Influence maximization and model reconstruction in spreading processes with dynamic message passing

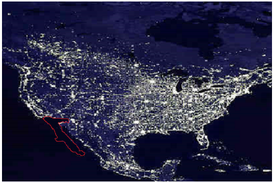
Andrey Lokhov

Theoretical Division and Center for Nonlinear Studies,
Los Alamos National Laboratory

University of Southern California

Motivation: control of spreading cascades

2.7 **million customers** left without power after cascading outages in Arizona and California in 2011



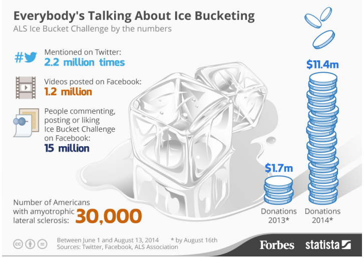
U.S. economy losses from cascading bankruptcies during the 2008 crisis estimated at the level of **$22 trillion**

As many as **579,000 people** could have been killed by H1N1 pandemic through transport-mediated spreading

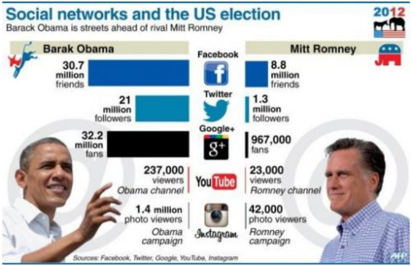
Cyber crime costs projected to reach **$2 trillion** globally by 2019

Motivation: control of spreading cascades

$115 million in donations generated by the ALS ice bucket challenge campaign in social networks



Win of the social media battle in 2012 presidential campaign in United States



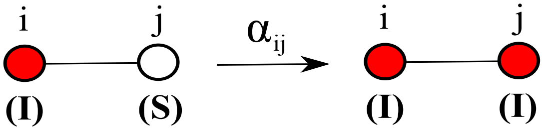
Additional examples: viral marketing, targeted chemically-induced control of dynamic biological processes, drug discovery, *etc.*

Dynamic Message-Passing Equations for Models with Unidirectional Dynamics

A.Y. Lokhov, M. Mézard, L. Zdeborová, PRE (2015)

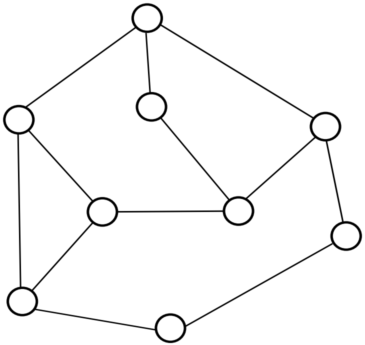
Modeling cascades: SI model as a typical example

In a graph $G = (V, E)$, node $i \in V$ at discrete time t is in state σ_i^t: "susceptible" $\sigma_i^t = S$ or "infected" $\sigma_i^t = I$. At each step $t \to t+1$:



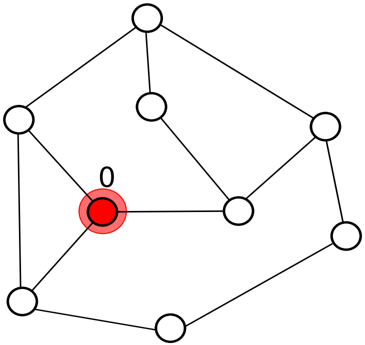
Example

A cascade: collection of activation times until horizon T



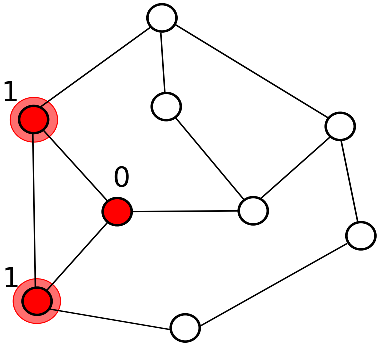
Example

A cascade: collection of activation times until horizon T



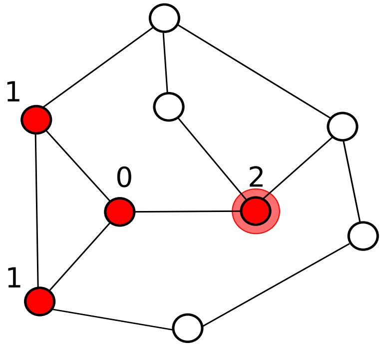
Example

A cascade: collection of activation times until horizon T



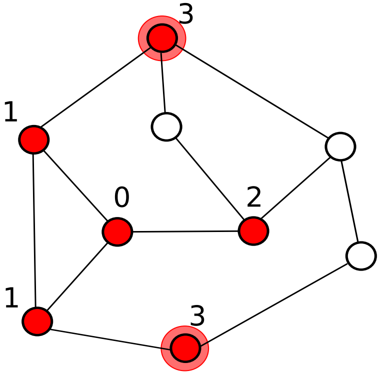
Example

A cascade: collection of activation times until horizon T



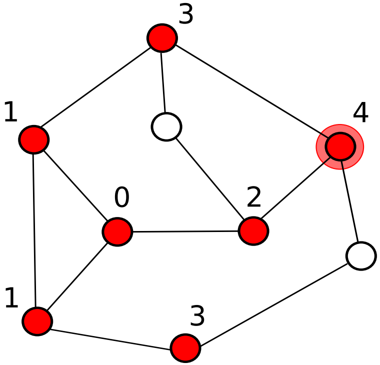
Example

A cascade: collection of activation times until horizon T



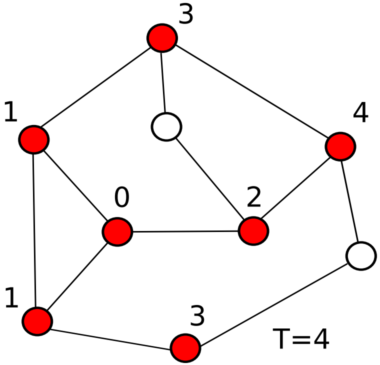
Example

A cascade: collection of activation times until horizon T



Example

A cascade: collection of activation times until horizon T



Plethora of other spreading dynamic models

- ✓ Independent Cascade (Susceptible-Infected-Recovered)
- ✓ Linear Threshold Model (Bootstrap Percolation, zero-temperature Random Field Ising Model)
- ✓ Rumor Spreading (Ignorant-Spreader-Stiffler), *etc.*
- ✓ Recurrent (non-progressive) dynamics: Voter Model, Glauber dynamics of spins, Susceptible-Infected-Susceptible, *etc.*

General question: Is it possible to **solve** the dynamics, i.e. estimate the marginal probabilities $\{P^i_\sigma(t)\}$ given G and $\{\alpha_{ij}\}$? [NP-hard]

Remark: Different from the data-mining approach; can be viewed as ML tool

A solution: Dynamic Message Passing for progressive dynamics

A.Y. Lokhov, M. Mézard, L. Zdeborová, PRE (2015)

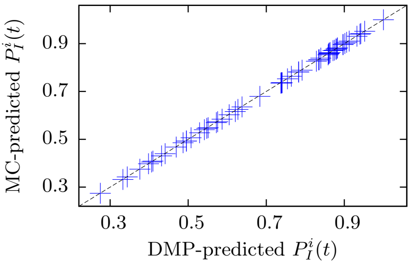
Example of the SI model

Estimation of $P_\sigma^i(t)$, $\sigma \in \{S, I\}$ on a given network instance with algorithmic complexity $O(|E|T)$.

$$P_\sigma^i(t+1) = F_i\left[\{P_\sigma^{j\to i}(t)\}_{j\in\partial i}\right].$$

Theorem: The quantities $P_S^i(t)$ are exact on tree graphs, and give lower bounds on values of marginal probabilities for general loopy graphs. (Accurate in practice).

Performance on a network of flights between major U.S. airports:



Dynamic message-passing equations

Marginal probabilities:

$$P_S^i(t) = P_S^i(0) \prod_{k \in \partial i} \theta^{k \to i}(t),$$
$$P_I^i(t) = 1 - P_S^i(t).$$

Dynamic "messages":

$$P_S^{k \to i}(t) = P_S^k(0) \prod_{l \in \partial k \setminus i} \theta^{l \to k}(t),$$
$$P_I^{k \to i}(t) = 1 - P_S^{k \to i}(t),$$
$$\theta^{k \to i}(t) = \theta^{k \to i}(t-1) - \alpha_{ki} \phi^{k \to i}(t-1),$$
$$\phi^{k \to i}(t) = (1 - \alpha_{ki}) \phi^{k \to i}(t-1) + P_I^{k \to i}(t) - P_I^{k \to i}(t-1).$$

Initial conditions:

$$\theta^{i \to j}(0) = 1, \qquad \phi^{i \to j}(0) = \delta_{\sigma_i^0, I} = P_I^i(0).$$

Dynamic message-passing equations

Marginal probabilities:

$$P_S^i(t) = P_S^i(0) \prod_{k \in \partial i} \theta^{k \to i}(t),$$

$$P_I^i(t) = 1 - P_S^i(t).$$

Dynamic "messages":

$$P_S^{k \to i}(t) = P_S^k(0) \prod_{l \in \partial k \setminus i} \theta^{l \to k}(t),$$

$$P_I^{k \to i}(t) = 1 - P_S^{k \to i}(t),$$

$$\theta^{k \to i}(t) = \theta^{k \to i}(t-1) - \alpha_{ki} \phi^{k \to i}(t-1),$$

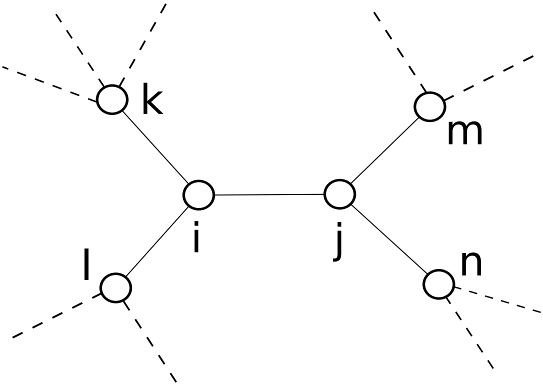
$$\phi^{k \to i}(t) = (1 - \alpha_{ki}) \phi^{k \to i}(t-1) + P_I^{k \to i}(t) - P_I^{k \to i}(t-1).$$

Initial conditions:

$$\theta^{i \to j}(0) = 1, \qquad \phi^{i \to j}(0) = \delta_{\sigma_i^0, I} = P_I^i(0).$$

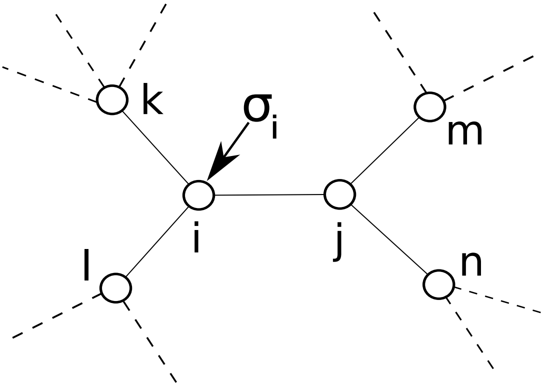
General framework: belief propagation

Static problem defined on a tree network $G = (V, E)$, where V is the set of nodes, and E is the set of connections between them:



General framework: belief propagation

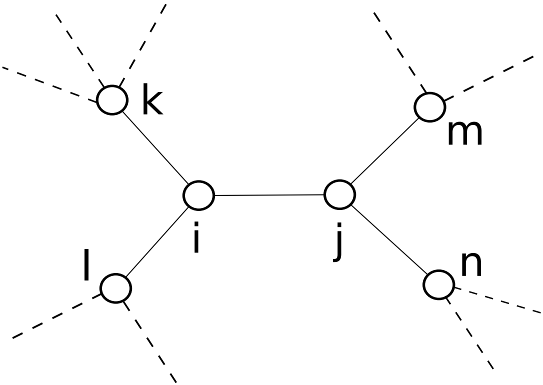
Variables σ_i (boolean variables, spins, colors, *etc.*) are associated with the nodes $i \in V$:



General framework: belief propagation

Joint probability distribution of the graphical model:

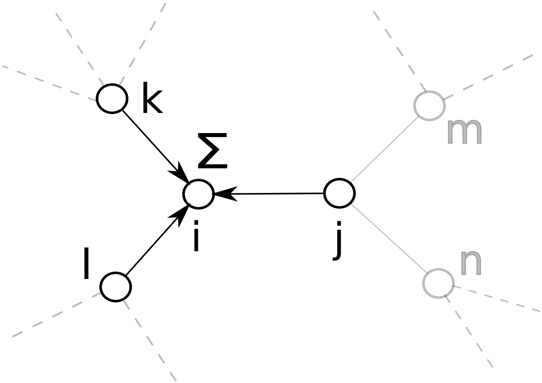
$$P(\{\sigma_i\}_{i \in V}) = \frac{1}{Z} \prod_{i \in V} \Psi_i(\sigma_i, \{\sigma_j\}_{j \in \partial i}).$$



General framework: belief propagation

Belief propagation algorithm in the static problem: efficient method for computing marginal probability distributions $m^i(\sigma_i)$:

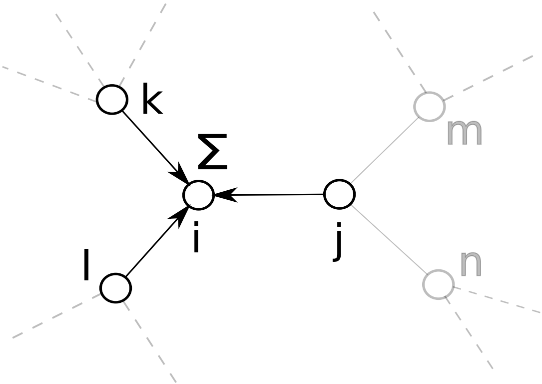
$$m^i(\sigma_i) = \sum_{\{\sigma_j\}_{j\in V, j\neq i}} P(\{\sigma_j\}_{j\in V}) \propto \prod_{j\in\partial i} m^{j\to i}(\sigma_j).$$



General framework: belief propagation

Belief propagation equation for marginal probability $m^i(\sigma_i)$:

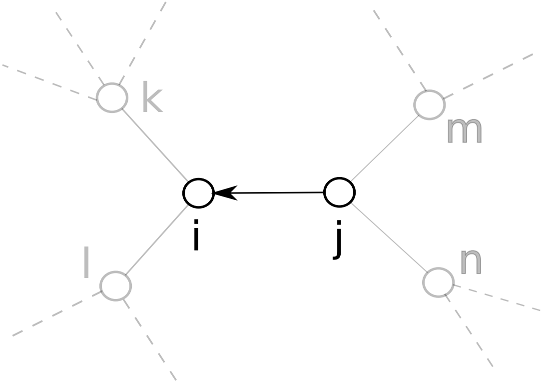
$$m^i(\sigma_i) = \frac{1}{Z^i} \sum_{\{\sigma_j\}_{j\in\partial i}} \Psi_i(\sigma_i, \{\sigma_j\}_{j\in\partial i}) \prod_{j\in\partial i} m^{j\to i}(\sigma_j).$$



General framework: belief propagation

How to compute the messages $m^{j\to i}(\sigma_j)$ in the right hand side?

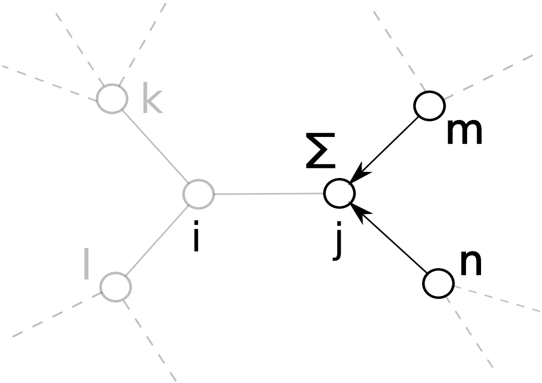
$$m^i(\sigma_i) = \frac{1}{Z^i} \sum_{\{\sigma_j\}_{j\in\partial i}} \Psi_i(\sigma_i, \{\sigma_j\}_{j\in\partial i}) \prod_{j\in\partial i} m^{j\to i}(\sigma_j).$$



General framework: belief propagation

Belief propagation equation for message $m^{j\to i}(\sigma_j)$:

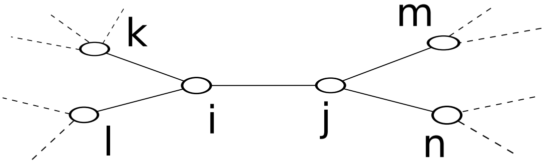
$$m^{j\to i}(\sigma_j) = \frac{1}{Z^{j\to i}} \sum_{\{\sigma_k\}_{k\in\partial j\setminus i}} \Psi_j(\sigma_j, \{\sigma_k\}_{k\in\partial j\setminus i}) \prod_{k\in\partial j\setminus i} m^{k\to j}(\sigma_k).$$



General framework: dynamic belief propagation

Belief propagation: asymptotically exact values of marginals on locally tree-like networks.

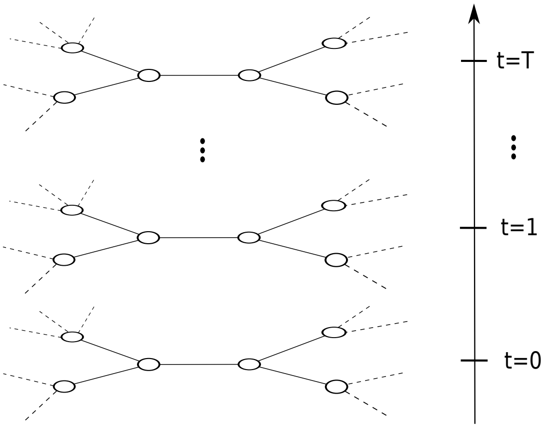
How to generalize for a dynamic process with transition probability between different discrete times $w_i(\sigma_i^{t+1} \mid \sigma_i^t, \{\sigma_j^t\}_{j \in \partial i})$?



Question: Is there a way to predict marginal probabilities $P^i_{\sigma_i}(t)$?

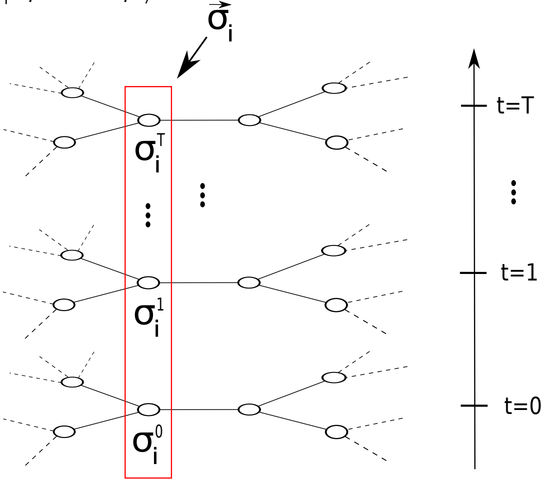
General framework: dynamic belief propagation

State of the network at each time step:



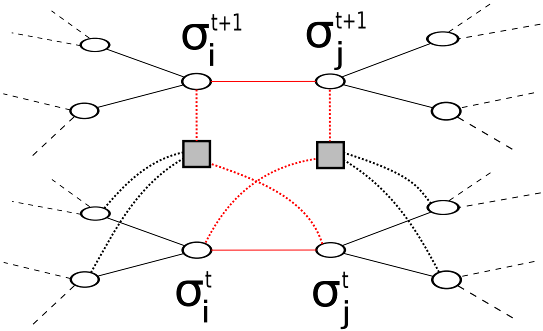
General framework: dynamic belief propagation

Idea: write belief propagation equations for node trajectories in time: $\vec{\sigma}_i = \left|\sigma_i^0, \ldots, \sigma_i^T\right\rangle$



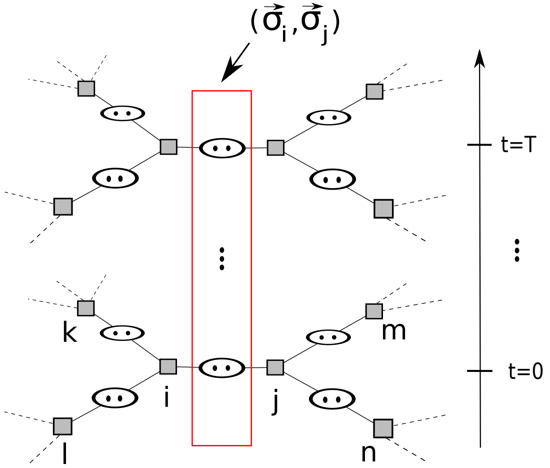
General framework: dynamic belief propagation

Technical moment: interaction of neighboring trajectories at nearest times: systematic short loops

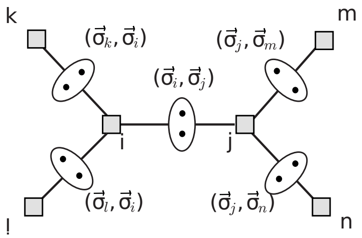


General framework: Dynamic belief propagation

Solution: working with pairs of trajectories (duality graph)



General framework: dynamic belief propagation



Dynamic belief propagation equation on conditional messages:

$$m^{i\to j}(\vec{\sigma}_i \mid \vec{\sigma}_j) = \sum_{\{\vec{\sigma}_k\}_{k\in\partial i\setminus j}} \left[\prod_{t=1}^{T-1} w_i(\sigma_i^{t+1} \mid \sigma_i^t, \{\sigma_k^t\}_{k\in\partial i}) \right]$$

$$\times P(\{\sigma_i^0\}_{i\in V}) \prod_{k\in\partial i\setminus j} m^{k\to i}(\vec{\sigma}_k \mid \vec{\sigma}_i).$$

Progressive dynamics

The trajectories are in one-to-one correspondence with the flipping times

$$\vec{\sigma}_i = \Big| S_0 SSSSSSI_{\tau_i} IIIIIII_T \Big\rangle \longleftrightarrow \tau_i$$

By definition:

$$P_S^i(t) = \sum_{\tau_i > t} m^i(\tau_i),$$
$$P_I^i(t) = \sum_{\tau_i \leq t} m^i(\tau_i).$$

The dynamic messages are weighted sums of marginals.

Dynamic Message-Passing Equations

Marginal probabilities:

$$P_S^i(t) = P_S^i(0) \prod_{k \in \partial i} \theta^{k \to i}(t),$$
$$P_I^i(t) = 1 - P_S^i(t).$$

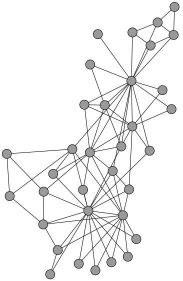
Dynamic "messages":

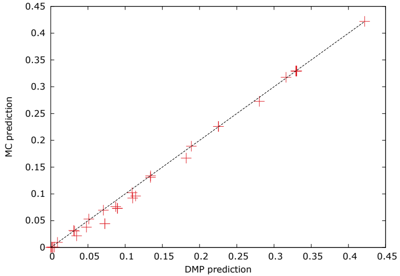
$$P_S^{k \to i}(t) = P_S^k(0) \prod_{l \in \partial k \setminus i} \theta^{l \to k}(t),$$
$$P_I^{k \to i}(t) = 1 - P_S^{k \to i}(t),$$
$$\theta^{k \to i}(t) = \theta^{k \to i}(t-1) - \alpha_{ki} \phi^{k \to i}(t-1),$$
$$\phi^{k \to i}(t) = (1 - \alpha_{ki}) \phi^{k \to i}(t-1) + P_I^{k \to i}(t) - P_I^{k \to i}(t-1).$$

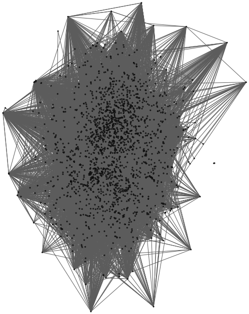
Initial conditions:

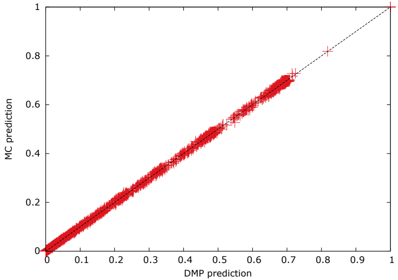
$$\theta^{i \to j}(0) = 1, \qquad \phi^{i \to j}(0) = \delta_{\sigma_i^0, I} = P_I^i(0).$$

In practice, works well on loopy graphs

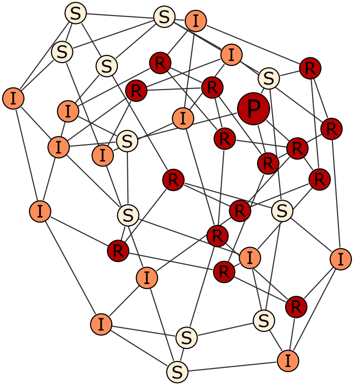








Application: inference of epidemic origin



Problem of estimating the origin of an epidemic outbreak, given a contact network and a snapshot of epidemic spread $\mathcal{O} \subset V$ at a certain time t_0 (in general, unknown). A.Y. Lokhov *et al.* PRE (2014)

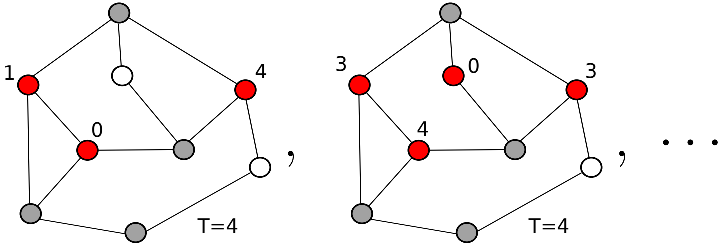
Reconstructing parameters of spreading models from partial observations

A.Y. Lokhov, NIPS (2016)

Formulation of the problem

Each cascade is divided into observed ($\mathcal{O}$) and hidden ($\mathcal{H}$) parts, $\Sigma^c = \Sigma^c_{\mathcal{O}} \cup \Sigma^c_{\mathcal{H}}$

Given M partially observed cascades $\Sigma_{\mathcal{O}} = \cup_{c=1}^{M} \Sigma^c_{\mathcal{O}}$, reconstruct model parameters $\{\alpha^*_{ij}\}_{(ij)\in E} \equiv G_{\alpha^*}$ used to generate data

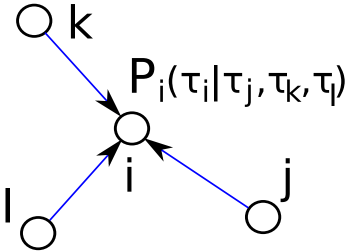


Special case: full observations

Full observations $\Sigma_{\mathcal{O}} = \Sigma$: easy
Maximize the likelihood function

$$P(\Sigma \mid G_\alpha) = \prod_{i \in V} \prod_{1 \le c \le M} P_i(\tau_i^c \mid \Sigma^c, G_\alpha)$$

$\widehat{G}_{\alpha^*} = \arg\min \left(-\ln P(\Sigma \mid G_\alpha)\right)$: local convex optimization $\forall i \in V$



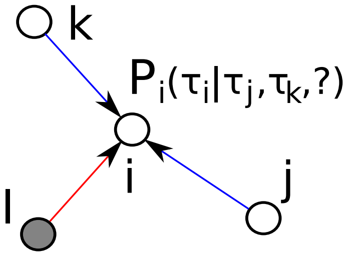
Maximization of the likelihood: intractable

Partial observations $\Sigma_{\mathcal{O}} \neq \Sigma$: **hard**

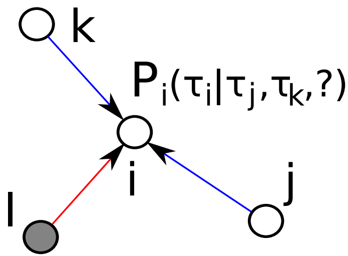
Maximization of the likelihood marginalized over unknown information:

$$P(\Sigma_{\mathcal{O}} \mid G_{\alpha}) = \sum_{\{\tau_h^c\}, h \in \mathcal{H}} P(\Sigma \mid G_{\alpha}),$$

computational complexity $\propto T^{H}$



Proposed speed-up: Heuristic Two-Stage Algorithm (HTS)



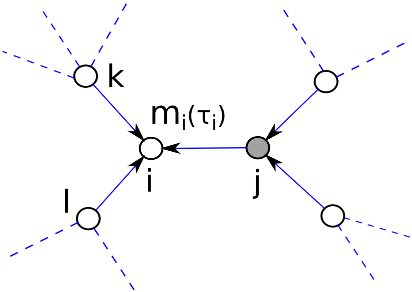
1. Complete missing $\{\tau_h^c\}_{h\in\mathcal{H}}$ by most probable values $\widehat{\Sigma}_{\mathcal{H}} = \arg\max P(\Sigma \mid \widehat{G}_\alpha)$ using MC sampling.
2. Solve the "full observations" problem using $\Sigma = \Sigma_{\mathcal{O}} \cup \widehat{\Sigma}_{\mathcal{H}}$.
3. Iterate steps 1 and 2 until global convergence of the algorithm.

(Still very slow).

DMP algorithm

Key idea: approximation of the likelihood with marginal probabilities

$$P(\Sigma_{\mathcal{O}} \mid G_\alpha) \approx \prod_{c=1}^{M} \prod_{i \in \mathcal{O}} \left[m^i(\tau_i^c \mid G_\alpha) \mathbb{1}_{\tau_i^c \leq T} + P_S^i(\tau_i^c \mid G_\alpha) \mathbb{1}_{\tau_i^c = T} \right]$$



Each $m^i(\tau_i^c)$ summarizes the effect of all possible propagation paths.

DMP algorithm

Minimization of the "free energy"

$$f_{\text{DMP}} = -\ln P(\Sigma_{\mathcal{O}} \mid G_{\alpha}) = \sum_{i \in \mathcal{O}} f^{i}_{\text{DMP}}$$

will yield the most likely consensus among the ensemble of parameters. Gradient $\partial f^{i}_{\text{DMP}} / \partial \alpha_{rs}$ through the DMP eqs. for derivatives $p^{k \to i}_{rs}(t) \equiv \partial \theta^{k \to i}(t) / \partial \alpha_{rs}$, e.g.

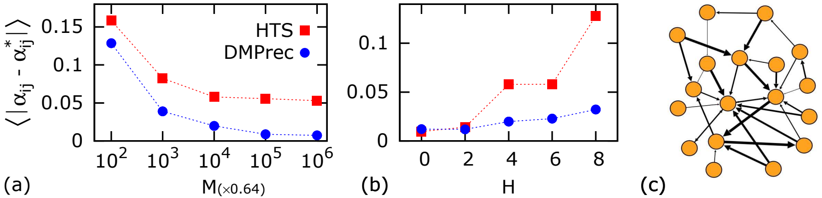
$$\frac{\partial P^{i}_{S}(t)}{\partial \alpha_{rs}} = P^{i}_{S}(0) \sum_{k \in \partial i} p^{k \to i}_{rs}(t) \prod_{l \in \partial i \setminus k} \theta^{l \to i}(t)$$

Claim: On tree graphs (regime in which DMP eqs. are derived), $\lim_{M \to \infty} \frac{\partial f_{\text{DMP}}}{\partial \alpha_{rs}} \big|_{G_{\alpha^*}} = 0$.

Remark: Also acts as a loss function for the influence learning

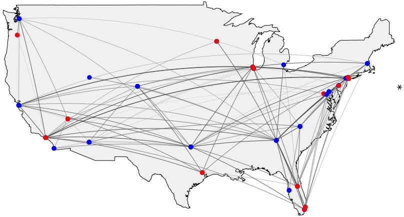
Comparison between HTS and DMPrec

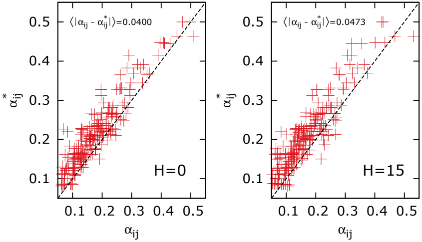
Days → minutes, for even more accurate results



Application to real data

Air-traffic mediated epidemic spreading





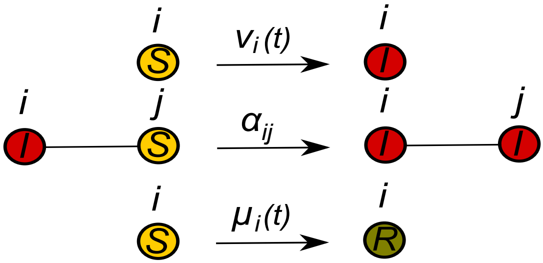
Results for a sub-network of flights ($|E| = 210$) between $N = 30$ major U.S. hubs. $M = 10,000$.

Optimal deployment of resources for maximizing impact in spreading processes

A.Y. Lokhov and D. Saad, arXiv:1608.08278 (2016)

Generalized SIR model

In a graph $G = (V, E)$, node $i \in V$ at discrete time t is in state σ_i^t: "susceptible" $\sigma_i^t = S$, "infected" $\sigma_i^t = I$ or "recovered" $\sigma_i^t = R$.



ν-**mechanism**: spontaneous activation (e.g. advertisement)
α-**mechanism**: spreading due to interactions with neighbors
μ-**mechanism**: effect of vaccination (e.g. immunity in epidemics)

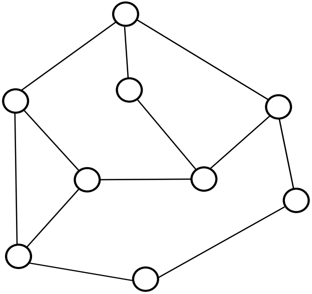
Example: propagation of information

Susceptible-Infected model in discrete time with external control:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$

$$S(i) \xrightarrow{\nu_i(t)} I(i)$$

at each time step for $i \in V$ and $(i, j) \in E$.

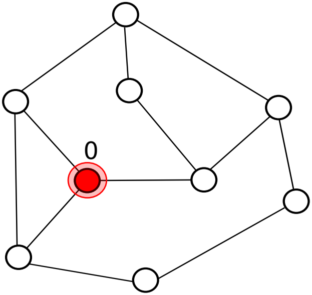


Example: propagation of information

Susceptible-Infected model in discrete time with external control:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$
$$S(i) \xrightarrow{\nu_i(t)} I(i)$$

at each time step for $i \in V$ and $(i, j) \in E$.



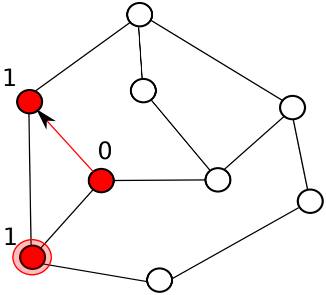
Example: propagation of information

Susceptible-Infected model in discrete time with external control:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$

$$S(i) \xrightarrow{\nu_i(t)} I(i)$$

at each time step for $i \in V$ and $(i, j) \in E$.

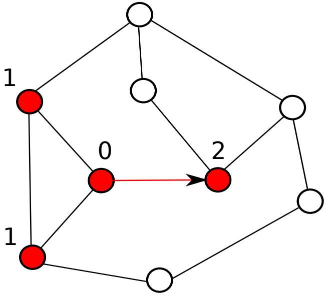


Example: propagation of information

Susceptible-Infected model in discrete time with external control:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$
$$S(i) \xrightarrow{\nu_i(t)} I(i)$$

at each time step for $i \in V$ and $(i, j) \in E$.



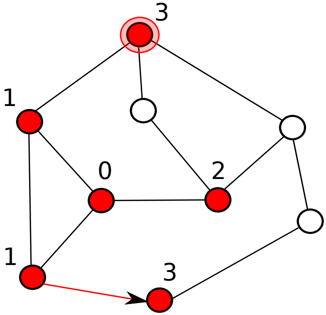
Example: propagation of information

Susceptible-Infected model in discrete time with external control:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$

$$S(i) \xrightarrow{\nu_i(t)} I(i)$$

at each time step for $i \in V$ and $(i, j) \in E$.

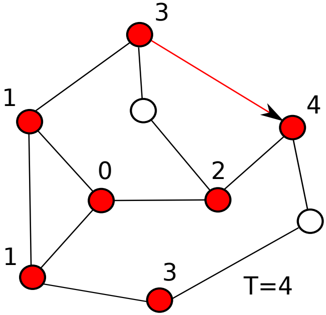


Example: propagation of information

Susceptible-Infected model in discrete time with external control:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$
$$S(i) \xrightarrow{\nu_i(t)} I(i)$$

at each time step for $i \in V$ and $(i, j) \in E$.

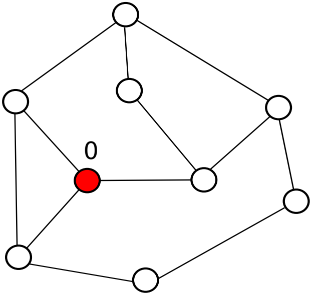


Example: epidemic spreading with vaccination

Susceptible-Infected-Recovered model with protection mechanism:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$
$$S(i) \xrightarrow{\mu_i(t)} R(i).$$

at each time step for $i \in V$ and $(i, j) \in E$.



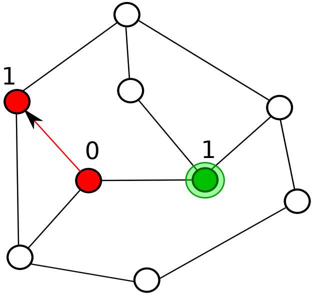
Example: epidemic spreading with vaccination

Susceptible-Infected-Recovered model with protection mechanism:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$

$$S(i) \xrightarrow{\mu_i(t)} R(i).$$

at each time step for $i \in V$ and $(i, j) \in E$.

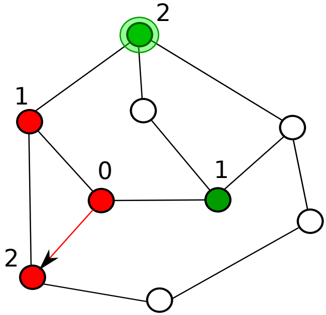


Example: epidemic spreading with vaccination

Susceptible-Infected-Recovered model with protection mechanism:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$
$$S(i) \xrightarrow{\mu_i(t)} R(i).$$

at each time step for $i \in V$ and $(i, j) \in E$.



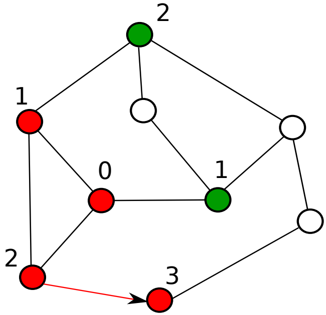
Example: epidemic spreading with vaccination

Susceptible-Infected-Recovered model with protection mechanism:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$

$$S(i) \xrightarrow{\mu_i(t)} R(i).$$

at each time step for $i \in V$ and $(i, j) \in E$.



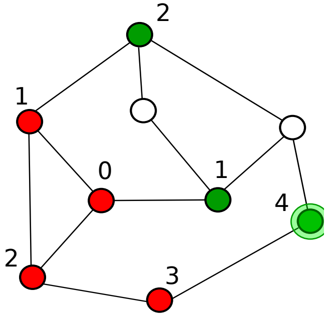
Example: epidemic spreading with vaccination

Susceptible-Infected-Recovered model with protection mechanism:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$

$$S(i) \xrightarrow{\mu_i(t)} R(i).$$

at each time step for $i \in V$ and $(i, j) \in E$.

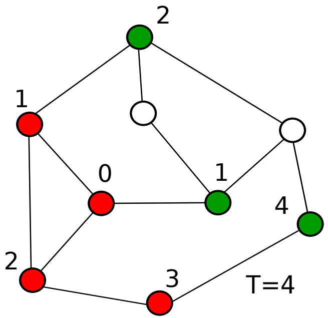


Example: epidemic spreading with vaccination

Susceptible-Infected-Recovered model with protection mechanism:

$$S(i) + I(j) \xrightarrow{\alpha_{ji}} I(i) + I(j),$$
$$S(i) \xrightarrow{\mu_i(t)} R(i).$$

at each time step for $i \in V$ and $(i, j) \in E$.



Typical formulation

Given restricted control budget, optimal control using $\nu_i(t)$ and $\mu_i(t)$ for maximizing or minimizing total spread at time T:

$$\mathcal{S}(T) = \mathbb{E}\left[\sum_{i \in V} \mathbb{1}[\sigma_i^T = I]\right] = \sum_{i \in V} P_I^i(T)$$

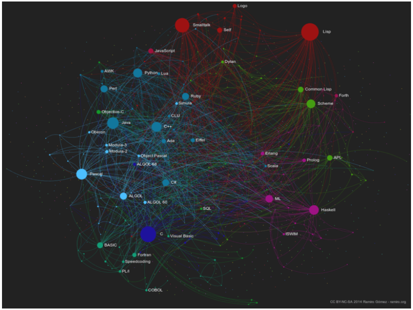
Seeding problem: find the smallest set of initial nodes which maximize the spread [NP-hard problem]

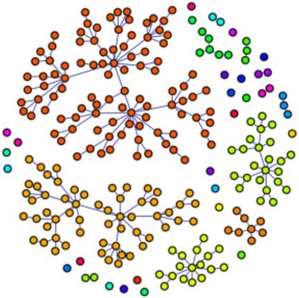
Domingos & Richardson, ACM SIGKDD (2001)

Existing approaches: neglecting dynamics, topology-based

Centrality measures: high-degree nodes, neighbors of random nodes, betweenness centrality, random walk centralities, k-shell decomposition, . . .

Network dismantling: optimal percolation, breaking the giant component, belief propagation algorithms, . . . [NP-complete]





Cohen *et al.*, Phys. Rev. Lett. (2003)
Chen *et al.*, Phys. Rev. Lett. (2008)
Kitsak *et al.*, Nature Physics (2010)

Altarelli *et al.*, Phys. Rev. X (2014)
Morone & Makse, Nature (2015)
Braunstein *et al.*, submitted to PNAS (2016)

Existing approaches: incorporating dynamics

Rigorous analysis of the seeding problem for the Independent Cascade model at $T = \infty$, greedy algorithm using submodularity

Kempe, Kleinberg & Tardos, ACM SIGKDD (2003)

Finite T greedy algorithm with sampling for the IC model

Du *et al.*, NIPS (2013)

Non-distributed mean-field solutions neglecting the topology

Survey: Nowzari *et al.*, IEEE Control Systems (2016)

Existing approaches: incorporating dynamics

Rigorous analysis of the seeding problem for the Independent Cascade model at $T = \infty$, greedy algorithm using submodularity

Kempe, Kleinberg & Tardos, ACM SIGKDD (2003)

Finite T greedy algorithm with sampling for the IC model

Du *et al.*, NIPS (2013)

Non-distributed mean-field solutions neglecting the topology

Survey: Nowzari *et al.*, IEEE Control Systems (2016)

Goal: solving the resource allocation problem on a given graph with finite T. Requires an explicit resolution of the dynamics.

Remark: DMP can be used directly in the $(1 - 1/e)$-approximation schemes

A.Y. Lokhov and D. Saad, to appear

Dynamic message-passing equations

Marginal probabilities:

$$P_S^i(t) = P_S^i(0)\left(\prod_{t'=0}^{t-1}(1-\nu_i(t'))\right)\left(\prod_{t'=0}^{t-1}(1-\mu_i(t'))\right)\prod_{k\in\partial i}\theta^{k\to i}(t),$$

$$P_R^i(t) = P_R^i(t-1) + \mu_i(t-1)P_S^i(t-1),$$

$$P_I^i(t) = 1 - P_S^i(t) - P_R^i(t).$$

Dynamic "messages":

$$P_S^{k\to i}(t) = P_S^k(0)\left(\prod_{t'=0}^{t-1}(1-\nu_k(t'))\right)\left(\prod_{t'=0}^{t-1}(1-\mu_k(t'))\right)\prod_{l\in\partial k\backslash i}\theta^{l\to k}(t),$$

$$P_R^{k\to i}(t) = P_R^{k\to i}(t-1) + \mu_k(t-1)P_S^{k\to i}(t-1),$$

$$P_I^{k\to i}(t) = 1 - P_S^{k\to i}(t) - P_R^{k\to i}(t),$$

$$\theta^{k\to i}(t) = \theta^{k\to i}(t-1) - \alpha_{ki}\phi^{k\to i}(t-1),$$

$$\phi^{k\to i}(t) = (1-\alpha_{ki})\phi^{k\to i}(t-1) + P_I^{k\to i}(t) - P_I^{k\to i}(t-1).$$

Initial conditions:

$$\theta^{i\to j}(0) = 1, \qquad \phi^{i\to j}(0) = \delta_{\sigma_i^0, I} = P_I^i(0).$$

Optimization framework

Constrained Lagrangian formulation:

$$\mathcal{L} = \underbrace{\mathcal{O}}_{\text{objective}} + \underbrace{\mathcal{B} + \mathcal{P} + \mathcal{I} + \mathcal{D}}_{\text{constraints}}$$

Optimization framework

Constrained Lagrangian formulation:

$$\mathcal{L} = \underbrace{\mathcal{O}}_{\text{objective}} + \underbrace{\mathcal{B} + \mathcal{P} + \mathcal{I} + \mathcal{D}}_{\text{constraints}}$$

Objective $\mathcal{O}$: targeting nodes at specific times

$$\mathcal{O} = \mathbb{E}\left[\sum_{i \in V} \mathbb{1}[\sigma_i^{t_i} = I]\right] = \sum_i P_I^i(t_i)$$

Optimization framework

Constrained Lagrangian formulation:

$$\mathcal{L} = \underbrace{\mathcal{O}}_{\text{objective}} + \underbrace{\mathcal{B} + \mathcal{P} + \mathcal{I} + \mathcal{D}}_{\text{constraints}}$$

Objective $\mathcal{O}$: targeting nodes at specific times

$$\mathcal{O} = \mathbb{E}\left[\sum_{i \in V} \mathbb{1}[\sigma_i^{t_i} = I]\right] = \sum_i P_I^i(t_i)$$

Budget constraints:

$$\mathcal{B} = \sum_{t=0}^{T-1} \lambda_B^\nu(t) \left[\sum_{i \in V} \nu_i(t) - B_\nu(t)\right]$$

Optimization framework

Constrained Lagrangian formulation:

$$\mathcal{L} = \underbrace{\mathcal{O}}_{\text{objective}} + \underbrace{\mathcal{B} + \mathcal{P} + \mathcal{I} + \mathcal{D}}_{\text{constraints}}$$

Objective $\mathcal{O}$: targeting nodes at specific times

$$\mathcal{O} = \mathbb{E}\left[\sum_{i \in V} \mathbb{1}[\sigma_i^{t_i} = I]\right] = \sum_i P_I^i(t_i)$$

Budget constraints:

$$\mathcal{B} = \sum_{t=0}^{T-1} \lambda_B^{\nu}(t)\left[\sum_{i \in V} \nu_i(t) - B_\nu(t)\right]$$

Control parameters variation:

$$\mathcal{P} = \epsilon \sum_{t=0}^{T-1} \sum_{i \in V} \left(\log\left[\nu_i(t) - \underline{\nu_i^t}\right] + \log\left[\overline{\nu_i^t} - \nu_i(t)\right]\right)$$

Optimization framework

Constrained Lagrangian formulation:

$$\mathcal{L} = \underbrace{\mathcal{O}}_{\text{objective}} + \underbrace{\mathcal{B} + \mathcal{P} + \mathcal{I} + \mathcal{D}}_{\text{constraints}}$$

Objective $\mathcal{O}$: targeting nodes at specific times

$$\mathcal{O} = \mathbb{E}\left[\sum_{i \in V} \mathbb{1}[\sigma_i^{t_i} = I]\right] = \sum_i P_I^i(t_i)$$

Budget constraints:

$$\mathcal{B} = \sum_{t=0}^{T-1} \lambda_B^{\nu}(t) \left[\sum_{i \in V} \nu_i(t) - B_\nu(t)\right]$$

Control parameters variation:

$$\mathcal{P} = \epsilon \sum_{t=0}^{T-1} \sum_{i \in V} \left(\log\left[\nu_i(t) - \underline{\nu_i^t}\right] + \log\left[\overline{\nu_i^t} - \nu_i(t)\right]\right)$$

Initial conditions $\mathcal{I}$ and dynamics $\mathcal{D}$ enforced by dual variables $\lambda_i^\sigma(t)$, $\lambda_{j\to i}^\sigma(t)$ with $t = 0, \ldots, T$.

Example: Lagrangian for SI model

$$
\mathcal{L} = \underbrace{\sum_{i\in V}\left(1 - P_S^i(t_i)\right)}_{\mathcal{O}} + \underbrace{\sum_{t=0}^{T-1}\lambda_B^{\nu}(t)\left[\sum_{i\in V}\nu_i(t) - B_\nu(t)\right]}_{\mathcal{B}} + \underbrace{\epsilon\sum_{t=0}^{T-1}\sum_{i\in V}\left(\log\nu_i(t) + \log\left[1-\nu_i(t)\right]\right)}_{\mathcal{P}}
$$

$$
\left.
\begin{aligned}
&+ \sum_{i\in V}\sum_{t=0}^{T-1}\lambda_i^S(t+1)\left[P_S^i(t+1) - P_S^i(t)(1-\nu_i(t))\prod_{k\in\partial i}\frac{\theta^{k\to i}(t+1)}{\theta^{k\to i}(t)}\right] \\
&+ \sum_{(ki)\in E}\sum_{t=0}^{T-1}\lambda_{k\to i}^S(t+1)\left[P_S^{k\to i}(t+1) - P_S^{k\to i}(t)(1-\nu_k(t))\prod_{l\in\partial k\setminus i}\frac{\theta^{l\to k}(t+1)}{\theta^{l\to k}(t)}\right] \\
&+ \sum_{(ki)\in E}\sum_{t=0}^{T-1}\lambda_{k\to i}^{\theta}(t+1)\left[\theta^{k\to i}(t+1) - \theta^{k\to i}(t) + \alpha_{ki}\phi^{k\to i}(t)\right] \\
&+ \sum_{(ki)\in E}\sum_{t=0}^{T-1}\lambda_{k\to i}^{\phi}(t+1)\left[\phi^{k\to i}(t+1) - (1-\alpha_{ki})\phi^{k\to i}(t) - P_S^{k\to i}(t) + P_S^{k\to i}(t+1)\right]
\end{aligned}
\right\}\mathcal{D}
$$

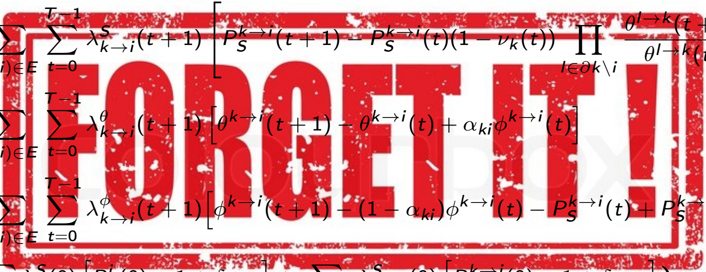
$$
\left.
\begin{aligned}
&+ \sum_{i\in V}\lambda_i^S(0)\left[P_S^i(0) - 1 + \delta_{\sigma_i^{\mathbf{0}},I}\right] + \sum_{(ki)\in E}\lambda_{k\to i}^S(0)\left[P_S^{k\to i}(0) - 1 + \delta_{\sigma_k^{\mathbf{0}},I}\right] \\
&+ \sum_{(ki)\in E}\lambda_{k\to i}^{\theta}(0)\left[\theta^{k\to i}(0) - 1\right] + \sum_{(ki)\in E}\lambda_{k\to i}^{\phi}(0)\left[\phi^{k\to i}(0) - \delta_{\sigma_k^{\mathbf{0}},I}\right].
\end{aligned}
\right\}\mathcal{I}
$$

Example: Lagrangian for SI model

$$
\mathcal{L} = \underbrace{\sum_{i \in V} \left(1 - P_S^i(t_i)\right)}_{\mathcal{O}} + \underbrace{\sum_{t=0}^{T-1} \lambda_B^{\nu}(t) \left[\sum_{i \in V} \nu_i(t) - B_\nu(t)\right]}_{\mathcal{B}} + \underbrace{\epsilon \sum_{t=0}^{T-1} \sum_{i \in V} \left(\log \nu_i(t) + \log\left[1 - \nu_i(t)\right]\right)}_{\mathcal{P}}
$$

$$
\left.
\begin{aligned}
&+ \sum_{i \in V} \sum_{t=0}^{T-1} \lambda_i^S(t+1) \left[P_S^i(t+1) - P_S^i(t)(1 - \nu_i(t)) \prod_{k \in \partial i} \frac{\theta^{k \to i}(t+1)}{\theta^{k \to i}(t)}\right] \\
&+ \sum_{(ki) \in E} \sum_{t=0}^{T-1} \lambda_{k \to i}^S(t+1) \left[P_S^{k \to i}(t+1) - P_S^{k \to i}(t)(1 - \nu_k(t)) \prod_{l \in \partial k \setminus i} \frac{\theta^{l \to k}(t+1)}{\theta^{l \to k}(t)}\right] \\
&+ \sum_{(ki) \in E} \sum_{t=0}^{T-1} \lambda_{k \to i}^{\theta}(t+1) \left[\theta^{k \to i}(t+1) - \theta^{k \to i}(t) + \alpha_{ki} \phi^{k \to i}(t)\right] \\
&+ \sum_{(ki) \in E} \sum_{t=0}^{T-1} \lambda_{k \to i}^{\phi}(t+1) \left[\phi^{k \to i}(t+1) - (1 - \alpha_{ki}) \phi^{k \to i}(t) - P_S^{k \to i}(t) + P_S^{k \to i}(t+1)\right]
\end{aligned}
\right\} \mathcal{D}
$$

$$
\left.
\begin{aligned}
&+ \sum_{i \in V} \lambda_i^S(0) \left[P_S^i(0) - 1 + \delta_{\sigma_i^0, I}\right] + \sum_{(ki) \in E} \lambda_{k \to i}^S(0) \left[P_S^{k \to i}(0) - 1 + \delta_{\sigma_k^0, I}\right] \\
&+ \sum_{(ki) \in E} \lambda_{k \to i}^{\theta}(0) \left[\theta^{k \to i}(0) - 1\right] + \sum_{(ki) \in E} \lambda_{k \to i}^{\phi}(0) \left[\phi^{k \to i}(0) - \delta_{\sigma_k^0, I}\right].
\end{aligned}
\right\} \mathcal{I}
$$



Optimization technique: forward-backward propagation

Extremization of the Lagrangian by variation with respect to primal $P_{\sigma}^{j\to i}(t)$, dual $\lambda_{j\to i}^{\sigma}(t)$ and control $\nu_i(t)$ variables:

$$\frac{\partial \mathcal{L}}{\partial \lambda_{j\to i}^{\sigma}(t)} = 0 \to \textit{Forward}\text{ DMP equations}$$

$$\frac{\partial \mathcal{L}}{\partial P_{\sigma}^{j\to i}(t)} = 0 \to \textit{Backward}\text{ message-passing equations on dual variables}$$

$$\frac{\partial \mathcal{L}}{\partial \nu_i(t)} = 0 \to \textit{Update}\text{ equations for control variables at time } t.$$

Optimization technique: forward-backward propagation

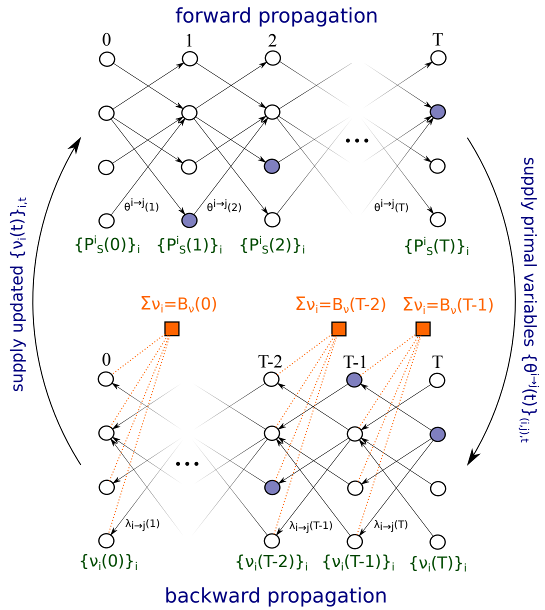
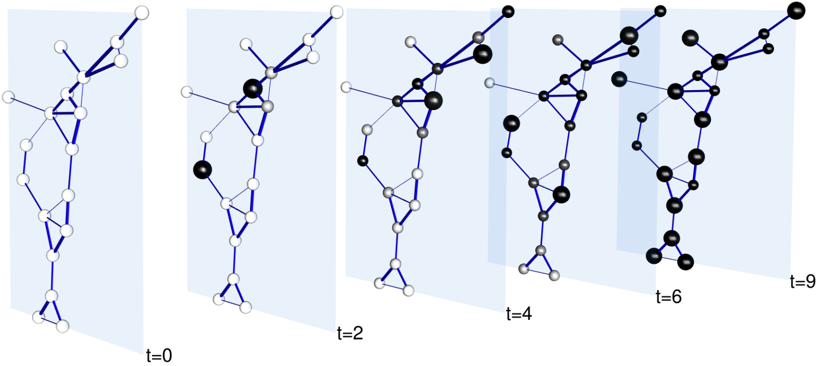


Illustration on a small network: targeting

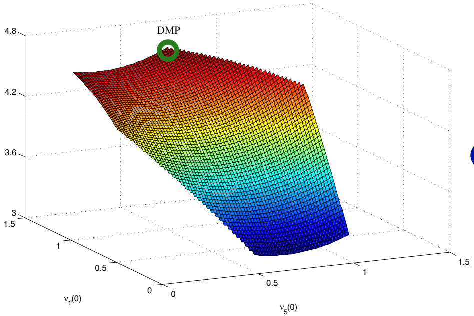
Dynamic resource allocation which aims at targeting specific nodes at required times (larger-sized nodes at $\{t_i\}$). Color intensity (gradually from white to black) indicates the values of $P_l^i(t)$.

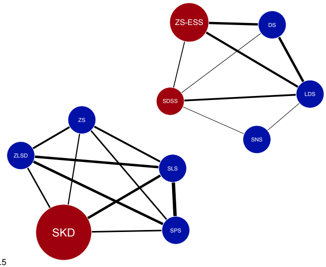


Convergence to the optimal solution in 7 iterations.

Illustration on a small network: seeding

Validation of the scheme in the seeding case on a small network with an explicit evaluation of the objective function $\mathcal{S}(T)$ via symbolic resolution of the DMP equations.



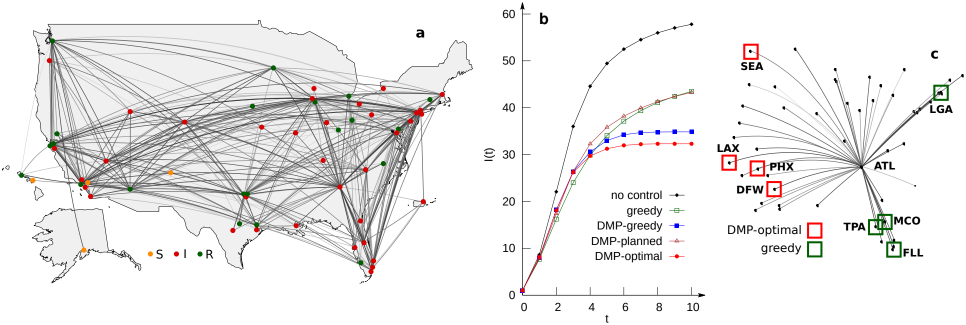


What happens on large networks?

Forward-backward propagation: iterate for a fixed number of swaps M outputting the best solution $\nu_i^*(t)$ with complexity $O(|E|TM)$.

Applications to very large networks with millions of nodes.

Illustration: online mitigation of an epidemic



- ✓ *DMP-planned* (offline dynamic resource allocation with T-horizon)
- ✓ *Greedy*: vaccination of nodes at "high risk" ranked by $P_i^t(S \to I) = 1 - \prod_{j \in \partial i}(1 - \alpha_{ji} \mathbb{1}[\sigma_j^t = I])$
- ✓ *DMP-greedy* (optimization at one-time step only)
- ✓ *DMP-optimal*: repeated re-evaluation of T-horizon problem based on the feedback from the current realization of dynamics

Path forward and extensions

- ✓ Continuous dynamics and temporal graphs via $\alpha_{ij}(t)$
- ✓ Optimization over edge-related control parameters $\beta_{ij}(t)$
- ✓ Application of the optimization to the percolation-like equations describing the $T \to \infty$ limit of the dynamics
- ✓ Solution of inverse problems from sparsely located sensor data
- ✓ Optimal placement of sensors for data collection
- ✓ Accelerated gradient-free learning. . .

Questions?