ETH zürich

Complexity theory and centrality measures

Giovanni Sansavini
Reliability and Risk Engineering Lab
ETH Zurich

DEMO UPCARMA

European Commission

Horizon 2020
European Union funding
for Research & Innovation

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Bundesamt für Energie BFE
Swiss Federal Office of Energy SFOE

SCCER – FURIES
Shaping the FUtuRe Swiss Electrical InfraStructure

KTI/CTI
DIE FÖRDERAGENTUR FÜR INNOVATION
THE INNOVATION PROMOTION AGENCY

FUTURE RESILIENT SYSTEMS 未来韧性系统

FNSNF
Fonds national suisse
Schweizerischer Nationalfonds
Fondo nazionale svizzero
Swiss National Science Foundation

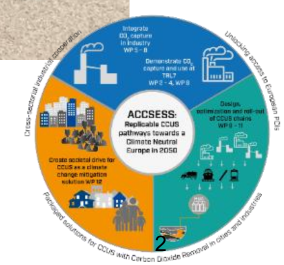
Energy Science Center

ETH RISKCENTER

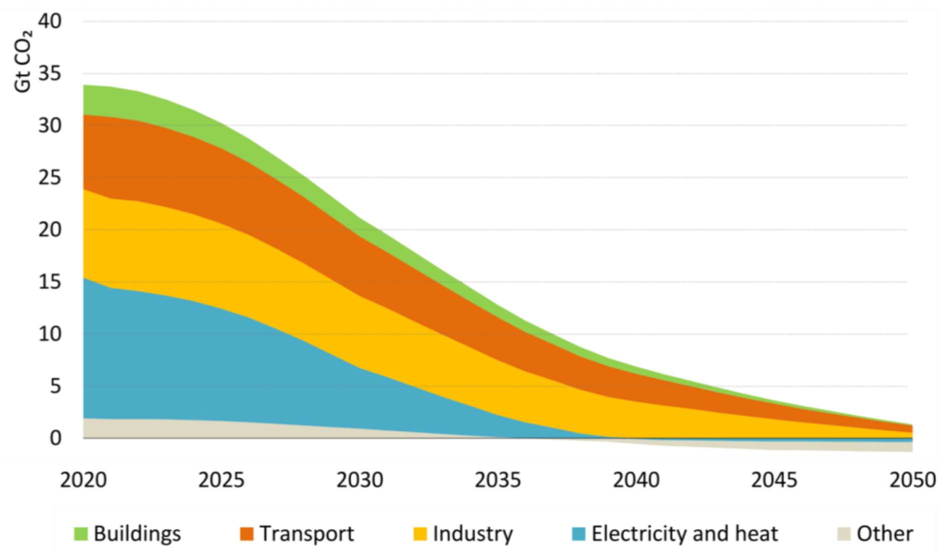
sweet swiss energy research for the energy transition

Swiss Re

Innosuisse

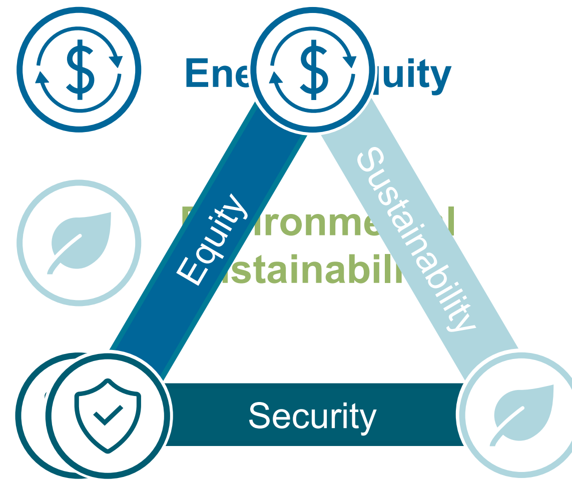


Climate Change Targets Require Transition to Sustainable Energy Systems

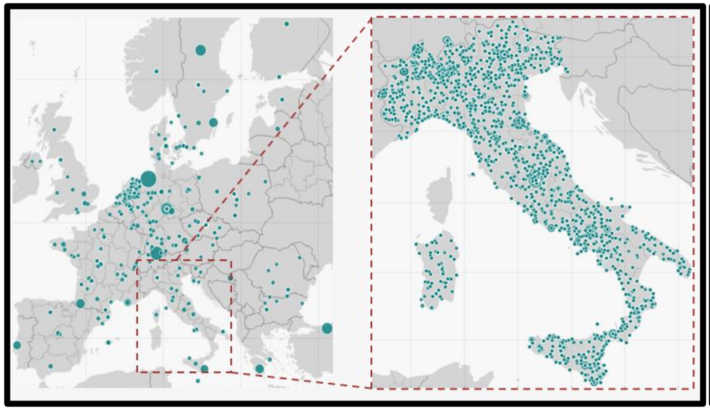


Projections for emission reductions until 2050

Source: IEA, 2021, Net Zero by 2050 – A Roadmap for the Global Energy Sector

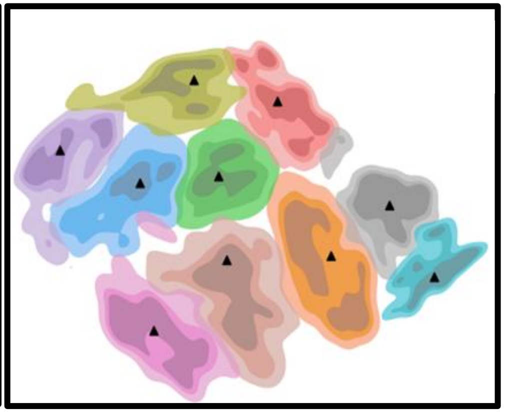


RRE
Reliability & Risk Engineering



Stankovski, Gjorgiev, Locher, Sansavini, **Joule** 2023

Zapparoli, Oneto, ..., Hug, Sansavini, **Scientific Data** 2025

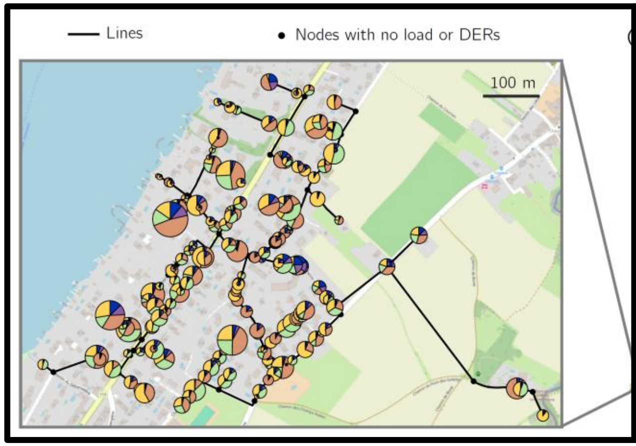


Oneto, Gjorgiev, Sansavini, **Pattern Recognition**, under review

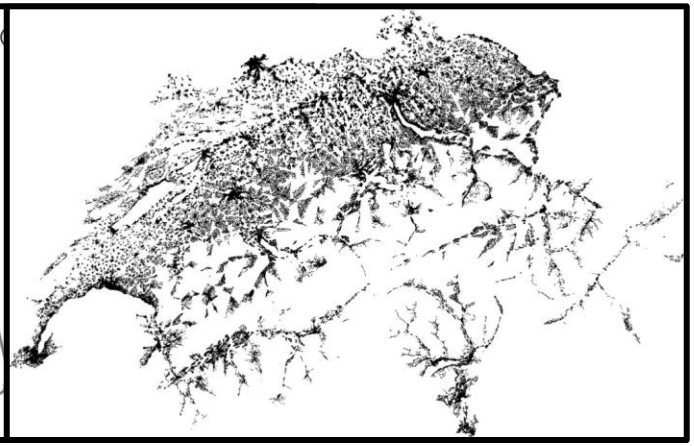
Empirical, theoretical and **computational** research activities in the fields of **Risk Analysis**, **Network Theory,** and **Operations Research** towards **safe transition to renewable and sustainable energy systems**

Safety challenges across scales from component to continent

Das, Gjorgiev, Sansavini, **Engineering Applications of Artificial Intelligence** 2025

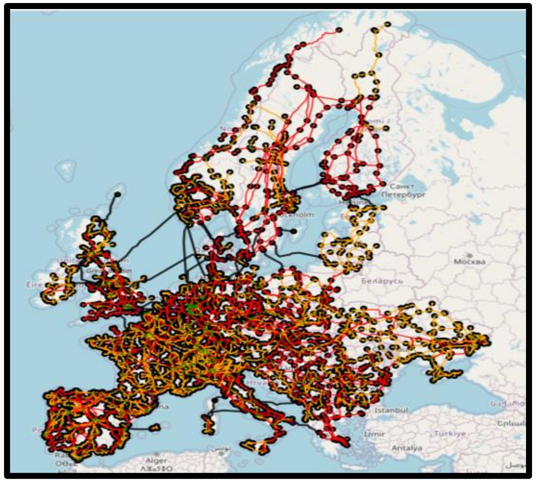


Zapparoli, Oneto, ..., Hug, Sansavini, **Scientific Data** 2025



Oneto, Gjorgiev, Tettamanti, Sansavini, **Sustainable Energy, Grids and Networks** 2025

Haegel, **MSc Thesis**, 2021

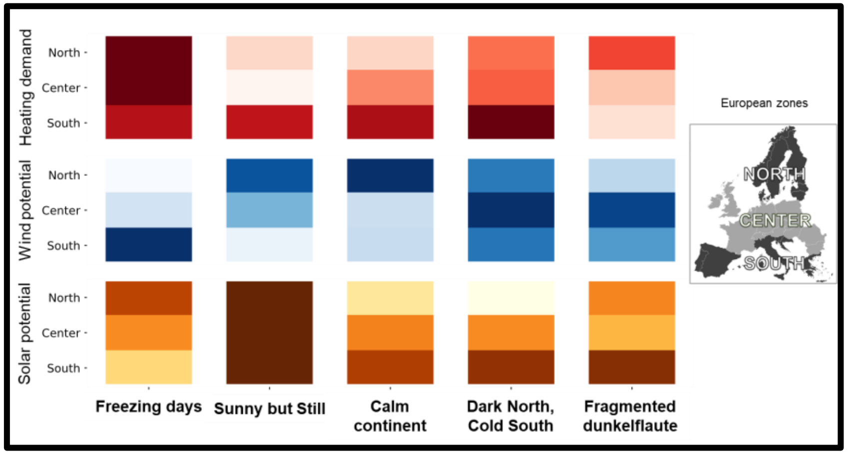


SFOE SWEET RECIPE project 2025

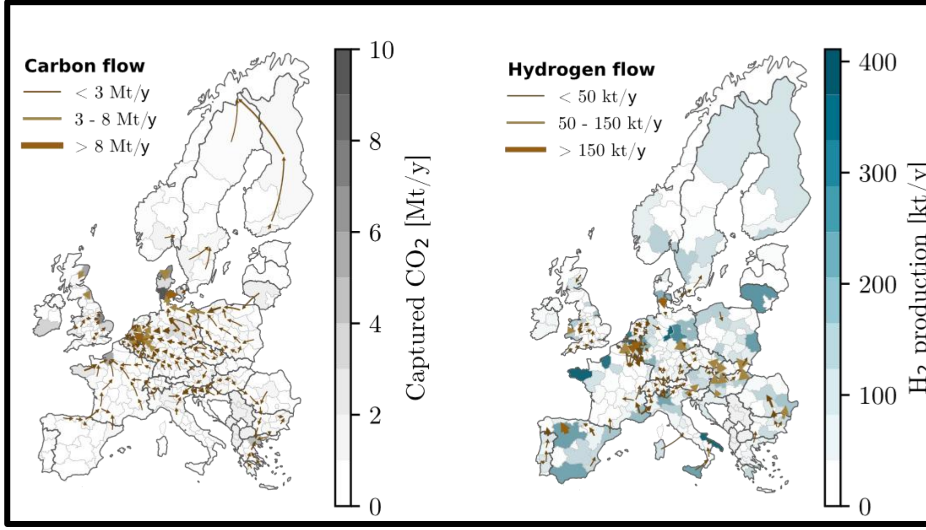
ETH zürich

Resilient transition of the energy infrastructure

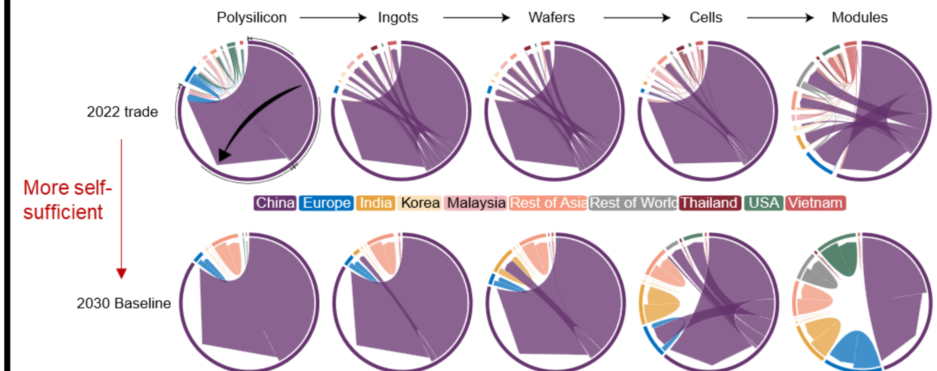
De Marco, Mannhardt, Oneto, Sansavini, **Advances in Applied Energy** 2025



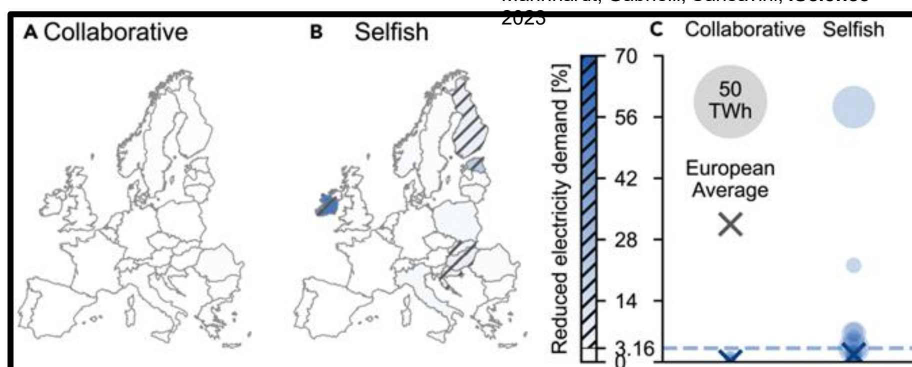
Ganter, Gabrielli, Goericke, Sansavini, **Environmental Science & Technology** 2025



Cui, Lonergan, Sansavini, **Nature Communications** 2025

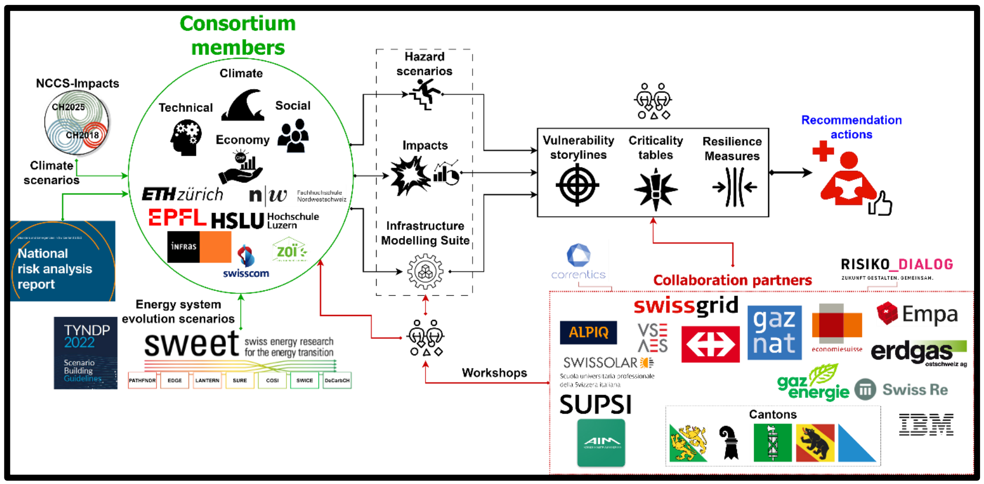


Mannhardt, Gabrielli, Sansavini, **iScience** 2023

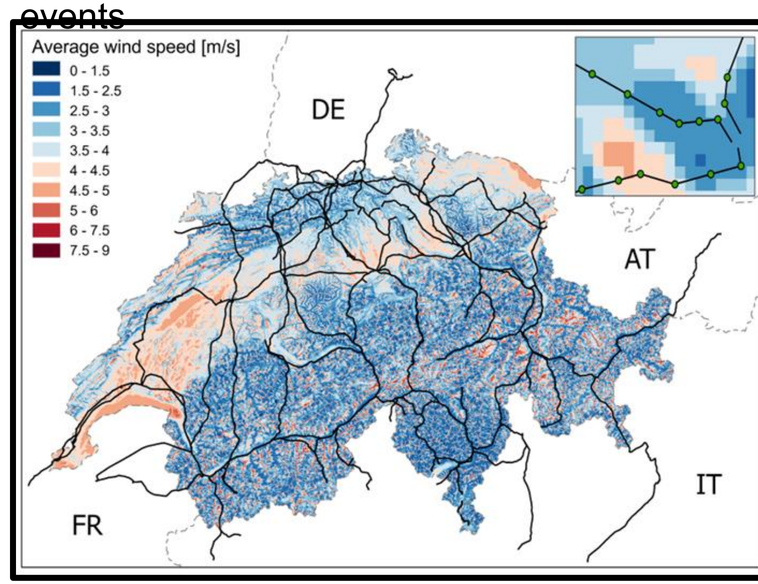


Resilient transition of the energy infrastructure

RECIPE – Resilient Infrastructure for the Swiss Energy Transition



Design changes under extreme weather events



SFOE SWEET RECIPE project
2025

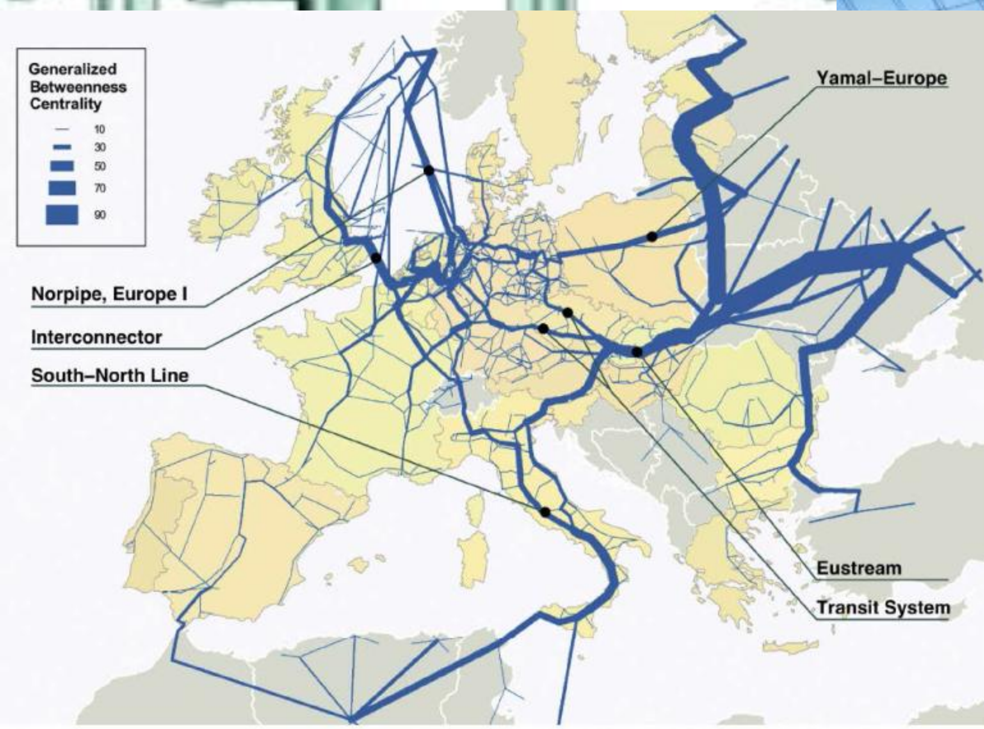
The Big Challenge Facing Risk Analysis

Dependencies Interdependencies	+	Uncertainties Unknown unknowns

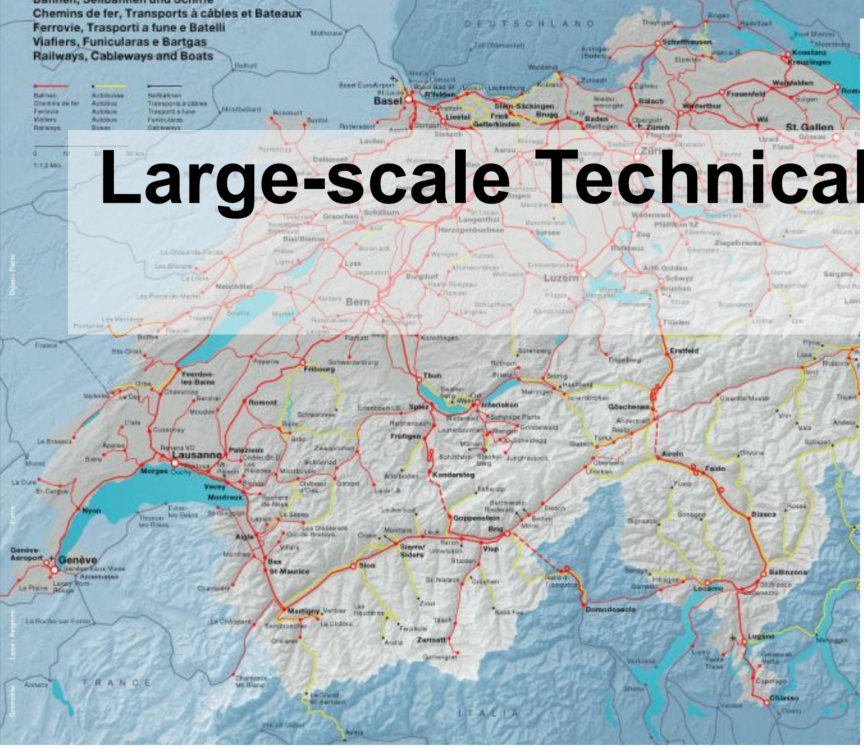
- The systems become more integrated and interconnected
- The consequences of some events have not been foreseen

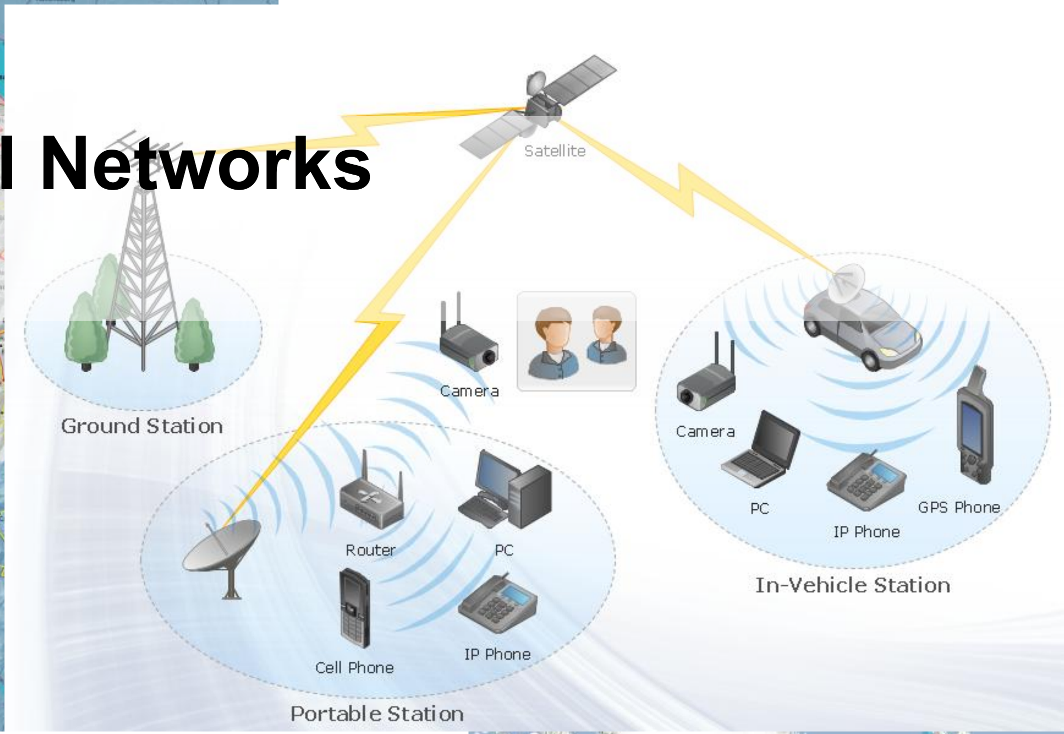

EVERGREEN

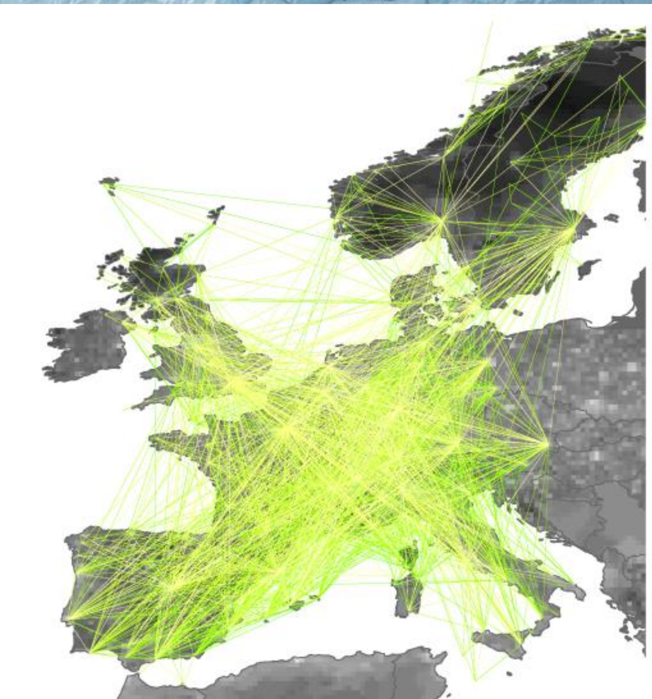
Critical Infrastructures

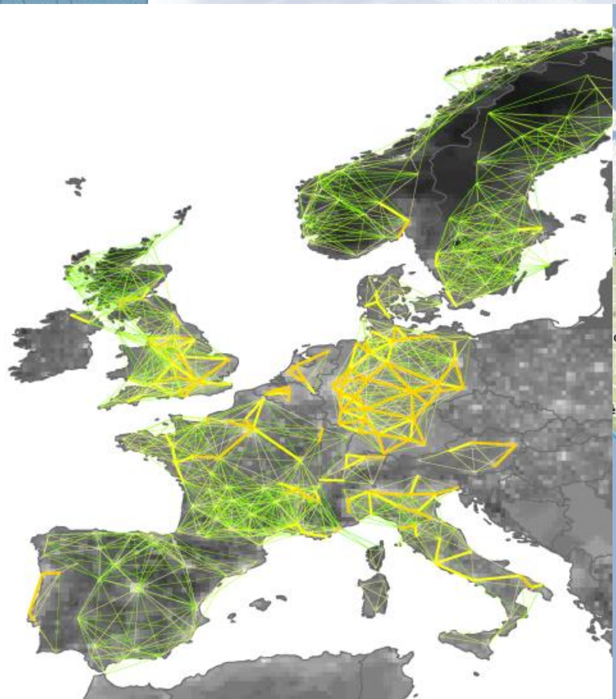


Large-scale Technical Networks









Introduction and Problem Description (I)

- Infrastructure systems provide essential goods and services to the industrialized society including **transport**, **water**, **communication** and **energy**
- A disruption or malfunction often has a significant economic impact and potentially **propagates** to other systems due to **interdependencies**
- Wide-area breakdowns of such large-scale engineering networks are often **caused by technical equipment failures** and their **coincidence in time** which eventually result in a series of fast cascading component outages
- Illustrative examples are a number of large **electric power blackouts** and **near-misses** as has been increasingly experienced in the last few years

Introduction and Problem Description (II)

How can we quantify the reliability of infrastructures and **assess the vulnerability** to such large-area breakdowns?

Basic problem: Infrastructures are highly complex and interdependent systems, consisting of an **enormous number of components.** Classic reliability analysis methods are limited due to the **state space explosion**

Example: Consider a system of N=20 components with up state and down state. A “state enumeration approach”, such as a complete Markovian chain would have to consider *$2^N = 2^{20} \sim 10^6$ system states!*

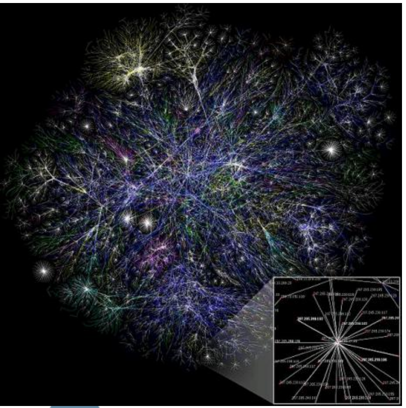
Approach 1: *Simulate the systems realistically by means of extensive modeling methods, including physical laws and operational dynamics.*

Introduction and Problem Description (III)

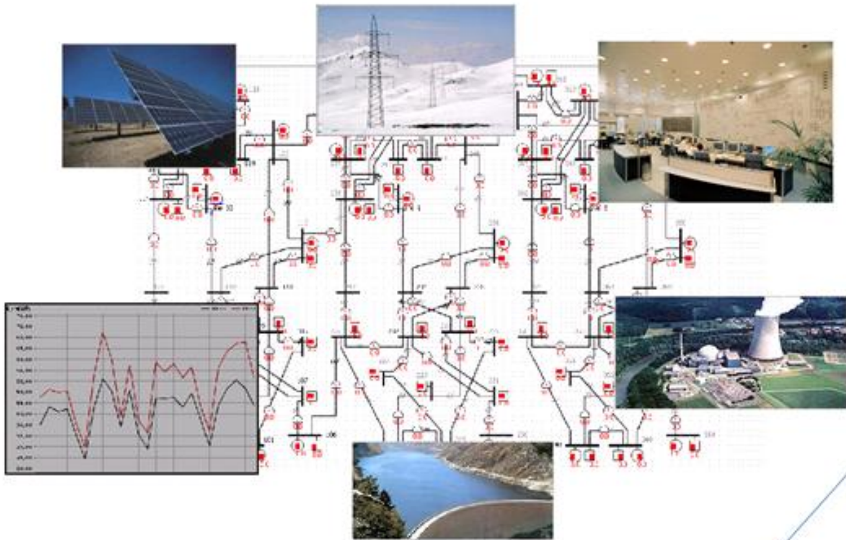
Approach 2: *Use highly simplified models in order to understand the* ***basic mechanisms*** *leading to infrastructure breakdowns. In this respect* ***network theory*** *allows for gaining valuable qualitative knowledge about the basic functioning of infrastructure systems, being networks in nature.*

However, due to its highly simplifying approach, network theory cannot replace more detailed reliability analysis methods.

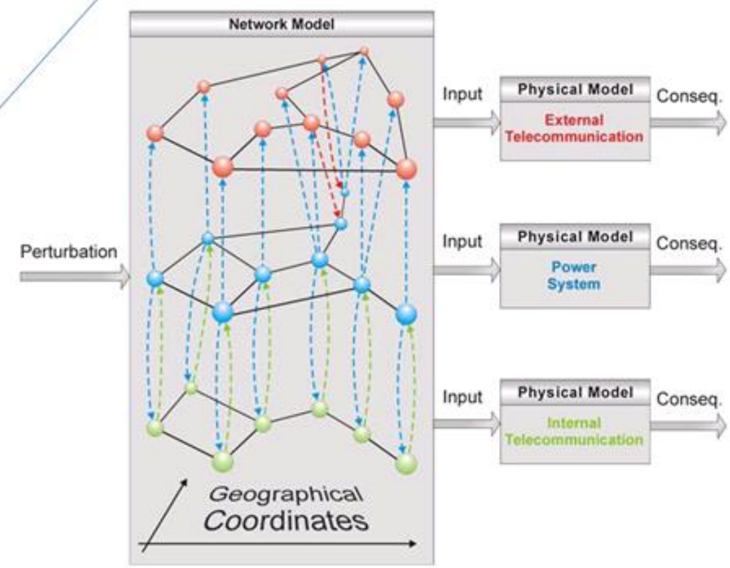
It rather serves as a first **screening analysis**, whereas the findings, e.g. robustness of topology, may serve as an *input* for detailed reliability studies.



Real Networks



Network Models



Many simplifying assumptions

What you will learn today…

- Know the characteristics and representation of complex networks by Complex Network Theory
  - Global and Local Properties
  - Exponential Networks vs. Scale-Free Networks
- How to perform the network static structural analysis
  - Unweighted
  - Weighted
- Compute the system vulnerability to element removal
- Identify the centrality of elements in the structure
  - Centrality measures, bottlenecks
- How to Model Cascading Failures Propagation in Network Systems
  - How to Identify Cascade-safe Operating Margins

Complex Graph/Network - Definitions

- A complex graph contains many different subgraphs
- A complex graph is a graph whose structure is irregular, complex and dynamically evolving in time
- All complex graphs contain a medium number of links: both very sparse graphs and nearly fully connected graphs are not complex
- Complex graphs have non-trivial topological features: heavy tail in the degree distribution, a high clustering coefficient, community structure, and hierarchical structure
- As opposed to regular lattices

Quantifying Complexity of Different Networks

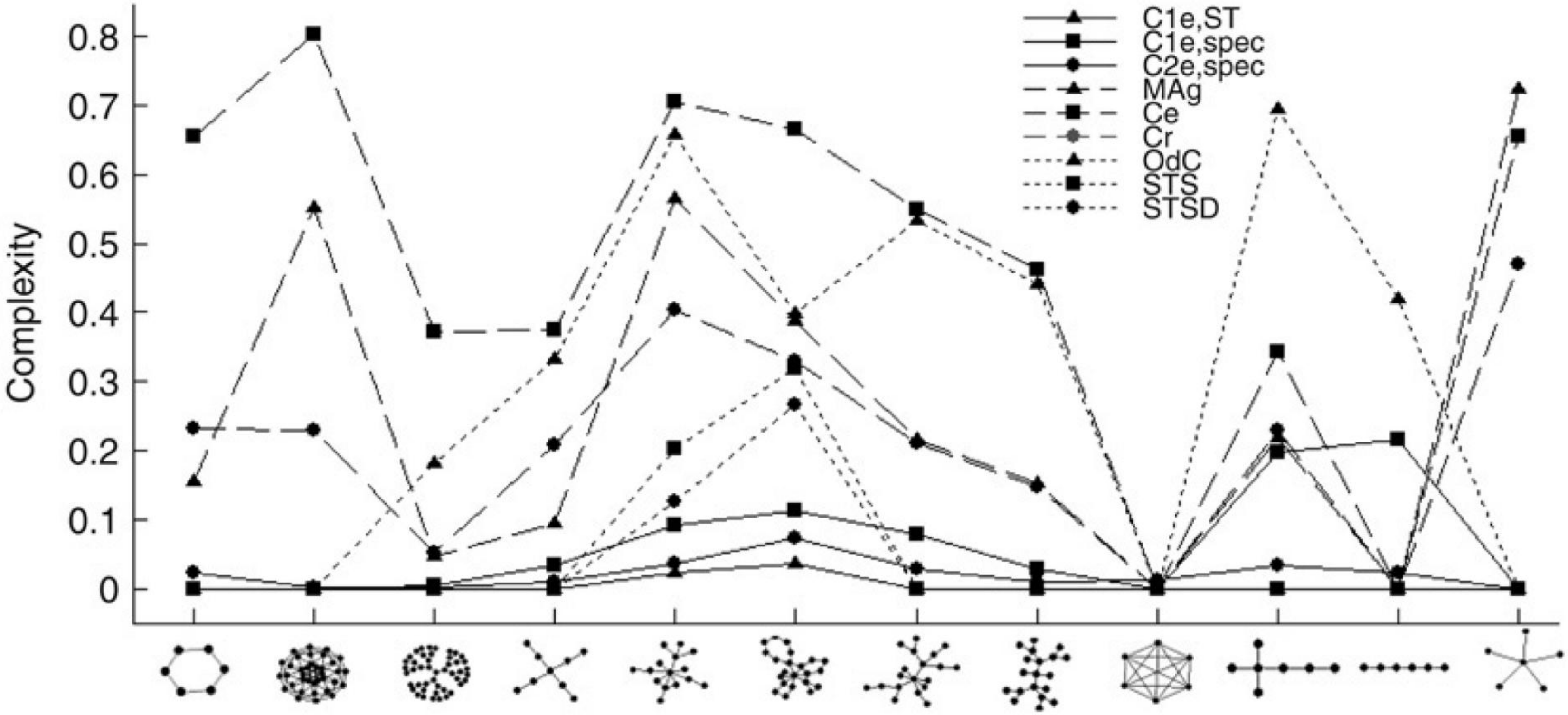


Fig. 4. Complexities of different test graphs.

T. Wilhelm, J. Kim (2008). "What is a complex graph?". *Physica A* **387**: 2637–2652.

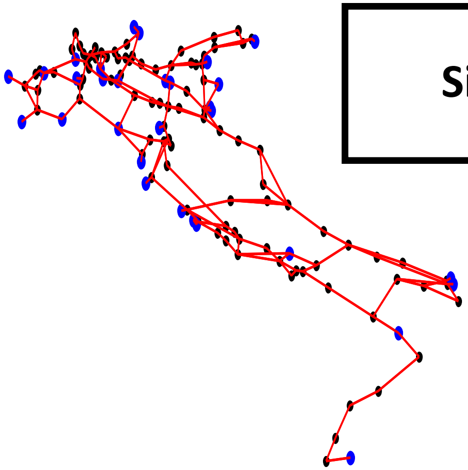
Different Types of Analysis

Network static structure analysis

- Topological analysis
- Weighted analysis

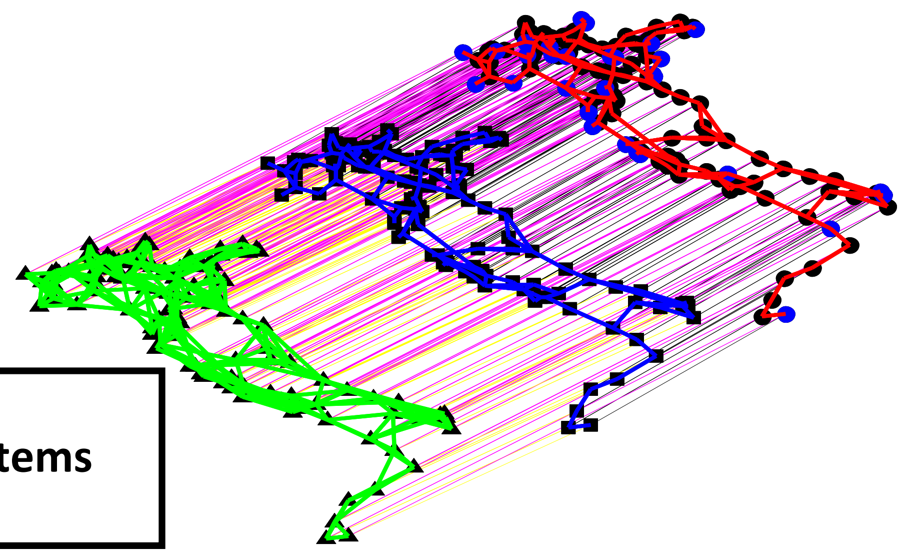
Dynamic modeling of cascading failures propagation in network systems

- Identification of cascade-safe operating margins
- Criticality indicators in failure cascade processes



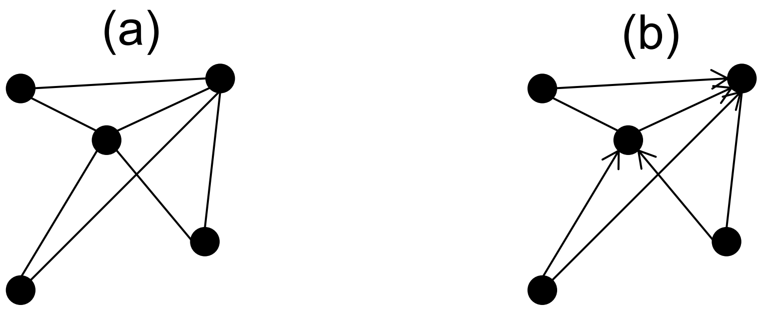
Single system

Systems of systems



Network Characteristics: Some Basic Notations

- A network (or **graph**) is a set of *N* **nodes** (or **vertices** or **sites**) connected by *L* **links** (or **edges** or **bonds**)
- **G***(N,L)*: arbitrary graph of **order** *N* and **size** *L*
- Networks with undirected links are called **undirected networks** (a), those with directed links are called **directed networks** (b)



- The total number of connections of a node *i* to its nearest neighboring nodes is called its **degree** k_i

Network Representation: Adjacency Matrix

- The adjacency matrix ***A*** provides a complete description of a network
- Consider a network with N nodes labelled by their index i (i=1,…,N).
  Then the adjacency matrix is a N x N matrix with elements a_{ij} :

if the network is undirected:

$a_{ij} = a_{ji}$, $\quad$ $a_{ij} = 1$ *if there exists a link between node i and j*

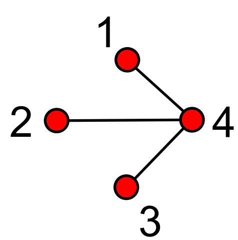
$a_{ij} = 0$ *otherwise*

if the network is directed:

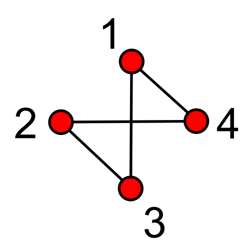
$a_{ij} \neq a_{ji}$, $\quad$ $a_{ij} = 1$ *if there exists a link leaving node i and going to node j*

$a_{ij} = 0$ *otherwise*

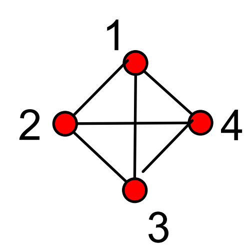
- Examples of undirected graphs:



$$\begin{array}{c} \\ \mathbf{a_1} \\ \mathbf{a_2} \\ \mathbf{a_3} \\ \mathbf{a_4} \end{array} \begin{array}{c} \mathbf{a_1}\ \mathbf{a_2}\ \mathbf{a_3}\ \mathbf{a_4} \\ \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array}$$



$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$



$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

$$k_i = \sum_j a_{ij}$$

Network Characteristics: Degree distribution

The **degree distribution** *P(k)* gives the probability that any randomly chosen vertex has degree *k.*

Poisson

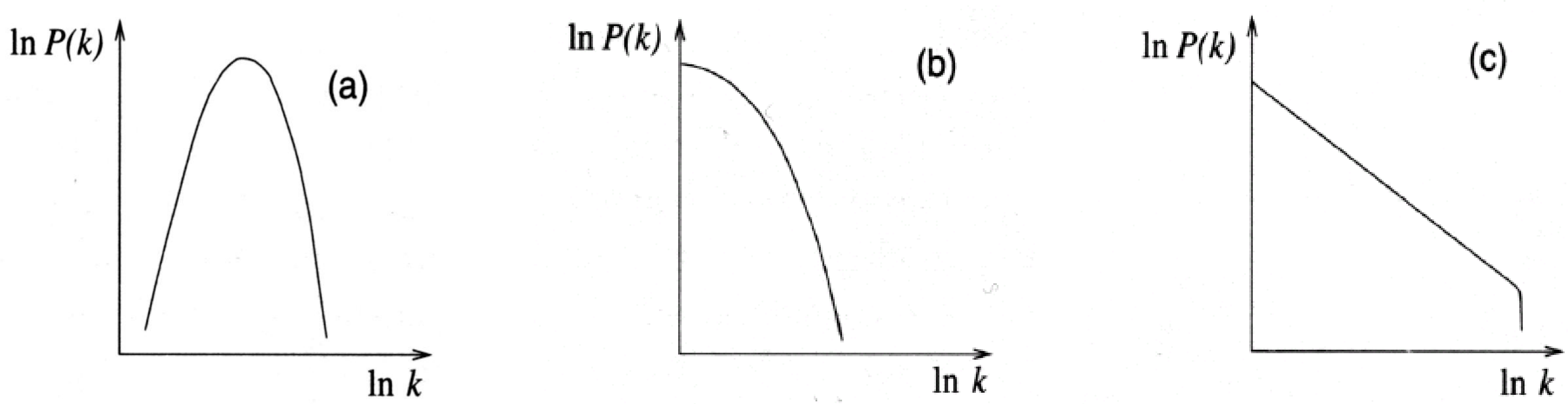
$$P(k) = \frac{e^{-\alpha}\alpha^k}{k!}, \text{ where } \alpha = \overline{k}$$

Exponential

$$P(k) \propto e^{-k/\alpha}, \text{ where } \alpha = \overline{k}$$

Power law

$$P(k) \propto k^{-\gamma}, \; k \neq 0$$



Source: Dorogovtsev, S. N. and Mendes (2003)

Examples of typical degree distributions

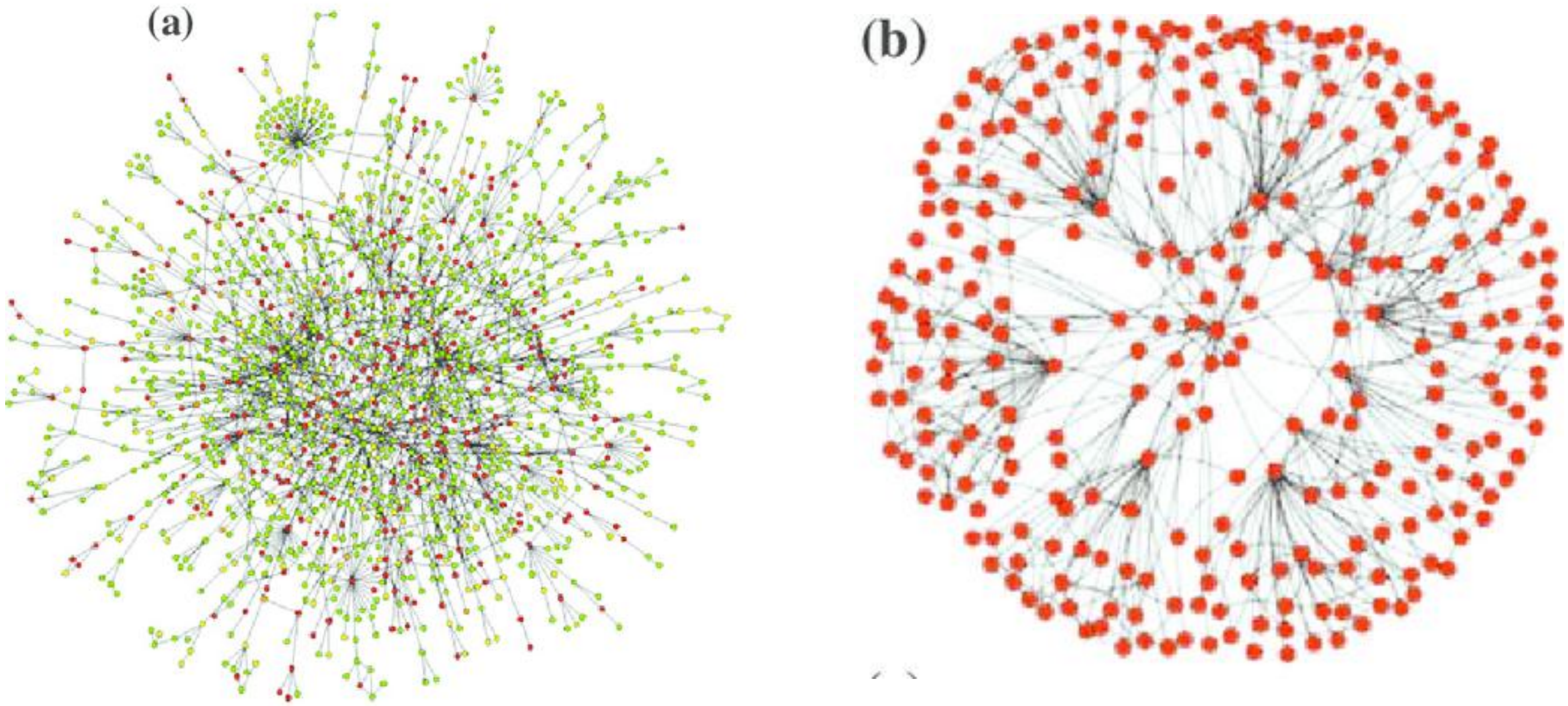


Illustration of network architectures. Left: scale-free network (**power law**), right: random graph (**Poisson**).

Source: Li et al., Web Intelligence Meets Brain Informatics (2007)

Degree distributions – WWW

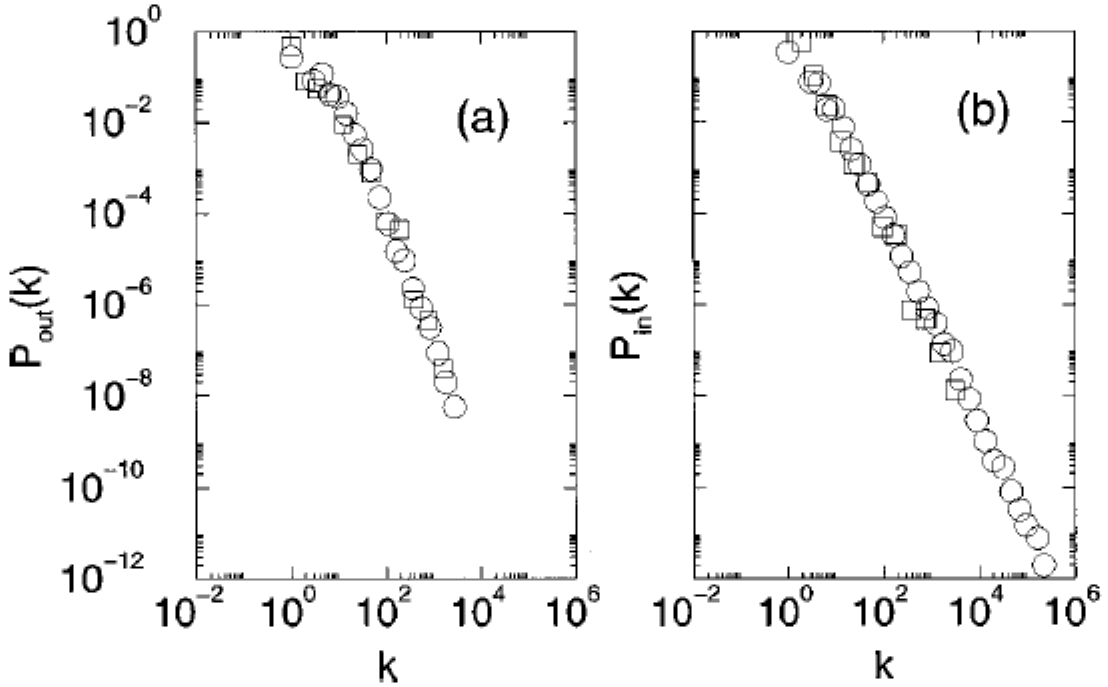
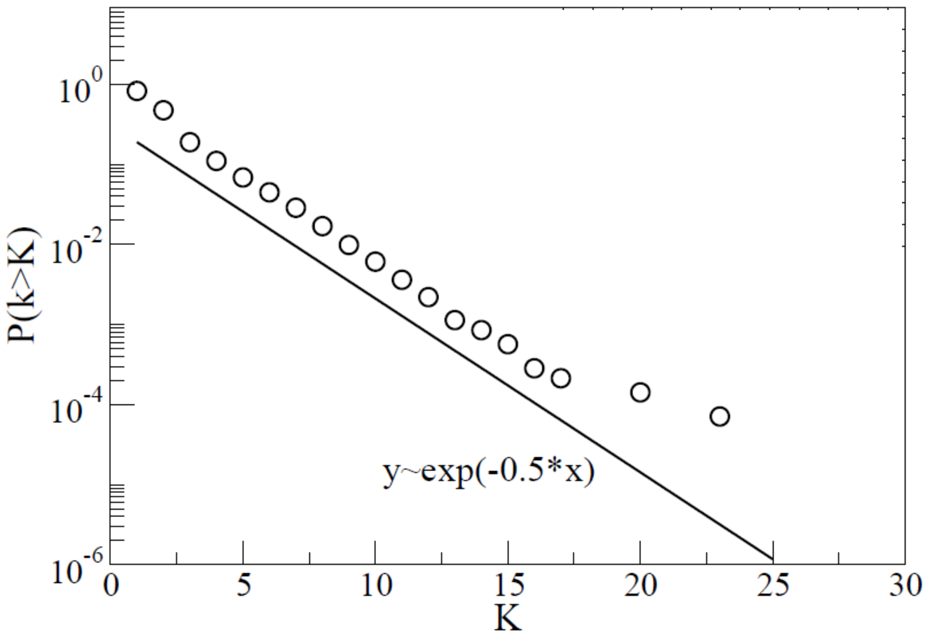


FIG. 2. Degree distribution of the World Wide Web from two different measurements: □, the 325 729-node sample of Albert *et al.* (1999); ○, the measurements of over 200 million pages by Broder *et al.* (2000); (a) degree distribution of the outgoing edges; (b) degree distribution of the incoming edges. The data have been binned logarithmically to reduce noise.

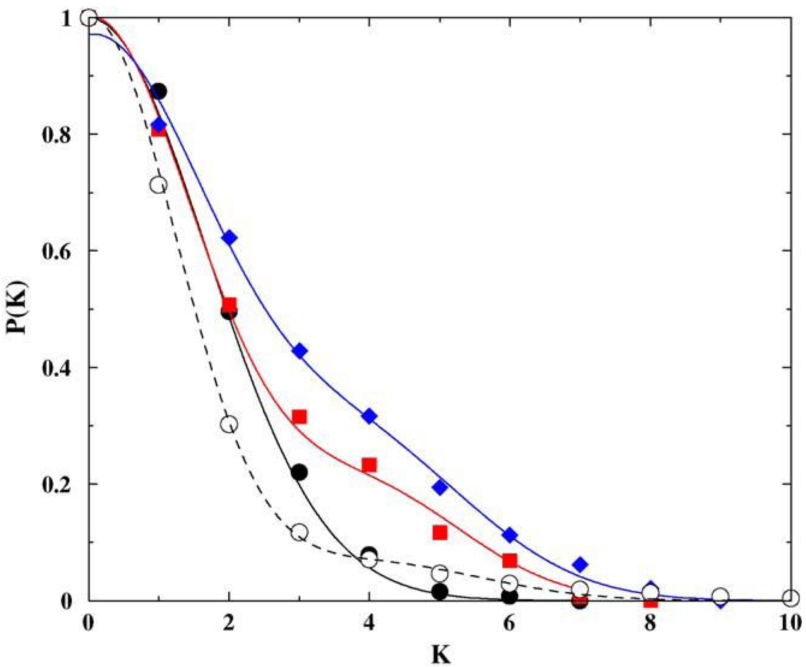
Source: Albert, R., Barabàsi, A.: Statistical Mechanics of Complex Networks, Rev. Mod. Phys., Vol. 74 (2002)

Degree distributions – Electric Power Systems



Cumulative distribution of the node degrees for the high-voltage 115 − 765 kV North American power grid. The model represents the power grid as a network of 14,099 nodes (substations) and 19,657 edges (transmission lines).

Source: R. Albert, I. Albert, G.L. Nakarado Structural vulnerability of the North American power grid Phys. Rev. E, 69 (2004), p. 025103(R)

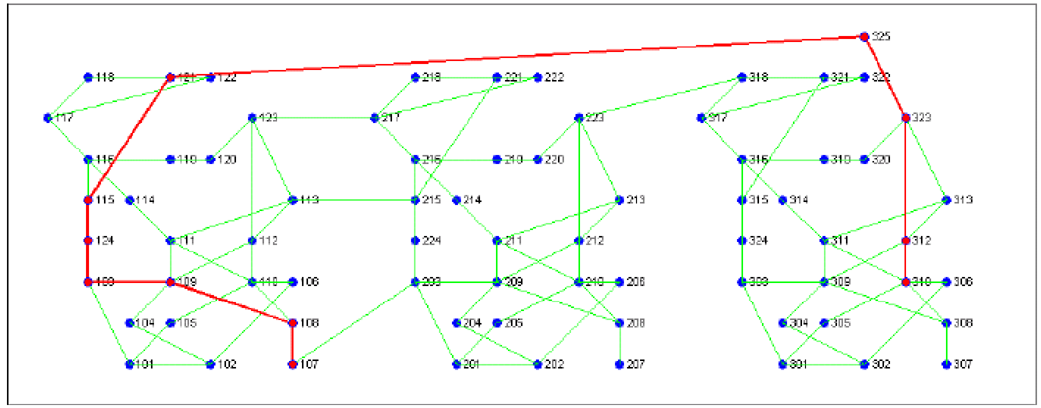


Cumulative distribution of the node degrees for the high-voltage transmission networks in Italy (full circles), Spain (diamonds) and France (squares). The empty circles represent the Italian „fine-grain“ network (from 380kV down to the distribution level).

Source: V. Rosato, *S. Bologna, F. Tiriticco*: Topological properties of high-voltage electrical transmission networks, Electric Power Systems Research, Vol. 77, 2007

Network Characteristics: Shortest Path (I)

Shortest Path between node 107 and 310



Assumption: the communication/service between two nodes is routed along the shortest path

Different *algorithms* are used to find the shortest path d_{ij} between two nodes *i* and *j*, e.g. Floyd-Warshall algorithm, Dijkstra‘s algorithm

Characteristic path length: average of all shortest paths in the network:

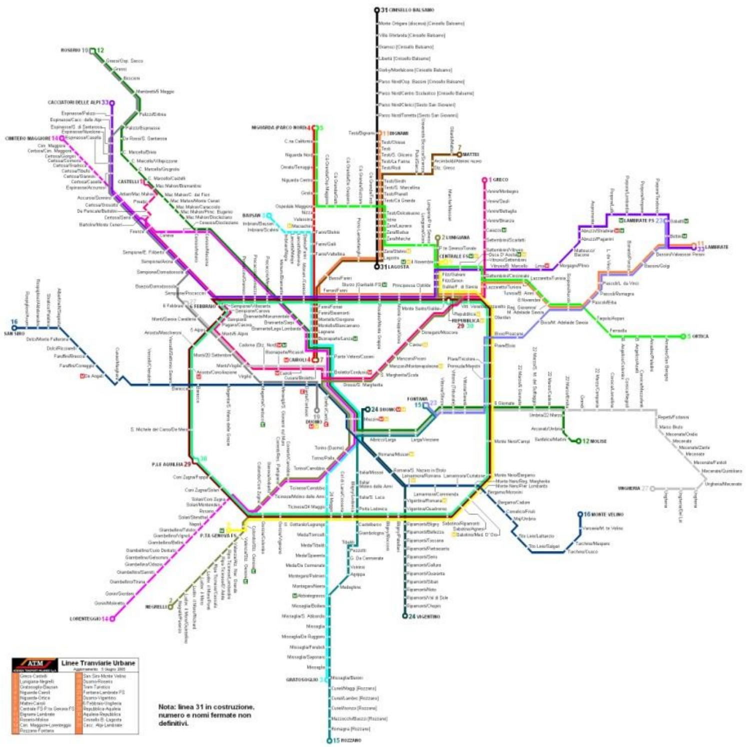
$$L = \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij}$$

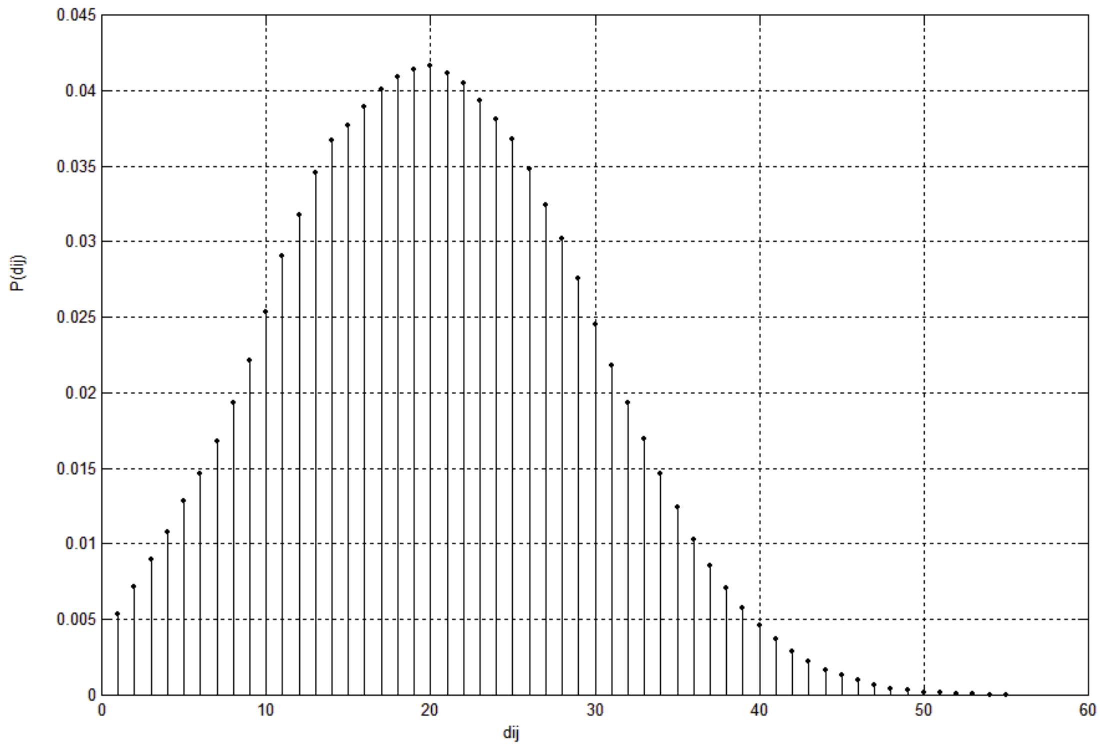
Average distance which has to be covered to reach the majority of nodes in the graph. Global property: typical separation between two nodes

Its value becomes infinity in case of a network splitting, due to e.g. disruption.

Network diameter: $d = \max_{i,j} d_{ij}$

Shortest path length distribution $P(d_{ij})$



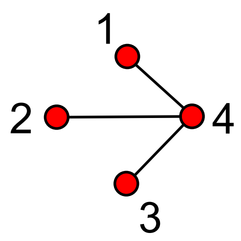


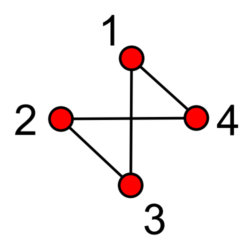
Most probable distance to travel : L = 20 stops

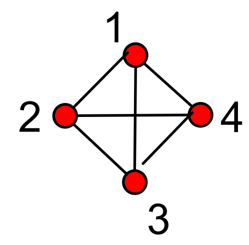
Maximum distance to travel: d = 55 stops

Network Characteristics: Shortest Path (II)

- Find the shortest path matrix for these undirected graphs:







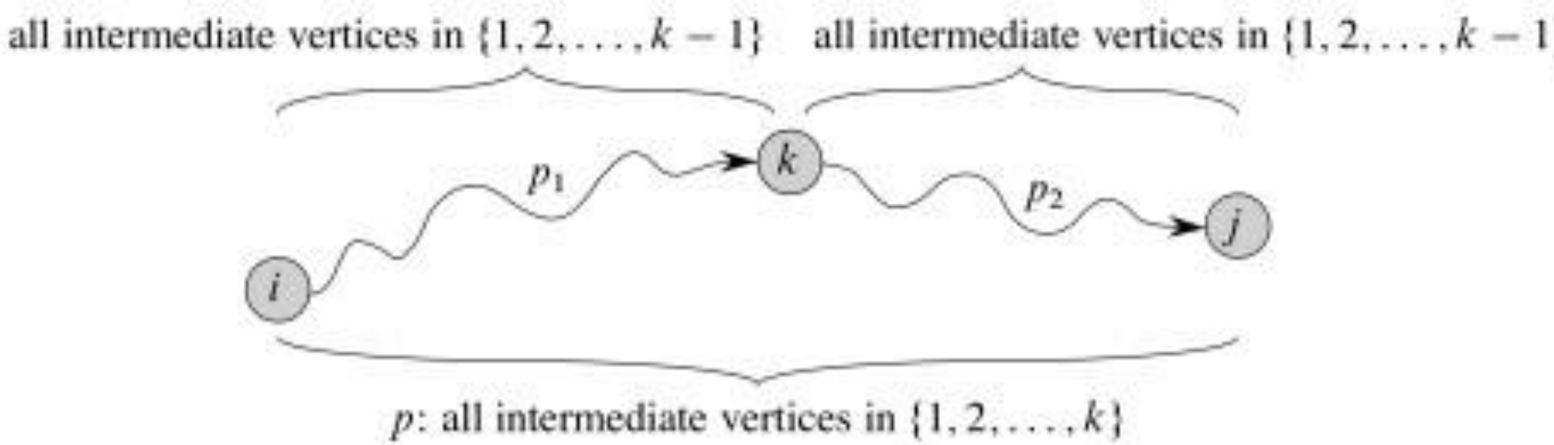
$$\begin{array}{c|cccc} & a_1 & a_2 & a_3 & a_4 \\ \hline a_1 & 0 & 0 & 0 & 1 \\ a_2 & 0 & 0 & 0 & 1 \\ a_3 & 0 & 0 & 0 & 1 \\ a_4 & 1 & 1 & 1 & 0 \end{array} \qquad \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

$$\begin{array}{c|cccc} & d_1 & d_2 & d_3 & d_4 \\ \hline d_1 & - & 2 & 2 & 1 \\ d_2 & 2 & * & 2 & 1 \\ d_3 & 2 & 2 & / & 1 \\ d_4 & 1 & 1 & 1 & + \end{array} \qquad \begin{pmatrix} 0 & & & \\ & 0 & & \\ & & 0 & \\ & & & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

Floyd-Warshall Algorithm

FLOYD-WARSHALL′(W)

```
n = W.rows
D = W
for k = 1 to n
    for i = 1 to n
        for j = 1 to n
            d_ij = min(d_ij, d_ik + d_kj)
return D
```



Network Characteristics: Clustering Coefficient (I)

- How interlinked are my friends?

The clustering coefficient of node *i*, C_i, measures the density of connections around a particular node *i*. Suppose you (node *i*) have k_i close friends (your first neighbors) If they all are again friends among themselves there will be

$$C_{max} = \binom{k_i}{2} = \frac{k_i!}{(k_i-2)!\cdot 2!} = \frac{k_i\cdot(k_i-1)}{2}$$

links between them. Suppose that in reality there are only *y* connections between them. C_i will be

$$C_i = \frac{y}{C_{max}} = \frac{2\cdot y}{k_i\cdot(k_i-1)}$$

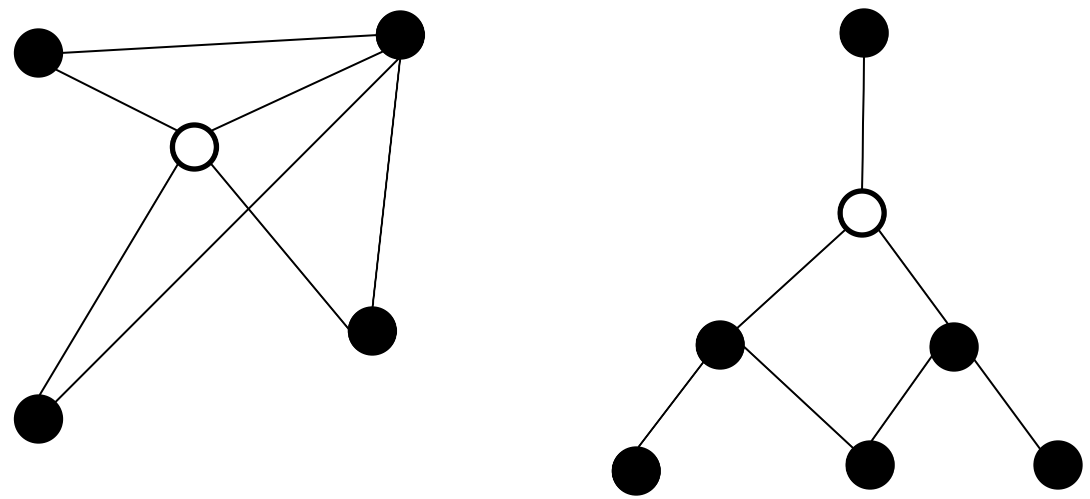
$$C_i = \frac{\text{Number of edges connecting the neighbours of } i}{\text{Max possible number of edges connecting the neighbours of } i,\ \frac{k_i(k_i-1)}{2}}$$

$$C = \frac{1}{N}\sum_i C_i$$

The **average** clustering coefficient *C* of a network is a measure of local connectivity. *C* = 1 for a fully-connected graph, and *C* = 0 for a complete sequential graph, i.e. a ring

Network Characteristics: Clustering Coefficient (II)

$$C_{max} = \binom{4}{2} = \frac{4!}{(4-2)! \cdot 2!} = \frac{4 \cdot 3 \cdot 2}{2 \cdot 2} = 6$$



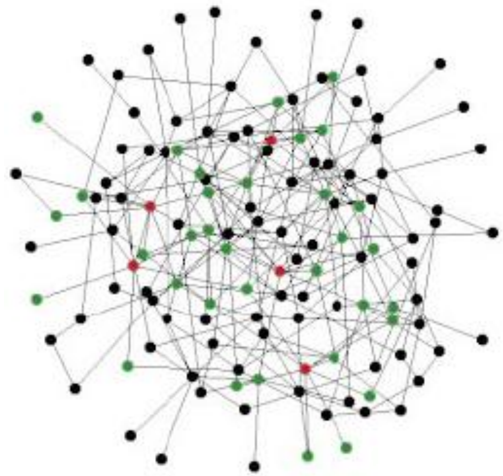
How would you calculate the clustering coefficient for the white node?

Large values of *C* would be welcome for the robustness of the connectivity: a node removal disconnecting two portions of the system would be overcome by simply passing onto adjacent working nodes through short-range neighboring nodes

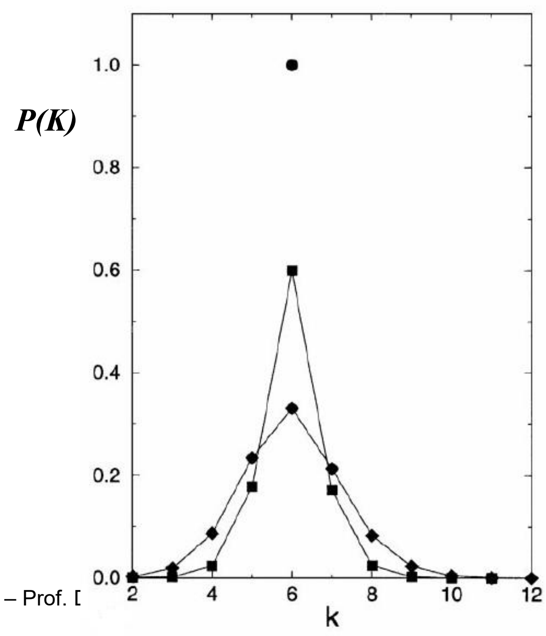
What you will learn today…

- ✓ Know the characteristics and representation of complex networks by Complex Network Theory
  - Exponential Networks vs. Scale-Free Networks
  - ✓ Global and Local Properties
- How to perform the network static structural analysis
  - Unweighted
  - Weighted
- Compute the system vulnerability to element removal
- Identify the centrality of elements in the structure
  - Centrality measures, bottlenecks
- How to Model Cascading Failures Propagation in Network Systems
  - How to Identify Cascade-safe Operating Margins

Exponential Networks - Small-World Property



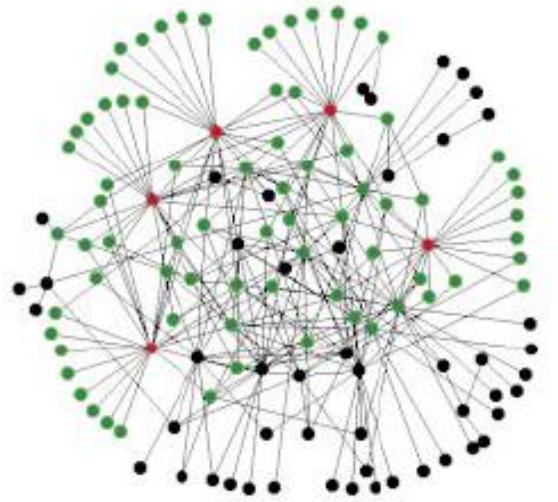
- Good global and local connectivity (local robustness due to clusters and global accessibility due to shortcuts among clusters)
- Characteristic scale (exponential decrease of the degree distribution)



- Less vulnerable to malicious attacks (no preferential nodes), tolerant to random faults

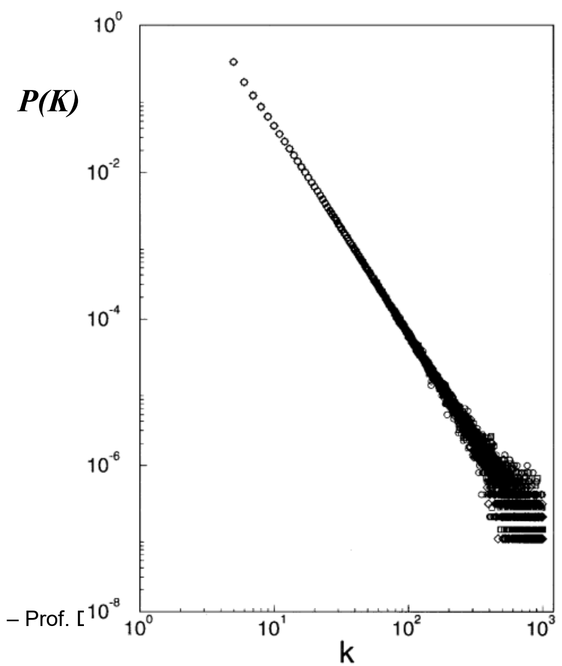
- They can be build using the Erdős-Rényi and Watts-Strogatz models

Power-Law Networks - Scale-Free Networks



- No characteristic scale in degree distribution (power law decrease) $P(k) \propto k^{-\gamma}\ ,\ \ k \neq 0$
- Few highly connected nodes

- <u>Robustness to random faults but vulnerability to targeted malicious attacks (due to the highly-connected hubs)</u>



- Scale-free networks are build from:
  - Growth process
  - Preferential attachment of new nodes to already well-connected nodes

Characteristics of Real-life Networks

	Network	Type	N	L	$\bar{k}$	ℓ	γ	C
Social	film actors	undirected	449 913	25 516 482	113.43	3.48	2.3	0.78
	company directors	undirected	7 673	55 392	14.44	4.60	–	0.88
	math coauthorship	undirected	253 339	496 489	3.92	7.57	–	0.34
	physics coauthorship	undirected	52 909	245 300	9.27	6.19	–	0.56
	biology coauthorship	undirected	1 520 251	11 803 064	15.53	4.92	–	0.60
	telephone call graph	undirected	47 000 000	80 000 000	3.16		2.1	
	email messages	directed	59 912	86 300	1.44	4.95	1.5/2.0	0.16
	email address books	directed	16 881	57 029	3.38	5.22	–	0.13
	student relationships	undirected	573	477	1.66	16.01	–	0.001
	sexual contacts	undirected	2 810				3.2	
Information	WWW `nd.edu`	directed	269 504	1 497 135	5.55	11.27	2.1/2.4	0.29
	WWW Altavista	directed	203 549 046	2 130 000 000	10.46	16.18	2.1/2.7	
	citation network	directed	783 339	6 716 198	8.57		3.0/–	
	Roget's Thesaurus	directed	1 022	5 103	4.99	4.87	–	0.15
	word co-occurrence	undirected	460 902	17 000 000	70.13		2.7	0.44
Technological	Internet	undirected	10 697	31 992	5.98	3.31	2.5	0.39
	power grid	undirected	4 941	6 594	2.67	18.99	–	0.080
	train routes	undirected	587	19 603	66.79	2.16	–	0.69
	software packages	directed	1 439	1 723	1.20	2.42	1.6/1.4	0.082
	software classes	directed	1 377	2 213	1.61	1.51	–	0.012
	electronic circuits	undirected	24 097	53 248	4.34	11.05	3.0	0.030
	peer-to-peer network	undirected	880	1 296	1.47	4.28	2.1	0.011
Biological	metabolic network	undirected	765	3 686	9.64	2.56	2.2	0.67
	protein interactions	undirected	2 115	2 240	2.12	6.80	2.4	0.071
	marine food web	directed	135	598	4.43	2.05	–	0.23
	freshwater food web	directed	92	997	10.84	1.90	–	0.48
	neural network	directed	307	2 359	7.68	3.97	–	0.28

Exponent γ is indicated only if the network is **scale-free**

Source: Newman, SIAM Rev. 45, 167 (2003)

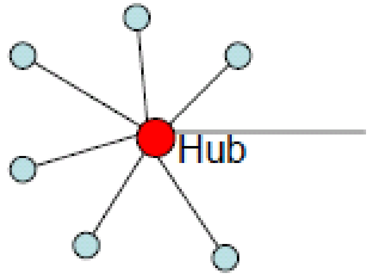
What you will learn today…

- ✓ Know the characteristics and representation of complex networks by Complex Network Theory
  - ✓ Exponential Networks vs. Scale-Free Networks
  - ✓ Global and Local Properties
- How to perform the network static structural analysis
  - Unweighted
  - Weighted
- Compute the system vulnerability to element removal
- Identify the centrality of elements in the structure
  - Centrality measures, bottlenecks
- How to Model Cascading Failures Propagation in Network Systems
  - How to Identify Cascade-safe Operating Margins

Random Failure and Attack Tolerance (I)

type of impact	exponential network	scale-free network
random	robust	extreme robust
malicious attack	robust	extreme vulnerable

scale-free network:



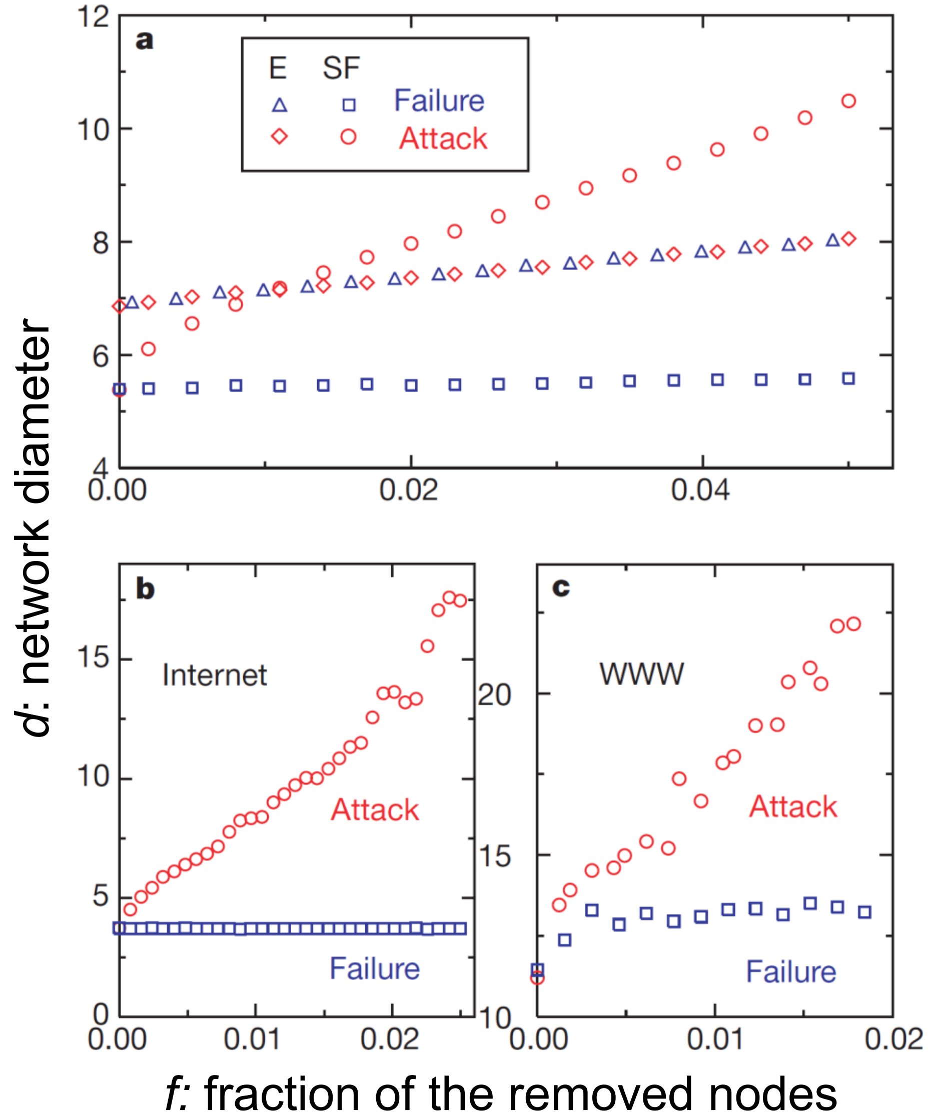
the chance to destroy the hub with a random attack is 1:7

a malicious attack to the hub destroys the connection to six nodes

Random Failure and Attack Tolerance (II)

E*: exponential*
SF*: scale-free*

Failure*: nodes are removed randomly*
Attack*: most connected nodes are removed*



Albert, Jeong, Barabasi, Nature 406, 378, 2000

What you will learn today…

- ✓ Know the characteristics and representation of complex networks by Complex Network Theory
  - ✓ Exponential Networks vs. Scale-Free Networks
  - ✓ Global and Local Properties
- ✓ How to perform the network static structural analysis
  - ✓ Unweighted
  - Weighted
- Compute the system vulnerability to element removal
- Identify the centrality of elements in the structure
  - Centrality measures, bottlenecks
- How to Model Cascading Failures Propagation in Network Systems
  - How to Identify Cascade-safe Operating Margins

Weighted Networks

- L and C are applicable only to unweighted networks in which only the topological information on the **existence or absence of a link**, with no reference to the physical length and capacity of the link, is retained
- The model of a realistic network could be weighed to account also for the physical properties of the systems
- In addition to the adjacency matrix $[a_{ij}]$, defined as for the unweighted graph, an additional matrix $[l_{ij}]$ of **weights**, e.g. physical distances, failure/accident probabilities, or 'electrical' distances can be introduced

- For an unweighted network, $l_{ij} = 1 = a_{ij}$ for $i \neq j$

Physical Weights l_{ij} – Coupling between nodes

- Physical length $$l_{ij} = length_{ij}$$
- Travel / communication time $$l_{ij} = \tau_{ij}$$
- Link reliability $p_{ij} = e^{-\lambda \cdot length_{ij}}$ $$l_{ij} = \frac{1}{p_{ij}} = e^{\lambda \cdot length_{ij}}$$
- Electric flow $$l_{ij} = \frac{\text{average flow of all lines}}{\text{flow of line } ij}$$

Idea to construct l_{ij} : long/unreliable/weakly-loaded/slow connections will be "longer" and less likely used in shortest paths

$$d_{ij} = \min_{\gamma_{ij}} \left(\sum_{mn \in \gamma_{ij}} l_{mn} \right)$$

If weight is time, flow or length

$$d_{ij} = \min_{\gamma_{ij}} \left(\prod_{mn \in \gamma_{ij}} (l_{mn}) \right)$$

If weight is reliability

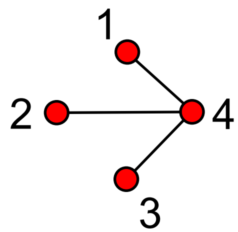
Weighted Analysis – Efficiency Matrix ε_{ij}

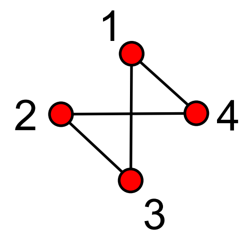
- The matrix of the shortest path lengths $[d_{ij}]$ is computed on the basis of both $[a_{ij}]$ and $[l_{ij}]$: the length d_{ij} of the shortest path linking *i* and *j* in the network is the smallest sum of the weights (e.g. physical distance in case l_{xy} is the physical length of the arc linking node *x* and *y*) throughout all the possible paths from *i* to *j*.
- Assuming that the network system is parallel, i.e. that every node concurrently sends information through its edges, a **measure of efficiency** in the communication between nodes *i* and *j* can be defined, inversely proportional to the shortest distance. Thus, the network is characterized also by an efficiency matrix $[\varepsilon_{ij}]$, whose entry is the efficiency in the communication between nodes *i* and *j*:

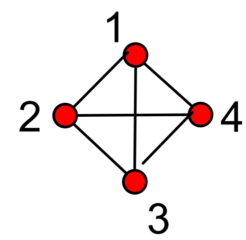
$$
\begin{aligned}
\varepsilon_{ij} &= \frac{1}{d_{ij}} \quad \text{if there is at least one path connecting } i \text{ and } j \\
&= 0 \quad \text{otherwise } (d_{ij} = \infty)
\end{aligned}
$$

Network Characteristics: Efficiency ε_{ij}

- Find the efficiency matrix ε_{ij} for these undirected **unweighted** graphs:







$$\begin{array}{c} \\ d_1 \\ d_2 \\ d_3 \\ d_4 \end{array} \begin{array}{c} \begin{array}{cccc} d_1 & d_2 & d_3 & d_4 \end{array} \\ \begin{pmatrix} 0 & 2 & 2 & 1 \\ 2 & 0 & 2 & 1 \\ 2 & 2 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array} \qquad \begin{pmatrix} 0 & 2 & 1 & 1 \\ 2 & 0 & 1 & 1 \\ 1 & 1 & 0 & 2 \\ 1 & 1 & 2 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

$$\begin{array}{c} \\ \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{array} \begin{array}{c} \begin{array}{cccc} \varepsilon_1 & \varepsilon_2 & \varepsilon_3 & \varepsilon_4 \end{array} \\ \begin{pmatrix} 0 & 0.5 & 0.5 & 1 \\ 0.5 & 0 & 0.5 & 1 \\ 0.5 & 0.5 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array} \qquad \begin{pmatrix} 0 & 0.5 & 1 & 1 \\ 0.5 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0.5 \\ 1 & 1 & 0.5 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

Global and Local Efficiency

- Global efficiency

~ Characteristic path length *L*

$$E_{glob}(G) = \frac{\sum_{i \neq j \in G} \varepsilon_{ij}}{N(N-1)} = \frac{\sum_{i \neq j \in G} \frac{1}{d_{ij}}}{N(N-1)}$$

- Efficiency of the subgraph G_i of the neighborhood of node *i*

~ Clustering coefficient C_i

$$E(G_i) = \frac{\sum_{n \neq m \in G_i} \varepsilon_{nm}}{k_i(k_i - 1)}$$

- Local efficiency

~ Clustering coefficient *C*

$$E_{loc}(G) = \frac{1}{N} \sum_{i \in G} E(G_i)$$

What you will learn today…

- ✓ Know the characteristics and representation of complex networks by Complex Network Theory
  - ✓ Exponential Networks vs. Scale-Free Networks
  - ✓ Global and Local Properties
- ✓ How to perform the network static structural analysis
  - ✓ Unweighted
  - ✓ Weighted
- Compute the system vulnerability to element removal
- Identify the centrality of elements in the structure
  - Centrality measures, bottlenecks
- How to Model Cascading Failures Propagation in Network Systems
  - How to Identify Cascade-safe Operating Margins

Vulnerability to Element Removal

Swiss 220/380 kV Power Network

Vulnerability Index for Links

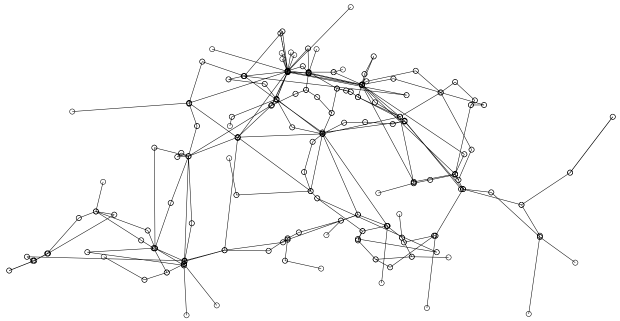
- Degradation of the global efficiency of the network due to the disconnection of a set of its links

$$V^* = \frac{E_{glob}(G) - E_{glob}(G^*)}{E_{glob}(G)}$$

- *G*: original network
- *G**: network after the disconnection of a link

Swiss 220/380 kV Power Network

Swiss 220/380 kV transmission system



N = 161 nodes (substations)

K = 219 edges (overhead lines)

0	1	0	0	1	0	0	0	0	0	0	0	0	0
1	0	1	1	1	0	0	0	0	0	0	0	0	0
0	1	0	1	0	0	0	0	0	0	0	0	0	0
0	1	1	0	1	0	1	0	1	0	0	0	0	0
1	1	0	1	0	1	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	1	1	1	0
0	0	0	1	0	0	0	1	1	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0	0	0	0	0
0	0	0	1	0	0	1	0	0	1	0	0	0	1
0	0	0	0	0	0	0	0	1	0	1	0	0	0
0	0	0	0	0	1	0	0	0	1	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0	1	0
0	0	0	0	0	1	0	0	0	0	0	1	0	1
0	0	0	0	0	0	0	0	1	0	0	0	1	0

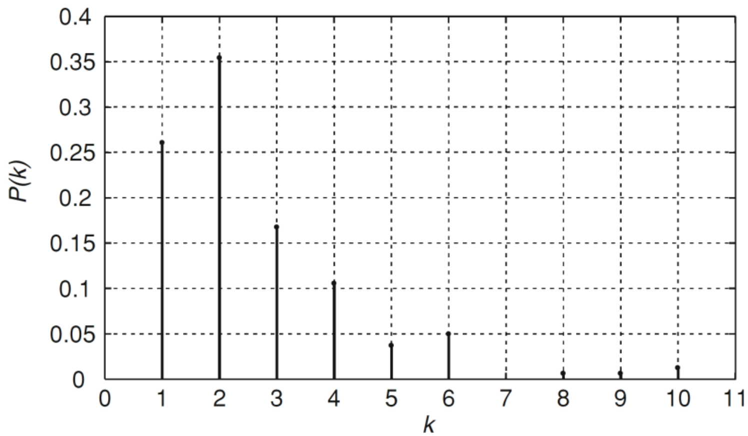
- l_{ij} weight matrix = 1 / p_{ij} (link reliability) (λ: failure rate per unit length)

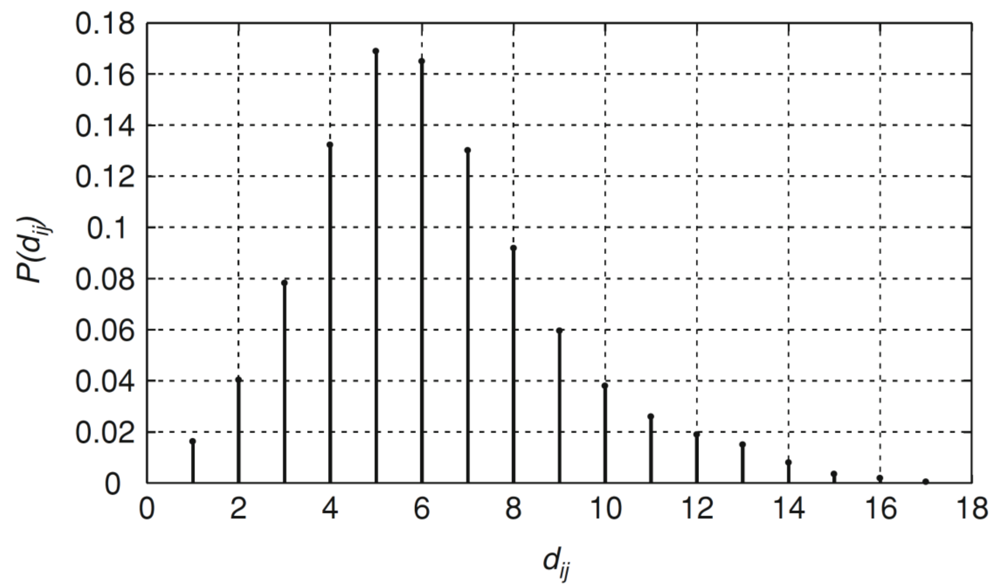
$$p_{ij} = e^{-\lambda \cdot length_{ij}}$$

- Most reliable path (Floyd iterative algorithm)

$$d_{ij} = \min_{\gamma_{ij}} \left(\frac{1}{\prod_{mn \in \gamma_{ij}} p_{mn}} \right)$$

Degree, Shortest Paths, Efficiencies



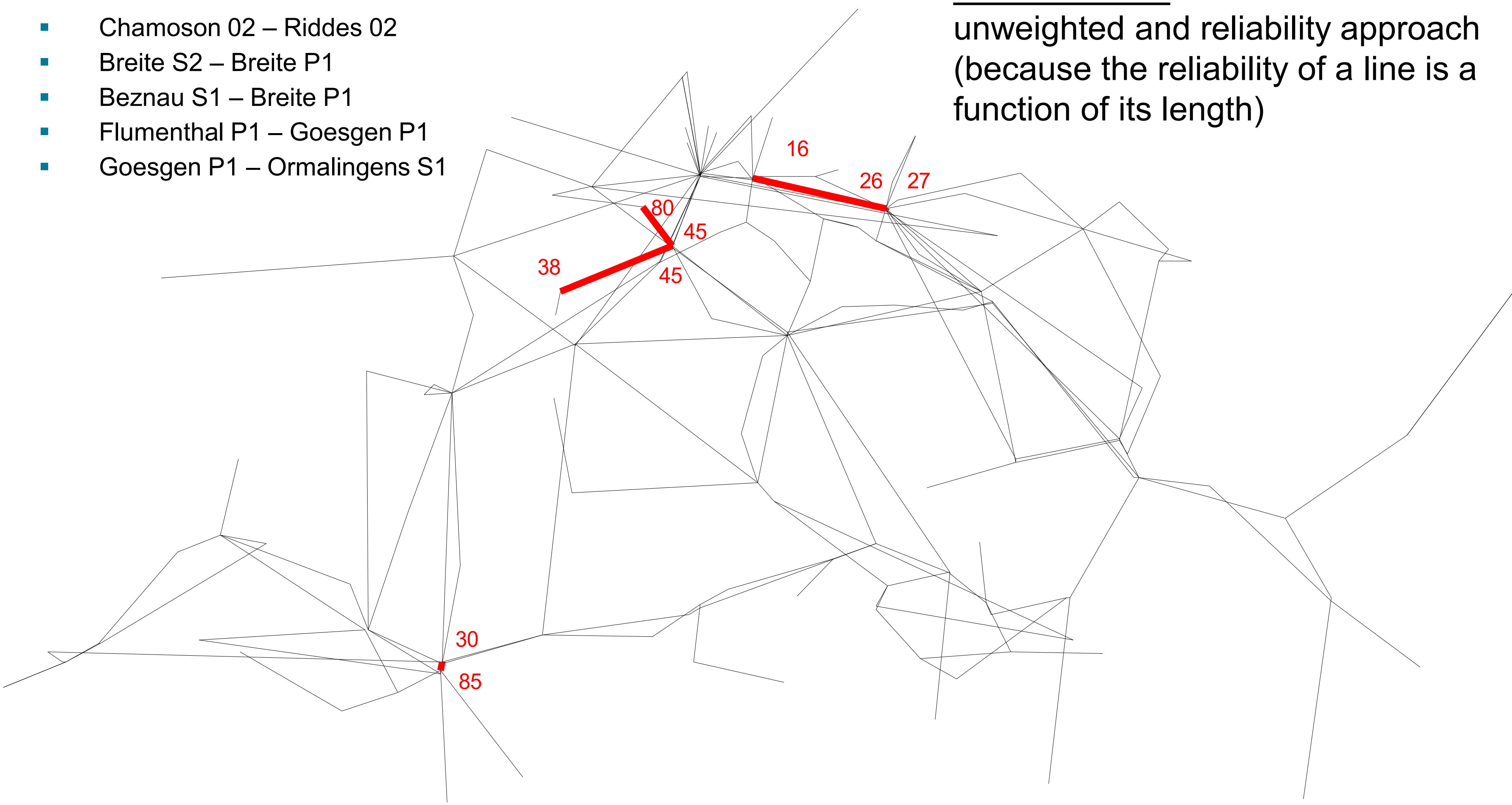


	Topological efficiency	Reliability efficiency
Global efficiency, $E_{\mathrm{glob}}(G)$	20.5×10^{-2}	9.30×10^{-2}
Local efficiency, $E_{\mathrm{loc}}(G)$	7.89×10^{-2}	4.72×10^{-2}

Most Vulnerable Lines - Reliability Weight

- Chamoson 02 – Riddes 02
- Breite S2 – Breite P1
- Beznau S1 – Breite P1
- Flumenthal P1 – Goesgen P1
- Goesgen P1 – Ormalingens S1

<u>No differences</u> in results between unweighted and reliability approach (because the reliability of a line is a function of its length)



Swiss 220/380 kV transmission system

Reliability + Electric Flow Weight

p_{mn}: link reliability
(λ: failure rate per unit length)

Reliability weight

$$l_{mn}^{reliability} = \frac{1}{p_{mn}} = e^{+\lambda \bullet length_{mn}}$$

$$e_{mn} = \frac{\text{average flow of line } mn}{\text{average flow of all lines}}$$

Electric flow weight

$$l_{mn}^{electric} = \frac{\text{average flow of all lines}}{\text{average flow of line } mn}$$

$$d_{ij} = \min_{\gamma_{ij}} \left(\prod_{mn \in \gamma_{ij}} \left(l_{mn}^{reliability} \cdot l_{mn}^{electric} \right) \right) = \min_{\gamma_{ij}} \left(\frac{1}{\prod_{mn \in \gamma_{ij}} \left(p_{mn} \cdot e_{mn} \right)} \right)$$

Most reliable path between *i* and *j* combined with the carried electric flow

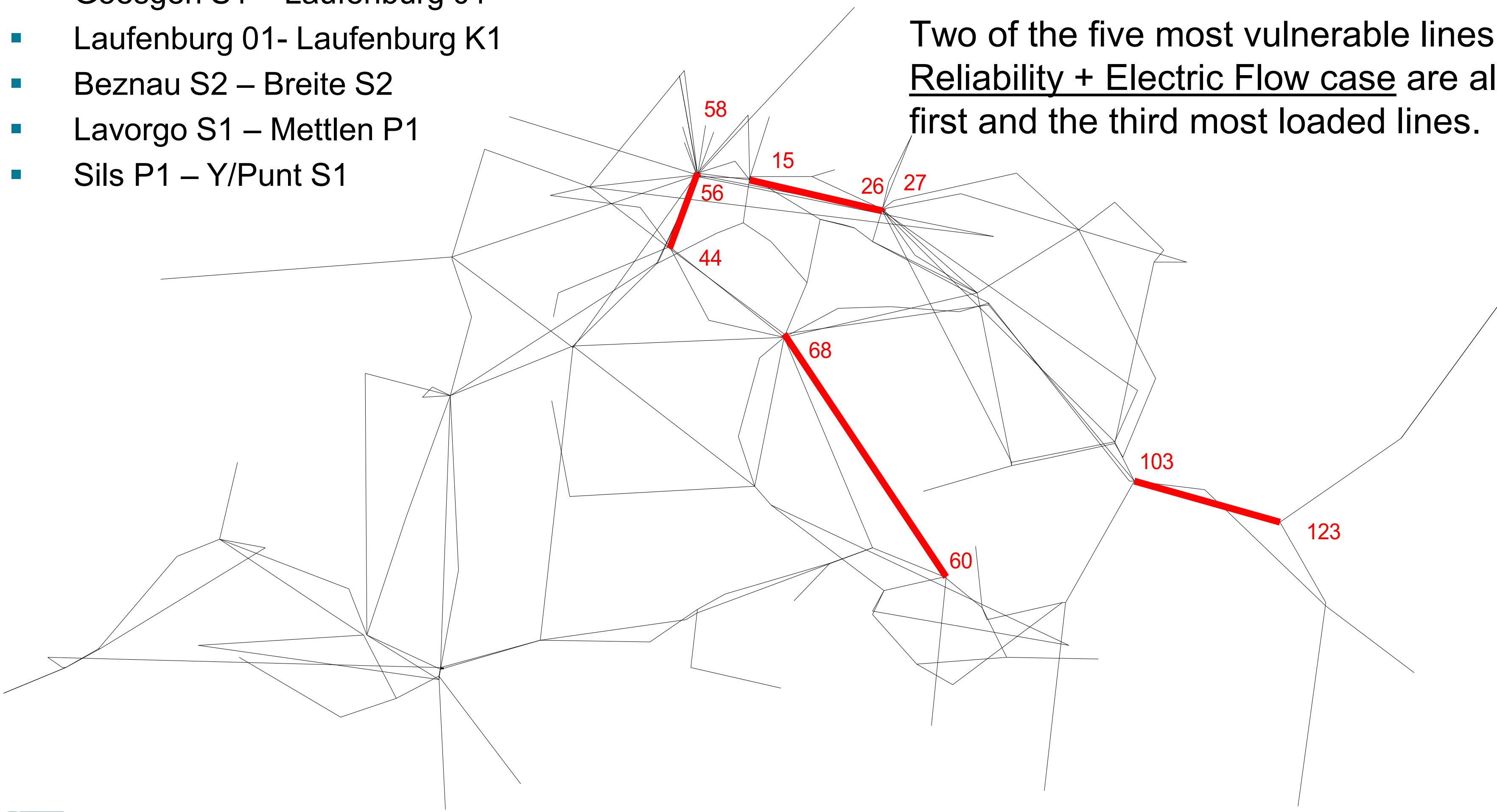
ETH zürich

Most Vulnerable Lines - Reliability + Electric Flow Weight

- Goesgen S1 – Laufenburg 01
- Laufenburg 01- Laufenburg K1
- Beznau S2 – Breite S2
- Lavorgo S1 – Mettlen P1
- Sils P1 – Y/Punt S1

Beznau S2 – Breite S2 connection is among the most vulnerable ones in any case.

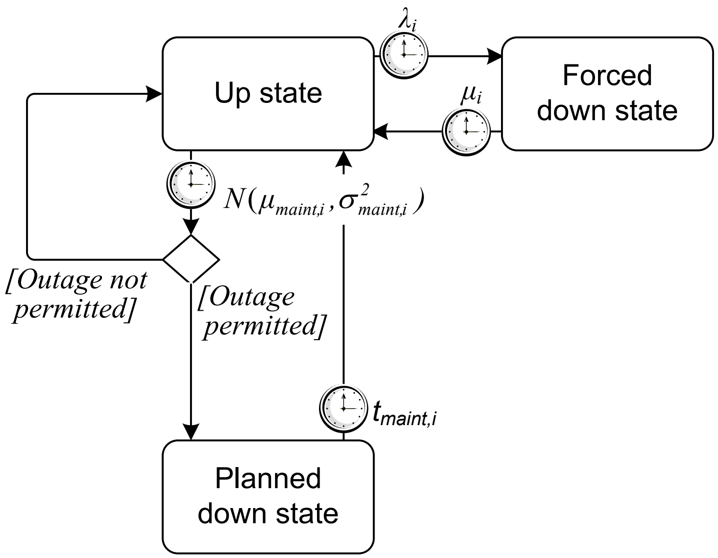
Two of the five most vulnerable lines in the Reliability + Electric Flow case are also the first and the third most loaded lines.



Swiss 220/380 kV transmission system

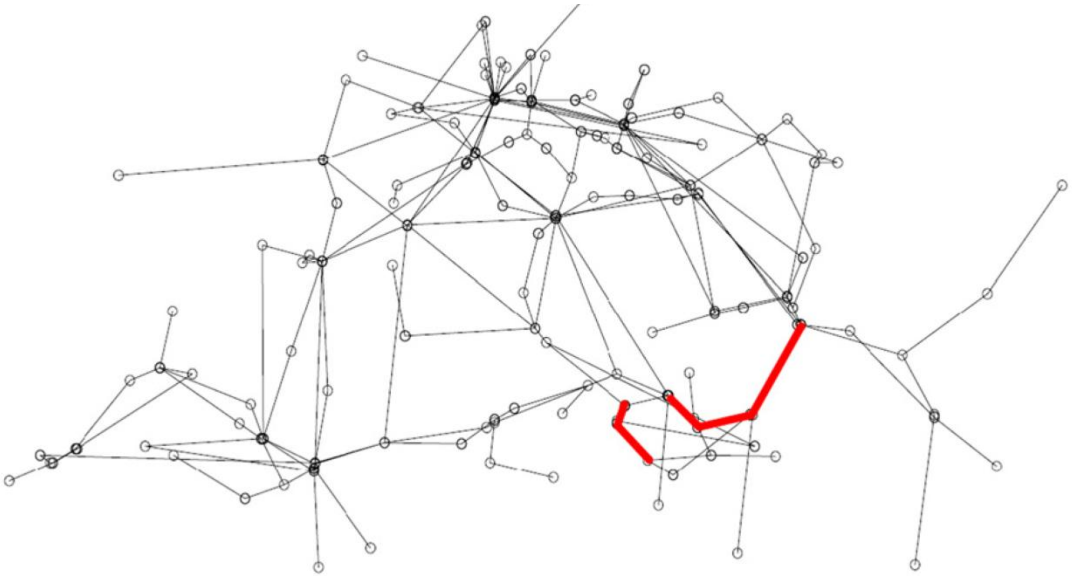
Comparison with Agent-Based+ Power Flow Modeling (I)

- Agents with state charts and analytical functions
  - Generators
  - Loads
  - Busbars
  - Transmission Line
  - Operator/Dispatcher



- Interactions
  - Physical Dependencies (node injections and line flows)
  - Human behavior (dispatch of the power system)

Comparison with Agent-Based+ Power Flow Modeling (II)



Most vulnerable lines

This much sophisticated model contains:

- Power flow equations (more than structure topology as in Complex Networks)
- Power exchange through Country boundaries
- Event-driven simulation, i.e. operational dynamics, line disconnection, operator actions

But requires more **data** and **computational resources**

What you will learn today…

- ✓ Know the characteristics and representation of complex networks by Complex Network Theory
  - ✓ Exponential Networks vs. Scale-Free Networks
  - ✓ Global and Local Properties
- ✓ How to perform the network static structural analysis
  - ✓ Unweighted
  - ✓ Weighted
- ✓ Compute the system vulnerability to element removal
- Identify the centrality of elements in the structure
  - Centrality measures, bottlenecks
- How to Model Cascading Failures Propagation in Network Systems
  - How to Identify Cascade-safe Operating Margins

Measures of Topological Centralities in Networks

Italian 380 kV Power Network

Centrality Measures

- Rely on topological information to qualify the importance of a network element
- Quantify the relevance of the element's location in the network with respect to a given network performance
- Originated in social network analysis, they qualify the role played by an element in the complex interaction and communication occurring in the network
- Classical topological centrality measures are: degree centrality, closeness centrality, betweenness centrality and information centrality

Measures of Topological Centrality in Networks (I)

1. Topological degree centrality, C^D

$$C_i^D = \frac{k_i}{N-1} = \frac{\sum_{j \neq i \in G} a_{ij}}{N-1}$$

Highest importance to the node with the largest number of first neighbors

2. Topological closeness centrality, C^C

$$C_i^C = \frac{N-1}{\sum_{j \neq i \in G} d_{ij}}$$

Identify the nodes which on average need fewer steps to communicate with the other nodes

Measures of Topological Centrality in Networks (II)

3. Topological betweenness centrality, C^B

$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j,k \in G, j \neq k \neq i} \frac{n_{jk}(i)}{n_{jk}}$$

A node is central if it is traversed by many of the shortest paths connecting pairs of nodes

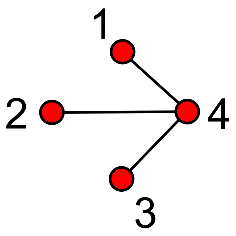
4. Topological information centrality, C^I

$$C_i^I = \frac{\Delta E(i)}{E} = \frac{E[G] - E[G^*(i)]}{E[G]}$$

Relates a node importance to the ability of the network to respond to the deactivation of the node

Network Characteristics: Betweeness Centrality

- Find the betweenness centrality of node 4, $C^B{}_4$, in this undirected graph.



$$C_i^B = \frac{1}{(N-1)(N-2)} \sum_{j,k \in G, j \neq k \neq i} \frac{n_{jk}(i)}{n_{jk}}$$

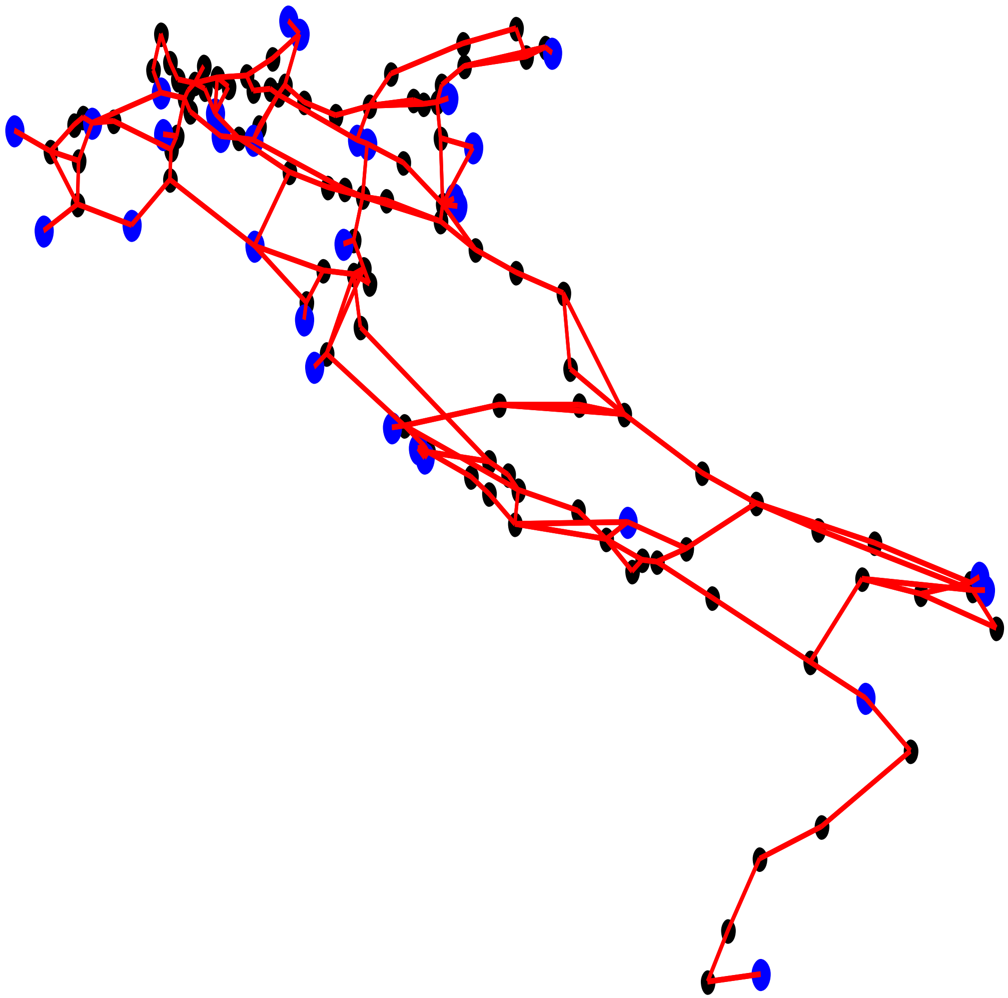
$$\begin{array}{c} \\ \mathbf{a_1} \\ \mathbf{a_2} \\ \mathbf{a_3} \\ \mathbf{a_4} \end{array} \begin{array}{c} \mathbf{a_1}\ \mathbf{a_2}\ \mathbf{a_3}\ \mathbf{a_4} \\ \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array}$$

n_{jk} number of shortest paths from node *j* to node *k*

$n_{jk}(i)$ number of shortest paths from node *j* to node *k* which pass through node *i*

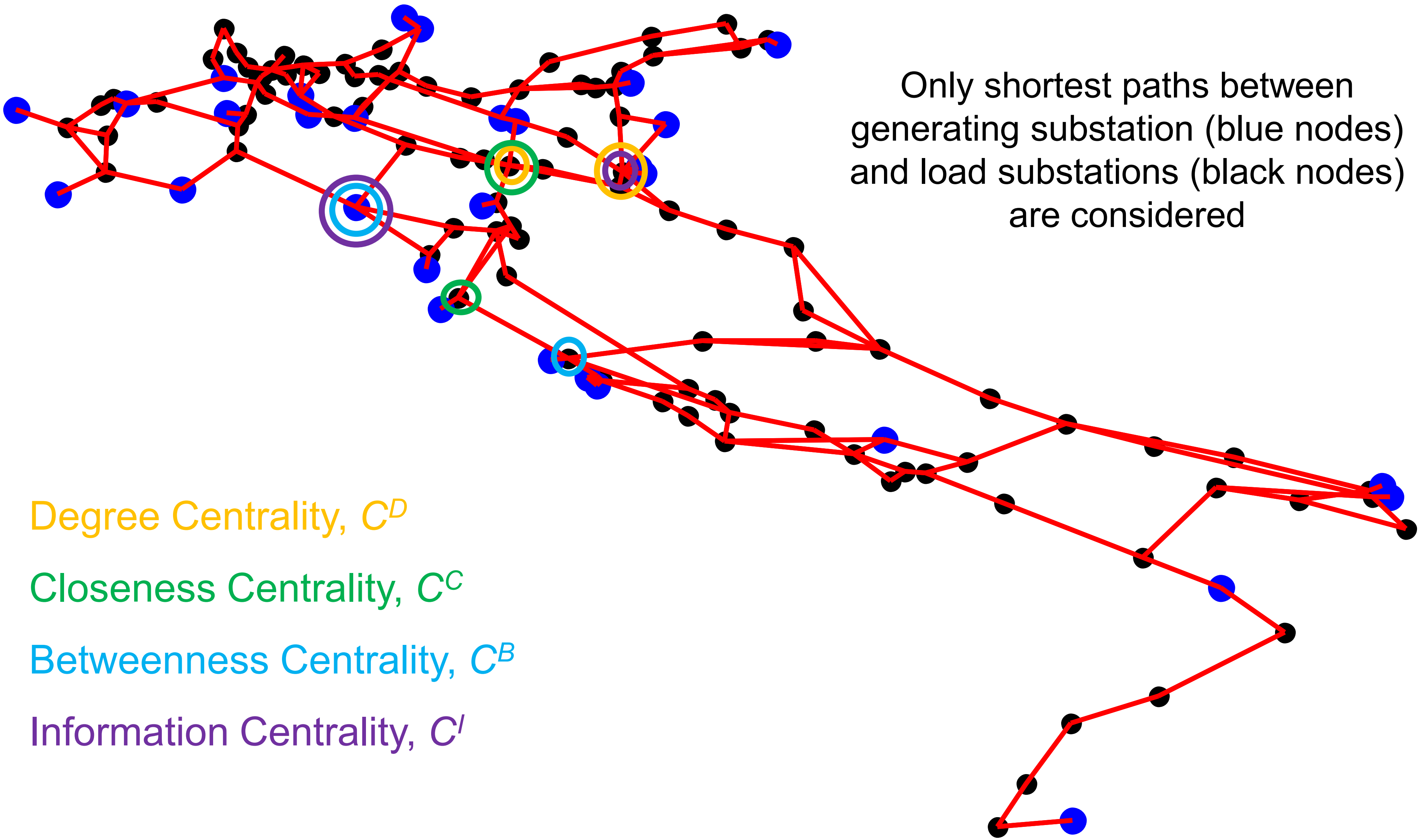
$$\begin{array}{c} \\ \mathbf{d_1} \\ \mathbf{d_2} \\ \mathbf{d_3} \\ \mathbf{d_4} \end{array} \begin{array}{c} \mathbf{d_1}\ \mathbf{d_2}\ \mathbf{d_3}\ \mathbf{d_4} \\ \begin{pmatrix} 0 & 2 & 2 & 1 \\ 2 & 0 & 2 & 1 \\ 2 & 2 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \end{array}$$

Application to the Italian 380 kV Power Network



N=127 nodes, *K*=342 links

Topological Centralities of Substations



Only shortest paths between generating substation (blue nodes) and load substations (black nodes) are considered

Degree Centrality, C^D

Closeness Centrality, C^C

Betweenness Centrality, C^B

Information Centrality, C^I

What you will learn today…

- ✓ Know the characteristics and representation of complex networks by Complex Network Theory
  - ✓ Exponential Networks vs. Scale-Free Networks
  - ✓ Global and Local Properties
- ✓ How to perform the network static structural analysis
  - ✓ Unweighted
  - ✓ Weighted
- ✓ Compute the system vulnerability to element removal
- ✓ Identify the centrality of elements in the structure
  - ✓ Centrality measures, bottlenecks
- How to Model Cascading Failures Propagation in Network Systems
  - How to Identify Cascade-safe Operating Margins

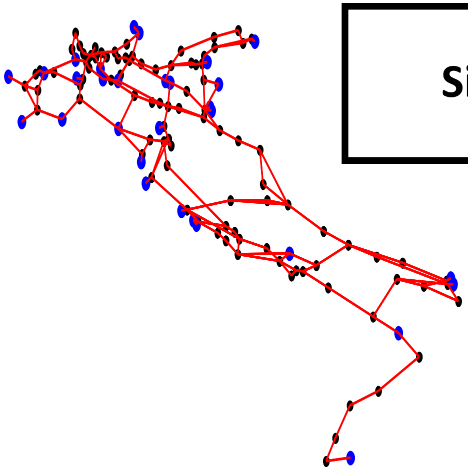
Different Types of Analysis

Network static structure analysis

- Topological analysis
- Weighted analysis

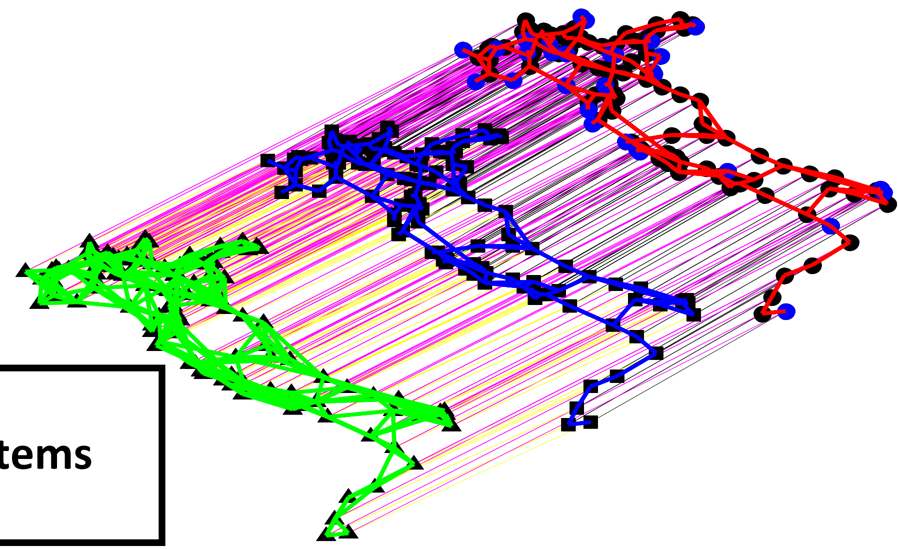
Dynamic modeling of cascading failures propagation in network systems

- Identification of cascade-safe operating margins
- Criticality indicators in failure cascade processes



Single system

Systems of systems



Two Models of “Abstract” Cascading Failures

1. Flow-based failure propagation
   - load = total number of shortest paths passing through that component
2. Local propagation of a fixed amount of load
   - fixed load transferred after a failure

Flow-based failure propagation

General Classification

- Common Cause Failures (CCF): multiple failures that result directly from a single common or shared root cause
  - Extreme environmental conditions
  - Failure of a piece of hardware external to the system
  - Human Error (operational or maintenance)

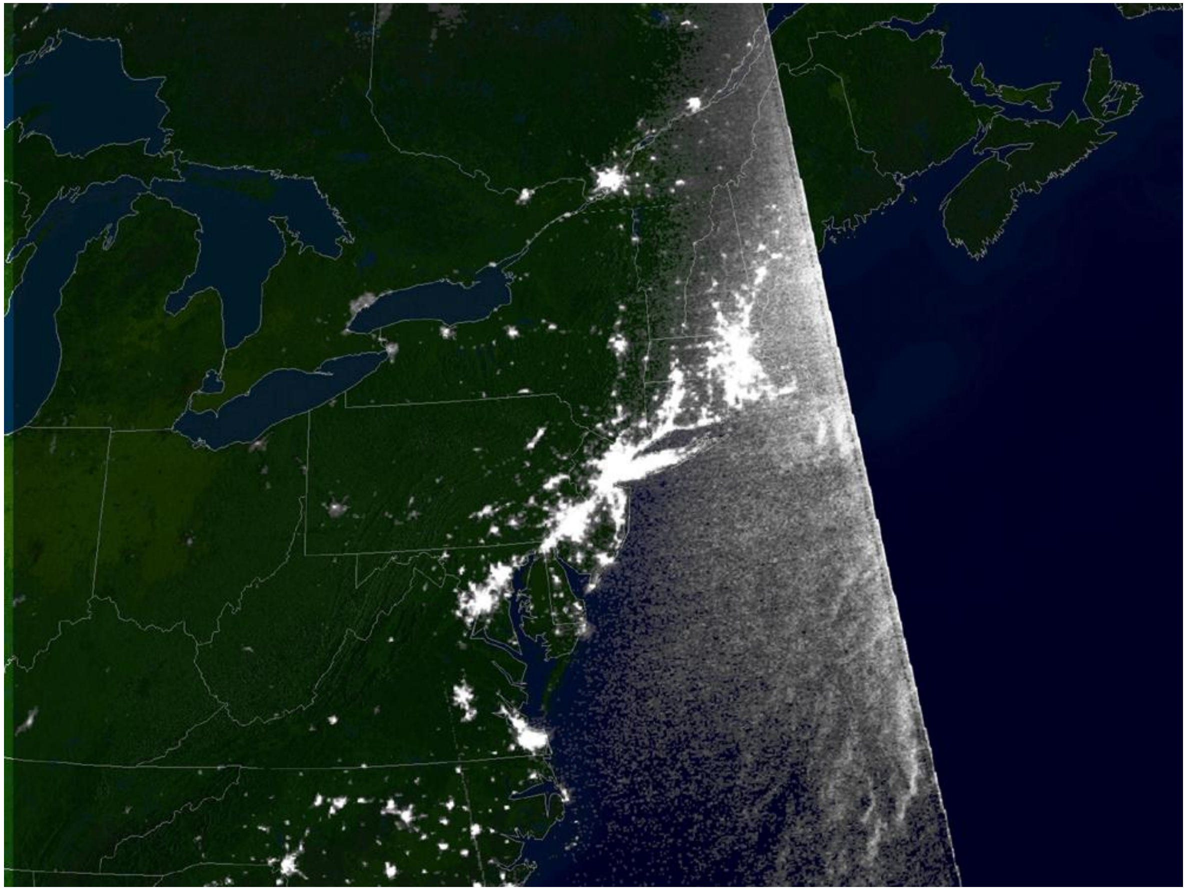
  e.g. fire at Browns Ferry Nuclear Power Plant (1975)

- Cascading Failures: several component share a common load → 1 component failure may lead to increase load on the remaining ones → increased likelihood of failure

  e.g. http://www.youtube.com/watch?v=Jj_K6bGQIfM

  e.g. 2003 northeast Blackout

2003 US Northeast Blackout



1. Software bug in GE energy management systems
2. Inadequate right-of-way for power lines
3. Poor coordination among TSO
4. Poor maintenance (Eastlake, Ohio generating plant shuts down)

Estimated 10 million people in Ontario and 45 million people in eight U.S. states

Electric Power Supply Systems: Recent Major Blackouts

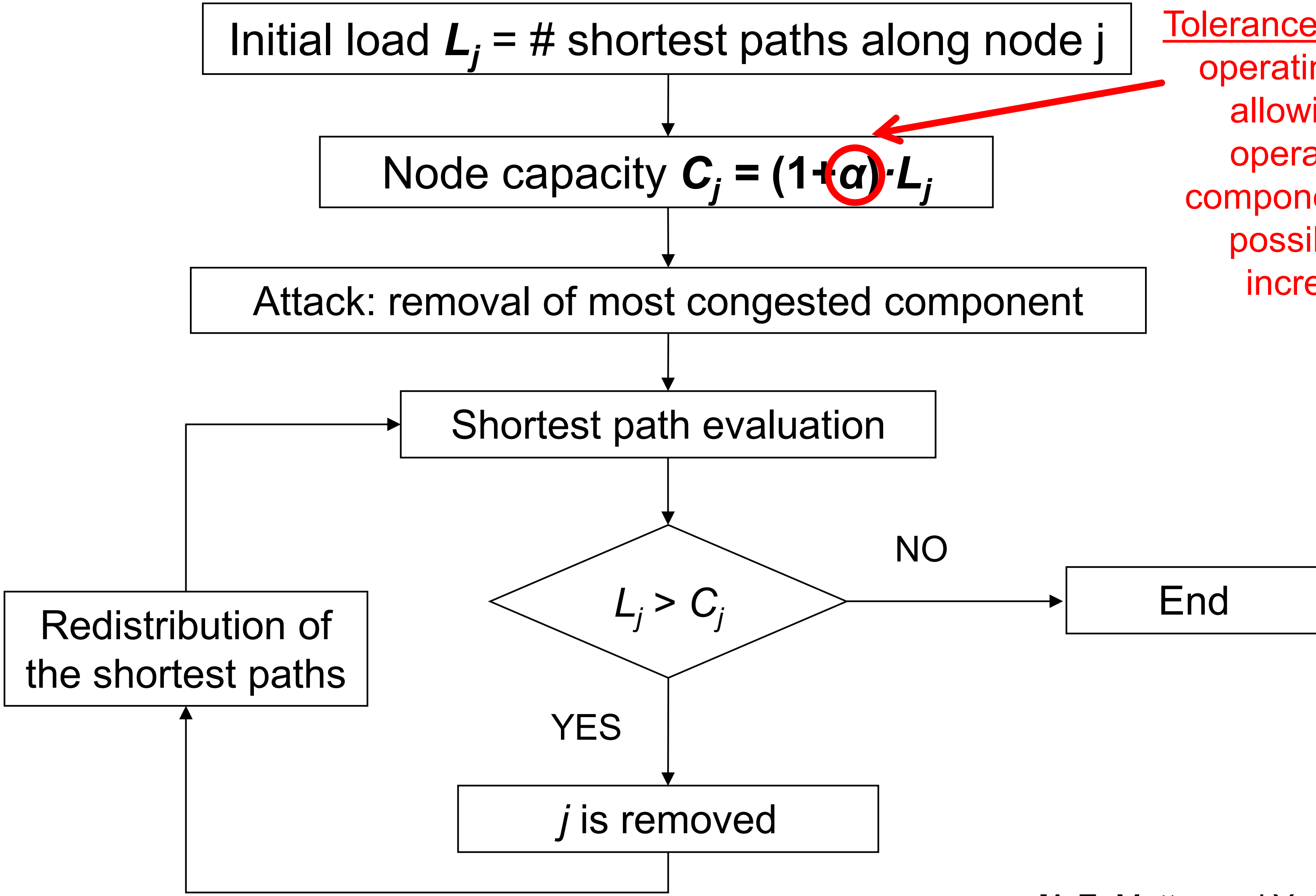
Systemic stress, e.g. high loading condition

Local event, e.g. line flashover

Systemic effect, i.e. cascading failure

Blackout		Load loss (GW)	Duration (h)	People affected	Main causes
Aug 14, 2003	Great Lakes, NYC	~60	~16	50 million	Inadequate right-of-way maintenance, EMS failure, poor coordination among neighboring TSOs
Aug 28, 2003	London	0.72	1	500,000	Incorrect line protection device setting
Sept 23, 2003	Denmark/Sweden	6.4	~7	4.2 million	Two independent component failures (not covered by N-1 rule)
Sept 28, 2003	Italy	~30	up to 18	56 million	High load flow CH-I, line flashovers, poor coordination among neighboring TSOs
July 12, 2004	Athens	~9	~3	5 million	Voltage collapse
May 25, 2005	Moscow	2.5	~4	4 million	Transformer fire, high demand leading to overload conditions
June 22, 2005	Switzerland (railway supply)	0.2	~3	200,000 passengers	Non-fulfillment of the N-1 rule, wrong documentation of line protection settings, inadequate alarm processing
Aug 14, 2006	Tokyo	?	–5	0.8 million households	Damage of a main line due to construction work
Nov 4, 2006	Western Europe ("controlled" line cut off)	~14	~2	15 million households	High load flow D-NL, violation of the N-1 rule, poor inter TSO–coordination
Nov 10, 2009	Brazil, Paraguay	~14	~4	60 million	Short circuit on key power line due to bad weather, Itaipu hydro plant (18 GW) shut down

1 - Flow-based Failure Propagation



Tolerance parameter: operating margin allowing safe operations of component *j* under possible load increments

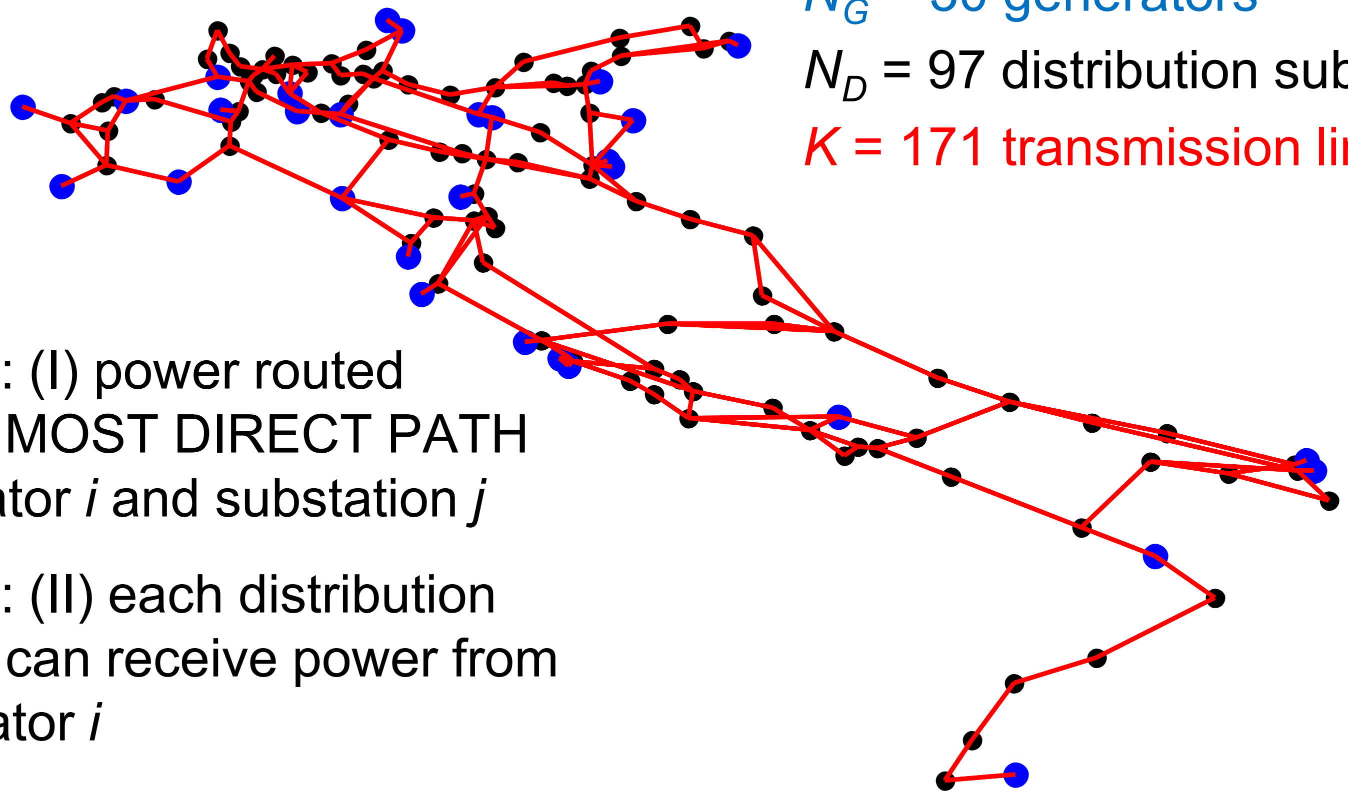
[A.E. Motter and Y.-C. Lai 2002]

Cascading Failures in Power Transmission Networks

N_G = 30 generators

N_D = 97 distribution substations

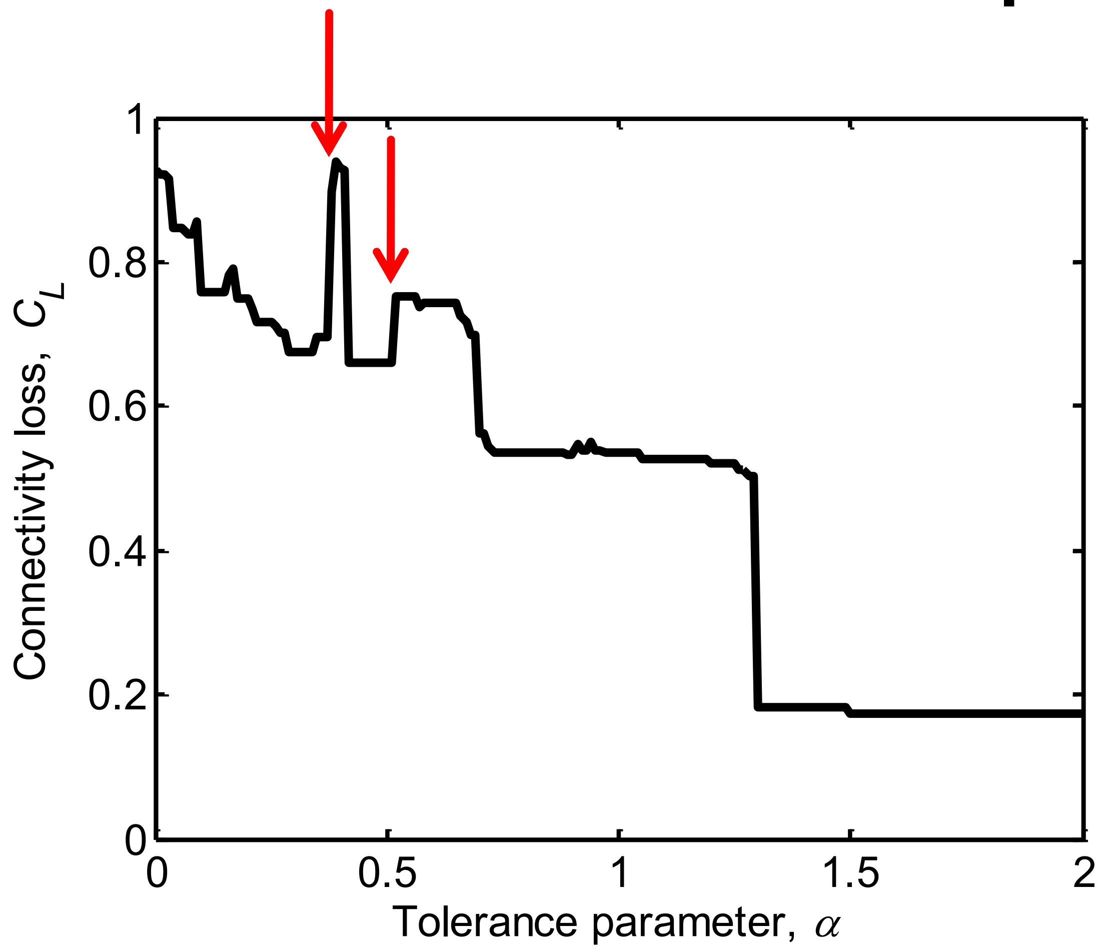
K = 171 transmission lines



- Assumption: (I) power routed through the MOST DIRECT PATH from generator *i* and substation *j*
- Assumption: (II) each distribution substation *j* can receive power from ANY generator *i*

L_j load on substation *j* = number of shortest paths connecting every generator to every distributor passing through *j*
i.e. node *j* betweenness

Identification of Cascade-Safe Operating Margins



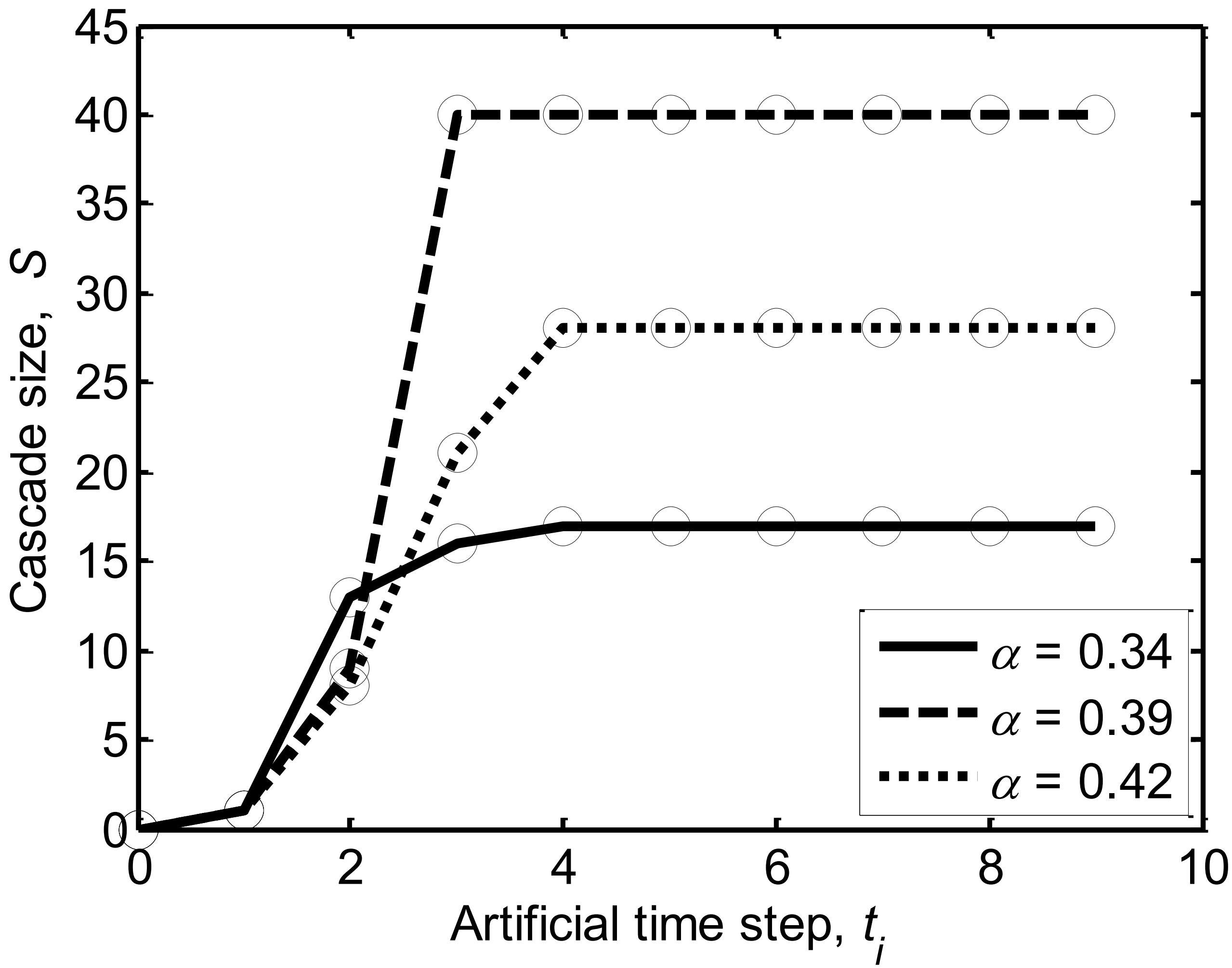
Removal of most congested node: worst case scenario

Islanding: failure of weak nodes splits the network into two isolated sections, reducing the loads and stabilizing the network

$$C_L = 1 - \frac{1}{N_D} \sum_{i=1}^{N_D} N_G^i / N_G$$

Connectivity loss: decrease of the ability of distribution substations to receive power from the generators

Cascade Evolution



Due to islanding effect, increasing the tolerance causes worst effects

Features of Model 1

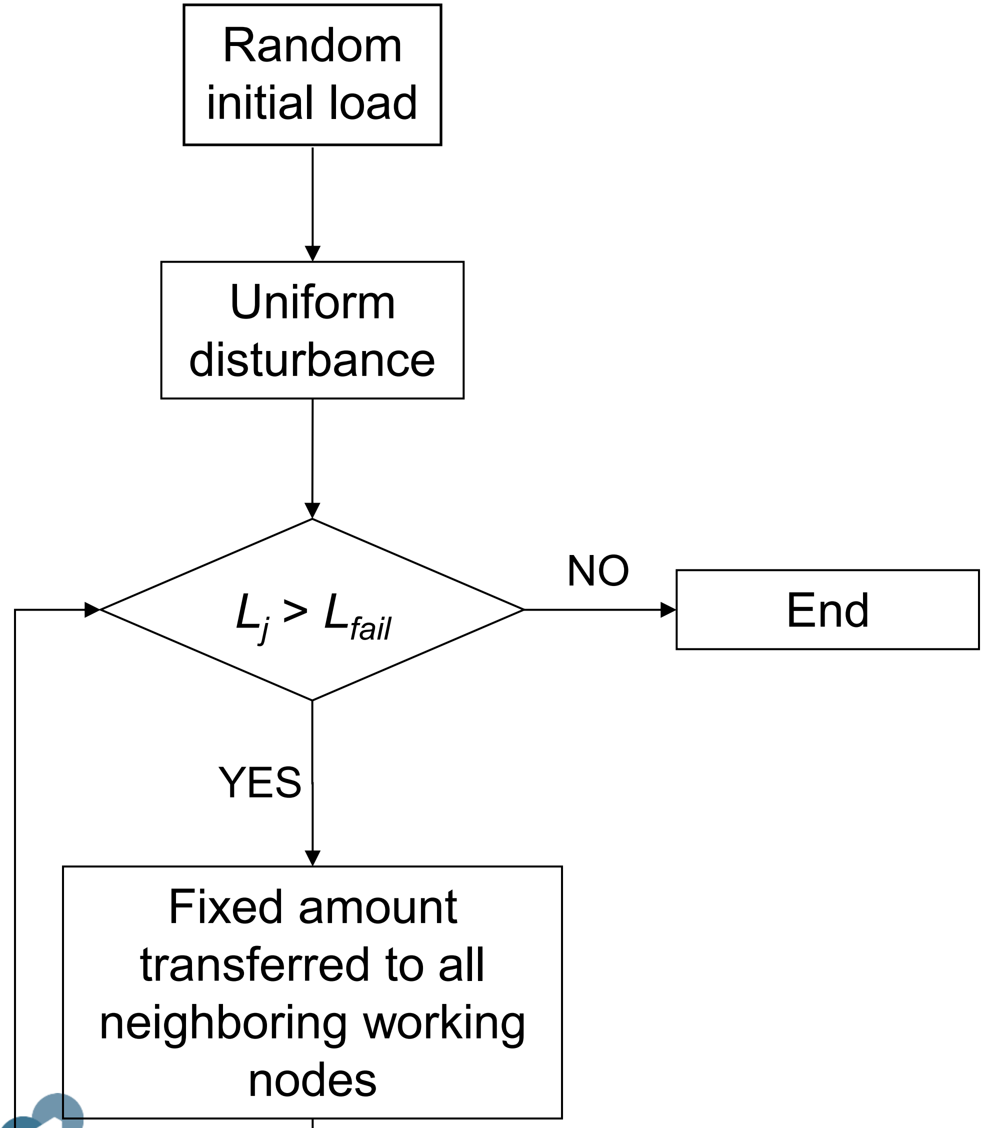
- The service is routed through the shortest path
- Therefore, the initial loading condition is dictated by the network structure (topology) (node specialization is yet possible)
- Long-distance effects as a result of a single propagation step due to the redistribution of shortest paths
- The cascade-triggering event is the removal of one/few component
- Discrete time steps for the propagation of the cascade
- Congestion is modeled by the increasing loading conditions

Applications of Model 1

- Modeling abstract flow-dependent cascades in:
  - power transmission grids
  - communication networks
  - fluid supply networks
  - supply chain networks

Local propagation of a fixed amount of load

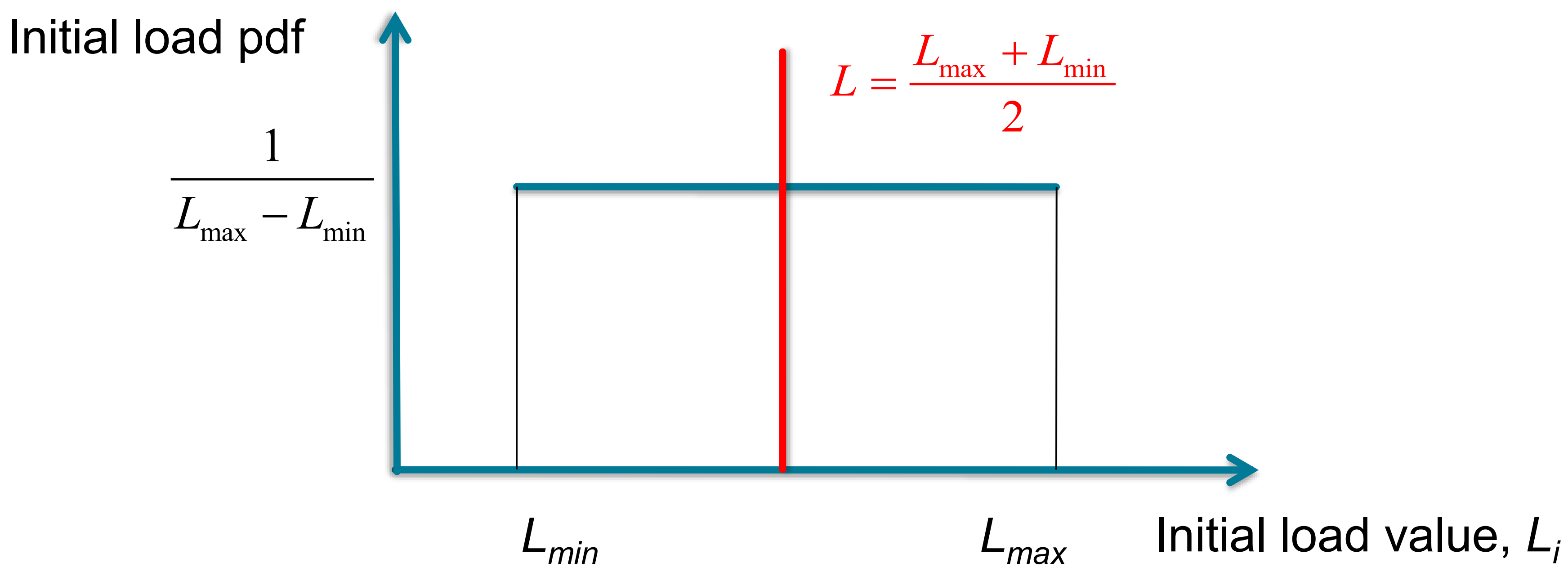
2 - Propagation of a Fixed Amount of Load



- The initial loads of each component are sampled from a uniform random distribution [L_{min}, L_{max}]
- A disturbance is applied to all the components. This may initiate a cascading failure
- The load L_j of every component *j* is compared to the failure load L_{fail}
- If L_j is greater than L_{fail}, the component *j* fails and a fixed amount of load is transferred to the neighboring working nodes (spreading rule)
- The cascading process continues until no new failure occurs

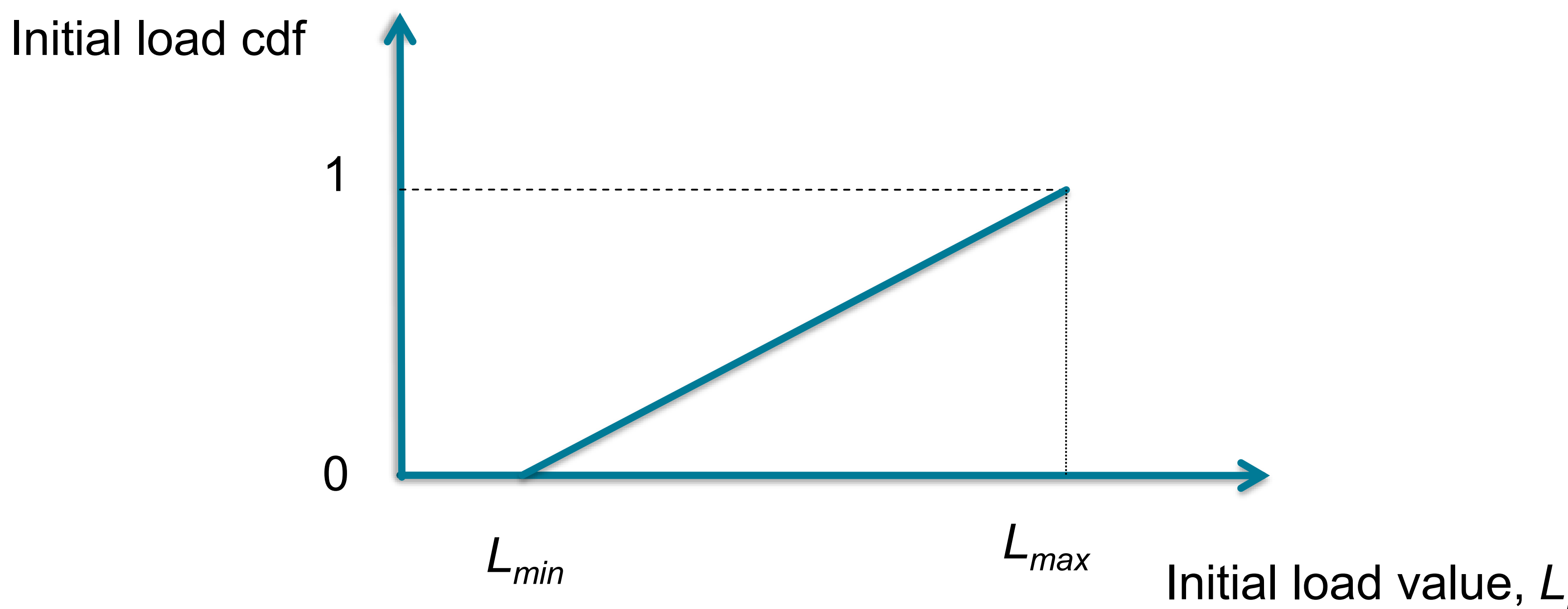
Initial load sampling (I)

- The initial loads are sampled from a uniform probability distribution function (pdf) between $[L_{min}, L_{max}]$
- The sampling process is repeated for every component
- L is the average load of the initial system configuration



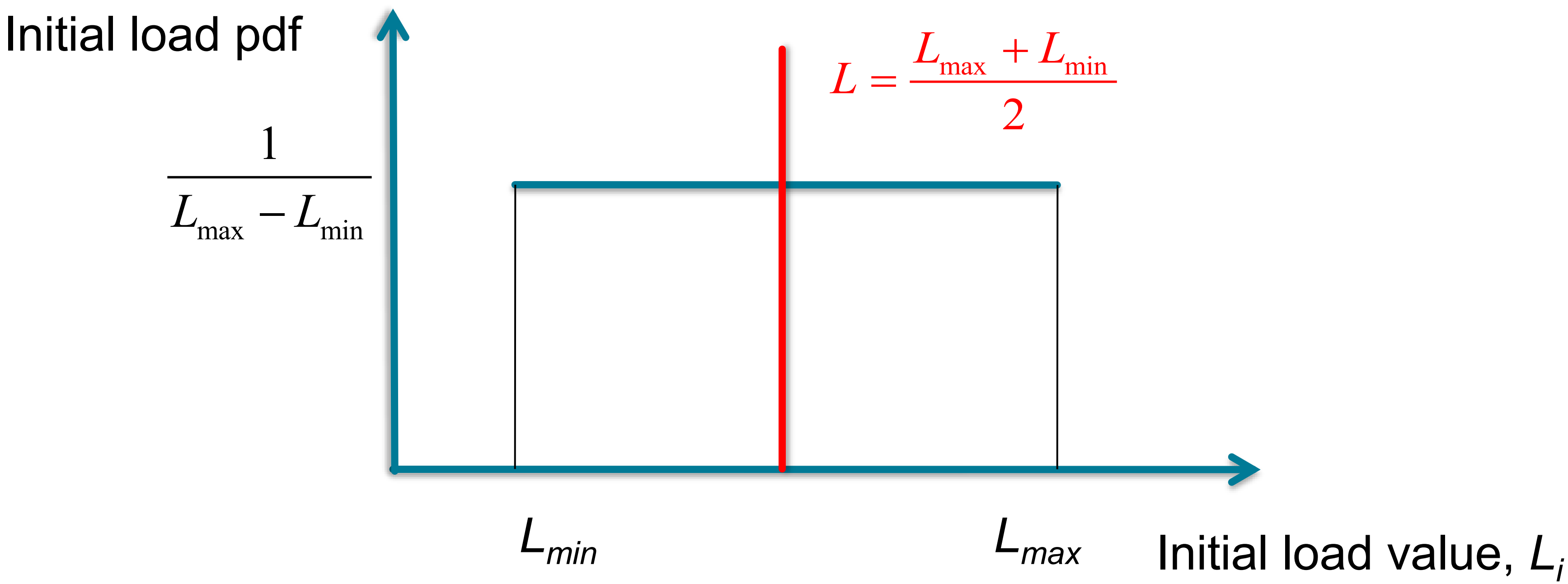
Initial load sampling (II)

- In order to sample from a pdf we have to resort to the correspondent cumulative distribution function (cdf) which is bounded between [L_{min}, L_{max}]



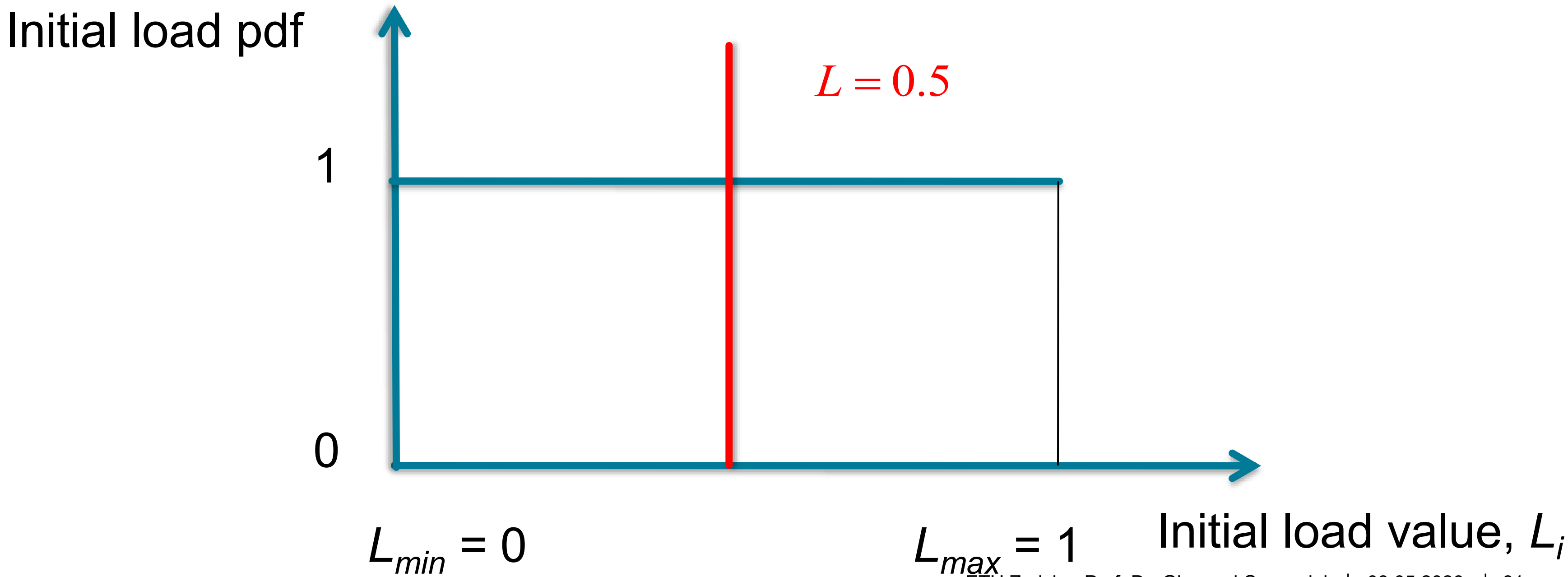
Initial load sampling (III)

- Our goal is to evaluate the robustness of the system against cascade failures
- We want to asses this under varying initial conditions
- We will vary initial condition by varying the initial load pdf



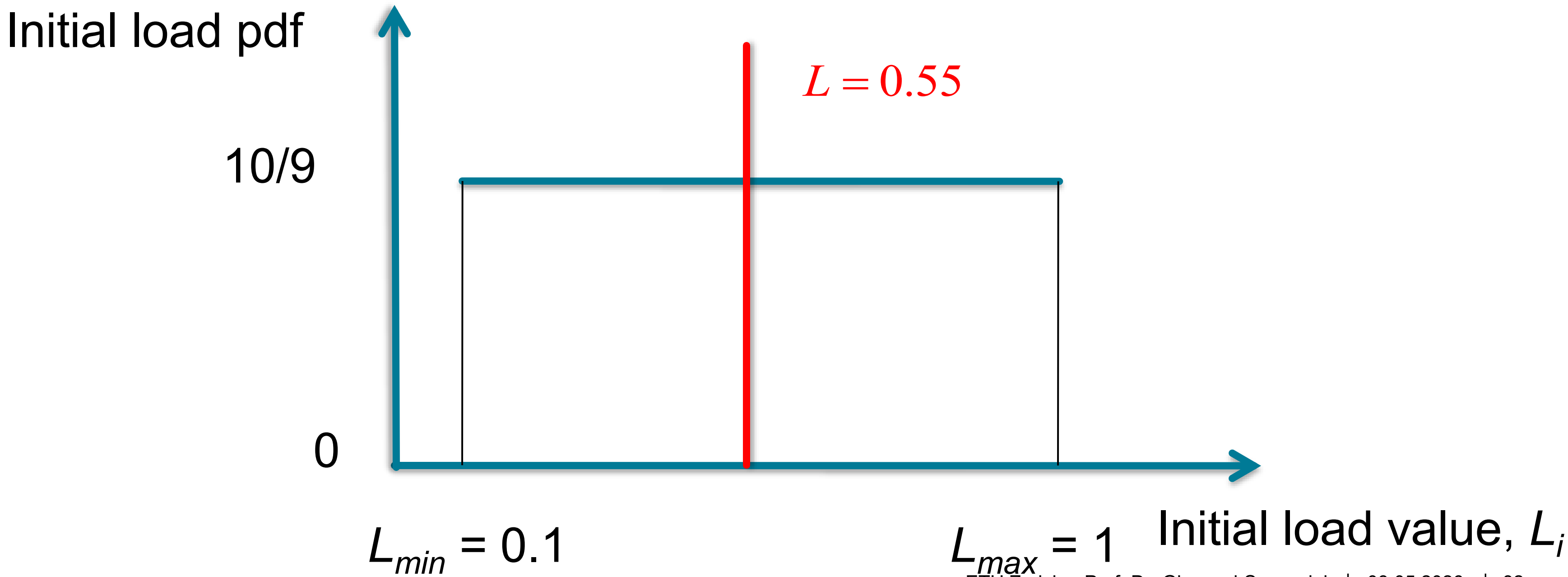
Initial load sampling (IV)

- Our goal is to evaluate the robustness of the system against cascade failures
- We want to asses this under varying initial conditions
- We will vary initial condition by varying the initial load pdf, i.e. by increasing L_{min} from 0 to 1



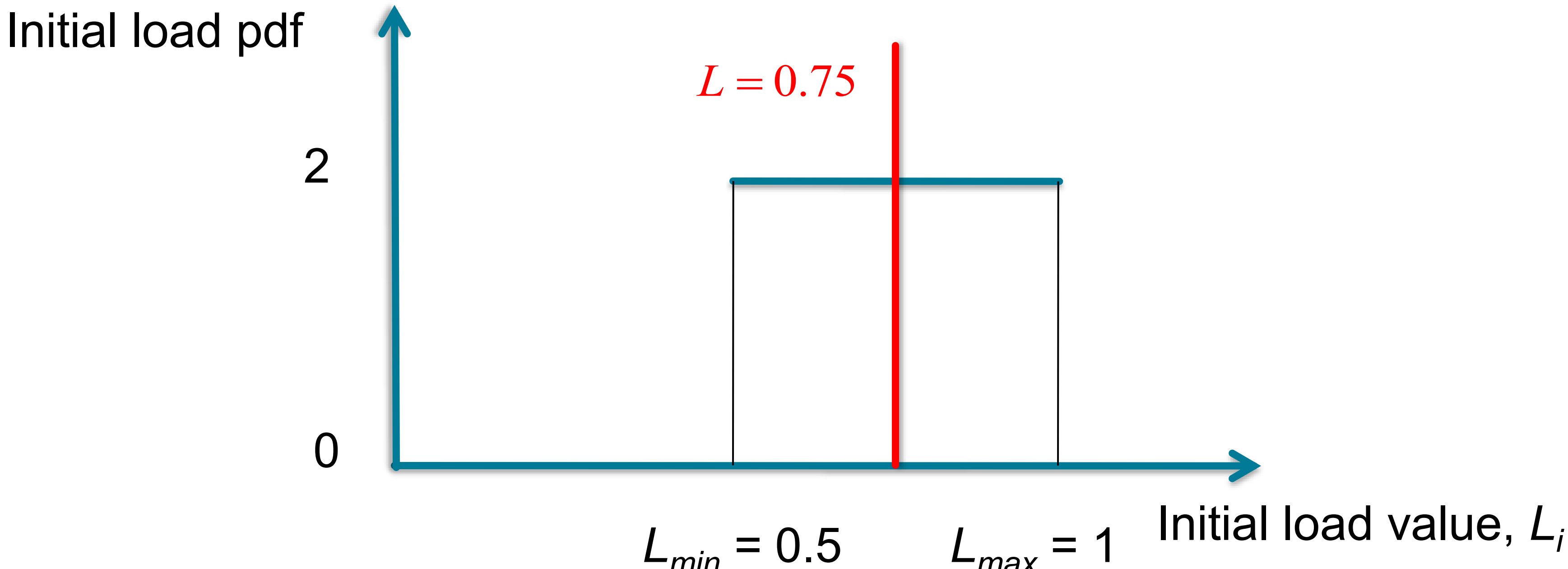
Initial load sampling (IV)

- Our goal is to evaluate the robustness of the system against cascade failures
- We want to asses this under varying initial conditions
- We will vary initial condition by varying the initial load pdf, i.e. by increasing L_{min} from 0 to 1



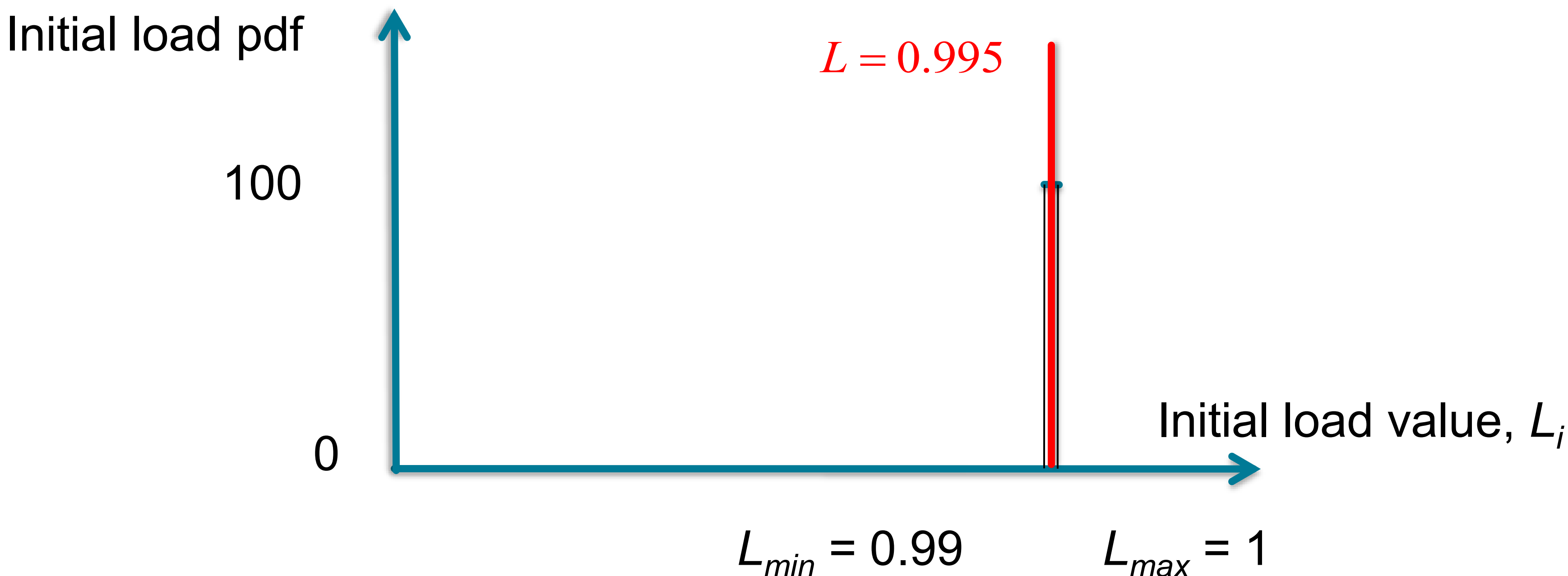
Initial load sampling (IV)

- Our goal is to evaluate the robustness of the system against cascade failures
- We want to asses this under varying initial conditions
- We will vary initial condition by varying the initial load pdf, i.e. by increasing L_{min} from 0 to 1



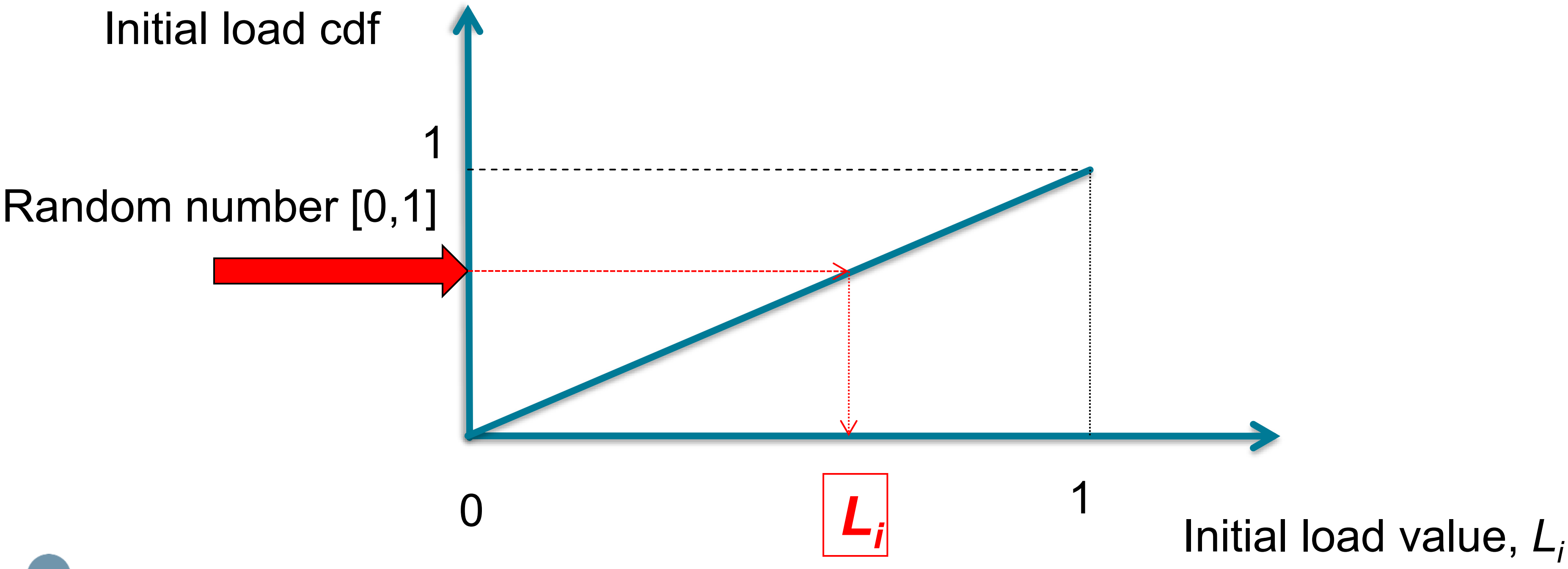
Initial load sampling (IV)

- Our goal is to evaluate the robustness of the system against cascade failures
- We want to asses this under varying initial conditions
- We will vary initial condition by varying the initial load pdf, i.e. by increasing L_{min} from 0 to 1



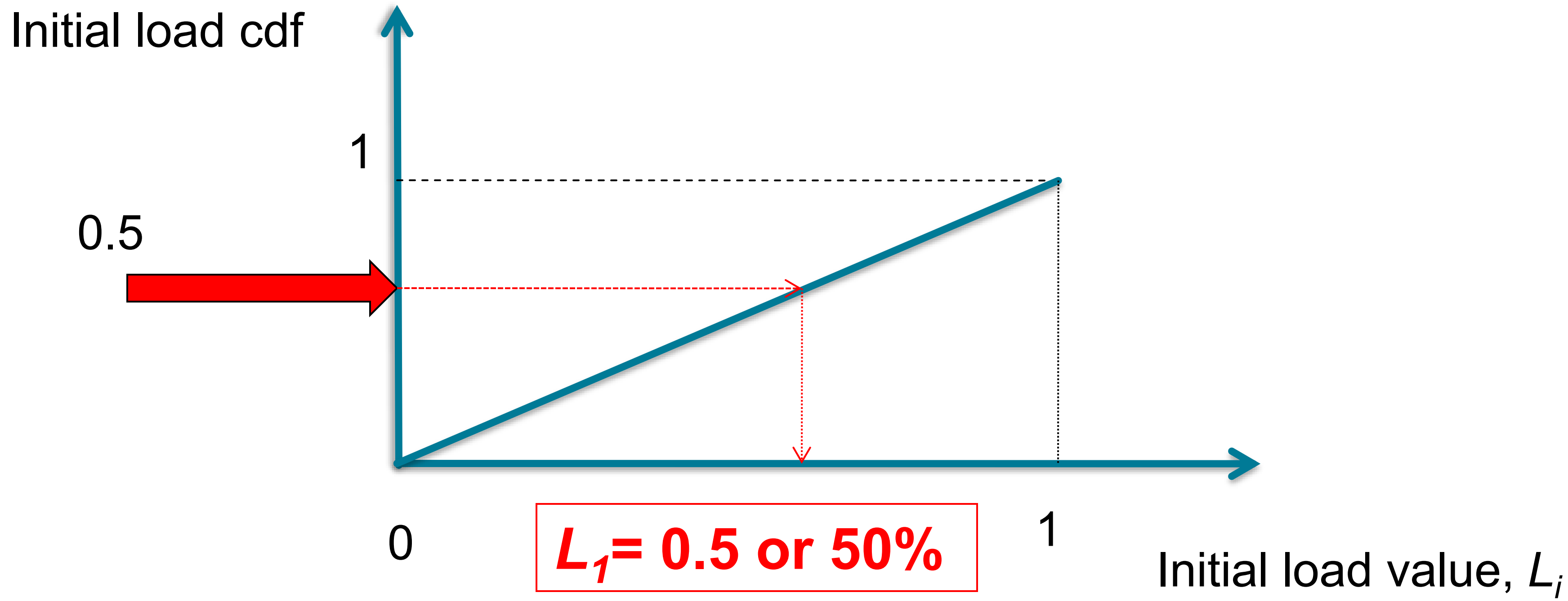
Initial load sampling (V)

- Generate a random number uniformly distributed between [0, 1], in our example case
- Use the inverse transform method to sample
- Trivial for uniform but not trivial for other distributions



Initial load sampling (VI)

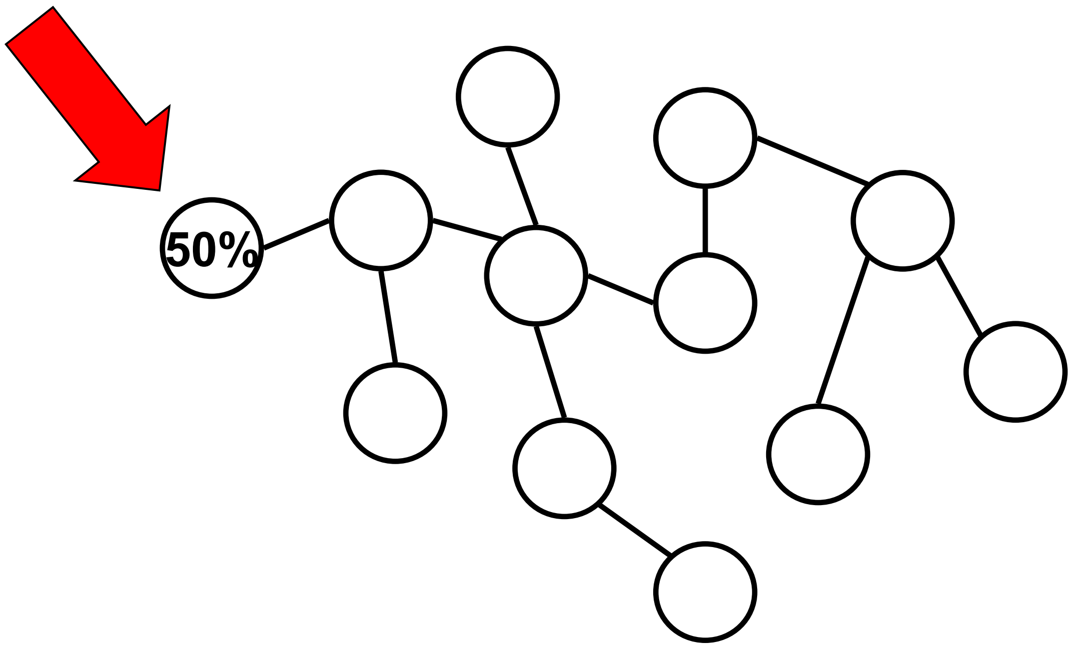
- Component 1:
- Generate a random number between [0, 1]: 0.5
- Apply the inverse transform to get L_1



L_1= 0.5 or 50%

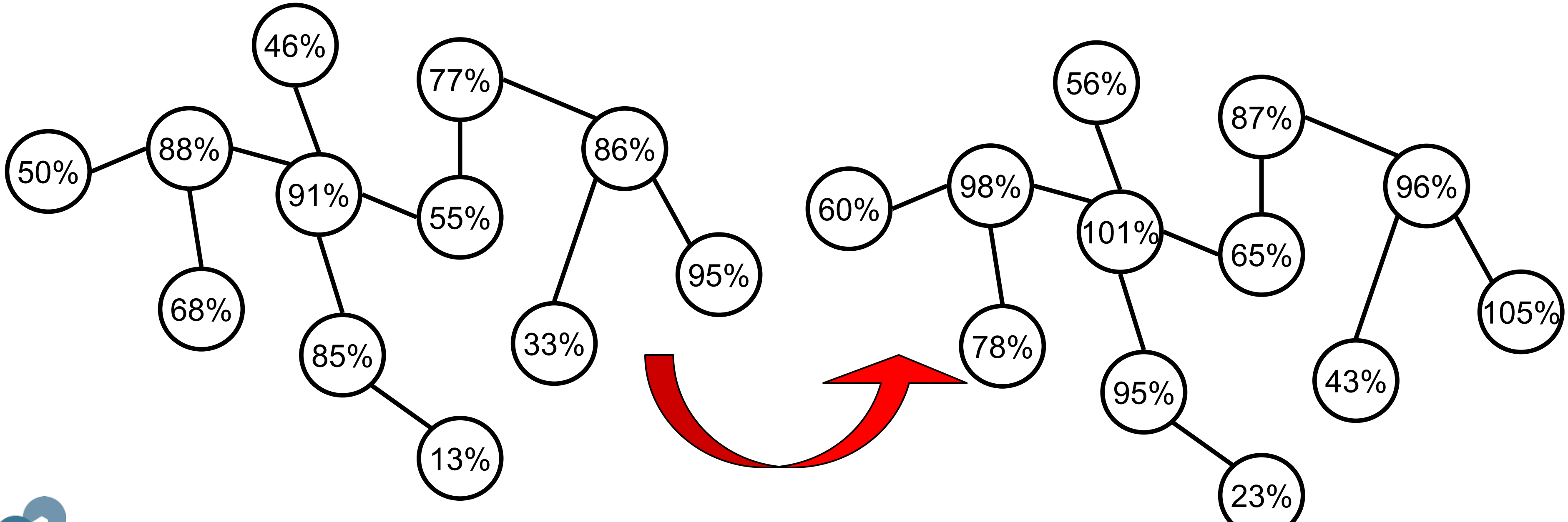
Initial load sampling (VII)

- Component 1:
- Generate a random number between [0, 1]: 0.5
- Apply the inverse transform to get L_1



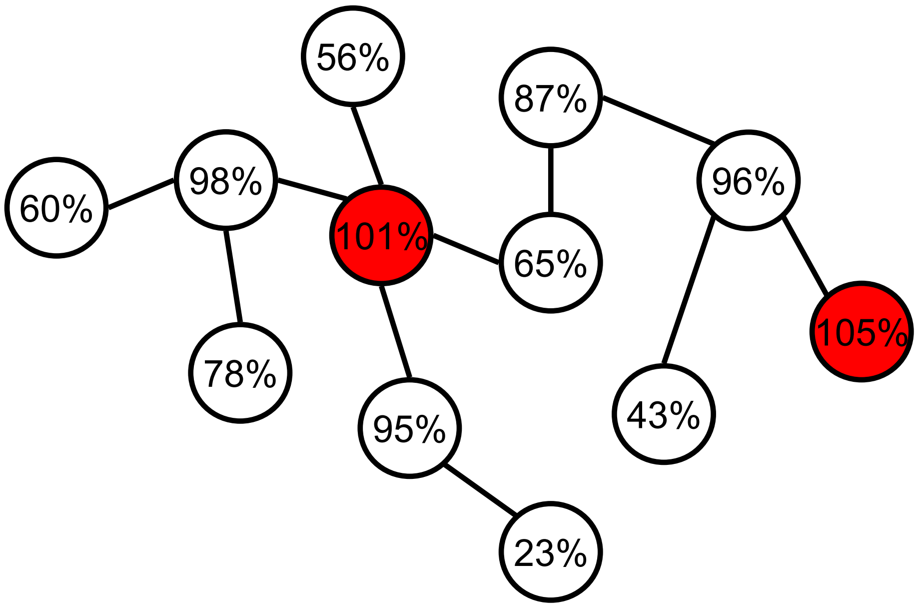
Initial Disturbance

- This step ends after the disturbance *d* has been added to the initial load of every component
- Thus, the initial load has increased from the initial sampling and, possibly, some load has exceeded the failure load L_{fail} = 1
- A failure in a node will then affect the neighboring nodes



Failure Detection (I)

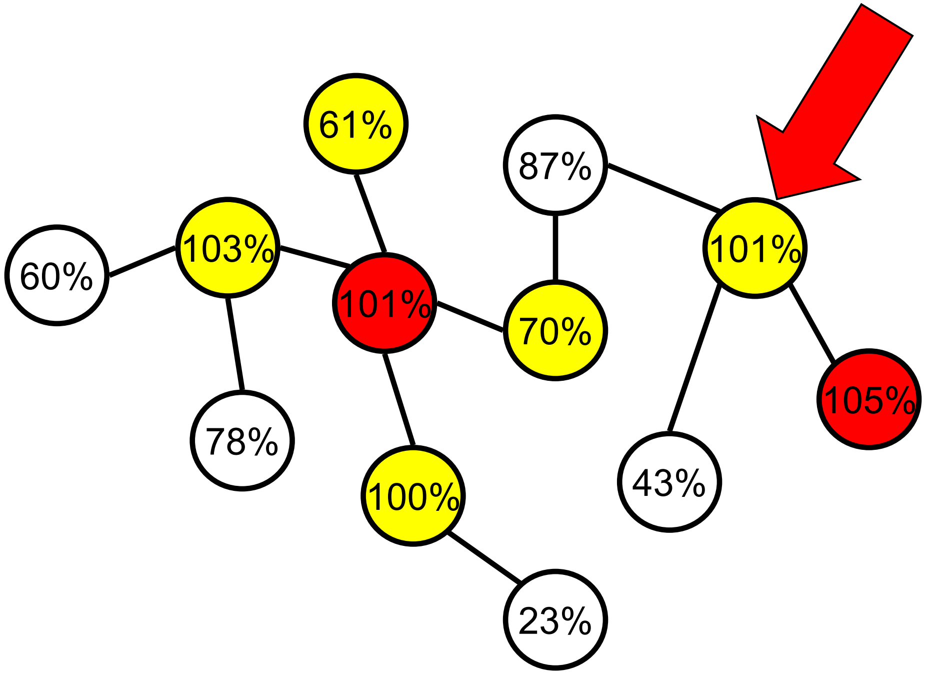
- Situation after the first failure detection:
  - 2 failed components



Failure counter S = S + 1 = 2

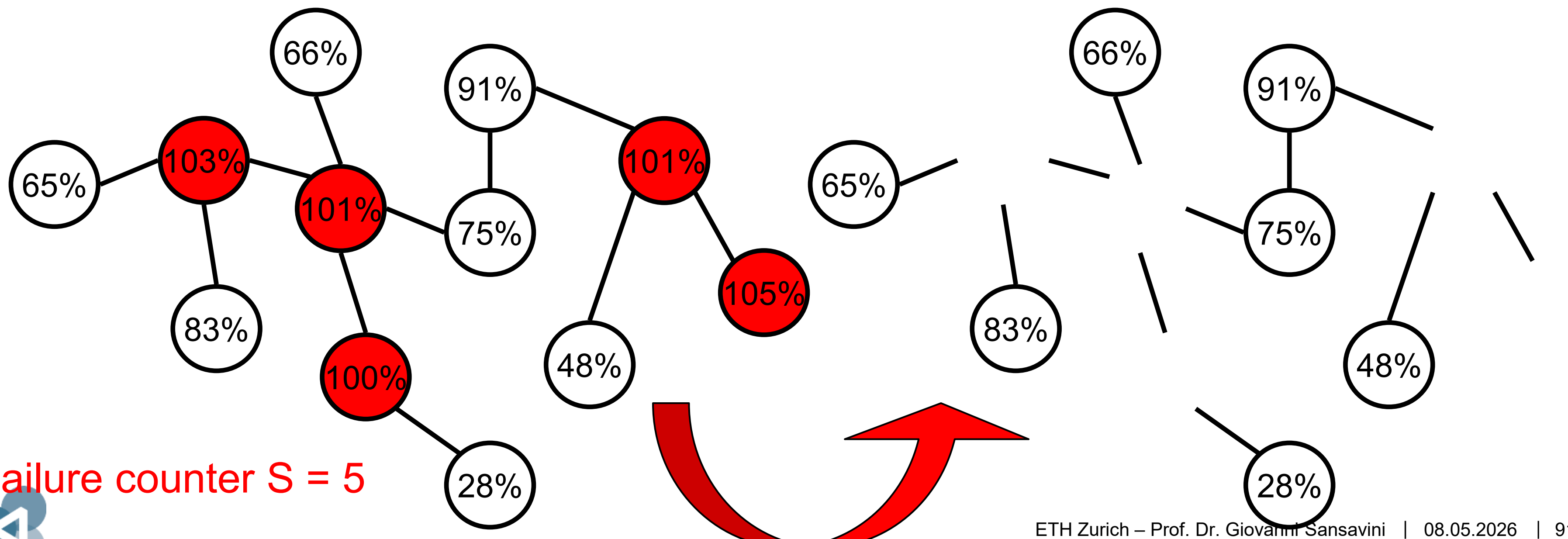
Spreading Rule: Fixed Amount

- <u>2 failed components</u>
- 5 neighboring components, i.e. for component 10:
  $L_{10} = L_{10} + t$, *i.e. 101% = 96% + 5%*



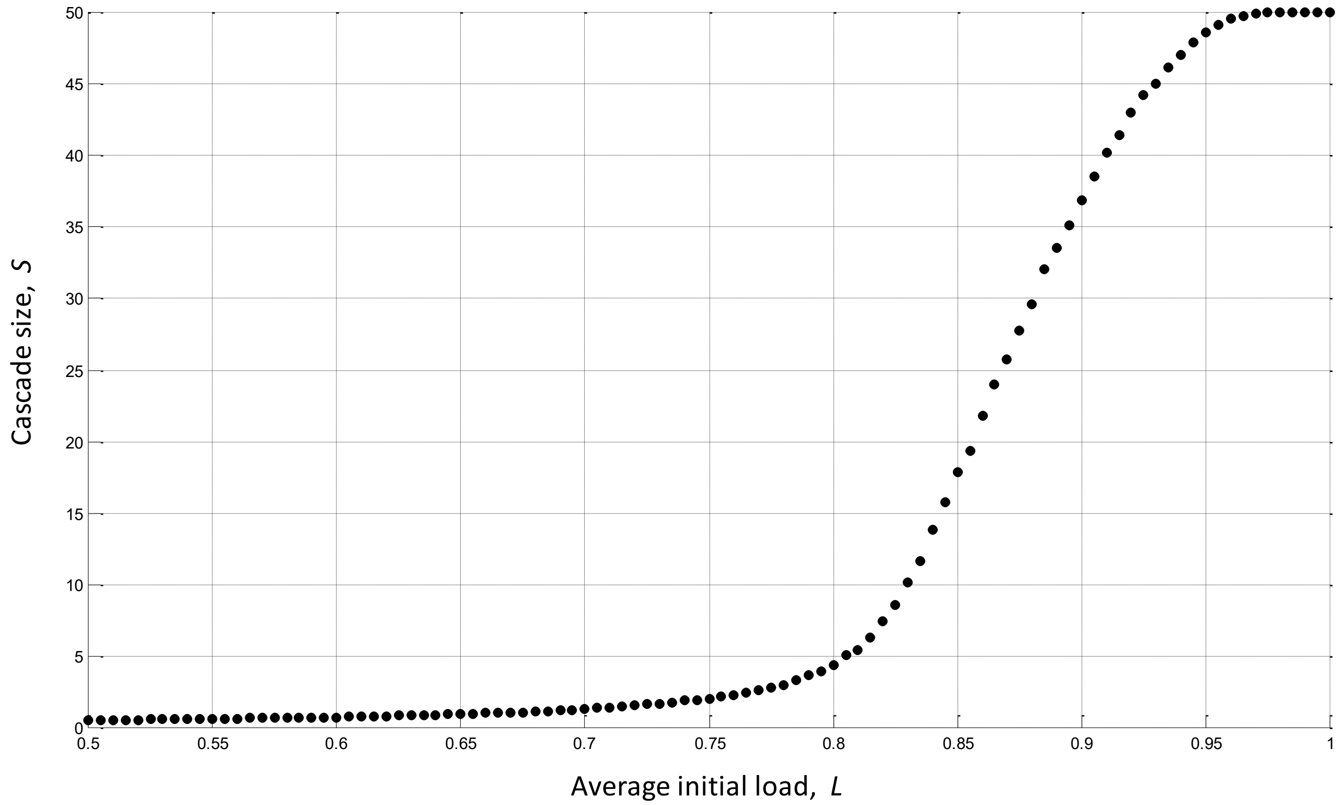
Failure Detection

- The failure detection $L_i > L_{fail}$ is repeated on every component which was not failed in the previous steps
- In this example no new failure is detected and the algorithm stops
- System after the cascade stops:
  - <u>5 failed components</u>



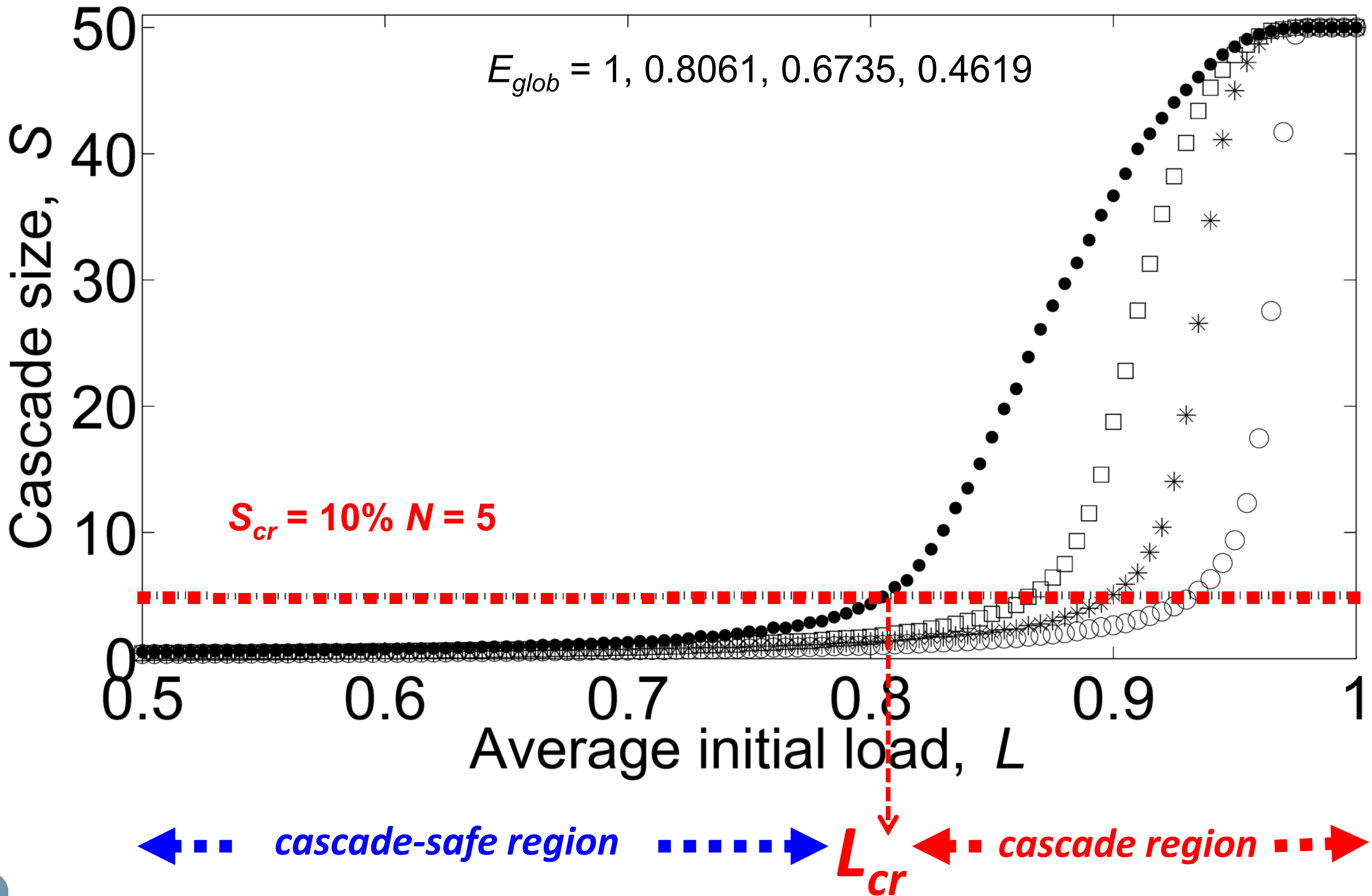
Failure counter S = 5

Presentation of the Results



- Each dot corresponds to an average over 100 simulations with identical initial conditions, i.e. L_i=[L_{min}, 1]
- For each of these simulations, the average load is always the same since it depends on L_i=[L_{min}, 1], i.e. the shape of the pdf but the single load values for each component will be different since they are the result of a random sampling

Identification of Cascade-Safe Operating Margins



Model 2 Features

- Local propagation of a fixed amount of load
- No long-distance effect as a result of a single propagation step
- The cascade-triggering event is applied simultaneously to every component
- The initial loading conditions are randomly sampled from any distribution. They do not result directly from the network structure
- Discrete time steps for the propagation of the cascade

Applications of Model 2

- Modeling probabilistic load-dependent cascades in:
  - power transmission grids
  - coupled power-communication systems
  - coupled power-market systems
  - coupled communication-transportation systems
  - coupled market-market systems
  - virus contamination in computer networks

What you will learn today…

- ✓ Know the characteristics and representation of complex networks by Complex Network Theory
  - ✓ Exponential Networks vs. Scale-Free Networks
  - ✓ Global and Local Properties
- ✓ How to perform the network static structural analysis
  - ✓ Unweighted
  - ✓ Weighted
- ✓ Compute the system vulnerability to element removal
- ✓ Identify the centrality of elements in the structure
  - ✓ Centrality measures, bottlenecks
- ✓ How to Model Cascading Failures Propagation in Network Systems
  - ✓ How to Identify Cascade-safe Operating Margins