Kemeny's constant for several families of graphs and real-world networks

Robert Kooij

7 July 2022

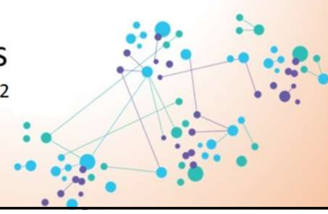
The Fifth Delft - Girona Workshop on

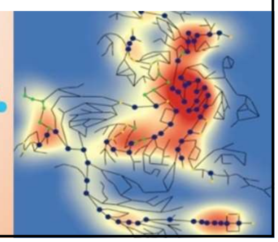
Robustness of Complex Networks

7 July 2022

TU Delft

Universitat de Girona





TNO innovation for life

TU Delft

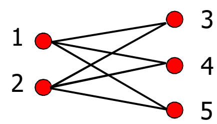
John George Kemeny

- János György Kemény (1926 - ??)
- Joined Manhattan Project aged 18!
- Invented BASIC programming language

TNO innovation for life

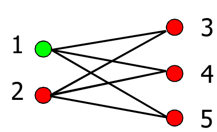
TU Delft

What is Kemeny's Constant?



- Undirected graphs
- Random walks

What is Kemeny's Constant?



- Undirected graphs
- Random walks

Pr(1→3) = 1/3
Pr(1→4) = 1/3
Pr(1→5) = 1/3

- Mean first passage time
  - $m_{13} = 5$

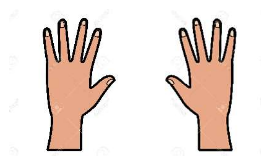
- Kemeny's Constant = mean first passage time from a given node to a randomly chosen other node

What is Kemeny's Constant?

- Kemeny's Constant = proxy for the robustness of a network
- Hence interest of the **Network Architectures and Services** group

- How do we increase the robustness of a network?

By increasing Kemeny's Constant

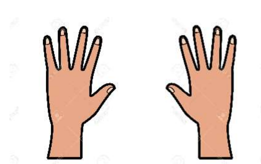


By decreasing Kemeny's Constant

What is Kemeny's Constant?

- Kemeny's Constant = proxy for the robustness of a network
- Hence interest of the **Network Architectures and Services** group

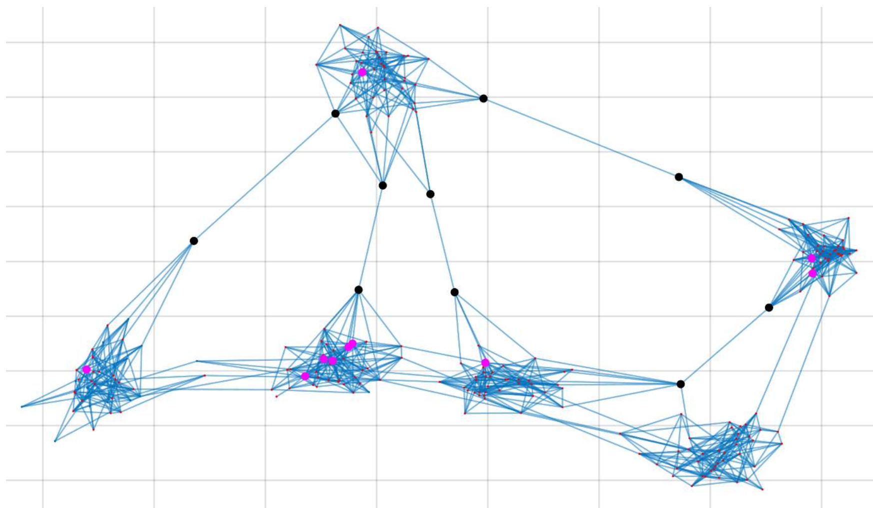
- How do we increase the robustness of a network?



By decreasing Kemeny's Constant

Applications

- Efficient testing for COVID in communities



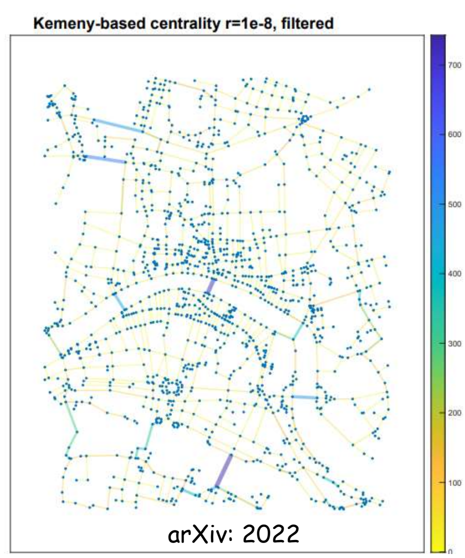
PLOS ONE: 2020

TNO innovation for life

TU Delft

Applications

- Detecting bottlenecks in road networks



arXiv: 2022

TNO innovation for life

TU Delft

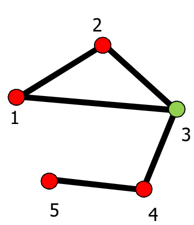
Background

- Joint work with Johan Dubbeldam (DIAM)

Discrete Applied Mathematics 285 (2020) 96–107

Background

- A: adjacency matrix of graph G(N,L)
- Δ: diagonal degree matrix



Pr(3→1) = 1/3
Pr(3→2) = 1/3
Pr(3→4) = 1/3

- P: transition probability matrix

$$P = \Delta^{-1} A$$

- π: steady state probability vector

$$\pi^T P = \pi^T$$

Background

- Q: Laplacian matrix $Q = \Delta - A$
- R_G: effective graph resistance $R_G = N \sum_{i=1}^{N-1} \mu_i$
- $Q^{\dagger}$: Moore-Penrose pseudo inverse of Laplacian $R_G = N trace(Q^{\dagger})$

$$Q^{\dagger} = (Q + \frac{1}{N}J)^{-1} - \frac{1}{N}J$$

TNO innovation for life

TU Delft

Background

$$K(P) = \sum_{i=1}^{N} \pi_i m_{ji} - 1$$

m_{ji}: mean first passage time

$$K(P) = \sum_{k-2}^{N} \frac{1}{1-\lambda_k}$$

λ_k: eigenvalues of $\Delta^{-1/2} A \Delta^{-1/2}$

$$K(P) = \zeta^T d - \frac{d^T Q^{\dagger} d}{2L}$$

ζ: vector with diagonal elements $Q^{\dagger}$

d: degree vector

Wang, Dubbeldam, Van Mieghem (LAA;2017)

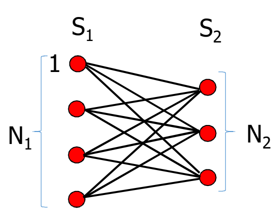
TNO innovation for life

TU Delft

Overview

- Complete bi-partite graphs and some trees
- Windmill graphs
- Relation with effective graph resistance
- A sharp upper bound
- Real-world networks
- Synergy with BDCS?
- Wrap-up

Complete bi-partite graphs and some trees



$K_{N1,N2}$

$$K(P) = N1 + N2 - \frac{3}{2}$$

$$P = \begin{pmatrix} 0_{N1\times N1} & \frac{1}{N_1}J_{N1\times N2} \\ \frac{1}{N_2}J_{N2\times N1} & 0_{N2\times N2} \end{pmatrix}$$

$$\pi^T = \left(\frac{1}{2N_1} \cdots \frac{1}{2N_1}; \frac{1}{2N_2} \cdots \frac{1}{2N_2}\right)$$

Complete bi-partite graphs and some trees

Mean passage time from node 1 to a specific node in S2

$$K(P) = \sum_{i=1}^{N1+N2} \pi_i m_{1i} - 1 \qquad = \frac{1}{2} m_{1S1} + \frac{1}{2} m_{1S2} - 1$$

Mean passage time from node 1 to a specific node in S_1

Condition on first jump

$$m_{1S} = \frac{1}{N2} + \left(1 - \frac{1}{N2}\right)(2 + m_{1S1}) \quad \Rightarrow \quad m_{1S2} = 2N2 - 1$$

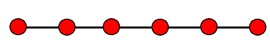
Similarly $m_{S21} = 2N1 - 1$

$$m_{1S1} = 1 + m_{S21} = 2N1$$

Mean passage time from specific node in S2 to node 1

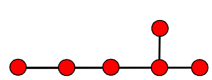
Complete bi-partite graphs and some trees

P_N



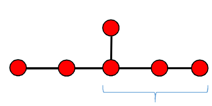
$$K(p) = \frac{1}{3}N^2 - \frac{2}{3}N + \frac{1}{2}$$

D_N



$$K(p) = \frac{1}{3}N^2 - \frac{2}{3}N + \frac{1}{2} - \frac{2(N-3)}{N-1}$$

$E_{N,M}$

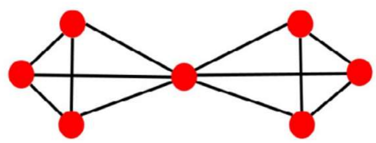


M

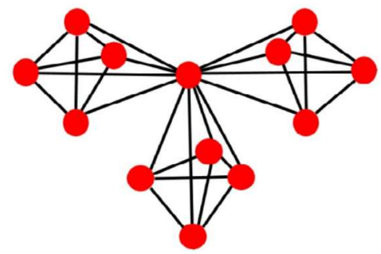
$$K(p) = \frac{1}{3}N^2 - \frac{2}{3}N + \frac{1}{2} - \frac{2(M(N-3) - M(M-1))}{N-1}$$

Windmill graphs

- Introduced by Estrada (2016)
  - η cliques of order k
  - connected through central node



W(2,3)



W(3,4)

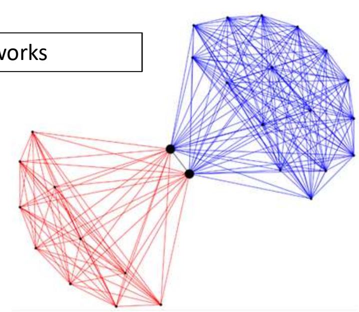
TNO innovation for life

TU Delft

Windmill graphs

- Generalized windmill Type I: W'(η,k,l)
  - η cliques of order k
  - connected through central clique of order l

Application: Public Transport Networks



TNO innovation for life

TU Delft

Windmill graphs

- K(P) for W'(η,k,l) using spectrum of $\Delta^{-1/2} A \Delta^{-1/2}$

$$K(P) = \frac{\eta(k-1)(k+l-1)}{k+l} + \frac{(\eta-1)(k+l-1)}{l} + \frac{(l-1)(\eta k+l-1)}{\eta k+l} + \frac{(\eta k+l-1)(k+l-1)}{\eta k(k+l-1)+l(\eta k+l-1)}.$$

Special case: l = 1; windmill graph W(η,k) $K(P) = \frac{k^2(2\eta-1)}{k+1}$

TNO innovation for life

TU Delft

Relation with effective graph resistance

- Palacios (2010) $K\,(P) = \frac{r}{N} R_G$
  - r-regular graphs on N nodes
- Approximation for non-regular graphs
  - Average degree D
  - Heterogeneity index H

$$D = \frac{2L}{N}$$

$$H = \frac{1}{N}\sum_{i=1}^{N}(d_i - D)^2$$

$$K^*(P) = \frac{D}{N} R_G + Hf(N,L)$$

- Choose f(N,L) such that K*(P) is exact for $K_{N1,N2}$

TNO innovation for life

TU Delft

Relation with effective graph resistance

- for $K_{N1,N2}$

$$K(P) = N1 + N2 - \frac{3}{2}$$

$$D = \frac{2N_1N_2}{N_1 + N_2} \qquad H = \frac{N_1N_2(N_1 - N_2)^2}{(N_1 + N_2)^2} \qquad R_G = (N_1 + N_2)(\frac{N_2 - 1}{N_1} + \frac{N_1 - 1}{N_2} + \frac{1}{N_1 + N_2})$$

$$f = \frac{1 - 2N_1 - 2N_2}{2N_1N_2}$$

- for $K_{N1,N2}$

$$N = N1 + N2$$

$$L = N1N2$$

$$f(N, L) = \frac{1 - 2N}{2L}$$

TNO innovation for life

TU Delft

Relation with effective graph resistance

$$K^*(P) = \frac{2L}{N^2} R_G + H \frac{1 - 2N}{2L}$$

- Exact for regular graphs
- Exact for complete bipartite graphs
- Also exact for windmill graphs!

TNO innovation for life

TU Delft

Relation with effective graph resistance

Table 1
Kemeny's constant and its approximations K^* for several graphs.

Graph	N	L	$K(P)$	$K^*(P)$
$K_{10,15}$	25	150	23.50	23.50
P_{10}	10	9	27.17	29.53
D_{10}	10	9	25.61	28.06
$E_{10,2}$	10	9	24.50	27.16
$W(3, 10)$	31	165	45.45	45.45
$W'(3, 10, 5)$	35	295	35.49	33.77
$W''(3, 10, 5)$	35	285	35.54	34.67

TNO innovation for life

TU Delft

A sharp upper bound

- K*(P) is not an upper bound

$$K(P) = \zeta^T d - \frac{d^T Q^\dagger d}{2L} \leq \zeta^T d - \frac{H}{D\mu_1} \equiv K_U(P)$$

Wang, Dubbeldam, Van Mieghem (LAA;2017)

TNO innovation for life

TU Delft

A sharp upper bound

Table 2
Kemeny's constant and the upper bound $K_U(P)$, for several graphs.

Graph	$K(P)$	$K_U(P)$
$K_{10,15}$	23.50	23.50
P_{10}	27.17	27.28
D_{10}	25.61	25.71
$E_{10,2}$	24.50	24.61
$W(3, 10)$	45.45	45.45
$W'(3, 10, 5)$	35.49	35.49
$W''(3, 10, 5)$	35.54	35.54

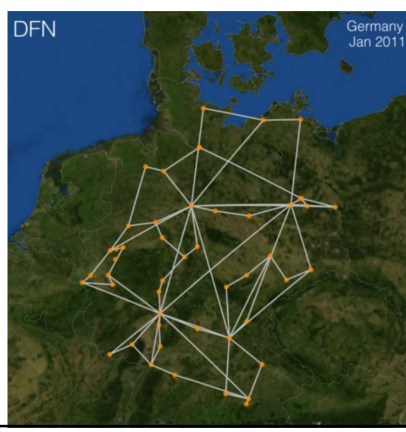
- $K_U(P)$ is tight for (generalized) windmill graphs

TNO innovation for life

TUDelft

Real-world networks

- Data from Internet Topology Zoo
  - 243 communication networks



TNO innovation for life

TUDelft

Real-world networks

Table 3
Kemeny's constant, the approximations K^* and the upper bound K_U, for the smallest and largest networks in the Internet Topology Zoo.

Graph	N	L	$K(P)$	K^*	K_U
Arpanet196912	4	4	2.54	2.73	2.60
Renam	5	4	3.50	3.50	3.50
Mren	6	5	4.50	4.50	4.50
UsCarrier	158	189	1175.99	1265.48	1176.68
Cogentco	197	245	1082.45	1197.24	1083.35
Kdl	754	899	5907.29	6264.78	5908.32

innovation for life

TU Delft

Real-world networks

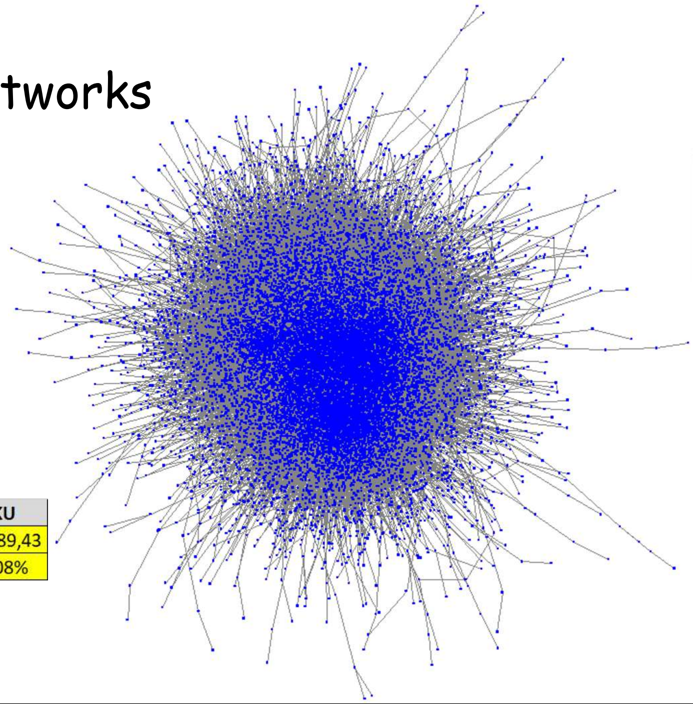
Table 4
Statistics for the absolute value of the relative errors for the approximations K^* and the upper bound K_U, for 243 real-world networks.

Metric	K^*	K_U
Average absolute rel. error	27.25%	0.73%
Maximum absolute rel. error	122.60%	8.05%

TNO innovation for life

TU Delft

Real-world networks



bio-CE-CX.edges (undirected)	
Summary Statistics	
Number of nodes	15063
Number of edges	245862

Network	K	K*	KU
bio-CE-CX	17375,46	128159,7	17389,43
	Rel. error	638%	0,08%

TNO innovation for life

TU Delft

Synergy with BDCS?

- Kemeny's Constant implemented in NRS
- Use Kemeny's Constant as a centrality indicator

International Journal of Critical Infrastructure Protection 33 (2021) 100422

Resilient backup controller placement in distributed SDN under critical targeted attacks☆,☆☆

Eusebi Calle[a,*], David Martínez[a], Mariusz Mycek[b], Michał Pióro[b]

[a] Institute of Informatics and Applications, Universitat de Girona, Girona, Spain
[b] Institute of Telecommunications, Warsaw University of Technology, Warsaw, Poland

TNO innovation for life

TU Delft

Wrap-up

- Explicit expressions for K for some graph families
- An approximation using effective graph resistance
- A sharp upper bound
- Validation on real-world networks
- Kemeny's constant as robustness metric

TNO innovation for life

TU Delft

Thanks for your attention!

- r.e.kooij@tudelft.nl

TNO innovation for life

TU Delft