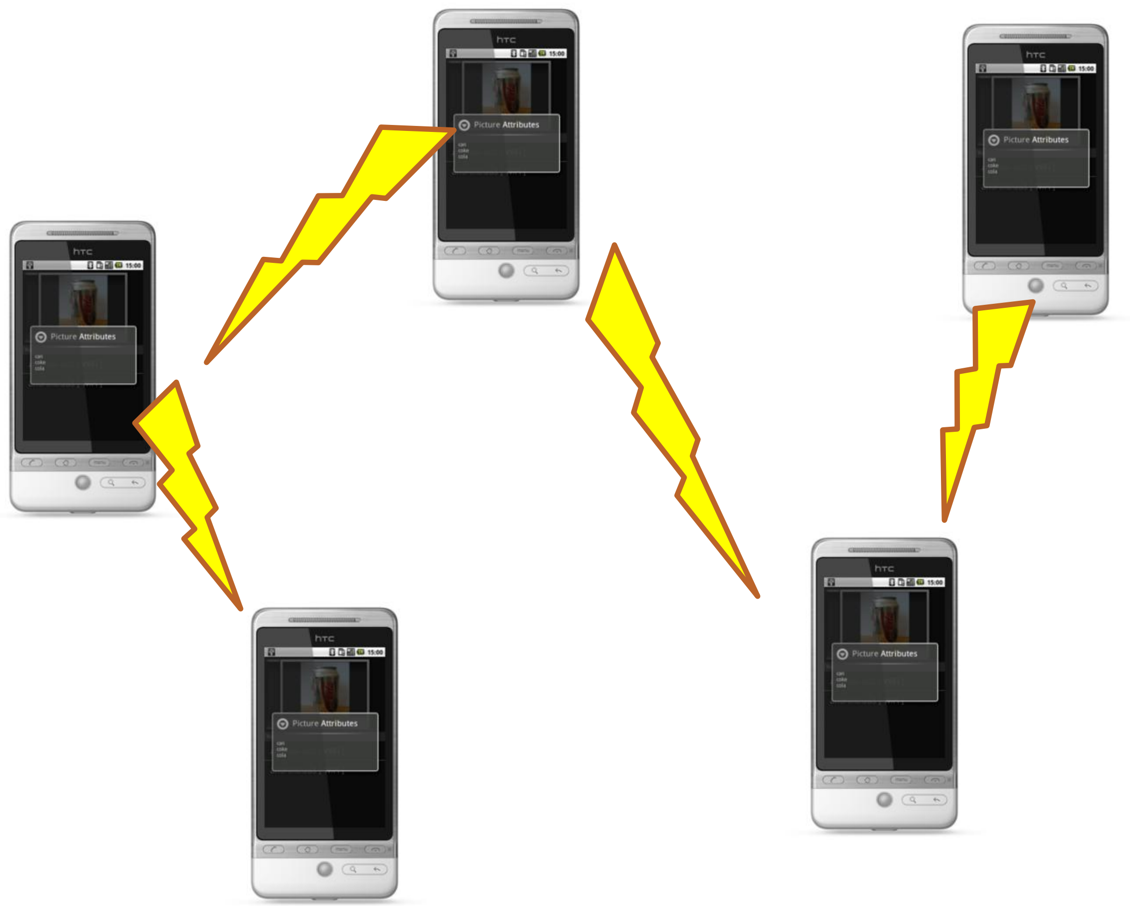
Social-aware Information Dissemination in Mobile Social Networks

M. Phil. Probation Talk

Hongxian Sun
Supervisor: Dr. C. Wu

Background

- The opportunistic contacts among mobile users bring lots of chances for information sharing.



Current Discussions & Motivation

- System performance

  discussed theoretically [Niyato2010, Ioannidis2009]

- Routing Protocols

  either with or without social awareness [Li2010, Hui2010]

- Assumption

  a user will exchange files with another **iff** they are *near*.

- A practical concern

  People always prefer interacting with their social friends.

How will message dissemination behave if the social ties of various strength are considered explicitly?

Our Angel of View

a MSN application with N users in an area A

- Metric

  end-to-end delivery delay between two randomly chosen users

- Structure of Friendship

➢ every two users are friends ***to some extent***

➢ **either one is a friend of another or not**

in this case, the social network of friendship forms a ***scale-free*** topology

Hopefully we could identify the role of 'social-hubs' and/or the affect of different social network models.

Related Work I

- Node Mobility Model

➢ Simple independent and ergodic processes
Random Way Point, Random Direction…
an **exponentially** distributed inter-contact time

➢ Human mobility pattern is complex
traces: a power-law fashion inter-contact time [Chaintreau2006]

➢ Simplification in our model
an exponentially distributed inter-contact time

➢ Rationality
Mathematical Tractability
partially consistent with empirical studies [Karagiannis2007]

Related Work II

- Delivery delay E[Td]

➢Markov Chain
number of copies as states [Groenevelt2005]

➢ODEs are fluid limit of Markov Chains
[Zhang2007, Banerjee2008]
not accurate because of many assumptions

➢Percolation theory
assume that message can propagate over a connected component of a network instantaneously
[Kong2008] discovers a critical value of node density ρ_c
if $\rho < \rho_c$, delay $\sim$ linear(init_dist(S-D))
else delay $\sim$ sublinear(init_dist(S-D))

Related Work III

- Epidemic Spreading in Scale-free Networks

  mean-field equations

➢ rich results on spatial behavior

  [Pastor-Satorras2000, Nekovee2007] show that an epidemic will never disappear in SIS dynamics (zero threshold for infection rate)

➢ limited result on temporal

  [Barthelemy2004] points out that in BA scale-free network, an epidemic pervades in a cascading fashion, from highest-degree nodes all the way down to lowest-degree ones.

➢ **No result obtained in terms of unicast delay**

➢ Starting point of our models

Model Framework

- Friendship

  a scale-free social graph with a power-law degree distribution

- Message Validity

➢ unlimited validity

  nodes never drop packets

➢ Limited validity

  a node may delete a message after carrying it for some finite time T

- Mean-field Equations

  govern the population of *spreaders* and *ignorants* in the system

  proposed by referencing the study on dynamics in scale-free networks

Friendship

- Social graph F*N*(V,E)
  - set of users $V = \{n_1, n_2, \ldots, n_N\}$
  - set of social connections E, $(n_i, n_l) \in E$ iff user n_i is an acquaintance of user n_l

- Degree distribution

$$P(k) = \begin{cases} 0, & k < m, \\ C(m,\gamma)k^{-\gamma}, & m \cdot \; k < N. \end{cases}$$

m is the smallest possible number of friends
γ depicts the skewness of the distribution

Mean-field Equations

$\langle k \rangle$	expected node degree in social graph $F_N(V,E)$
$i(k,t)$	percentage of ignorants of social degree k at time t, $i(t)=\sum_{k=m}^{N-1} i(k,t)$
$s(k,t)$	percentage of spreaders of social degree k at time t, $s(t)=\sum_{k=m}^{N-1} s(k,t)$
$\theta(t)$ (or $\theta(k,t)$)	probability that a friend of a given ignorant is a spreade
$\alpha(t)$	auxiliary function, $\alpha(t)=\int_0^t \theta(x)dx$
$e(k,t)$	percentage of timed-out nodes of social degree k at tim $e(t)=\sum_{k=m}^{N-1} e(k,t)$
$E[T_d]$	the expected end-to-end unicast delay
T	the timeout value

- Unlimited validity

$$\frac{di(k,t)}{dt} = -\lambda k i(k,t)\theta(k,t)$$
$$k = m, m+1, \ldots N-1.$$

- Limited validity

while t > T, writing g(k,t) = i(k,t)θ(t), we have

$$\frac{di(k,t)}{dt} = -\lambda k g(t) \quad (2)$$
$$\frac{ds(k,t)}{dt} = \lambda k (g(t) - g(t-T)) \quad (3)$$
$$\frac{de(k,t)}{dt} = \lambda k g(t-T) \quad (4)$$

- θ(k,t) is the probability that a friend of a given ignorant is a spreader. Neglecting degree correlations of F_N, θ(k,t)≡ θ(t) has the form

$$\theta(k,t) \approx \sum_{k'=m}^{N-1} \frac{k'-1}{\langle k \rangle} s(k',t) \qquad (5)$$

- Eqn. (1) ~ (5) together describes the populations of spreaders and ignorants. **But**

- Computing each i(k,t) is very difficult if N is very large
- No intuitive insights can be got from these equations

A Small Transformation

- Using a variable limit integral $\alpha(t) = \int_0^t \theta(x)dx$
- In unlimited validity case, an ODE holds

$$\frac{d\alpha(t)}{dt} = \theta(t) = \sum_{k=m}^{N-1} \frac{k-1}{\langle k \rangle}(P(k) - i(k,t))$$

$$= \sum_{k=m}^{N-1} \frac{(k-1)P(k)}{\langle k \rangle}(1 - \frac{N-1}{N} e^{-\lambda k \alpha(t)}) \qquad (6)$$

- For the case of limited message validity, we have the DDE

$$t < 0 : \alpha(t) = 0,$$

$$t > T : \frac{d\alpha(t)}{dt} =$$

$$\sum_{k=m}^{N-1} \frac{(k-1)P(k)}{\langle k \rangle}(\frac{1}{N} + \frac{N-1}{N}(e^{-\lambda k \alpha(t-T)} - e^{-\lambda k \alpha(t)}) \qquad (7)$$

- Further the expected delay goes like

$$E[T_d] = \int_0^{\infty} e^{-\lambda \langle k \rangle \alpha(t)} dt \qquad (8)$$

- It’ s valid for both two cases.

- With Eqn. (7) and (9), we can rigorously prove that the delay increases with γ and decreases with m monotonically.

The models

- many assumptions and simplifications
- moderate accuracy
- good enough to capture the performance trend

Empirical Studies

- Steady behavior of delay w.r.t. network size

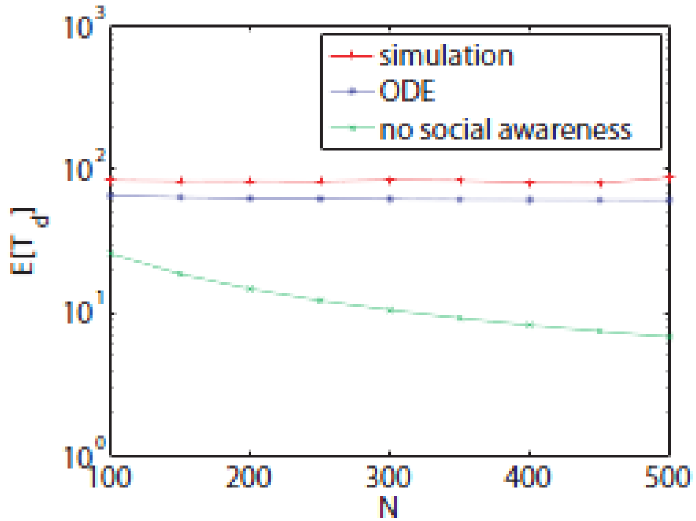


Figure 1. Expected delivery delay vs. network size.

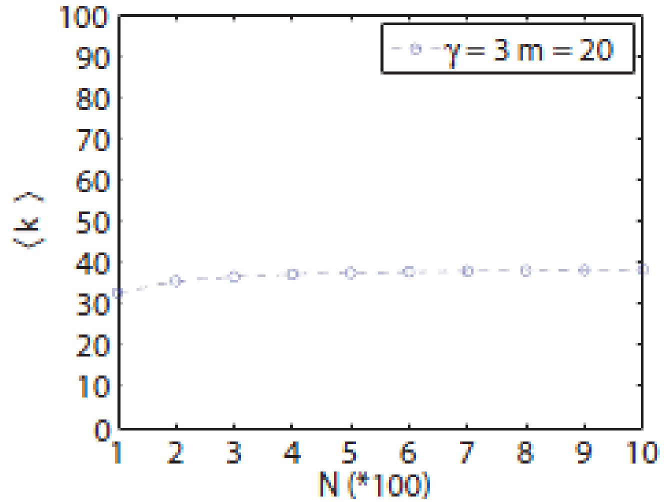


Figure 2. Average number of friends per user vs. N.

- Delay vs. m and γ

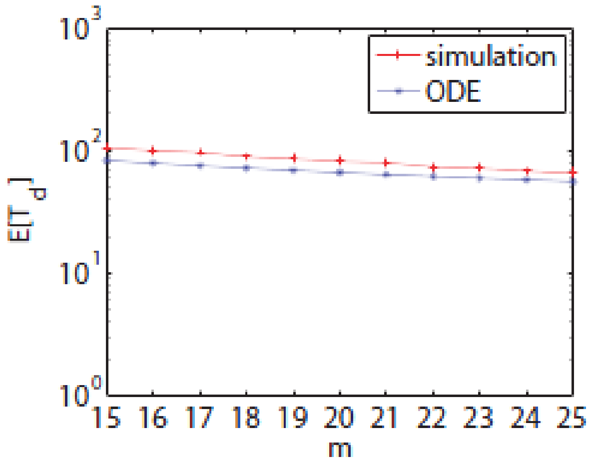


Figure 3. $E[T_d]$ vs. m.

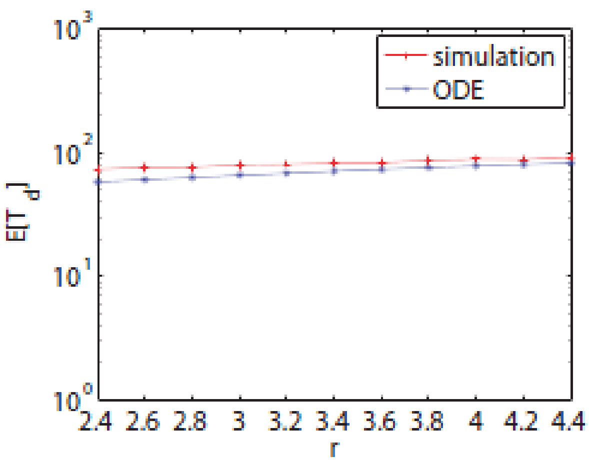


Figure 4. $E[T_d]$ vs. γ.

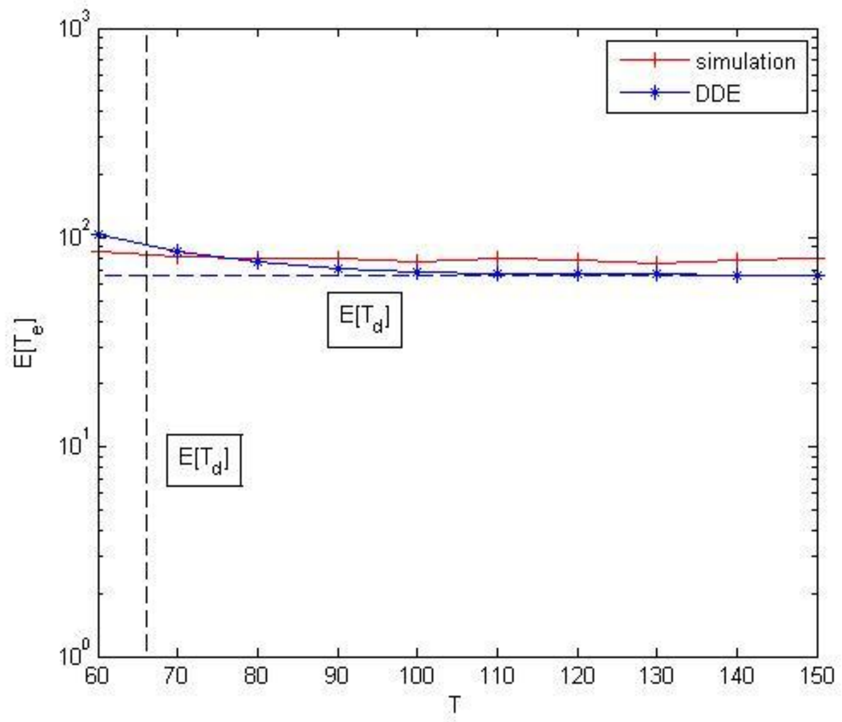
- Verify the theoretical result

Limited Message Validity

The red line corresponds to

$$\frac{\text{expected delay of successful dilveries}}{\text{delivery ratio}}$$

While the blue line is the solution of the DDE.



With a timeout value T slightly larger than the expected delay obtained under the unlimited validity case, we achieve perfect trade-off between successful delivery ratio and power consumption.

Outlook

- Continue on message propagation

- Utimate Goal:

  give a clear yet vivid picture on how inhomogeneous friendship affects message dissemination

- A continuous characterization of friendship

  one possible way to go:

  a node n_i is willing to exchange message with any other node with probability p_i. In a system $\{p_i\}$ are samples from some distribution.

  nontrivial, nodes are stochastically different