

# Proposed Formalization Framework

Below is a proposed starting point for formalizing definitions, metrics, criteria, and parameters. The approach is deliberately domain-agnostic, focusing on abstract representations and mathematical formalisms that can be adapted to various fields. Subsequent iterations can tailor these definitions and metrics to specific applications (e.g., infrastructure, cybersecurity, environmental systems).

## 1 System Representation

### Definition (System):

A system  $S$  is defined as a set of components  $C$ , a set of interactions  $I$ , and a set of external parameters or conditions  $E$ . Formally,

$$S = (C, I, E)$$

- **Components ( $C$ ):** Individual units (hardware elements, nodes in a network, financial instruments, etc.) that constitute the system. Each component  $c_i \in C$  may have states represented by a state variable  $x_i(t)$  at time  $t$ .
- **Interactions ( $I$ ):** Defined relationships or functional mappings that describe how the state of one component influences another. Interactions can be represented as functions  $f_{ij}$  or as edges in a graph structure  $G$ .
- **External Conditions ( $E$ ):** Environmental parameters, boundary conditions, or other contextual factors that influence system behavior. Examples include weather conditions for infrastructure systems, regulatory environments for financial systems, or threat landscapes for cybersecurity systems.

**State Representation:** The global system state at time  $t$  can be represented as a vector:

$$X(t) = [x_1(t), x_2(t), \dots, x_n(t)]$$

## 2 Failure Criteria and Distributed Failures

### Definition (Component Failure):

A component  $c_i$  is considered to have failed at time  $t$  if its state  $x_i(t)$  crosses a predefined failure threshold. Let  $F_i$  be a threshold or condition that determines failure (e.g.,  $x_i(t) > F_i$ )

or  $x_i(t) \in F_i$  where  $F_i$  is a failure set). Formally:

$$\text{Fail}(c_i, t) = \begin{cases} 1 & \text{if } x_i(t) \in F_i, \\ 0 & \text{otherwise.} \end{cases}$$

### Definition (Distributed Failure):

A Distributed Failure (DF) is not a single isolated component failure, but a set of failures occurring within a specified spatial, structural, or temporal window that collectively degrades overall system performance. Let  $T$  be a time horizon and  $k$  a minimum scale parameter (e.g., fraction of components or a minimum number of distinct subsystem failures). Formally, a Distributed Failure event occurs if there exists a subset  $C' \subseteq C$  such that:

$$|C'| \geq k \quad \text{and} \quad \text{Fail}(c_i, t) = 1 \text{ for multiple } c_i \in C'.$$

## 3 Frequency and Impact Metrics

### Frequency Metrics:

- **Component-Level Failure Probability ( $p_i$ ):** The probability that component  $c_i$  fails within a given period:

$$p_i = P(\text{Fail}(c_i, t) = 1 \text{ for some } t \in [0, T]).$$

- **Distributed Failure Frequency ( $\lambda_{DF}$ ):** The expected number of Distributed Failures over a specified time horizon. If  $N_{DF}(T)$  is the random variable denoting the number of DF events in the interval  $[0, T]$ , then:

$$\lambda_{DF} = \frac{E[N_{DF}(T)]}{T}.$$

Alternatively, define:

$$P_{DF}(T) = P(\text{At least one Distributed Failure in } [0, T]).$$

### Impact Metrics:

- **Component Impact ( $I_i$ ):** A scalar or vector representing the cost, damage, or severity associated with the failure of component  $c_i$ . Examples include economic loss (\$), downtime (hours), safety impact (injuries/fatalities), or environmental damage:

$$I_i : \text{Fail}(c_i, t) = 1 \implies I_i \in \mathbb{R}^+.$$

- **System Impact Function ( $\mathcal{I}$ ):** When a set of components  $C'$  fail, the system-level impact is an aggregation function  $\mathcal{I}$ , which could be additive, multiplicative, or more complex (e.g., nonlinear, based on network centrality of components):

$$\mathcal{I}(C', t) = F(\{I_i : c_i \in C'\}).$$

- **High-Impact Threshold ( $\Gamma$ ):** A parameter distinguishing “high impact” from “low impact” scenarios. For example:

$$\text{High Impact if } \mathcal{I}(C', t) > \Gamma.$$

## 4 Low Frequency High Impact Failures (LFHI)

### Definition (LFHI Event):

A Low Frequency High Impact (LFHI) event is a DF event that meets both a low-frequency and high-impact criterion. Let  $\delta_f$  be a frequency threshold and  $\Gamma$  the impact threshold:

- **Low Frequency:** The probability of occurrence in a given time frame is less than  $\delta_f$ :

$$P_{DF}(T) < \delta_f.$$

- **High Impact:** The system impact exceeds  $\Gamma$ :

$$\mathcal{I}(C', t) > \Gamma.$$

## 5 Distributed vs. Cascading Failures

### Distinguishing Distributed and Cascading Failures:

- **Cascading Failure:** A failure where one component’s failure directly triggers subsequent failures in a chain-like manner, creating a temporal sequence.
- **Distributed Failure:** May not follow a single causal chain. Failures arise simultaneously or from different triggers without a clear linear sequence.

## 6 Parameters and Adaptability

### Parameter Categories:

- Structural Parameters: Number of components ( $n$ ), graph structure, interaction strengths ( $W_{ij}$ ).
- Failure Thresholds:  $F_i$  for components,  $\Gamma$  for high impact.
- Time Scales: Analysis horizon ( $T$ ), time-step granularity ( $\Delta t$ ).
- Probability Distributions: For component states and failure triggers.
- Uncertainty Parameters: Variance, confidence intervals.

## 7 Summary

This initial formalization creates a foundation for probability models, resilience indicators, and optimization frameworks. Future work involves refining functional forms for  $\mathcal{I}$ , distributions for component failures, and algorithms for parameter estimation.