POLITECNICO DI MILANO





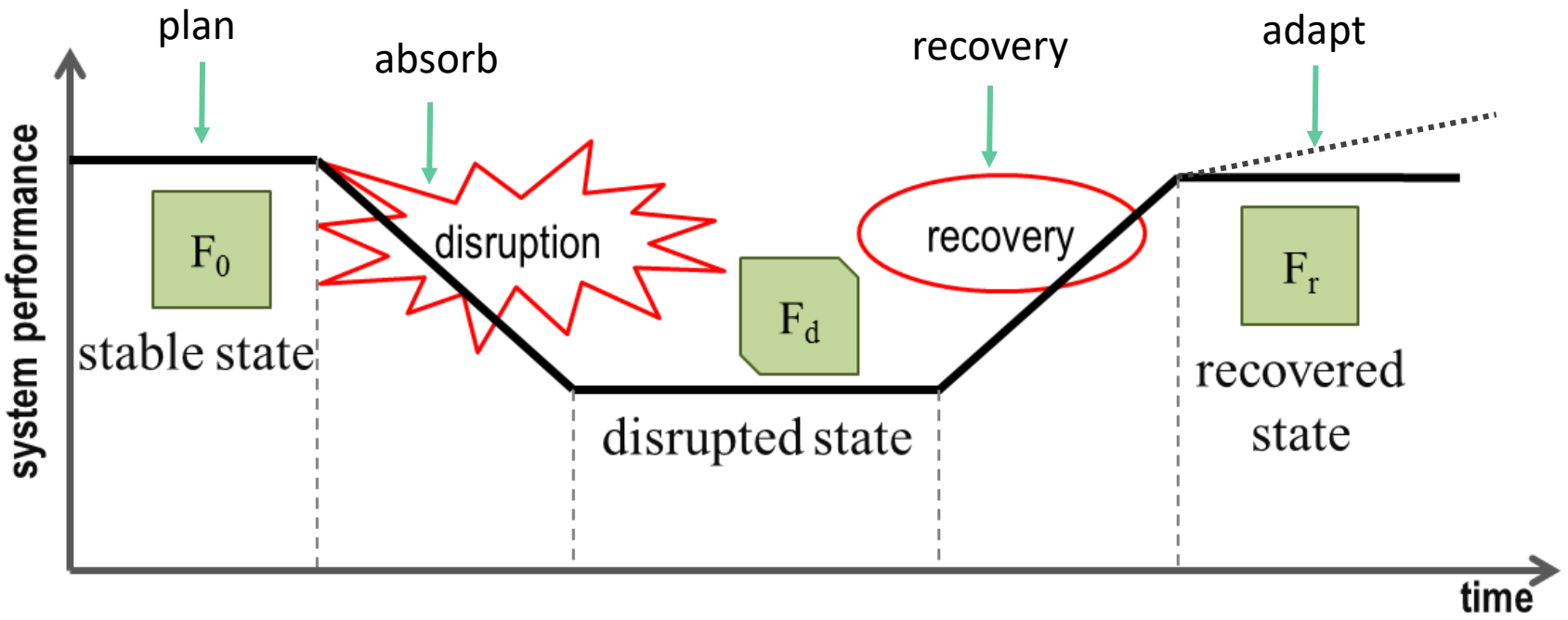
Decision analysis for resilience

(Game theory, Adversarial Risk Analysis, Reinforcement Learning,…)

Prof. Francesco Di Maio
Dipartimento di Energia
Via La Masa 34, B12

francesco.dimaio@polimi.it

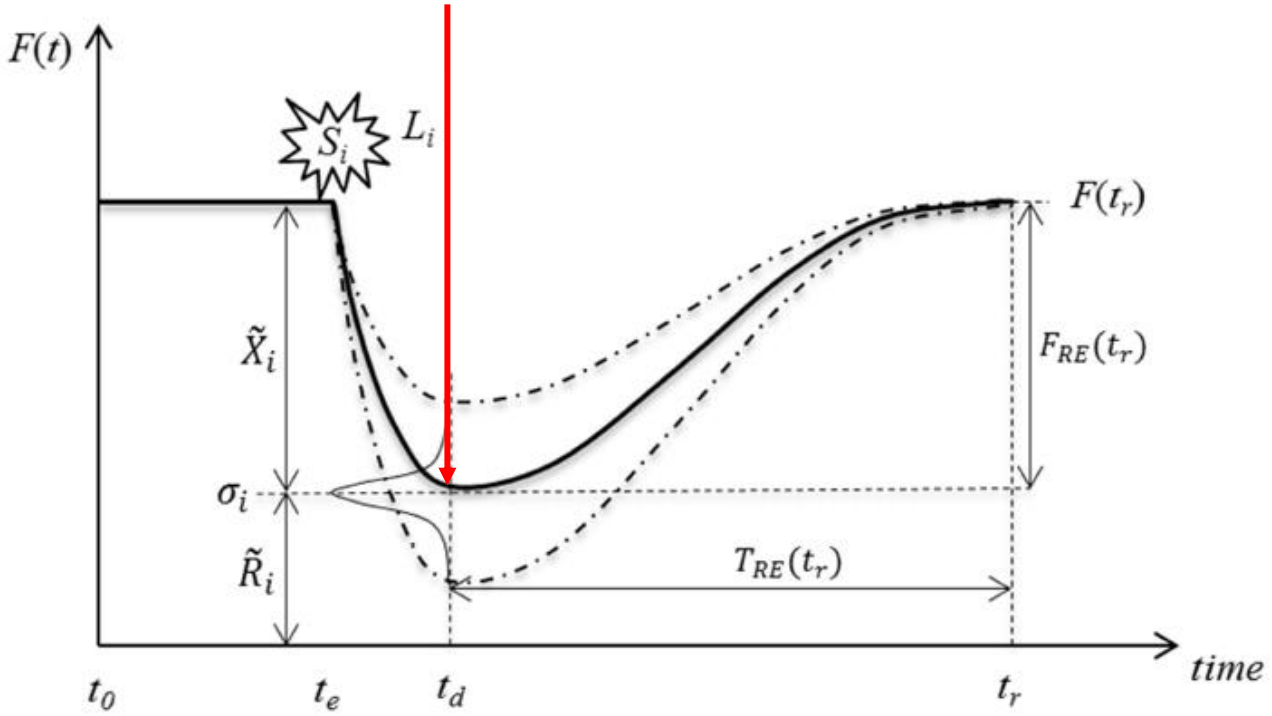
Concept: what is system resilience?



- Resilience concerns the **whole evolution dynamics** of a system to disruptions

Concept: difference with risk

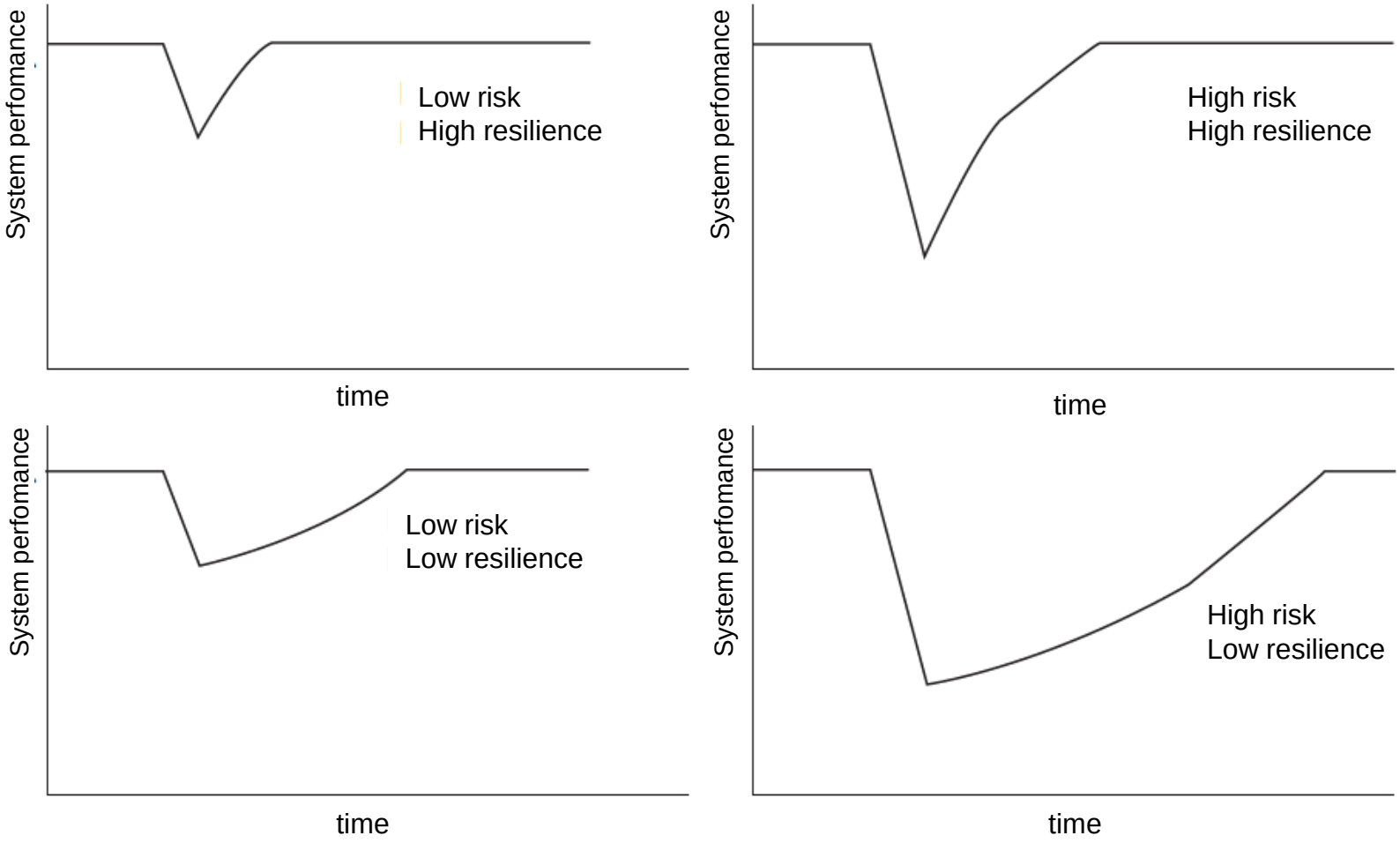
Risk = $\{\langle S_i, L_i, \tilde{X}_i(\sigma_i)\rangle\}$



Resilience = $\{\langle S_i, L_i, \tilde{R}_i, T_{RE}(t_r), F_{RE}(t_r)\rangle\}$

Robustness **Recovery rapidity**

Concept: difference with risk



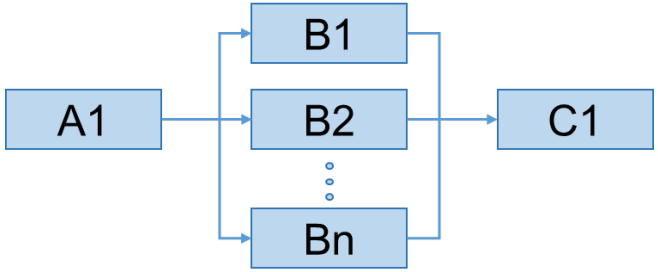
Concept: properties

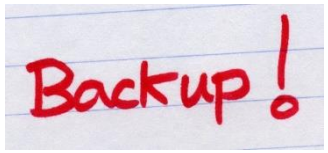
4R framework

Robustness, **R**apidity

Redundancy

- Physical
- Operational





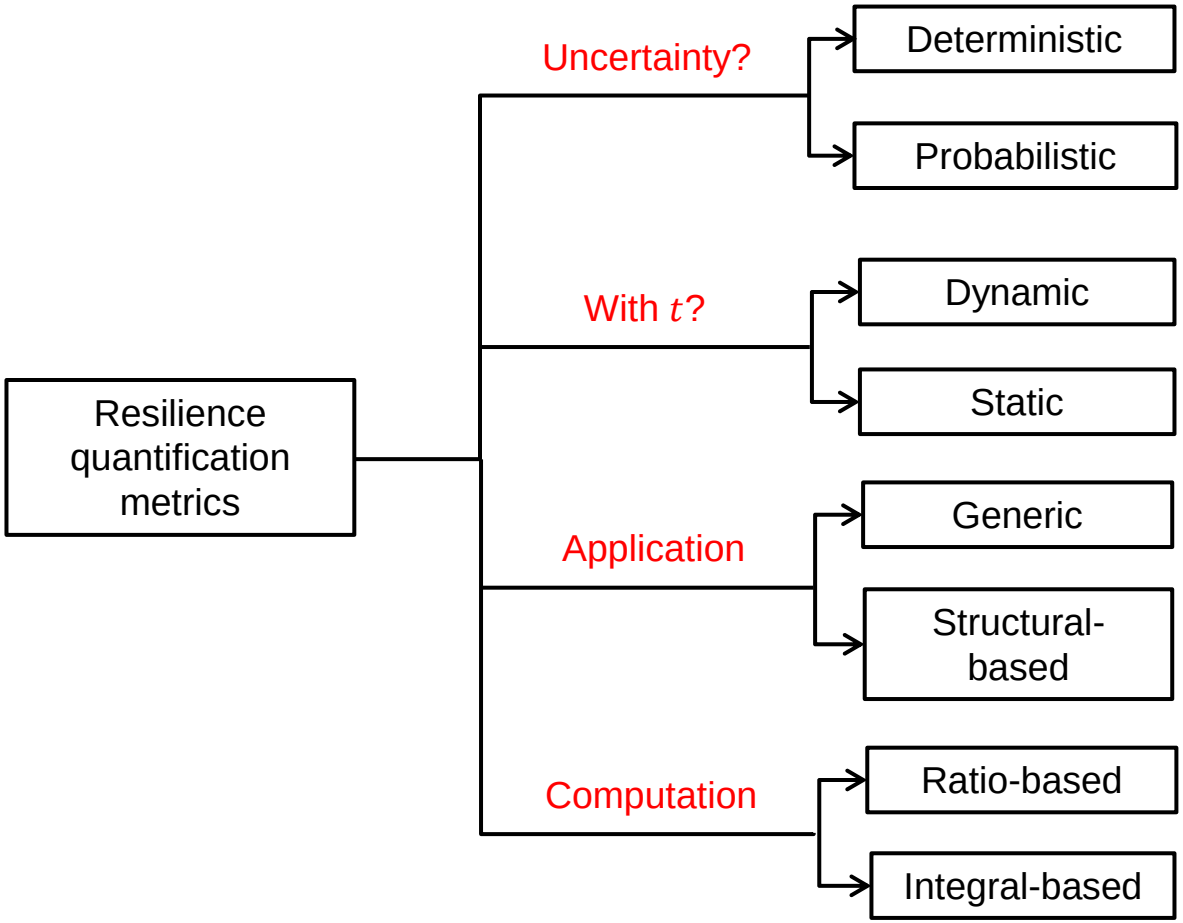
Resourcefulness

- Technical
- Organizational

Resilience quantification

Resilience metrics

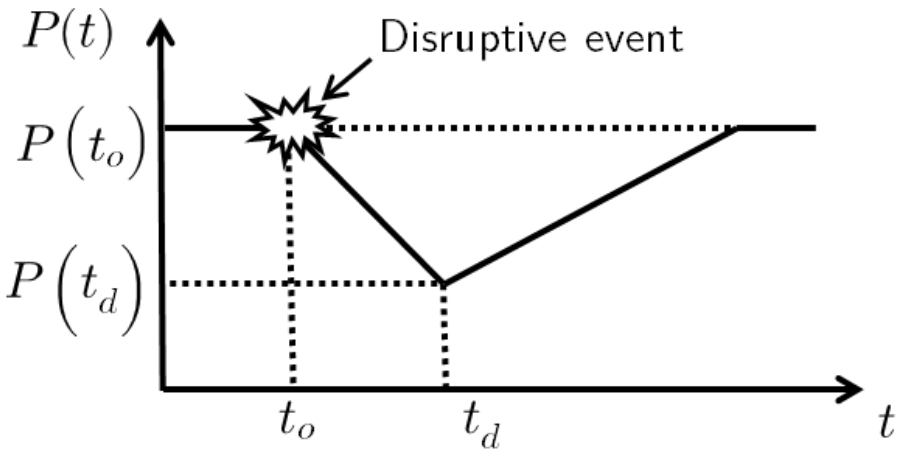
- Classification: depend on the applications of interest, calculation methods, available data, etc.



Resilience metric

Resilience metric 1 (Henry et al. 2012)

$$\boldsymbol{R}(t) = \frac{P(t) - P\left(t_d\right)}{P\left(t_o\right) - P\left(t_d\right)}, t \geq t_d$$



→ $\boldsymbol{R}(t)$: system resilience at time t during the recovery process

→ $P(t)$ system performance function

→ t_o: the time when the external disruptive event occurs

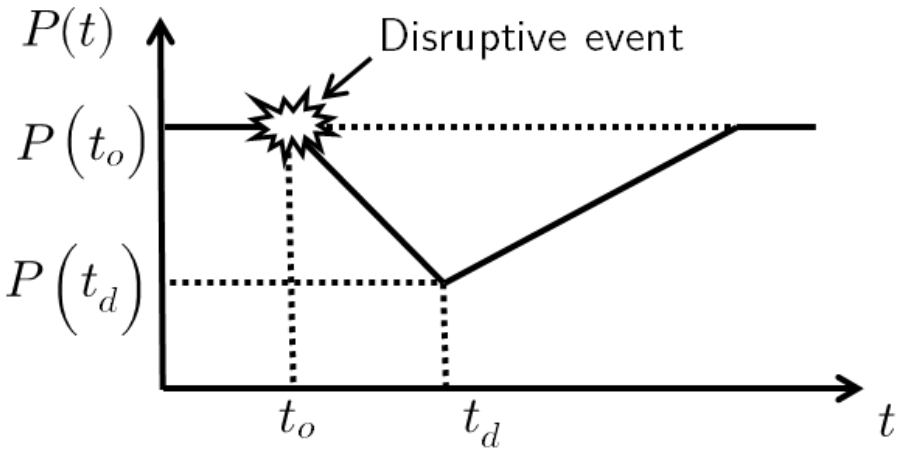
→ t_d: the time when system performance reach its lowest level

Measures the ratio between the recovered performance up to time t and the maximal loss of system performance due to a disruptive event.

Resilience metric

Resilience metric 1 (Henry et al. 2012)

$$\boldsymbol{R}(t) = \frac{P(t) - P\left(t_d\right)}{P\left(t_o\right) - P\left(t_d\right)}, t \geq t_d$$

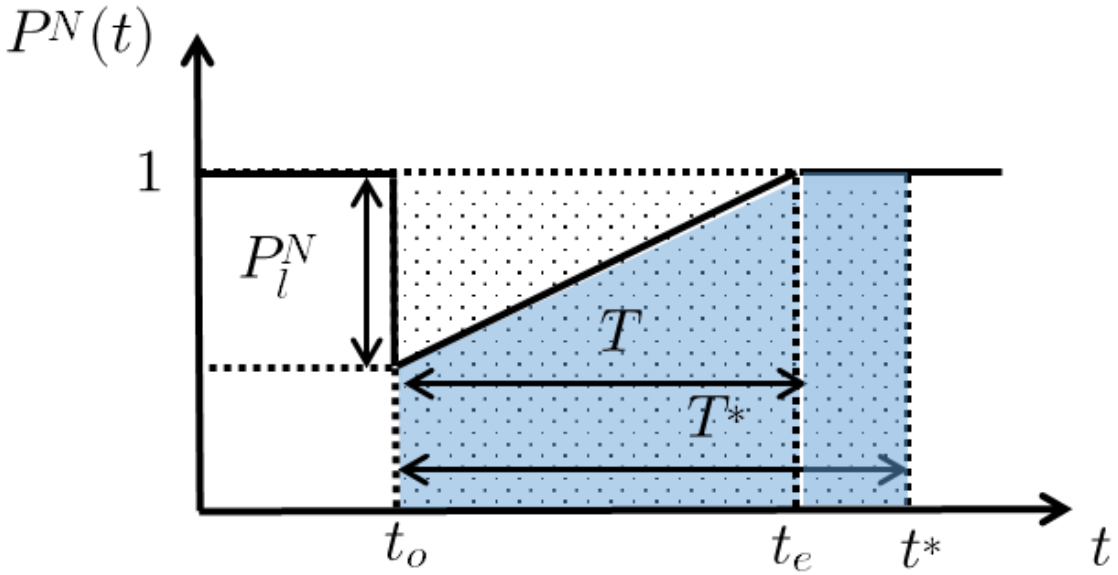


- Deterministic, dynamic
- $P(t)$ could be sometimes high than $P(t_o)$, therefore $\boldsymbol{R}(t)$ could >1
- Consider the lowest level of system performance, while not embrace the failure process, $t < t_d$, in the failure phase

Resilience metric

Resilience metric 2 (Zobel 2011)

$$\boldsymbol{R} = \frac{T^* - P_l^N \cdot T/2}{T^*} = 1 - \frac{P_l^N T}{2T^*}$$



→ $P^N(t)$: normalized system performance function

→ P_l^N: loss of normalized system performance after a disruption

→ t_o: the instant when P(t) reaches its minimum

→ t_e: the instant when the system performance returns to original level

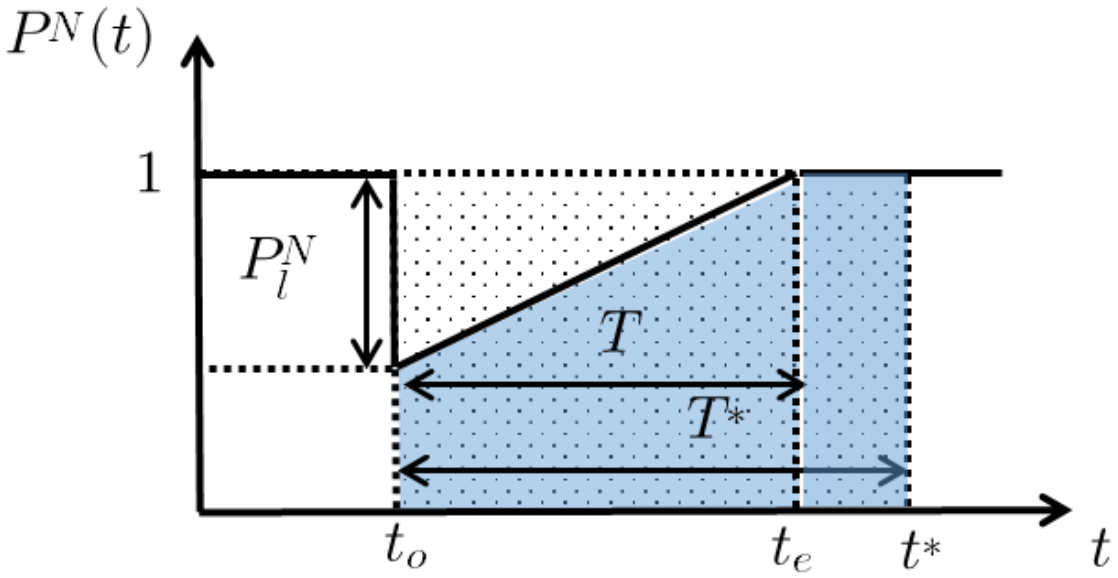
→ t^*: a **strict upper bound** for restoration time t_e

→ $T = t_e - t_o$ and $T^* = t^* - t_o$.

Resilience metric

Resilience metric 2 (Zobel 2011)

$$\boldsymbol{R} = \frac{T^* - P_l^N \cdot T/2}{T^*} = 1 - \frac{P_l^N T}{2T^*}$$

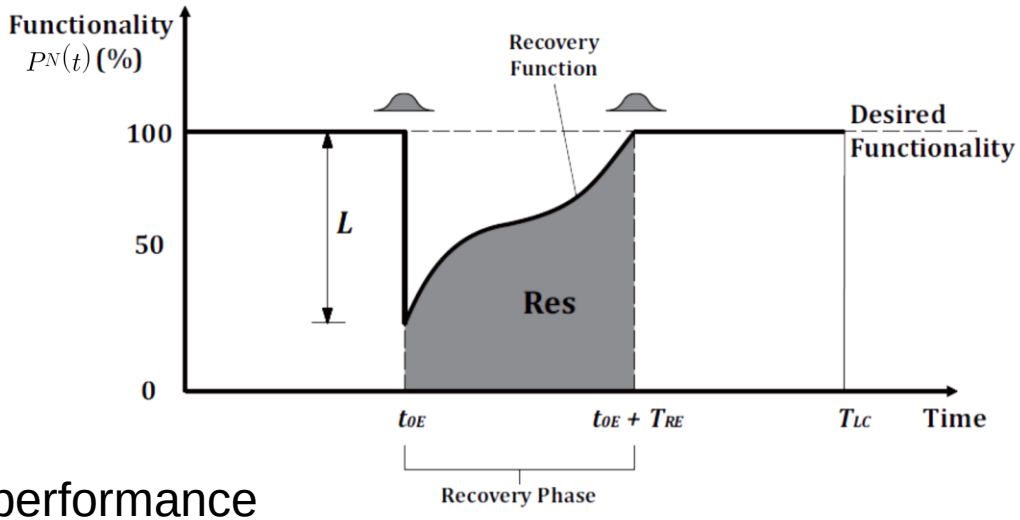


- Measures the ratio between the area between the actual (simplified) system performance curve and the desidered function (=1) and the area below the desired function (=1) and the time axis
- Deterministic, static
- Considers both the performance loss and the length of recovery

Resilience metric

Resilience metric 3 (Bruneau 2003; Bruneau and Reinhorn 2007; Cimellaro 2010 from MCEER)

$$R = \int_{t_{0E}}^{t_{0E}+T_{RE}} \frac{P^N(t)}{T_{RE}}\, dt$$



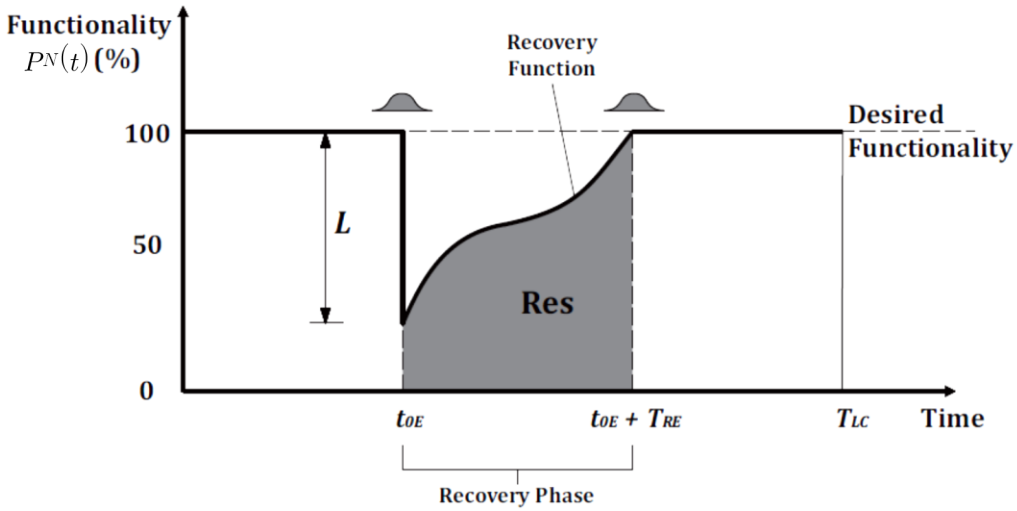
→ $P^N(t)$: normalized system performance
→ t_{0E}: the time when P(t) reaches its minimum
→ T_{RE}: the total recovery time

- The area between the system functionality and the time axis, normalized by the recovery time

Resilience metric

Resilience metric 3 (Bruneau 2003; Bruneau and Reinhorn 2007; Cimellaro 2010 from MCEER)

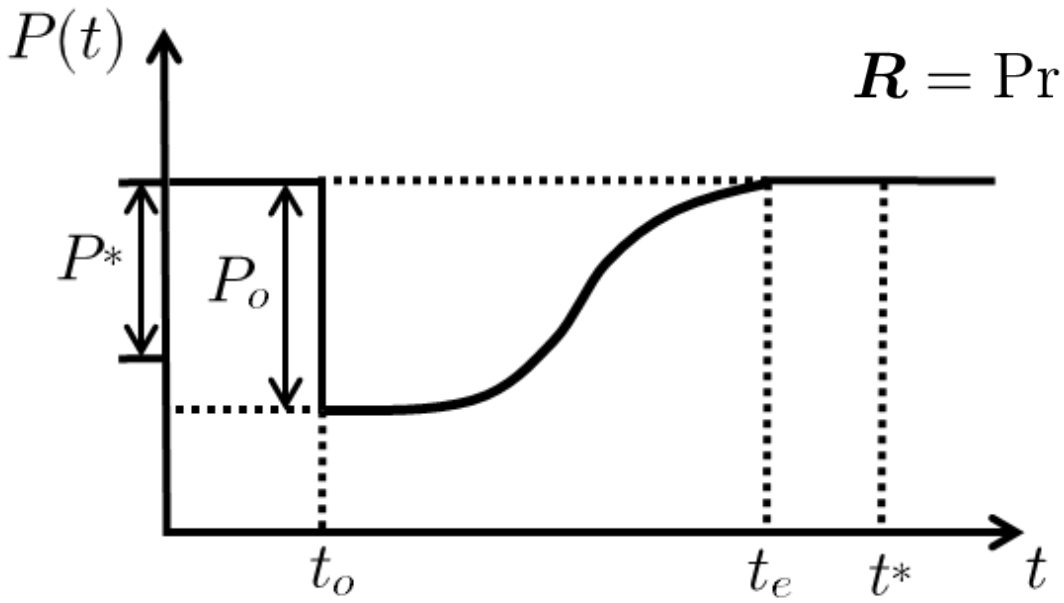
$$R = \int_{t_{0E}}^{t_{0E}+T_{RE}} \frac{P_N(t)}{T_{RE}}\, dt$$



- One of the fundamental metrics
- Deterministic, static
- Give an more accurate evaluation of the resilience level, by considering the true evolution of system performance

Resilience metric

Resilience metric 4 (Chang & Shinozuka 2004)

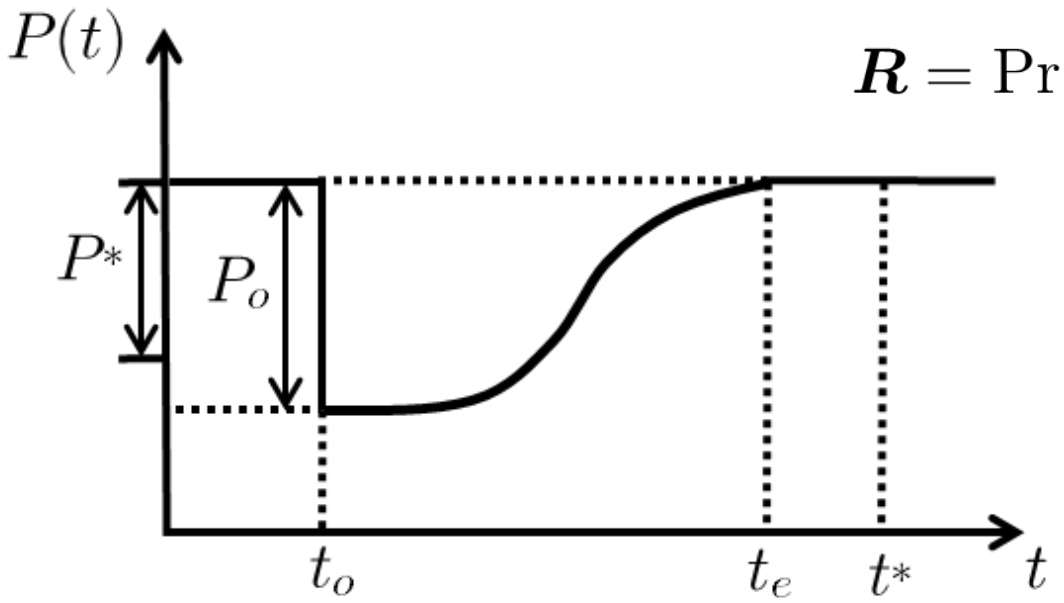


$$\boldsymbol{R} = \Pr(P_o < P^* \text{ and } t_e < t^*)$$

→ P_o: the initial system performance loss after a disruption, i.e., the largest loss during the disruptive event
→ P^*: the maximum acceptable loss of performance
→ t_e: the time when the performance returns to its original level
→ t^*: the maximum acceptable system recovery time

Resilience metric

Resilience metric 4 (Chang & Shinozuka 2004)



$$\boldsymbol{R} = \Pr(P_o < P^* \text{ and } t_e < t^*)$$

- Probabilistic metric, static
- Take into account both the loss of performance and the length of recovery, but not detailed restoration curve.
- Consider uncertainties within the process

Resilience metric

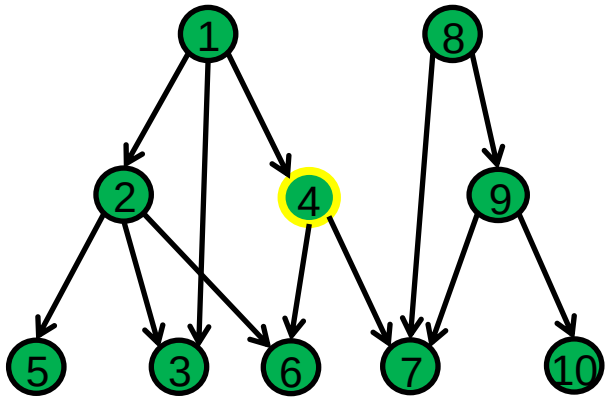
- The use of resilience metrics depends on the **applications** and **available** data, e.g., the types of systems, types of disruptive events.

Metrics	**Value**	**Data required in system performance (SP) dimension**	**Data required in time dimension**	**Feature/advantage in the application**
Metric 1 $\boldsymbol{R}(t) = \frac{P(t) - P(t_d)}{P(t_o) - P(t_d)}$	Function of time	$P(t)$: SP function	t_o and t_d	It is actually a normalized SP function
Metric 2 $\boldsymbol{R} = 1 - \frac{P_l^N T}{2T^*}$	Single value $\in [0,1]$	The lowest value of the normalized SP.	t_o, t_e, and t^*	Provide a rough estimation based on relatively less information
Metric 3 $\boldsymbol{R} = \int_{t_0}^{t_e} \frac{P^N(t)}{T_{RE}} dt$	Single value $\in [0,1]$	$P^N(t)$: the normalized SP function of time	t_o and t_e	Including more information of SP
Metric 4 $\boldsymbol{R} = \Pr(P_o < P^* \cap t_e < t^*)$	Single value Probability	SP value at a critical time instant	t_e and t^*	Taking into account the uncertainty of the event or system

Exercise

A network with nodes and directed links

- Each node has two states: operation vs. failure
- The direction of links indicates the functional dependency
- A node operates when all the nodes it depends on are functioning well

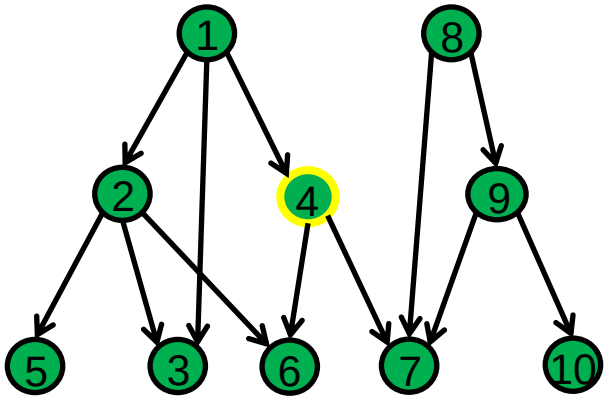


- System performance $P(t)$ is defined as the proportion of operational nodes at each time t.
- At $t = t_o$, a disruption occurs and leads to the failure of certain nodes, and the failure propagates to other nodes.
- It takes Δt_r=1 for a node to shift from operation (failure) to failure (operation) when its dependent nodes are failed (repaired)
- Nodes having redundancy (yellow circle) are able to sustain disruptions

Exercise

Failure scenario

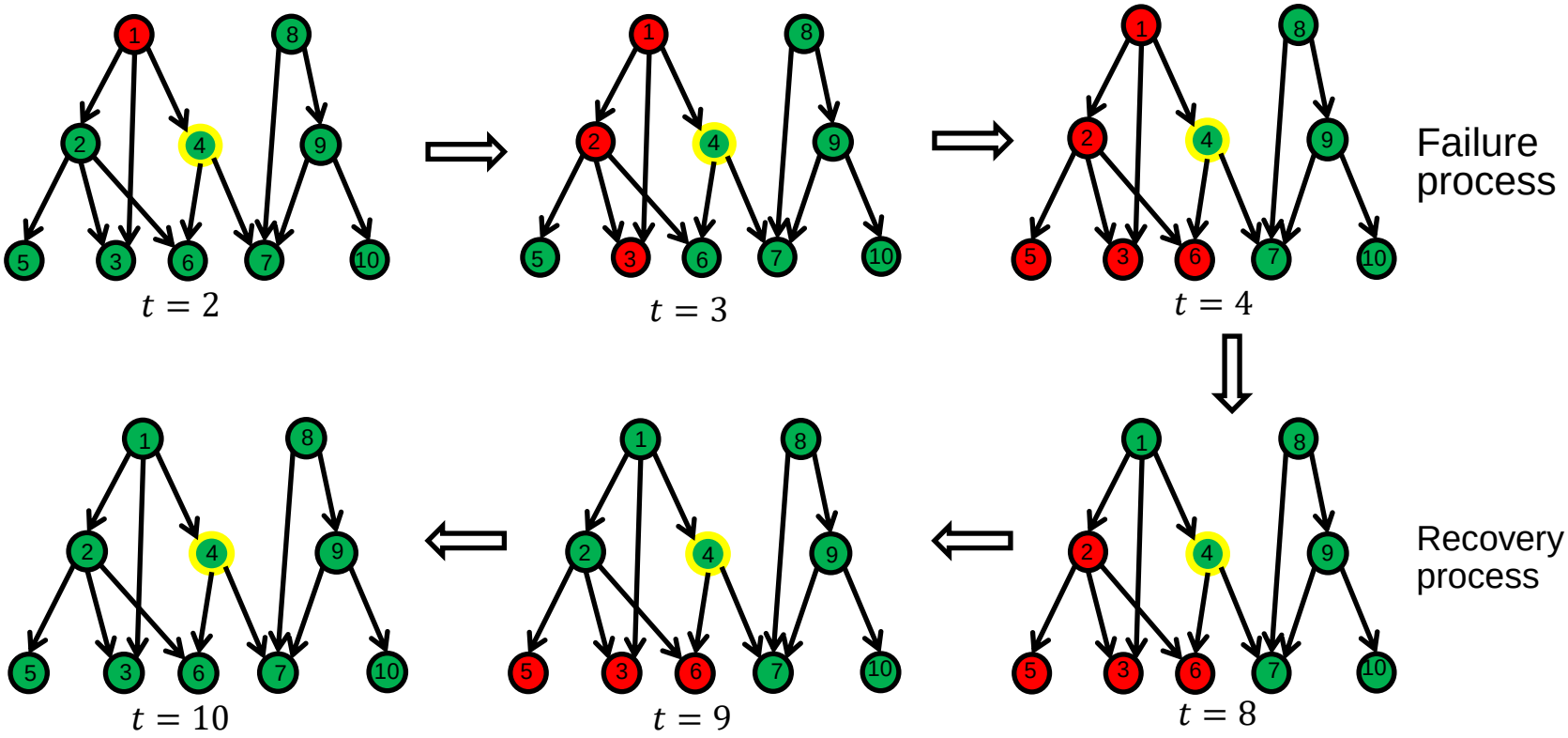
→ At $t_o = 1$, a disruptive event strikes the system, node 1 is failed at t=2
→ At $t_r = 8$, the failed node (node 1) is fixed;
→ The maximum acceptable system recovery instant $t^* = 12$.



Exercise

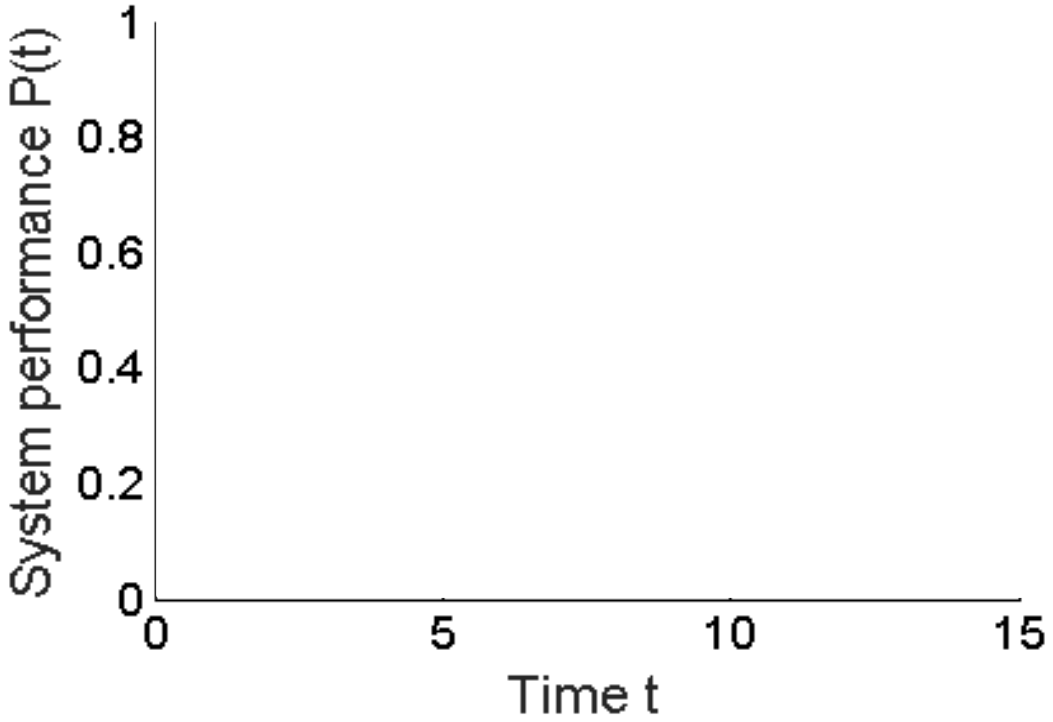
Failure scenario

→ At $t_o = 1$, a disruptive event strikes the system, node 1 is failed at t=2
→ At $t_r = 8$, the failed node (node 1) is fixed;
→ The maximum acceptable system recovery instant $t^* = 12$.



Exercise (solutions)

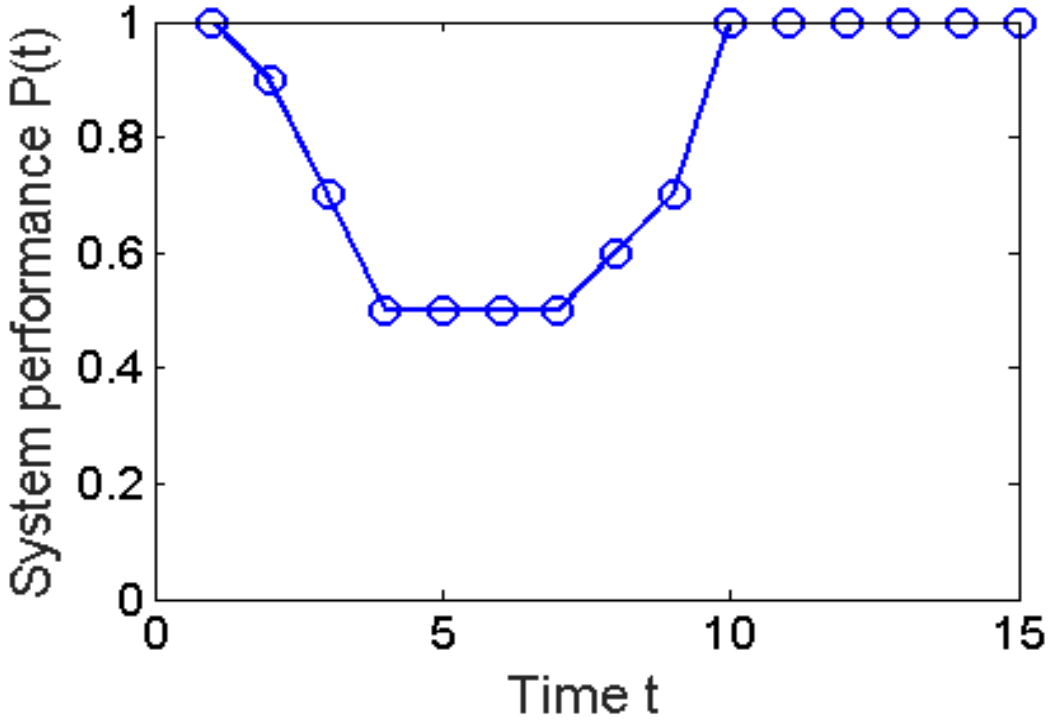
System performance curve calculation



t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P															

Exercise (solutions)

System performance curve calculation



t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Exercise (solutions)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 1 (Henry)

$$\boldsymbol{R}(t) = \frac{P(t) - P\left(t_d\right)}{P\left(t_o\right) - P\left(t_d\right)}, t \geq t_d$$

- The data required are:
  - → $P(t)$
  - → $t_o = 0$
  - → $t_d = 4$
  - → $P(t_o) = P(t = 0) = 1$
  - → $P(t_d) = P(t = 0) = 0.5$

- Essentially, it is a normalized system performance function for $t \geq t_d$

Exercise (solutions)

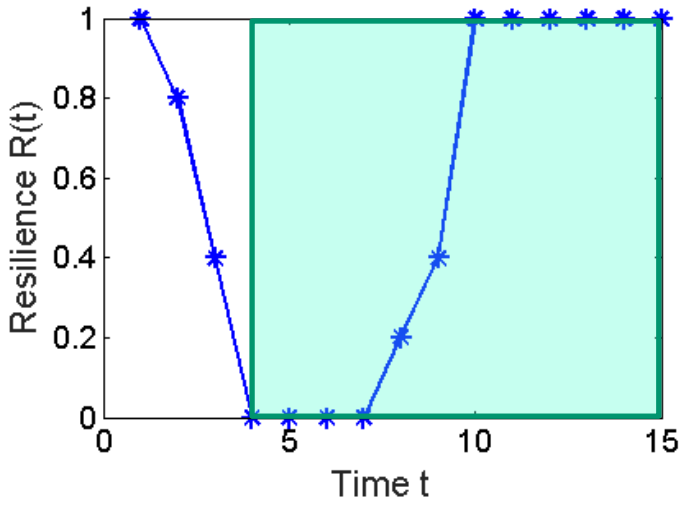
t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 1 (Henry)

$$\boldsymbol{R}(t) = \frac{P(t) - P\left(t_d\right)}{P\left(t_o\right) - P\left(t_d\right)}, t \geq t_d$$

- The data required are:
  - → $P(t)$
  - → $t_o = 0$
  - → $t_d = 4$
  - → $P(t_o) = P(t = 0) = 1$
  - → $P(t_d) = P(t = 0) = 0.5$

$$\boldsymbol{R}(t) = \frac{P(t) - P(t_d)}{P(t_o) - P(t_d)} = 2P(t) - 1$$



- Essentially, it is a normalized system performance function for $t \geq t_d$

Exercise (solutions)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 2 (Zoebel)

$$\boldsymbol{R} = \frac{T^* - P_l^N \cdot T/2}{T^*} = 1 - \frac{P_l^N T}{2T^*}$$

- The data required are:
  - → $P_l^N = 0.5$
  - → $t_o =$
  - → $t_e = 10$
  - → $t^* = 12$
  - → $T = t_e - t_o$
  - → $T^* = t^* - t_o$

Exercise (solutions)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 2 (Zoebel)

$$R = \frac{T^* - P_l^N \cdot T/2}{T^*} = 1 - \frac{P_l^N T}{2T^*}$$

- The data required are:
  - → $P_l^N = 0.5$
  - → $t_o = 4$
  - → $t_e = 10$
  - → $t^* = 12$
  - → $T = t_e - t_o = 6$
  - → $T^* = t^* - t_o = 8$

$$\Rightarrow \quad R = 1 - \frac{0.5 * (10 - 4)}{2 * (12 - 4)} = 0.81$$

- A single value in the range of [0, 1]
- Only require field data about P_l^N and T, a **rough estimation**

Exercise (solutions)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 2 (Zoebel)

$$R = \frac{T^* - P_l^N \cdot T/2}{T^*} = 1 - \frac{P_l^N T}{2T^*}$$

- The data required are:
  - → $P_l^N = 0.5$
  - → $t_o = 1$
  - → $t_e = 10$
  - → $t^* = 12$
  - → $T = t_e - t_o = 9$
  - → $T^* = t^* - t_o = 11$

$$\Rightarrow \quad R = 1 - \frac{0.5 * (10 - 1)}{2 * (12 - 1)} = 0.8$$

- A single value in the range of [0, 1]
- Only require field data about P_l^N and T, a **rough estimation (more conservative → considers the failure process)**

Exercise (solutions)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 3 (Bruneau)

$$R = \int_{t_{0E}}^{t_{0E}+T_{RE}} \frac{P^N(t)}{T_{RE}} dt$$

- The data required are:
  - → $P^N(t)$
  - → $t_{oE} = 4$
  - → $t_e = 10$
  - → $T_{RE} = t_e - t_{oE}$
- A single value in the range of [0, 1]
- Require data of system performance $P^N(t)$ for the whole recovery process

Exercise (solutions)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 3 (Bruneau)

$$R = \int_{t_{0E}}^{t_{0E}+T_{RE}} \frac{P^N(t)}{T_{RE}} dt$$

- The data required are:
  - → $P^N(t)$
  - → $t_{oE} = 4$
  - → $t_e = 10$
  - → $T_{RE} = t_e - t_{oE}$

$$\Rightarrow \quad R = \int_{t=4}^{10} \frac{P^N(t)}{T_{RE}} dt = 0.55$$

- A single value in the range of [0, 1]
- Require data of system performance $P^N(t)$ for the whole recovery process

Exercise (solutions)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 3 (Cimellaro)

$$R = \int_{t_{0E}}^{t_{0E}+T_{RE}} \frac{P^N(t)}{T_{RE}} dt$$

- The data required are:
  - → $P^N(t)$
  - → $t_{oE} = 1$
  - → $t_e = 10$
  - → $T_{RE} = t_e - t_{oE}$

- **More realistic → considers the failure process**

Exercise (solutions)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 3 (Cimellaro)

$$R = \int_{t_{0E}}^{t_{0E}+T_{RE}} \frac{P^N(t)}{T_{RE}} dt$$

- The data required are:
  - → $P^N(t)$
  - → $t_{oE} = 1$
  - → $t_e = 10$
  - → $T_{RE} = t_e - t_{oE}$

$$\Rightarrow \quad R = \int_{t=1}^{10} \frac{P^N(t)}{T_{RE}} dt = 0.59$$

- **More realistic → considers the failure process**

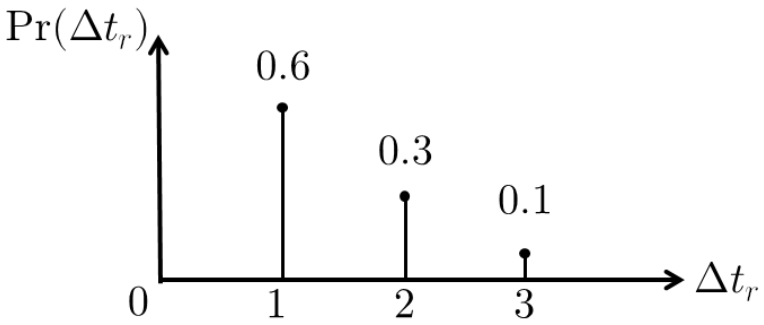
Exercise (solutions)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 4 (Chang)

$$\boldsymbol{R} = \Pr(P_o < P^* \text{ and } t_e < t^*)$$

- The data required are:
  - → **Consider uncertainty:** the recovery time Δt_r of failed node is a discrete random variable



  - → The maximum acceptable performance loss P^*=0.5
  - → The maximum acceptable recovery time t^*= 12

Exercise (solutions)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 4 (Chang)

$$\boldsymbol{R} = \Pr(P_o < P^* \text{ and } t_e < t^*)$$

- Monte Carlo simulation for the recovery process
- Example of system performance curves

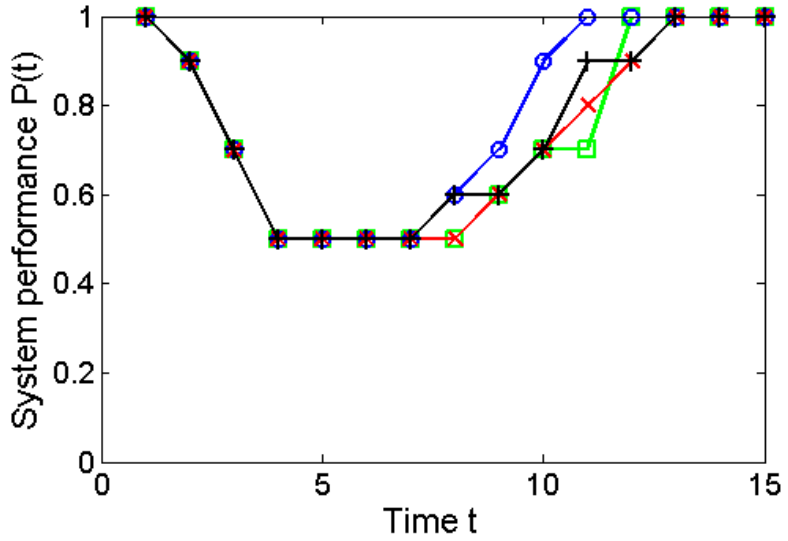
Exercise (solutions)

t	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P	1	0.9	0.7	0.5	0.5	0.5	0.5	0.6	0.7	1	1	1	1	1	1

Resilience metric 4 (Chang)

$$\boldsymbol{R} = \Pr(P_o < P^* \text{ and } t_e < t^*)$$

- Monte Carlo simulation for the recovery process
- Example of system performance curves



$$\boldsymbol{R} = 0.68$$

- we only need to check if the time that the system is recovered to 100% is earlier than t=12.

Resilience assessment

- **Objective**: quantitatively evaluate the resilience of a system against **specific hazards**
- **Methods:**
  - **Statistical** methods based on historical events data if available
  - **Simulation**-based computational methods if hazards are predictable and calculable, e.g., can be modeled via probabilistic approaches
  - **Worst-case analysis** methods for deep uncertain hazards
    - → Nondeliberate hazards, e.g., extreme disasters
    - → Climate Change
    - → Deliberate threats (e.g., vandalism, sabotage, and terrorism)

A game-theory worst-case assessment framework

"WORST-CASE" DISRUPTION MODEL

- Hypothetical intelligent adversary
- Disruption: loss of sets of system components
- Worst-case adversary behavior: use **limited** capability to inflict maximum damage

SYSTEM RESPONSE MODEL

- Model the function of a system of interest
- Operator makes decisions about system activities after disruption
- What we want (objectives) vs. What is feasible (constraints)

Attacker-Defender Interdiction Model

$$\min_{x \in X} \max_{y \in Y(x)} P(y)$$

→ X: limited capability of the attacker
→ $Y(x)$: defender's feasible operation space, in a function of x
→ $P(y)$: What we want, system performance function
→ Min or max depending on the definition of P(y)

A game-theory worst-case assessment framework

Attacker-Defender Interdiction Model

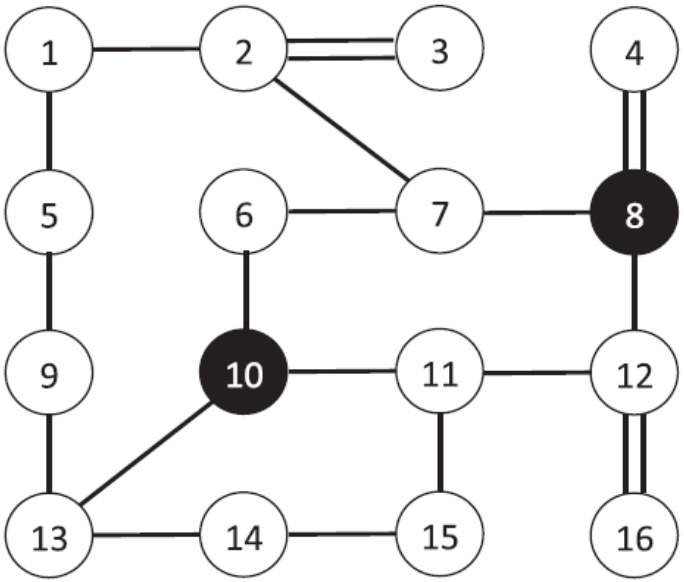
$$\min_{x \in X} \max_{y \in Y(x)} P(y)$$

- A Stackelberg (leader-follower) game
- Solution is given by its Stackelberg Equilibrium
- ***Theorem***: in a finite game with 2 players, i.e., X and $Y(x)$ are finite, there is **always** a Stackelberg equilibrium

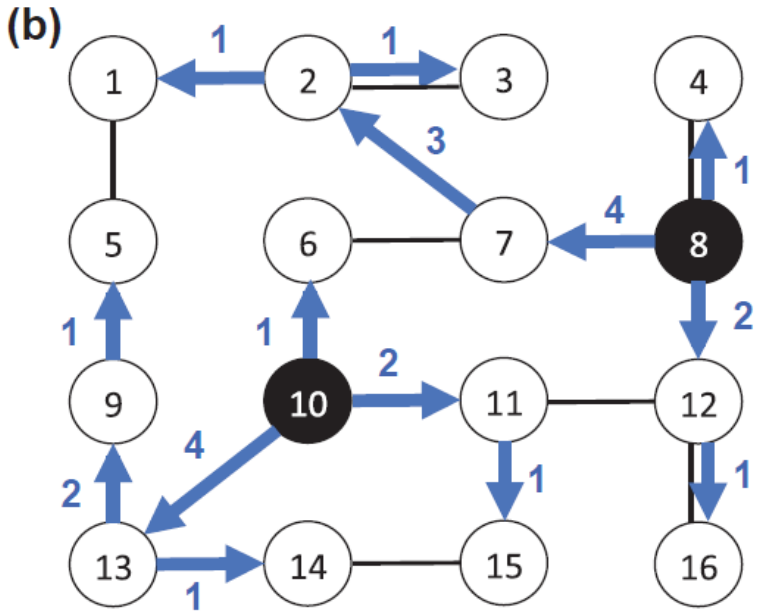
A game-theory worst-case assessment framework

An example: fuel supply network

(a)



(b)



- White → demand =1 barrel of fuel
- Black → suppliers = 10 barrels
- Links → bidirectional, capacity = 15, transmission cost = 1€/barrel
- Penalty of unsatisfied demand: 10€/barrel
- We care about **the total cost**

$$\min_{y \in Y} Cost(y)$$

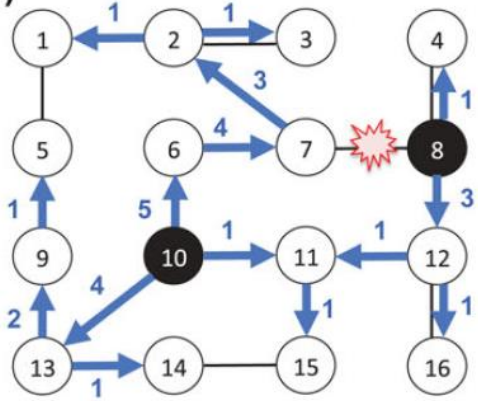
- A baseline flow pattern corresponding the minimum cost flow solution
- Total cost = **25€**

A game-theory worst-case assessment framework

An example: fuel supply network

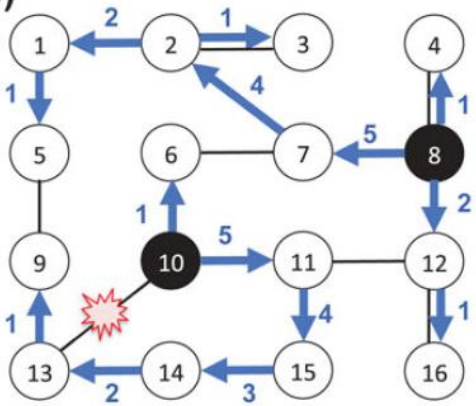
$$\max_{x \in X} \min_{y \in Y(x)} Cost(y)$$

(a)



Total cost = **30**€

(b)



Total cost = **33**€

(c)

Rank	Consequence	Links
1	+8	[10,13]
2	+7	[2,7]
3	+5	[7,8]
4	+4	[9,13]
T5	+2	[8,12], [10,11]
T7	+1	[1,2], [5,9], [6,10], [11,15], [13,14]

All other links result in zero consequence if broken:
[1,5], [2,3], [4,8], [6,7], [11,12], [12,16], [14,15]

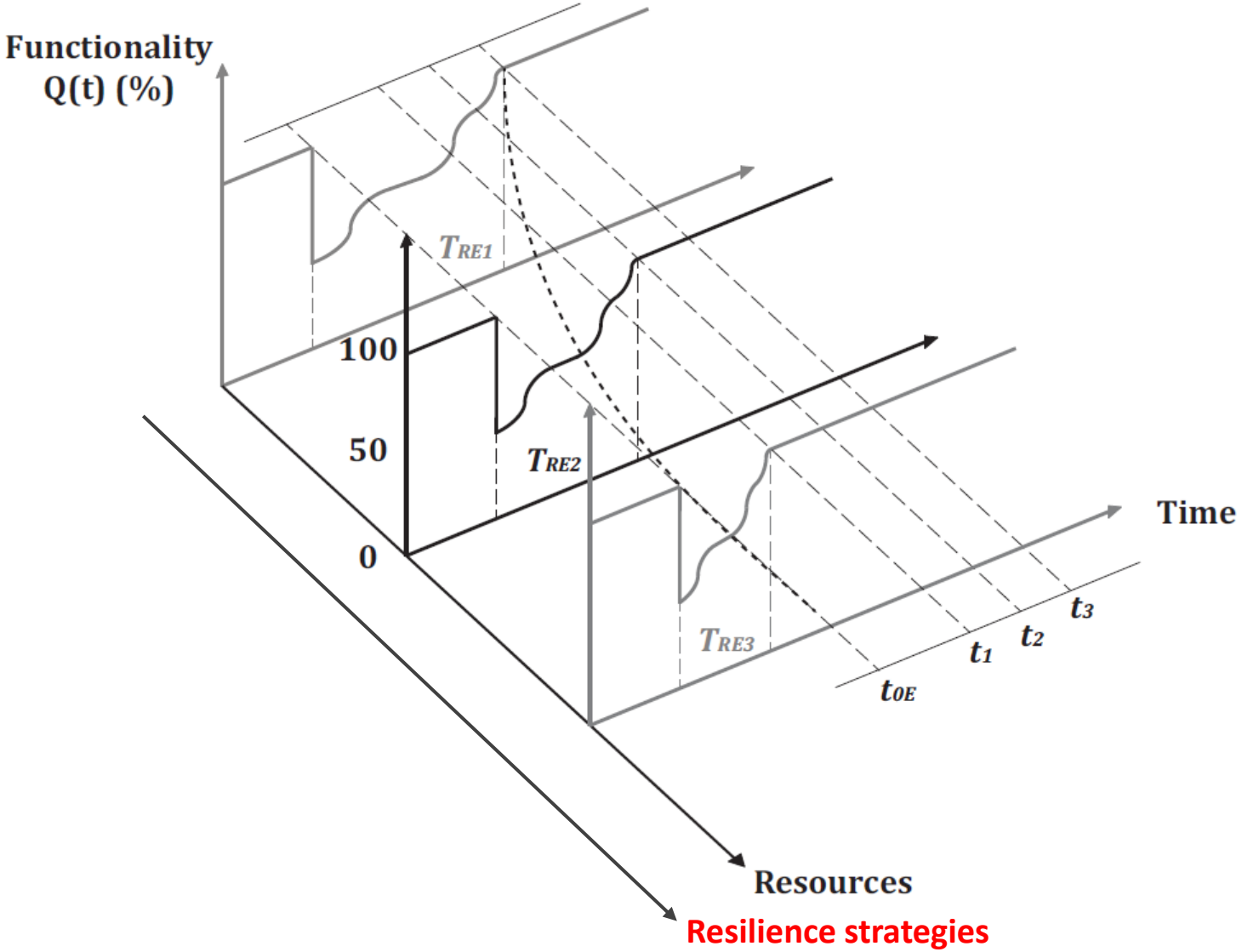
- x→only single link break is allowed
- Recourse actions y: re-dispatching network flow
- Incurs additional operating cost, but does not prevent fuel from being delivered to each demand
- Worst-case single disruption: **[10, 13]** → increase **8€** cost

A game-theory worst-case assessment framework

How does the worst-case assessment help?

- Avoid the two biggest gotchas:
  - → "We didn't know that X would cause Y…"
  - → "We never thought that could happen…"
- We don't have to guess at scenarios
  - → (or try to assess the intent of bad guys)

Resilience-oriented decision making



Resilience strategies

- To enhance resilience, **resilience strategies** include:
  - → Enhance the resilience awareness
  - → Share information
  - → Make integrated decision makings
  - → Train staff and managers
  - → Harden system components
  - → Adjust system topology
  - → Control system demand level
  - → Deploy backup systems (redundancy)
  - → Optimize repair sequence
- Resilience strategies are **system specific**
- Generally, in the time domain: **pre-event strategies** vs. **post-event strategies**, e.g., hardening vs. repair crew scheduling

Resilience-oriented decision making

Optimal pre-event planning for resilience enhancement against threats

- The system operators **make decisions about the pre-event resilience strategies** in order to reconcile the **objectives** (enhance resilience) with its **constraints** in an intelligent and efficient manner.

Constrained optimization:

Defender-Attacker-Defender models

$$\max_{w \in W} \min_{x \in X(w)} \max_{y \in Y(w,x)} P(y)$$

Pre-event resilience strategies: defensive investments and budget constraints for the "defender"

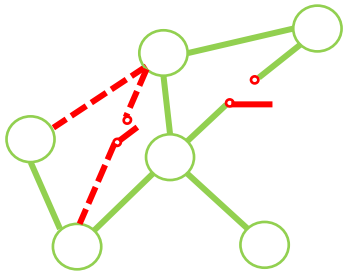
The worst-case resilience assessment framework

Resilience-oriented decision making

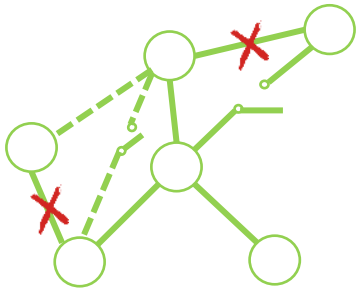
Defender-Attacker-Defender models

$$\max_{w \in W} \min_{x \in X(w)} \max_{y \in Y(w,x)} P(y)$$

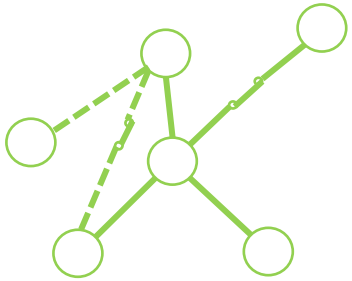
System defender: determine the pre-event resilience strategies, w, pursuing maximal system performance



Disruptive agent: maximize the damage considering his ability $X(w)$



System operator: mitigate the negative effect of the disruptive attack by recourse actions $y \in Y(w, x)$



Resilience-oriented decision making: ALFRED

The Advanced Lead-cooled Fast Reactor European Demonstrator (ALFRED)

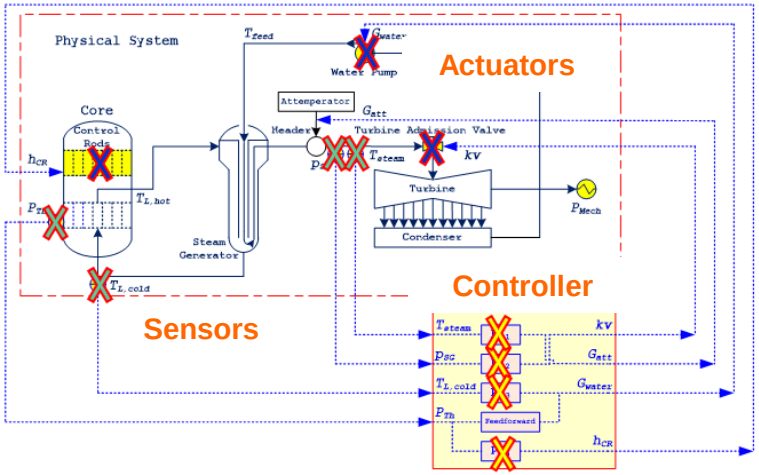
(Attacker)

Single attack a

- DoS attack to 4 Sensors;
- Stuck failure to 3 actuators;
- FDI attack to 4 PIs.

UNCERTAIN to defender (subjective judgment)

a∈A



ALFRED subjected to cyber attacks

d*

(Defender)

Defensive barriers d

Prevention:
- Firewall;
- Intrusion detection system;
- Operators;
- Security software;

Recovery:
- Mainframes computers;
- Database servers;
- Security engineers.

Objective: to provide a one-sided (i.e., defender) prescriptive support strategy d* for optimizing allocation of resources for the defensive barriers based on a subjective expected utility model.

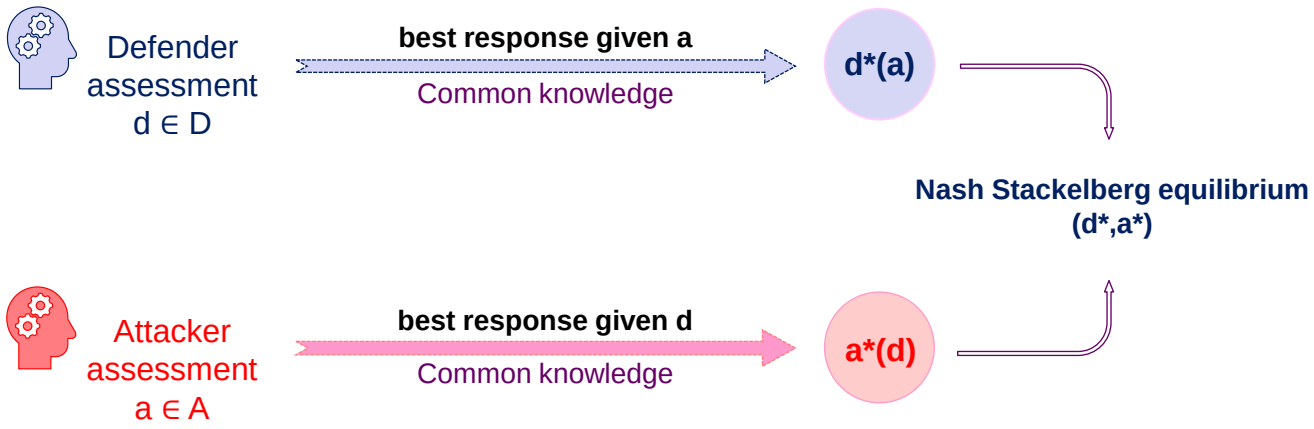
Resilience-oriented decision making: ALFRED

State of the art: traditional game-theoretical defend-attack modeling

- The players (defender and attacker) rely on shared knowledge;

The defender:

- √ knows own beliefs and preferences;
- √ knows those of the attacker (and vice versa);



Resilience-oriented decision making: ALFRED

[J] Wang W., Di Maio F., Zio E. Risk Analysis, (2018)

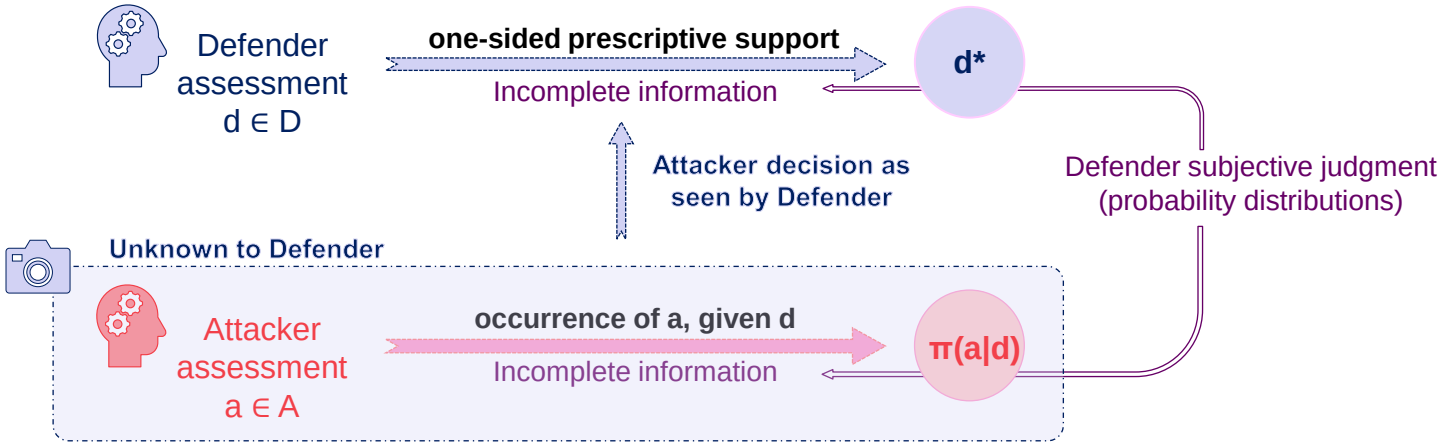
Approach: Adversarial Risk Analysis (ARA) game-theoretical defend-attack modeling

- incomplete information,;

The defender:

√ only know own beliefs and preferences;

X only speculates those of the attacker;



Resilience-oriented decision making: ALFRED

[J] Wang W., Di Maio F., Zio E. Risk Analysis, (2018) under review.

Approach: Monte Carlo (MC) scheme embedded within the ARA modeling

- **Attacker decision for π(a|d), as seen by defender:**

Unknown to Defender

OUTER: **given d**, propagate uncertainty among N_m runs, for estimating **frequency of occurrence of a being the optimal $a^{*,m}(d)$, π(a|d):**

INNER: calculate attacker expected utility of each a, for finding the optimal one:

$$a^{*,m}(d) = \text{argmax}_a \Psi_A^m(a|d)$$

- sample incomplete information (e.g. costs) from subjective distributions

π(a|d) for each d

- **Defender decision for d*:**

OUTER: propagate uncertainty among N_v runs, for estimating frequency of d being the optimal $d^{*,v}$, and finding **the optimal allocation d* of resource for defensive barriers** by:

$$d^* = \text{argmax}_d f(d)$$

d*

INNER: calculate defender expected utility of each d, for finding the optimal one:

$$d^{*,v} = \text{argmax}_d \Psi_D^v(d)$$

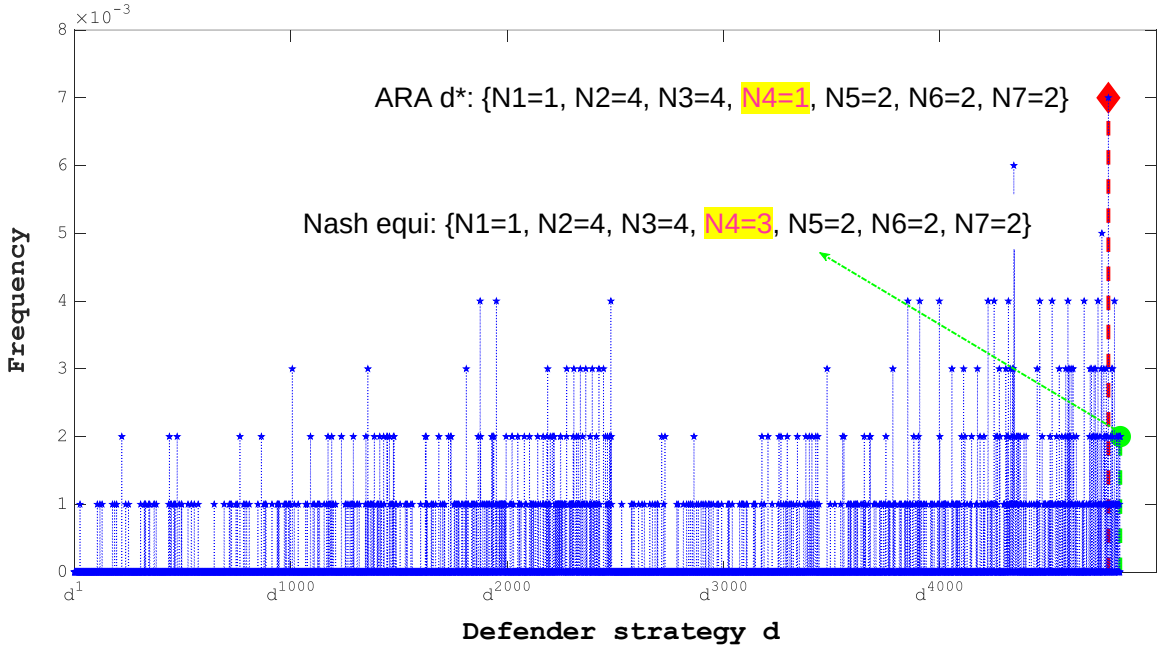
where,

$$\Psi_D^v(d) = \sum_a \left[\sum_s p_D(s|d,a) \cdot u_D(d,a,s) \right] \cdot \pi(a|d)$$

- sample incomplete information (e.g. costs) from subjective distributions

Resilience-oriented decision making: ALFRED

Results: ARA model



Prevention:
- N1=Firewall;
- N2=Intrusion detection system;
- N3=Operators;
- N4=Security software;

Recovery:
- N5=Mainframes computers;
- N6=Database servers;
- N7=Security engineers.

With maximum defense budget = 2,000 k€

What we have done:

- Optimization of the allocation of defensive barriers against cyber attacks by Adversarial Risk Analysis (ARA);

What we have found (case study):

- ARA beats Nash equilibrium in ALFRED;

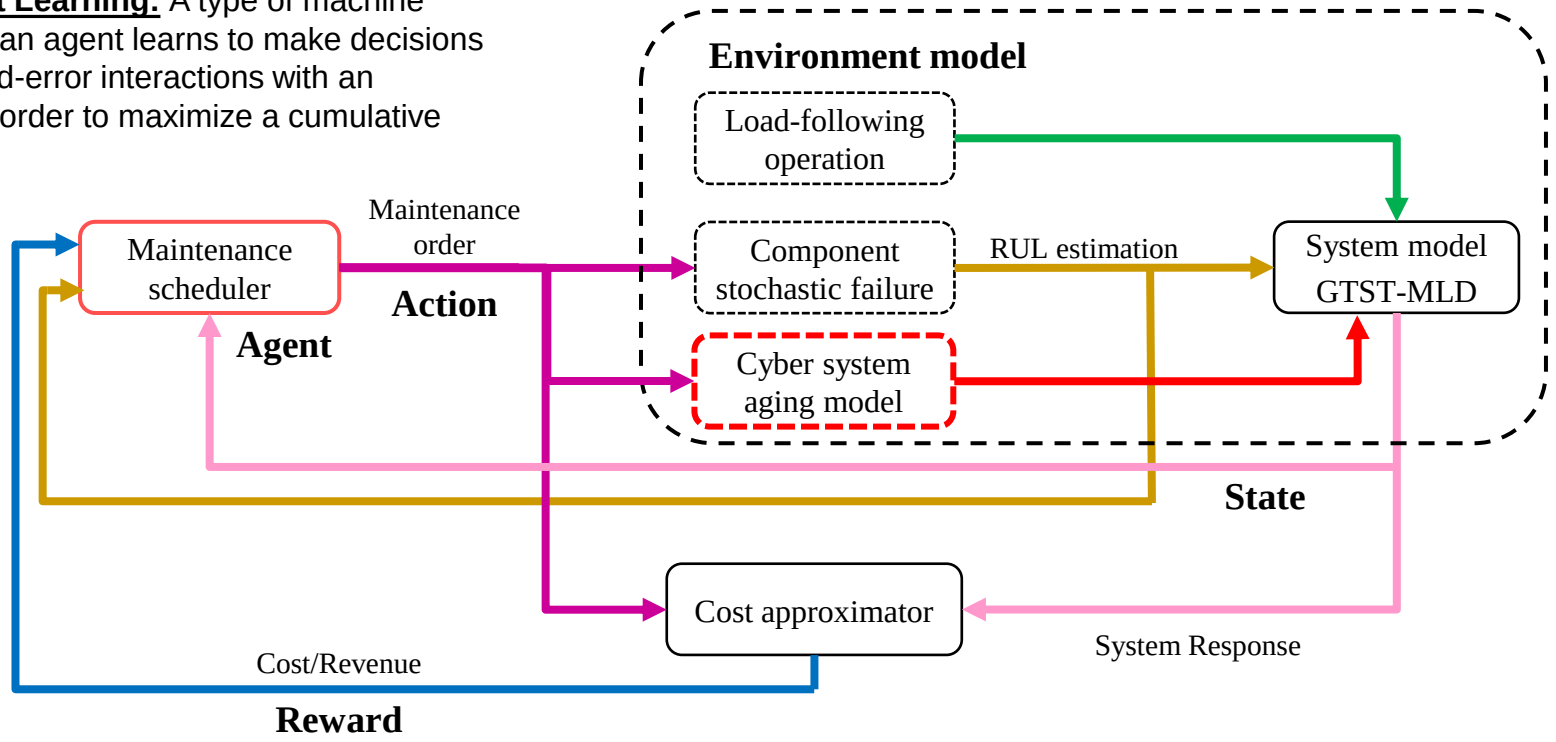
[J] W. Wang, F. Di Maio, E. Zio. Adversarial Risk Analysis to Allocate Optimal Defense Resources for Protecting Nuclear Power Plants from Cyber Attack. Risk Analysis,

Resilience-oriented decision making

Optimal pre-event planning for resilience enhancement against threats

- The system operators **make decisions about the pre-event resilience strategies** in order to reconcile the **objectives** (enhance resilience) with its **constraints** in an intelligent and efficient manner.

Reinforcement Learning: A type of machine learning where an agent learns to make decisions through trial-and-error interactions with an environment in order to maximize a cumulative reward.



Profit-driven-reward Reinforcement Learning

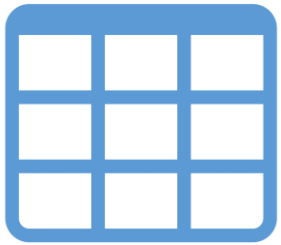
Approach: Deep Reinforcement Learning

State-value

or

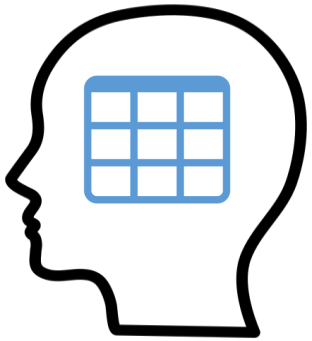
Action-value

Value Tabular

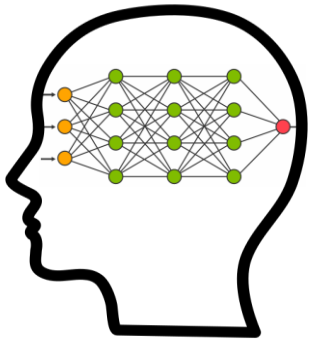


Tabular RL

For continuous state space or complex system, it's impossible to list all the combinations



Tabular



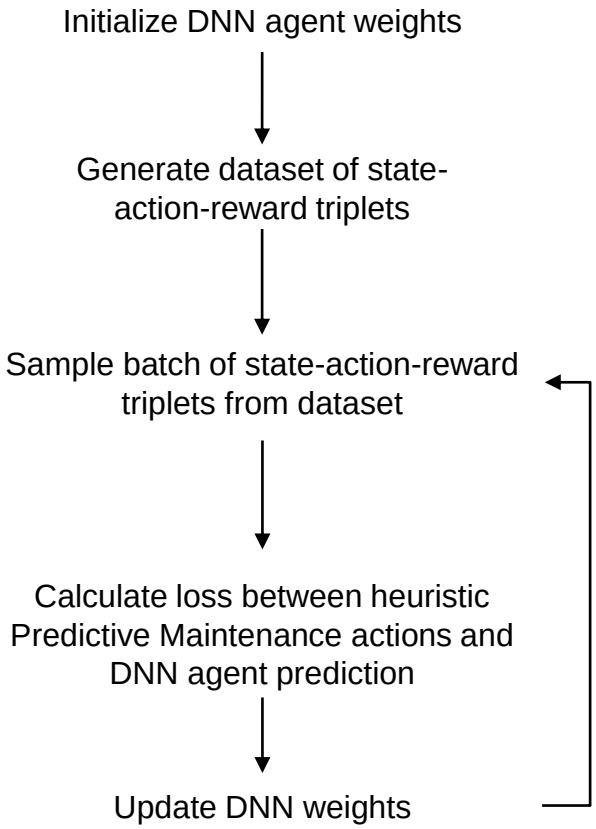
Deep Neural Network (DNN)

Profit-driven-reward Reinforcement Learning

Approach: Imitation Learning (IL)

Imitation Learning (IL) is a type of supervised learning in which an agent learns to perform a task by mimicking the behavior of an expert.

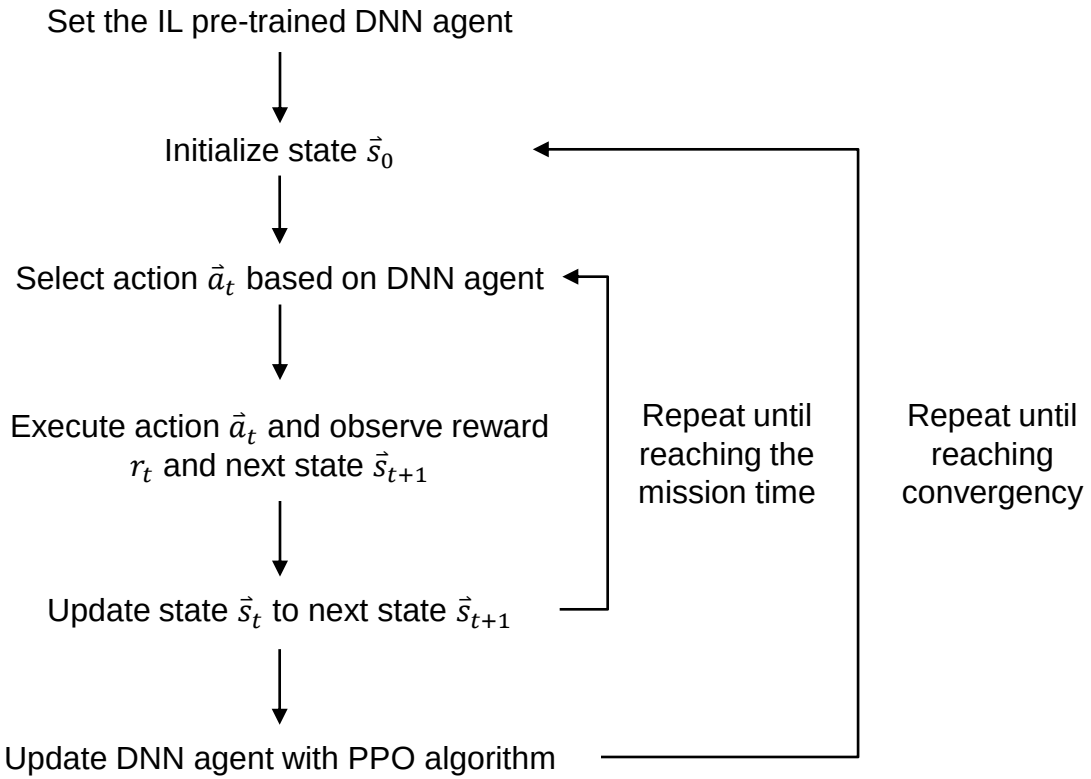
IL can help speed up the convergency when the state space is large and it is hard to find global optimum.



Profit-driven-reward Reinforcement Learning

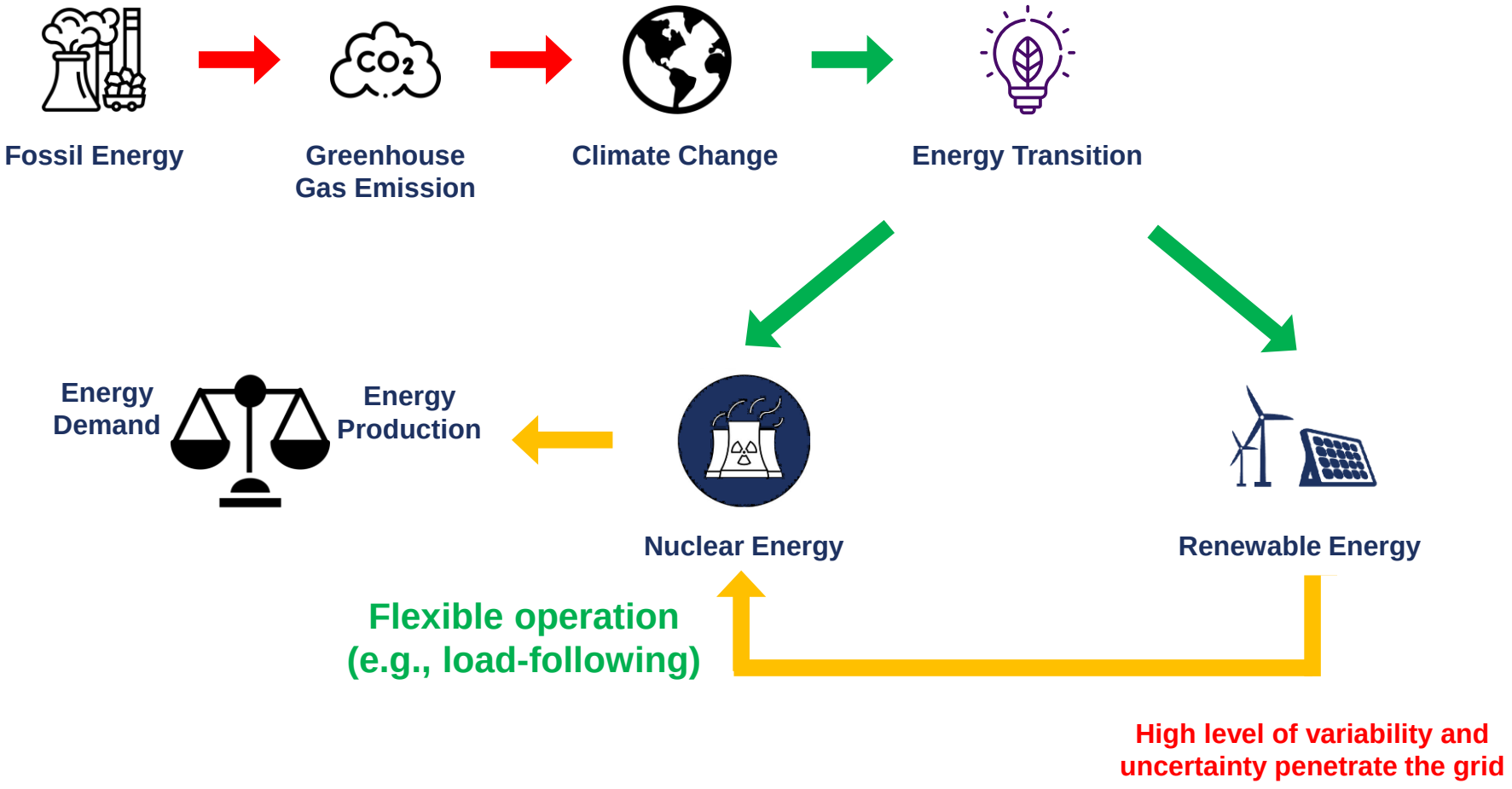
Approach: Proximal Policy Optimization (PPO)

Proximal Policy Optimization (PPO) is a popular actor-critic RL algorithm, which aims at stabilizing the policy optimization by optimizing a surrogate objective function that is a compromise between the current policy and a new candidate policy, and constraining/clipping the gradient updates, in the attempt to monotonically improve the policy.



Application: Cyber-Physical Energy Systems (CPESs)

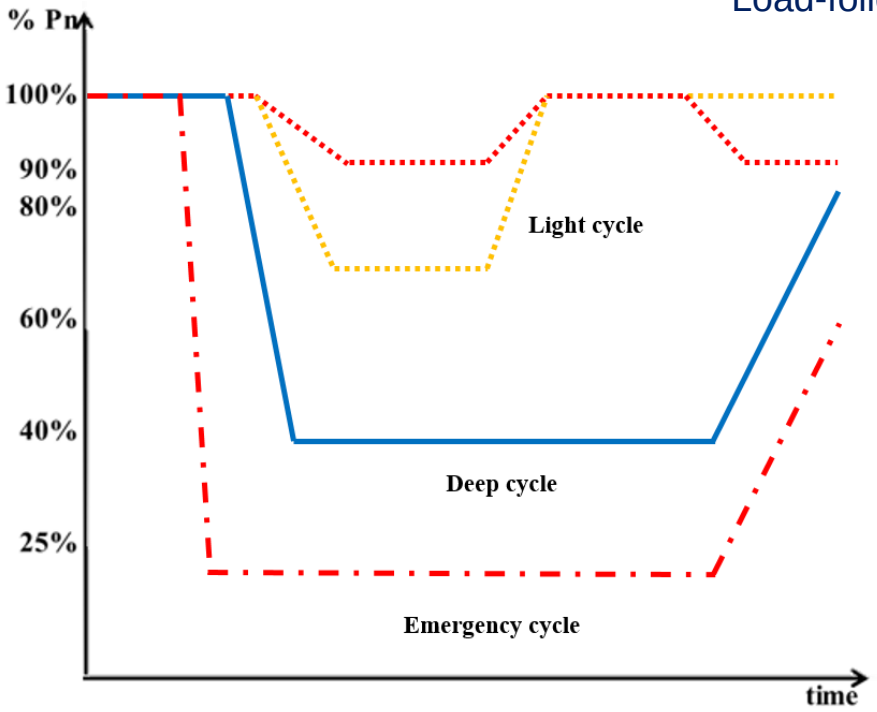
Context: The Flexibility of Energy Production



Context : CPES Flexible Operation

Flexible operation: The ability of a plant to adjust its power output to match fluctuations in electricity demand, while maintaining safety and efficiency (load-following includes three phases: a power decrease, a low power duration and a power ramp).

Load-following scheme



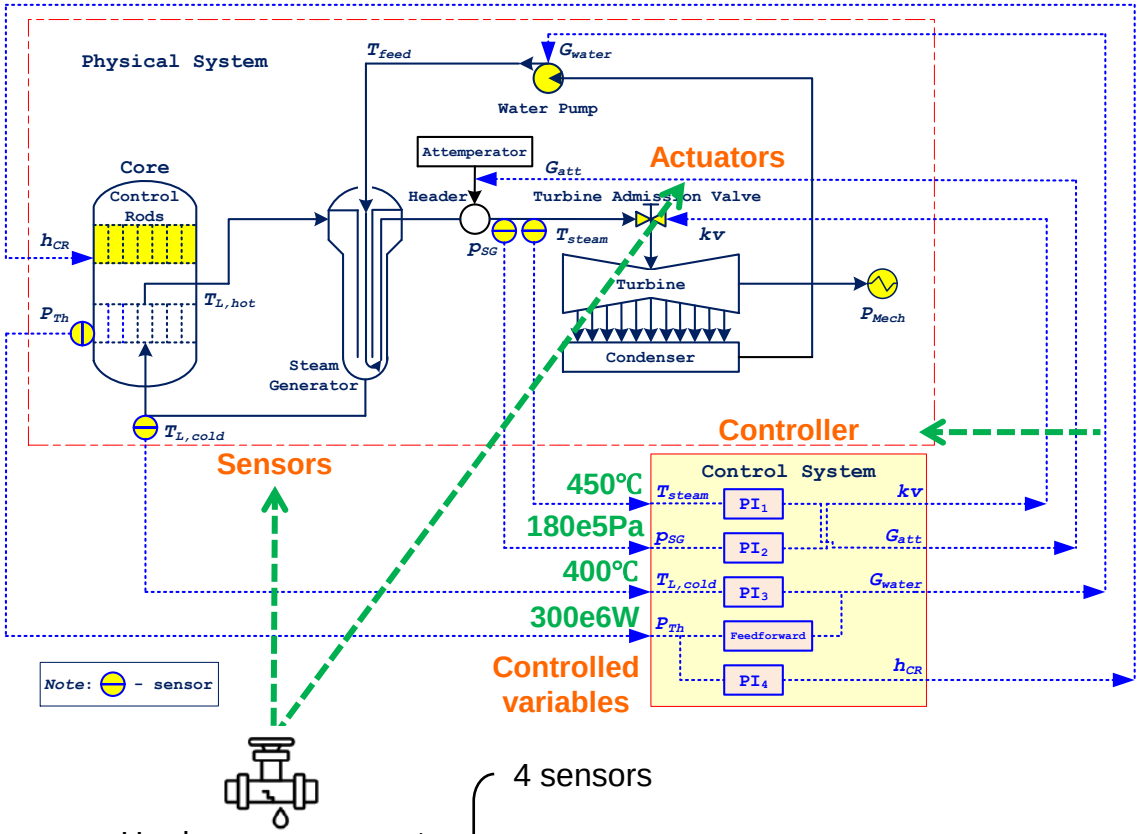
Type		Load Cycle (%)	Maximum Number of Load Cycles	Probability
Normal	Light cycle	100-90-100	100,000	0.163
		100-80-100	100,000	0.163
		100-60-100	15,000	0.0245
	Deep cycle	100-40-100	12,000	0.0196
Emergency		100-20-100	100	1.65e-4
Base-load		---	---	0.6297

Case Study: The ALFRED

The Advanced Lead Fast Reactor European Demonstrator (ALFRED)

ALFRED multi-loop control scheme

4 Single-Input-Single-Output control loops



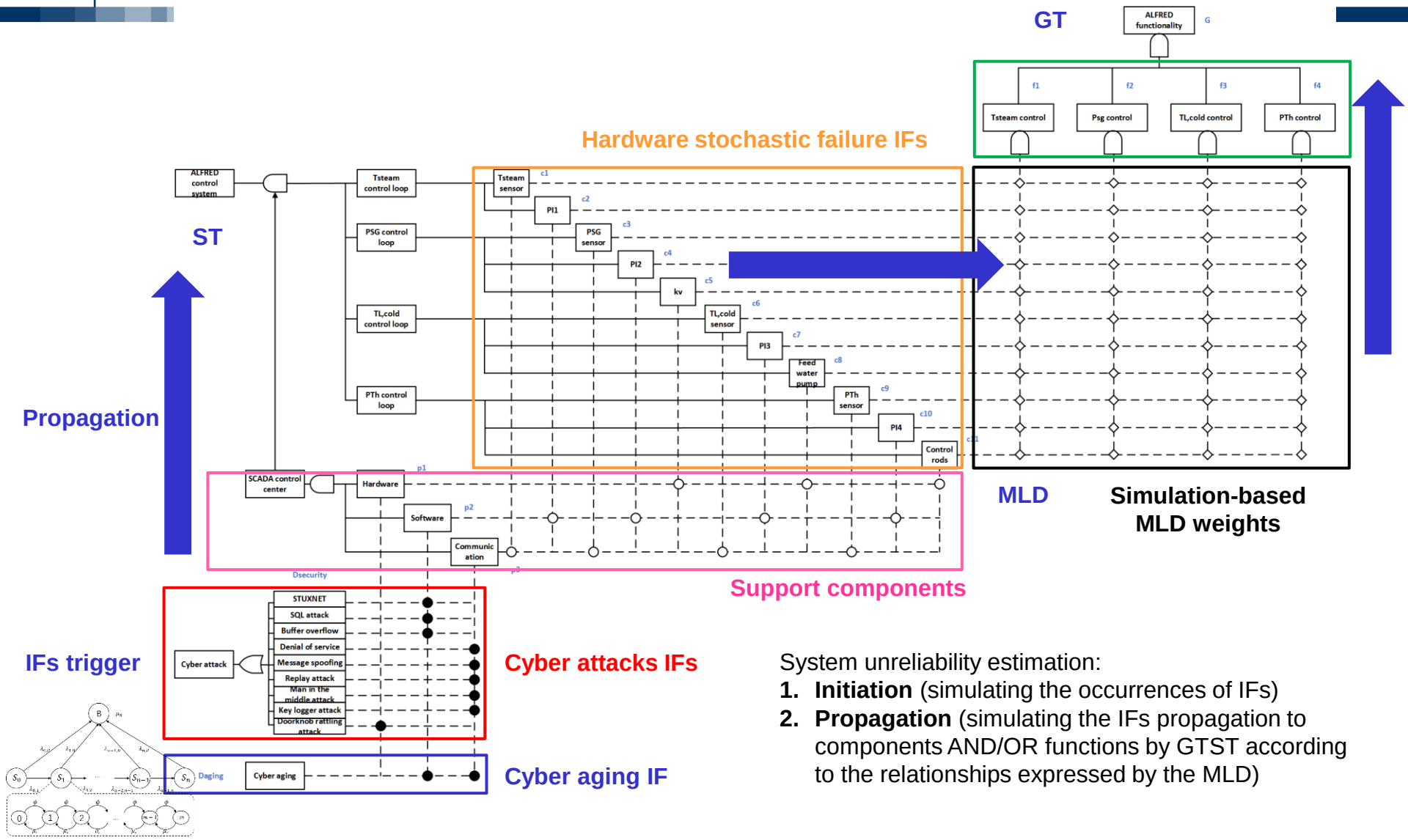
Cyber controller aging

4 PIs

Hardware component stochastic failures

- 4 sensors
- 3 actuators
  - Turbine Admission Valve
  - Feedwater pump
  - Control rod

Dynamic Reliability Assessment Framework



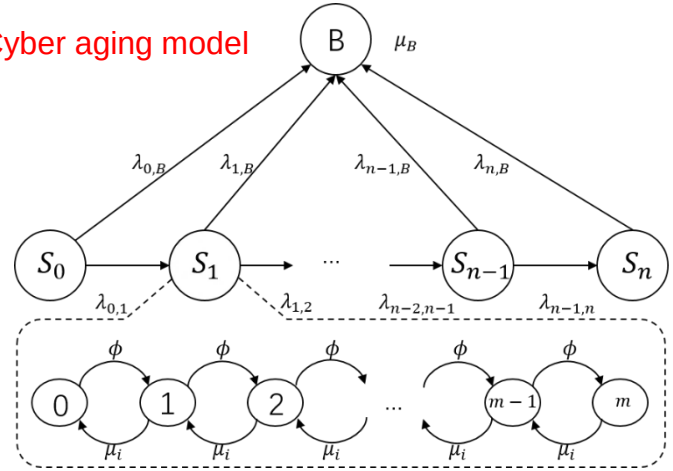
System unreliability estimation:

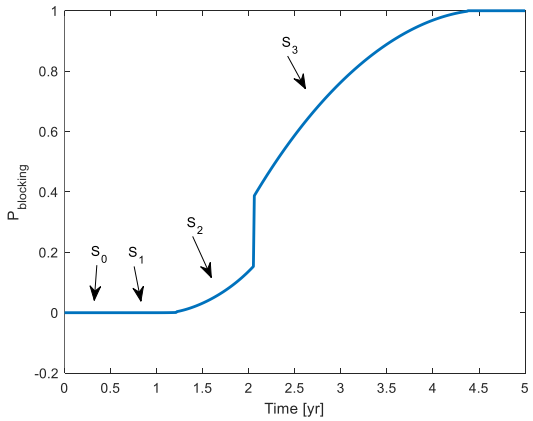
1. **Initiation** (simulating the occurrences of IFs)
2. **Propagation** (simulating the IFs propagation to components AND/OR functions by GTST according to the relationships expressed by the MLD)

[1] Di Maio F, Mascherona R, Zio E. Risk analysis of cyber-physical systems by GTST-MLD[J]. IEEE Systems Journal, 2019, 14(1): 1333-1340.

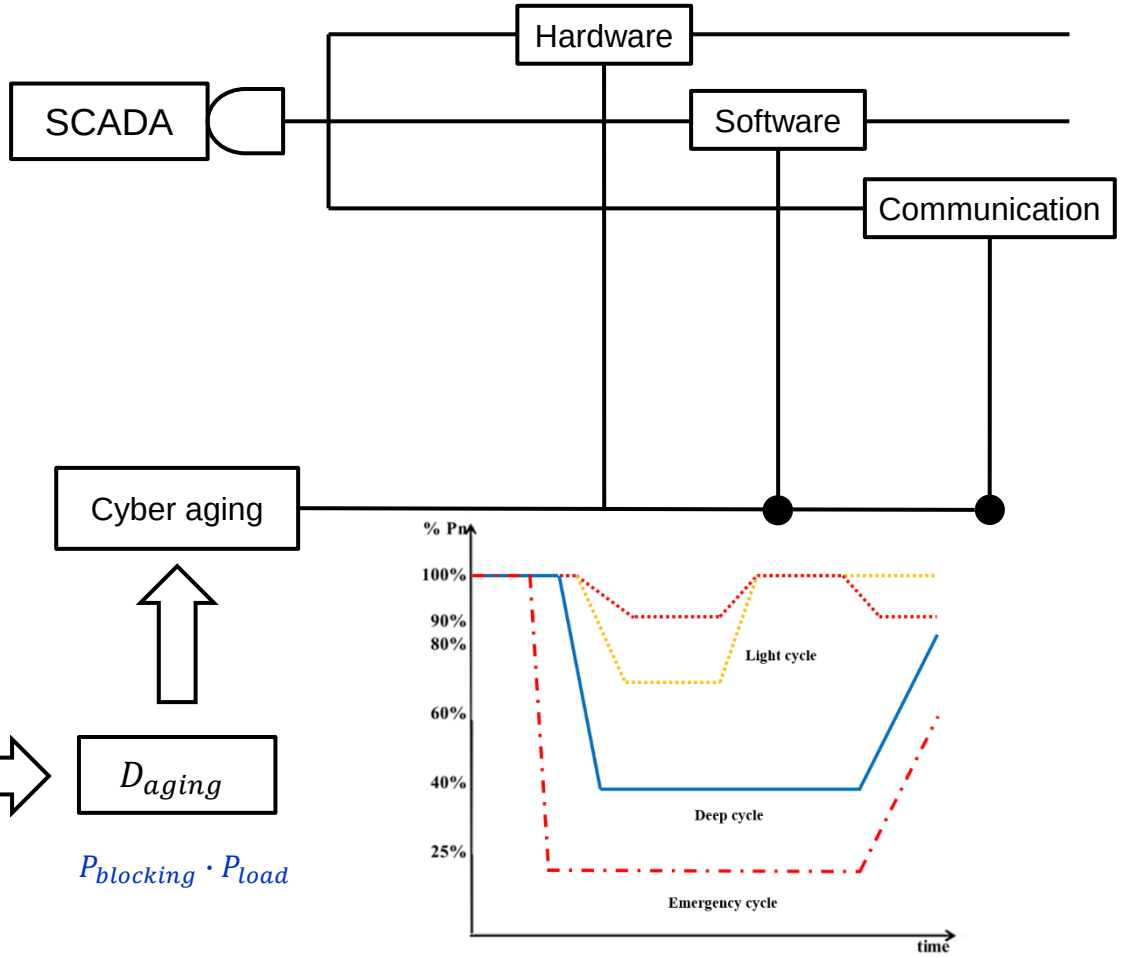
Dynamic Reliability Assessment Framework

Cyber aging model





System blocking probability $\boldsymbol{P_{blocking}}$

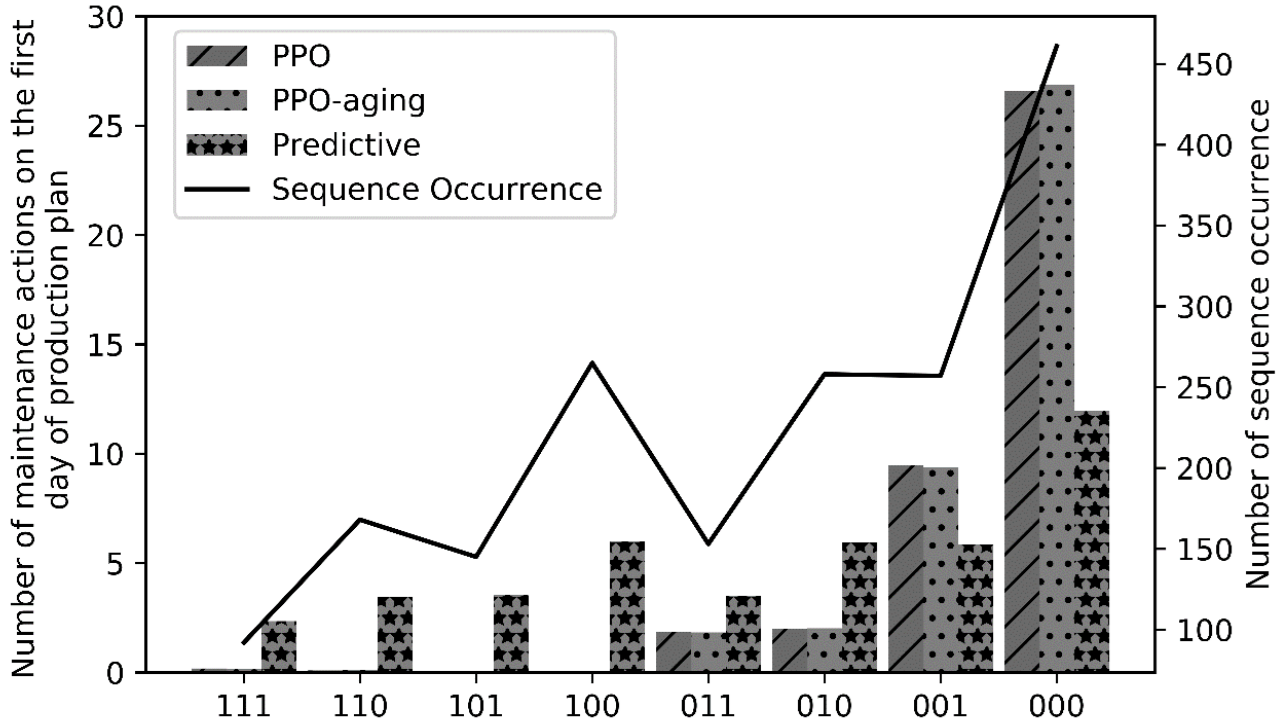


P_{load} refers to probability of load-following operation

Cyber aging IF models the controller aging under flexible load-following operation

Case Study: The ALFRED

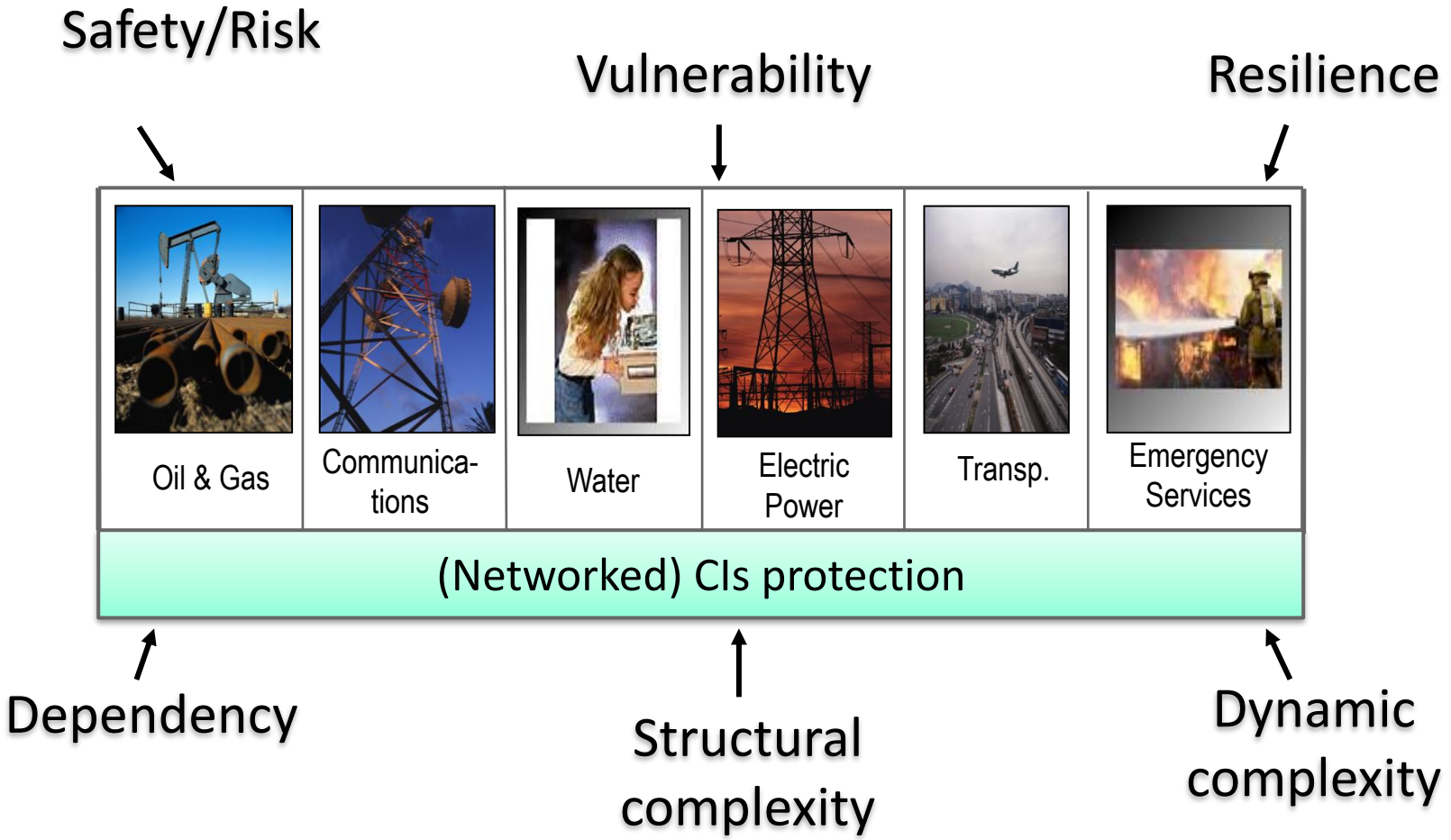
Result: Maintenance timing comparison



- Predictive strategy randomly arranges maintenance activities.
- RL strategies arrange maintenance intervention mostly on 000 and 001 sequence days to satisfy load-following operation as much as possible.

Z. Hao, F. Di Maio, E. Zio, “A Sequential Decision Problem Formulation and Deep Reinforcement Learning Solution of the Optimization of O&M of Cyber-Physical Energy Systems (CPESs) for Reliable and Safe Power Production and Supply”, Reliability Engineering & System Safety, Vol. 235, 109231, 2023

Protection and resilience of critical infrastructures: scientific and technical issues



Resilience of critical infrastructures