Mean Field Analysis of Gossip Protocols

Lucia Cloth, Boudewijn Haverkort

University of Twente
Design & Analysis of Communication Systems
http://dacs.ewi.utwente.nl/

November 13, 2007

Properties of gossip protocols [Bahkshi et al. 2007]

general

- simplicity
- scalability
- symmetry

functional

- connectivity
- convergence
- shortest path length

non-functional

- time complexity
- message complexity
- elasticity

$N \to \infty$ instead of N large?

- gossip protocols are most suitable if the number of nodes N is large
- model-based simulation or analytical evaluation of such systems is in general very costly
- is it possible to calculate the limit of the measure of interest for $N \to \infty$?
- yes — sometimes it is!
- convergence to *mean field*, a deterministic dynamical system which constitutes a limit for $N \to \infty$ [Le Boudec et al. 2007]

$N \longrightarrow \infty$ instead of N large?

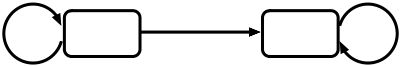
- gossip protocols are most suitable if the number of nodes N is large
- model-based simulation or analytical evaluation of such systems is in general very costly
- is it possible to calculate the limit of the measure of interest for $N \to \infty$?
- yes — sometimes it is!
- convergence to *mean field*, a deterministic dynamical system which constitutes a limit for $N \to \infty$ [Le Boudec et al. 2007]

$N \longrightarrow \infty$ instead of N large?

- gossip protocols are most suitable if the number of nodes N is large
- model-based simulation or analytical evaluation of such systems is in general very costly
- is it possible to calculate the limit of the measure of interest for $N \rightarrow \infty$?
- yes — sometimes it is!
- convergence to *mean field*, a deterministic dynamical system which constitutes a limit for $N \rightarrow \infty$ [Le Boudec et al. 2007]

Interacting objects

- N interacting objects
- each with discrete (finite) state space S
- time is discrete, $T = 0, 1, 2, \ldots$
- random variable $X_n^N(t)$ state of object n at time t



Occupancy measure & global memory

- random variable $M_i^N(t)$: **fraction of objects** that is in state i at time t

$$M_i^N(t) = \frac{1}{N} \sum_{n=1}^{N} 1_{\{X_n^N(t)=i\}}$$

- random variable $\overline{R}^N(t)$: **history** of the occupancy measure

$$\overline{R}^N(t+1) = g(\overline{R}^N(t), \overline{M}^N(t+1))$$

Transitions

- transition probabilities for each object depend on
  - the current state i
  - the current global memory $\bar{r}$

$$P^N_{i,j}(\bar{r}) = \Pr\{X^N_n(t+1) = j \mid \overline{R}^N(t) = \bar{r}, X^N_n(t) = i\}$$

$$\lim_{N \to \infty} P^N_{i,j}(\bar{r}) = P_{i,j}(\bar{r})$$

Convergence to mean field

assume

- $\overline{M}^N(0)$ converges almost surely to deterministic limit $\overline{\mu}(0)$,
- $\overline{R}^N(0)$ converges almost surely to deterministic limit $\overline{\rho}(0)$.

define

$$\left.\begin{array}{rcl} \overline{\mu}(t+1) & = & \overline{\mu}(t) \cdot \mathbf{P}(\overline{\rho}(t)), \\ \overline{\rho}(t+1) & = & g(\overline{\rho}(t), \overline{\mu}(t+1)). \end{array}\right\} \text{mean field}$$

then for any t, almost surely,

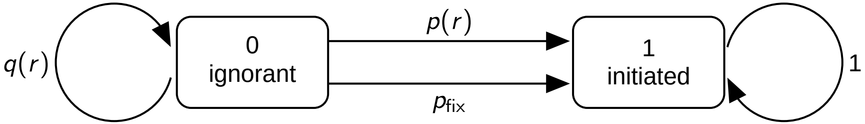
$$\lim_{N\to\infty} \overline{M}^N(t) = \overline{\mu}(t) \quad \text{and} \quad \lim_{N\to\infty} \overline{R}^N(t) = \overline{\rho}(t).$$

(almost surely = with probability one)

An Example

- P2P network with N nodes
- random overlay, in-degree = out-degree = C
- we want to model the dissemination of a message
- with probability p_{fix} a nodes receives the message from outside the network
- if a node is initiated it sends the message to K nodes in its partial view per time step

A single node



occupancy measure

$$\overline{M}^N(t) = (M_0^N(t), M_1^N(t))$$

global memory

$$R^N(t) = M_1^N(t+1)$$

transition probabilities

$$\mathbf{P}(r) = \begin{pmatrix} q(r) & p(r) + p_{\text{fix}} \\ 0 & 1 \end{pmatrix}$$

Transition probabilities

assume: N nodes, $r \cdot N$ nodes are infected

- look at a single, non-infected node and take a single in-link
  - probability that this in-link comes from an ignorant node

$$q_1(r) = \frac{N - rN}{N} = 1 - r$$

  - probability that this in-link comes from an initiated node but no message is sent (since only K are selected)

$$q_2(r) = \frac{rN}{N} \cdot \frac{C - K}{C} = r \cdot \frac{C - K}{C}$$

- probability that no message reaches this node from another node

$$q'(r) = \sum_{i=0}^{C} q_1^i \cdot q_2^{C-i}$$

- probability to get infected

$$q(r) = q'(r) - p_{\text{fix}}$$

Mean field

all nodes are ignorant at the beginning

$$\overline{\mu}(0) = (1, 0)$$

global memory corresponds to fraction of initiated nodes

$$\rho(0) = 0$$

equations for the mean field

$$\overline{\mu}(t+1) = \overline{\mu}(t) \cdot \begin{pmatrix} q(\rho(t)) & 1 - q(\rho(t)) \\ 0 & 1 \end{pmatrix}$$

$$\rho(t+1) = \mu_1(t+1)$$

Simulation for finite N

- modelling each object individually would result in 2^N states
- it suffices to keep track of the number of infected nodes ($N+1$ states)
- transition probability from i to $j \geq i$ for $r \cdot N$ infected nodes

$$\Pr(i \to j) = \binom{N-i}{j-i} \cdot q(r)^{N-j}(1-q(r))^{j-i}$$

Time complexity