Opportunistic Routing Algorithms in Delay Tolerant Networks

Eyuphan Bulut

Rensselaer Polytechnic Institute
Department of Computer Science and Network Science and Technology (NeST) Center

PhD Thesis Defense
Feb 4th, 2011

Outline

- Introduction to DTNs
  - Challenges of routing
- Proposed Algorithms
  1) Multi-period Spray and Wait routing
  2) Multi-period erasure coding based routing
  3) Efficient single-copy routing utilizing correlation between node meetings
  4) Social relation based routing

  Simulation Results Based on Real and Synthetic Traces
- Summary of Contributions

Outline

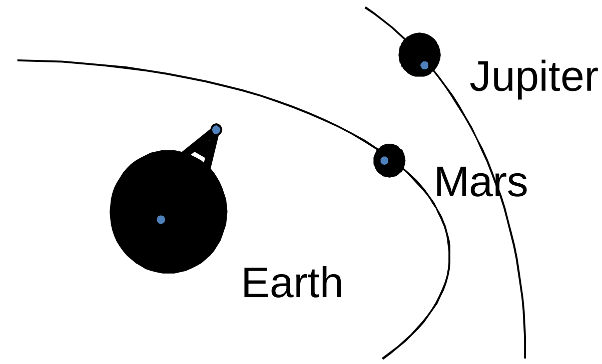
- **Introduction to DTNs**
  - **Challenges of routing**
- Proposed Algorithms
  1) Multi-period Spray and Wait routing
  2) Multi-period erasure coding based routing
  3) Efficient single-copy routing utilizing correlation between node meetings
  4) Social relation based routing
- Summary of Contributions

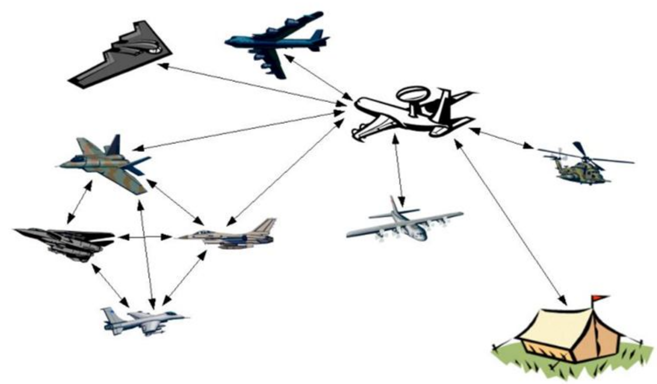
Delay Tolerant Networks

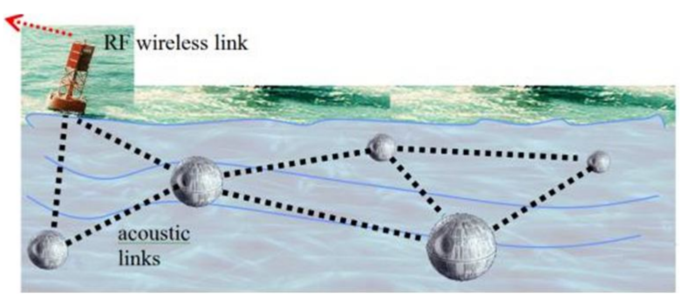
- Intermittently connected mobile networks
  - Sparse mobile networks
    - Main difference from MANETs
  - Lack of continuous end-to-end connectivity
  - Utilizes “store-carry-and-forward” paradigm in routing

Applications of DTNs

- Space networks
  - Satellites and planets
- Military Networks
  - Soldiers, aircrafts
- Social Networks
  - People, base stations
- Vehicular Networks
- Underwater networks

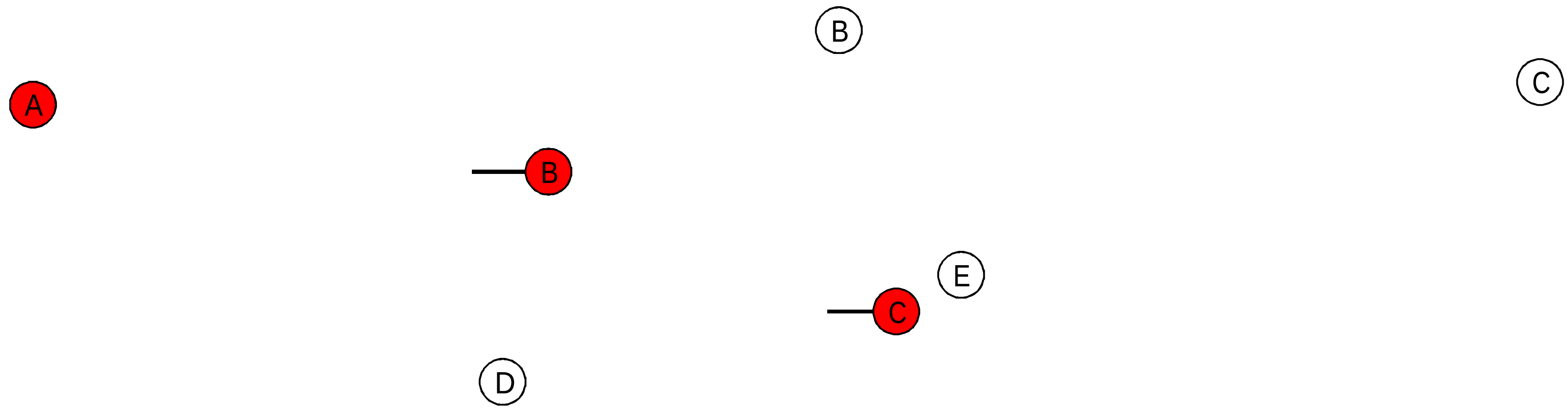




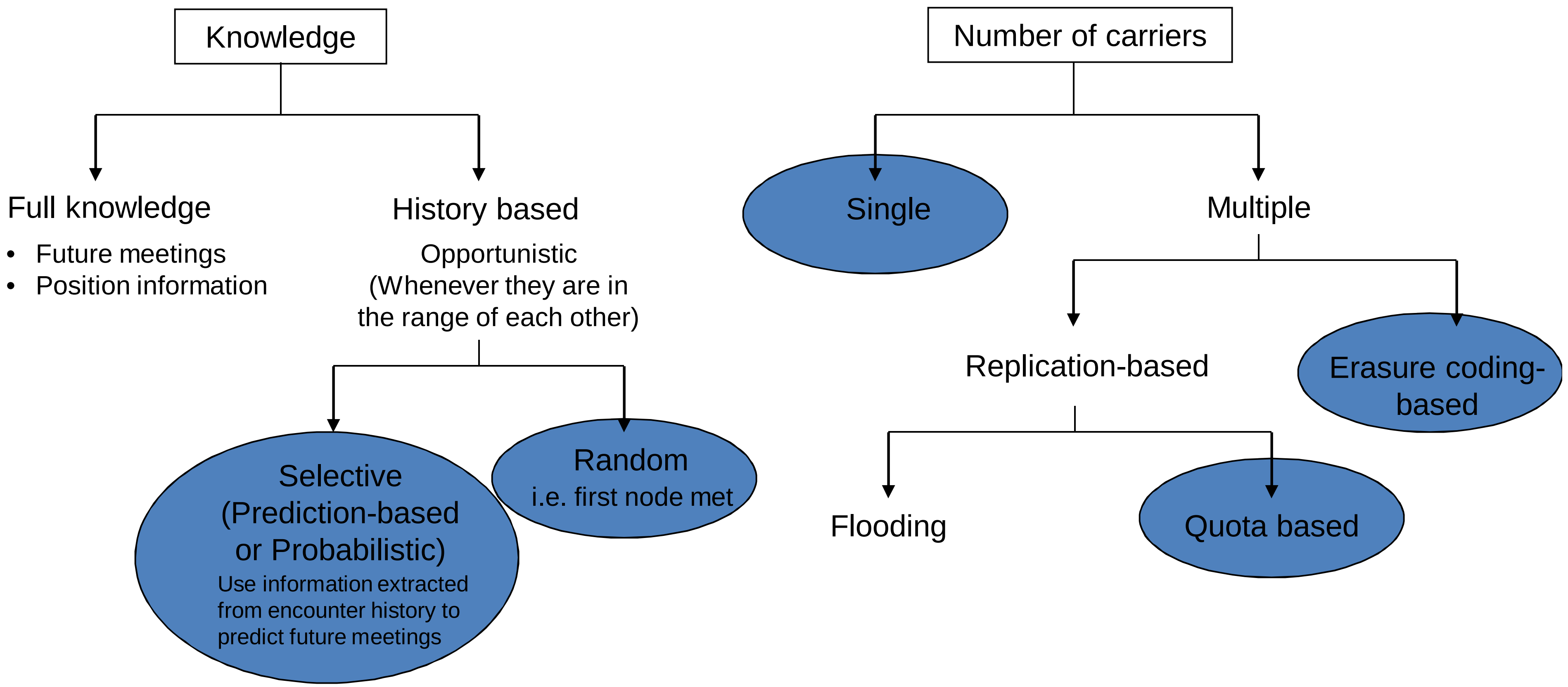


Routing in DTNs

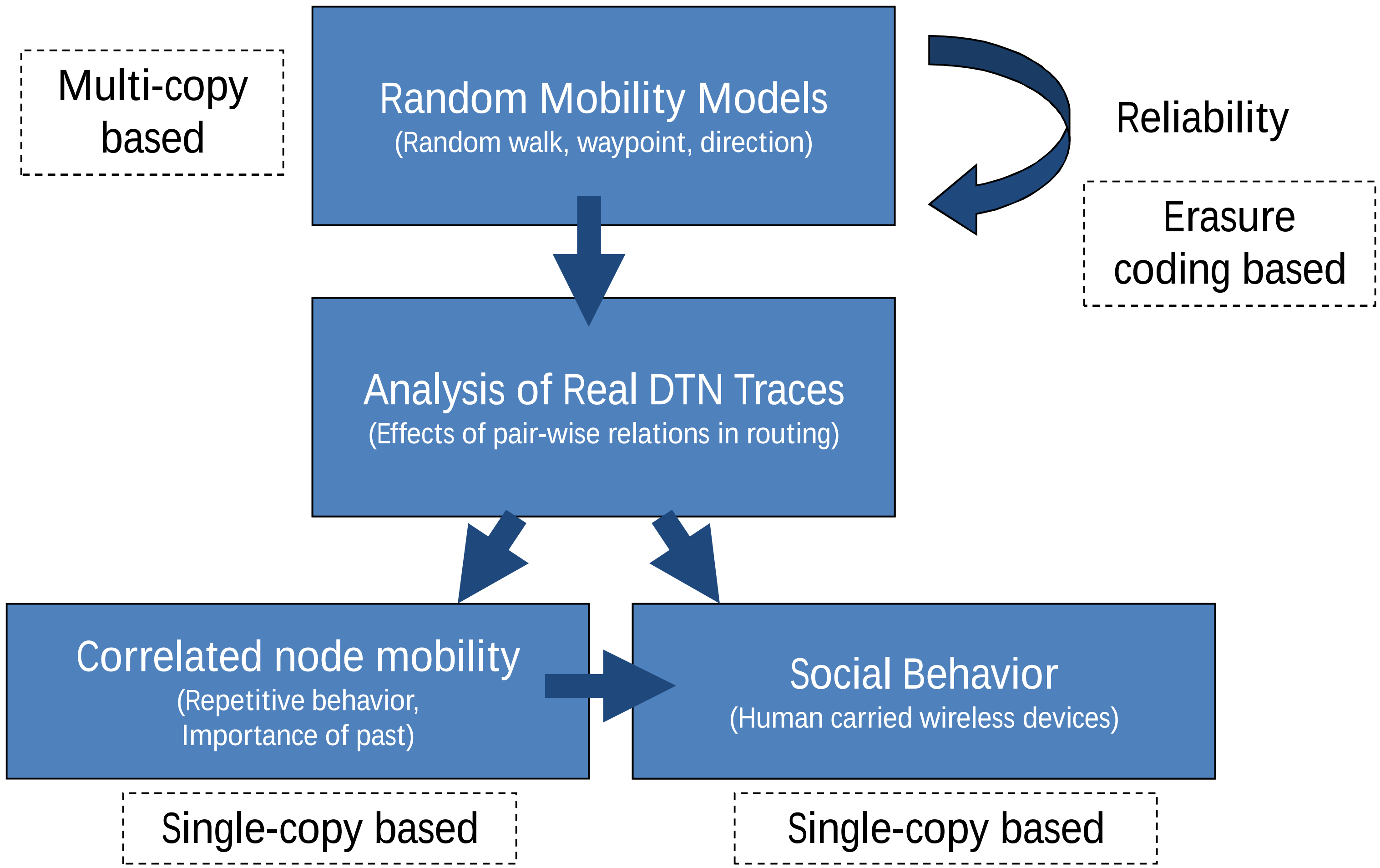
- Challenges:
  - Dynamic and sparse topology
  - Low probability of end-to-end connectivity
  - How to locate destination with local knowledge?
  - Opportunistic message exchanges
    - When nodes come to the range of each other
    - How to decide whom to forward/copy a message?



Routing in DTNs



Our Research Path



Outline

- Introduction to DTNs
  - Challenges of routing
- **Proposed Algorithms**

  1) Multi-period Spray and Wait routing

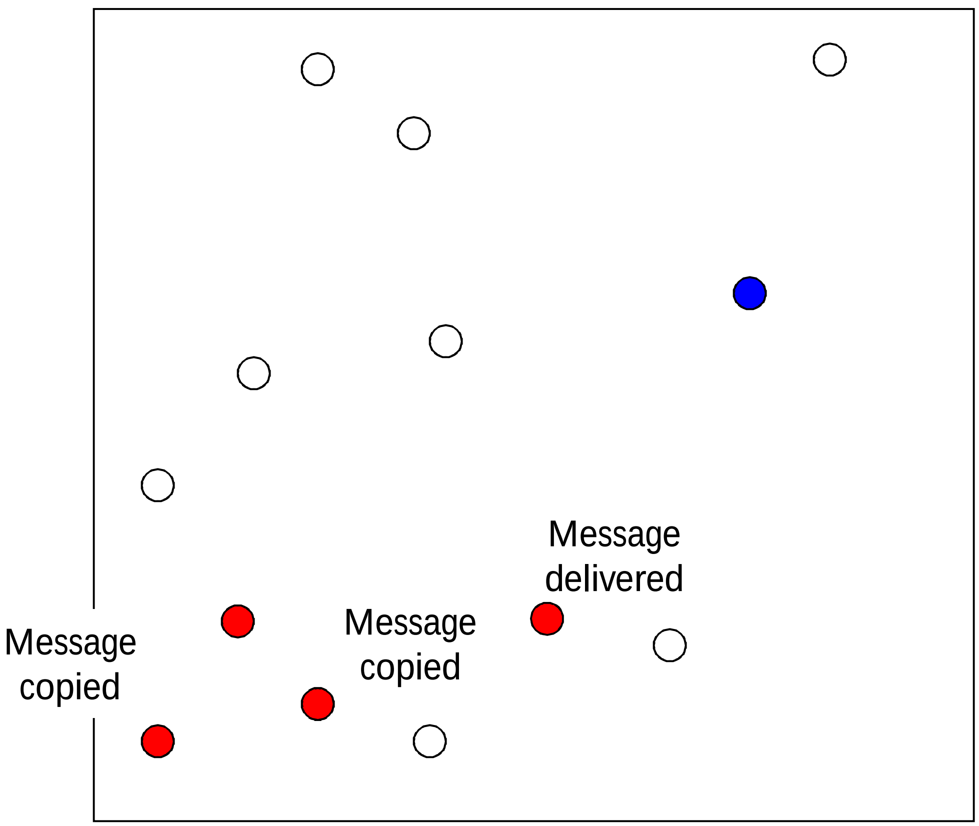
  2) Multi-period erasure coding based routing

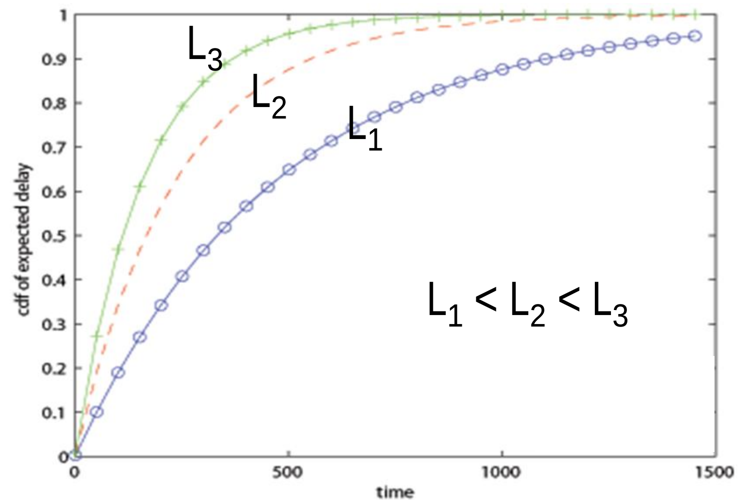
  3) Efficient single-copy routing utilizing correlation between node meetings

  4) Social relation based routing
- Summary of Contributions

Spray and Wait*

- Random mobility model
  - Exp. dist. intermeeting times





Destination

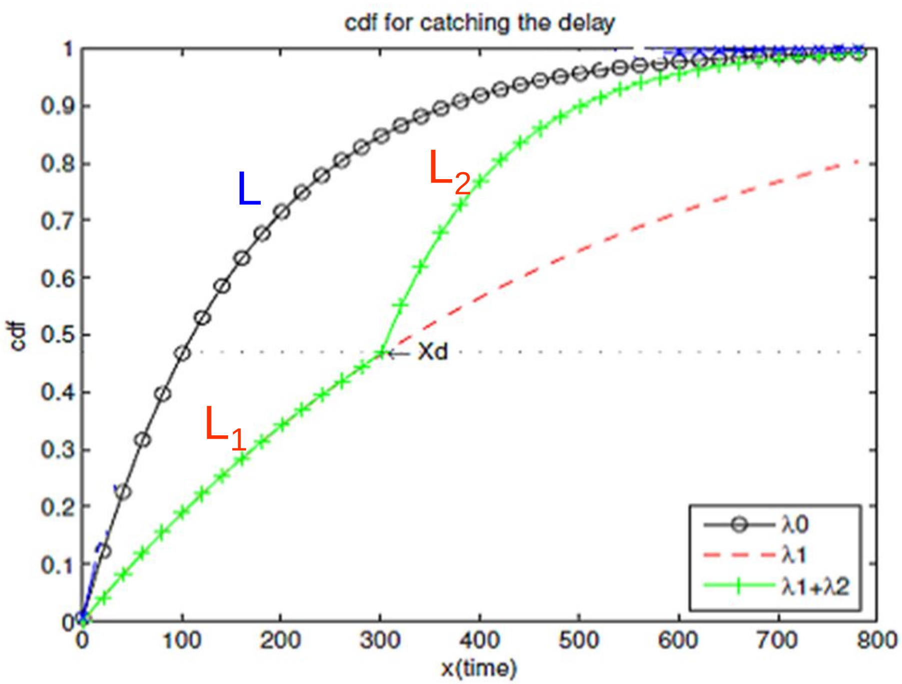
Node with message copy

Node without message copy

*Spyropoulos et. al.
Transactions on Networking, 08

Two Period Spray & Wait

- Spray L_1 copies at the beginning
- Spray additional L_2 - L_1 copies at time x_d (start of second period)



GOALS:

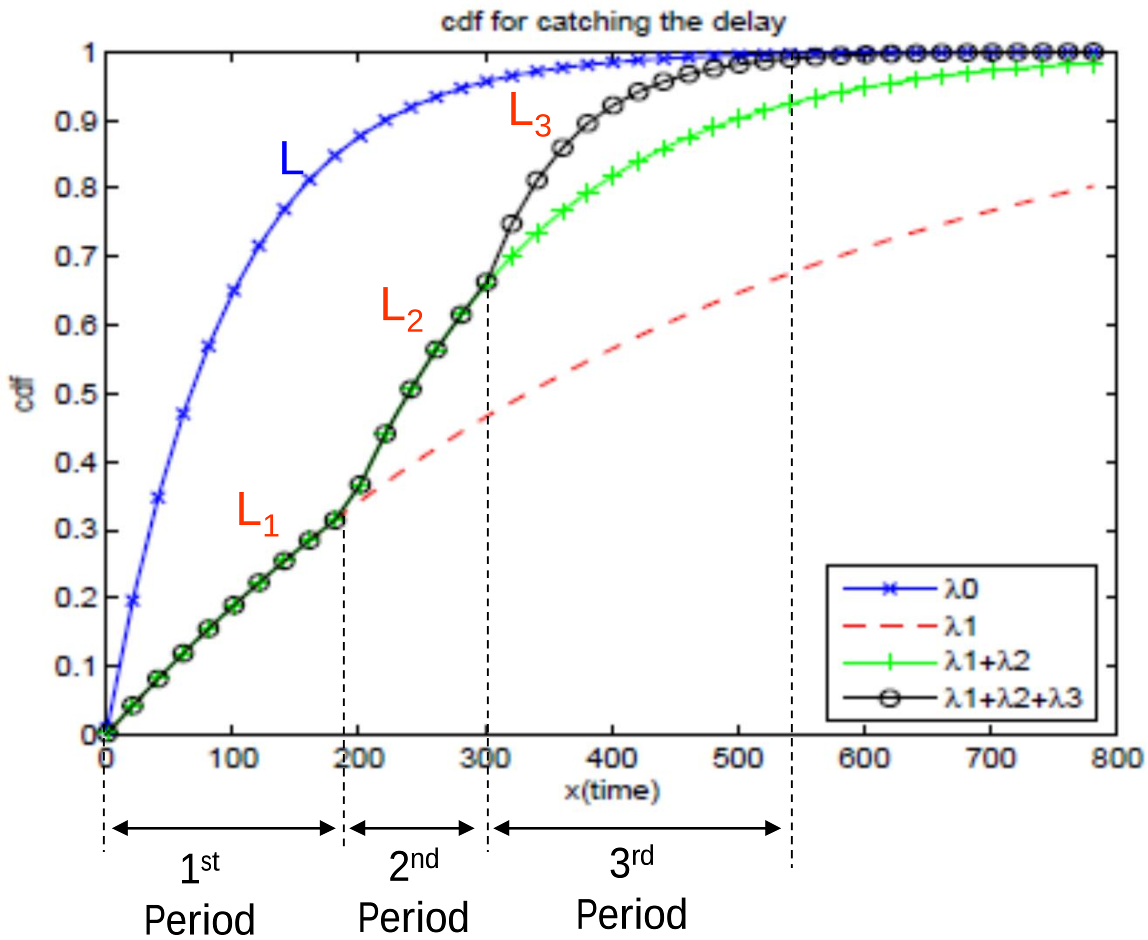
1) Maintain the same delivery rate by deadline (td)

2) Lower the average cost

$$L_1(P)+ L_2(1-P) < L$$

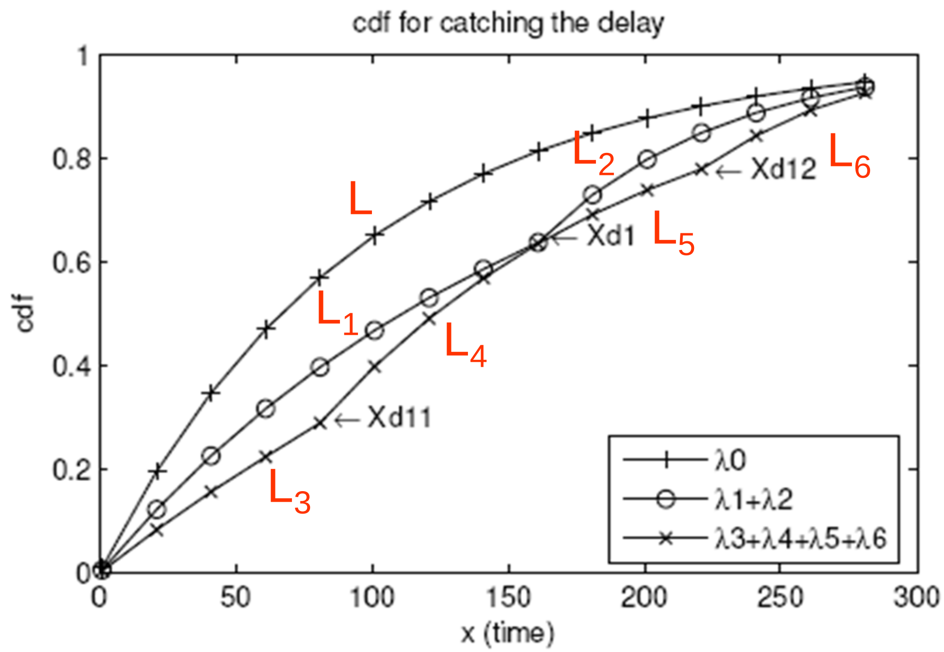
Delivery probability in first period

Three Period Spray & Wait



Multiple Period Spray & Wait

- If we currently have k spray and wait periods, to obtain k+1 periods:
  - Partition each period into two sub-periods optimally
  - Take the one which makes the overall cost minimum



2 periods: L_1 and L_2

3 periods:

Select either

a) L_1, L_5 and L_6

b) L_3, L_4 and L_2

Acknowledgment of delivery

- Two types:
  - Type I: Acknowledgment by flooding
    - Pros: Acks are small, lower cost
    - Cons: Takes time to reach all nodes, thus extra copying may occur
  - Type II: Single broadcast with powerful radio
    - Pros: Immediate acknowledgment
    - Cons: Cost of powerful radio

Optimum L_i's from Analysis

TABLE I

OPTIMUM L_i COPY COUNTS THAT MINIMIZE THE AVERAGE NUMBER OF COPIES WHILE PRESERVING THE DESIRED PROBABILITY OF DELIVERY.

	Random Walk			Random Waypoint		
t_d	L_{min} in 1p	2p Optimum L_i's	3p Optimum L_i's	L_{min} in 1p	2p Optimum L_i's	3p Optimum L_i's
200	12	7,22	6,12,27	9	5,16	4,8,20
250	9	[illegible]	5,9,19	7	4,13	3,6,15
300	8	5,14	4,8,18	6	3,11	3,6,14
400	6	4,11	3,6,14	[illegible]	3,10	2,4,12
500	5	3,9	2,4,11	4	2,8	2,4,10
600	4	[illegible]	2,4,9	3	2,5	1,2,6
700	4	2,8	2,4,10	3	2,6	1,2,7
800	3	2,5	1,2,6	2	2,7	1,3,10
900	3	2,6	1,2,7	2	1,4	1,2,5

Cost=5.87

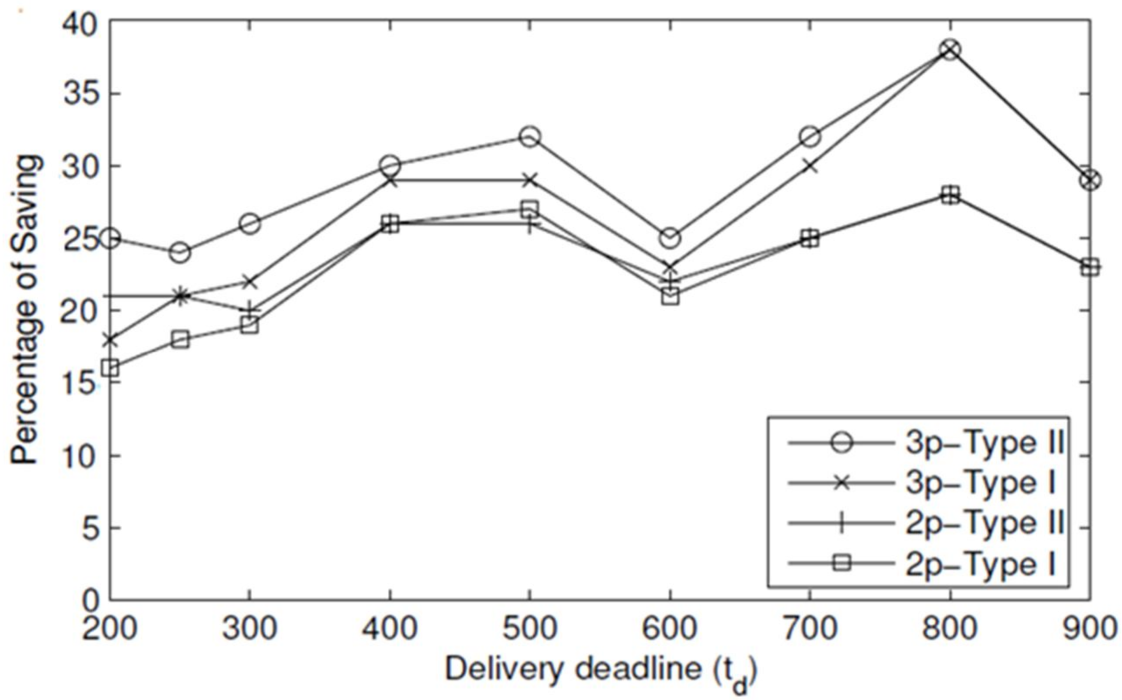
Cost=4.64

X_d=285s

Cost=4.28

Simulation Results

- Percentage of Saving
  - While achieving the same delivery ratio by deadline



Extensive Simulations

- Theoretical results are matching with simulation results
- Effect of different number of nodes, different desired delivery ratios etc.
- Results on Real Traces
  - Demonstrates benefit, but still needs careful analysis due to heterogeneous meeting behavior of different nodes

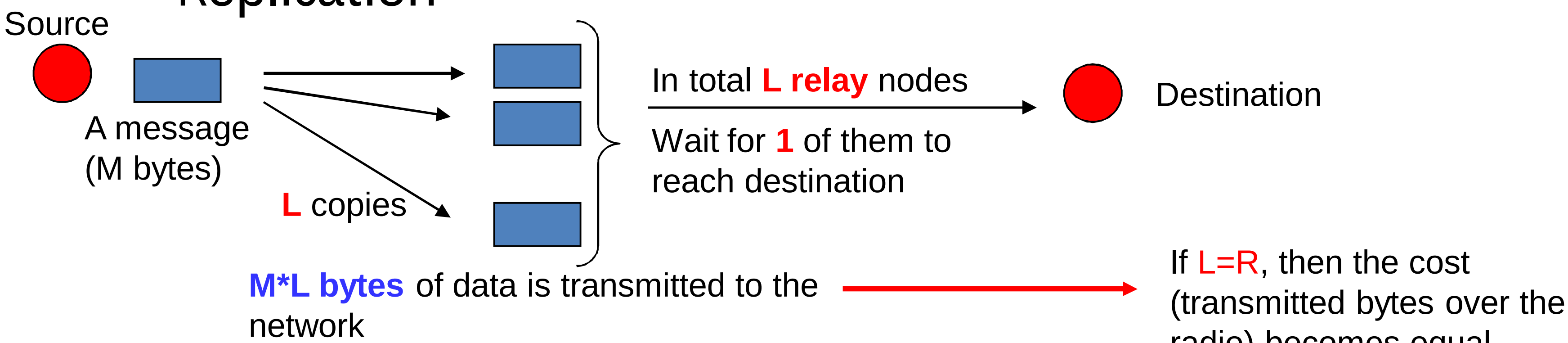
[IEEE/ACM Transactions on Networking'10],
[Globecom'08], [ACITA'08]

Outline

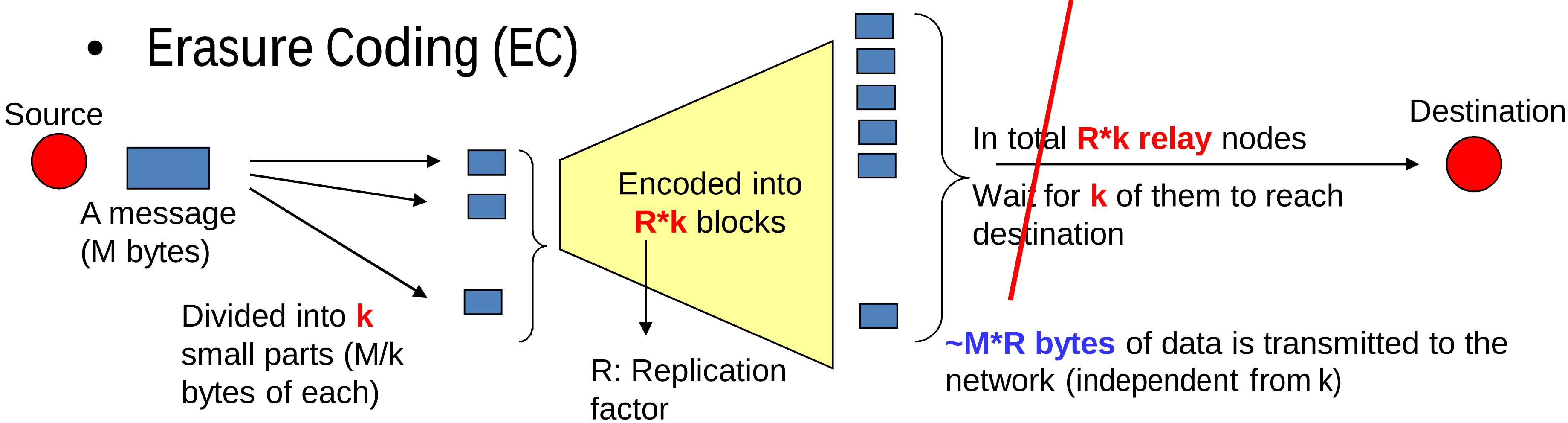
- Introduction to DTNs
  - Challenges of routing
- Proposed Algorithms
  1) Multi-period Spray and Wait routing
  2) **Multi-period erasure coding based routing**
  3) Efficient single-copy routing utilizing correlation between node meetings
  4) Social relation based routing
- Summary of Contributions

Replication vs. Erasure Coding

- Replication



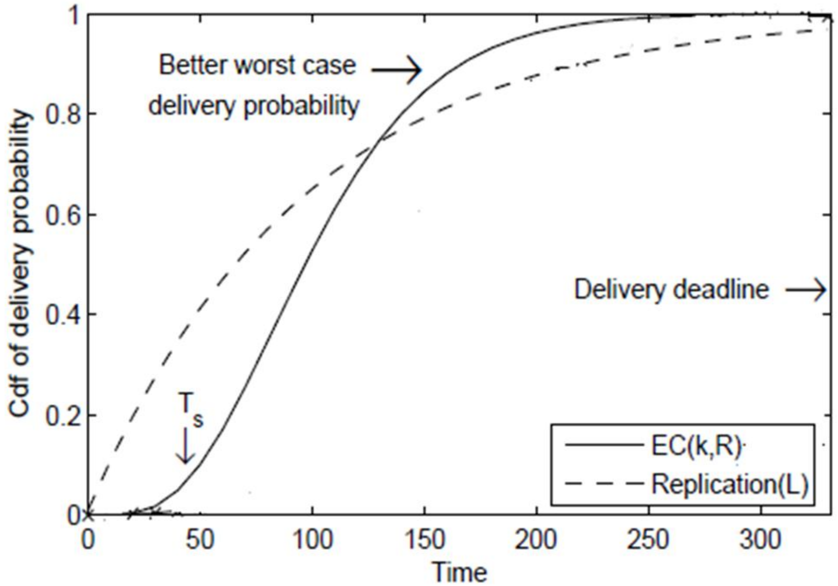
- Erasure Coding (EC)



Replication vs. Erasure Coding

- Which one is better?
  1. Spraying L messages and waiting for 1 (to reach destination)?
     - Spraying duration takes less time than the second one
  2. Spraying Φ=R*k messages and waiting for k?

Costs are the same when R=L



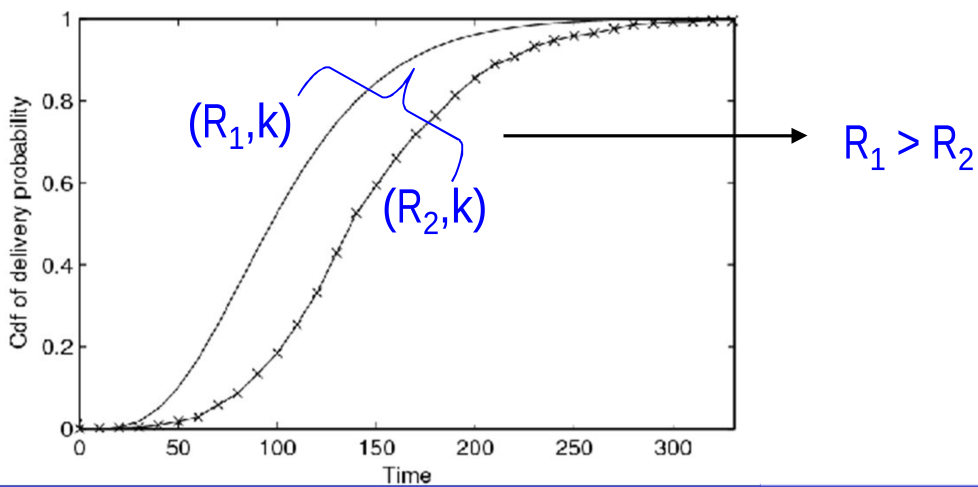
“EC” also provides more reliable routing:

In a failure of one packet, the performance of “replication” routing is affected more than the performance of “erasure coding” based routing.

Replication vs. Erasure Coding

- If desired delivery rate is higher, we can achieve more cost saving with erasure coding based routing compared to replication routing.

- Optimum Single-period erasure coding based routing:
  - Try to minimize Replication factor (R) since cost is proportional to R.
  - Maintain delivery rate by deadline

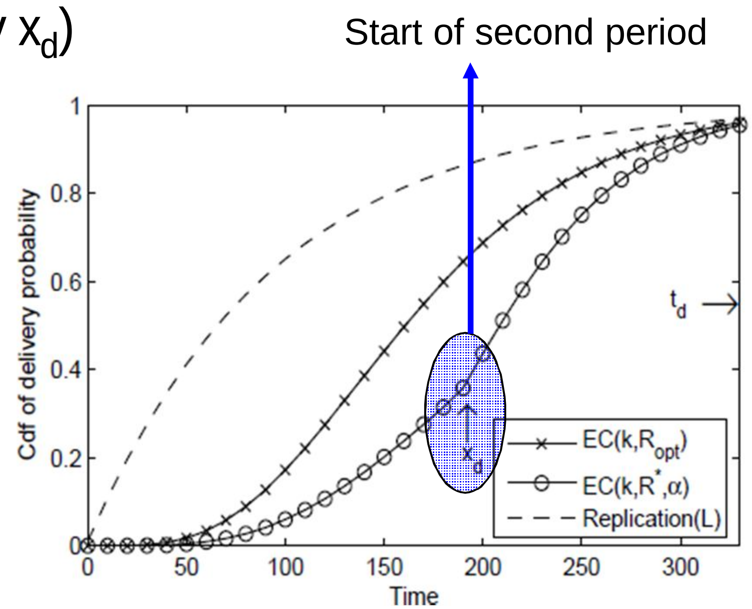


Multi-period Erasure Coding-based Routing

- In 1st period:
  - Create Φ_2=kR* coded blocks, where $R^* > R_{opt}$ (optimum R in single period)
    - Linear time complexity of creating these packets (Tornado codes)
  - Spray $\Phi_1 = \alpha k R^*$ of them and try delivery with them

- If delivery doesn't happen in 1st period (by x_d)
  - Spray remaining Φ_2-Φ_1 of them in 2nd period

- Same goals:
  - Maintain delivery rate by deadline
  - Achieve lower cost on average



Simulations

- Message size 100Kb
- Costs at the delivery (Type II) and after all nodes are acknowledged (Type I):

1 period — 2 periods

t_d	Cost of EC-1p-SS (Kbytes)			Cost of EC-2p-SS(Kbytes)		
sec	Opt(R,k)	Type I	Type II	Opt(R^*,α, x_d)	Type I	Type II
600	(3,2)	343	342	(5, 0.4, 410)	323	319
500	(3,3)	357	356	(5, 0.4, 345)	299	295
400	(4,3)	445	443	(6, 0.5, 270)	412	398
300	(5,3)	[illegible]	[illegible]	(7, 0.5, 200)	[illegible]	[illegible]
250	(5,5)	[illegible]	[illegible]	(8, 0.5, 185)	[illegible]	[illegible]

1 period Replication Based Cost=587

1 period Replication Based Cost=578

2 period Replication Based Cost=478

2 period Replication Based Cost=464

[illegible] II

MINIMUM AVERAGE COSTS OF SINGLE AND TWO PERIOD ERASURE CODING ALGORITHMS.

[ICC'10]

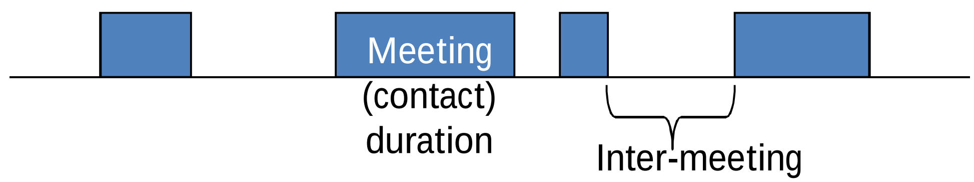
Lower cost than Replication based Routing

Outline

- Introduction to DTNs
  - Challenges of routing
- Proposed Algorithms
  1) Multi-period Spray and Wait routing
  2) Multi-period erasure coding based routing
  3) **Efficient single-copy routing utilizing correlation between node meetings**
  4) Social relation based routing
- Summary of Contributions

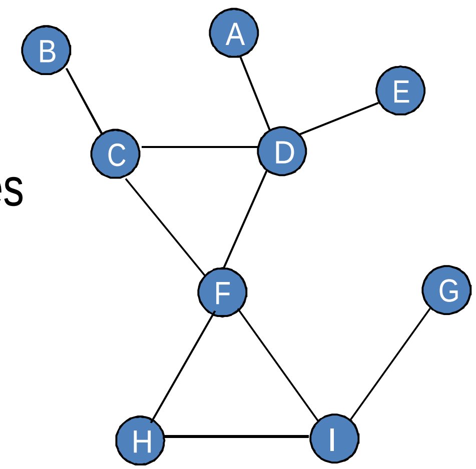
Analysis of Real DTN Traces

- Haggle Project (people, conference, Imote)
- MIT Reality Project (campus, phone)
- UMass Diesel-Net Project (bus meetings)
- RollerNet Traces (roller skate tour, Imote)
- Others:
  - Zebra, taxi etc.



- Extracted information:
  - Pair-wise and aggregate inter-meeting and contact duration

Single-copy DTN Routing



- Shortest-path based routing
  - DTN Graph Model:
    - Vertices are nodes, edges are links between nodes
    - Weights of edges are average inter-meeting times
  - Ex: MED, MEED etc.
- Metric-based (utility) routing
  - When two nodes meet, one forwards its message to other if the other's metric suggests more delivery chance with destination
  - Ex: Prophet, Fresh, SimBet etc.

Motivation

- Conclusions from Analysis of Real-traces (human-based):
  - Pair-wise intermeeting times follow log-normal distribution
    - NOT memory-less (as opposed to exp. dist.)
  - Non-deterministic but cyclic mobility is frequent
    - Periodic meetings of same nodes

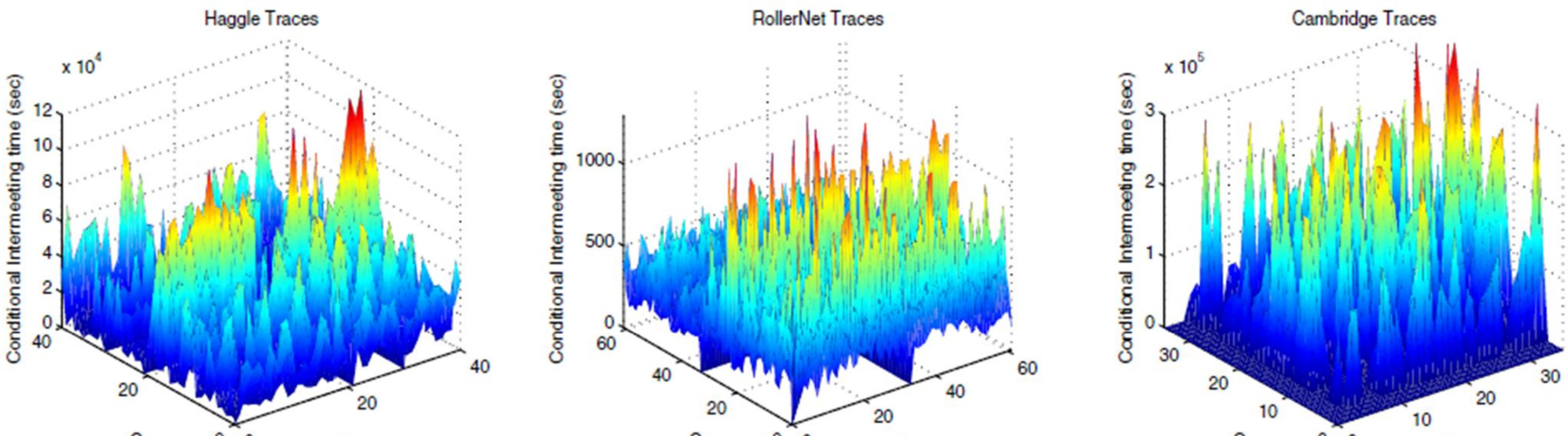
- Conclusion from Current Metric-based Algorithms
  - Forwarding decision based on <u>only individual relations</u> of nodes with destination
    - "Meeting with each other" is not used.
    - Meetings of nodes are assumed independent (uncorrelated) from each other

    - BUT node meetings may be correlated
      - Ex: Family (home), security guard (gate), office friends (work)

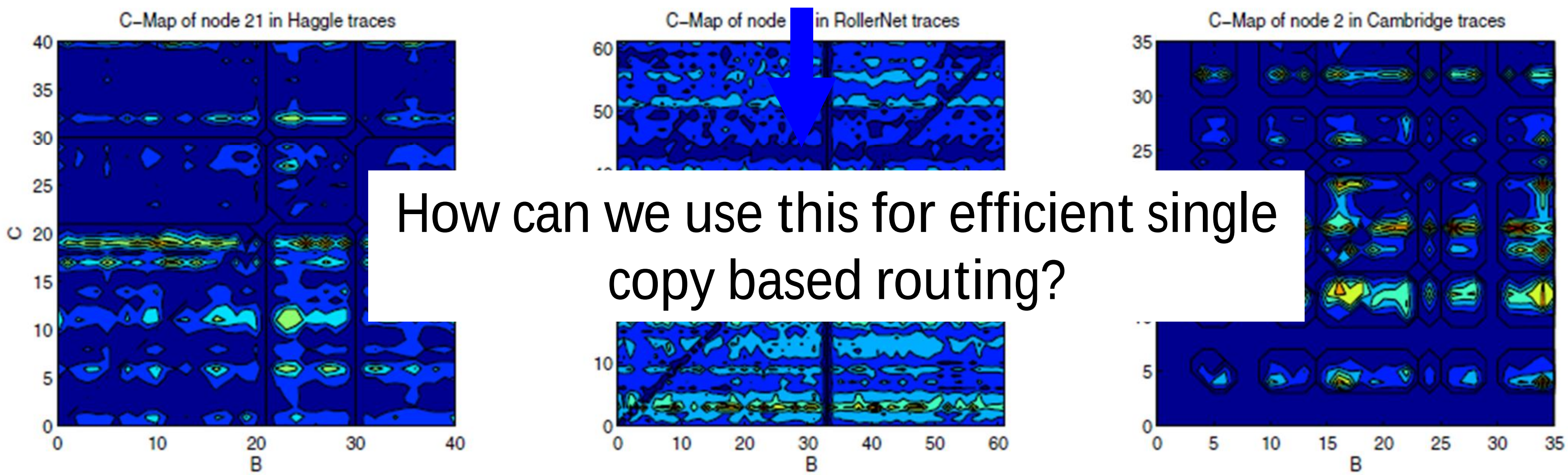
Proposed Metric

- Conditional Intermeeting Time:
  - $\tau_A(C|B)$ = Average time it takes for node A to meet node C after the time node A meets B (condition).
    - Can be computed from contact history
    - $\tau_A(C|C)$ standard intermeeting time
  - If the meetings of a node with other nodes are correlated (if the identity of B matters):
    - Residual time to a node's next meeting with other nodes can be predicted more accurately.

Statistics from Real Traces ($\tau_A(C|B)$)



Meetings of a node with other nodes are correlated



How can we use this for efficient single copy based routing?

(a) C-Map of node 21 in Haggle traces (b) C-Map of node 33 in RollerNet traces (c) C-Map of node 2 in Cambridge traces

Modification in Shortest-path based routing

- Conditional Shortest Path Routing (CSPR)
  - Finds path with minimum CSP and sends message over that path:

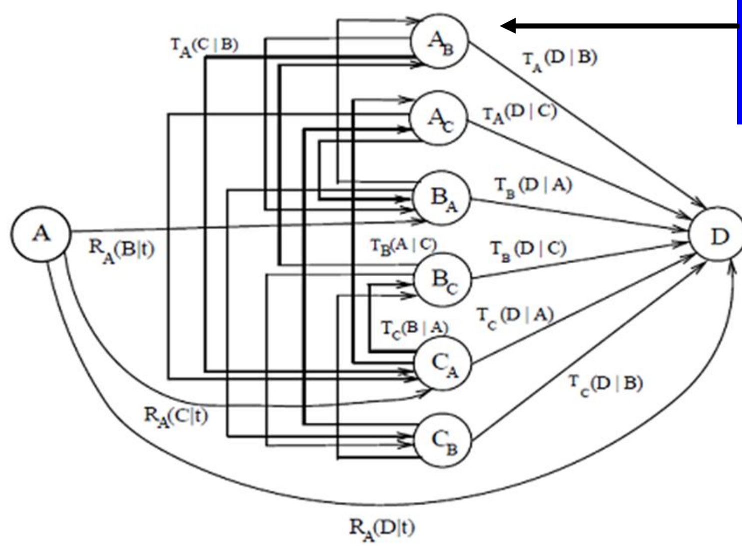
First hop weight

$$CSP(n_0, n_d) = \{n_0, n_1, \ldots, n_{d-1}, n_d \mid \Re_{n_0}(n_1|t) + \sum_{i=1}^{d-1} \tau_{n_i}(n_{i+1}|n_{i-1}) \text{ is minimized.}\}.$$

Conditional intermeeting time

Node A that has last met node B

Updated DTN Graph Model:



Modification in Metric based Routing Algorithms

- Current Design:
  - When node A meets node B, A forwards its message (which is destined to D) to B:
    - In **Prophet**: If A's delivery probability to D is smaller than B's delivery probability.
    - In **Fresh**: If B has a more recent meeting with D than A.

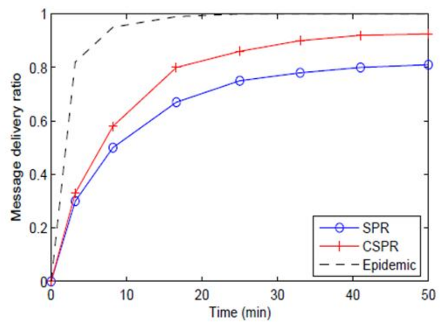
- Modification:
    - In **C-Prophet** and **C-Fresh**, we add <u>one more condition</u> for forwarding: $\tau_A(D|B) > \tau_B(D|A)$

Simulations

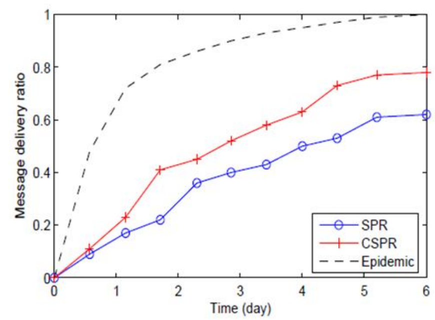
- Data Sets:
  - Real DTN Traces
    - Cambridge Traces
    - RollerNet Traces
    - Haggle Project Traces
  - Synthetic Traces based on Community Models
- Algorithms in Comparison:
  - Group 1: Shortest path based routing
    - 1) SPR (MEED, MED) 2) CSPR
  - Group 2: Metric based routing
    - 1) Fresh 2) Prophet 3) C-Fresh 4) C-Prophet
    - Epidemic (Optimal delivery ratio, delay)

Simulation Results (Group 1)

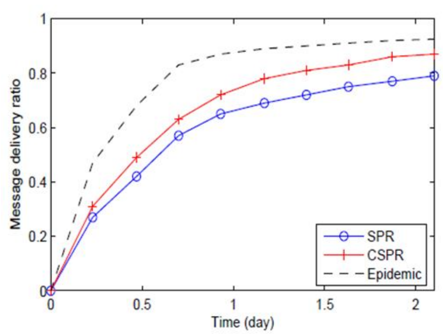
- Higher delivery ratio, lower delay



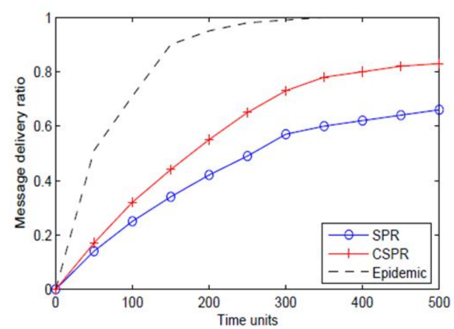
(a) RollerNet Traces



(b) Cambridge Traces



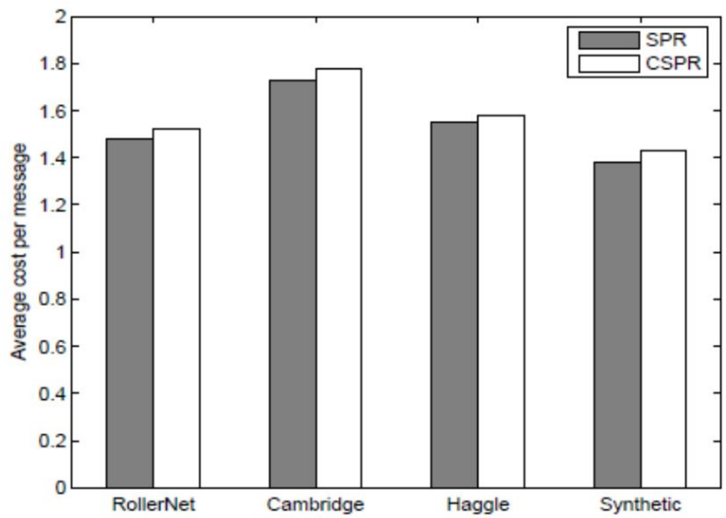
(c) Haggle Project Traces



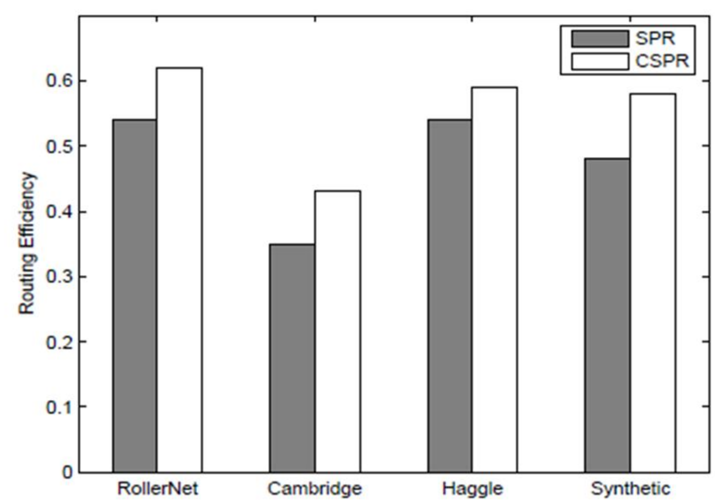
(d) Synthetic Data

Simulation Results (Group 1)

- Close Average Cost:
  - Number of forwardings per message
- Better Routing Efficiency (10%-23% improvement):
  - Delivery Ratio/Average Cost



(e) Average Cost

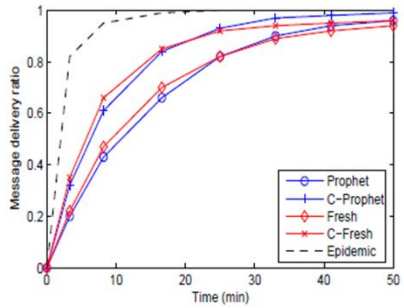


(f) Routing Efficiency

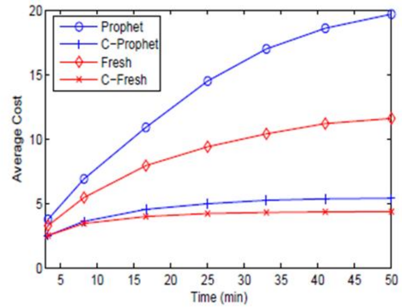
Simulation Results (Group 2)

- Modified versions are better in all performance metrics:

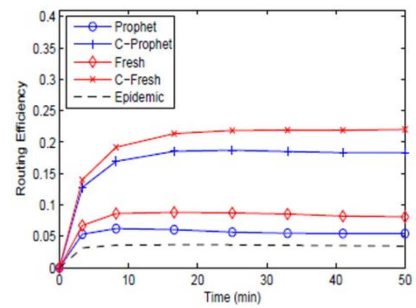
RollerNet Traces



(a) Message delivery ratio vs. time

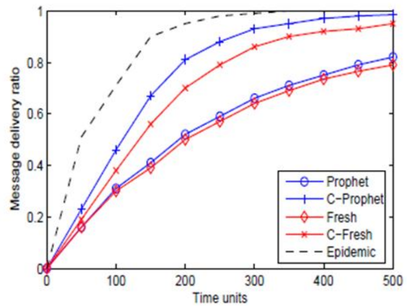


(b) Average cost vs. time

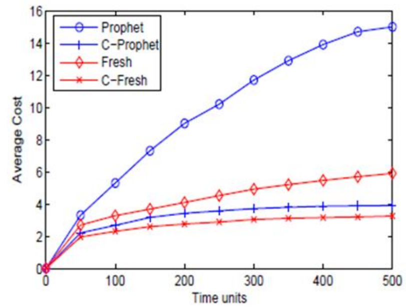


(c) Routing Efficiency vs. time

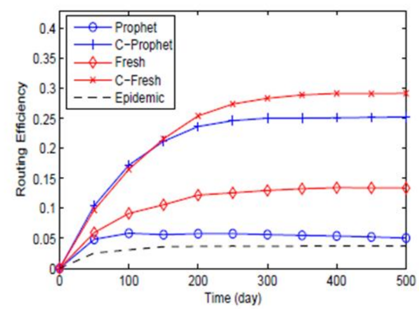
Synthetic Traces



(a) Message delivery ratio vs. time



(b) Average cost vs. time

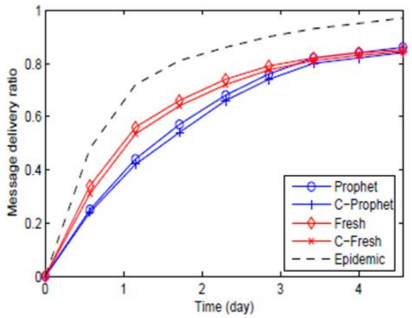


(c) Routing Efficiency vs. time

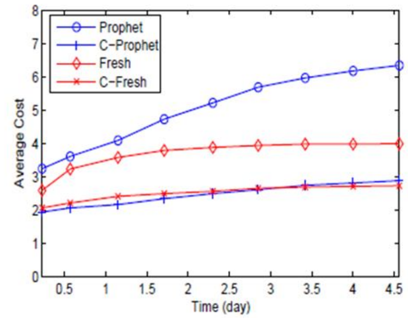
Simulation Results (Group 2)

- Modified versions provide smaller cost and better routing efficiency but same delivery ratio (due to not so clear repetitive meetings between nodes):

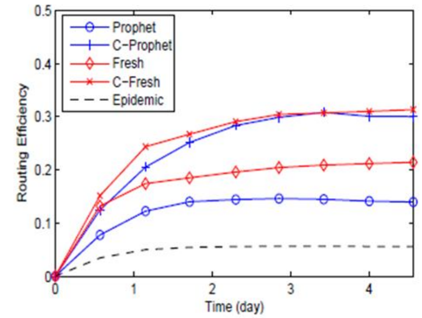
Cambridge Traces



(a) Message delivery ratio vs. time

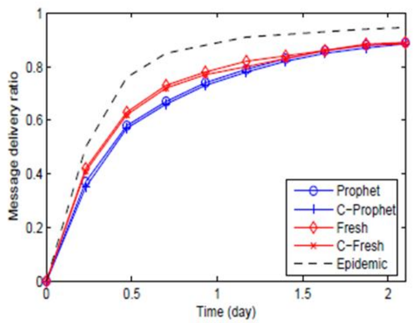


(b) Average cost vs. time

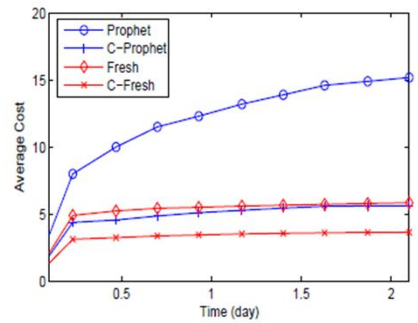


(c) Routing Efficiency vs. time

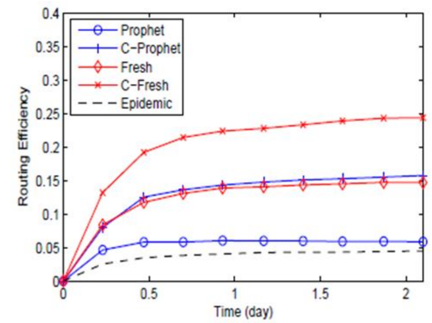
Haggle Traces



(a) Message delivery ratio vs. time



(b) Average cost vs. time



(c) Routing Efficiency vs. time

Outline

- Introduction to DTNs
  - Challenges of routing
- Proposed Algorithms
  1) Multi-period Spray and Wait routing
  2) Multi-period erasure coding based routing
  3) Efficient single-copy routing utilizing correlation between node meetings
  4) Social relation based routing
- Summary of Contributions

Social relation based routing

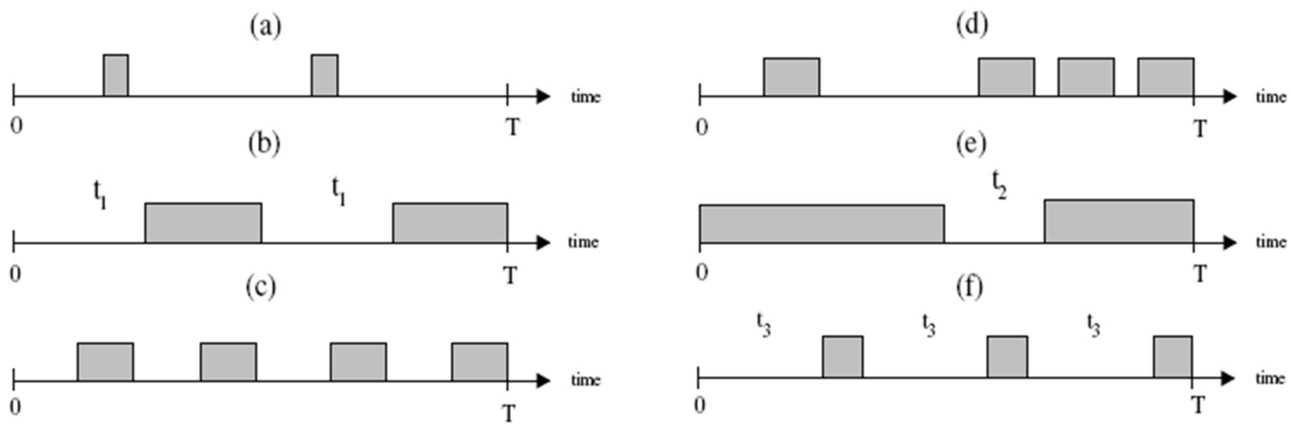
- Social network metrics are used to understand contact relations between nodes

- Recent work:
  - Similarity and betweenness: SimBet [Mobihoc 07]
  - Community Detection: BubbleRap [MobiHoc 08], LocalCom [Secon 09]

- Deficiencies:
  - Lack of a metric to detect the node relations (also the opportunities for message exchanges) accurately:
  - Improper handling of indirect relations
  - Improper handling of temporal variations in node relations

Analysis of “Message Exchange Opportunity” using Current Metrics

- Average Contact frequency
  - Can not differentiate cases a & b
- Average Contact duration
  - Can not differentiate cases b & c
- Average Separation Period
  - Can differentiate cases c & d utilizing variance (irregularity) between different separation periods
  - Can not differentiate cases b & e (also cases b & f)



Proposed Metric

- Social Pressure Metric (SPM)
  - Measures "**how actual is the knowledge of a person about his/her friend?**"
  - What would be the average message forwarding delay to node **j** if node **i** has a new message for node **j** at each time unit?

$$SPM_{i,j} = \frac{\int_{t=0}^{T} f(t)dt}{T} \text{ and } w_{i,j} = \frac{1}{SPM_{i,j}}$$

Node i's delivery metric (weight) to j

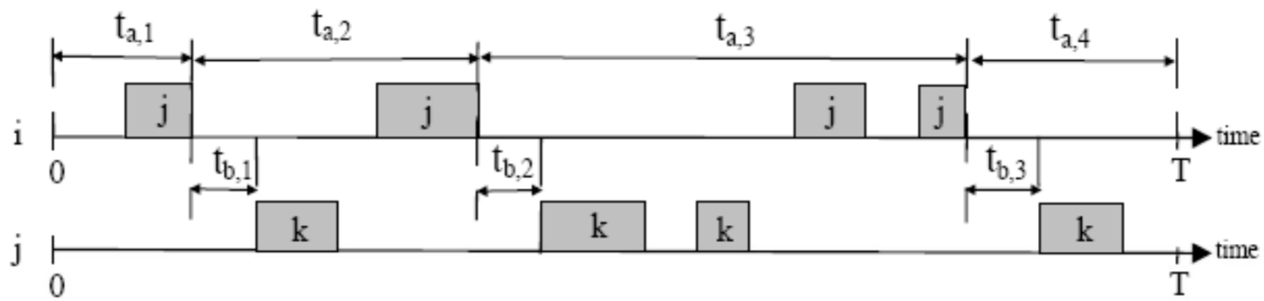
$$SPM_{i,j} = \left(\sum_{x=1}^{n} t_{inter,x}^{2}\right) / (2T)$$

Ex: 

SPM = $[(t_1)^2 + (t_2)^2 + (t_3)^2]/2T$

Indirect Relations

- Relative-SPM (RSPM)
  - What would be the average delivery delay of node i's continuously generated messages (for k) if they follow the path <i,j,k>?
    - Better than ($SPM_{i,j}$ + $SPM_{j,k}$) due to a possible correlation between node j's meetings with i and k



$$RSPM_{i,k|j} = \sum_{x=1}^{n} \int_0^{t_{a,x}} (t_{b,x} + t_{a,x} - t)dt = \sum_{x=1}^{n} \frac{2t_{b,x}t_{a,x} + t_{a,x}^2}{2}$$

$$w_{i,j,k} = 1/RSPM_{i,k|j}$$

Node i's indirect delivery metric (weight) to k through j

Friendship Community Formation

- Each node i:
  - computes its $SPM_{i,j}$ values with each **j**
  - receives its $RSPM_{i,j,k}$ from **j** periodically

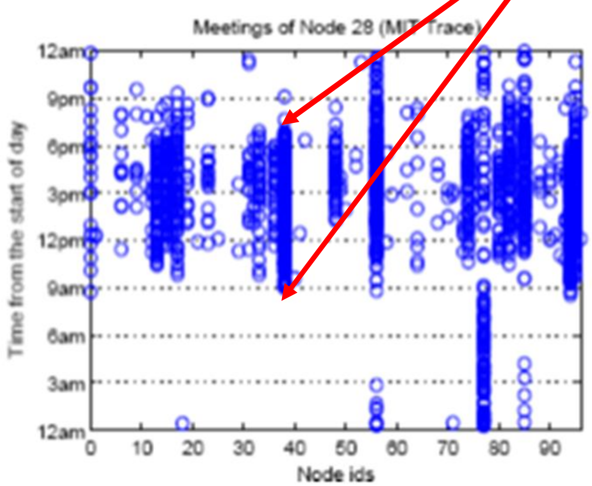
- Node i forms its friendship community as the nodes having direct or indirect weight larger than a threshold (τ):

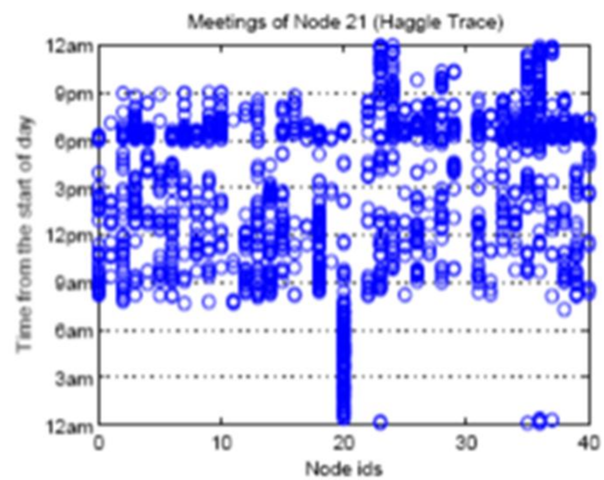
$$
\begin{aligned}
F_i &= \{j \mid w_{i,j} > \tau \text{ and } i \neq j\} \cup \\
& \quad \{k \mid w_{i,j,k} > \tau \text{ and } w_{i,j} > \tau \text{ and } i \neq j \neq k\}
\end{aligned}
$$

Temporal Variations in Node Relations

- Aging of weights may cause wrong decisions
- <u>Proposed Solution:</u> Generating different communities at different times of the day
  - Ex: 3 hour ranges

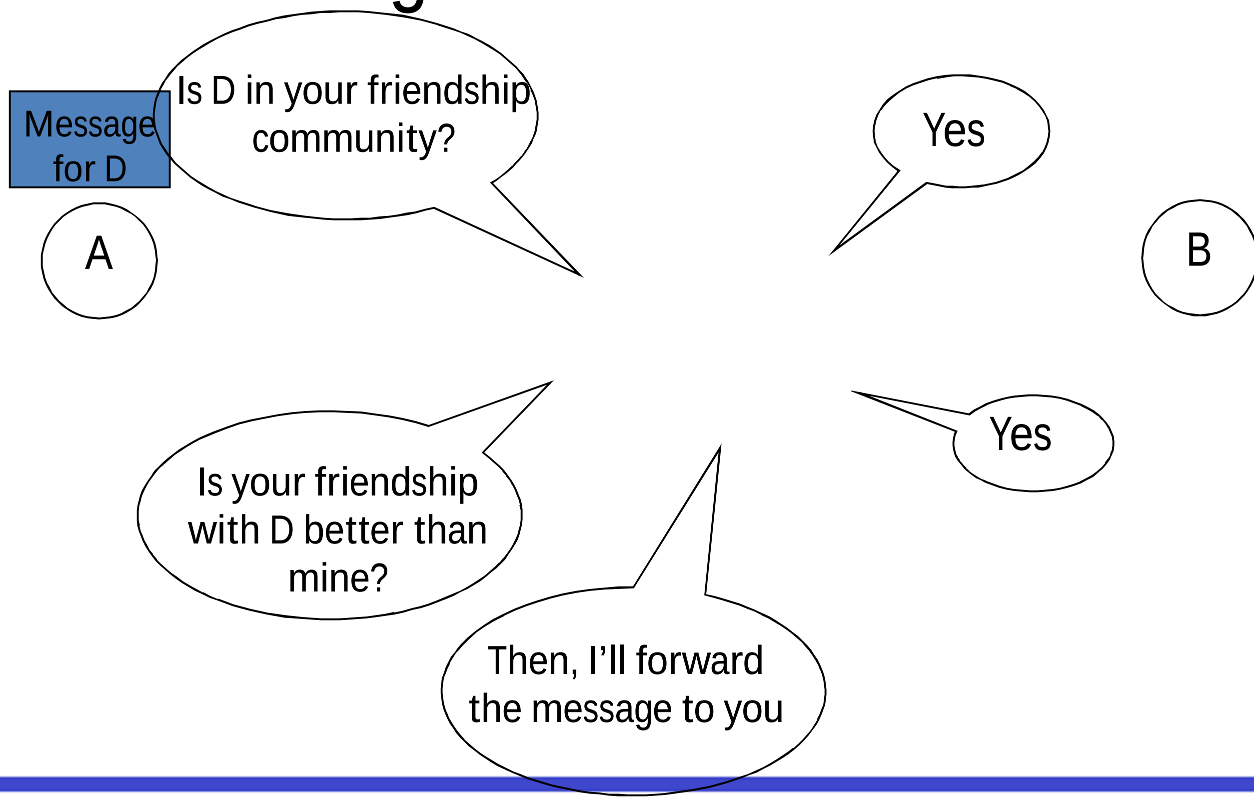
node 28 meets node 38
between 9am-7pm





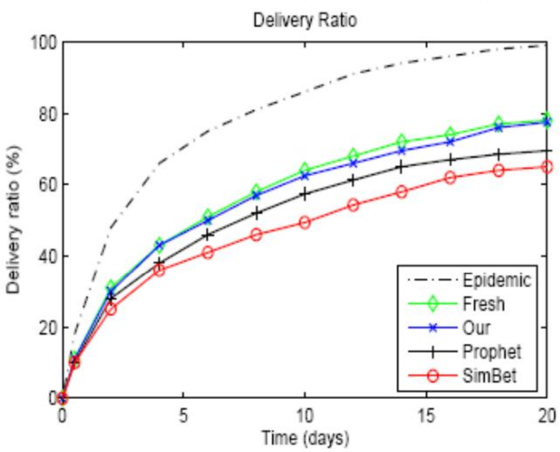
Forwarding Algorithm

- When two nodes meet, message is forwarded if the following case occurs:

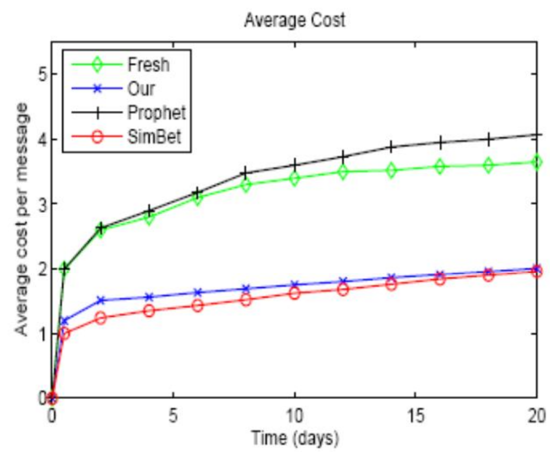


Simulation Results (MIT traces)

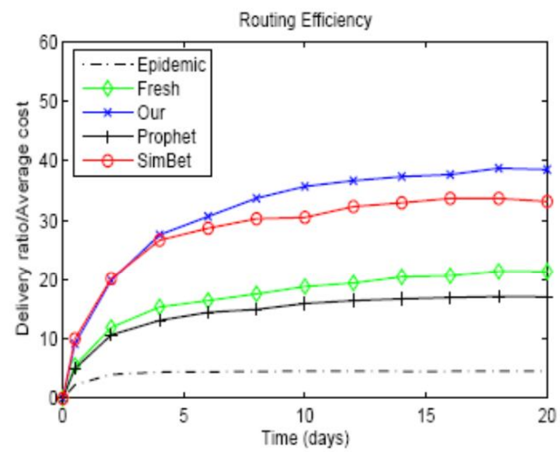
- Higher delivery ratio, lower delay
- Lowest cost, best routing efficiency (delivery rate/cost)



(a) Delivery ratio vs. time



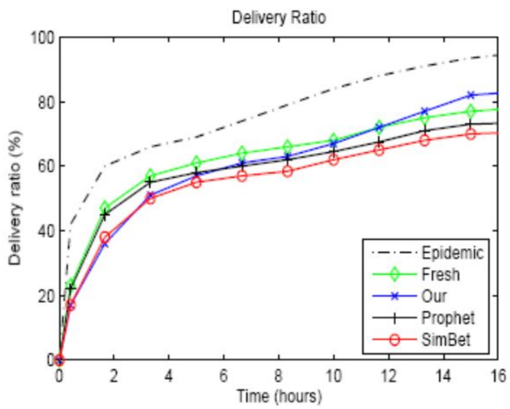
(b) Average cost vs. time



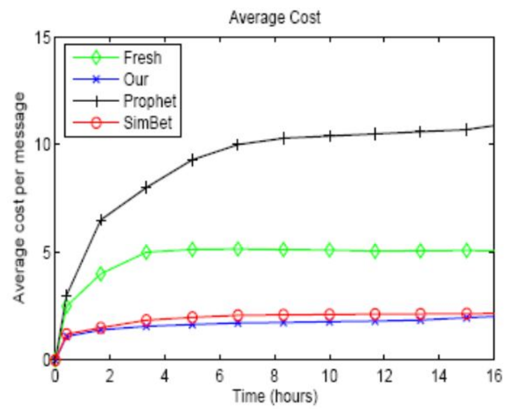
(c) Routing Efficiency vs. time

[Globecom 2010], [TPDS Journal in submission]

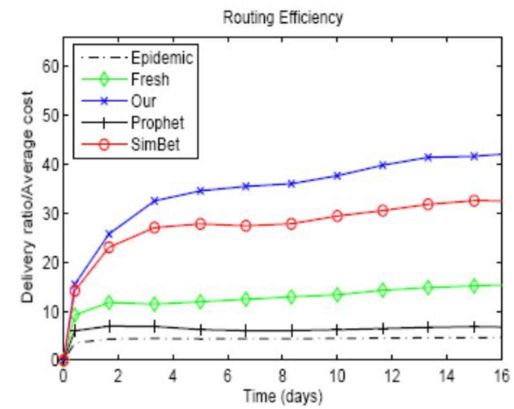
Simulation Results (Haggle Traces)



(a) Delivery ratio vs. time

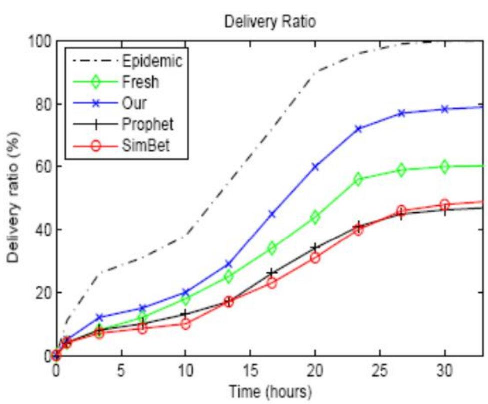


(b) Average cost vs. time

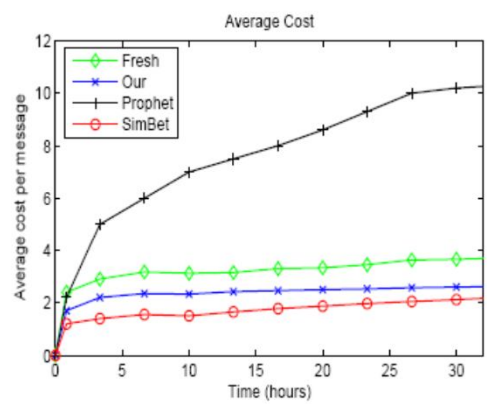


(c) Routing Efficiency vs. time

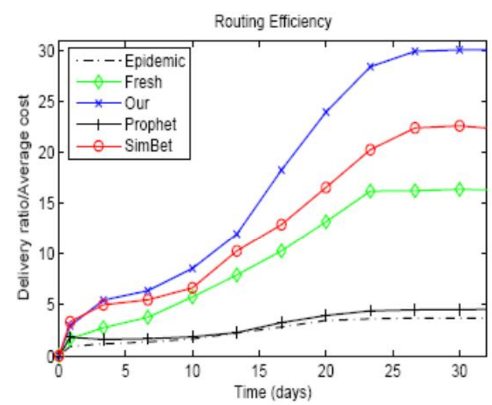
Simulation Results (Synthetic Traces)



(a) Delivery ratio vs. time



(b) Average cost vs. time



(c) Routing Efficiency vs. time

Outline

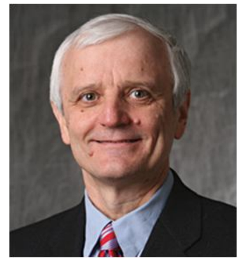
- Introduction to DTNs
  - Challenges of routing
- Proposed Algorithms
  1) Multi-period Spray and Wait routing
  2) Multi-period erasure coding based routing
  3) Efficient single-copy routing utilizing correlation between node meetings
  4) Social relation based routing
- **Summary of Contributions**

Summary of Contributions

- Multi-period spray and wait algorithm
  - Fewer average copies while keeping the same delivery rate by the deadline
- Cost efficient routing with erasure coded messages
  - Single and multi-period (reliability)
- Efficient single-copy routing
  - Conditional shortest path routing
  - Modifying metric based algorithms using conditional intermeeting time (correlated node mobility)
- Social relation based routing
  - Friendship based routing
  - (Relative) Social-pressure metric

Acknowledgment

- My advisor:
  - Prof. Boleslaw Szymanski



- My colleagues:
  - Zijian Wang

Sahin Cem Geyik

- Grants and Research Centers:
  - ITA International Technology Alliance in Network and Information Science
  - SCNARC Social and Cognitive Networks Academic Research Center

THANK YOU