured, archieved domain knowledge; limited to the information of the file this is exportet from; machine readable

Stage 5 — Experiment file
Executable experiment file for the ARL toolchain. Contains the information given in the XML file

Trajectories from (Mis-) Use Cases

- Annotate UML diagrams to allow sampling; construct:
  - Experiment file
  - State machine from transitions

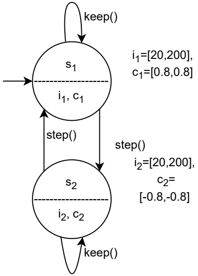
$$\begin{aligned} M_{tg} &= (Q, \Sigma, \delta, q0, F) \\ &\text{with} \\ &(q, (\{c_q\} \in \mathit{ActuatorSetpoints}, \\ &\{i_q\} \in \mathit{TimeStepIntervals})) \in Q \\ &(i \in Q, n \in Q, \{sc\} \in \mathit{StepConstraints}) \in \delta \end{aligned} \tag{9}$$

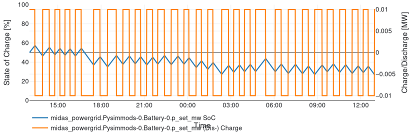
Relevant properties:

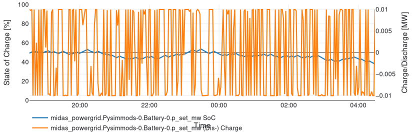
- Non-determinism
- State/actuator constraints c_q (think Gymnasium spaces)
- Time step intervals (sync to simulation semantics)
- Constrained steps (e. g., grid codes)

Carl von Ossietzky Universität Oldenburg

How to Put This to Use



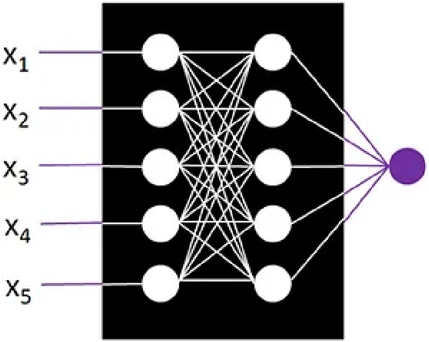




- Agent learns a specific strategy
- Some strategies are hard to discover from reward alone (cf. oscillating behavior)
- Replicates AND advances the strategy

Wellßow, Logemann, and Veith [7]

“The Black Box”



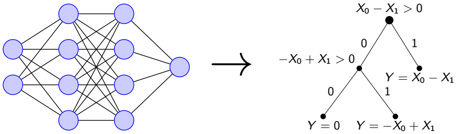
Explanation goals

- Motivation: No trust without explanation of learned strategies of agents
- Idea: Use Decision Trees (DTs) with extraction of rulesets for explanation
  - DTs are transparent and somewhat interpretable
  - They can be trained directly (no need for black-box Deep Neural Network (DNN) models)
  - But DNNs are better regularized, which increases trainability [2]
- Conflicting goals:
  - Construction of powerful (Deep Reinforcment Learning (RL) (RL)) learning system
  - (Post-hoc) Explainability with comprehensible model (e. g. DTs)

Carl von Ossietzky Universität Oldenburg

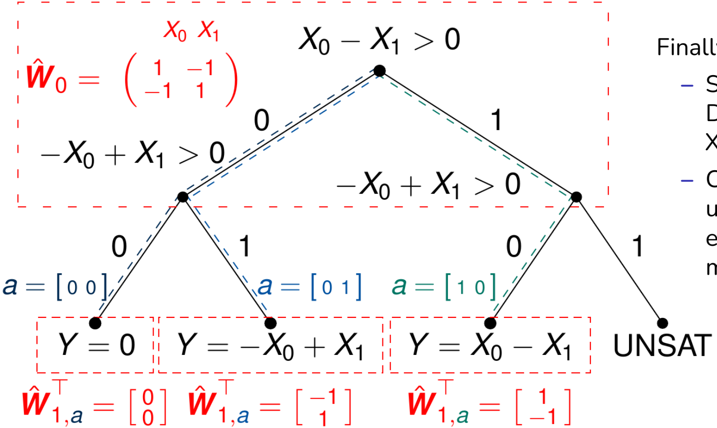
Learned Policy Explanation

- Equivalent transformation of efficient-learnable Feed-Forward DNNs (DNNs) into compressed DTs



- NN2EQCDT algorithm heavily relies on equivalence description of DNNs and DTs [1], but still addressed research gaps to better use it for explainability:
  - Transformation algorithm and actual implementation proposed for PyTorch models
  - Exponential growth is addressed by lossless pruning
  - Dynamic compression reduces computation time significantly and may reduce inference time
  - Option to directly include global constraints for further pruning

XOR model: DT Construction



Finally:

- Simple example of an DT representing an XOR function
- Construction of DT using calculated effective weight matrices

Carl von Ossietzky Universität Oldenburg

XOR model: DT Pruning

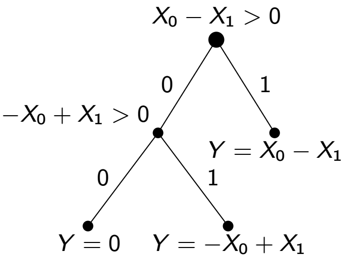


Figure: Simple pruning example

Pruning UNSAT node by

- remove parent and
- connecting sibling subgraph to parent of parent

Comparison of construction methods

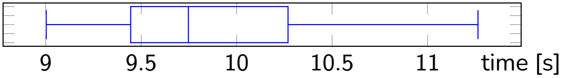


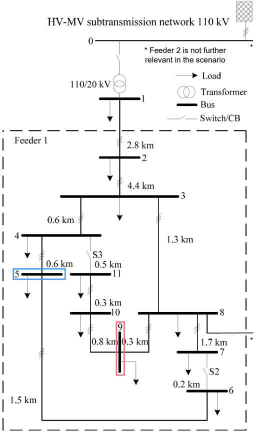
Figure: Boxplot ($n = 30$) for the computation time of the NN2EQCDT algorithm for the simple model

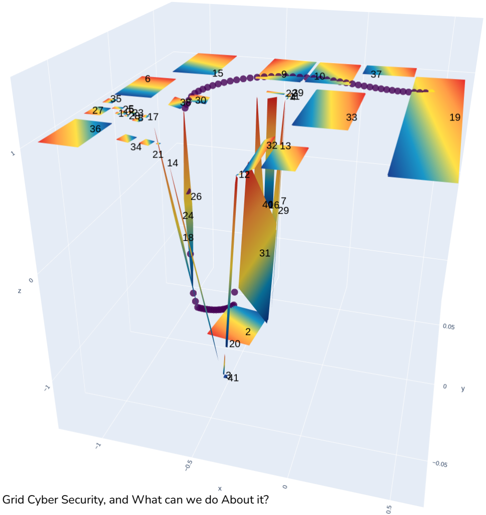
Table: Comparison of results or calculations for the construction of a DT from the simple model without and with compression of the NN2EQCDT algorithm

Pruning	#nodes	Computation time
☐	262143	$>$ 1.5h
☑	83	9.75s

– Pruning ratio (amount of nodes) of 99.97%

Applications in Practice





Carl von Ossietzky Universität Oldenburg

IARIA

Carl von Ossietzky Universität Oldenburg

Going Beyond Deep Reinforcement Learning

Deep Reinforcement Learning is not the Only Answer

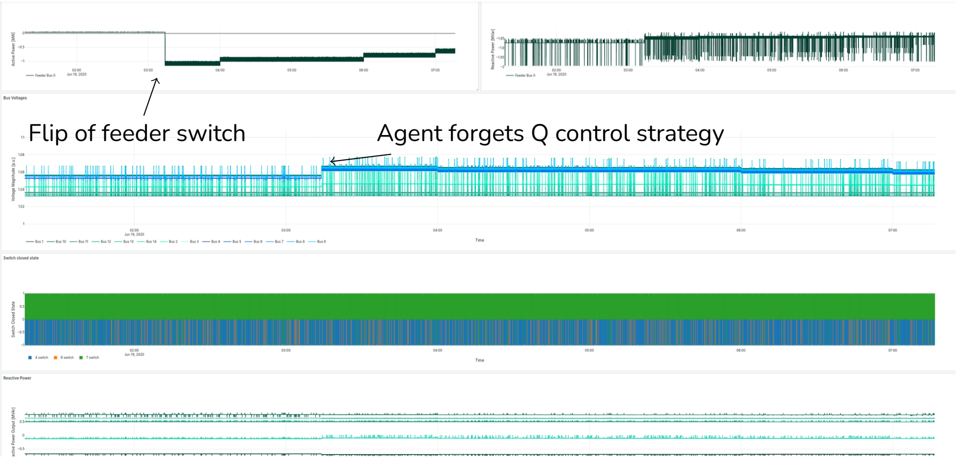
The state of the art has many nice features:

- Offline learning (learning from domain knowledge)
- Imitation learning (learn existing control strategies by example)
- Model-based and model-free DRL
- eXplainable Reinforcement Learning to explain each action with low computational overhead

… however, this agent is still far from being safe.

Carl von Ossietzky Universität Oldenburg

“Good” Agents Fail to Apply Learned Strategies



Carl von Ossietzky Universität Oldenburg

Catastrophic Forgetting on Topology Changes

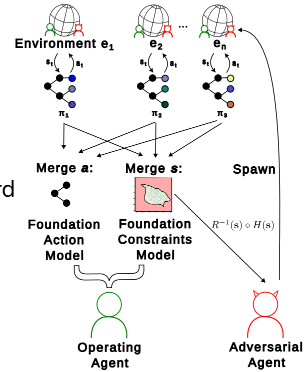
A simple topology change screws the agent completely. Countermeasures:

1. Train the agent on as many scenarios as possible.
2. Verify the DRL agent.
3. Create a *Foundation Model* for actions
4. … Combine all of the above!

Carl von Ossietzky Universität Oldenburg

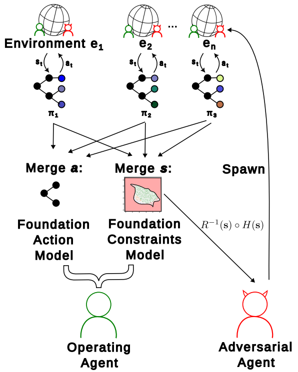
Training Strategy

- Take the autocurriculum approach to spawning environments
- Two adversaries: One “spawner” and *n* “workers”
- Operator agent trains on all of them (traditional multi-worker)
- Adversary spawns environments based on inverted reward and entropy $R^{-1}(s) \circ H(s)$



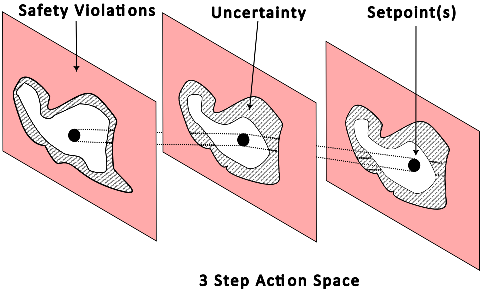
Latent Spaces

- Core idea: Use Graph Neural Networks to learn representations of underlying grids
- Graph space is our feature space: $\mathcal{X} = G(\boldsymbol{A}, \boldsymbol{K})$
- Train encoder for latent space representation of all G_i, where i is an environment instance we encountered: $\gamma : \mathcal{X} \mapsto \mathcal{L}$
- Use transformer to work directly on latent space
- Result: A foundation model for actions



Carl von Ossietzky Universität Oldenburg

And Verification...?



- Alongside the Foundation Action Model, train a foundation model for contraints: *Foundation Constraints Model*
- Use the Foundation Constraints Model for N-step verification of trajectories to provide safety guarantees

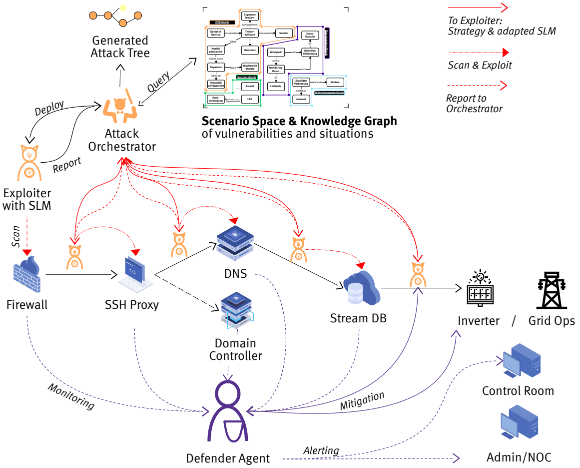
Carl von Ossietzky Universität Oldenburg

How to Discover Spillover Effects with DRL/ARL

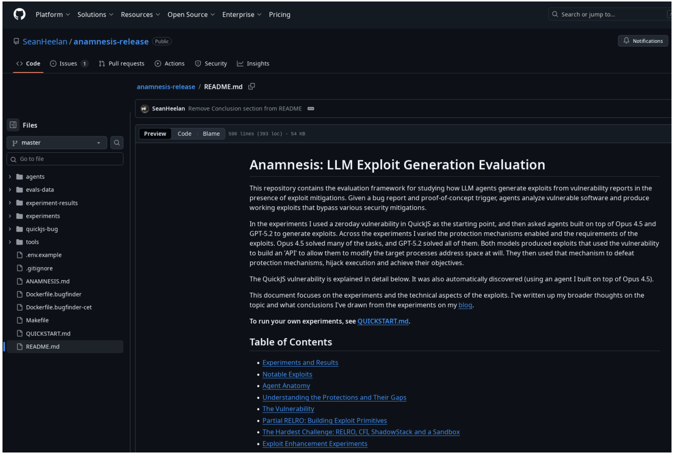
Carl von Ossietzky Universität Oldenburg

IARIA

LACRIMA: Discovering Spillovers using "AI Worms"

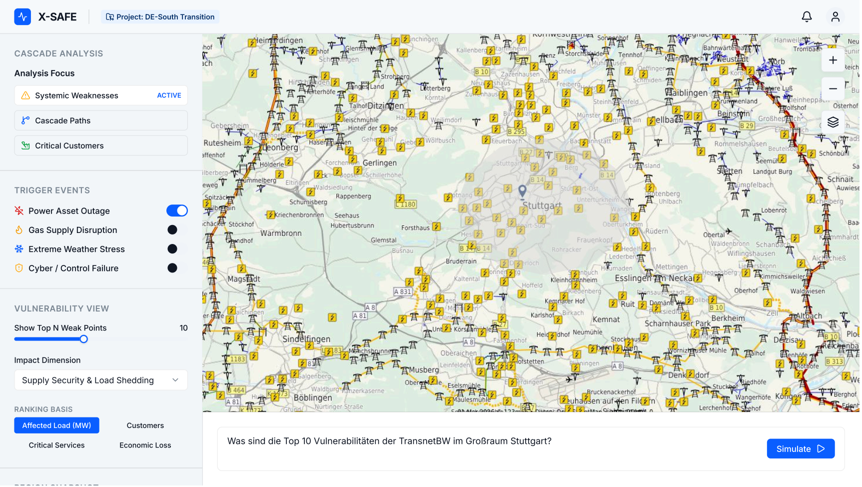


“AI Worms” — Is That a Thing?!

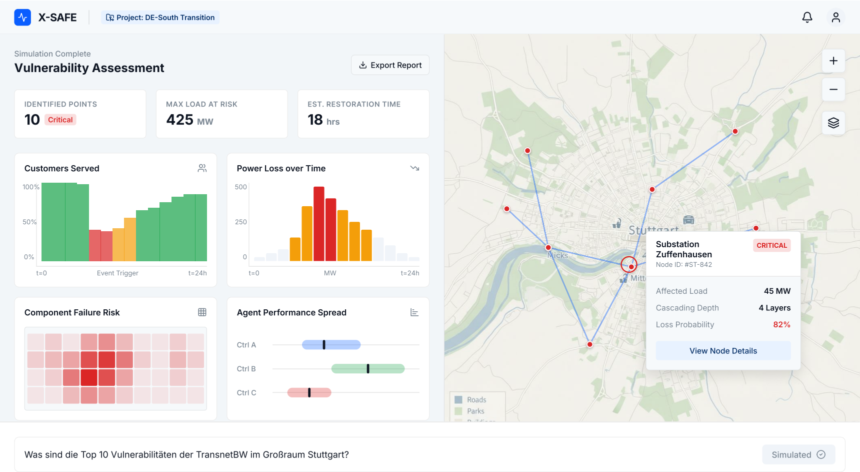


https://github.com/SeanHeelan/anamnesis-release

Input UI

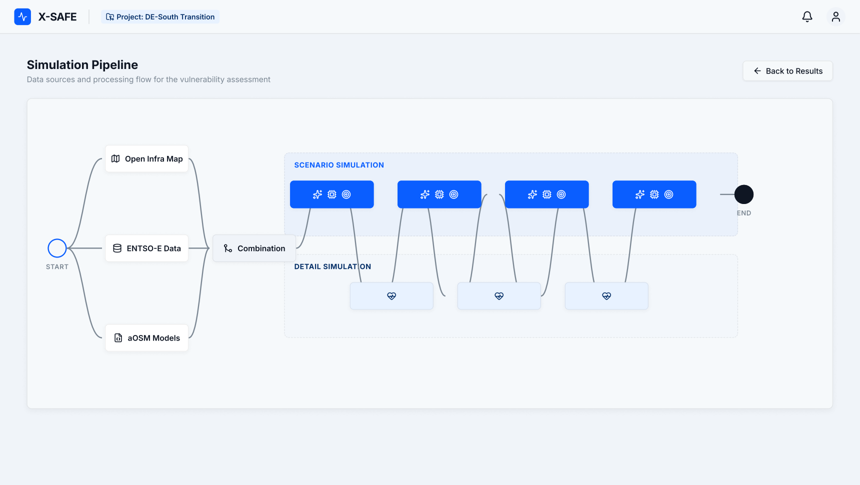

X-SAFE
Project: DE-South Transition
CASCADE ANALYSIS
Analysis Focus
Systemic Weaknesses
ACTIVE
Cascade Paths
Critical Customers
TRIGGER EVENTS
Power Asset Outage
Gas Supply Disruption
Extreme Weather Stress
Cyber / Control Failure
VULNERABILITY VIEW
Show Top N Weak Points
10
Impact Dimension
Supply Security & Load Shedding
RANKING BASIS
Affected Load (MW)
Customers
Critical Services
Economic Loss
Stuttgart
Was sind die Top 10 Vulnerabilitäten der TransnetBW im Großraum Stuttgart?
Simulate

Results UI


X-SAFE
Project: DE-South Transition
Simulation Complete
Vulnerability Assessment
Export Report
IDENTIFIED POINTS
10 Critical
MAX LOAD AT RISK
425 MW
EST. RESTORATION TIME
18 hrs
Customers Served
100%
50%
0%
t=0
Event Trigger
t=24h
Power Loss over Time
500
250
0
t=0
MW
t=24h
Component Failure Risk
Agent Performance Spread
Ctrl A
Ctrl B
Ctrl C
Substation Zuffenhausen
CRITICAL
Node ID: #ST-842
Affected Load
45 MW
Cascading Depth
4 Layers
Loss Probability
82%
View Node Details
Roads
Parks
Was sind die Top 10 Vulnerabilitäten der TransnetBW im Großraum Stuttgart?
Simulated

Carl von Ossietzky Universität Oldenburg

eXplainability UI



Carl von Ossietzky Universität Oldenburg

A Lookout

- The journey towards highly automated grid operation has just begun.
- Current practical approaches try to analyses CPES for weaknesses, uncover cascades & spillover effects
- AI can help testing future grids, be part of certification processes
- AI itself needs safeguards: Rulesets, explainability, and eventually certification, too. (Insurance. . . ?)
- We will see sophisticated agent architectures in the near future.
- If you want to see interesting code, head over to `http://palaestr.ai` or shout out to `eric.veith@uol.de`!

Carl von Ossietzky Universität Oldenburg

Bibliography I

[1] Çaglar Aytekin. "Neural Networks are Decision Trees". In: *CoRR* abs/2210.05189 (2022). [retrieved: 05, 2023], pp. 1–8. arXiv: 2210.05189. URL: https://arxiv.org/abs/2210.05189.

[2] Jimmy Ba and Rich Caruana. "Do deep nets really need to be deep?" In: *Advances in Neural Information Processing Systems* 27 (2014), pp. 2654–2662.

[3] Emilie Frost, Eric Veith, and Lars Fischer. "Robust and Deterministic Scheduling of Power Grid Actors". In: *2020 7th International Conference on Control, Decision and Information Technologies (CoDIT)*. Vol. 1. 2020, pp. 100–105. DOI: 10.1109/CoDIT49905.2020.9263948.

[4] Tohid Shekari, Alvaro A Cardenas, and Raheem Beyah. "MaDIoT 2.0: modern high-wattage IoT botnet attacks and defenses". In: *31st USENIX security symposium (USENIX Security 22)*. 2022, pp. 3539–3556.

[5] Eric Veith, Arlena Wellßow, and Mathias Uslar. "Learning new attack vectors from misuse cases with deep reinforcement learning". In: *Frontiers in Energy Research* (2023).

Carl von Ossietzky Universität Oldenburg

Bibliography II

[6] Eric MSP Veith. *Universal Smart Grid Agent for Distributed Power Generation Management*. Logos Verlag Berlin GmbH, 2017.

[7] Arlena Wellßow, Torben Logemann, and Eric Veith. "Trajectory Generation Model: Building a Simulation Link Between Expert Knowledge and Offline Learning". In: *Proceedings of the 14th International Conference on Simulation and Modeling Methodologies, Technologies and Applications - SIMULTECH*. INSTICC. SciTePress, 2024, pp. 91–102. ISBN: 978-989-758-708-5. DOI: 10.5220/0012764900003758.

[8] Thomas Wolgast, Eric MSP Veith, and Astrid Nieße. "Towards Reinforcement Learning for Vulnerability Analysis in Power-Economic Systems". In: *DACH+ Energy Informatics 2021: The 10th DACH+ Conference on Energy Informatics*. Freiburg, Germany, Sept. 2021, pp. 1–20.