MINIMUM LATENCY DATA DIFFUSION IN INTERMITTENTLY CONNECTED MOBILE NETWORKS

Summary and Discussion for SE4960 by Steve Mazza

Introduction

Authors, Publications, References, and Citations

Authors

- Sathiamoorthy, M.
  - Ming Hsieh Dept. of Electr. Eng., Univ. of Southern California, Los Angeles, CA, USA
- Wei Gao
  - Department of Computer Science and Engineering, Penn State Univ., State College, PA
- Krishnamachari, B.
  - Ming Hsieh Dept. of Electr. Eng., Univ. of Southern California, Los Angeles, CA, USA
- Guohong Cao
  - Department of Computer Science and Engineering, Penn State Univ., State College, PA

Publication

- Vehicular Technology Conference (VTC Spring), 2012 IEEE 75th
  - **Date of Conference:** 6-9 May 2012
  - **Location:** Yokohama
  - Page(s): 1 - 5
  - ISSN : 1550-2252
  - E-ISBN : 978-1-4673-0988-2
  - Print ISBN: 978-1-4673-0989-9
  - INSPEC Accession Number: 12866080

References

1. D. Bertsekas, "Dynamic programming and optimal control," Athena Scientific Belmont, MA, 1995.
2. A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," CS-2000-06, Duke University, Tech. Rep., 2000.
3. T. Spyropoulos, K. Psounis, and C. Raghavendra, "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in Proc. of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking. ACM, 2005, pp. 252-259.
4. -, "Spray and focus: Efficient mobility-assisted routing for heterogeneous and correlated mobility," in PerCom Workshops '07. IEEE, 2007, pp. 79-85.
   Abstract | Full Text: PDF (165KB)
5. Y. Zhang, W. Gao, G. Cao, T. La Porta, B. Krishnamachari, and A. Iyengar, "Social-aware data diffusion in delay tolerant manets," Handbook of Optimization in Complex Networks Communication and Social Networks, 2011.
   [CrossRef]
6. M. McPherson, L. Smith-Lovin, and J. Cook, "Birds of a feather: Homophily in social networks," Annual review of sociology, pp. 415-444, 2001.
   [CrossRef]
7. W. Gao and G. Cao, "User-centric data dissemination in disruption tolerant networks," in Proc. of Infocom, 2011.
   Abstract | Full Text: PDF (687KB) | Full Text: HTML
8. W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: a social network perspective," in Proc. of MobiHoc. ACM, 2009, pp. 299-308.
   Full Text: Access at ACM
9. C. Boldrini, M. Conti, and A. Passarella, "Contentplace: social-aware data dissemination in opportunistic networks," in Proc. of MSWiM. ACM, 2008, pp. 203-210.
   Full Text: Access at ACM
10. V. Lenders, G. Karlsson, and M. May, "Wireless ad hoc podcasting," in Proc. of SECON, june 2007, pp. 273-283.
    Abstract | Full Text: PDF (6794KB)
11. E. Yoneki, P. Hui, S. Chan, and J. Crowcroft, "A socio-aware overlay for publish/subscribe communication in delay tolerant networks," in Proc. of MSWiM. ACM, 2007, pp. 225-234.
    [CrossRef]

Citations

- Citation data not available
- 47 usages since July of 2012
- APA citation
  - Sathiamoorthy, M., Gao, W., Krishnamachari, B. & Cao, G. (2012). Minimum Latency Data Diffusion in Intermittently Connected Mobile Networks.. *VTC Spring* (p./pp. 1-5), : IEEE. ISBN: 978-1-4673-0989-9

Summary

Abstract and brief look at the problem space.

Abstract

"…diffusing cached content in an intermittently connected mobile network, starting from a given initial configuration to a desirable goal state where all nodes interested in particular contents have a copy of their desired contents."

Goal: minimize time from initial condition to the desired state.

Why Do We Care?

- Increasing number of mobile devices.
- Commercial interest in vehicular networks.
- Facilitate data dissemination in disconnected, intermittent, limited (DIL) networks.
- Good model for information sharing on a peer model.
- Interesting application in federation of remote data stores (e.g., TiGR, CPoF, tactical marketplace).

Characteristics

- Node storage constraints are considered.
- Bandwidth is presumed to be limited.
- Transactions are atomic.
- Selfish strategies are not generally effective.
- Nodes participate in data routing.
- Modeled as a stochastic shortest path (SSP) problem.

Related Work

Literature Search

Epidemic Routing

- Do not tend to be sensitive to bandwidth and storage constraints.
- Tend to work best when routing to a single node.

Socio-aware Diffusion

- Nodes share similar content with “friends” and dissimilar content with “strangers.”
- Similar people tend to become friends (homophily phenomenon).
- Friends meet often; strangers meet infrequently.

PodNet

Podcasting system for mobile devices using wireless and ad-hoc networks.

Models and assumptions

Scoping the work.

Setup

- Nodes (or Users): i, (i = 1 .. N)
- Data Objects: d, (d_1 .. d_K)
- Each user, i, has capacity C.
- E: data incidence matrix (which node has what data object).
  - Element (i, k) indicates if node i has datum k
  - Values are 1 or 0
- E_t: interest matrix (also the desired end state)

Contact Process

- At each time step, 2 nodes are chosen at random with some pre-defined probability.
- Thus there are $N(N - 1)$ possible encounters.
- During each encounter, only 1 of the 2 nodes will be allowed to alter its cache.
- Use ordered pair notation (i, j) to show that i modifies its cache based on data from j.
- Each (i, j) is assigned a non-negative probability p_{ij} such that $\sum_{i \neq j} p_{ij} = 1$.

Stochastic Shortest Path

Discussion of the model implementation.

Representation

SSP problem is represented by a 5-tuple containing the following information:

1. S: The state space
2. t, in S: Representation of the terminating state
3. A: Finite set of actions
4. P_a: Probability matrix controlling state transitions
5. $C_a(s, s')$: Cost associated with each state transition (i.e., from s to s')

State Space

- Denoted by
  - $s = \{E,(i,j)\}$ for all non-terminal states, and
  - E_t for the terminal state.
- Some non-terminal states are illegal
  - State results in lost datum
  - Node stores more than its capacity

Actions

- For all non-terminal states
  - $a = [d_{remove}, d_{download}]$ where
    - $D_{remove} = 0 .. k$, and
    - $D_{download} = 0 .. K$, where
      - 0 represents no action, and
      - k indexes the datum on which the action occurs.
- For the terminal state, $\{E_t\}$,
  - No further action takes place

Costs

The cost of transition from any state, *s*, to any other state, s', is set to 1 except as follows:

- Cost is set to 0 if *s* is a terminating state, and
- Cost is set to ∞ if s' is an illegal state.

Markovian Dynamics

- Defines the probabilities of transitioning from one state to another as based on
  - The probability of pair selection,
  - Values in the data incidence matrix, and
  - Values in the data interest matrix.

4-node Example

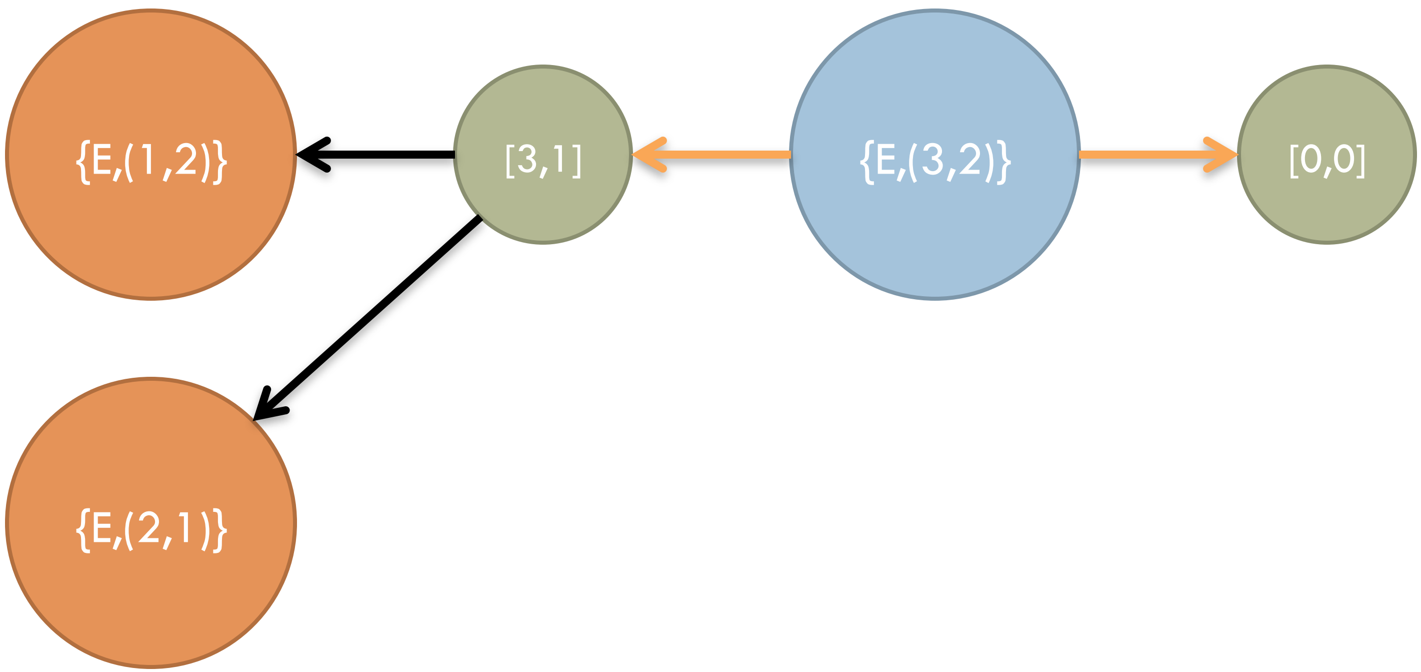
- Let $N = 4$
- Let $K = 4$
- Let $C = 2$
- Let nodes 1 & 2 desire d1 & d2
- Let node 3 desire d_3 & d_4
- Let node 4 desire d_2 & d_3

4-node Example (cont.)

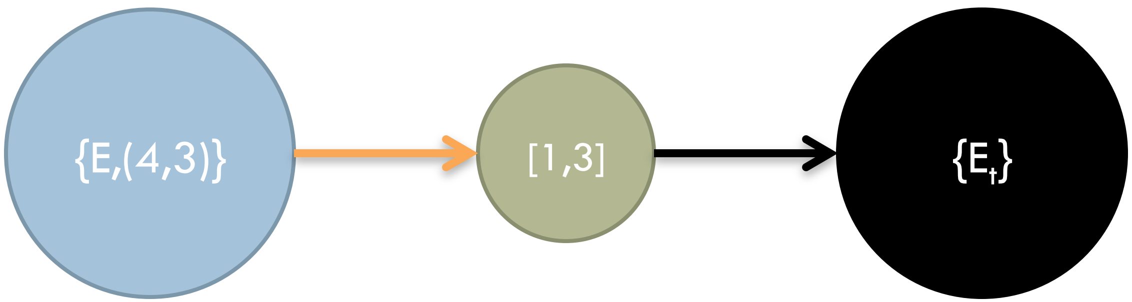
Interest Matrix (terminating state)			
1	1	0	0
1	1	0	0
0	0	1	1
0	1	1	0

Data Incidence Matrix			
1	1	0	0
1	1	0	0
0	0	1	1
1	1	0	0

4-node Example (cont.)



4-node Example (cont.)



Dynamic Program Issues

- Network configurations (connectivity) may prevent convergence on the terminal state.
- Value fragmentation in large networks may prevent convergence in Polynomial time.
- $s \to s'$ with probability $P_a(s,s')$
- $s \to s'$ with cost $C_a(s,s')$
- Refer to the paper for more notes on program formulation.

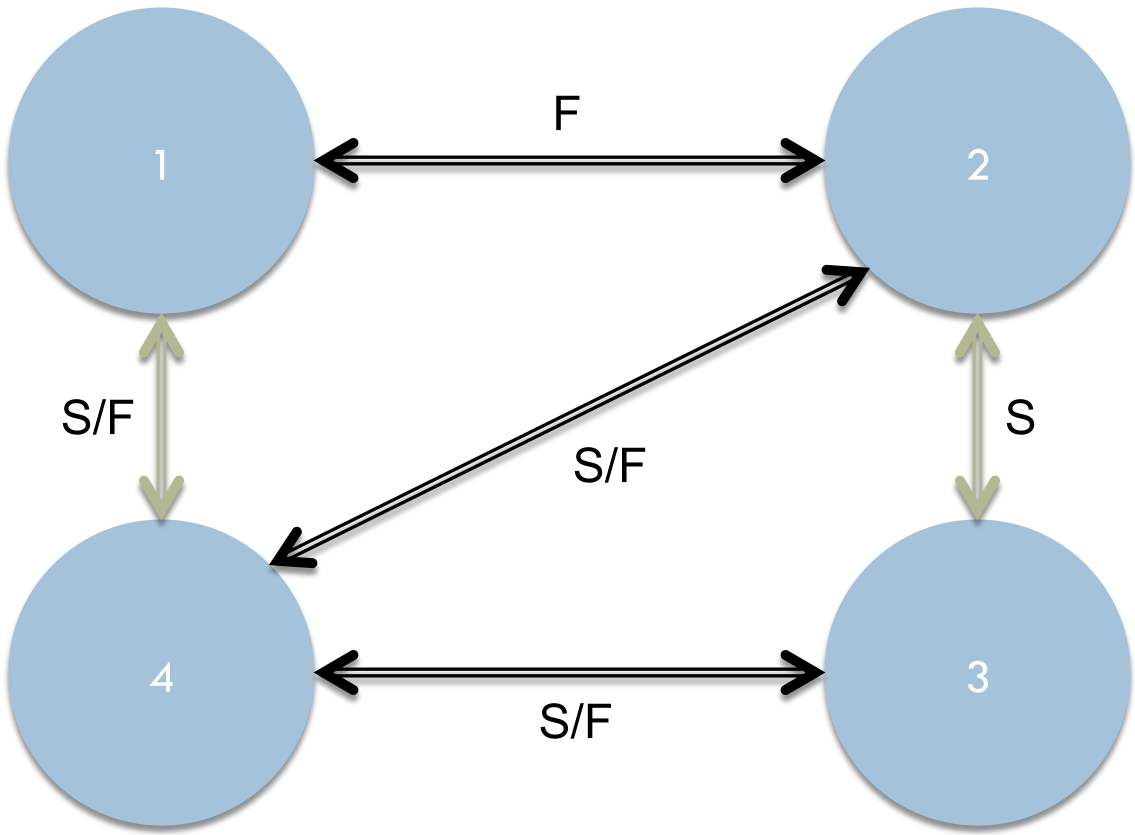
Simulation and Results

What did the model tell us?

Simulation Setup

- 2 identical networks
  - $N = 4$
  - $K = 4$
  - $C = 2$
- Vary contact probabilities
- Treat (i,j) and (j,i) equal
- Actions*
  - i downloads content it needs
  - i downloads content it does not need
  - i does not download any content

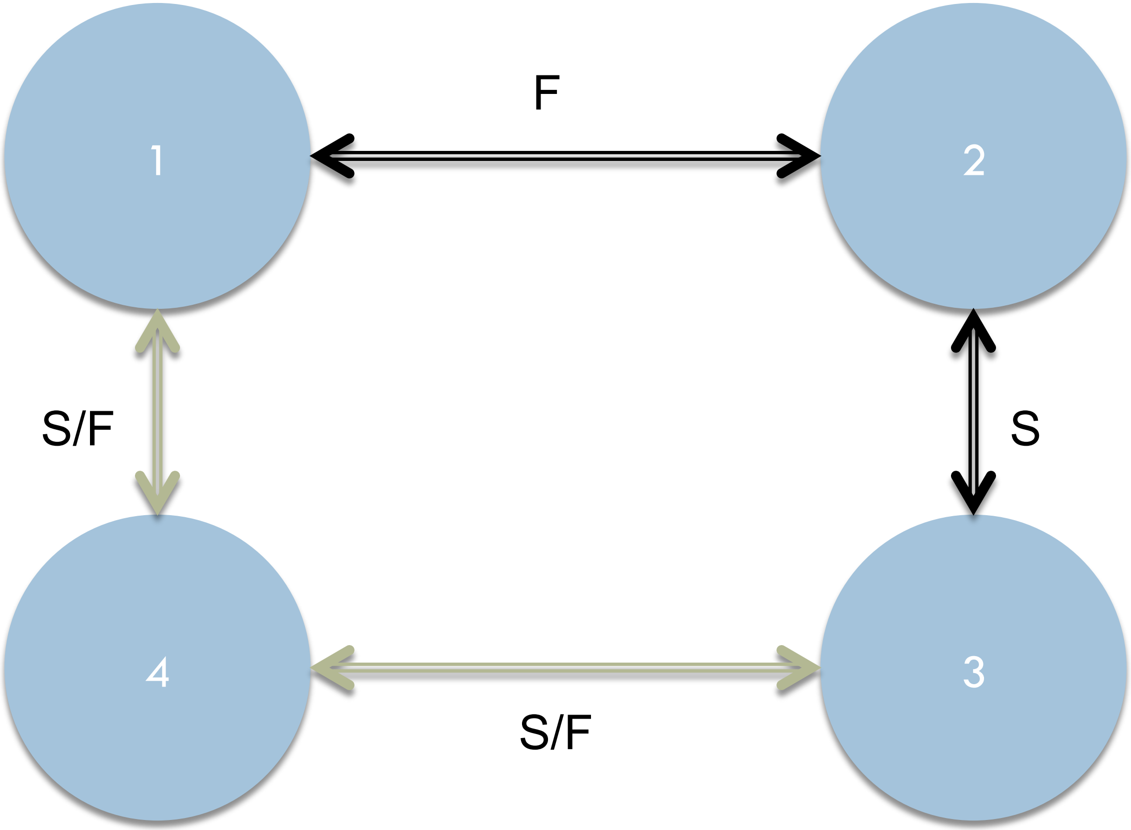
First Setup



First Setup Notes

- Node 1 is mostly independent (i.e., no buffering)
  - Shares all interests with node 2
  - Only meets node 2 frequently
  - Downloads content it needs with frequency 0.4445.
- Node 4 buffers data for node 3
  - Downloads data it does not need even when it has not yet satisfied its own needs with frequency 0.1136.
  - This same frequency for node 1 is 0.

Second Setup



Second Setup Notes

- This setup removes correlation between friendship and contact frequency by allowing some Strangers to meet more often.
- Node 2 buffers data for node 3 from node 1 with frequency 0.2087
  - This is considered a relatively high frequency for this type of event.
  - Prevents a slow route through node 4.
- Conclude that contact probabilities matter, not friendships.

Conclusion

Takeaway and opportunity for follow-on work.

Takeaway

Data diffusion in intermittently connected mobile networks can be effectively modeled as a stochastic shortest path problem.

Follow-on Work

- Derive better heuristics from the solution.
- Identify faster solutions for SSP.