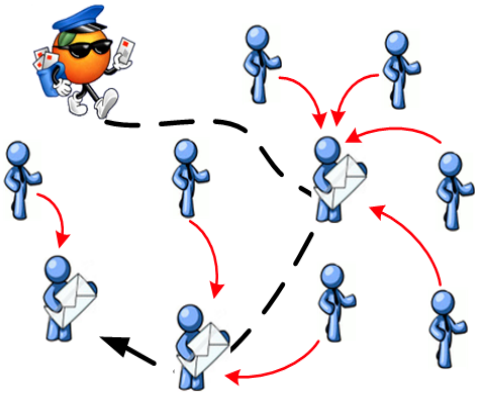
Collaborative Data Collection in Global Sensing Systems

Greg Bigwood†‡, Aline Carneiro Viana‡, Marcelo Dias de Amorim*, and Mathias Boc◊
†University of St Andrews, ‡INRIA, *UPMC Sorbonne Universités, ◊CEA LIST
gjb4@st-andrews.ac.uk, aline.viana@inria.fr, marcelo.amorim@lip6.fr,
michael.boc@cea.fr

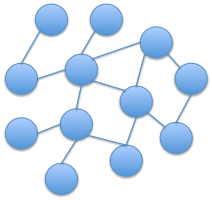


What? We compute subsets of mobile nodes from which a mobile collector can collect messages such that every node's messages are collected.

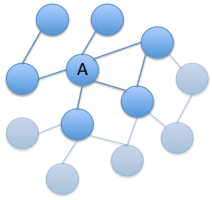
Why? Collecting messages from a subset of delegates rather than all nodes improves efficiency and reduces cost of collection for the collector.

How? We compute the delegate sets using metrics inspired by social networking. Nodes use local information to predict which of their neighbours to use as a delegate for their messages during a timeslot, based on historical information[1]. We compare this approach against computing subsets using global information and against the optimal subset of nodes. We evaluate the strategies using trace-driven analysis of college campus and taxi-cab networks.

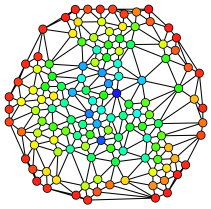
Strategy We use Betweenness-Centrality[2] and Degree-Centrality[3] as our delegate set selection metrics. We contrast the results of using either Ego-centric or Complete-network graphs from which to derive the metric scores.



Complete network



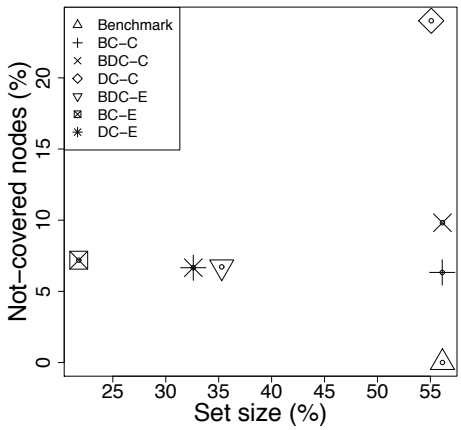
Ego-network of **A**



Betweenness of nodes[4]

Degree centrality: number of ties a nodes has
Betweenness centrality: the sum of the number of shortest paths on which a nodes lies between a pair of nodes, divided by the number of paths between the pair; for all pairs.
Benchmark: derived from MPRset[5]

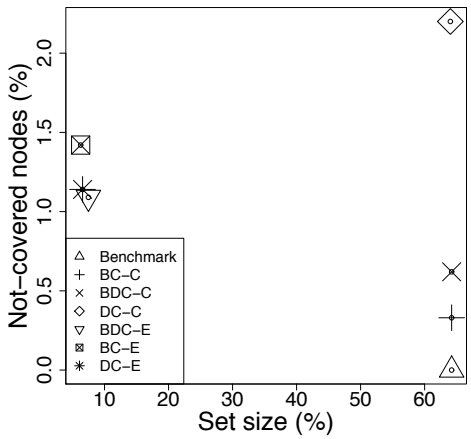
Results We observe that much can be extracted from the inherent mobility of the nodes and that our prediction strategy is effective when used for delegate selection. We were surprised to note that local knowledge of the network is more than enough to achieve high collection ratios, with values that are close to those obtained with full knowledge of the topology.



Dartmouth campus trace

To assess the performance of the delegate-selection schemes we compare the average percentage of nodes not covered by the collection strategy per timeslot.

KEY:
BC= Betweenness centrality
DC= Degree centrality
BDC= 50%BC and 50%DC mix
-C= Operates on complete network
-E= Operates on ego-network



Taxi-cab trace

Future Work We are applying this approach to synthetic mobility traces. We intend to compute the appropriate collector trajectory to decide in which order to visit delegates. We hope to demonstrate further how local information and collaborative data collection can be useful when infrastructure access is prohibitively expensive.

[1] "**Opportunistic data collection through delegation,**" G. Bigwood, A. C. Viana, M. Boc, and M. D. de Amorim, INRIA, Research Report, 2010, rR-7361. [Online]. Available: http://hal.inria.fr/inria-00508273/en
[2] **"A set of measures of contrality based on betweenness,"** L. C. Freeman, Sociometry, vol. 40, no. 1, 1977
[3] **"Centrality in social networks conceptual clarification,"** L.C. Freeman, Social Networks, vol. 1, no. 3, 1979
[4] **"Hue scale representing node betweenness on a graph",** Claudio Rocchini, http://commons.wikimedia.org/wiki/File:Graph_betweenness.svg
[5] **"Computing connected dominated sets with multipoint relays,"** P. Jacquet, and L. Viennot,, INRIA, Tech. Rep. 4597, 2002

UPMC SORBONNE UNIVERSITÉS
informatics mathematics Inria
ceal
University of St Andrews