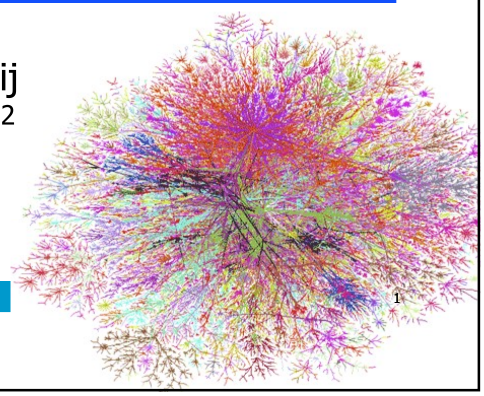
Robustness of Complex Networks

presentation at

ATLANTIC GRAPH THEORY SEMINAR

Robert Kooij
19 January 2022
TU Delft



1

About me

TU Delft

1988 - 1993

1994 - 1996

kpn

Royal Dutch Telecom

1997 - 2003

Applied → TNO

SUTD SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

Theoretical → TU Delft

I'm Back

2018 - 2019

2005 - now

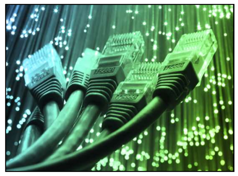
2

Robustness of Complex Networks

TU Delft

Introduction (1/2)

- Society is critically depending on complex networks



- Robustness: extent to which a complex network can cope with disruptions
  - failures of its nodes and/or links
- **Use graph theory to deal with robustness**

3

Introduction (2/2)

- How to quantify network robustness?
- What part of the network is most vulnerable?
- How to make the network more robust?



Critical Infrastructures

4

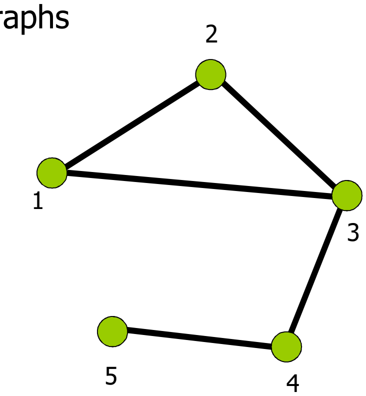
Outline

- Terminology
- Robustness w.r.t. malware spread
- Robustness of a gas distribution network
- Robustness of network controllability
- Wrap-up
- Bonus

TUDelft

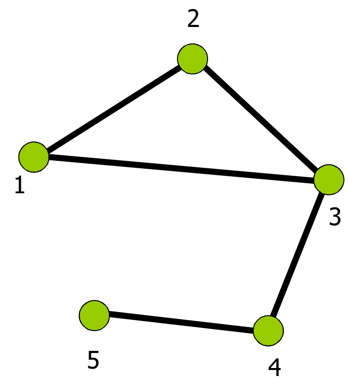
Terminology (1/4)

- Networks are represented as graphs
- Graph G(N,L)
  - N = number of nodes
  - L = number of links
- Graphs can be
  - undirected or directed
  - unweighted or weighted

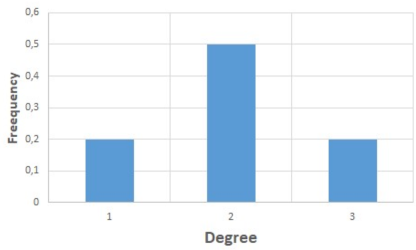


TUDelft

Terminology (2/4)



- degree D_i of node i
  - number of neighbours of node i
- degree distribution



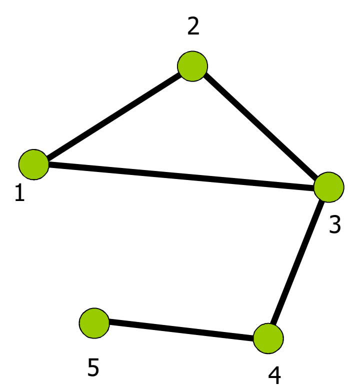
7

TU Delft

Terminology (3/4)

- Adjacency matrix

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



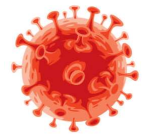
- ρ = spectral radius = largest eigenvalue of A

8

TU Delft

Terminology (4/4)

- The objects we study are NOT static
- Dynamical processes ON network
    -



- Network elements subject to stochastic process

- Methods from statistical physics
    - Mean-field approach
    - Simulations vs. models

9

TU Delft

Robustness w.r.t. malware spread (1/10)

- Spread of malware (malicious software)

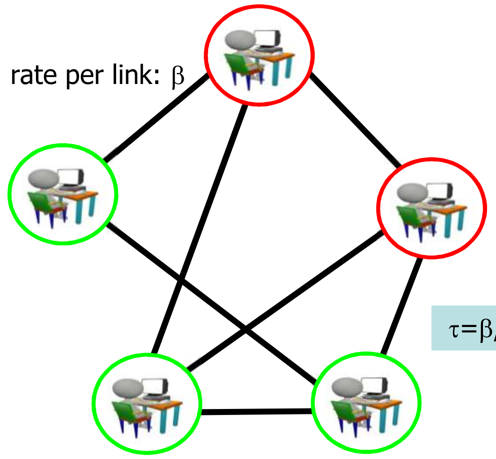
- Relation malware spread and network structure?

10

TU Delft

Robustness w.r.t. malware spread (2/10)

Infection rate per link: β



Curing rate per node: δ

τ=β/δ: effective spreading rate

Susceptible-Infected-Susceptible (SIS) model

TU Delft

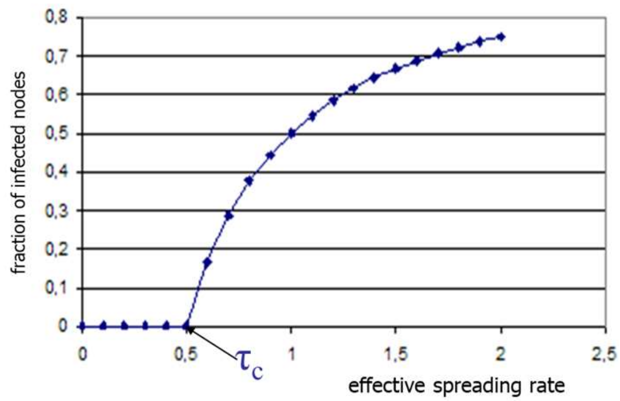
Robustness w.r.t. malware spread (3/10)

- Epidemic threshold τ_c
  - Effective spreading rate ≤ τ_c → malware dies
  - Effective spreading rate > τ_c → malware survives

$$\tau_c = \frac{1}{spectral\ radius}$$

TU Delft

Robustness w.r.t. malware spread (4/10)

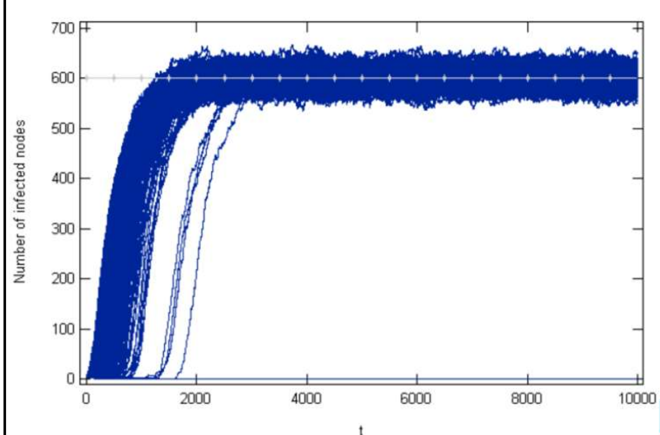


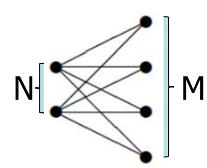
13

TUDelft

Robustness w.r.t. malware spread (5/10)

- Complete bi-partite graphs: $K_{N,M}$



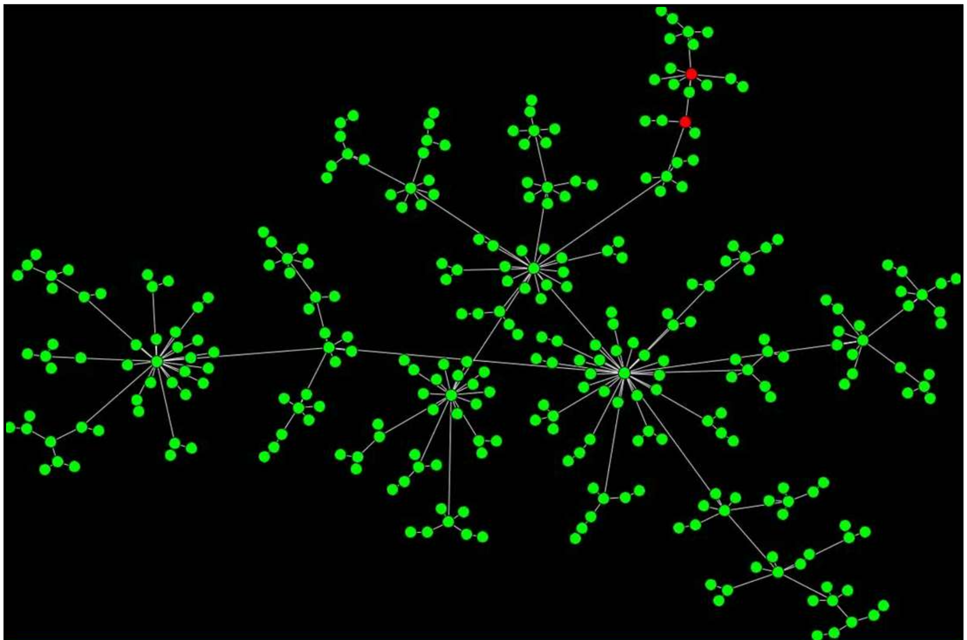


$$\tau_c = \frac{1}{\sqrt{MN}}$$

14

$K_{10,990}$: $\tau = 0.15 > 0.0101 = \tau_c$

TUDelft



Robustness w.r.t. malware spread (7/10)

- smaller ρ: more robustness against malware spread
- connected graphs: which topology has the smallest ρ ?
  - the path P_N 

$$\rho(P_N) = 2\cos(\frac{\pi}{N+1})$$

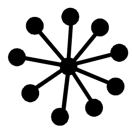
- what if we pose extra conditions?

16

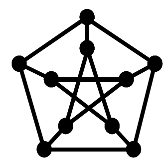
Robustness w.r.t. malware spread (8/10)

- Relation between minimal ρ and diameter of a graph?

- Graphs on N nodes with diameter 2:
  Minimal $\rho = \sqrt{N-1}$
  - Star topology
  - 3 additional cases : regular graphs (N = 5, 10, 50)



N =10



ρ = 3

TU Delft

Robustness w.r.t. malware spread (9/10)

- Found minimal ρ for *Diameter* $\in$ {⌊N/2⌋,N-3, N-2, N-1}

- And for nearly all graphs on at most 20 nodes

$D \setminus n$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2		1.4142	1.7321	2	2.2361	2.4495	2.6458	2.8284	3	3.1623	3.3166	3.4641	3.6056	3.7417	3.8730	4	4.1231	4.2426	4.358
3			1.6180	1.8478	2	2	2.3028	2.2784	2.4728	2.4860	2.5616	2.6970	2.7817	2.7321	2.8779	$\leq$ 2.9755	$\leq$ 3	$\leq$ 3.0742	$\leq$ 3
4				1.7321	1.9021	2	2	2	2.2361	2.2361	2.2230	2.3686	2.3778	2.3989	2.4303	2.5335	$\leq$ 2.7024	$\leq$ 2.7498	$\leq$ 2.79
5					1.8019	1.9319	2	2.0840	2	2	2.1701	2.2105	2.1987	2.1907	2.3028	2.3167	2.3228	2.3536	$\leq$ 2.54
6						1.8478	1.9499	2	2.0743	2.0743	2	2	2.1463	2.1940	2.1829	2.1753	2.1701	2.2688	$\leq$ 2.33
7							1.8794	1.9616	2	2.0684	2.1067	2.0684	2	2	2.1285	2.1693	2.1723	2.1649	2.15
8								1.9021	1.9696	2	2.0647	2.1010	2.1010	2.0647	2	2	2.1149	2.1505	2.16
9									1.9190	1.9754	2	2.0623	2.0912	2.1149	2.0912	2.0623	2	2	2.10
10										1.9319	1.9796	2	2.0608	2.0840	2.1120	2.1120	2.0840	2.0608	2
11											1.9419	1.9829	2	2.0598	2.0785	2.1054	2.1183	2.1054	2.07
12												1.9499	1.9854	2	2.0592	2.0743	2.1010	2.1169	2.11
13													1.9563	1.9874	2	2.0588	2.0710	2.0981	2.11
14														1.9616	1.9890	2	2.0586	2.0684	2.09
15															1.9659	1.9904	2	2.0584	2.06
16																1.9696	1.9915	2	2.05
17																	1.9727	1.9924	2
18																		1.9754	1.99
19																			1.97

Minimal ρ for *Diameter* = 3?

TU Delft

Robustness w.r.t. malware spread (10/10)

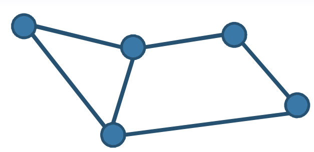
Virus spread in networks P Van Mieghem, J Omic, RE Kooij
IEEE/ACM Transactions On Networking 17 (1), 1-14, 2009

The minimal spectral radius of graphs with a given diameter
ER van Dam, RE Kooij
Linear Algebra and its Applications 423 (2-3), 408-419, 2008

19

TU Delft

Robustness of a gas distribution network

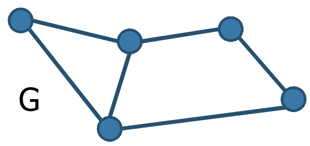


- N nodes
- L links
- undirected graph

- Network availability = Pr {network is **connected**}
  - Nodes always operational
  - Each link interdependently operational with probability *p*
  - All-terminal reliability

20

Reliability polynomial



$R_G(p)$ = Pr {G is connected}

$$R_G(p) = F_0 p^L + F_1(1-p)p^{L-1} + F_2(1-p)^2 p^{L-2} + .. + F_{L-N+1}(1-p)^{L-N+1} p^{N-1}$$

F_i: # of sets of i links, whose removal leave G connected

F_1 = 6

A case study

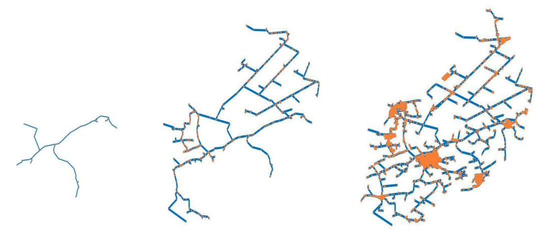
- Links: gas pipes
- Nodes: points where pipes connect

Network details

7240 households

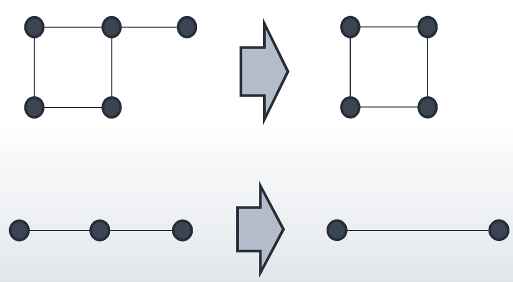
- One entry point for gas
- Network consists of three parts
  - 8, 3 and 0.1 bar

Network	8 bar	8 & 3 bar	Full network
Nodes	256	1845	20567
Links	255	1851	20749



Reductions on the network

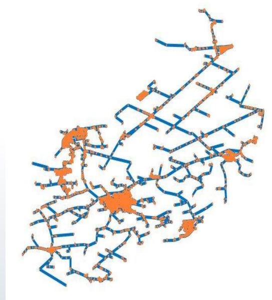
- Network is too large to process
- Reduce its size without loss of relevant information

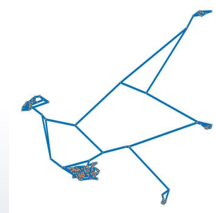


Shooman (1995)

Full network: significant reduction

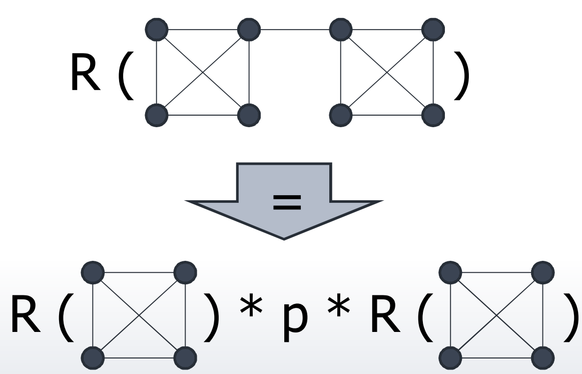
	Before	After
Nodes	20567	262
Edges	20749	393





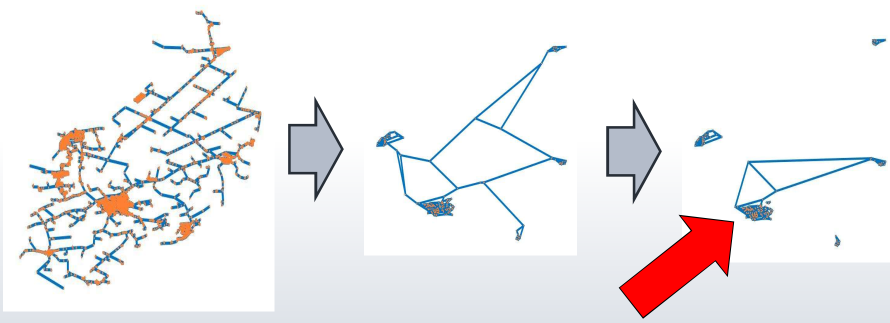
- Still too large for direct computation

Additional reductions



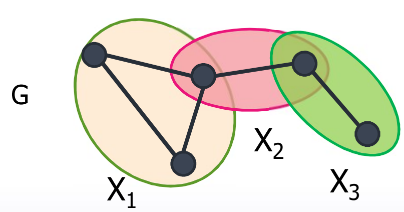
Result of split: 5 sub-networks

	Before	Sub 1	Sub 2	Sub 3	Sub 4	Sub 5
Nodes	20567	34	12	186	4	12
Edges	20749	51	18	279	6	18



Largest sub-network: decomposition

- Decomposition based upon **pathwidth** of graph



- Computation time polynomial in pathwidth(G)

Results

Can we compute the exact availability of our gas network?

– Computation takes about **2 minutes**
– Individual p values depend on
  - Soil type
  - Length of pipes

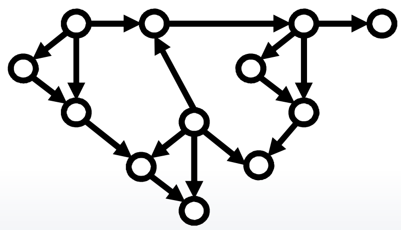
– Availability = 0.9919
  - 70 hours per year at least one node is disconnected
  - Assume every non-availability influences 3 households
  - Mean gas outage per household: 70*3600*3/7240 ≈ 104 seconds

Robustness of a gas distribution network

The reliability of a gas distribution network: A case study
W Pino, D Worm, R van der Linden, R Kooij
2016 International Conference on System Reliability and Science (ICSRS), 122-129

Robustness of network controllability

- Directed networks



- number of nodes = *N*
- number of links = *L*

Introduction to network control

$$\frac{d\boldsymbol{x}(t)}{dt} = A\boldsymbol{x}(t) + B\boldsymbol{u}(t)$$

$\mathbf{x}(t) = (x_1(t),\ldots\ldots,x_N(t))^T$: state of system at time t

$\mathbf{u}(t) = (u_1(t),\ldots\ldots,u_M(t))^T$: control input vector

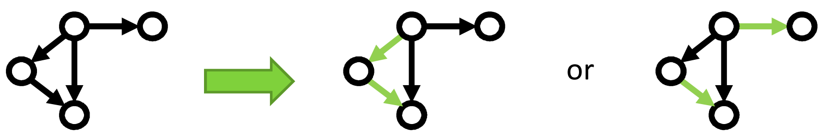
A: *NxN* matrix, describing systems connections

B: *NxM* input matrix, identifying nodes under outside control

- What is the minimum number of nodes that need to be controlled, to bring the system to a desired state?

Introduction to network control

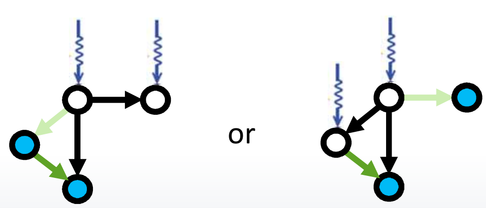
- How to find minimum number of driver nodes N_D?
- Through 'maximum matching' of network
  - maximum set of links that do not share start or end nodes



- Number of links in maximum matching is unique
- Maximum matching itself is NOT unique
- $O(N^{1/2}L)$ algorithm (Hopcraft-Karp) to find maximum matching

Introduction to network control

- Matched links point to matched nodes
- N_D = number of unmatched nodes



$N_D = 2$

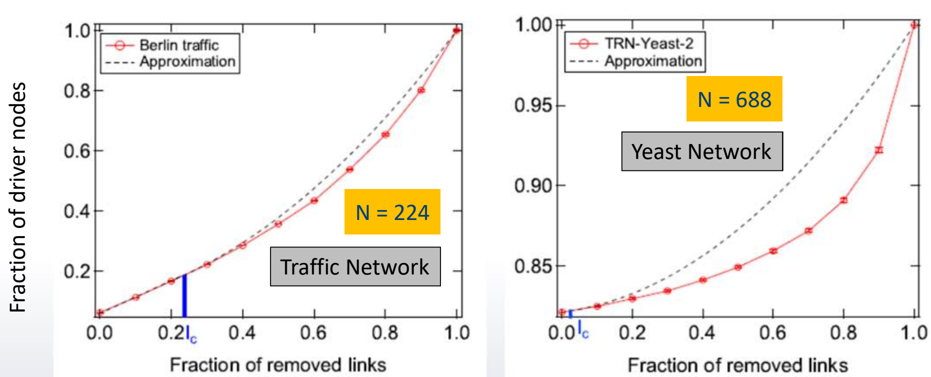
- Critical link: appears in every maximum matching
- l_C = fraction of critical links

Robustness of network controllability

- Assume links are removed from network
  - Random removal (failures)
  - Targeted removal (attacks)
- Number of driver nodes N_D will increase
- Analytic approximations for the increase in N_D
- Approximation
  - fraction of removed links $\leq l_C$: N_D linear in fraction of removed links
  - fraction of removed links $> l_C$: N_D quadratic in fraction of removed links

Robustness of network controllability

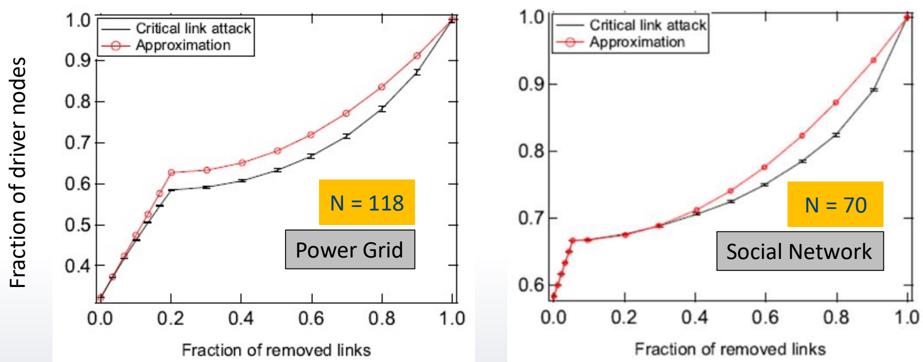
- Random link removal



Good results for small number of removals

Robustness of network controllability

- Targeted link removal



- Good results for small number of removals
- Approximation is worst-case

Robustness of network controllability

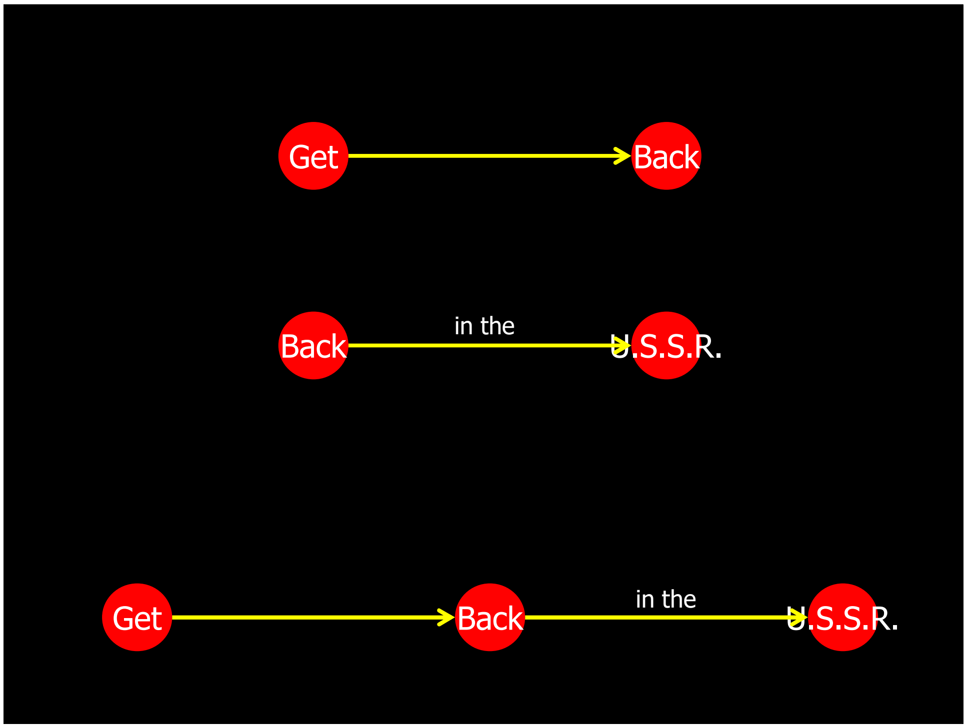
Quantifying the robustness of network controllability
P Sun, RE Kooij, Z He, P Van Mieghem
2019 4th International Conference on System Reliability and Safety

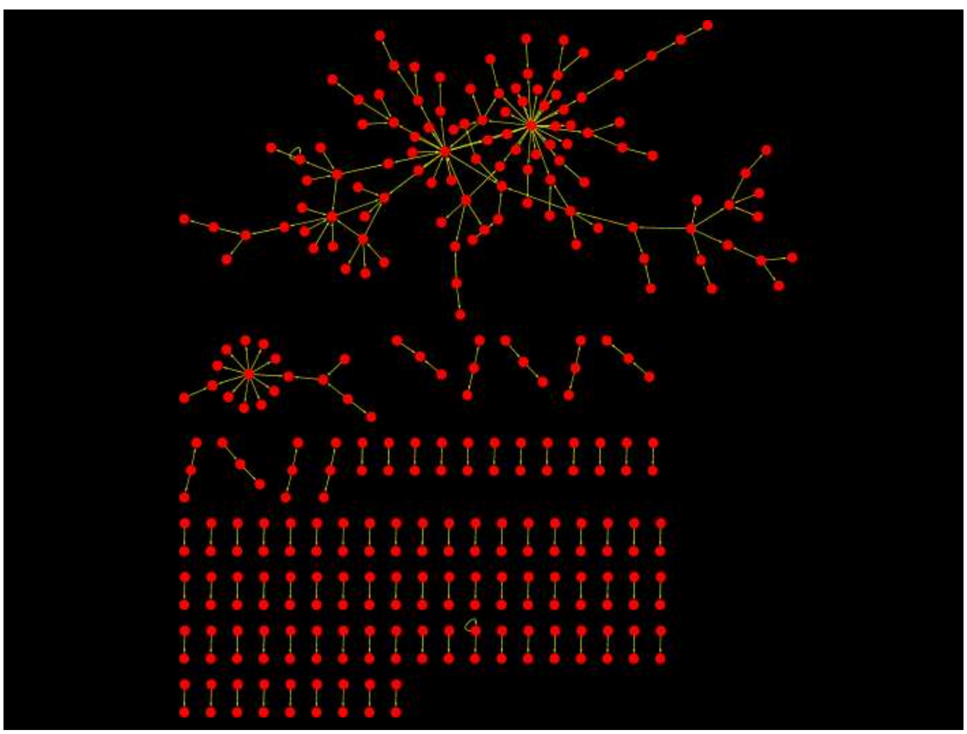
Wrap-up

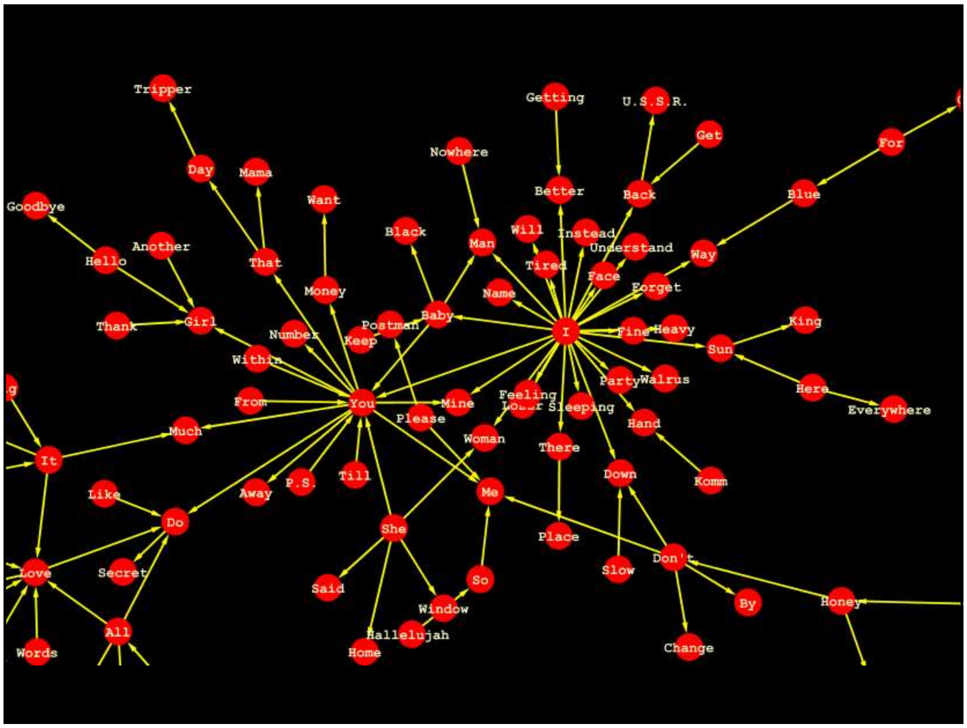
- Robustness of complex networks
- Societal relevance
- Quantification of robustness
  - Malware spread
  - Availability in gas distribution network
  - Network controllability
- Methods from statistical physics

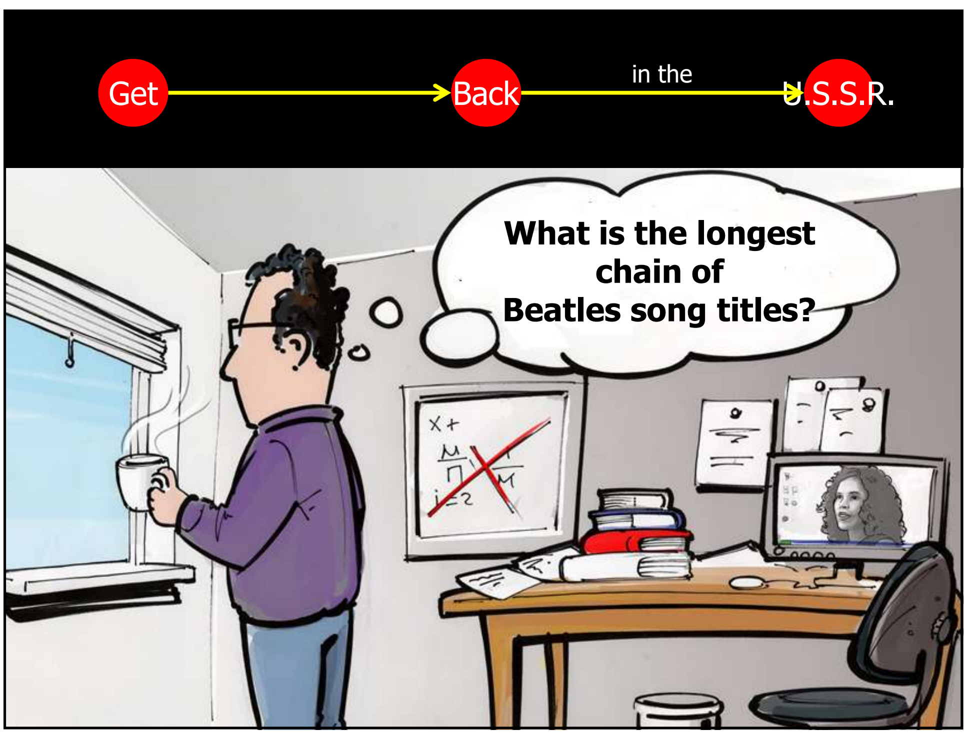
r.e.kooij@tudelft.nl

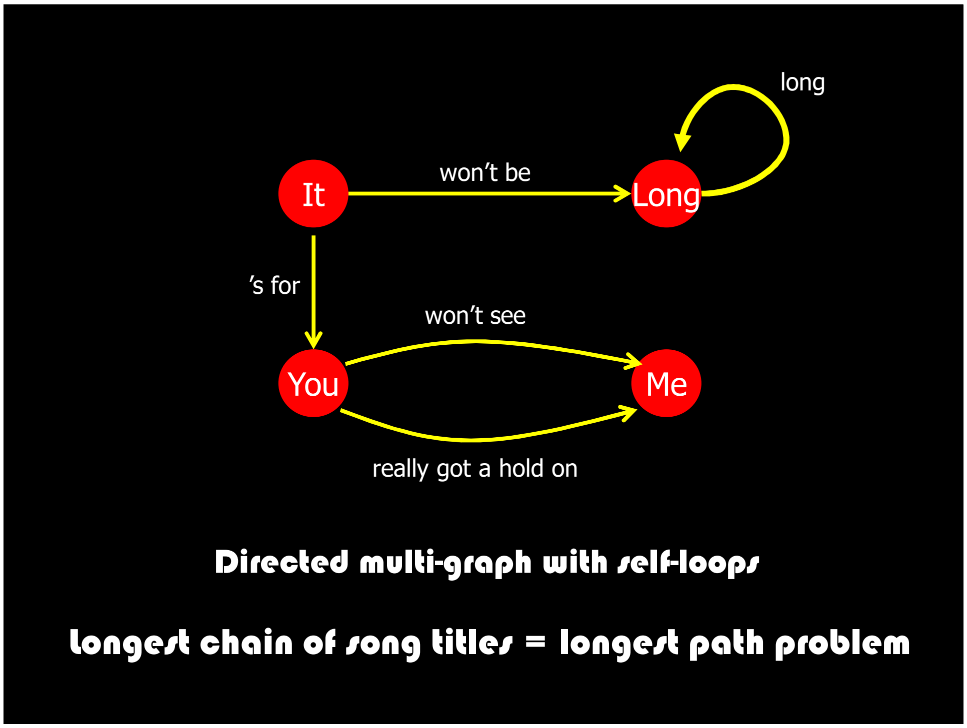
nas.ewi.tudelft.nl

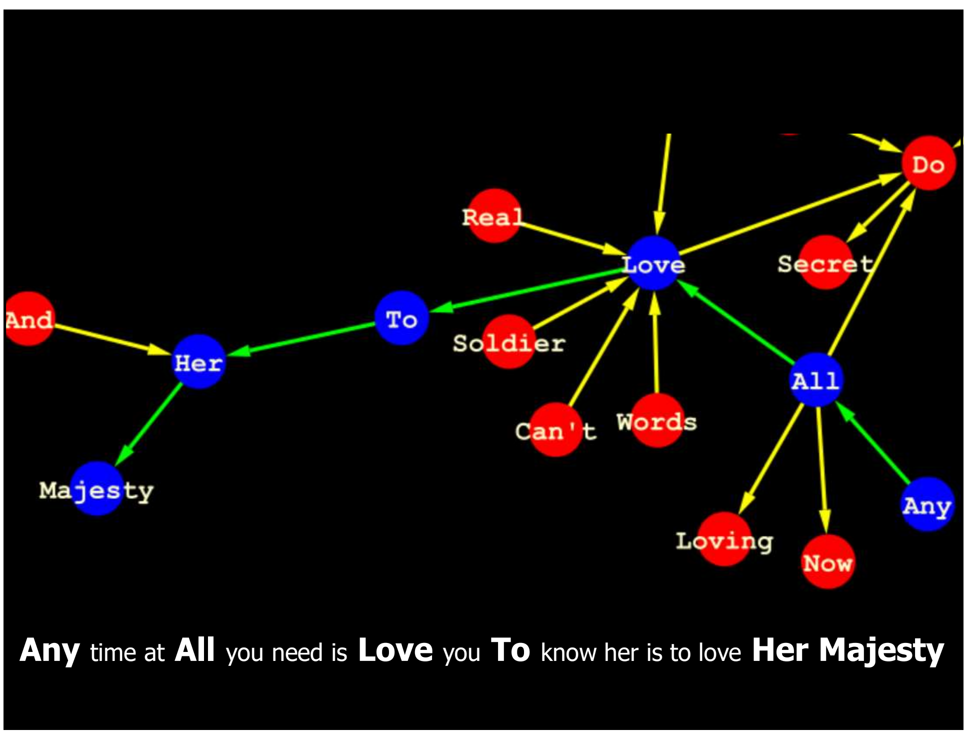

Get
Back
Back
in the
U.S.S.R.
Get
Back
in the
U.S.S.R.

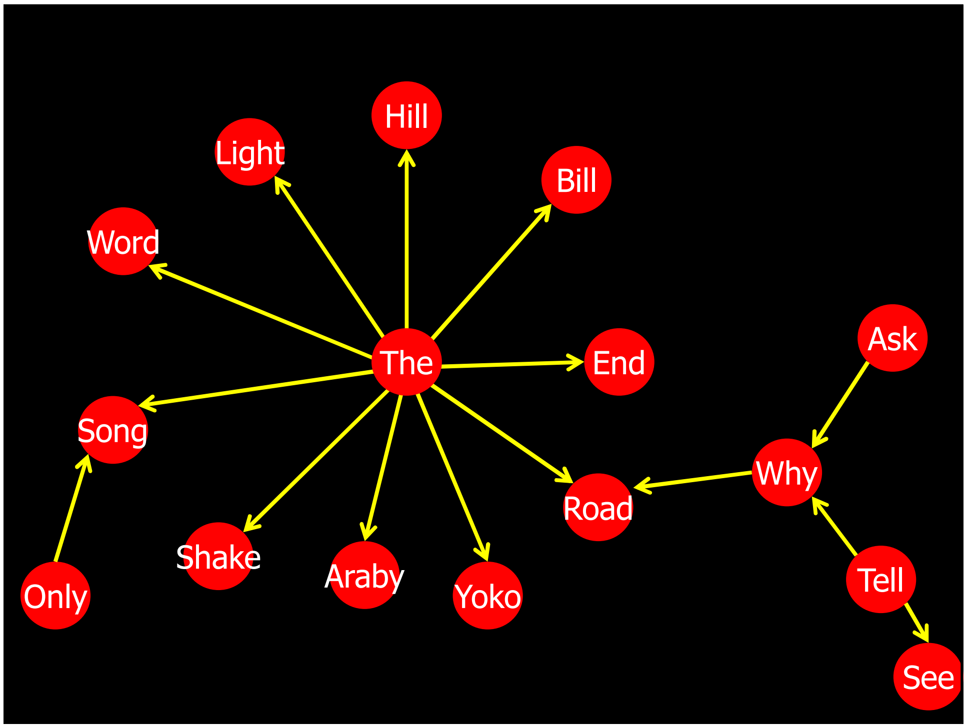


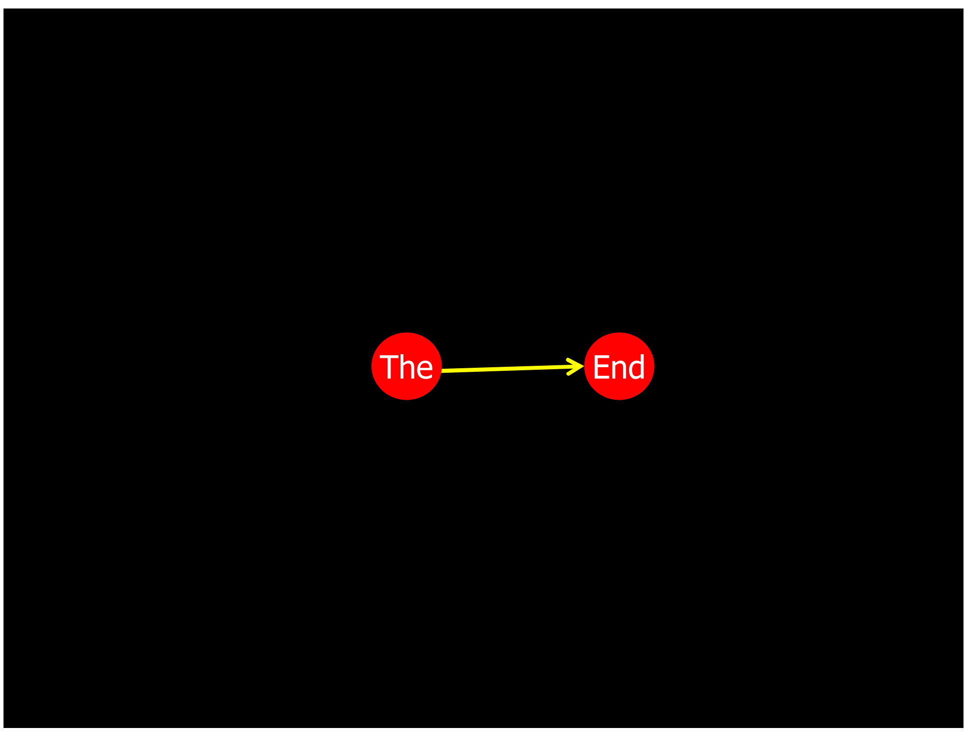

Tripper
Getting
U.S.S.R.
Get
For
Nowhere
Day
Mama
Better
Back
Blue
Want
Goodbye
Black
Will
Instead
Man
Understand
Way
Another
Hello
That
Tired
Face
Money
Name
Forget
Baby
Thank
Girl
Postman
Number
Keep
I
Fine
Heavy
King
Sun
Within
Party
Walrus
Here
Feeling
From
You
Mine
Sleeping
Everywhere
Please
Hand
Much
Woman
There
It
Down
Komm
Like
P.S.
Till
Away
Me
Do
She
Place
Don't
Slow
Secret
Love
So
Said
By
Honey
Window
All
Hallelujah
Words
Home
Change


Get
Back
in the
U.S.S.R.
What is the longest chain of Beatles song titles?


long
It
won't be
Long
's for
won't see
You
Me
really got a hold on
Directed multi-graph with self-loops
longest chain of song titles = longest path problem


Do
Real
Love
Secret
And
To
Her
Soldier
All
Can't
Words
Majesty
Any
Loving
Now
Any time at All you need is Love you To know her is to love Her Majesty


Hill
Light
Bill
Word
The
End
Ask
Song
Why
Road
Shake
Araby
Yoko
Tell
Only
See


The
End