Social-Aware Routing in Delay Tolerant Networks

Jie Wu

Dept. of Computer and Info. Sciences

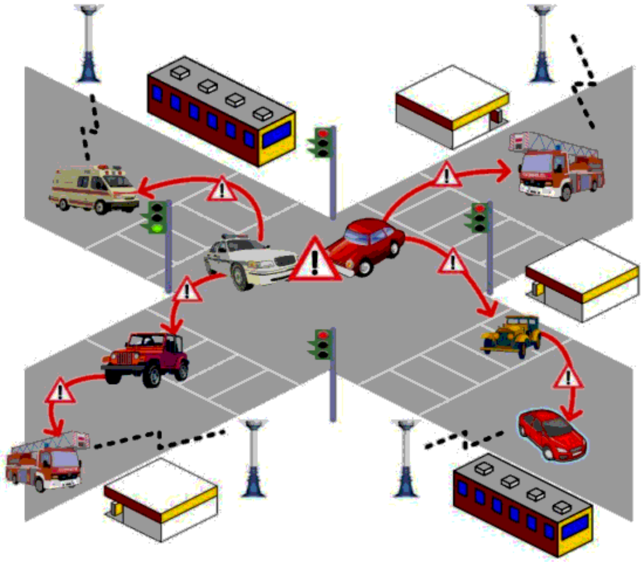
Temple University

Challenged Networks

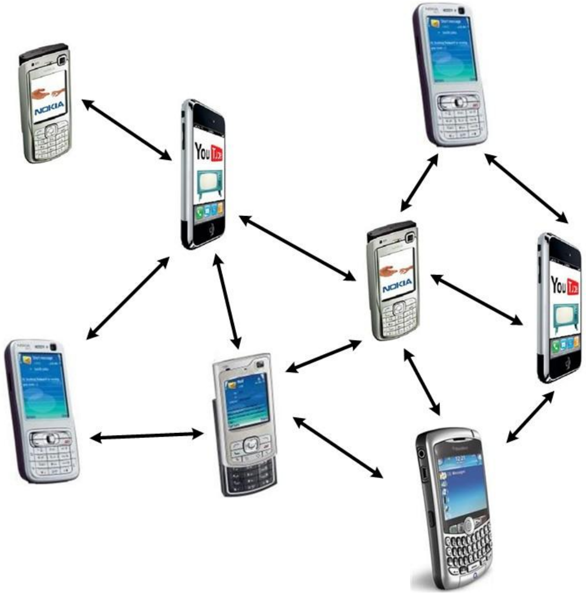
- Assumptions in the TCP/IP model are violated
  - DTNs
    - Delay-Tolerant Networks (also Disruption-Tolerant Networks)
  - Limited end-to-end connectivity
    - Due to mobility, power saving, or unreliable networks
  - Activities
    - IRTF's DTRNRG (Delay Tolerant Net. Research Group)
    - EU's Haggle project

Examples of DTNs

- Vehicular communication



- Social contact networks

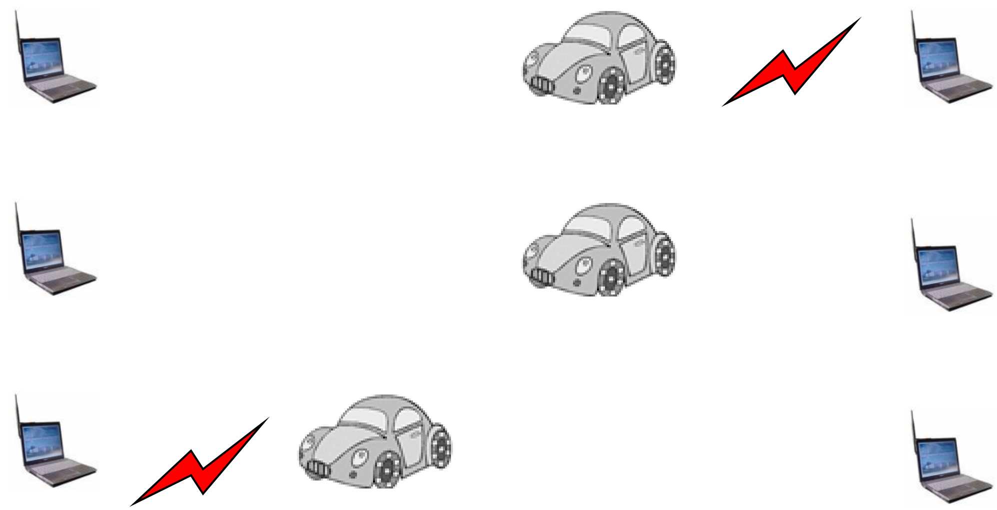


Store-Carry-Forward

- Movement-Assisted Routing

  Views node movement as a desirable feature

  - Store
  - Carry
  - Forward



Two Paradigms

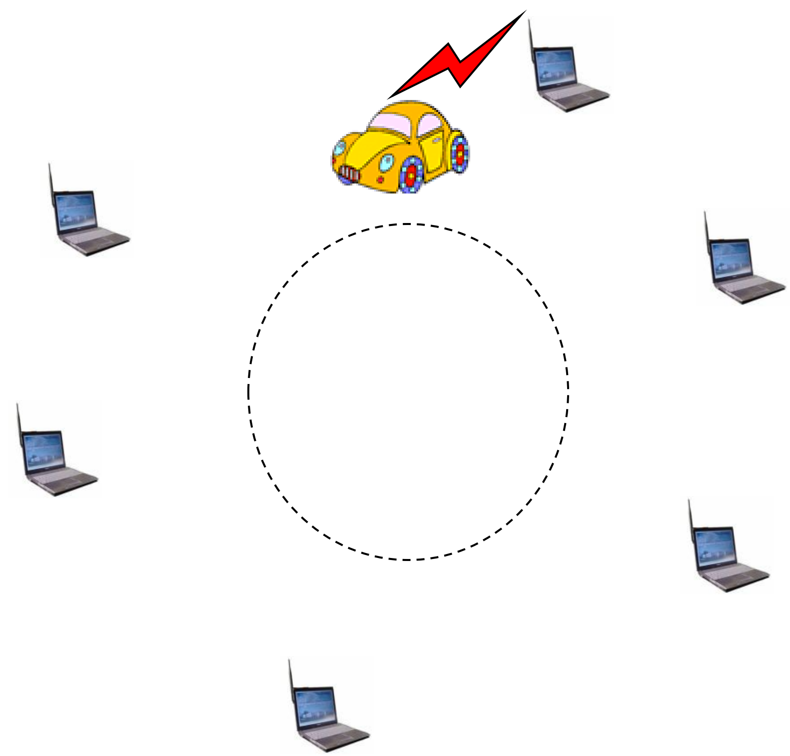
- Random Mobility
  - E.g., epidemic routing
  - Sightseeing cars (random movement)
- Controlled Mobility
  - E.g., message ferrying
  - Taxi (destination-oriented)
  - Public transportation (fixed route)

Epidemic Routing (Vahdat & Becker '00)

- Nodes store data and exchange it when they meet
- Data is replicated through a random talk

Message Ferrying (Zhao & Ammar '03)

- Special nodes (ferries) have completely predictable routes throughout the geographic area



Key Techniques

- Knowledge
  - Global vs. local information (Jain et al. '04)
  - Zero information
- Replication
  - Single vs. multiple copy: spray-and-wait (-focus)
  - Controlled copy: delegation forwarding
- Closeness (to destination)
  - Location information (of contacts and dest.)
  - Similarity (between intermediate nodes and dest.)

Two Fundamental Questions

- How to determine the appropriate number of copies
- How to efficiently distribute copies once determined

Delegation Forwarding (Erramilli et al. '08)

- The holder forwards the message to an encounter with a higher quality than those in all previous nodes seen so far
- The expected cost of the algorithms
  - Flooding: $O(n)$
  - Delegation forwarding: $O(\sqrt{n})$
- Extensions
  - Probabilistic delegation forwarding (Chen, Jian, Graves & Wu '09)
  - Delegation forwarding in multicasting (Wang & Wu '10)

Spray-and-Wait (-Focus)

(Spyropoulous, Psounis & Raghavendra '05 & '07)

Two phases

- Sprays a number of copies into the network
- Each copy waits until meeting with the destination
  - Wait : forward only to its destination
  - Focus : forward to destination or a node "closer" to the destination
- Mobility model: random walk

Key Challenges

- Existing DTN routing relies on
  - Contact history
  - Mobility pattern
- Collecting such information is costly
- Both contact and mobility are
  - Highly dynamic
  - Unstructured

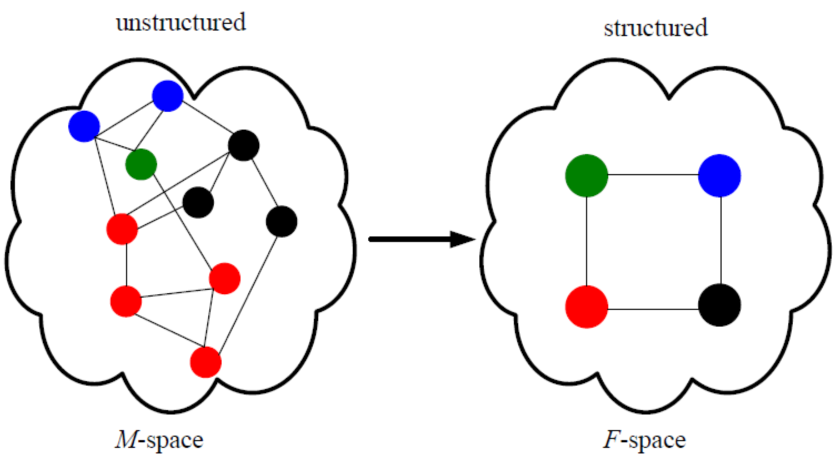
Social-Aware Routing

- Based on coarse-grain social-aware approaches
  - Social features-based: semi-structured contacts (Wu & Wang '12, INFOCOM 2012)
  - Social home-based: semi-structured mobility (Wu, Xiao & Huang '13, INFOCOM 2013)
- Differ from fine-grain social-aware approaches
  - Community, centrality, betweenness, and strong and weak ties, …
  - …

Social Feature-Based (Wu & Wang '12)

- Mobile & unstructured contact space (M-space) ⟶

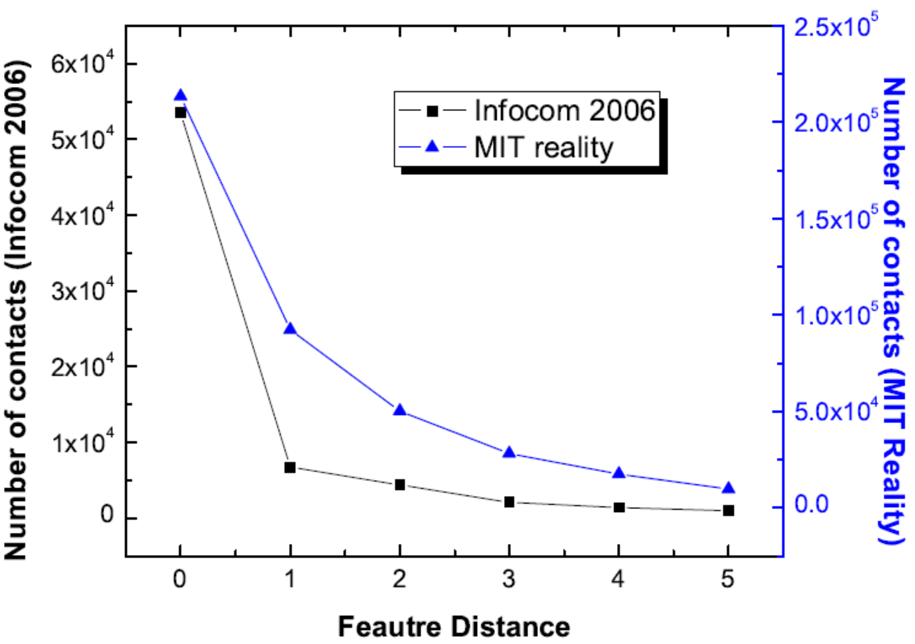
  Static & structured feature space (F-space)



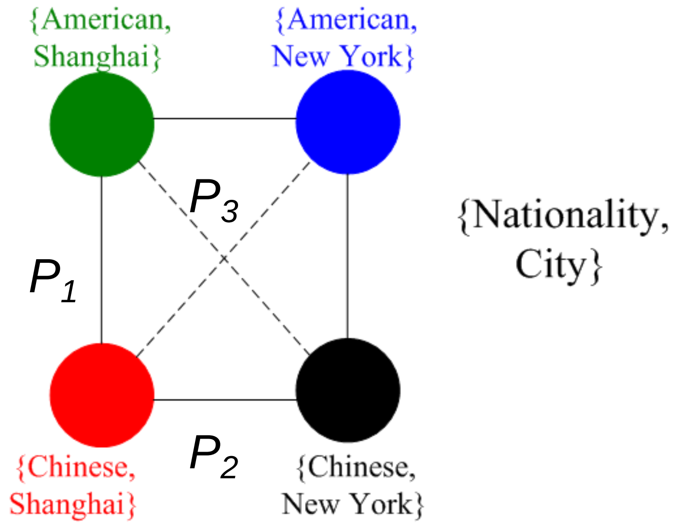
- Each individual with a social feature profile *{F_1, F_2,…}*
- Individuals with the same features mapped to a group

Resolve Feature Difference

- Our approach
  - Feature-based (race, gender, language, …)
  - Resolve feature difference one at a time !
- Increase delivery rate: multiple copy
  - Flooding: *O(n)*
  - Delegation forwarding: $O(\sqrt{n})$
  - Our approach: *log n**
  - n* is feature space size

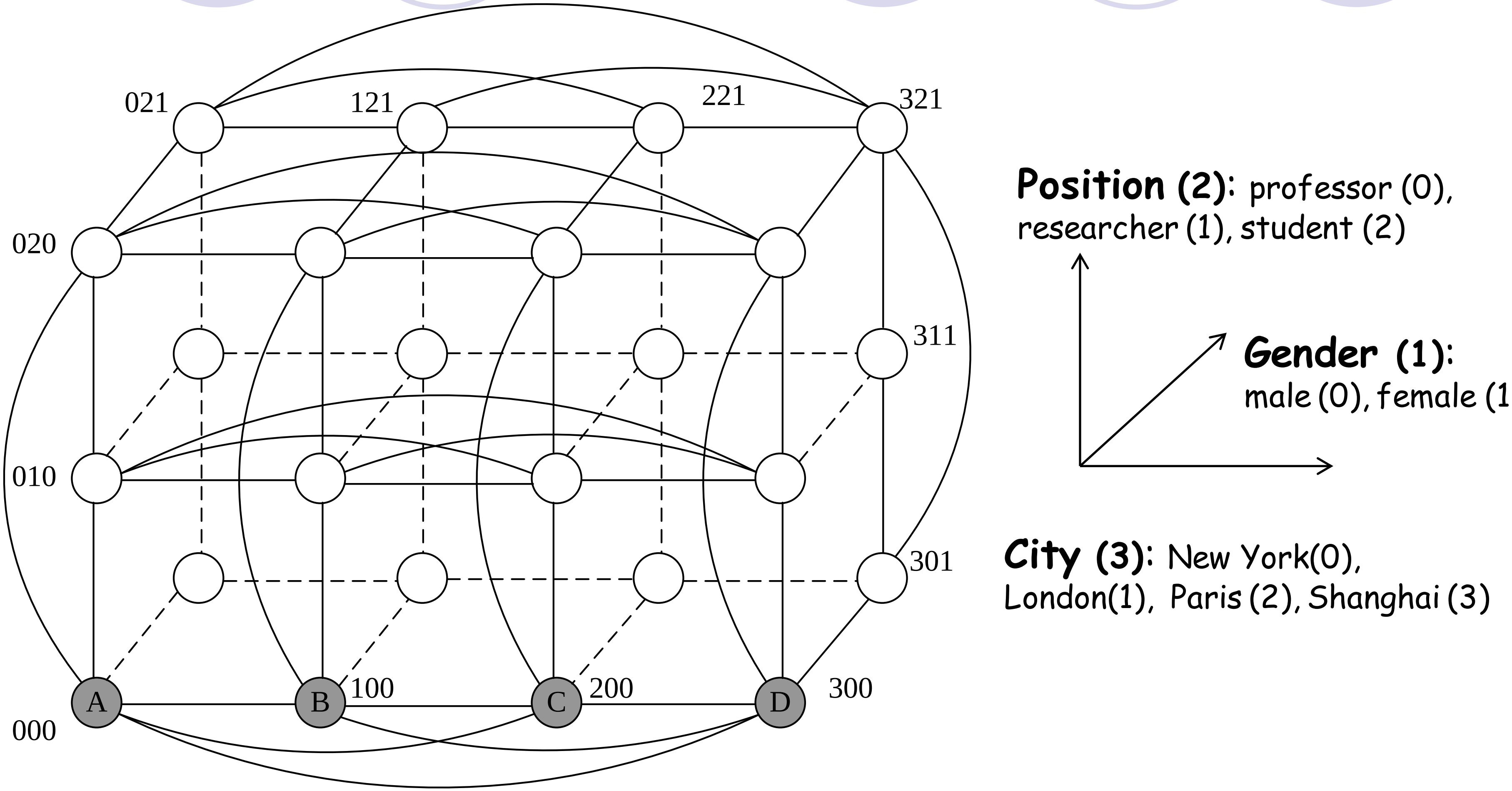


Feature-based Grouping Example



- People come in contact with each other more frequently if they have more social features in common ($P_1 > P_3$, $P_2 > P_3$)

A 3-dimensional (3-D) Hypercube



Example: "*311*": a female researcher lives in Shanghai

Feature Extraction

- Extract *m* most important features based on Shannon's entropy

$$E(F) = \sum_{i=1}^{l} p(x_i) \log p(x_i)$$

  - *E(F)*: entropy of feature *F*
  - *p:* feature probability mass function
  - $\{x_1, x_2, \ldots, x_l\}$: value set of feature *F*

Social Feature Entropy

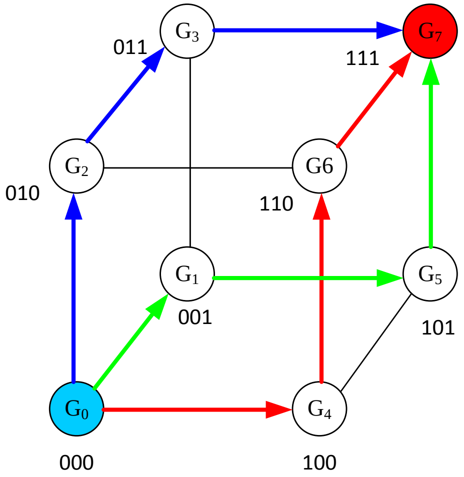
- The entropy of each social feature (Infocom 2006 trace)

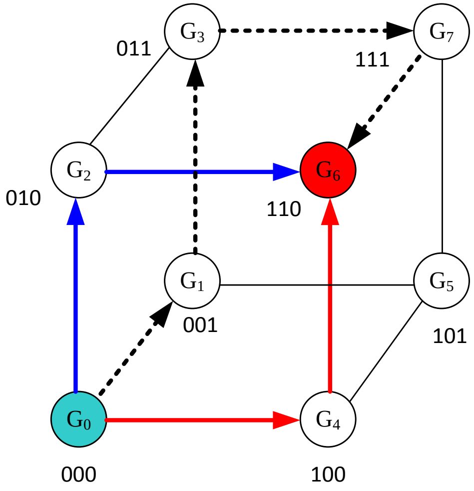
Social Feature	Entropy
Affiliation	*4.64*
City	*4.45*
Nationality	*4.11*
Language	*4.11*
Country	*3.59*
Position	*1.37*

Property of Hypercubes

Efficient routing in an *m*-d binary cube (*S* and *D* differ in *k* features)

- *k* shortest paths of length *k*
- *m-k* non-shortest paths of length *k+2*



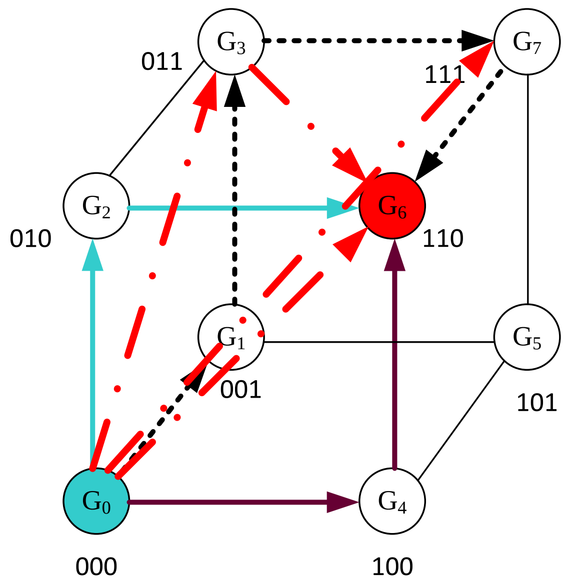


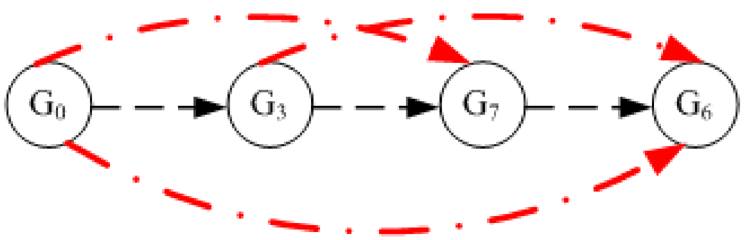
Hypercube Routing

- The relative address of the current group and destination group (a small string in the header)
  - calculated through XOR on S and D
  - sent, along with the packet, to the next node
- Any node in the group can forward to any node in the adjacent group
- Special treatment is needed at the destination group: spread-and-wait (-focus)

Shortcuts

- Feature matching shortcut can resolve the feature distance more than one at a time





- Shortcut reduces the number of forwardings while ensuring the path disjointness property

Simulation

- Synthetic trace
  - A node *A* has *m* contact frequencies, $p_1, p_2, \ldots, p_m$, with its *m* neighbors in the *m*-D F-space.
  - 128 individuals
  - *m* is 4, 5, 7, 8, and 11.
- Real trace
  - Infocom 2006 trace.
  - 61 participants
  - 6 social features, *m* is 3, 4, 5, and 6.

7 routing schemes comparison

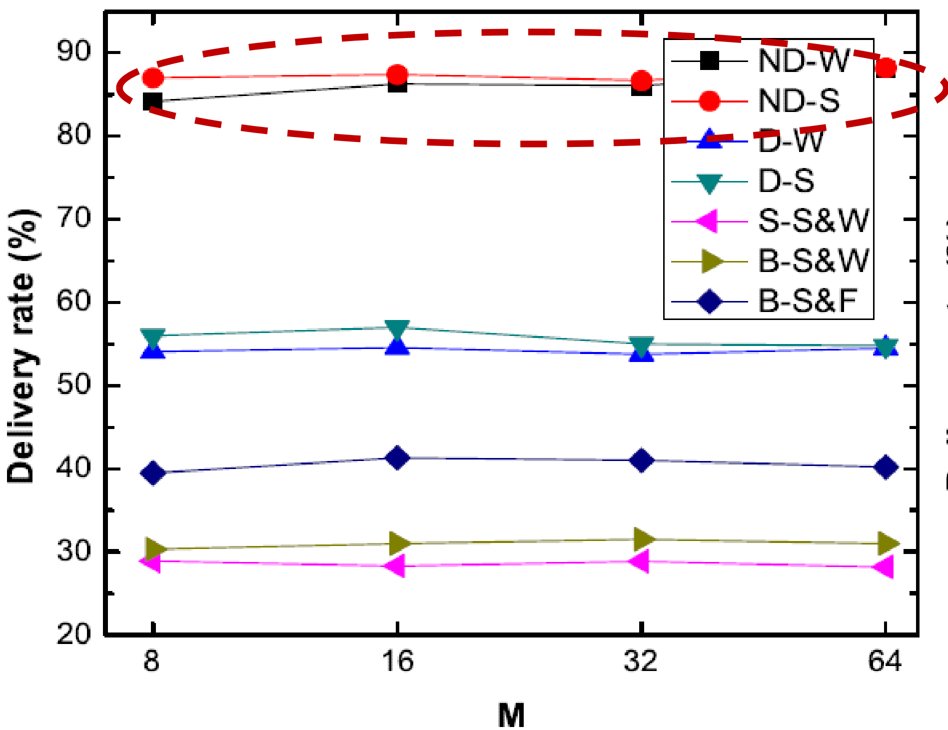
- Node-disjoint-based with wait-at-destination (ND-W)
  - Waiting for the destination after the packet enters the destination group
- Node-disjoint-based with spray-at-destination (ND-S)
  - Spraying N/(2M) copies into the destination group after the packet enters the group
- Delegation-based with wait (**D-W**)
- Delegation-based with spray (**D-S**)
  - Delegation-based: The copies of a packet is only forwarded to the individual with a smallest feature distance to the destination it has met so far

- Note: spray is needed at the destination group to increase the chances to meet the actual destination!

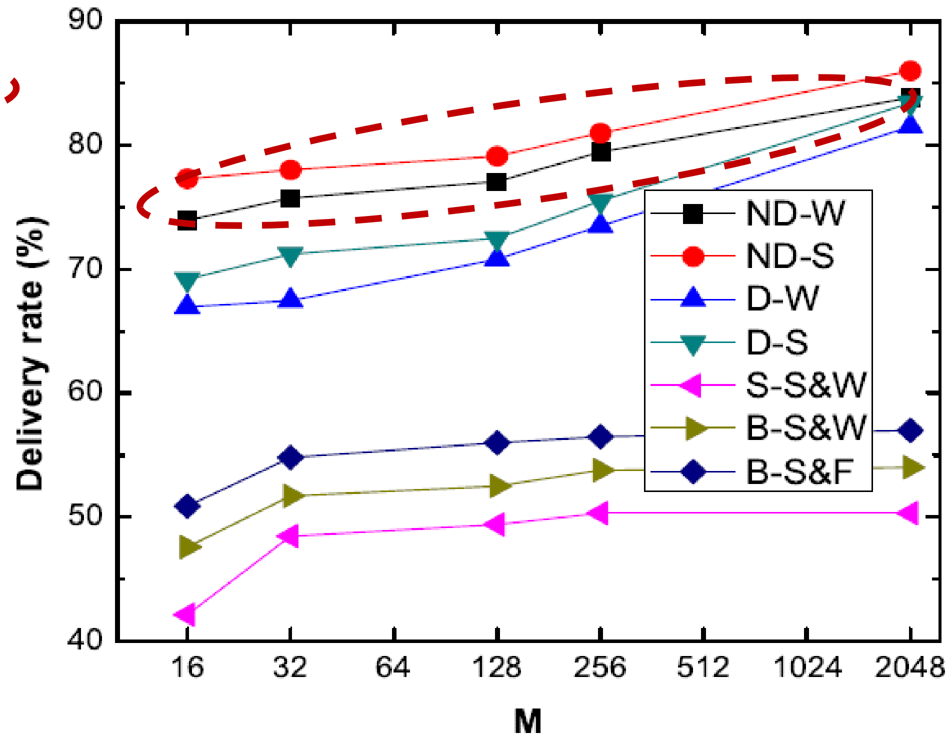
7 routing schemes comparison-cont'd

- Source spray-and-wait (**S-S&W**)
  - The source forwards copies to the first m distinct nodes it encounters. If the destination is not found, the copy carriers wait for the destination.
- Binary spray-and-wait (B-S&W)
  - Any node with copies will forward half of the copies to the encountered node with no copy.
- Binary spray-and-focus (B-S&F)
  - The copy carriers forward the copy to the encountered node with a smaller feature distance to the destination.

Varying node density – delivery rate

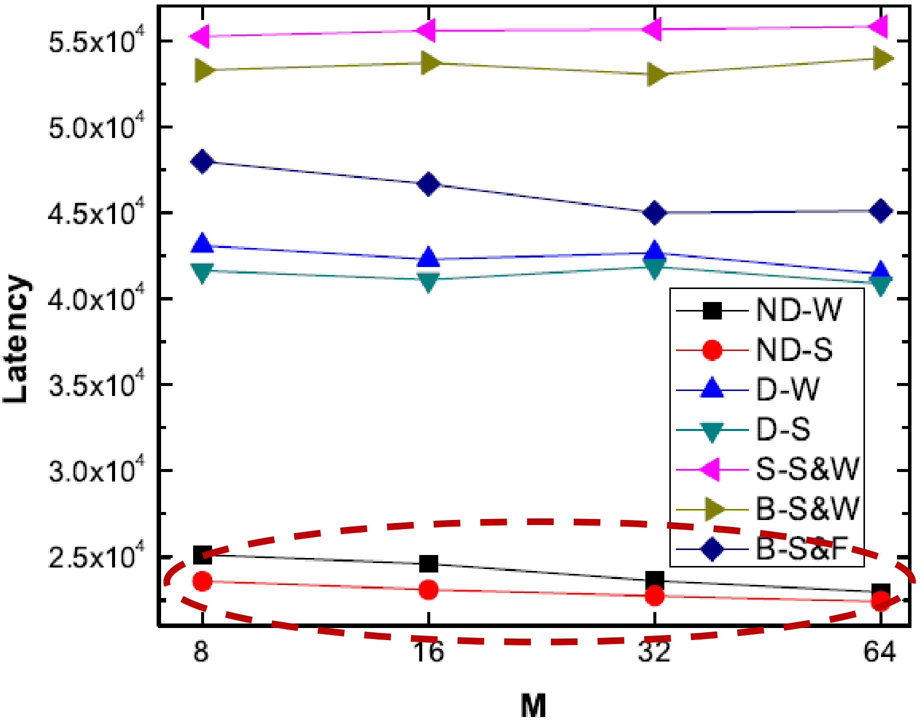


Real trace

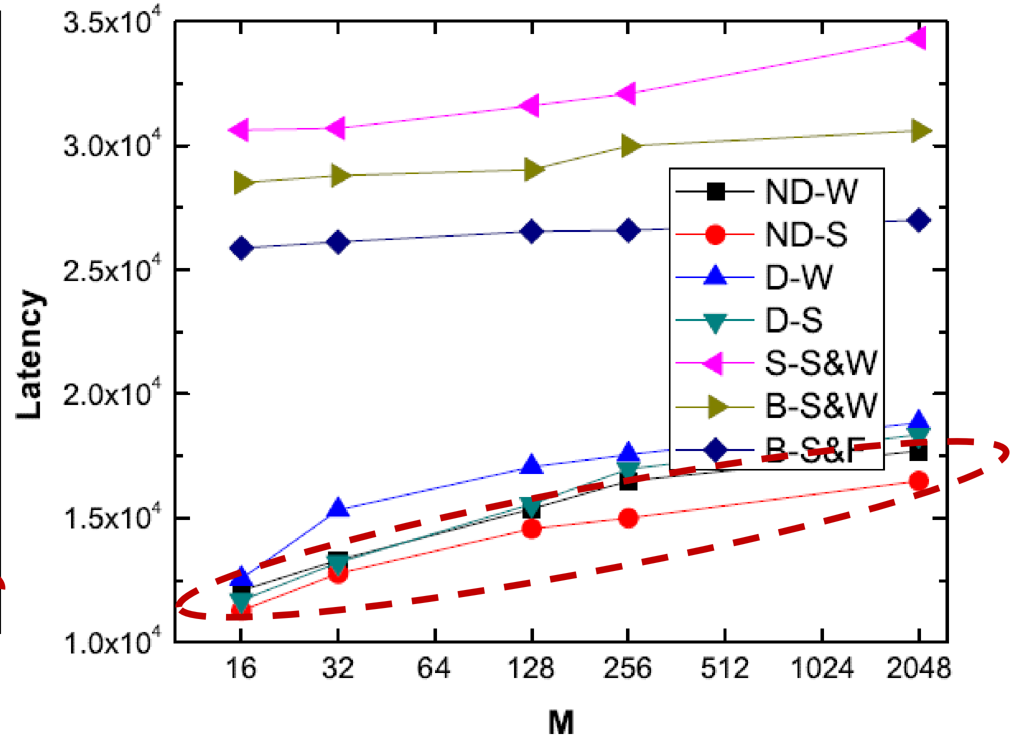


Synthetic trace

Varying node density - latency

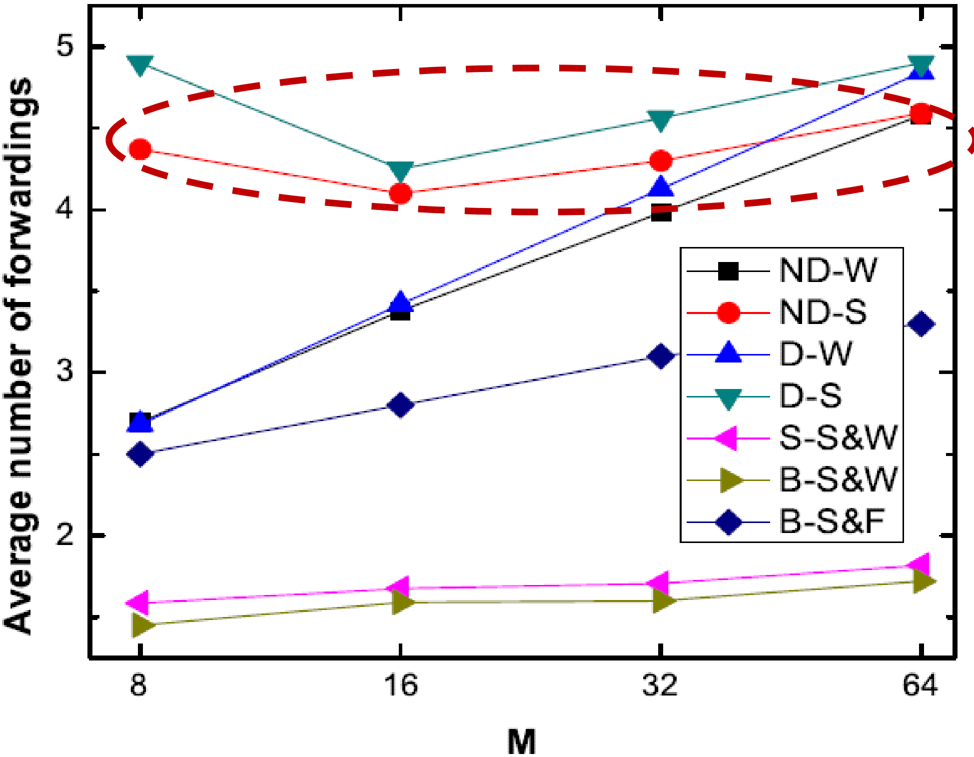


Real trace

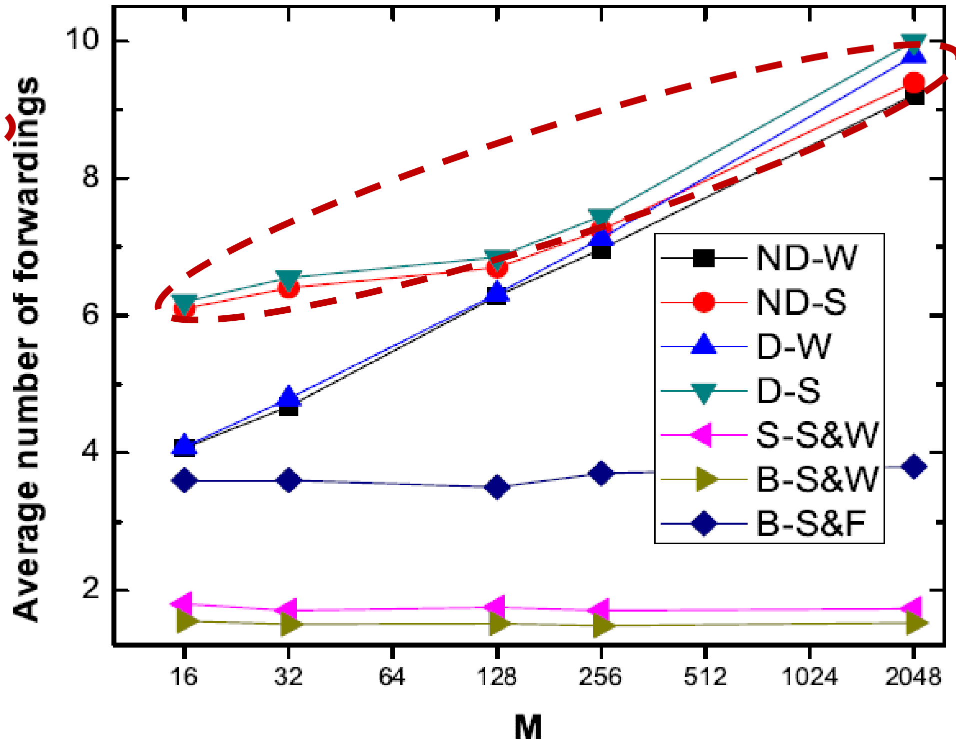


Synthetic trace

Varying node density - # of forwardings



Real trace



Synthetic trace

Recap and Extensions

- Multi-path routing in hypercubes
  - Feature-based, efficient copy management, and node-disjoint-based

- Extensions
  - General hypercubes

    (Wu & Wang 2014, IEEE TC)
  - Analytical results: delivery rate and latency

Social Home-based

(Wu, Xiao, & Huang '13)

- Social characteristics

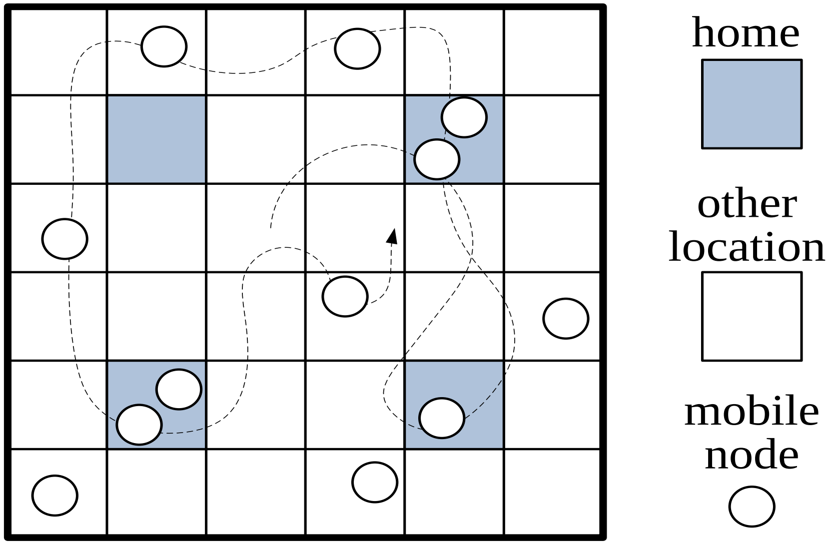
  Nodes visit some locations (homes) frequently, while visiting other locations less frequently

- Real or virtual "throwbox"

  Each community home is equipped with a real or virtual "throwbox" so that it can store and forward messages

Problem

- Network Model
  - n mobile nodes in V
  - h homes: H = {1, 2,…, h} and other locations



Italian student: {Italian town, school, dorm}

Female student: {shopping center, school, dorm}

Solution

- Homing Spread (HS)
  - Homing phase
    - The source sends message copies to homes quickly
  - Spreading phase
    - Homes with multiple copies spread them to other homes and mobile nodes
  - Fetching phase
    - The destination fetches the message when it meets any message holder

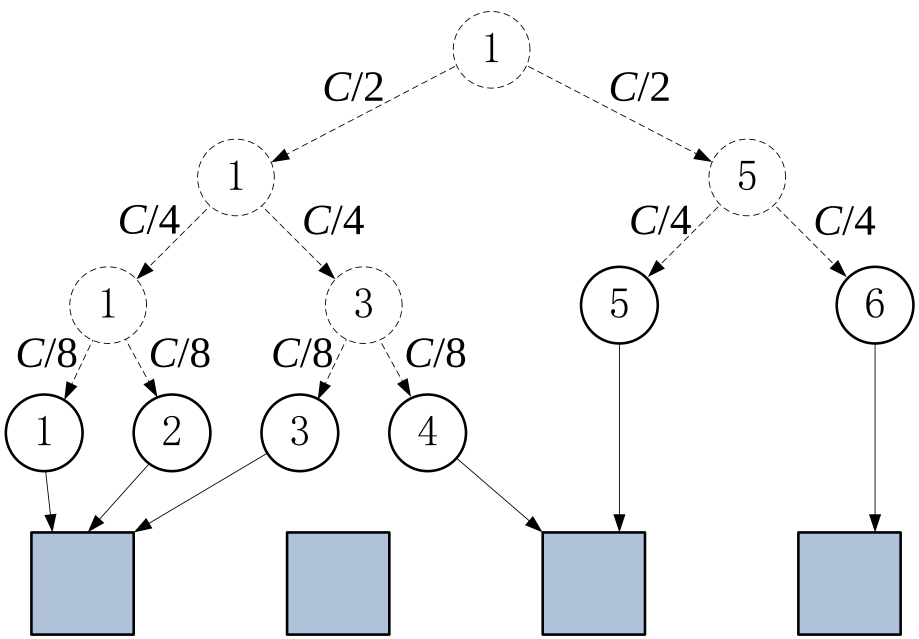
(Homing: meet at a home site. Roaming: at a non-home site.)

Challenges

- Given a fixed number of message copies C
  - What is the optimal way for a message holder to spread copies during homing and roaming?
  - Once a home receives some message copies, how should it further spread these copies?
  - What is a general way for a mobile destination to obtain a copy?

Homing Phase

- Binary Homing Scheme:
  - Each message holder sends all of its copies to the first (visited) home.
  - If the message holder meets another node before it visits a home, it binary splits the copies between them.



Homing Phase

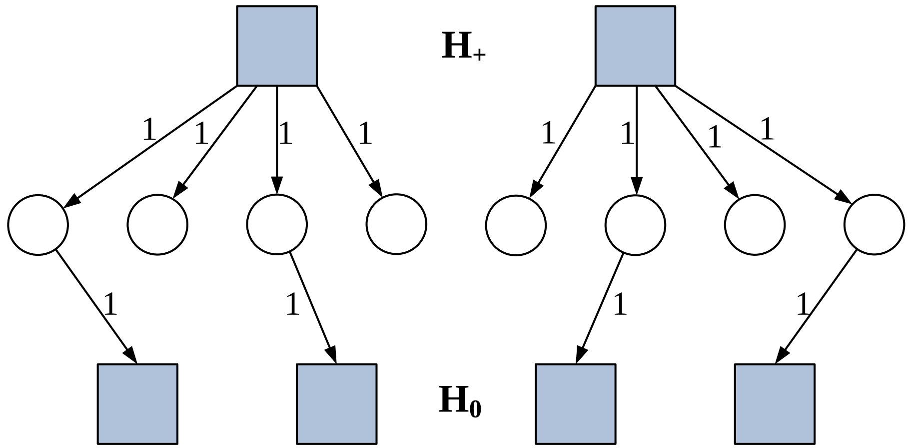
- Assume
  - Inter-meeting time follows the exponential distribution: λ (between two mobile nodes) and Λ (a mobile node to a home)
- Lemma 1
  - The binary homing scheme can spread the C message copies to the maximum number of nodes before they reach the homes
- Analysis
  - The expected delay of each message copy is always $1/h\Lambda$ no matter which splitting scheme is adopted
  - The maximum number of nodes (homes) received the message copies

Spreading Phase

- 1-Spreading Scheme
  - Each home with more than one message copy spreads a copy to each visiting node until only one copy remains
  - If a node with one copy later goes and visits a new home, the node sends the copy to that home
- Analysis
  - Each home has at most one copy
  - If $C > h$, there are $C - h$ nodes outside the homes that have a copy
  - Home is always more important than a regular node as it can spread the message faster

Spreading Phase

- 1-Spreading Scheme

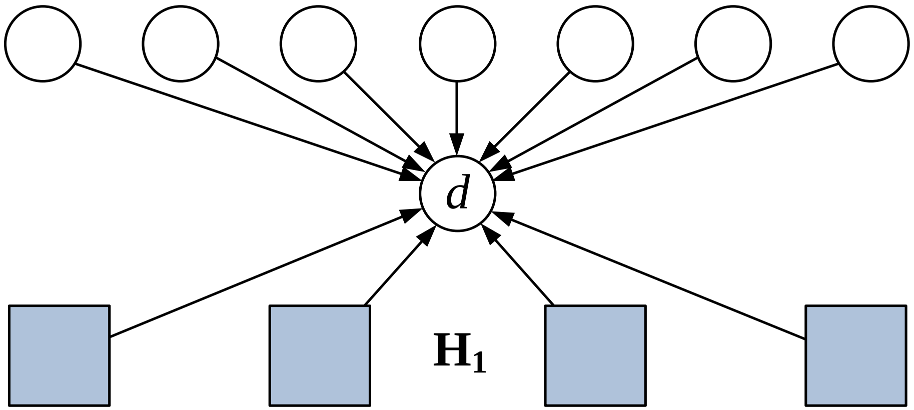


(H+: multiple copies, H_1: one copy, H_0: no copy)

- Lemma 2
  - The 1-spreading scheme can spread copies from a home to the maximum number of nodes with the fastest speed.

Fetching Phase

- Fetching Scheme
  - The destination fetches the message once it meets a message holder



Distributed HS Algorithm

Algorithm 1 The Homing Spread (HS) algorithm

for each mobile node i **do**
    if node i encounters another node j **then**
        if node j is the destination **then**
            node i sends the message to j;
        if nodes i and j have r_i and r_j message copies **then**
            node i holds $\lceil r_i/2 \rceil + \lfloor r_j/2 \rfloor$ copies through exchange with node j;
    if node i visits a home h **then**
        node i sends all its copies to h;
        if $h \in H_+$ or i is the destination **then**
            h sends a copy to node i.

Continuous Markov Chain

- Expected delivery delay
- Number of copies needed for a given delay bound

Network State

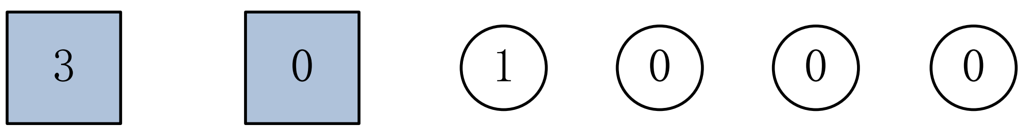
- State s is a vector with h+n components

$$s = \langle s_1, s_2, \ldots, s_h, s_{h+1}, \ldots, s_{h+n} \rangle,$$

where s_i: copies held by i-th home (if i≤h) or (i–h)-th node (if i>h)

and $$\sum_{i=1}^{h+n} s_i = C \ (s_1 \geq s_2 \geq \Lambda \geq s_h; s_{h+1} \geq s_{h+2} \geq \Lambda \geq s_{h+n})$$

- For example: $s = \langle 3, 0, 1, 0, 0, 0 \rangle$



- Start state: $s_t = \langle 0, 0, \ldots, 0, C, 0, \ldots, 0 \rangle$
- Optimal state: $s_o = \langle 1, 1, \ldots, 1, 0, \ldots, 0 \rangle$

Optimality of HS

- HS follows binary homing and 1-spreading schemes
- Lemmas 1 and 2 show that the HS is the fastest way to turn a network state into the optimal state
- Each state transition based on binary homing and 1-spreading schemes can turn the current state into the best next state that has the minimum expected delivery delay

Expected Delivery Delay

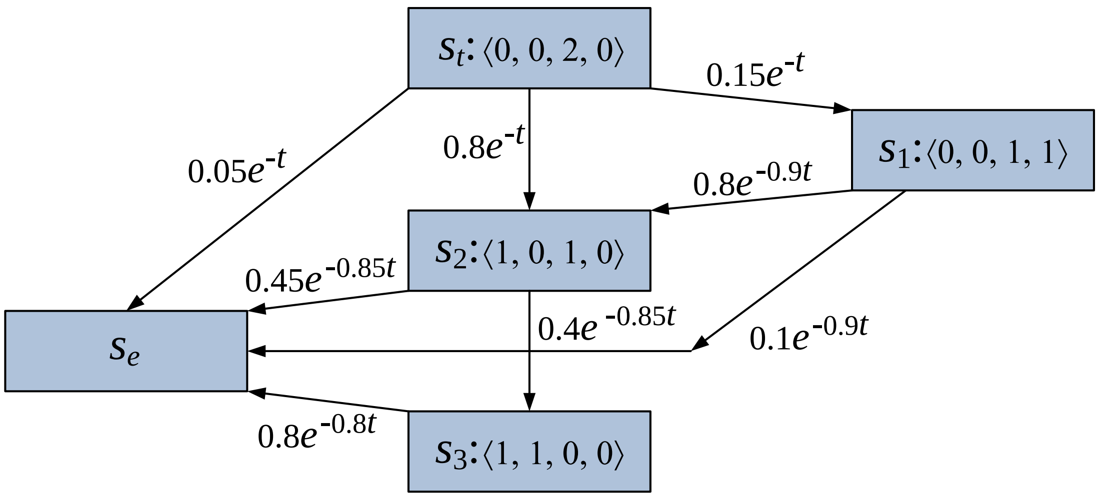
Continuous Markov Chain

- State space $s = \langle s_1, s_2, \ldots, s_h, s_{h+1}, \ldots, s_{h+n} \rangle$
- State transition graph
  - Binary homing scheme and 1-spreading scheme
  - A directed acyclic graph
  - State transition function $\rho_{s,s'}(t)$

    Probability density function about the time t that it takes for state transition from s to s'

Expected Delivery Delay

- Derive the cumulative probability density function for the state transition from start to end state
- Calculate the expected delivery delay



exponential density function

$$p(t) = \lambda e^{-\lambda t}$$

h=2, n=5, C=2, λ = 0.005, and Λ = 0.4

Three Phases

Homing phase: average delay for the first copy reaching a home

$$D^{(1)} = \frac{1}{h\Lambda}$$

Spreading phase: average delay for each home to receive a copy

$$D^{(2)} \leq \frac{2}{\Lambda}$$

Fetching phase: average delay after the first two phases in message fetching

$$D^{(3)} = \begin{cases} \frac{1}{C\Lambda} & , C \leq h \\ \frac{1}{h\Lambda + (C-h)\lambda} & , C > h \end{cases}$$

Upper Bound

- Define

$$D' = \frac{1}{h\Lambda} + \frac{2}{\Lambda}$$

homing (first term) plus spreading (second term) delay

- The expected delay D of HS satisfies:

$$D \le \begin{cases} D' + \frac{1}{C\Lambda} & , \ C \le h \\ D' + \frac{1}{h\Lambda + (C-h)\lambda} & , \ C > \mathrm{h} \end{cases}$$

- *C* copies needed for a given delay bound $\Theta(\ge D')$

$$C = \begin{cases} \frac{1}{\Lambda(\Theta - D')} & , \quad \Theta \ge \frac{1}{h\Lambda} + D' \\ \frac{1}{\lambda} \cdot \left(\frac{1}{\Theta - D'} - h\Lambda\right) + h & , \quad D' < \Theta < \frac{1}{h\Lambda} + D' \end{cases}$$

Simulation

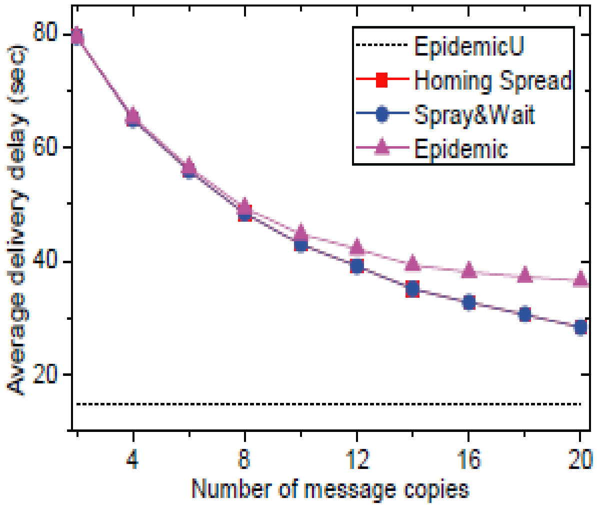
- Trace
  - Synthetic trace: Time-Variant Community Model

- Settings

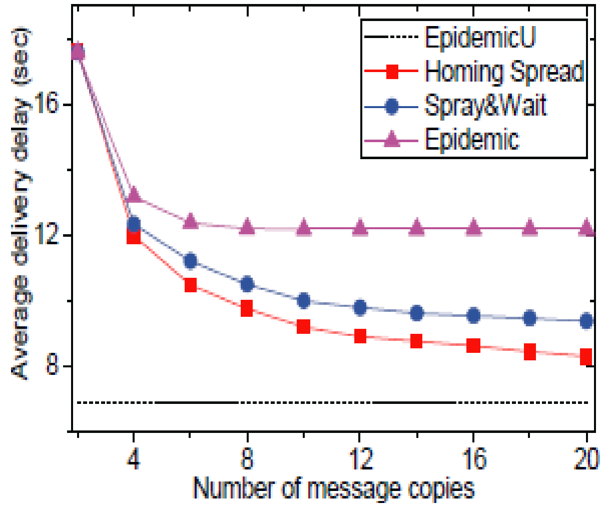
Parameter name	Default value	Range
Deployment area	20×20	-
Number of nodes n	200	100-400
Number of homes h	5	0-15
Homing probability per sec Λ	0.04	0.04-0.16
Number of messages	10,000	-
Allowed message copies C	10	2-20

Average Delay vs. Number of Homes

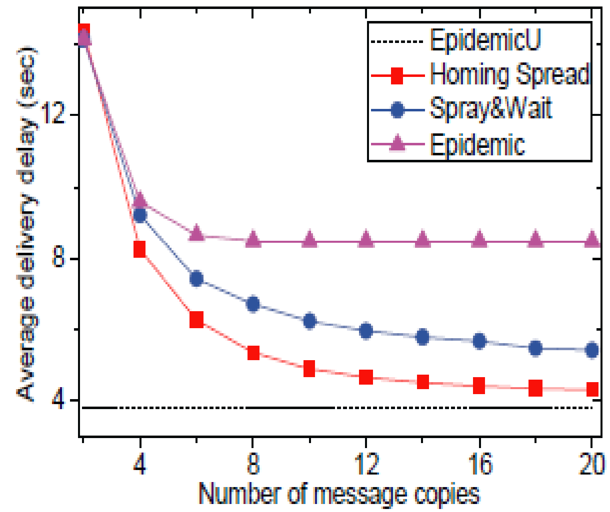
- Epidemic (*C* copies)
- EpidemicU (unlimited copies)



(a) Number of homes: $h = 0$

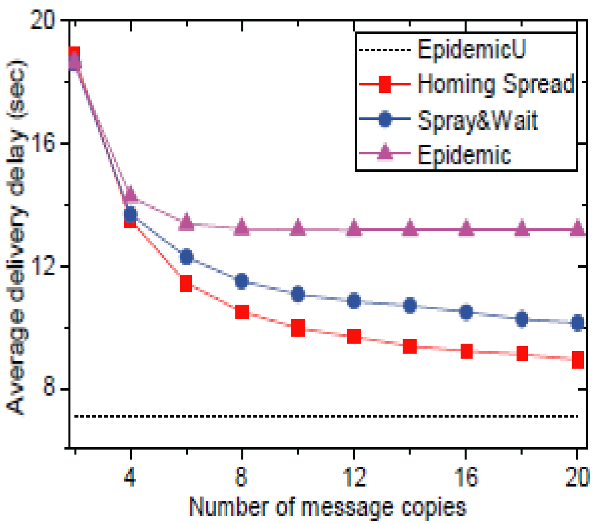


(b) Number of homes: $h = 5$

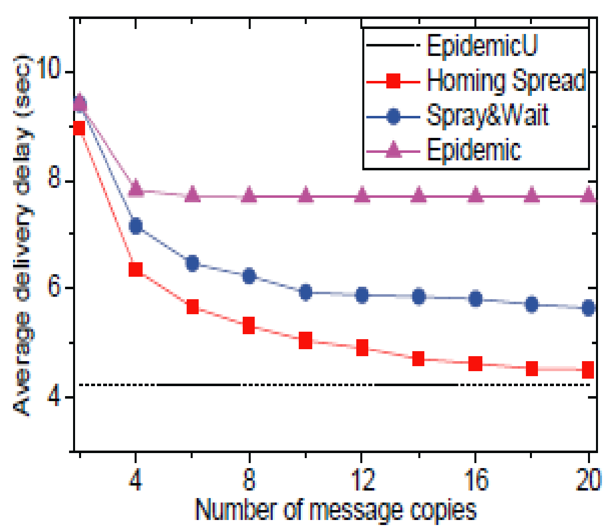


(d) Number of homes: $h = 15$

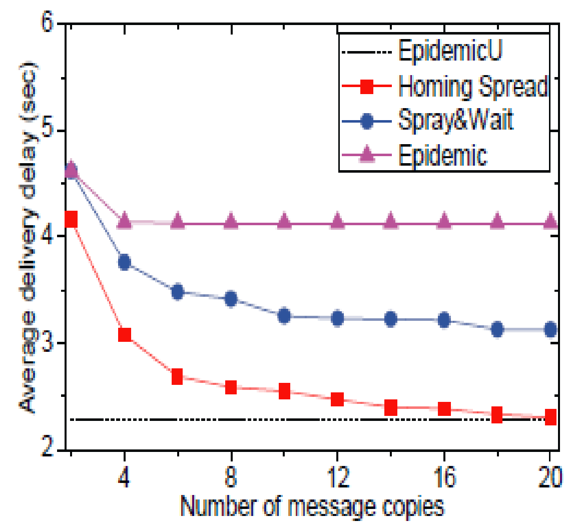
Average Delay vs. Homing Probability



(a) Homing probability: $\Lambda = 0.04$

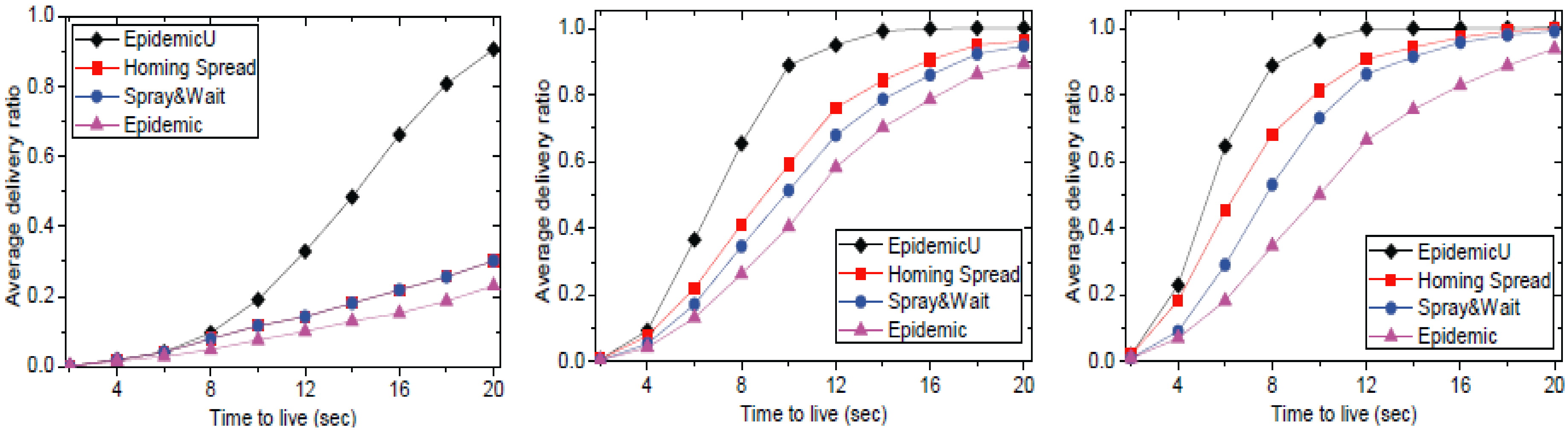


(b) Homing probability: $\Lambda = 0.08$



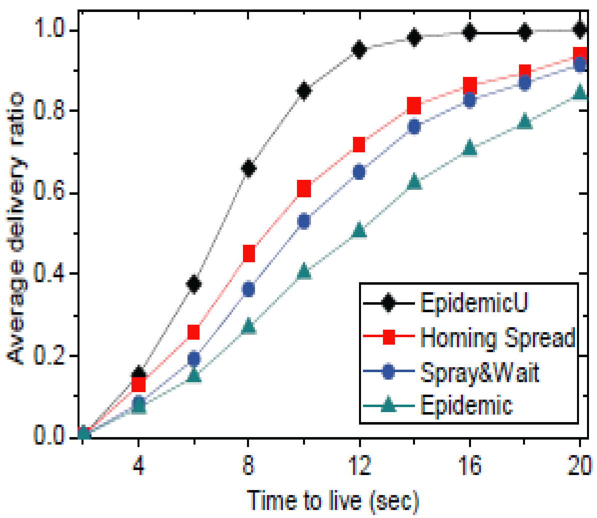
(d) Homing probability: $\Lambda = 0.16$

Average Delivery Ratio vs. Number of Homes

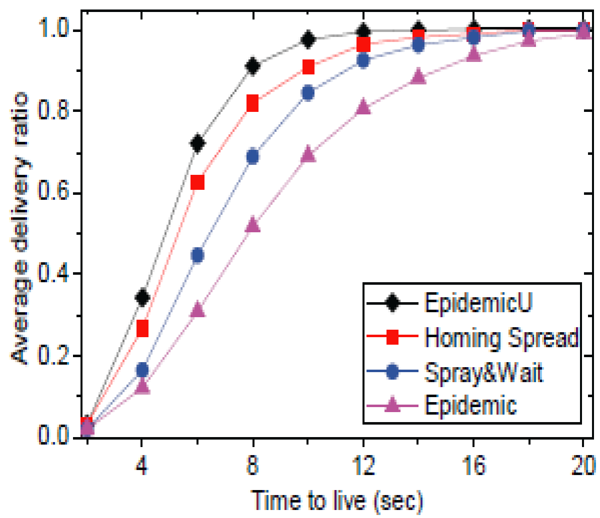


(a) Number of homes: $h = 0$ (b) Number of homes: $h = 5$ (c) Number of homes: $h = 10$

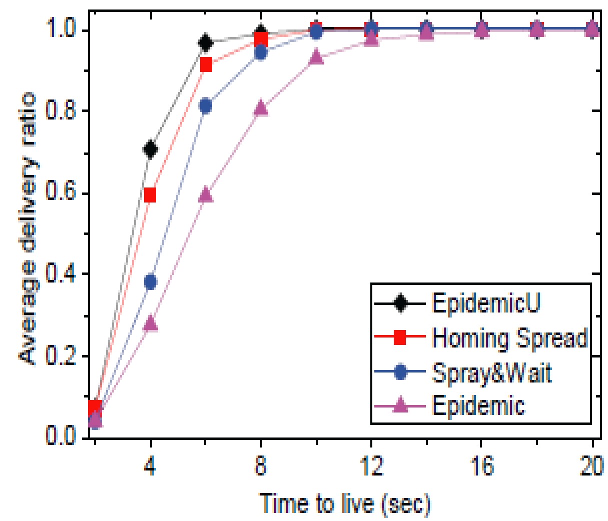
Average Delivery Ratio vs. Homing Probability



(a) Homing probability: $\Lambda = 0.04$



(b) Homing probability: $\Lambda = 0.08$



(c) Homing probability: $\Lambda = 0.12$

Recap and Extensions

- Homing-spread
  - Optimal multi-copy routing in three phases: homing, spreading, and fetching
- Extensions
  - Heterogeneous homes: different homes
    (Xiao, Wu & Huang '14, IEE TPDS)
  - Heterogeneous visiting: different visit rates

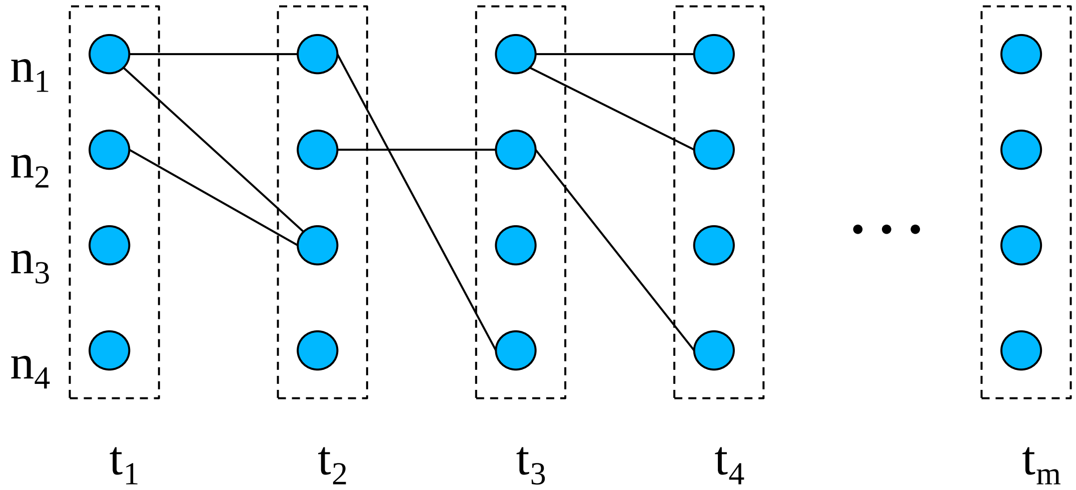
Big Picture: Network Science

- How the Overall Network System will Behave
  - Static network structure
  - Dynamic networks structure (more challenging)

- Epidemic and Other Spreading Processes
  - Epidemic is well studied in static networks
    - Susceptible-Infected (SI), SIR (recovered), SIS, and SIRS
  - Epidemic, especially controlled one, in dynamic networks
    - More challenging and not well understood

Evolving Graphs

- Time-space view



- Social-aware contacts
  - A special temporal-spatial link summary
- Activity level (AL)
  - AL of a node evolves based on ALs of neighbors

Collaborators

- Yunsheng Wang

  Kettering Univ.

- Mingjun Xiao

  Univ. of Sci. & Tech. of China