HELSINKI JUNE 2008

Project contract no. 043363

MANMADE

Diagnosing vulnerability, emergent phenomena and volatility in man-made networks

Work Package 6 – D6.1

Vulnerability of interconnected networks

D6. 1 A method to calculate interoperability matrices

Revision 1

Macedonian academy of sciences and arts,
Bul. Krste Misirkov 2, 1000 Skopje, R Macedonia

CONTENTS

HELSINKI JUNE 2008

- Executive summary
- Tasks – defining vulnerability, spectral analysis, attack simulations, influence model, flow model
- Results
- Conclusions

EXECUTIVE SUMMARY

HELSINKI JUNE 2008

- A detailed extensive study over the theoretical background
- Defining frame for further activities – the influence model accepted as the most suitable for describing interdependencies
- Spectral analysis
- Studying vulnerability of generic and manmade networks
- Flow model – LP algorithm
- Summary and conclusions

Vulnerability – the concept

Definitions

- Einarsson and Rausand (industry) – the properties of an industrial system that may weaken or limit its ability to endure threats and survive accidental events that originate both within and outside the system boundaries
- Berdica (transportation) – a susceptibility to incidents that can result in considerable reductions in road network serviceability
- Morakis et al. – measure of the exploitability of a weakness
- Barefoot et al. – a lack of resistance of the graph to the deletion of vertices and edges

Vulnerability – the concept

Robustness and resilience – complement to vulnerability

- Robustness – the ability of the system to retain its structure (function) intact when exposed to perturbations
- Resilience – the ability of the system to adapt to regain a new stable position (to recover) after perturbations
- Hansson and Helgesson – robustness can be treated as a special case of resilience (the recovery time equals zero)

Vulnerability – the concept

Network vulnerability

- Quantitative measures – connectivity and efficiency

$$v_{loc} = E(G) - E(G \setminus \{i\})$$

$$v_{glob} = \frac{1}{N} \sum_{i \in G} \left| E(G) - E(G \setminus \{i\}) \right|$$

v_{loc} – local vulnerability
v_g – global efficiency
E – efficiency
N – number of nodes

- An axiomatic approach
  - invariance under isomorphism
  - normalization
  - computational costs – polynomial time
  - scale invariance

$$v^* = \exp\left\{ \frac{M - m}{N} + N - L + \frac{2}{N} \right\}$$

v^* - network vulnerability
M, m – the maximum and minimum degree
L – number of links

HELSINKI JUNE 2008

Vulnerability – the concept

Static and dynamic robustness

- Static robustness – deleting nodes without need to redistribute the quantity transferred through the networks
- Dynamic robustness – redistribution of flows after a removal of nodes and/or lines

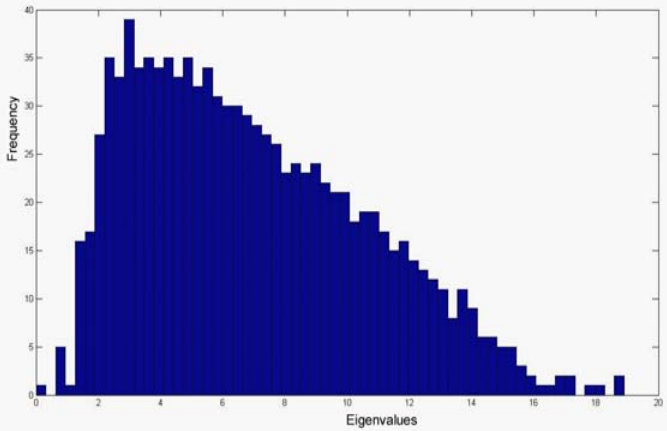
Spectral analysis of networks

Introduction

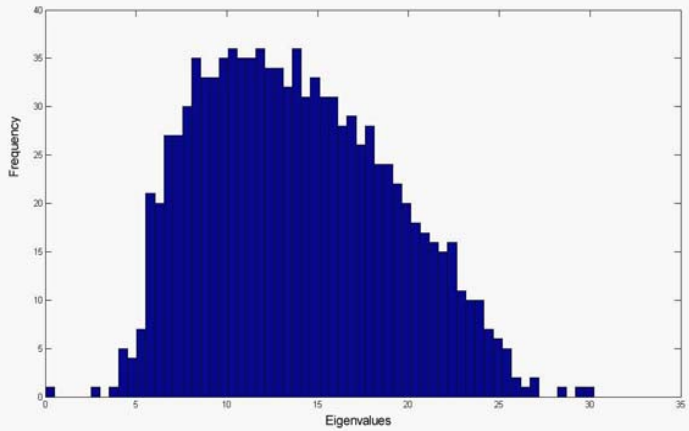
- Adjacency matrix
- Laplacian $L = D - A$
- Normalized Laplacian $I - D^{-1/2}AD^{-1/2}$

Spectral analysis of networks

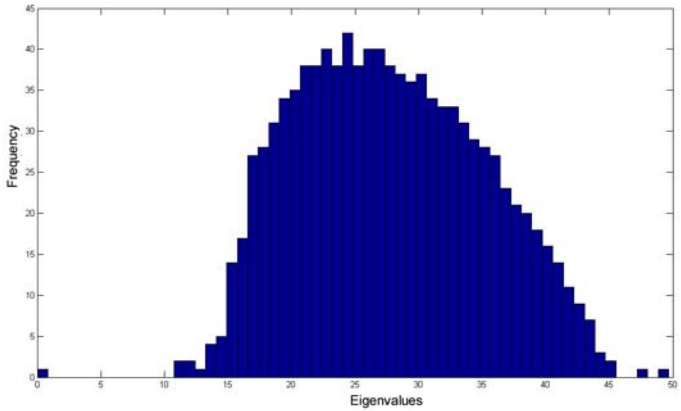
Laplacian spectra of generic graphs – Random Erdos-Renyi graph (*N*=1024)



a) Frequency vs. eigenvalues ($p=p_c$)



b) Frequency vs. eigenvalues ($p=2p_c$)

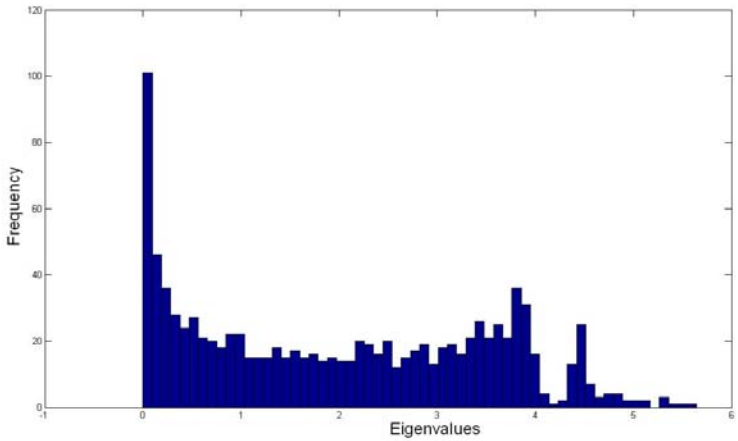


c) Frequency vs. eigenvalues ($p=4p_c$)

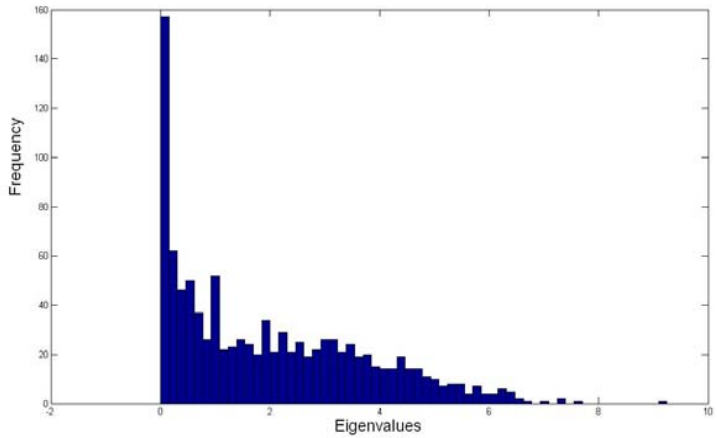
Spectral analysis of networks

HELSINKI JUNE 2008

Laplacian spectra of generic graphs – small-world Watts-Strogatz graph (*N*=1024, k=1)



a) Frequency vs. eigenvalues (*p*=0.1)

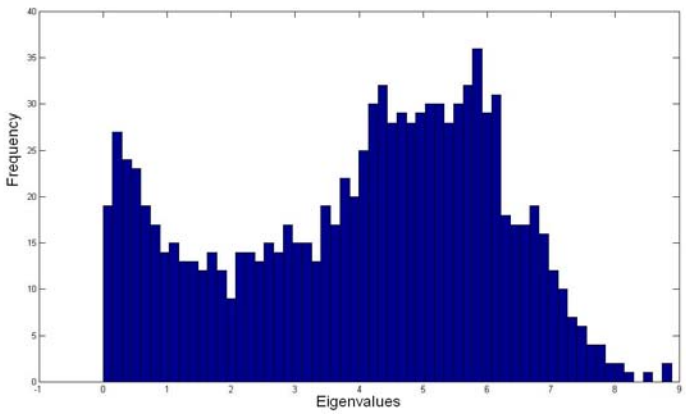


b) Frequency vs. eigenvalues (*p*=0.5)

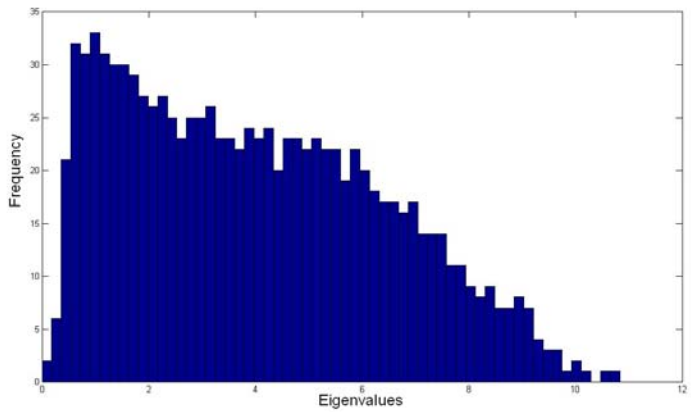
Spectral analysis of networks

HELSINKI JUNE 2008

Laplacian spectra of generic graphs – small-world Watts-Strogatz graph (*N*=1024, k=2)



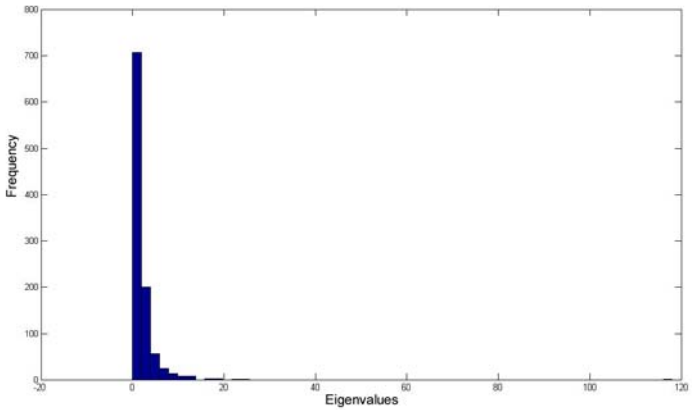
a) Frequency vs. eigenvalues (*p*=0.1)



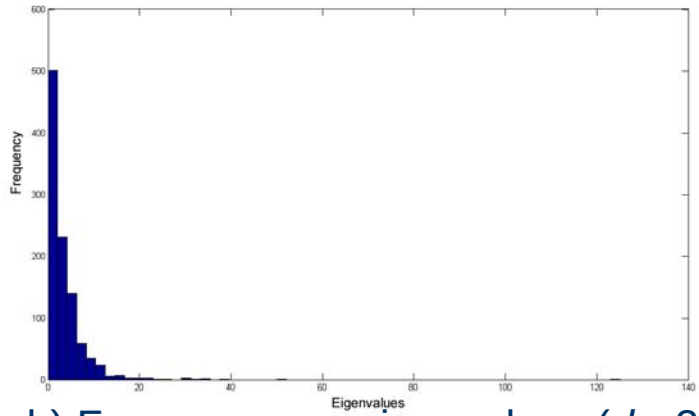
b) Frequency vs. eigenvalues (*p*=0.5)

Spectral analysis of networks

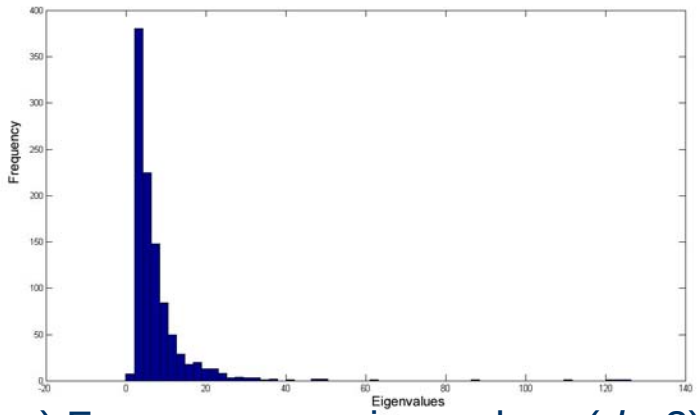
Laplacian spectra of generic graphs – scale-free Barabasi - Albert graph (N=1024)



a) Frequency vs. eigenvalues (d_0=1)



b) Frequency vs. eigenvalues (d_0=2)

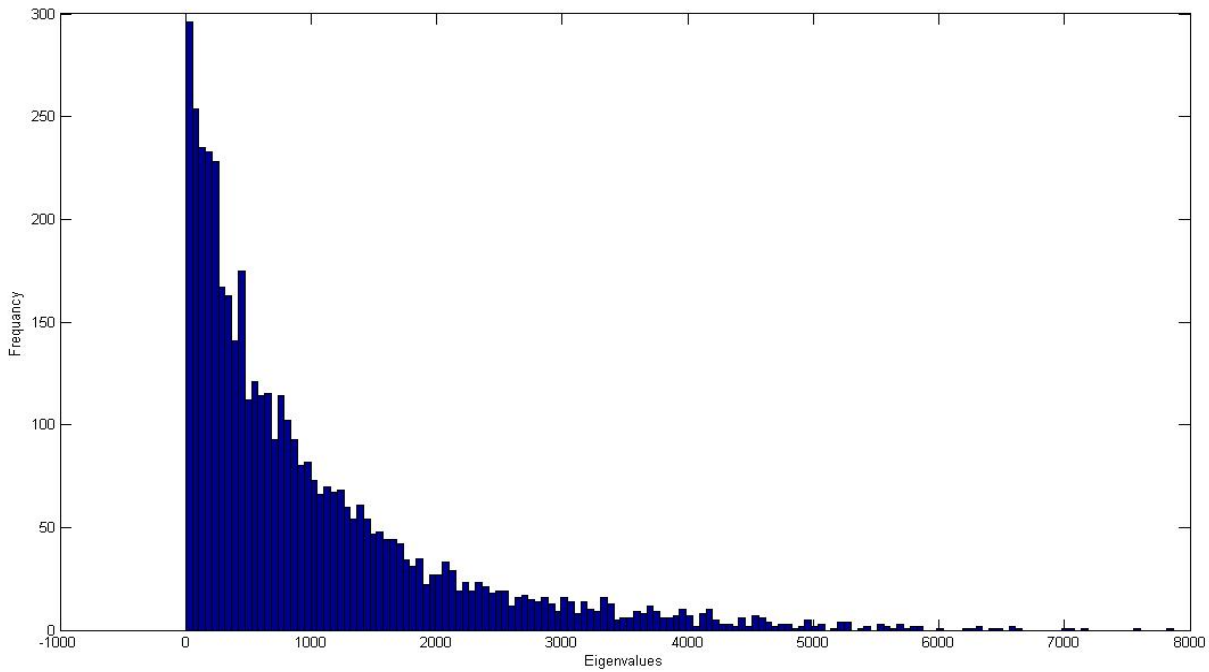


c) Frequency vs. eigenvalues (d_0=2)

Spectral analysis of networks

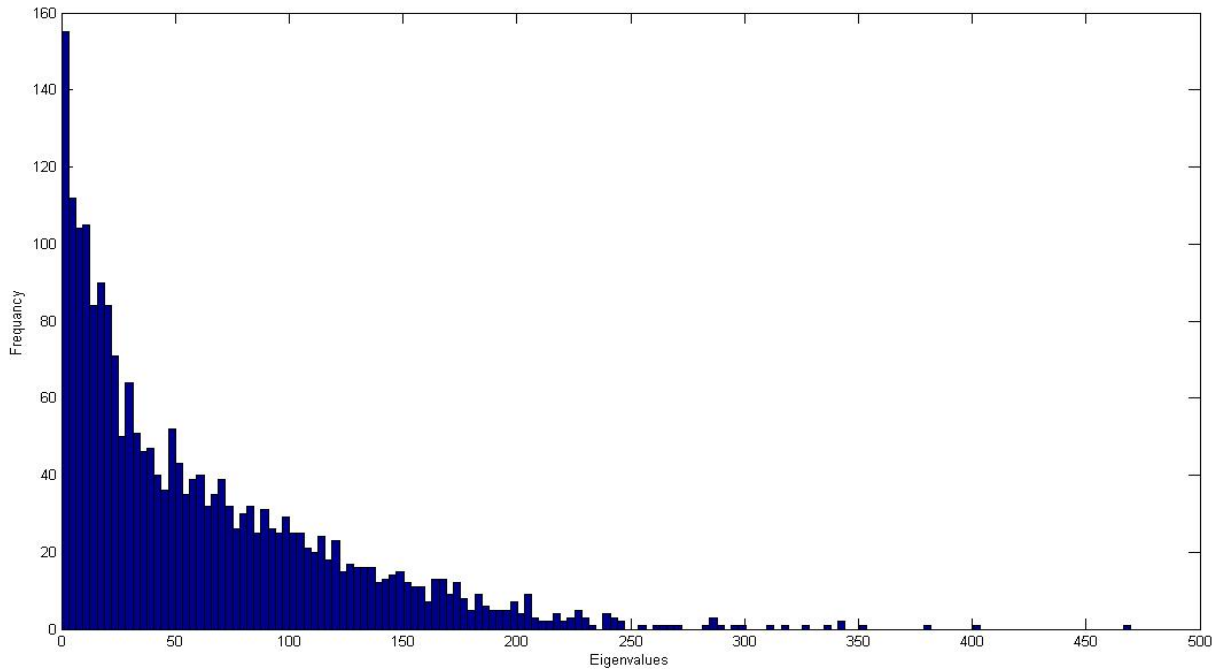
HELSINKI JUNE 2008

Laplacian spectra of manmade networks – EU power grid (undirected weighted graph)



Spectral analysis of networks

Laplacian spectra of manmade networks – EU gas network



Vulnerability of MANMADE networks

Introduction

- Static tolerance to attacks (preferential removal of nodes)
- Methods to assess busyness of nodes and lines
  - betweenness centrality
  - modal weight

$$w_i = \sum_{j=1}^{N} \left| \gamma_{ij} \right|$$

$$l_{i,j} = \sum_{k=1}^{N} \left| \gamma_{ik} - \gamma_{jk} \right|$$

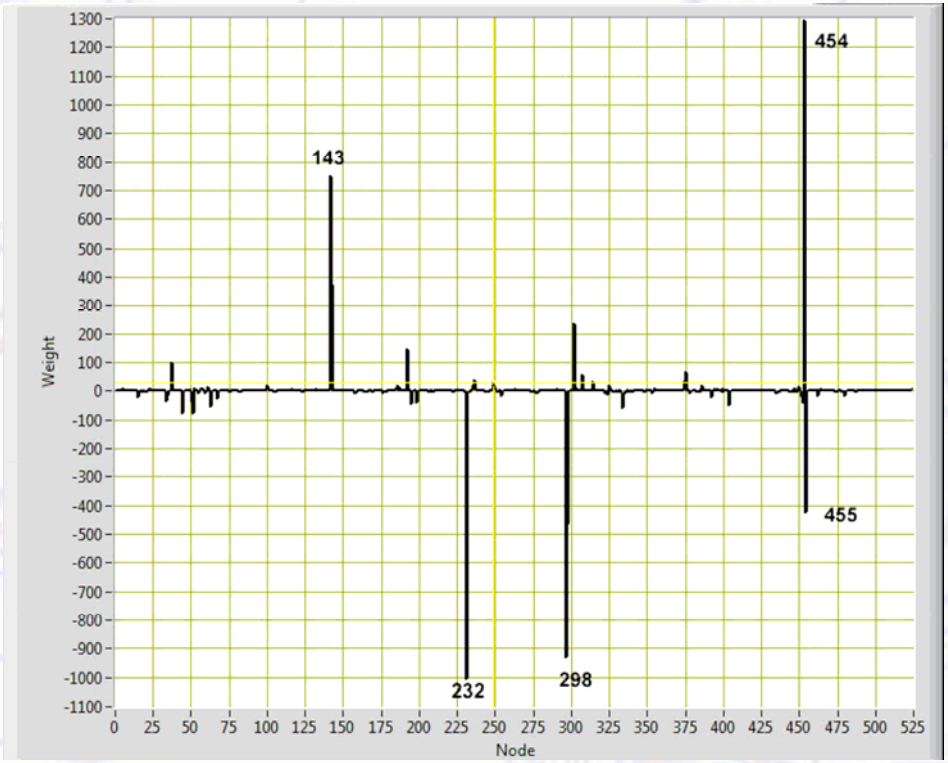
w_i – modal weight of the node *i*
$l_{i,j}$ – modal weight of the line *ij*
γ – entries of the modal connectivity matrix Γ

$\Gamma = \mathbf{L'\Phi}$, **L'** - transposed Laplacian, Φ - eigenvectors of the Laplacian

Vulnerability of MANMADE networks

HELSINKI JUNE 2008

Nodal distribution of the modal weight for one of the possible modes obtained by the modal analysis



Vulnerability of MANMADE networks

Attack vulnerability - simulation details

- Different topologies were investigated
  - Random graphs (Erdos – Renyi model)
  - Scale – free (Barabasi – Albert)
  - Manmade - a segment of the European power grid
- Attack strategy – nodes deletion according to their
  - Degree
  - Betweenness centrality
  - Modal weight
- Adaptive and non-adaptive strategy

HELSINKI JUNE 2008

Vulnerability of MANMADE networks

Attack vulnerability - simulation details

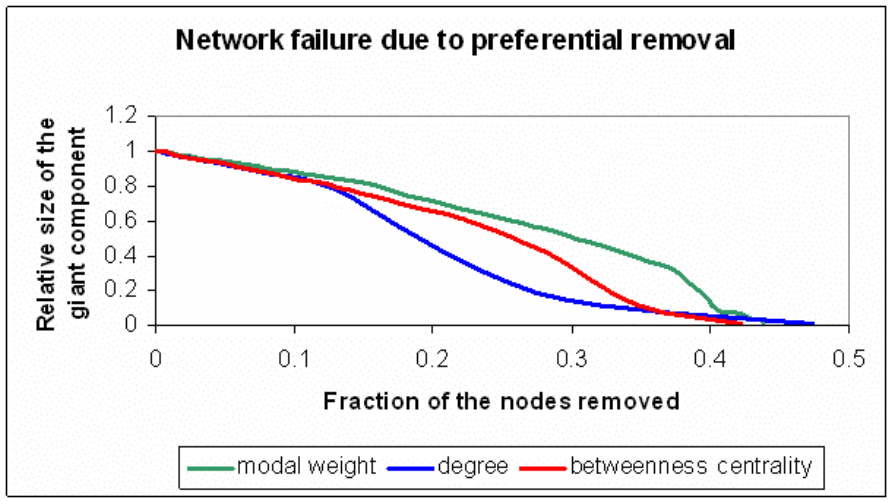
- Properties analyzed – dependence of the
  - network fragmentation (number of isolated islands)
  - relative size of the giant component
  - Diameter

on the fraction of nodes that are removed

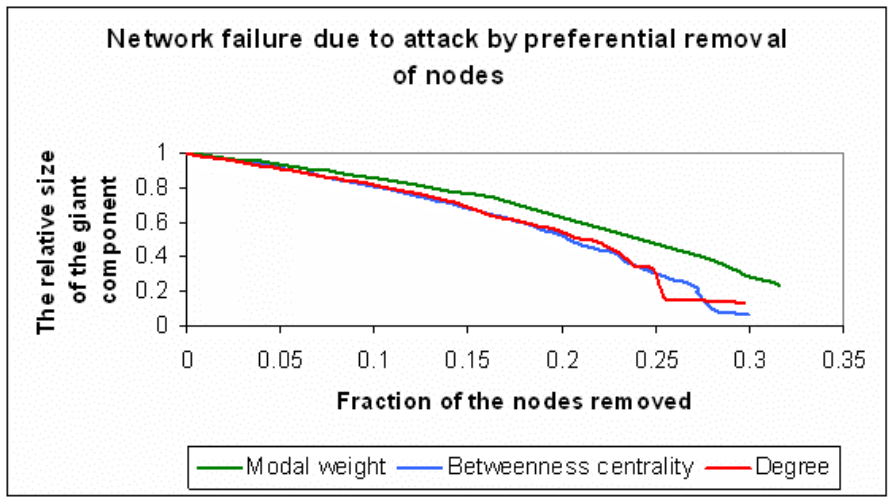
Vulnerability of MANMADE networks

Attack vulnerability – results (non-adaptive strategy)

HELSINKI JUNE 2008



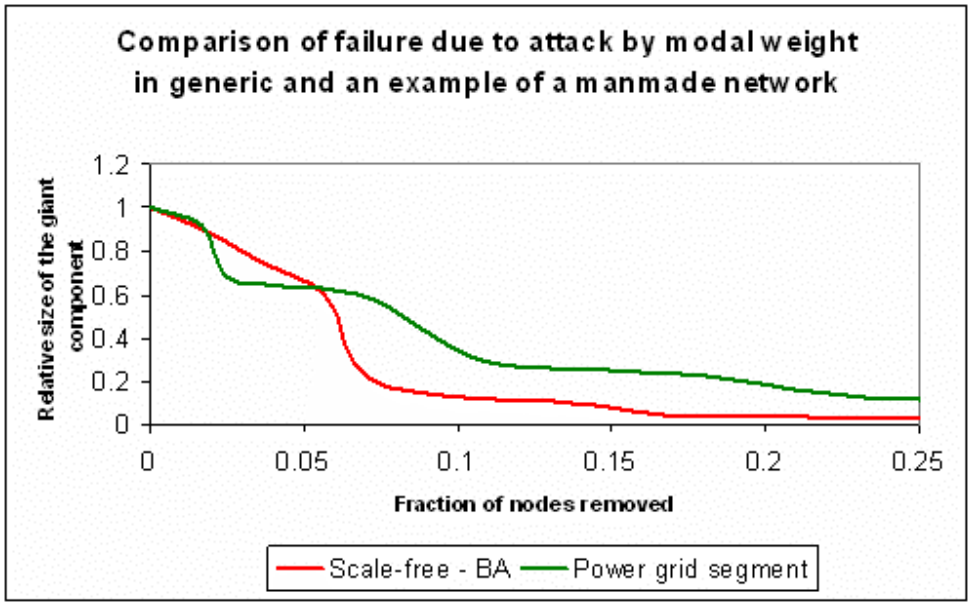
a) The dependence of the relative size of the giant component on the fraction of removed nodes for a random ER graph with 1000 nodes and average degree 3.5



b) The dependence of the relative size of the giant component on the fraction of removed nodes for a random SF graph with 1000 nodes and average degree 3.5

Vulnerability of MANMADE networks

Attack vulnerability – results (non-adaptive strategy)

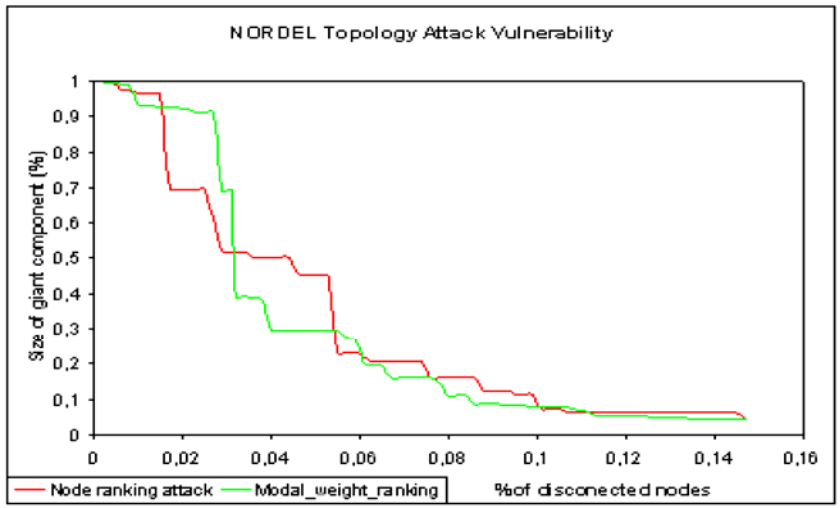


- Attack vulnerability of a manmade network compared to an adequate generic BA graph, both networks has 524 nodes, average degree 3.5 and similar degree distribution

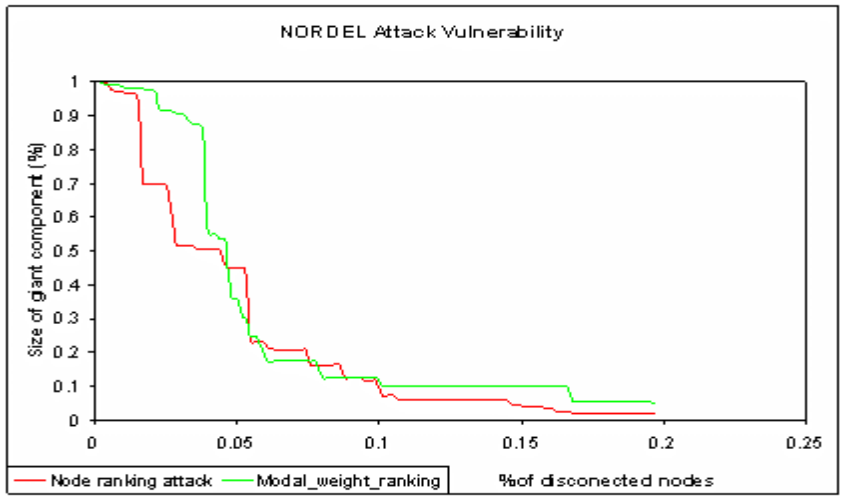
HELSINKI JUNE 2008

Vulnerability of MANMADE networks

Attack vulnerability – results (adaptive strategy)



a) Attack vulnerability of a manmade network, simulated by an adaptive strategy, for unweighted graph, having 524 nodes and average degree of 2.4

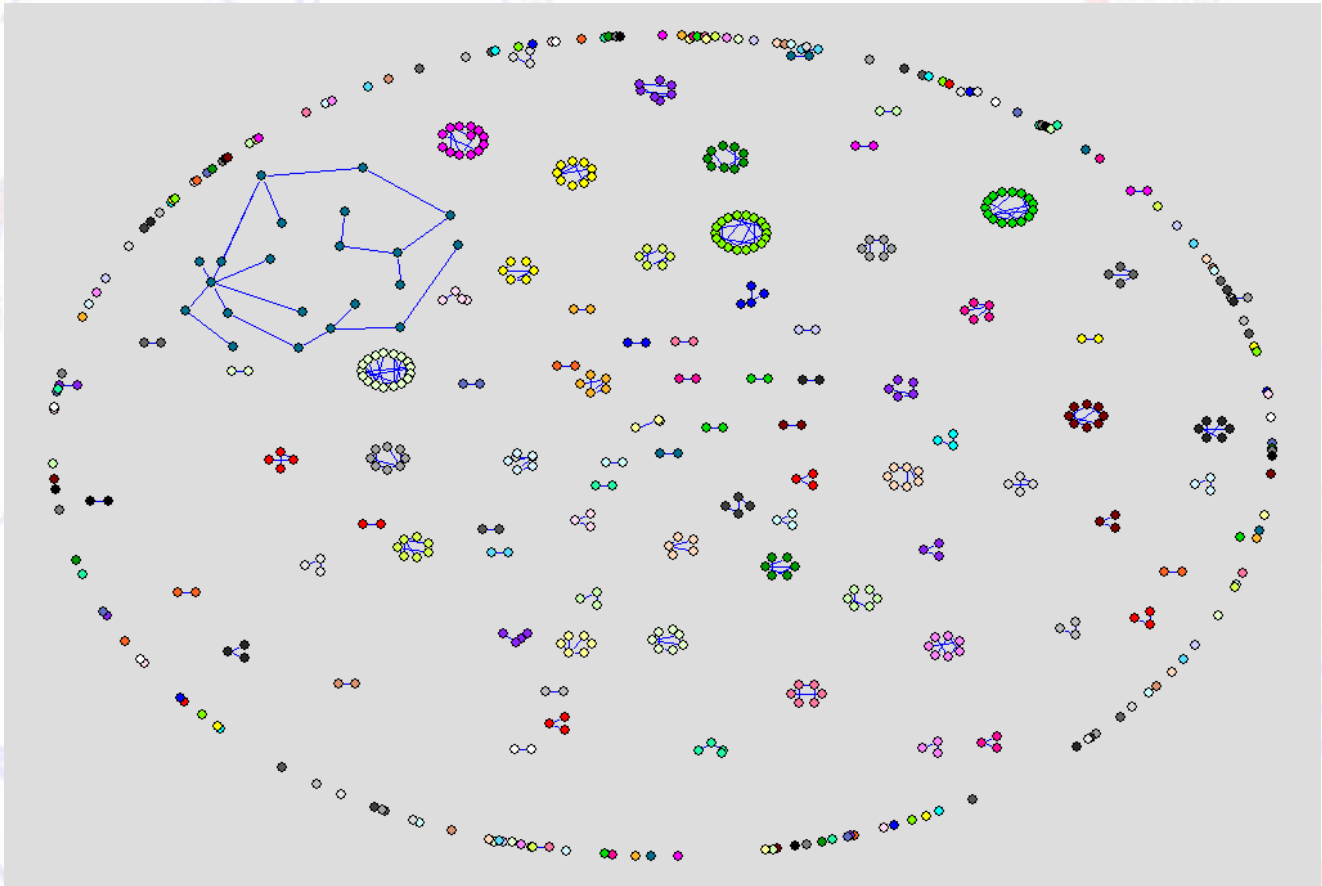


b) Attack vulnerability of a manmade network, simulated by an adaptive strategy, for weighted graph, having 524 nodes and average degree of 2.4

HELSINKI JUNE 2008

Vulnerability of MANMADE networks

Attack vulnerability – a snapshot of the simulation



HELSINKI JUNE 2008

Vulnerability of MANMADE networks

Attack vulnerability – conclusions

- Modal analysis was applied to asses nodes busyness
- The results were compared with standard topological node assessment tools (node degree and betweenness centrality)
- The analyses of the attack vulnerability of generic networks (ER, BA) suggested that node removal according to modal weight is not an efficient strategy for network disintegration, compared to standard node deletion methods
- When applied to manmade networks (segments of the EU power grid) node ranking based on modal analysis, proved to be an efficient strategy for network disintegration

LP Power Flow Model

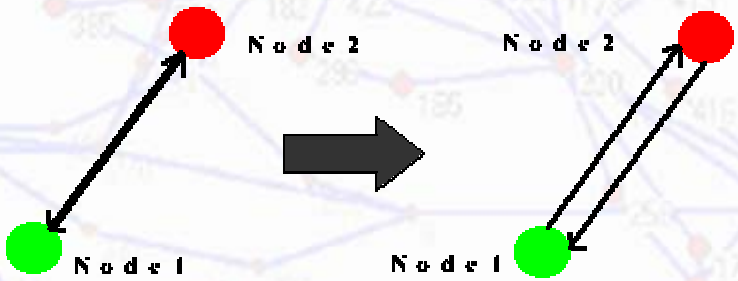
Introduction

- The simplest method of calculating flow trough a network, requiring only generation and demand of each node, as well as lines capacities, as input parameters;
- Applicable to all types of networks (electricity, gas, transportation, internet) with minor changes in the constrains;
- Based on the simplex method: given the constrains, the flow is calculated on the bases of optimizing (maximizing) a certain function called **Objective Function**.

LP Power Flow Model

The Model

- Three types of nodes:
  - Generation Node
  - Demand Node
  - Transmission Node.
- Lines:
  - All lines in the model are considered as bi-directional;
  - However, the definition of the simplex method in linear is based on presumption that all parameters in the model are positive. In our case that means that the lines should be unidirectional.
  - Solution:



Each bidirectional line is substituted with 2 unidirectional

LP Power Flow Model

Constraints

- **Lines**:

$$0 \le x_i, x_j \le C$$

$$\left| x_i - x_j \right| \le C$$

x_i , x_j - the flows trough directed lines, that form a bidirectional line with capacity C

- **Generation nodes**

$$\sum x_i \le G$$

x_i connecting the observed node

G – Node generation capacity

- **Demand nodes**

$$\sum x_i \le D$$

x_i connecting the observed node

D – Node demand

LP Power Flow Model

Constraints

- **Transmission nodes**

$$\sum x_i = 0$$

x_i connecting the observed node

- **Objective function: sum of the product of demand nodes consumption and priority index:**

$$Z = \sum_{\text{All Demand Nodes}} P_{node} C_{node}$$

P_{node} - the priority of certain node,

C_{node} - node consumption

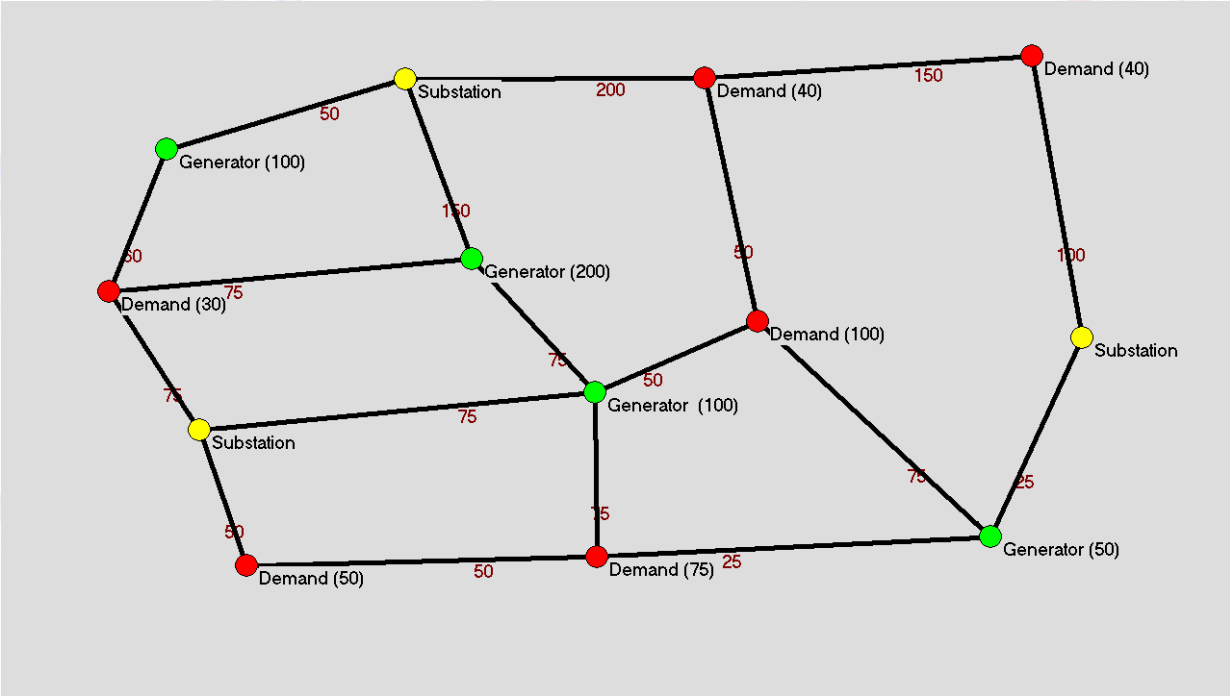
LP Power Flow Model

Vulnerability analysis of a power grid

- Attacks simulated by preferential removal of lines;
- Adaptive strategy: after each step parameters recalculated and the line with highest ranking, according the accepted criteria is removed.
- Ranking criteria:
  - Nominal line capacities: after each step, line with the highest nominal capacity removed. If two lines have the same nominal capacity, the decision is made randomly;
  - Modal weight of the line, according to the nominal capacity: after each step, the modal weights of the remaining lines are recalculated and the line with the highest modal weight is removed;
  - Actual flow trough the line: after each step the redistribution of flow trough the network is recalculated and the line with the highest flow is removed. If two or more lines have the same values, different scenarios observed;
  - Modal weight of the line, according to the actual flow trough the line: after each step the redistribution of flow trough the network is recalculated, after which the modal analysis using actual flows trough the lines is performed. Line with the highest modal weight is removed.
- Network disintegration indicator: the maximum of the Objectivity Function – maximal consumption in the network.

LP Power Flow Model

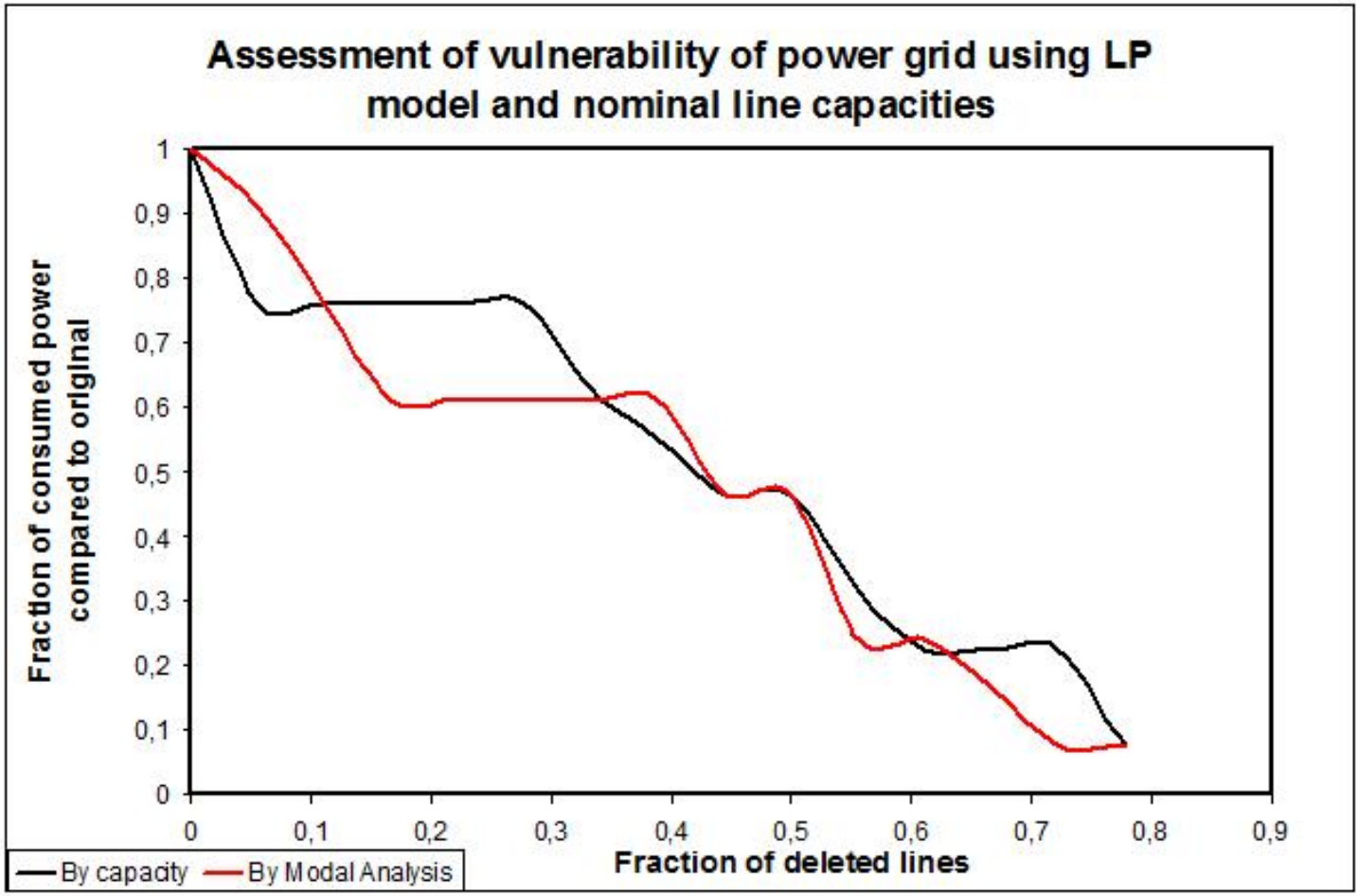
The analyzed network:



- 13 Nodes (4 Generating, 6 Demand, 3 Transmission)
- 18 Lines

LP Power Flow Model

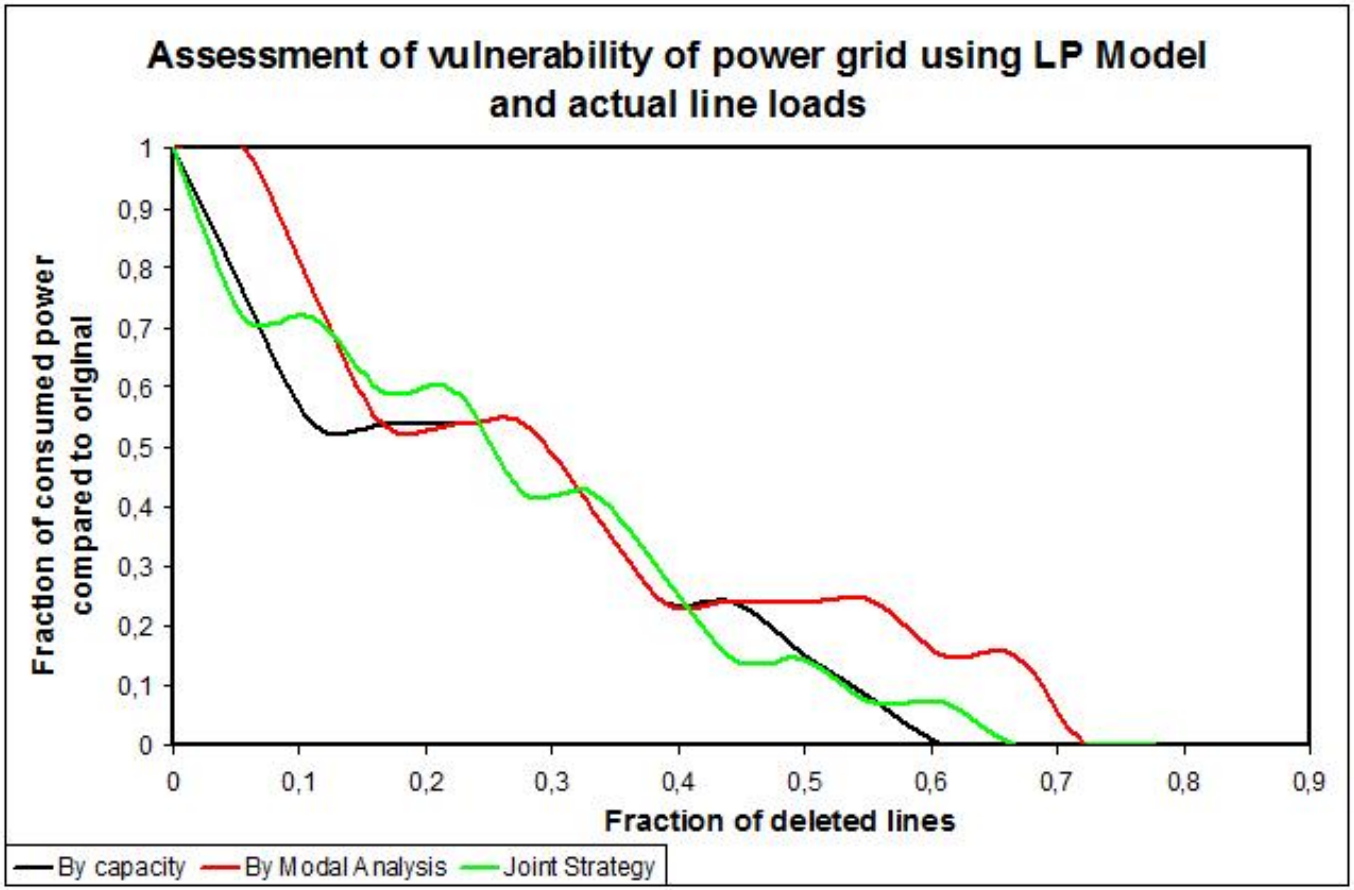
Results



HELSINKI JUNE 2008

LP Power Flow Model

Results

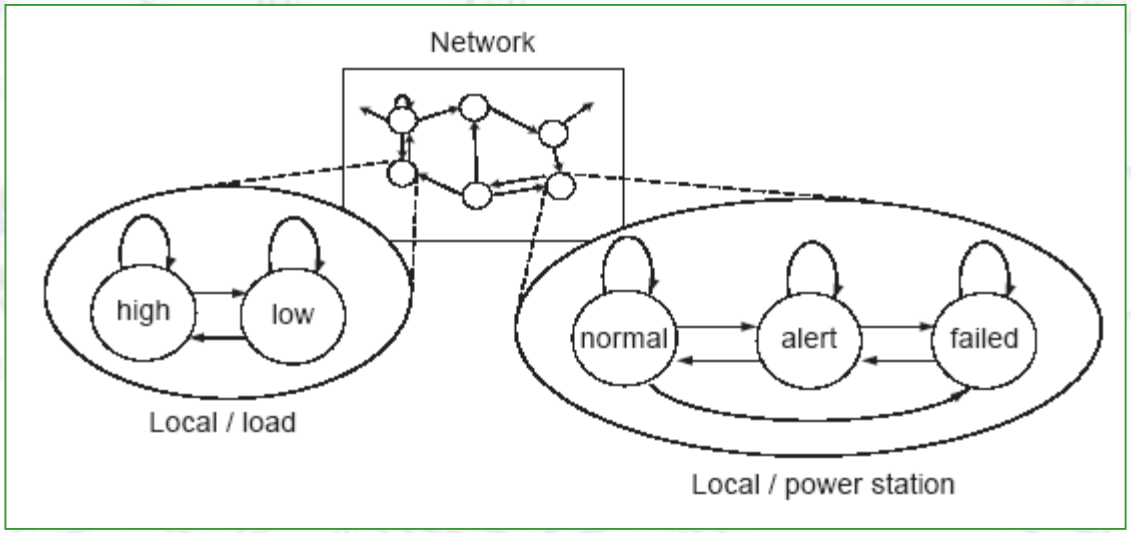


Note on Joint strategy: when 2 or more lines have the same flow trough, the line with the highest modal weight is removed.

HELSINKI JUNE 2008

Influence model

Influence model is a stochastic dynamical system defined on a graph, and is described at two levels: the network level and the local level.



At the network level, each node can be treated as one active entity, and is called a site.

Example: A site can either be a power station (generator) or a load. A power station may be represented as being in one of three possible statuses at any given time: normal, alert or failed. The loads, which can be cities or factories, might be in either high or low status, depending on the present level of demand.

Influence model

Directed graphs

Let $A = [a_{ij}]$ be an $n \times n$ matrix

Directed graph $\Gamma(A)$:

Directed edge from i to j exsits if and only if $a_{ij} \neq 0$

Edge weight is given by a_{ij}

Stochastic matrix, Transpose matrix,
Directed (weighted) graph = Network

Influence model

Binary influence model: the status of each node at any given time step is assumed to be 0 or 1, which may represent any two different statuses such as ‘on’ vs. ‘off’, ‘healthy’ vs. ‘sick’, or ‘normal’ vs. ‘failed’.

The binary influence model can potentially illuminate our understanding of the qualitative behavior of a number of systems.

In power systems, this model can be used as a highly simplified paradigm for cascading blackouts. Here the network graph would represent the power grid, and each node would be a substation or a power plant whose status value is amenable to a binary label. To simulate cascading failure, we can start with a network in which every node is in ‘normal’ state and then initiate a node failure by turning the status at some node to ‘failed’.

Binary Influence Model

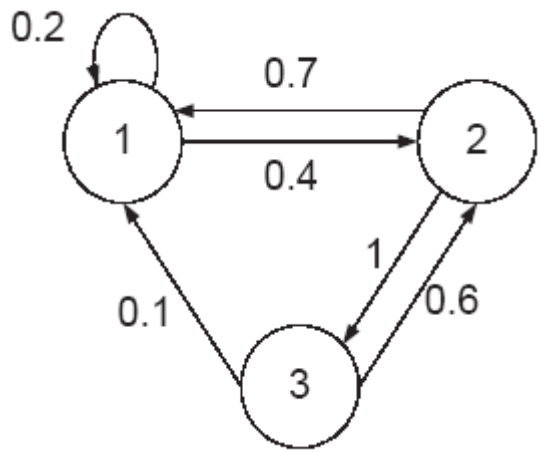
D $n \times n$ stochastic matrix

$\Gamma(D^T)$ network influence model

The sum of edges pointing into a site is 1. This feature allows us to treat each edge weight as the ***relative amount of influence*** from the source node to the destination.

$$D' = \begin{bmatrix} .2 & .4 & 0 \\ .7 & 0 & 1 \\ .1 & .6 & 0 \end{bmatrix}$$

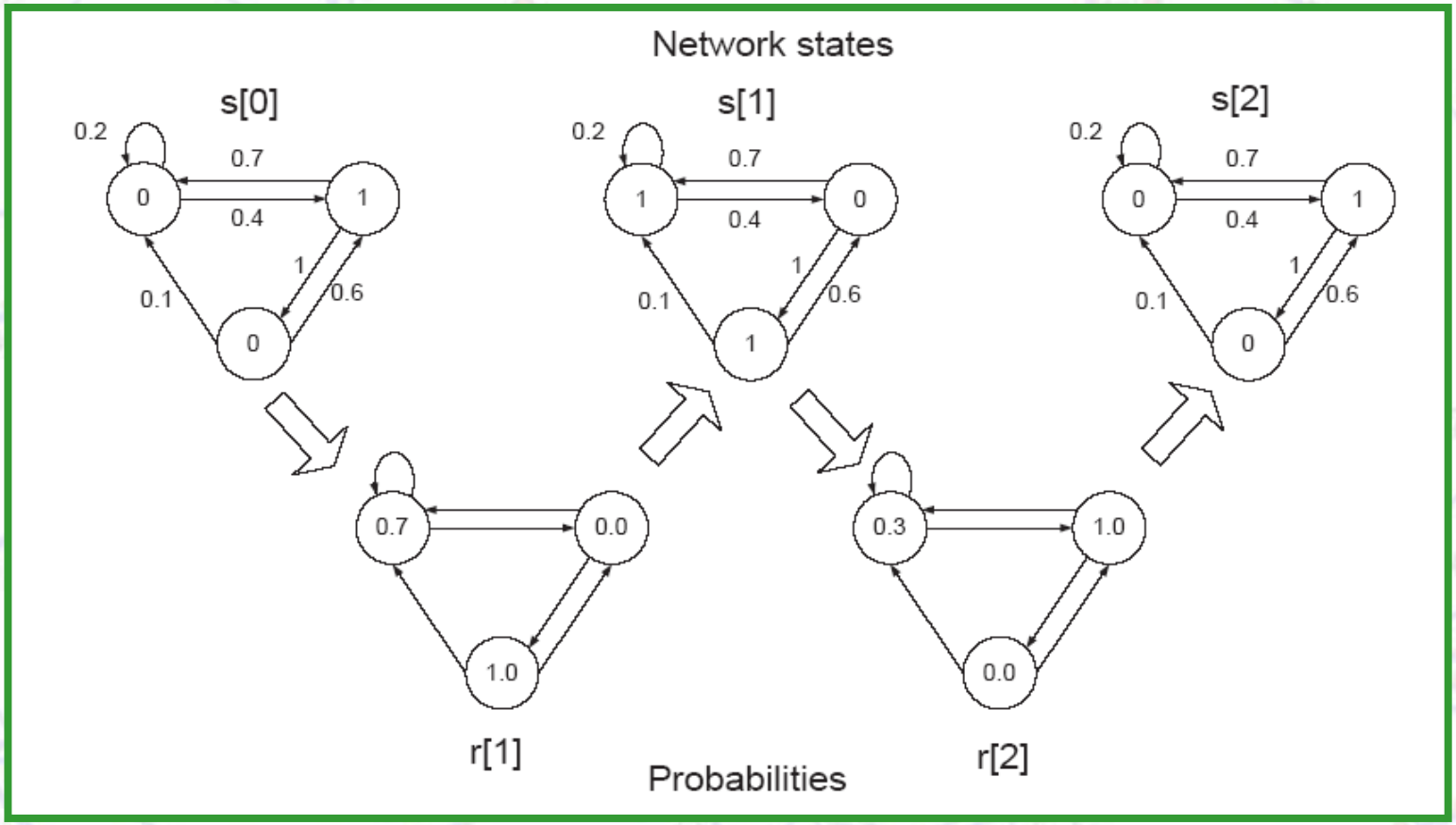
Transpose matrix



Binary Influence Model

$$\mathbf{s}[k] \triangleq [\, s_1[k] \;\ldots\; s_n[k] \,]'$$

$$\begin{aligned} \mathbf{r}[k+1] &= D\,\mathbf{s}[k] \\ \mathbf{s}[k+1] &= Bernoulli(\mathbf{r}[k+1]) \end{aligned}$$



Vulnerability Rank

If D is ergodic, then $\lim_{k\to\infty} D^k = \mathbf{1}\pi^T$

π is the left eigenvector corresponding to the eigenvalue at 1, which has been normalized so that

$$\pi^T \mathbf{1} = 1$$

$\Gamma(D^T)$ is ergodic

The probability that the influence process starting from initial state $s[0]$ will eventually settle in the all-ones consensus state is

$$\pi^T s[0]$$

The probability of reaching the all-zeros consensus state is

$$1 - \pi^T s[0]$$

Vulnerability Rank

We define ***vulnerability*** of the network as the ***stationary distribution*** of its influence matrix, which is the normalized left eigenvector associated with the eigenvalue 1.

$$\pi = [\pi_1, \pi_2, ..., \pi_n]^T$$

$$0 < \pi_{j_n} < ... < \pi_{j_2} < \pi_{j_1} < 1$$

$$s_{j_n}(0) = ... s_{j_2}(0) = 0; s_{j_1}(0) = 1$$

The probability that the influence process starting from initial state $s(0)$ will eventually settle in the ***all-ones consensus*** state is

$$\pi^T s(0) = \pi_{j_1}$$

The site j_1 is the most *vulnerable* site. ***Vulnerability Rank*** describes what is the influence of each cite i to the failure of the network

Vulnerability Rank

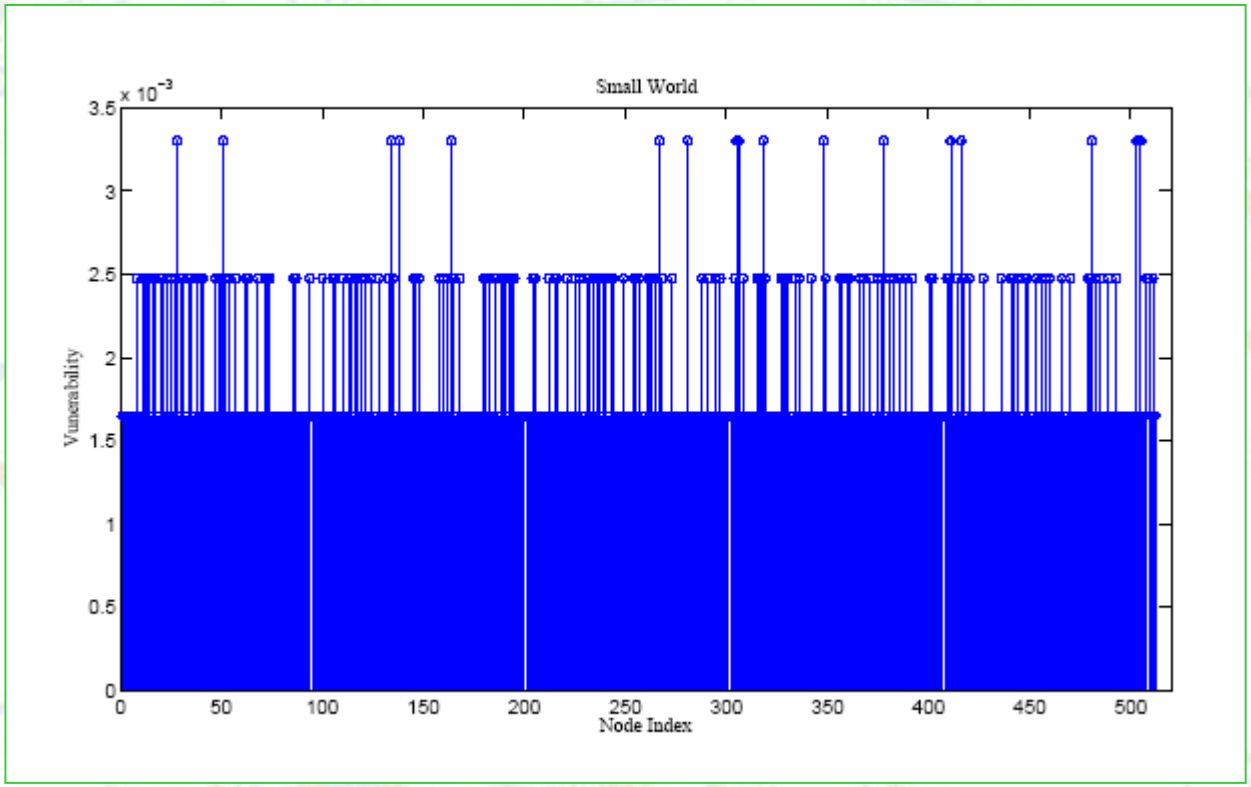
$$D = \begin{bmatrix} 1/60 & 7/15 & 7/15 & 1/60 & 1/60 & 1/60 \\ 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 \\ 19/60 & 19/60 & 1/60 & 1/60 & 19/60 & 1/60 \\ 1/60 & 1/60 & 1/60 & 1/60 & 7/15 & 7/15 \\ 1/60 & 1/60 & 1/60 & 7/15 & 1/60 & 7/15 \\ 1/60 & 1/60 & 1/60 & 11/12 & 1/60 & 1/60 \end{bmatrix}$$

$$\pi^T = [\ .03721 \quad .05396 \quad .04151 \quad .3751 \quad .206 \quad .2862\ ]$$

The site 4 is the most vulnerable site and $\pi_4 + \pi_6 > 0.5$

HELSINKI JUNE 2008

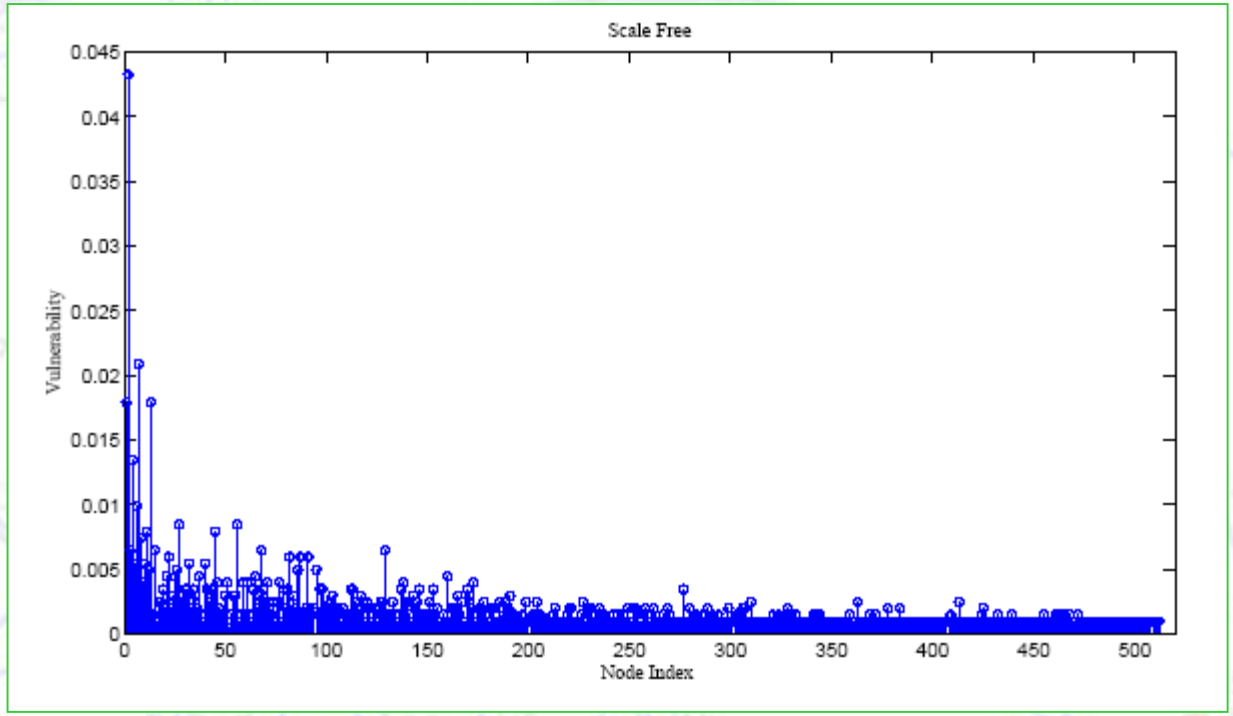
Vulnerability Rank



Vulnerability Rank for binary influence model on small world graph.

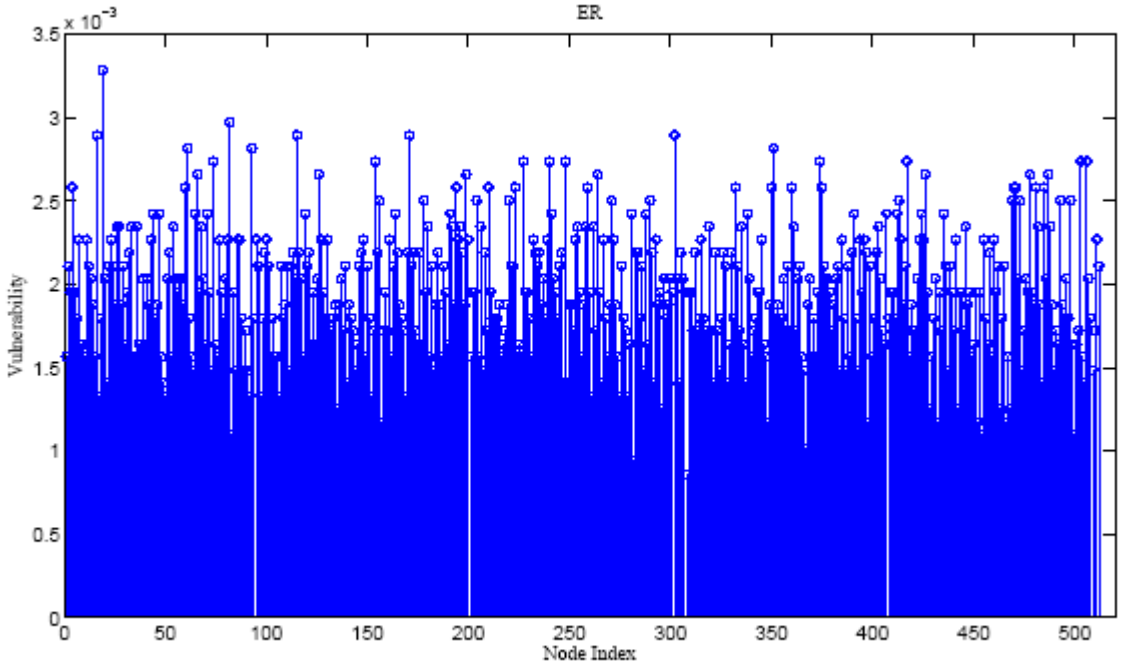
HELSINKI JUNE 2008

Vulnerability Rank



Vulnerability Rank for binary influence model on scale free graph.

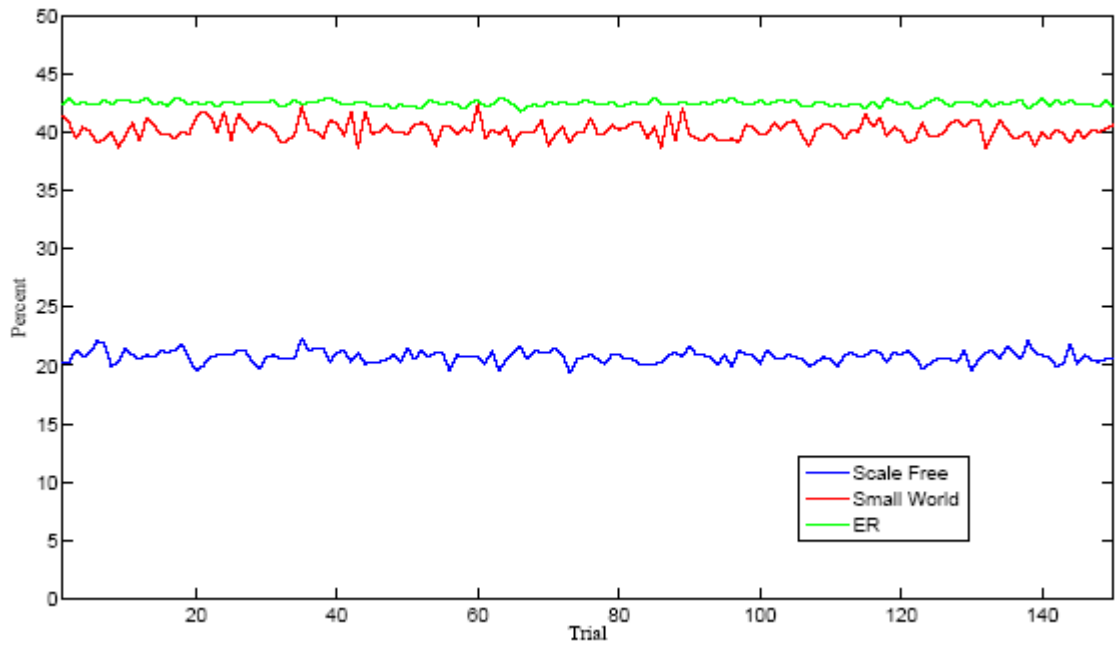
Vulnerability Rank



Vulnerability Rank for binary influence model on ER graph.

HELSINKI JUNE 2008

Vulnerability Rank



Percentage of the nodes that need to be in the status off at time 0 so that, when time goes to infinity, the probability of a site to be in the status 1 (off) is greater or equal to 0.5 for 150 different realizations of the corresponding graph.

General influence model

The influence model is a discrete-time Markov process whose state space is the tensor product of the statuses of all the local Markov chains.

$$H \triangleq D' \otimes \{A_{ij}\} = \begin{bmatrix} d_{11}A_{11} & \cdots & d_{n1}A_{1n} \\ \vdots & & \vdots \\ d_{1n}A_{n1} & \cdots & d_{nn}A_{nn} \end{bmatrix}$$

$$\begin{aligned} \mathbf{p}'[k+1] &\triangleq \mathbf{s}'[k]\,H \\ \mathbf{s}'[k+1] &\triangleq MultiRealize(\,\mathbf{p}'[k+1]\,) \end{aligned}$$

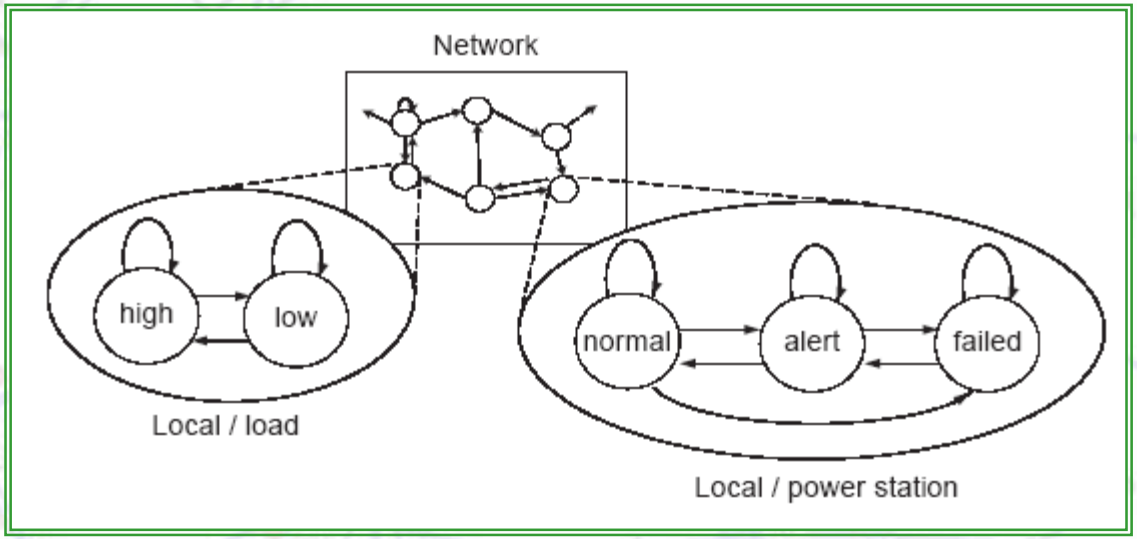
General influence model

The influence matrix H is, in general, not stochastic. However, its dominant eigenvalue is one. Assuming for simplicity that all its eigenvalues are distinct, the steady-state value of the evolution of the status probability mass-function (PMF) approaches the left eigenvector corresponding to eigenvalue 1.

$$E(s^T(k) = E(s^T(0))H^k \rightarrow \pi$$

Vulnerability Rank of the network is defined as the ***stationary distribution*** of its influence matrix, which is the normalized left eigenvector associated with the eigenvalue 1.

General influence model



Normal Alert Failed

High Low

$$A_{11} = \begin{bmatrix} 1-p & p \\ 1-q & q \end{bmatrix}$$

$$A_{22} = \begin{bmatrix} a & b & 1-a-b \\ c & d & 1-c-d \\ e & f & 1-e-f \end{bmatrix}$$

General influence model

Load Plant

$$D^T = \begin{bmatrix} 1-t & 1 \\ t & 0 \end{bmatrix}$$

Load

Plant

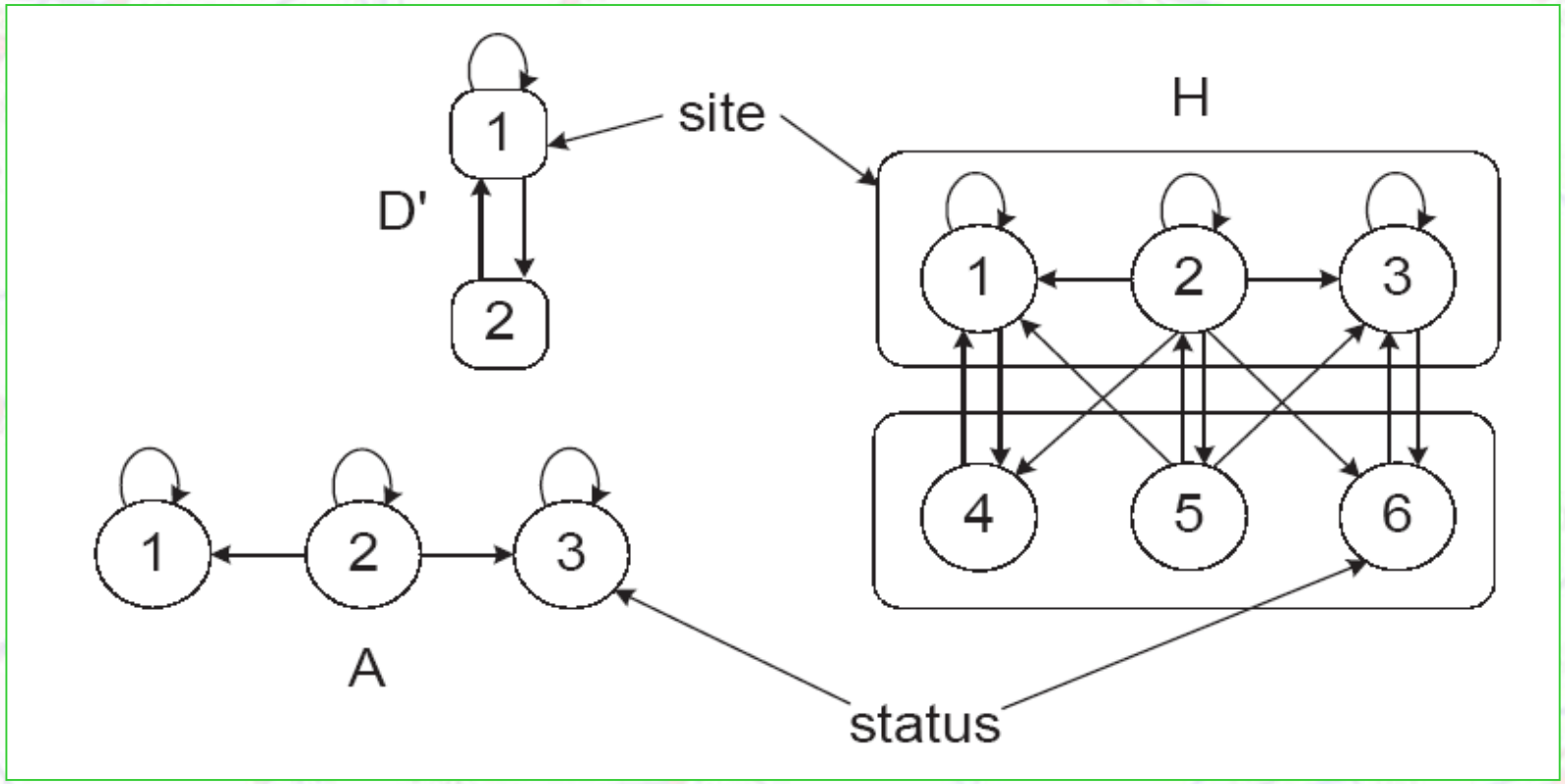
Influence matrix

$$H = D^T \otimes \{A_{ij}\}$$

$$H = \begin{bmatrix} (1-t)(1-p) & (1-t)p & 0 & m & 1-m \\ (1-t)(1-q) & (1-t)q & n & 1-n & 0 \\ tx & t(1-x) & 0 & 0 & 0 \\ ty & t(1-x) & 0 & 0 & 0 \\ 0 & t & 0 & 0 & 0 \end{bmatrix}$$

HELSINKI JUNE 2008

General influence model



HELSINKI JUNE 2008

General influence model

$$t = p = q = m = n = x = y = 0.5$$

$$a = b = c = d = e = f = 1/3$$

High	Low	Normal	Alert	Failed

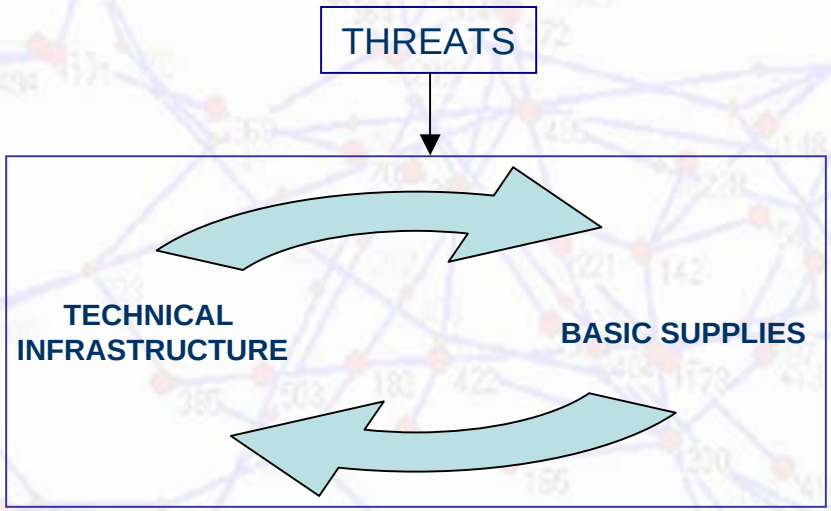
$$\pi = [0.4444 \quad 0.5556 \quad 0.2778 \quad 0.5 \quad 0.2222]$$

The probability that in the steady-state the power plan will be in normal status is 0.2778

Interdependencies of critical infrastructures

Basic concept

- Assessment of the compound risk of failures
- Interoperating items considered



Statistical analysis of the failures interdependence (Sivonen)

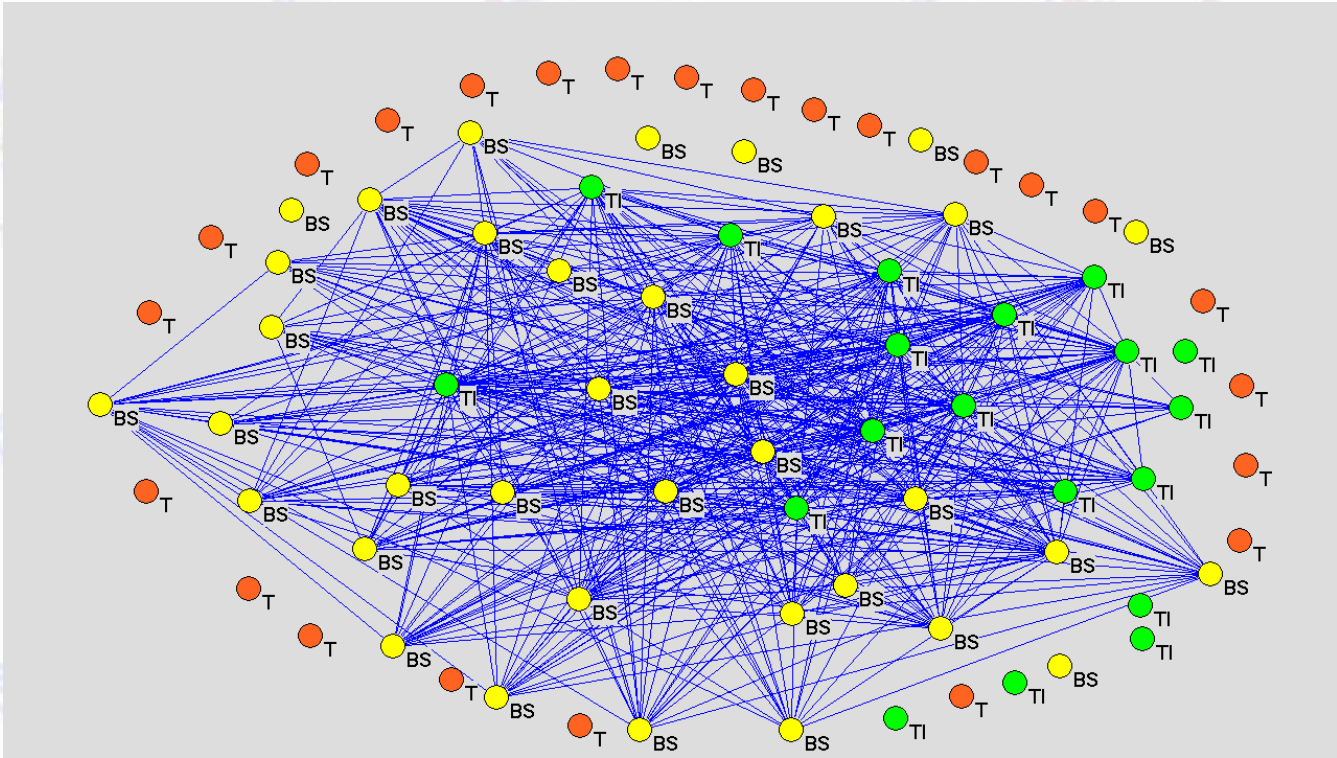
-The mean time between failures (probability)
-The durations of failures
-The effect of a one-day-long failure

Compound risk = effect × probability × duration

Interdependencies of critical infrastructures

HELSINKI JUNE 2008

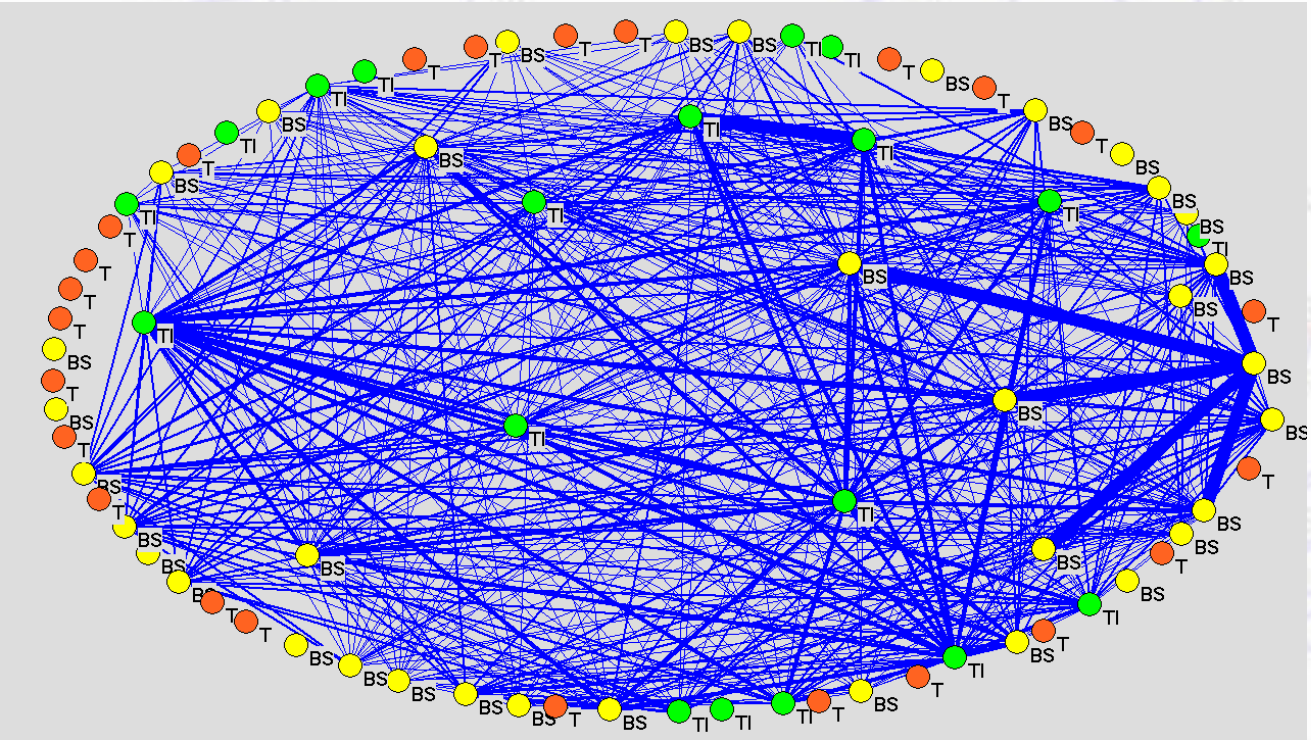
Graph of the interdependent networks consisting of basic supplies, technical infrastructures and threats



Interdependencies of critical infrastructures

HELSINKI JUNE 2008

The influence model simulation outcome (the width of the lines is proportional to the probability the corresponding nodes to be in the same status (ON or OFF), TI stands for technical infrastructure, BS - basic supplies, T - threats)



In calculating the Vulnerability Rank for the graph representing the Network of Infrastructures, we face the following problem: ***the graph is not irreducible***.

Irreducibility is a desirable property because it is precisely the feature that guarantees that a influence model possesses a unique (and positive) stationary distribution vector.

$$\pi_i = (1/n)\sum_{i=1}^{n} \lim_{k\to\infty} D^k s_i(0)$$

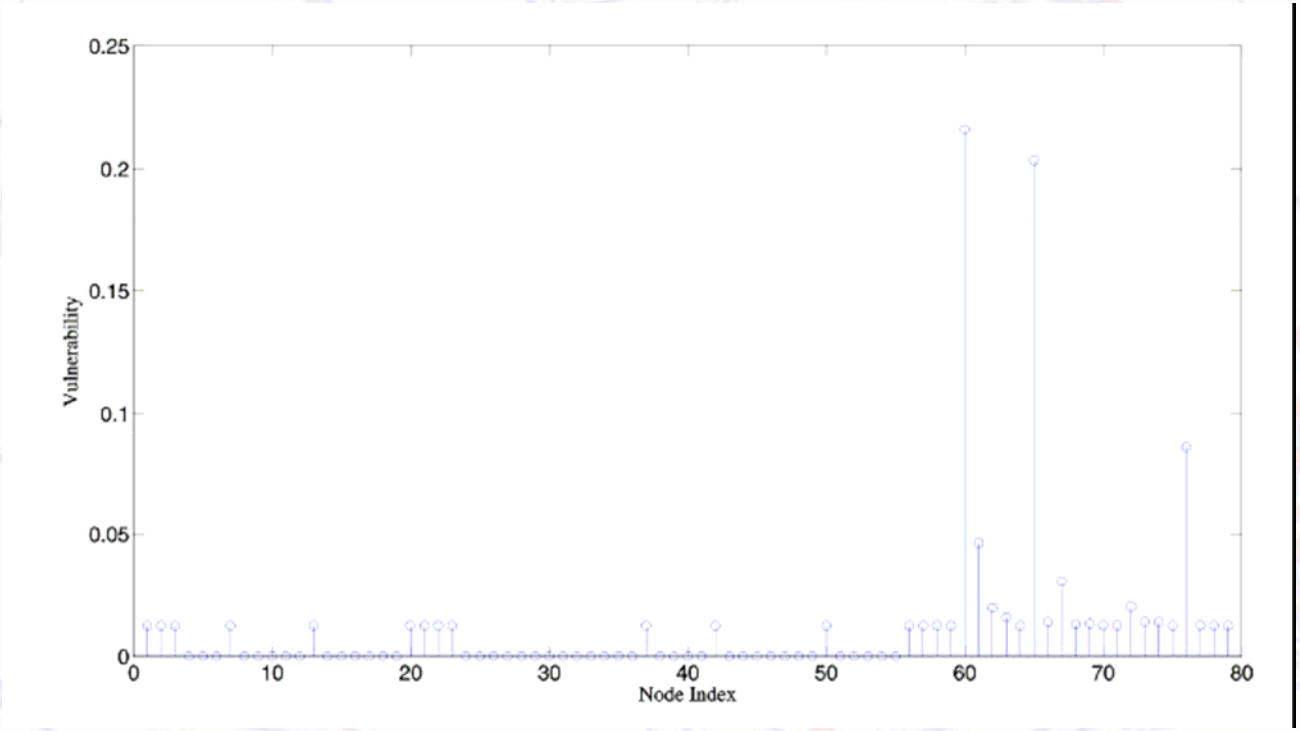
$$s_i(0) = [0\ldots\ s_{ii}(0) = 1\ \ldots 0]^T$$

$$D^* = \alpha D + (1-\alpha)\frac{E}{n}$$

$$\pi \approx \pi^*$$

Interdependencies of critical infrastructures

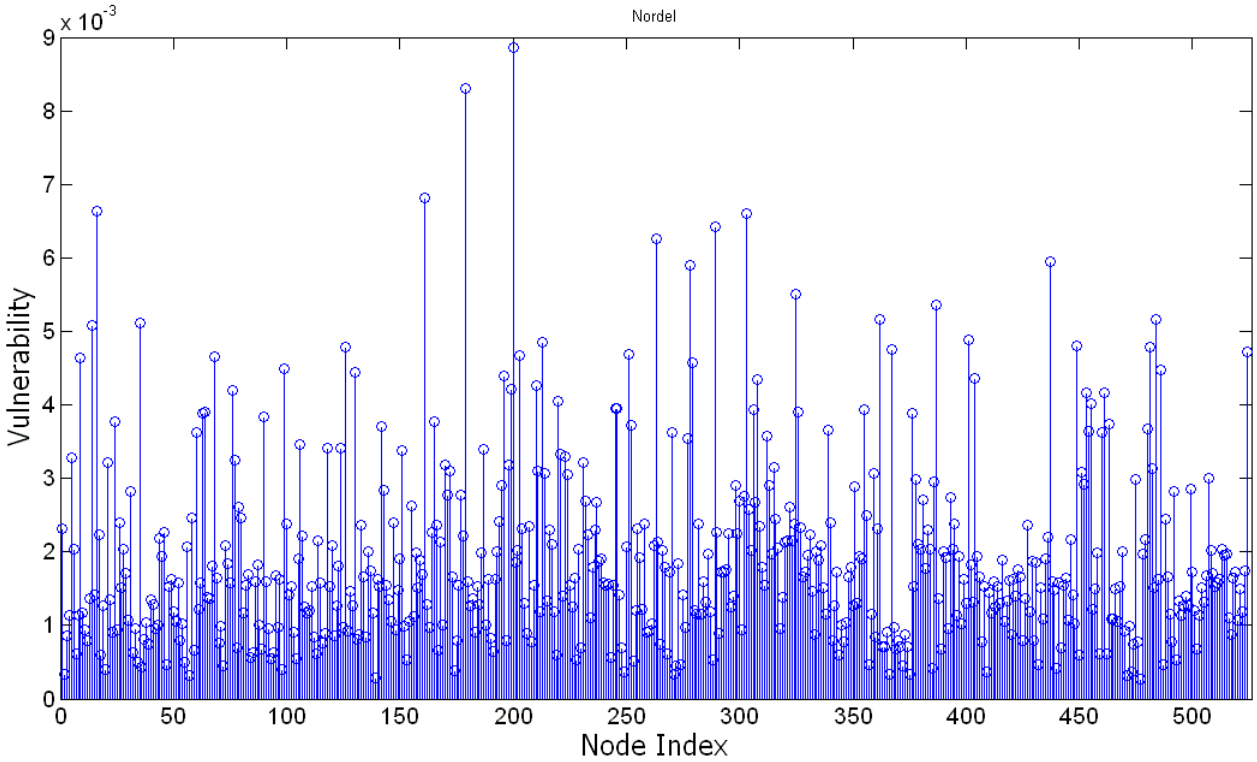
Vulnerability Rank for infrastructure network influence graph



The most vulnerable sites are the sites representing the following threats (out of 17 threats grouped in 4 groups: Causes for severe disturbances, Economic threats, Environment and health treats, Political security threats):

1. Weather phenomenon,
2. Threats to data systems,
3. Crime and terrorism,
4. Strike, and
5. International logistics crisis.

The threats 1 and 4 belong to the group: Causes for severe disturbances, threats 2 i 5 to the group Economic threats and 3 to Political security threats.



Vulnerability Rank for NORDEL power grid.

CONCLUSIONS

- Several definitions of vulnerability are presented
- The attack vulnerability was analyzed from static and dynamic aspects
- The dynamic studies were based on influence model and flow models
- Influence matrix was introduce as a possible device to quantify the interoperability of networks
- A method for calculating *Vulnerability Rank* for networks of Markov chains, applicable to critical infrastructures was suggested