Universal mean-field framework (UMFF) for SIS epidemics on networks

Piet Van Mieghem

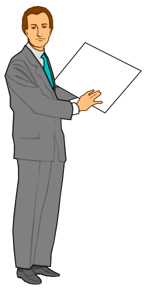
in collaboration with Karel Devriendt

Advances on Epidemics in Complex Networks
Delft
31 August-1 September 2017

NAS

TUDelft

Outline



Exact SIS prevalence

Universal mean-field framework (UMFF)

- Idea
- Framework

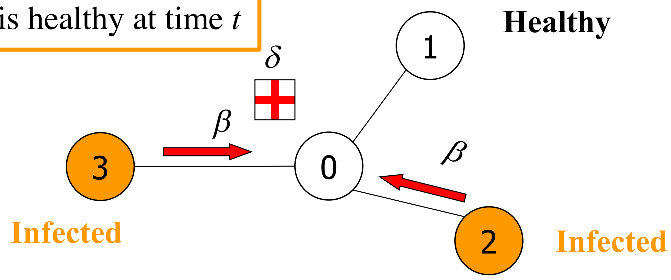
TUDelft

Continuous-time SIS model on networks

- Constant infection rate β on all links
- Constant curing rate δ for all nodes

  $\tau = \beta / \delta$: effective spreading rate

$X_j(t) = 1$	node j is infected at time t
$X_j(t) = 0$	node j is healthy at time t



Infection and curing are independent Poisson processes

P. Van Mieghem, J. Omic, R. E. Kooij, "Virus Spread in Networks", IEEE/ACM Transaction on Networking, Vol. 17, No. 1, pp. 1-14, (2009).

TUDelft

Governing SIS equation for node *j*

$$\frac{dE[X_j]}{dt} = E\left[-\delta X_j + (1 - X_j)\beta \sum_{k=1}^{N} a_{kj} X_k\right]$$

time-change of $E[X_j] = \Pr[X_j = 1]$, probability that node j is infected

if *infected*: probability of curing per unit time

if *not infected (healthy)*: probability of infection per unit time

$$\frac{dE[X_j]}{dt} = -\delta E[X_j] + \beta \sum_{k=1}^{N} a_{kj} E[X_k] - \beta \sum_{k=1}^{N} a_{kj} E[X_j X_k]$$

4

R. Pastor-Satorras, C. Castellano, P. Van Mieghem and A. Vespignani, "Epidemic processes in complex networks", Review of Modern Physics, 2015

TUDelft

SIS Prevalence

- Fraction of infected nodes in the graph G

$$S(t) = \frac{1}{N}\sum_{j=1}^{N} X_j(t) \qquad \text{(random variable!)}$$

- Prevalence: Expected fraction of infected nodes in G

$$y(t) = E[S(t)] = \frac{1}{N}\sum_{j=1}^{N} \Pr[X_j(t) = 1]$$

P. Van Mieghem, 2016, "Approximate formula and bounds for the time-varying SIS prevalence in networks", Physical Review E, Vol. 93 No. 5, p. 052312.

Differential equation prevalence

Summing $\frac{dE[X_j]}{dt} = E\left[-\delta X_j + (1-X_j)\beta\sum_{k=1}^{N} a_{kj}X_k\right]$ over all nodes:

$$\frac{d}{\delta dt}\left(\frac{1}{N}\sum_{j=1}^{N} E[X_j]\right) = E\left[-\frac{1}{N}\sum_{j=1}^{N} X_j + \frac{\tau}{N}\sum_{j=1}^{N}\sum_{k=1}^{N}(1-X_j)a_{kj}X_k\right]$$

Using the definition of prevalence $y(t) = \frac{1}{N}\sum_{j=1}^{N} E[X_j]$ and

$\sum_{j=1}^{N}\sum_{k=1}^{N}(1-X_j)a_{kj}X_k = (u-w)^T Aw$ where w is the nodal random Bernoulli vector $w = (X_1, X_2, \ldots, X_N)$

$$\frac{dy(t^*)}{dt^*} = -y(t^*) + \frac{\tau}{N}E[(u-w)^T Aw]$$

Finally, if A is symmetric, the SIS prevalence is written in terms of the Laplacian $Q = \Delta - A$ and the normalized time $t^* = \delta t$

$\frac{dy(t^*)}{}$

P. Van Mieghem, F. Darabi Sahneh and C. Scoglio, 2014, "Exact Markovian SIR and SIS epidemics on networks and an upper bound for the epidemic threshold", Proceedings of the 53rd IEEE Conference on Decision and Control (CDC'14), December 15-17, Los Angeles, CA, USA (also on http://arxiv.org/abs/1402.1731).

“Local rule - global emergent properties” class

$$\frac{dE[X_j(t)]}{dt} = E\left[-\delta X_j(t) + (1 - X_j(t))\beta \sum_{k=1}^{N} a_{kj} X_k(t)\right]$$

Local SIS rule

Global emergent SIS spread

$$\frac{dy(t^*)}{dt^*} = -y(t^*) + \frac{\tau}{N} E\left[w^T(t^*) Q w(t^*)\right]$$

The Laplacian $Q = \Delta - A$
The normalized time t* = δ t
Bernoulli state vector
$w(t^*) = (X_1(t^*), X_2(t^*), \ldots, X_N(t^*))$

P. Van Mieghem, 2016, "Approximate formula and bounds for the time-varying SIS prevalence in networks", Physical Review E, Vol. 93 No. 5, p. 052312.

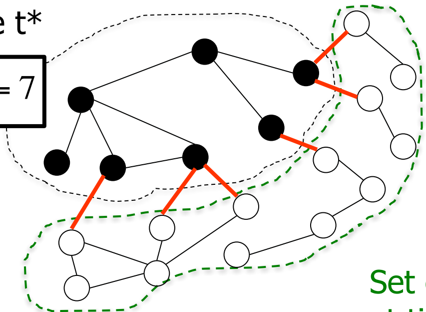
TU Delft

SIS prevalence dynamics

$$\frac{dy(t^*)}{dt^*} = -y(t^*) + \frac{\tau}{N} E\left[w^T(t^*) Q w(t^*)\right]$$

Set of infected nodes at time t*

$NS(t^*) = 7$



$w^T(t^*) Q w(t^*) = 6$

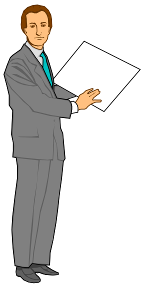
Cut-Set: set of links with 1 infected node at time t*

Set of susceptible nodes at time t*

P. Van Mieghem, 2016, "Universality of the SIS prevalence in networks", Delft University of Technology, report20161006 (http://arxiv.org/abs/1612.01386).

TU Delft

Outline



Exact SIS prevalence

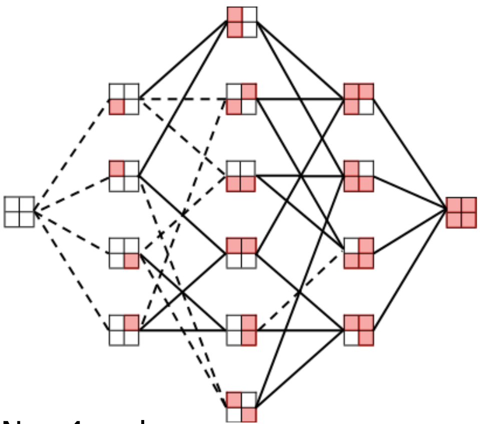
Universal mean-field framework (UMFF)

- **Idea**
- Framework

Devriendt, K. and P. Van Mieghem, 2017, "Universal mean-field framework for SIS epidemics on networks, based on graph partitioning and the isoperimetric inequality", arXiv 1706.10132

TUDelft

Continuous-time Markovian SIS epidemics on networks



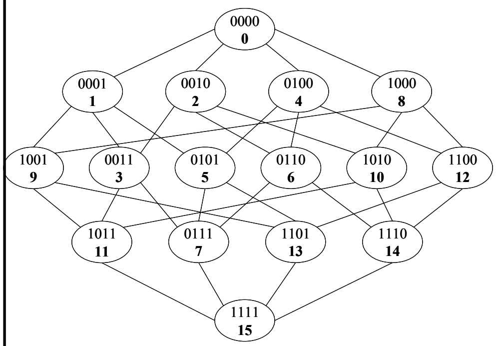
N = 4 nodes

Problems:
- 2^N states
- Complex structure
- Insight?

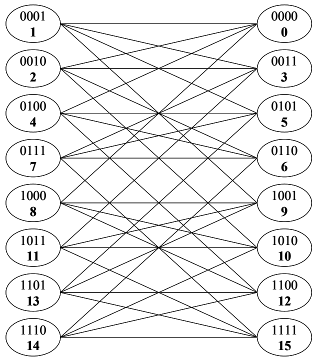
Approximate

TUDelft

Markov theory



Regular bipartite Markov graph



Recursive structure of infinitesimal general Q_N

A. Economou, A. Gómez-Corral, M. López-García, A stochastic SIS epidemic model with heterogeneous contacts, Physica A, Volume 421, 1 March 2015, Pages 78-97

Simon, P., M. Taylor and I. Z. Kiss, Exact epidemic models on graphs using graph-automorphism driven lumping, Mathematical Biology, Vol. 62, pp. 479-508, 2011

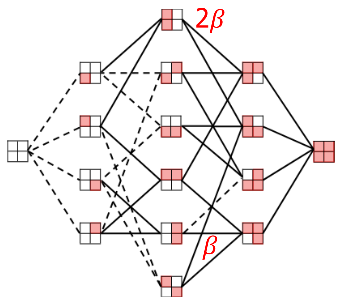
11

Van Mieghem, P. and E. Cator, ε-SIS epidemics and the epidemic threshold, Physical Review E, vol. 86, No. 1, July, p. 016116, 2012

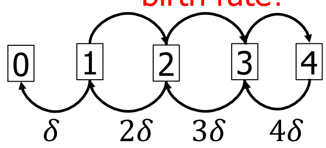
TUDelft

UMFF (1): New variable = #infected nodes

2^N States



N+1 states



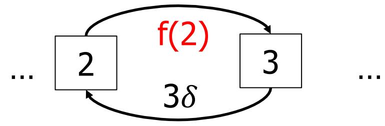
Problem:
Infection rate = β x # infective links

12

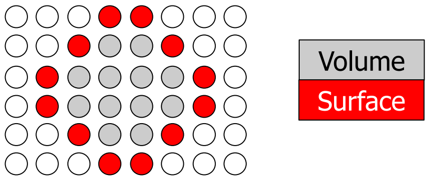
TUDelft

UMFF (2): approximation

$$\# \, Infective\ links \approx f(\# Infected\ nodes)$$



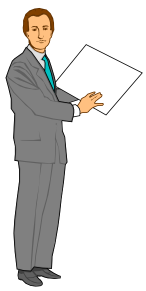
idea: isoperimetric problem in geometry



13

TU Delft

Outline



Exact SIS prevalence

Universal mean-field framework (UMFF)

- Idea
- **Framework**

14

TU Delft

Universal Mean-Field framework

- UMFF = General approximation framework for SIS:
  - Contains existing methods (NIMFA, HMF, pQMF, …)
  - Bounds on approximations

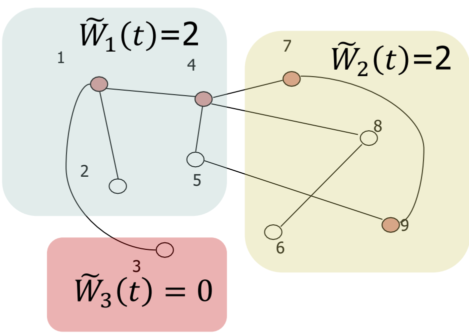
- **UMFF principles**:
  - Graph partitioning
  - Two approximation steps:
    - Topological approximation: **isoperimetric inequality**
    - Moment-closure approximation

Devriendt, K. and P. Van Mieghem, 2017, "Universal mean-field framework for SIS epidemics on networks, based on graph partitioning and the isoperimetric inequality", arXiv 1706.10132

Partitioning

- Graph partitioned into K non-overlapping subgraphs
- Count #infected in each partition

$$W(t) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$



$$\widetilde{\boldsymbol{W}}(t) = \begin{bmatrix} \widetilde{W}_1(t) \\ \widetilde{W}_2(t) \\ \widetilde{W}_3(t) \end{bmatrix}$$

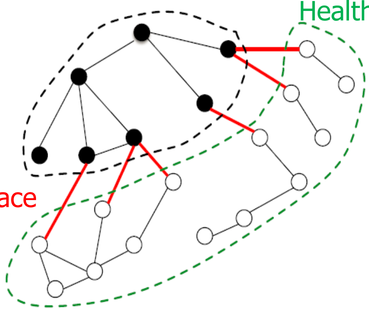
$$0 \le \widetilde{W}_1(t) \le 4$$
$$0 \le \widetilde{W}_2(t) \le 4$$
$$0 \le \widetilde{W}_3(t) \le 1$$

Closing the equations: idea

$$\frac{d\mathbf{E}[y(t)]}{dt} = -\delta\mathbf{E}[y(t)] + \frac{\beta}{N}\mathbf{E}[w^T Qw]$$

$$\frac{d\mathbf{E}[y(t)]}{dt} = -\delta\mathbf{E}[y(t)] + \beta f(\mathbf{E}[y(t)])$$

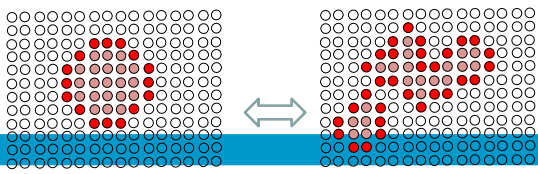
isoperimetric idea: "Volume" instead of "surface"



TUDelft

The isoperimetric inequality (1)

- iso-perimetric (η ιση περιοδος) → *"same perimeter"*
  - **Ancient Greeks**: what is the maximal area A that can be enclosed by a curve with a given perimeter P
  - **Solution**: In the plane, $P^2 \geq 4\pi A$ holds with equality for the circle
- *Generalizations* (20^{th} century)
  - Higher dimensions, curved space, manifolds, **graphs**, …
- *Cut-set*: volume vs surface!



V. Blasjo. "The evolution of the isoperimetric problem" Mathematical Association of America, 2005

TUDelft

The isoperimetric inequality (2)

$$\left|w^TQw - \frac{d_{av}}{N-1}Y(N-Y)\right| \le \frac{\max_{1\le i\le N-1}\left|\frac{N}{N-1}d_{av}-\mu_i\right|}{N}Y(N-Y)$$

exact for complete graph (all non-zero $\mu_i = N$)

$$\frac{d\mathbf{E}[y(t)]}{dt} = -\delta\mathbf{E}[y(t)] + \frac{\beta}{N}\mathbf{E}[w^TQw]$$

$$\frac{dY}{dt} = -\delta Y + \beta Y(N-Y)$$

logistic diff. eq.

Isoperimetric ineq. relates to Szemeredi's regularity theorem

Eigenvalues of Laplacian Q: $\mu_N = 0 \le \mu_{N-1} \le \cdots \le \mu_1$; Average degree: d_{av};
Number of infected nodes: Y

19

Chung, F. Discrete Isoperimetric Inequalities. DMTCS, 1996

TUDelft

UMFF

- SIS dynamics of **Y (=number of infected nodes)**
  - ✓ Split nodes in K partitions
  - ✓ Graph consists of disconnected and bipartite graphs

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} 0 & A_{12} \\ A_{21} & 0 \end{bmatrix} + \begin{bmatrix} A_{11} & 0 \\ 0 & A_{22} \end{bmatrix}$$

- **UMFF equations for the prevalence of partition k**

$$\frac{d\mathbf{E}[Y_k]}{dt} = -\delta\mathbf{E}[Y_k] + \beta\sum_{m=1}^{K}\frac{L_{km}}{N_kN_m}(N_k - \mathbf{E}[Y_k])\mathbf{E}[Y_m]$$

Curing

Infection due to each other partition m
- $m \neq k$: Bipartite approximation
- $m = k$: Classical approximation

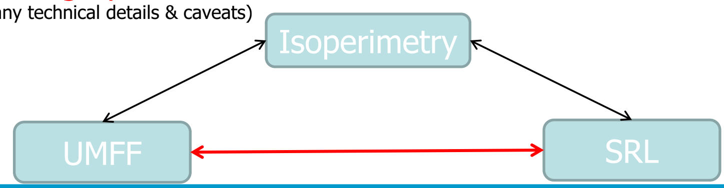
TUDelft

Szemerédi's regularity lemma (SRL)

SRL: Deep mathematical theorem: "Stone of Rosetta"
any large graph has an ϵ-regular partitioning

ϵ-regular partitioning relates to isoperimetry
⇒ good overall UMFF approximation

Translated: SIS on large graphs is ϵ-similar to "scaled" SIS on smaller graphs
(+ many technical details & caveats)



25

TU Delft

UMFF contains NIMFA

- N-Intertwined Mean-Field Approximation (NIMFA)
  - Only moment-closure approximation
- NIMFA equations:

$$\frac{dE[\widetilde{W}_i]}{dt} = -\delta E[\widetilde{W}_i] + \beta \sum_{m=1}^{N} a_{im} E[\widetilde{W}_m](1 - E[\widetilde{W}_i])$$

i = node index
$E[\widetilde{W}_i]$ = infection probability of node i
a_{ij} = adjacency element; a_{ij}=1 if node i and j are linked

$$\frac{d\boldsymbol{E}[Y_k]}{dt} = -\delta \boldsymbol{E}[Y_k] + \beta \sum_{m=1}^{K} \frac{L_{km}}{N_k N_m} (N_k - \boldsymbol{E}[Y_k]) \boldsymbol{E}[Y_m]$$

UMFF with K=N partitions is NIMFA

26

P. Van Mieghem, **2011, "The N -Intertwined SIS epidemic network model", Computing (Springer), Vol. 93, Issue 2, p. 147-169.**

TU Delft

UMFF contains HMF

Heterogeneous mean-field (HMF)

- Based completely on degree distribution

HMF equations:

$$\frac{d\rho_k}{dt} = -\delta\rho_k + \beta k(1-\rho_k)\Theta \qquad \Theta = \sum_{d_m=d_{min}}^{d_{max}} \rho_k \frac{d_m Pr[D=d_m]}{\sum_{d_l=d_{min}}^{d_{max}} d_l Pr[D=d_l]}$$

HMF equations (rewritten into UMFF form):

$$\frac{dE[\widetilde{W}_{d_k}]}{dt} = -\delta E[\widetilde{W}_{d_k}] + \beta \sum_{d_m=d_{min}}^{d_{max}} \tilde{a}_{d_k d_m} E[\widetilde{W}_{d_m}](N_{d_k} - E[\widetilde{W}_{d_k}])$$

N_{d_k} = number of nodes with degree d_k
$E[\widetilde{W}_{d_k}]$ = expected number of infected nodes of degree d_k
$\tilde{a}_{d_k d_m}$ = connection probability between nodes of degree d_m and d_k

UMFF with degree-partitions is equivalent to HMF

27

Pastor-Satorras, R. and Vespignani, A., 2001, " Epidemic dynamics and endemic states in complex networks ", Physical Review E, vol. 63, pp. 066117.

TU Delft

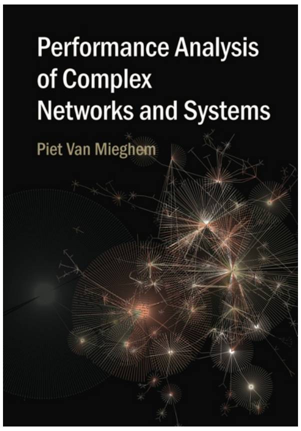
Conclusion

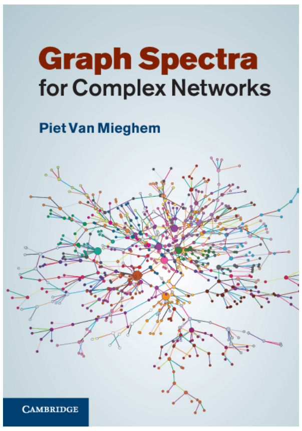
- The prevalence $y(t)$ in networks:
  - premier indicator of epidemic spread in networks
  - time-dependence hardly studied
  - mainly determined by the cut-set, i.e. the number of infective links
- **UMFF**: universal mean-field framework based on isoperimetric inequality and graph partitioning
  - with links to Szemerédi's regularity lemma
  - contains a.o. both HMF (Pastor-Satorras & Vespignani) and NIMFA
  - general bounds of mean-field approximations by isoperimetric inequality

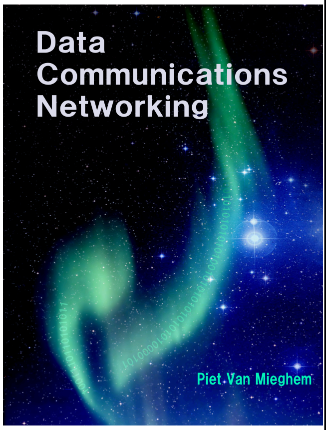
30

TU Delft

Books







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

31

TUDelft

32

TUDelft