Best spreader node in a graph

Piet Van Mieghem

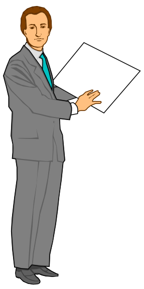
in collaboration with Karel Devriendt and Hale Cetinay

1

Second workshop on Critical and Collective Effects in Graphs and Networks (CCEGN 2017)
Moscow, May 15-19, 2017

TU Delft

Outline



Background: Electrical matrix equations

New graph metrics

Best nodal spreader

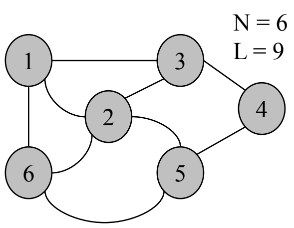
TU Delft

Function of network

- Usually, the function of a network is related to the *transport of items over its underlying graph*
- In man-made infrastructures: two major types of transport
  - Item is a flow (e.g. electrical current, water, gas,…)
  - Item is a packet (e.g. IP packet, car, container, postal letter,…)
- Flow equations (physical laws) determine transport (Maxwell equations (Kirchoff & Ohm), hydrodynamics, Navier-Stokes equation (turbulent, laminar flow equations, etc.)
- Protocols determine transport of packets (IP protocols and IETF standards, car traffic rules, etc.)

TU Delft

Adjacency matrix A



$$A_{N\times N} = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

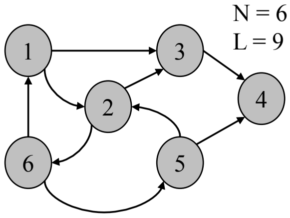
For an undirected graph: $A = A^T$ is symmetric

Number of neighbors of node *i* is the degree: $d_i = \sum_{k=1}^{N} a_{ik}$

if there is a link between node i and j, then $a_{ij} = 1$
else $a_{ij} = 0$

TU Delft

Incidence matrix B



N = 6
L = 9

- Label links (e.g.: l_1 = (1,2), l_2 = (1,3), l_3 = (1,6), l_4 =(2,3), l_5 =(2,5), l_6 =(2,6), l_7 =(3,4), l_8 =(4,5), l_9 =(5,6))
- Col *k* for link l_k = (i,j) is zero, except:
  source node i = 1 → b_{ik} = 1
  destination node j = -1 → b_{jk} = -1

$$B_{N\times L} = \begin{bmatrix} 1 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix}$$

Col sum B is zero: $u^T B = 0$

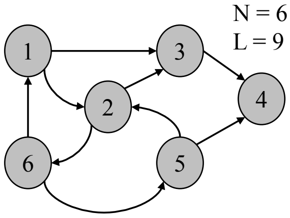
where the all-one vector u = (1,1,…,1)

5

B specifies the directions of links

TUDelft

Laplacian matrix Q



N = 6
L = 9

$$Q_{N\times N} = \begin{bmatrix} 3 & -1 & -1 & 0 & 0 & -1 \\ -1 & 4 & -1 & 0 & -1 & -1 \\ -1 & -1 & 3 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & -1 & 0 & -1 & 3 & -1 \\ -1 & -1 & 0 & 0 & -1 & 3 \end{bmatrix}$$

$$Q = BB^T = \Delta - A$$

$$\Delta = diag(d_1 \quad d_2 \quad \ldots \quad d_N)$$

Since BB^T is symmetric, so are *A* and *Q*. Although *B* specifies directions, *A* and *Q* lost this info here.

Basic property: $Qu = 0$

u is an eigenvector of *Q*
Belonging to eigenvalue μ = 0

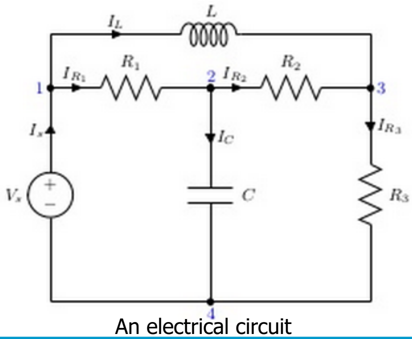
6

$Qu = BB^T u = 0$ because $0 = u^T B = B^T u$

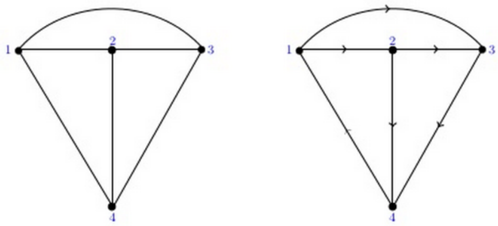
TUDelft

Electrical Circuits as Graphs

- An electrical circuit can be represented as a graph $G(N,L)$:
  - Terminals, buses ⇨ Nodes
  - Lines, branches ⇨ Links



An electrical circuit



Its undirected & directed graph

North-Eastern Hill University, EE-304 Electrical Network Theory [Class Notes1] - 2013

TUDelft

Analysis of Electrical Circuits

- Find all link currents (y_{ij}) and node voltages (v_i) from the known parameters in the network
- Two important laws apply:

Kirchhoff's Current Law (a conservation law): The sum of all the directed currents at each node *i* in a circuit equals zero:

$$\sum_{j \in \text{neighbors}(i)} y_{ij} = \sum_{j=1}^{N} a_{ij} y_{ij} = 0$$

Ohm's Law: The current through a conductor equals the potential difference over the conductor multiplied by its resistance. For any link between two nodes *i* and *j* with a resistance r_{ij}:

$$v_i - v_j = r_{ij} y_{ij}$$

Recall that $A = A^T$: thus $a_{ij} = a_{ji}$; *A* represents link existence;
B specifies link directionality and the link current obeys $y_{ij} = -y_{ji}$.
Both the vector *y* and *v* can have negative components.

TUDelft

Current-voltage relation for any graph

Ohm's Law: $v_i - v_j = r_{ij} y_{ij}$ → $y_{ij} = \frac{1}{r_{ij}}(v_i - v_j)$

Kirchhoff's law (conservation law): $x_i = \sum_{j \in \text{neighbors(i)}} y_{ij} = \sum_{j=1}^{N} a_{ij} y_{ij}$

Substituting Ohm's law into Kirchhoff's law:

$$x_i = \sum_{j=1}^{N} \frac{a_{ij}}{r_{ij}}(v_i - v_j) = v_i \sum_{j=1}^{N} \frac{a_{ij}}{r_{ij}} - \sum_{j=1}^{N} \frac{a_{ij}}{r_{ij}} v_j$$

We define a weighted Laplacian by $\tilde{Q} = \tilde{\Delta} - \tilde{A}$, where the elements in the weighted adjacency matrix Ã are $\tilde{a}_{ij} = \frac{1}{r_{ij}} a_{ij}$

Finally: $x = \tilde{Q} v$

Clearly, if all r_{ij} = 1, then $\tilde{Q} = Q$

TUDelft

Pseudo-inverse of the Laplacian (1)

Since det Q = 0, we **cannot** solve $x = Qv$ as $v = Q^{-1}x$

The spectral decomposition of the Laplacian: $Q = ZMZ^T$
where Z is the orthogonal eigenvector matrix ($Z^T Z = ZZ^T = I$)
and M is the diagonal matrix with the ordered eigenvalues

$$\mu_1 \geq \mu_2 \geq \ldots \geq \mu_{N-1} \geq \mu_N = 0$$

Expanding $Q = ZMZ^T$: $Q = \sum_{k=1}^{N} \mu_k z_k z_k^T = \sum_{k=1}^{N-1} \mu_k z_k z_k^T + \frac{\mu_N}{N} uu^T = \sum_{k=1}^{N-1} \mu_k z_k z_k^T$

Since G is connected, $\mu_k > 0$ for 0<k<N, so that the matrix

$$\hat{Q}^{-1} = \sum_{k=1}^{N-1} \frac{1}{\mu_k} z_k z_k^T$$

exists and is called the pseudo-inverse of Q.

10

TUDelft

Pseudo-inverse of the Laplacian (2)

The pseudo-inverse of Q

$$\hat{Q}^{-1} = \sum_{k=1}^{N-1} \frac{1}{\mu_k} z_k z_k^T$$

obeys

$$\hat{Q}^{-1} Q = Q \hat{Q}^{-1} = I - \frac{1}{N} J$$

Indeed, use orthogonality $z_k^T z_j = \delta_{kj}$ and $Z^T Z = Z Z^T = I$

$$\hat{Q}^{-1} Q = \sum_{k=1}^{N-1} \frac{1}{\mu_k} z_k z_k^T \sum_{j=1}^{N-1} \mu_j z_j z_j^T = \sum_{k=1}^{N-1} \sum_{j=1}^{N-1} \mu_j \frac{1}{\mu_k} z_k \left(z_k^T z_j \right) z_j^T = \sum_{k=1}^{N-1} z_k z_k^T = I - \frac{u}{\sqrt{N}} \frac{u^T}{\sqrt{N}} = I - \frac{1}{N} J$$

An alternative form is $\hat{Q}^{-1} = (Q + \alpha J)^{-1} - \frac{1}{N^2} J$ for any $\alpha \neq 0$

11

P. Van Mieghem, Graph Spectra for Complex Networks, Cambridge University Press, 2010, p.203-206

TUDelft

Matrix Representations of the Basic Laws

Using the pseudo-inverse on $x = Qv$

$$\hat{Q}^{-1} x = \hat{Q}^{-1} Q v = \left(I - \frac{1}{N} J \right) v$$

Further $$\left(I - \frac{1}{N} J \right) v = v - u \left(\frac{u^T v}{N} \right) = v - v_{average} u$$

Hence, from the spectral decomposition of the Laplacian matrix

$$Q = \sum_{k=1}^{N} \mu_k z_k z_k^T$$

the pseudo-inverse is constructed as $\hat{Q}^{-1} = \sum_{k=1}^{N-1} \frac{1}{\mu_k} z_k z_k^T$

The potential vector v is determined from the injected current vector x

$$v - v_{average} u = \hat{Q}^{-1} x$$

- Only potential differences matter!
- Choose e.g. $v_{average} = 0$

12

Observe that $u^T \hat{Q}^{-1} = 0$ so that $u^T v - v_{average} u^T u = 0$ $\rightarrow$ $v_{average} = \frac{u^T v}{N}$

Effective Resistance Matrix Ω

We consider a network in which a flow with magnitude I_c is injected in node *a* and the flow leaves the network at node *b*

The injected current vector $x = I_c(e_a - e_b)$, where the basic vector e_k has components $(e_k)_j = \delta_{kj}$

The general voltage-injected current vector relation $v = \hat{Q}^{-1}x$

shows $$v = I_c \hat{Q}^{-1}(e_a - e_b)$$

The aim is to determine the effective resistance matrix Ω with elements ω_{ab} that satisfy $v_a - v_b = I_c\, \omega_{ab}$

Since $v_a - v_b = (e_a - e_b)^T v$, we have $v_a - v_b = I_c (e_a - e_b)^T \hat{Q}^{-1}(e_a - e_b)$

$$\omega_{ab} = (e_a - e_b)^T \hat{Q}^{-1}(e_a - e_b)$$

P. Van Mieghem, Graph Spectra for Complex Networks, Cambridge University Press, 2010, p.203-206

TU Delft

Effective Resistance Matrix Ω

Multiplying out $\omega_{ab} = (e_a - e_b)^T \hat{Q}^{-1}(e_a - e_b)$

yields $$\omega_{ab} = \left(\hat{Q}^{-1}\right)_{aa} + \left(\hat{Q}^{-1}\right)_{bb} - 2\left(\hat{Q}^{-1}\right)_{ab}$$

so that the effective resistance matrix is

$$\Omega = u\zeta^T + \zeta u^T - 2\hat{Q}^{-1}$$

where the vector ζ is

$$\zeta = \left(\left(\hat{Q}^{-1}\right)_{11}, \left(\hat{Q}^{-1}\right)_{22}, \ldots, \left(\hat{Q}^{-1}\right)_{NN}\right)$$

TU Delft

Effective Resistance Matrix Ω

Properties of $\Omega = u\zeta^T + \zeta u^T - 2\hat{Q}^{-1}$

1. Symmetric: $\Omega^T = \Omega$
2. All diagonal elements of Ω are zero
3. Triangular inequality: $\omega_{ij} \leq \omega_{ik} + \omega_{kj}$

} Distance matrix

The effective graph resistance is defined as $R_G = \frac{1}{2} u^T \Omega u$

R_G measures the difficulty of transport in a graph G: low R_G corresponds to a "good conducting" network (i.e. low resistance to flows)

Since $u^T \hat{Q}^{-1} u = 0$ (each eigenvector z_k with k< N is orthogonal to z_N = u)

we find $R_G = N u^T \zeta = N \text{trace}\left(\hat{Q}^{-1}\right) = N \sum_{k=1}^{N-1} \frac{1}{\mu_k}$

Ellens, W. , F. A. Spieksma, P. Van Mieghem, A. Jamakovic and R. E. Kooij, 2011, "Effective Graph Resistance", Linear Algebra and its Applications, Vol. 435, pp. 2491-2506.

More general than resistor networks

$x(s) = Q(s)v(s)$ for link impedances $r_l + sL_l + \frac{1}{s}C_l$

Processes 'proportional' to the network topology (graph)

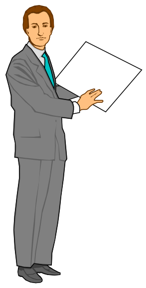
The analogy with voltage is in brackets:

- water flow networks (height of water column)
- spring-mass systems (energy stored in spring)
- gas networks (pressure)
- warmth diffusion in networks (temperature)

Processes described by a weighted Laplacian

- Any Markovian continuous-time processes, where the infinitesimal generator is minus a weighted Laplacian of the Markov state graph)

Outline



Background: Electrical matrix equations

New graph metrics

Best nodal spreader

P. Van Mieghem, K. Devriendt and H. Cetinay, "Pseudo-inverse of the Laplacian and best spreader node in a network", unpublished, 2017

TU Delft

New electrical matrix equations

Recall the inverse current-voltage relations

$x = Qv$ $\qquad$ $v = Q^{\$}x$

Matrix Ohm law: $v = -\frac{1}{2N}(N\,I - J)\Omega x$

where the effective resistance matrix is denoted by Ω

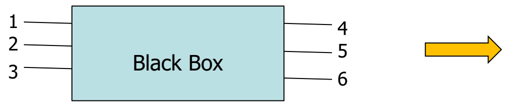
Power $\wp$ dissipated in the network: $\wp = -\frac{1}{2}x^T \Omega x$

Current/flow eigenvalue equation: $Q\Omega x = -2x$

19

TU Delft

Relation between "function" and "structure" of a network



$\Omega_{6\times 6}$

Block inverse (*inspired by Fiedler, "Geometry of Laplacian", LAA, 1998*):

$$\begin{bmatrix} 0 & u^T \\ u & \Omega \end{bmatrix} = -2 \begin{bmatrix} \zeta^T Q \zeta + \frac{4R_G}{N^2} & -\left(Q\zeta + \frac{2}{N}u\right)^T \\ -\left(Q\zeta + \frac{2}{N}u\right) & Q \end{bmatrix}^{-1}$$

where $\left(Q\zeta + \frac{2}{N}u\right)$ is the eigenvector of $Q\Omega$ belonging to the zero eigenvalue.

$\zeta = \left(\left(Q^{\$}\right)_{11}, \left(Q^{\$}\right)_{22}, \ldots, \left(Q^{\$}\right)_{NN} \right)$

Geometry

Spectral decomposition: $Q^{\$} = ZM^{\$}Z^T = \left(Z\sqrt{M^{\$}}\right)\left(Z\sqrt{M^{\$}}\right)^T$

The matrix $S^{\$} = \left(Z\sqrt{M^{\$}}\right)^T$ has rank *N-1* (row *N* =0 due to $\mu_N = 0$)

The *i*-th column vector s_i = point p_i in (*N-1*)-dim space AND

$$\left\| s_i - s_j \right\|_2^2 = \omega_{ij}$$

All *N* points p_j form a hyperacute simplex with volume V_G

$$V_G = \frac{1}{N!\sqrt{2\xi}}$$

where the number of (weighted) spanning trees ξ is

$$\xi = \frac{1}{N}\prod_{k=1}^{N-1} \mu_k$$

K. Menger, "New foundation of Euclidean geometry", American Journal of Mathematics, 53(4):721-745, 1931

Graph metrics

The volume V_G (or the weighted complexity)

$$\xi = \frac{1}{N}\prod_{k=1}^{N-1}\mu_k = \frac{1}{2(N!\,V_G)^2}$$

provides similar information as the effective graph resistance

$$R_G = N\sum_{k=1}^{N-1}\mu_k$$

by concavity of log(x) function:

Small volume → large number of spanning trees
→ good conductance
→low effective graph resistance

22

TU Delft

Graph metrics

Average $\frac{1}{N}\sum_{k=1}^{N}\zeta_k = \frac{R_G}{N^2}$ ⟶ $R_G = N\sum_{k=1}^{N-1}\frac{1}{\mu_k}$

R_G is the *effective graph resistance*

Variance $\frac{1}{N}\sum_{k=1}^{N}\left(\zeta_k - \frac{R_G}{N^2}\right)^2 \le \frac{\nabla R}{N^2}$ ⟶ $\nabla R = (N-1)\sum_{k=1}^{N-1}\frac{1}{\mu_k^2} - \left(\sum_{k=1}^{N-1}\frac{1}{\mu_k}\right)^2$

Most ζ_k lie in the interval $\left(\frac{R_G}{N^2} - \frac{\sqrt{\nabla R}}{N}, \frac{R_G}{N^2} + \frac{\sqrt{\nabla R}}{N}\right)$

⟶ ∇R reflects the nodal spreading heterogeneity in the graph

23

$$\zeta = \left((Q^\$)_{11}, (Q^\$)_{22}, \ldots, (Q^\$)_{NN}\right)$$

TU Delft

Open problem: when is $Q^\$$ a (weighted) Laplacian?

1. $(Q^\$)_{ii} \geq (Q^\$)_{ij}$

2. $(Q^\$)_{ii}\,(Q)_{ii} \geq \left(1-\frac{1}{N}\right)^2$ → Similar to Heisenberg's uncertainty relation

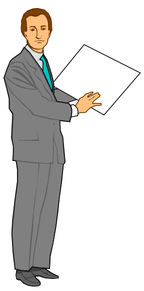
   → $R_G \geq (N-1)^2 E\left[\frac{1}{D}\right]$

3. $(\widehat{Q}^{-1})_{ii} \leq \frac{1}{d_i}\left(1-\frac{1}{N}\right) + \max_{k\neq i} (\widehat{Q}^{-1})_{ik}$

4. If $\widehat{Q}^{-1}$ is a weighted Laplacian then

$$\frac{1}{d_i}\left(1-\frac{1}{N}\right)^2 \leq (\widehat{Q}^{-1})_{ii} \leq \frac{1}{d_i}\left(1-\frac{1}{N}\right)$$

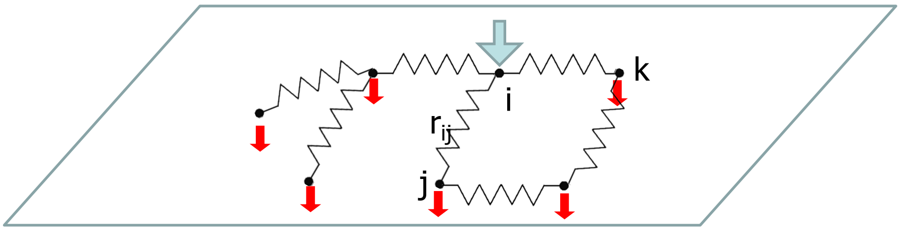
Outline



Background: Electrical matrix equations

New graph metrics

Best nodal spreader



Inverses: $x = Qv$ and $v = Q^{\$}x$ with the convention for voltages $u^T v = 0$

Q : weighted Laplacian of the graph
$Q^{\$}$: pseudo-inverse of the weighted Laplacian: $Q.Q^{\$} = Q^{\$}.Q = I - 1/N\ J$
x: vector with external nodal current
v: vector with nodal potentials

$v_i = (Q^{\$})_{ii}$ If $x = e_i - 1/N\ u$

The best spreader is the node *k* with minimum $(Q^{\$})_{kk}$

$\zeta = \left((Q^{\$})_{11}, (Q^{\$})_{22}, \ldots, (Q^{\$})_{NN} \right)$

TUDelft
Technische Universiteit Delft

Interpretation: best spreader node

Inverses: $x = Qv$ and $v = Q^{\$}x$ Clearly: $u^T x = 0$ and $u^T v = 0$

$v_i = (Q^{\$})_{ii}$ is same as $(Q^{\$})_{ii} = v_i - u^T v = \frac{1}{N}\sum_{k=1}^{N}(v_i - v_k)$

The best spreader minimizes the sum of potential differences *between its and all other node potentials*

→ "closeness" minimization of average distance to all other nodes

→ best spreader lies in center of "gravity"

$\zeta = \left((Q^{\$})_{11}, (Q^{\$})_{22}, \ldots, (Q^{\$})_{NN} \right)$

TUDelft
Technische Universiteit Delft

Closeness Cl_i and $(Q^\$)_{ii}$

Closeness ***Cl_i*** **of a node** ***i*** is the reciprocal of the total hopcount of all shortest path at this node *i* to all other nodes in the graph *G*:

$$Cl_i = \frac{1}{\sum_{j \in G \backslash \{i\}} H\left(P^*_{i \to j}\right)}$$

Concept of distance in both Cl_i and $(Q^\$)_{ii}$

- All paths [$(Q^\$)_{ii}$] versus only the shortest path [Cl_i]
- $(Q^\$)_{ii}$ lowerbounds Cl_i
  ($(Q^\$)_{ii} = Cl_i$ in trees with unit link weights)

$(Q^\$)_{ii}$ is more analytically tractable than a shortest path computation (closeness)

28

P. Van Mieghem, Performance Analysis of Complex Networks and Systems, Cambridge University Press, 2014, p. 370

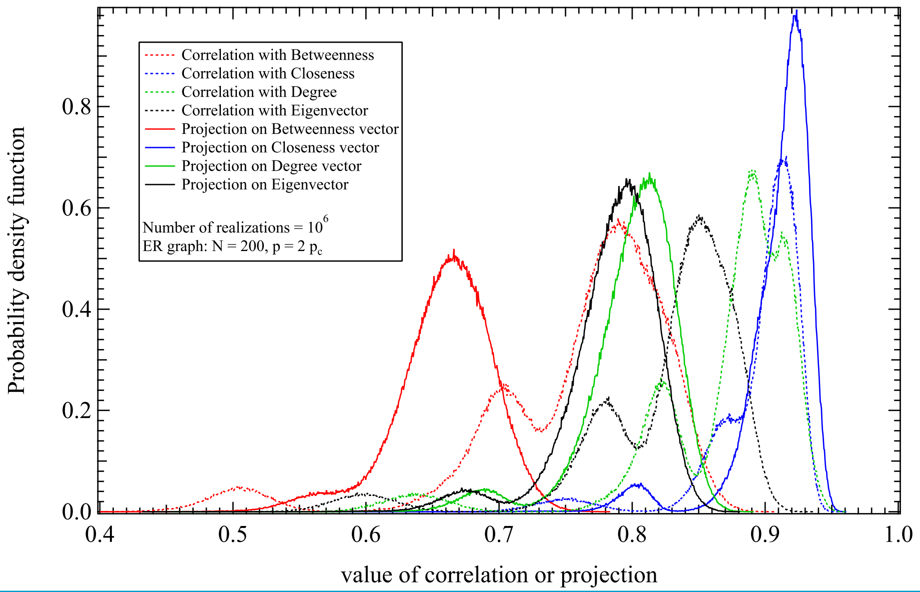
TU Delft

Contemplation: best spreader node

- *Valid for any weighted Laplacian*
  - Not heuristic, but based on the **law of conservation**
  - If resistances r_{ij} = 1: pure graph focus
  - The infinitesimal generator of a continuous-time Markov chain (MC) is minus a weighted Laplacian (which is not necessarily symmetric!)
    - Huge potential as nearly all processes can be approximated by a MC, provided the state space is sufficiently large
- *Interpretations*:
  - Ranking of nodes according to "diffusive centrality" or dynamic connectivity to all others
  - Resilience/Robustness: Safe-guarding nodes in this ranking to secure dynamic processes

29

TU Delft

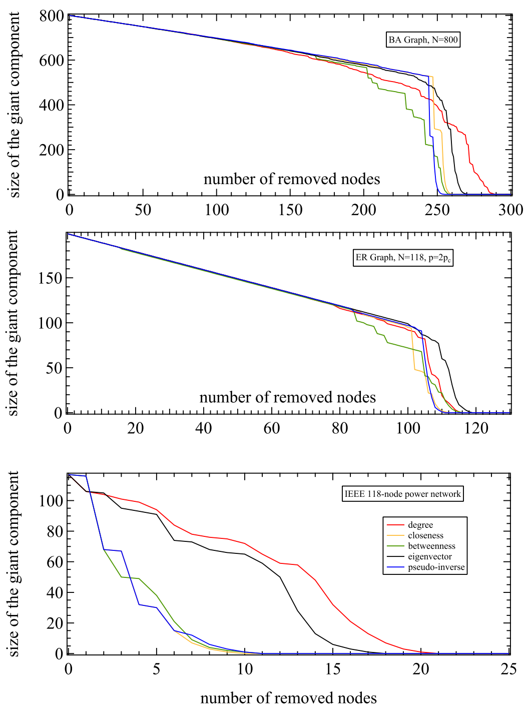


30

$$\zeta = \left(\left(Q^{\$}\right)_{11}, \left(Q^{\$}\right)_{22}, \ldots, \left(Q^{\$}\right)_{NN} \right)$$

TU Delft

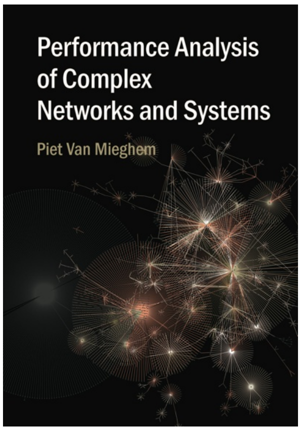
Destroying a network

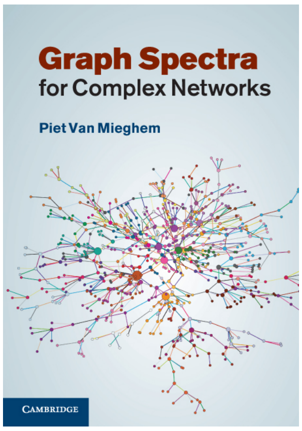


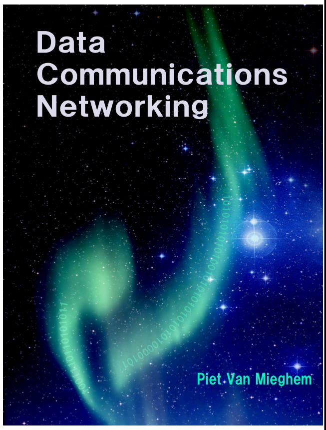
31

TU Delft

Books







Articles: **http://www.nas.ewi.tudelft.nl**

32

33