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

Piet Van Mieghem

in collaboration with Karel Devriendt

Outline

Exact SIS prevalence

Universal mean-field framework (UMFF)

- Idea
- Framework
Continuous-time SIS model on networks

- Constant infection rate $\beta$ on all links
- Constant curing rate $\delta$ for all nodes

$$\tau = \frac{\beta}{\delta} : \text{effective spreading rate}$$

\[ X_j(t) = 1 \quad \text{node } j \text{ is infected at time } t \]
\[ X_j(t) = 0 \quad \text{node } j \text{ is healthy at time } t \]

Infection and curing are independent Poisson processes


Governing SIS equation for node $j$

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

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

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

SIS Prevalence

- Fraction of infected nodes in the graph $G$
  \[ S(t) = \frac{1}{N} \sum_{j=1}^{N} X_j(t) \] (random variable!)

- Prevalence: Expected fraction of infected nodes in $G$
  \[ y(t) = E[S(t)] = \frac{1}{N} \sum_{j=1}^{N} \text{Pr}[X_j(t) = 1] \]

Differential equation prevalence

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

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

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

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

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

\[ \frac{dy(t^*)}{dt^*} = -y(t^*) + \frac{\tau}{N} E[(u - w)^T Aw] \]
"Local rule - global emergent properties" class

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

Local SIS rule

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

Global emergent SIS spread

\[
The \text{Laplacian } Q = \Delta - A
\]

The normalized time \( t^* = \frac{\tau}{t} \)

Bernoulli state vector \( w(t^*) = (X_1(t^*), X_2(t^*), ..., X_N(t^*)) \)


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 \( t^* \)

\( NS(t^*) = 7 \)

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

\textbf{Cut-Set}: set of links with 1 infected node at time \( t^* \)

Set of susceptible nodes at time \( t^* \)

Outline

Exact SIS prevalence

Universal mean-field framework (UMFF)

- Idea
- Framework


Continuous-time Markovian SIS epidemics on networks

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

Approximate

N = 4 nodes
Markov theory

Regular bipartite Markov graph

Recursive structure of infinitesimal general $Q_N$


UMFF (1): New variable = #infected nodes

2^N States

N+1 states

Problem:
Infection rate = $\beta \times \#$ infective links
UMFF (2): approximation

$\# \text{ Infective links} \approx f(\# \text{Infected nodes})$

\[ f(2) \]

\[ 3\delta \]

... \[ 2 \]

\[ 3 \]

\[ ... \]

idea: isoperimetric problem in geometry

Outline

Exact SIS prevalence

Universal mean-field framework (UMFF)

- Idea
- Framework
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


Partitioning

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

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

\[ \bar{W}_1(t) = 2, \quad \bar{W}_2(t) = 2, \quad \bar{W}_3(t) = 0 \]

\[ \bar{W}(t) = \begin{bmatrix} \bar{W}_1(t) \\ \bar{W}_2(t) \\ \bar{W}_3(t) \end{bmatrix} \]

\[ 0 \leq \bar{W}_1(t) \leq 4 \\
0 \leq \bar{W}_2(t) \leq 4 \\
0 \leq \bar{W}_3(t) \leq 1 \]
The isoperimetric inequality (1)

- iso-perimetric (ἱσοπερίμετρος) → "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 (20th 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
The isoperimetric inequality (2)

\[ w^T Q w - \frac{d_{av}}{N-1} Y(N - Y) \leq \max_{1 \leq i \leq N-1} \frac{\sum_{i=1}^{N} d_{av} - \mu_i}{N} Y(N - Y) \]

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

\[ \frac{dE[y(t)]}{dt} = -\delta E[y(t)] + \frac{\beta}{N} E[w^T Q w] \]
\[ \frac{dy}{dt} = -\delta Y + \beta Y(N - Y) \]

logistic diff. eq.

Isoperimetric ineq. relates to Szemeredi’s regularity theorem

Eigenvalues of Laplacian \(Q\): \(\mu_N = 0 \leq \mu_{N-1} \leq \ldots \leq \mu_1\); Average degree: \(d_{av}\);

Number of infected nodes: \(Y\)

Chung, F. Discrete Isoperimetric Inequalities. DMTCS, 1996
**Szemerédi’s regularity lemma (SRL)**

SRL: Deep mathematical theorem: “Stone of Rosetta”
any large graph has an $\epsilon$-regular partitioning

$\epsilon$-regular partitioning relates to isoperimetry
→ good overall UMFF approximation

*Translated:* SIS on large graphs is $\epsilon$-similar to “scaled” SIS on smaller graphs
(+ many technical details & caveats)

---

**UMFF contains NIMFA**

- N-Intertwined Mean-Field Approximation (NIMFA)
  - Only moment-closure approximation
- NIMFA equations:
  \[
  \frac{dE[\tilde{W}_i]}{dt} = -\delta E[\tilde{W}_i] + \beta \sum_{m=1}^{N} a_{im} E[\tilde{W}_m](1 - E[\tilde{W}_i])
  \]
  
i = node index
  $E[\tilde{W}_i]$ = infection probability of node $i$
  $a_{ij}$ = adjacency element; $a_{ij}=1$ if node $i$ and $j$ are linked

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

UMFF with $K=N$ partitions is NIMFA
UMFF contains HMF
Heterogeneous mean-field (HMF)
  - Based completely on degree distribution

HMF equations:
\[
\frac{d\rho_k}{dt} = -\delta \rho_k + \beta k (1 - \rho_k) \Theta
\]
\[
\Theta = \sum_{d_m=d_{\min}}^{d_{\max}} \rho_k \frac{d_m Pr[D = d_m]}{\sum_{d_1=d_{\min}}^{d_{\max}} d_1 Pr[D = d_1]}
\]

HMF equations (rewritten into UMFF form):
\[
\frac{dE[\tilde{W}_{d_k}]}{dt} = -\delta E[\tilde{W}_{d_k}] + \beta \sum_{d_m=d_{\min}}^{d_{\max}} \tilde{a}_{d_k,d_m} E[\tilde{W}_{d_m}](N_{d_k} - E[\tilde{W}_{d_k}])
\]

\(N_{d_k}\) = number of nodes with degree \(d_k\)
\(E[\tilde{W}_{d_k}]\) = 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


Conclusion

- The prevalence \(\gamma(t)\) 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