

Trace-driven analysis of data forwarding in opportunistic networks

Merkourios Karaliopoulos, Panagiotis Pantazopoulos, Eva Jaho, Ioannis Stavrakakis
Department of Informatics and Telecommunications
University of Athens, Greece.
ANR group <http://anr.di.uoa.gr/>



presented by
Panagiotis Pantazopoulos

Overview

- exclusive use of mobile phone traces for assessing the performance of different opportunistic forwarding schemes
 - approach that relies on graph-expansion techniques to extract space-time graph constructs out of the contact sequence (datasets)
- assessment of a centrality-based forwarding scheme
- how about enriching these traces with users' interests?



Traces from mobile node encounters

Datasets (contact traces):

5 well-known iMote-based real traces available in the CRAWDAD archive*

- Bluetooth sightings by users carrying iMotes, gathered the last 5 years

Configuration	Intel	Cambridge	Infocom05	Content	Infocom06
Device type	iMote	iMote	iMote	iMote	iMote
Network type	B/T	B/T	B/T	B/T	B/T
Duration (days)	6	6	4	24	4
Scan time (sec)	5-10	5-10	5-10	5-10	5-10
Granularity (sec)	120	120	120	120-600	120
Mobile Devices	8	12	41	36	78
Stationary Dev.	1	0	0	18	20
External Dev.	119	211	233	11368	4421
Average internal contacts/pair/day	9.09	12.09	8.60	0.66	9.03
# of Contacts	2766	6732	28216	41330	227657

* <http://crawdad.cs.dartmouth.edu/>



Computation of shortest paths over traces

Input are time-ordered traces of node contacts

Step 1. from the original contact trace to the sequence of forwarding contacts

—*inflating the original trace*

Why? for forwarding opportunities to become visible to the parser (no need for backtracking)

contact id	involved nodes	contact start time	contact end time	contact id	involved nodes	contact start time	contact end time
...
C ₀	1 3	589	940	C _{0,0}	1 3	589	940
C ₁	2 3	589	619	C ₁	2 3	589	619
C ₂	6 7	639	699	C _{0,1}	1 3	589	940
C ₃	1 5	700	816	C ₂	6 7	639	699
C ₄	5 6	702	818	C _{0,2}	1 3	639	940
C ₅	2 7	816	1185	C _{3,0}	1 5	700	816
C ₆	4 8	819	909	C _{0,3}	1 3	700	940
C ₇	8 6	938	1182	C _{4,0}	5 6	702	818
...	C _{0,4}	1 3	702	940
				C _{3,1}	1 5	702	816
				C _{5,0}	2 7	816	1185
				C _{0,5}	1 3	816	940
				C _{3,2}	1 5	816	816
				C _{4,1}	5 6	816	818
				C ₆	4 8	819	909
				C _{0,5}	1 3	819	940
				C _{5,1}	2 7	819	1185
				C ₇	8 6	938	1182
				C _{0,6}	1 3	938	940
				C _{5,2}	2 7	938	1185
			

Each record C_k is replicated as many times as the number of contact records that start later than C_k but before its end time



Computation of shortest paths over traces

Input are time-ordered traces of node contacts

Step 1. from the original contact trace to the sequence of forwarding contacts

—*inflating the original trace*

Why? for forwarding opportunities to become visible to the parser (no need for backtracking)

contact id	involved nodes	contact start time	contact end time		contact id	involved nodes	contact start time	contact end time
...
C ₀	1 3	589	940	→	C _{0,0}	1 3	589	940
C ₁	2 3	589	619		C ₁	2 3	589	619
C ₂	6 7	639	699	→	C _{0,1}	1 3	589	940
C ₃	1 5	700	816	→	C ₂	6 7	639	699
C ₄	5 6	702	818	→	C _{0,2}	1 3	639	940
C ₅	2 7	816	1185	→	C _{3,0}	1 5	700	816
C ₆	4 8	819	909	→	C _{0,3}	1 3	700	940
C ₇	8 6	938	1182	→	C _{4,0}	5 6	702	818
...		C _{0,4}	1 3	702	940
					C _{3,1}	1 5	702	816
					C _{5,0}	2 7	816	1185
				→	C _{0,5}	1 3	816	940
					C _{3,2}	1 5	816	816
					C _{4,1}	5 6	816	818
					C ₆	4 8	819	909
				→	C _{0,5}	1 3	819	940
					C _{5,1}	2 7	819	1185
					C ₇	8 6	938	1182
				→	C _{0,6}	1 3	938	940
					C _{5,2}	2 7	938	1185
				

Each record C_k is replicated as many times as the number of contact records that start later than C_k but before its end time



Computation of shortest paths over traces

Step 1 cont'd. from the original contact trace to the sequence of forwarding contacts

—*Contact-filtering*

Why?: To retain those contacts that can result in data forwarding

use criteria that account for the different opportunistic schemes

t_s : when a message is available at source node s for destination d

Example

Two-hops forwarding: nodes other than the message source can forward it only to the destination node

$$\text{Message } m = (n1, n6, t_g < t_{0s})$$



Computation of shortest paths over traces

Step 1 cont'd. from the original contact trace to the sequence of forwarding contacts

contact id	involved nodes		contact start time	contact end time
...
C ₀	n1	n2	t _{0,s}	t _{0,e}
C ₁	n3	n4	t _{1,s}	t _{1,e}
C ₂	n4	n5	t _{2,s}	t _{2,e}
C ₃	n2	n5	t _{3,s}	t _{3,e}
C ₄	n5	n3	t _{4,s}	t _{4,e}
C ₅	n3	n2	t _{5,s}	t _{5,e}
C ₆	n3	n6	t _{6,s}	t _{6,e}
C ₇	n1	n4	t _{7,s}	t _{7,e}
C ₈	n2	n6	t _{8,s}	t _{8,e}
...



Computation of shortest paths over traces

Step 1 cont'd. from the original contact trace to the sequence of forwarding contacts

contact id	involved nodes		contact start time	contact end time
...
C ₀	n1	n2	t _{0,s}	t _{0,e}
C ₁	n3	n4	t _{1,s}	t _{1,e}
C ₂	n4	n5	t _{2,s}	t _{2,e}
C ₃	n2	n5	t _{3,s}	t _{3,e}
C ₄	n5	n3	t _{4,s}	t _{4,e}
C ₅	n3	n2	t _{5,s}	t _{5,e}
C ₆	n3	n6	t _{6,s}	t _{6,e}
C ₇	n1	n4	t _{7,s}	t _{7,e}
C ₈	n2	n6	t _{8,s}	t _{8,e}
...

records up to the first one $c_0 = (n1, *, t_s)$ are excluded



Computation of shortest paths over traces

Step 1 cont'd. from the original contact trace to the sequence of forwarding contacts

contact id	involved nodes	contact start time	contact end time
...
→ C ₀	n1 n2	t _{0,s}	t _{0,e}
C ₁	n3 n4	t _{1,s}	t _{1,e}
C ₂	n4 n5	t _{2,s}	t _{2,e}
C ₃	n2 n5	t _{3,s}	t _{3,e}
C ₄	n5 n3	t _{4,s}	t _{4,e}
C ₅	n3 n2	t _{5,s}	t _{5,e}
C ₆	n3 n6	t _{6,s}	t _{6,e}
C ₇	n1 n4	t _{7,s}	t _{7,e}
C ₈	n2 n6	t _{8,s}	t _{8,e}
...

Forwarding list=[n1, n6]



Computation of shortest paths over traces

Step 1 cont'd. from the original contact trace to the sequence of forwarding contacts

contact id	involved nodes	contact start time	contact end time
...
→ C ₀	n1 n2	t _{0,s}	t _{0,e}
C ₁	n3 n4	t _{1,s}	t _{1,e}
C ₂	n4 n5	t _{2,s}	t _{2,e}
C ₃	n2 n5	t _{3,s}	t _{3,e}
C ₄	n5 n3	t _{4,s}	t _{4,e}
C ₅	n3 n2	t _{5,s}	t _{5,e}
C ₆	n3 n6	t _{6,s}	t _{6,e}
C ₇	n1 n4	t _{7,s}	t _{7,e}
C ₈	n2 n6	t _{8,s}	t _{8,e}
...

Forwarding list=[n1, n6]

Forwarding opportunities:
one of the two nodes is
in the forwarding list



Computation of shortest paths over traces

Step 1 cont'd. from the original contact trace to the sequence of forwarding contacts

contact id	involved nodes	contact start time	contact end time
...
C ₀	n1 n2	t _{0,s}	t _{0,e}
C ₁	n3 n4	t _{1,s}	t _{1,e}
C ₂	n4 n5	t _{2,s}	t _{2,e}
C ₃	n2 n5	t _{3,s}	t _{3,e}
C ₄	n5 n3	t _{4,s}	t _{4,e}
C ₅	n3 n2	t _{5,s}	t _{5,e}
C ₆	n3 n6	t _{6,s}	t _{6,e}
C ₇	n1 n4	t _{7,s}	t _{7,e}
C ₈	n2 n6	t _{8,s}	t _{8,e}
...

Forwarding list=[n1, n6]



Computation of shortest paths over traces

Step 1 cont'd. from the original contact trace to the sequence of forwarding contacts

contact id	involved nodes	contact start time	contact end time
...
C₀	n1 n2	t_{0,s}	t_{0,e}
C₁	n3 n4	t_{1,s}	t_{1,e}
C₂	n4 n5	t_{2,s}	t_{2,e}
C₃	n2 n5	t_{3,s}	t_{3,e}
C₄	n5 n3	t_{4,s}	t_{4,e}
C₅	n3 n2	t_{5,s}	t_{5,e}
C₆	n3 n6	t_{6,s}	t_{6,e}
→ C ₇	n1 n4	t _{7,s}	t _{7,e}
C ₈	n2 n6	t _{8,s}	t _{8,e}
...

Forwarding list=[n1, n6, n4]



Computation of shortest paths over traces

Step 1 cont'd. from the original contact trace to the sequence of forwarding contacts

contact id	involved nodes	contact start time	contact end time
...
C₀	n1 n2	t_{0,s}	t_{0,e}
C₁	n3 n4	t_{1,s}	t_{1,e}
C₂	n4 n5	t_{2,s}	t_{2,e}
C₃	n2 n5	t_{3,s}	t_{3,e}
C₄	n5 n3	t_{4,s}	t_{4,e}
C₅	n3 n2	t_{5,s}	t_{5,e}
C₆	n3 n6	t_{6,s}	t_{6,e}
C ₇	n1 n4	t _{7,s}	t _{7,e}
→ C ₈	n2 n6	t _{8,s}	t _{8,e}
...

Forwarding list=[n1, n6, n4, n2]



Computation of shortest paths over traces

Step 1 cont'd. from the original contact trace to the sequence of forwarding contacts

—*Contact-filtering*

Two-hops forwarding

Message $m=(n1,n6,t_g < t_{qs})$

contact id	involved nodes		contact start time	contact end time	additional fields
...
C ₀	n1	n2	t _{0,s}	t _{0,e}	...
C ₁	n3	n4	t _{1,s}	t _{1,e}	...
C ₂	n4	n5	t _{2,s}	t _{2,e}	...
C ₃	n2	n5	t _{3,s}	t _{3,e}	...
C ₄	n5	n3	t _{4,s}	t _{4,e}	...
C ₅	n3	n2	t _{5,s}	t _{5,e}	...
C ₆	n3	n6	t _{6,s}	t _{6,e}	...
C ₇	n1	n4	t _{7,s}	t _{7,e}	...
C ₈	n2	n6	t _{8,s}	t _{8,e}	...
...

Original trace

contact id	involved nodes		contact start time	contact end time	additional fields
...
C ₀	n1	n2	t _{0,s}	t _{0,e}	...
C ₁	n3	n4	t _{1,s}	t _{1,e}	...
C ₂	n4	n5	t _{2,s}	t _{2,e}	...
C ₃	n2	n5	t _{3,s}	t _{3,e}	...
C ₄	n5	n3	t _{4,s}	t _{4,e}	...
C ₅	n3	n2	t _{5,s}	t _{5,e}	...
C ₆	n3	n6	t _{6,s}	t _{6,e}	...
C ₇	n1	n4	t _{7,s}	t _{7,e}	...
C ₈	n2	n6	t _{8,s}	t _{8,e}	...
...

Forwarding contacts for the two-hops forwarding scheme



Computation of shortest paths over traces

Step 2. building the forwarding contact graph

Outcome of first step: a reduced set of contact records corresponding to forwarding contacts.

Aim : to derive the graph construct $G_c = (V_c, E_c)$ that can capture these contacts and their timing relationship.

Controlled flooding schemes: $G_c = (V_c, E_c)$ is built out of

- the first contact c_0 and
- forwarding contacts occurring thereafter.



Computation of shortest paths over traces

Step 2. building the forwarding contact graph

contact id	involved nodes		contact start time	contact end time
...
C_0	n1	n2	$t_{0,s}$	$t_{0,e}$
C_1	n3	n4	$t_{1,s}$	$t_{1,e}$
C_2	n4	n5	$t_{2,s}$	$t_{2,e}$
C_3	n2	n5	$t_{3,s}$	$t_{3,e}$
C_4	n5	n3	$t_{4,s}$	$t_{4,e}$
C_5	n3	n2	$t_{5,s}$	$t_{5,e}$
C_6	n3	n6	$t_{6,s}$	$t_{6,e}$
C_7	n1	n4	$t_{7,s}$	$t_{7,e}$
C_8	n2	n6	$t_{8,s}$	$t_{8,e}$
...



Computation of shortest paths over traces

Step 2. building the forwarding contact graph

contact id	involved nodes		contact start time	contact end time
...
→ C ₀	n1	n2	t _{0,s}	t _{0,e}
C ₁	n3	n4	t _{1,s}	t _{1,e}
C ₂	n4	n5	t _{2,s}	t _{2,e}
C ₃	n2	n5	t _{3,s}	t _{3,e}
C ₄	n5	n3	t _{4,s}	t _{4,e}
C ₅	n3	n2	t _{5,s}	t _{5,e}
C ₆	n3	n6	t _{6,s}	t _{6,e}
C ₇	n1	n4	t _{7,s}	t _{7,e}
C ₈	n2	n6	t _{8,s}	t _{8,e}
...



Computation of shortest paths over traces

Step 2. building the forwarding contact graph

contact id	involved nodes		contact start time	contact end time
...
→ C ₀	n1	n2	t _{0,s}	t _{0,e}
C ₁	n3	n4	t _{1,s}	t _{1,e}
C ₂	n4	n5	t _{2,s}	t _{2,e}
C ₃	n2	n5	t _{3,s}	t _{3,e}
C ₄	n5	n3	t _{4,s}	t _{4,e}
C ₅	n3	n2	t _{5,s}	t _{5,e}
C ₆	n3	n6	t _{6,s}	t _{6,e}
C ₇	n1	n4	t _{7,s}	t _{7,e}
C ₈	n2	n6	t _{8,s}	t _{8,e}
...

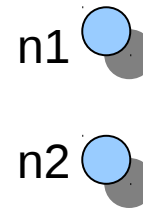
n1 



Computation of shortest paths over traces

Step 2. building the forwarding contact graph

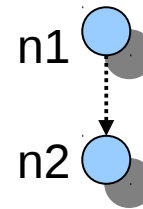
contact id	involved nodes		contact start time	contact end time
...
→ C ₀	n1	n2	t _{0,s}	t _{0,e}
C ₁	n3	n4	t _{1,s}	t _{1,e}
C ₂	n4	n5	t _{2,s}	t _{2,e}
C ₃	n2	n5	t _{3,s}	t _{3,e}
C ₄	n5	n3	t _{4,s}	t _{4,e}
C ₅	n3	n2	t _{5,s}	t _{5,e}
C ₆	n3	n6	t _{6,s}	t _{6,e}
C ₇	n1	n4	t _{7,s}	t _{7,e}
C ₈	n2	n6	t _{8,s}	t _{8,e}
...



Computation of shortest paths over traces

Step 2. building the forwarding contact graph

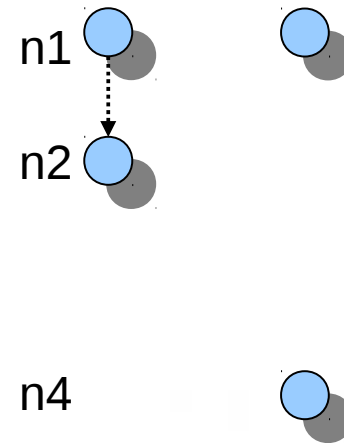
contact id	involved nodes		contact start time	contact end time
...
→ C ₀	n1	n2	t _{0,s}	t _{0,e}
C ₁	n3	n4	t _{1,s}	t _{1,e}
C ₂	n4	n5	t _{2,s}	t _{2,e}
C ₃	n2	n5	t _{3,s}	t _{3,e}
C ₄	n5	n3	t _{4,s}	t _{4,e}
C ₅	n3	n2	t _{5,s}	t _{5,e}
C ₆	n3	n6	t _{6,s}	t _{6,e}
C ₇	n1	n4	t _{7,s}	t _{7,e}
C ₈	n2	n6	t _{8,s}	t _{8,e}
...



Computation of shortest paths over traces

Step 2. building the forwarding contact graph

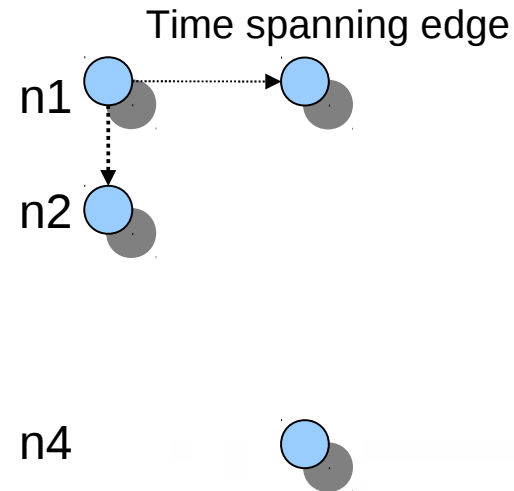
contact id	involved nodes		contact start time	contact end time
...
C_0	n1	n2	$t_{0,s}$	$t_{0,e}$
C_1	n3	n4	$t_{1,s}$	$t_{1,e}$
C_2	n4	n5	$t_{2,s}$	$t_{2,e}$
C_3	n2	n5	$t_{3,s}$	$t_{3,e}$
C_4	n5	n3	$t_{4,s}$	$t_{4,e}$
C_5	n3	n2	$t_{5,s}$	$t_{5,e}$
C_6	n3	n6	$t_{6,s}$	$t_{6,e}$
→ C_7	n1	n4	$t_{7,s}$	$t_{7,e}$
C_8	n2	n6	$t_{8,s}$	$t_{8,e}$
...



Computation of shortest paths over traces

Step 2. building the forwarding contact graph

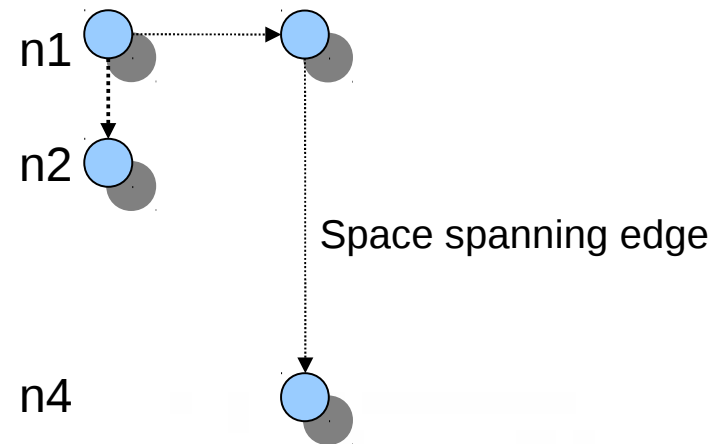
contact id	involved nodes		contact start time	contact end time
...
C_0	n1	n2	$t_{0,s}$	$t_{0,e}$
C_1	n3	n4	$t_{1,s}$	$t_{1,e}$
C_2	n4	n5	$t_{2,s}$	$t_{2,e}$
C_3	n2	n5	$t_{3,s}$	$t_{3,e}$
C_4	n5	n3	$t_{4,s}$	$t_{4,e}$
C_5	n3	n2	$t_{5,s}$	$t_{5,e}$
C_6	n3	n6	$t_{6,s}$	$t_{6,e}$
→ C_7	n1	n4	$t_{7,s}$	$t_{7,e}$
C_8	n2	n6	$t_{8,s}$	$t_{8,e}$
...



Computation of shortest paths over traces

Step 2. building the forwarding contact graph

contact id	involved nodes		contact start time	contact end time
...
C_0	n1	n2	$t_{0,s}$	$t_{0,e}$
C_1	n3	n4	$t_{1,s}$	$t_{1,e}$
C_2	n4	n5	$t_{2,s}$	$t_{2,e}$
C_3	n2	n5	$t_{3,s}$	$t_{3,e}$
C_4	n5	n3	$t_{4,s}$	$t_{4,e}$
C_5	n3	n2	$t_{5,s}$	$t_{5,e}$
C_6	n3	n6	$t_{6,s}$	$t_{6,e}$
→ C_7	n1	n4	$t_{7,s}$	$t_{7,e}$
C_8	n2	n6	$t_{8,s}$	$t_{8,e}$
...



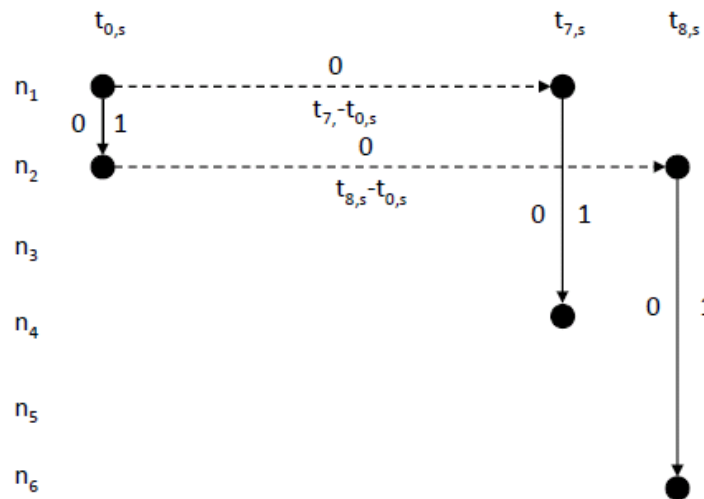
Computation of shortest paths over traces

Step 2. building the forwarding contact graph

Construct G_c is weighted:

-minimum-delay: the time-spanning edge weights express the time over which a message is stored and carried by a given node.

-minimum-hopcount: time-spanning edges are assigned zero weights and space-spanning ones unit weights.



Two-hop forwarding graph construct G_c



Computation of shortest paths over traces

Step 3. computing shortest paths over the forwarding contact graph

Standard shortest paths algorithms are run over the G_c construct

-to yield the space-time path resulting minimum message delivery

It can be shown that:

-graph constructs G_c are directed acyclic graphs (DAGs)
-running Dijkstra will yield the source destination minimum-hopcount space-time path in $O(|V|^2 \log_2 |V|)$



Assessment of centrality-based forwarding

SNA-based opportunistic forwarding :

- Forwarding decisions based on the Betweenness Centrality values (perfect info)
- How good (close-to-optimal) can centrality-based routing perform under perfect knowledge?



Assessment of centrality-based forwarding

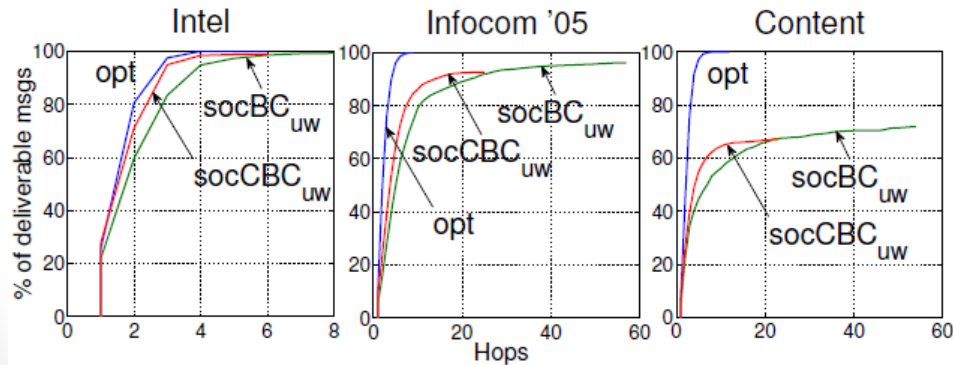
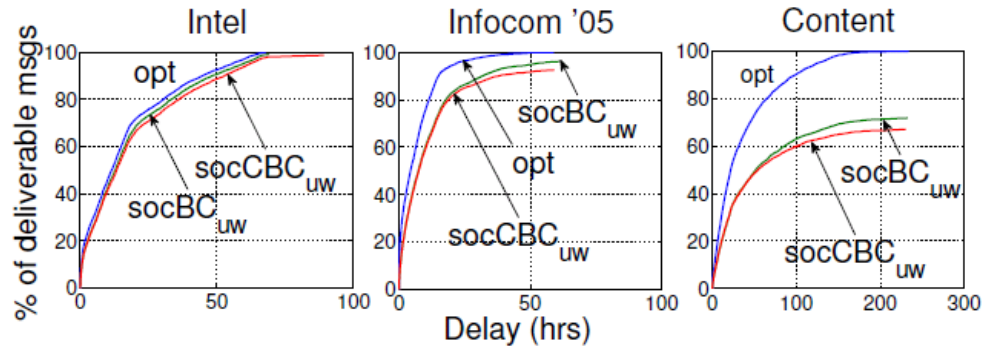
Experimentation methodology

- Emulation of optimal (*opt*) forwarding
 - minimum delay and hopcount paths computed directly out of the contact sequence
- Emulation of BC-based forwarding :
 - replaying the trace we compute centrality (variants) for each contact record and take forwarding decisions



Indicative set of results

- BC against the destination-aware Conditional BC (CBC) as forwarding criterion
 - CBC: maintains per-destination centrality values for each node



-CBC results in fewer hops for successful delivery



How about enriching the traces with users' interests?

- Interest-based forwarding improves considerably the information dissemination in opportunistic settings
 - requires social information about user's preferences
- ISCoDe framework that identifies communities of nodes with similar interests.
 - end-user interests can be inferred out of a real online social networking (OSN) application
- Option to be considered: Enriching mobile phone datasets with preferences of users.
 - encoded as user preference distributions over a set of thematic areas (i.e., music, sports)
 - inferred out of tags annotating data that users save in their mobile phones.



Future work across the three threads

- To expand the application of our trace-driven analysis on a wide range of protocols that assess the relaying utility of encountered nodes
- To explore the impact of weighted contact graphs on the centrality-based routing and expand our study to “many-to-many” communication settings.
- To infer social information from OSNs trying to correlate traces of encounters with online user profiles



NKUA published work of relevance

M. Karaliopoulos and C. Rochner. Trace-based performance analysis of opportunistic forwarding. Technical report, Online:<http://www.csg.ethz.ch/people/karaliom/traceanalysis.pdf> July 2011.

P. Nikolopoulos, T. Papadimitriou, P. Pantazopoulos, M. Karaliopoulos, and I. Stavrakakis, "How much off-center are centrality metrics for routing in opportunistic networks," in ACM MobiCom 2011 CHANTS Workshop), Las Vegas, NV, USA, Sep 2011.

S. Allen, M. Chorley, G. Colombo, E. Jaho, M. Karaliopoulos, I. Stavrakakis, and R. Whitaker, "Exploiting user interest similarity and social links for microblog forwarding in mobile opportunistic networks," in Elsevier Pervasive and Mobile Computing (submitted), 2011.

E. Jaho, M. Karaliopoulos, and I. Stavrakakis, "ISCoDe: A framework for interest similarity-based community detection in social networks," in Proc. 3rd Int'l Workshop on Network Science for Communication Networks, 2011.

P. Pantazopoulos, M. Karaliopoulos, and I. Stavrakakis. Centrality-driven scalable service migration. In the 23rd International Teletraffic Congress (ITC 2011), San Francisco, CA, USA, Sept. 2011



Questions ? ? ?



Back up slides

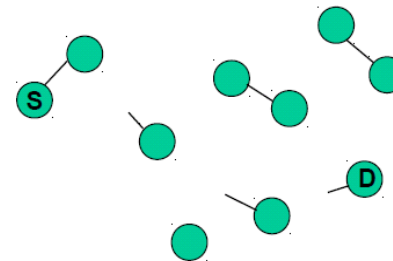


The opportunistic networking paradigm

Opportunistic Networks (OppNets)

-typically, a complete end-to-end path is highly unstable and may break soon after it has been discovered

OppNets



The opportunistic networking paradigm

How to deliver a message?

- Method 1: replication-based message forwarding
 - Epidemic: full replication, min delay, severe contention
 - Probabilistic, two-hop message routing, spray-and-*
- Method 2: heuristic forwarding
 - assign utilities to individual nodes to capture their relaying significance
 - the utility of a node may be related to some SNA metric

