Part 3. Research Themes

- Social-based Communication
- Epidemiology
- Complex Networks
- Human Mobility
- Social Phenomena
- DTN Capacity

Impact of Altruism on Opportunistic Communication

- Opportunistic Communication
  - Largely relies on human as relays
  - Problems (battery life)
  - Requires altruistic behaviour
- Human Communication Pattern
  - Community-biased
- Findings
  - Opportunistic networks are very robust to the distributions of altruism

Methodology

- Simulation of asynchronous messaging on Static social network topologies

Community generation model	ER	SF	Simple	Caveman	NP	Kumpula
Number of nodes	1000	1000	1000	1000	991	1000
Number of edges	4492	4509	4492	4500	4503	4511
Diameter	6	4	7	10	21	7
Clustering coefficient	0.0105	0.3991	0.3487	0.7399	0.1136	0.2965
Max degree	19	634	17	13	28	33
Number of communties	8	13	28	48	4	17
Average size of communities	125	76.9231	35.7143	20.8333	247.75	58.8235

- Emulation on real human mobility traces

Experimental data set	Cambridge	Infocom05	Infocom06	Reality
Device	iMote	iMote	iMote	Phone
Network type	Bluetooth	Bluetooth	Bluetooth	Bluetooth
Duration (days)	11	3	3	246
Granularity (seconds)	600	120	120	300
Number of experimental devices	54	41	98	97
Number of internal contacts	10,873	22,459	191,336	54,667
Average no. of contacts/pair/day	0.345	4.6	6.7	0.024

- Performance metrics
  - Successful delivery ratio
  - Delivery delay

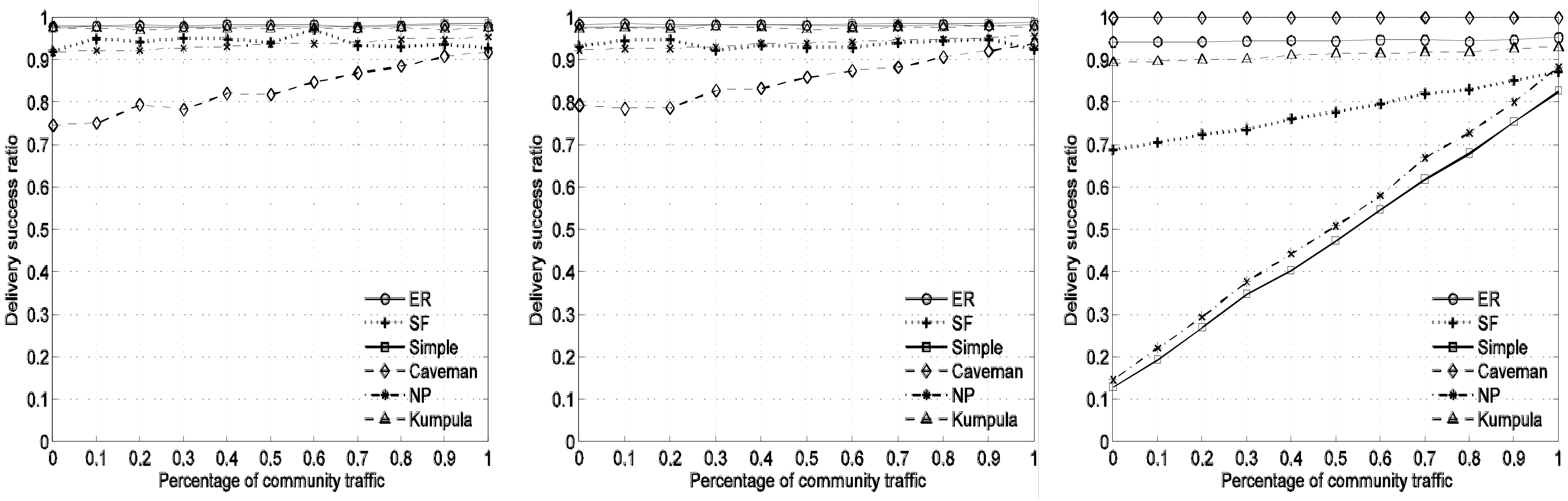
Altruism and Traffic Models

- Altruism Models
  - Percentage of selfishness
  - Uniform, normal
  - Geometric $P(X = k) = (1 - p)^{(1-k)} p$
  - Community-biased $P_{intra} + P_{inter} = 1$
  - Degree-biased

$$a_i = \frac{(k_i - k_{min})^\alpha}{(k_{max} - k_{min})^\alpha} \quad \text{with} \quad \alpha > 0 \qquad a_i = \frac{(k_{max} - k_i)^\alpha}{(k_{max} - k_{min})^\alpha} \quad \text{with} \quad \alpha > 0$$

- Traffic Models
  - Uniform
  - Community-biased

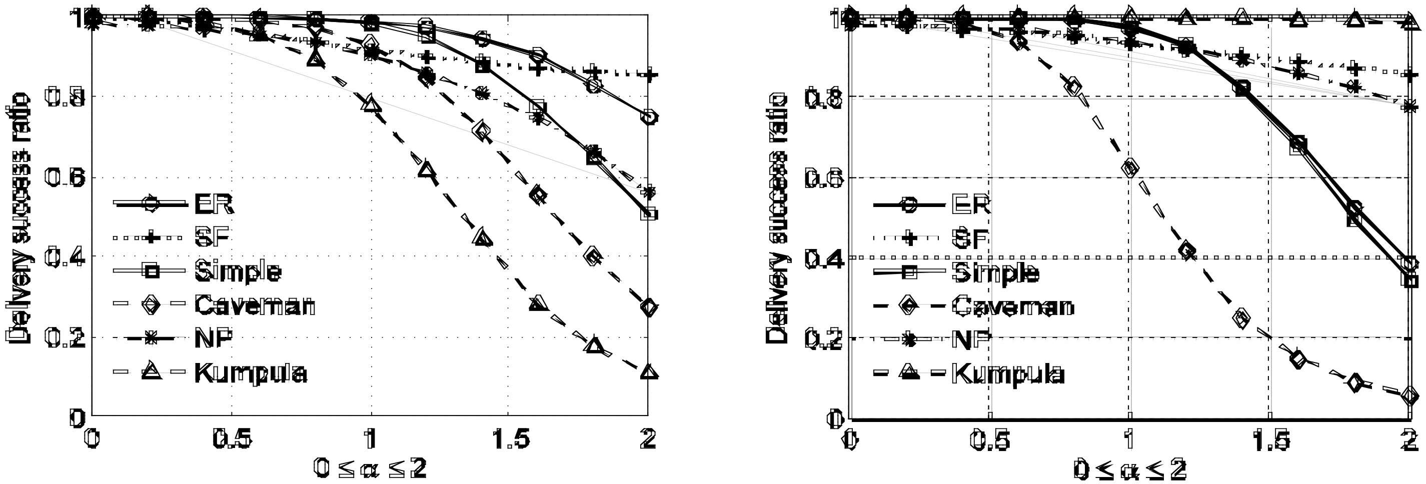
Results and Evaluations



(a) Uniform (b) Normal (c) Geometric (p = 0.5)

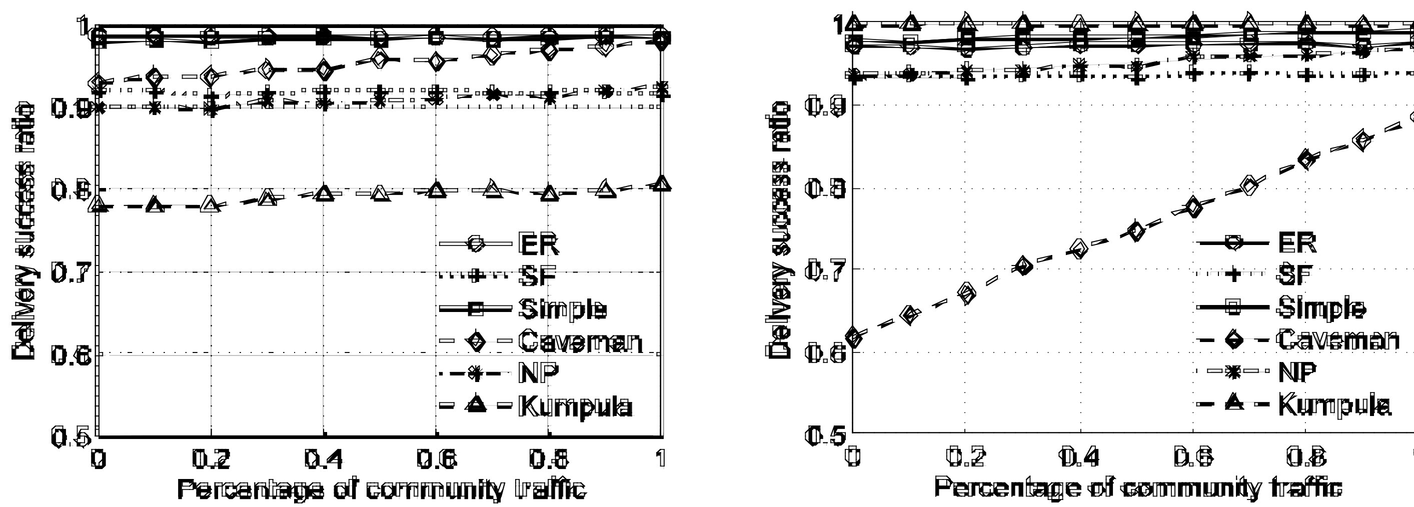
Fig. 1. Delivery ratio with percentage of community-biased traffic using different altruism distribution

Results and Evaluations (2)



(a) High-degree-biased (b) Low-degree-biased

Fig. 2 Delivery ratio with parameter alpha using uniform traffic



(a) High-degree-biased (b) Low-degree-biased

Fig. 3. Delivery ratio with percentage of community biased traffic using degree biased altruism distribution

Results and Evaluations (3)

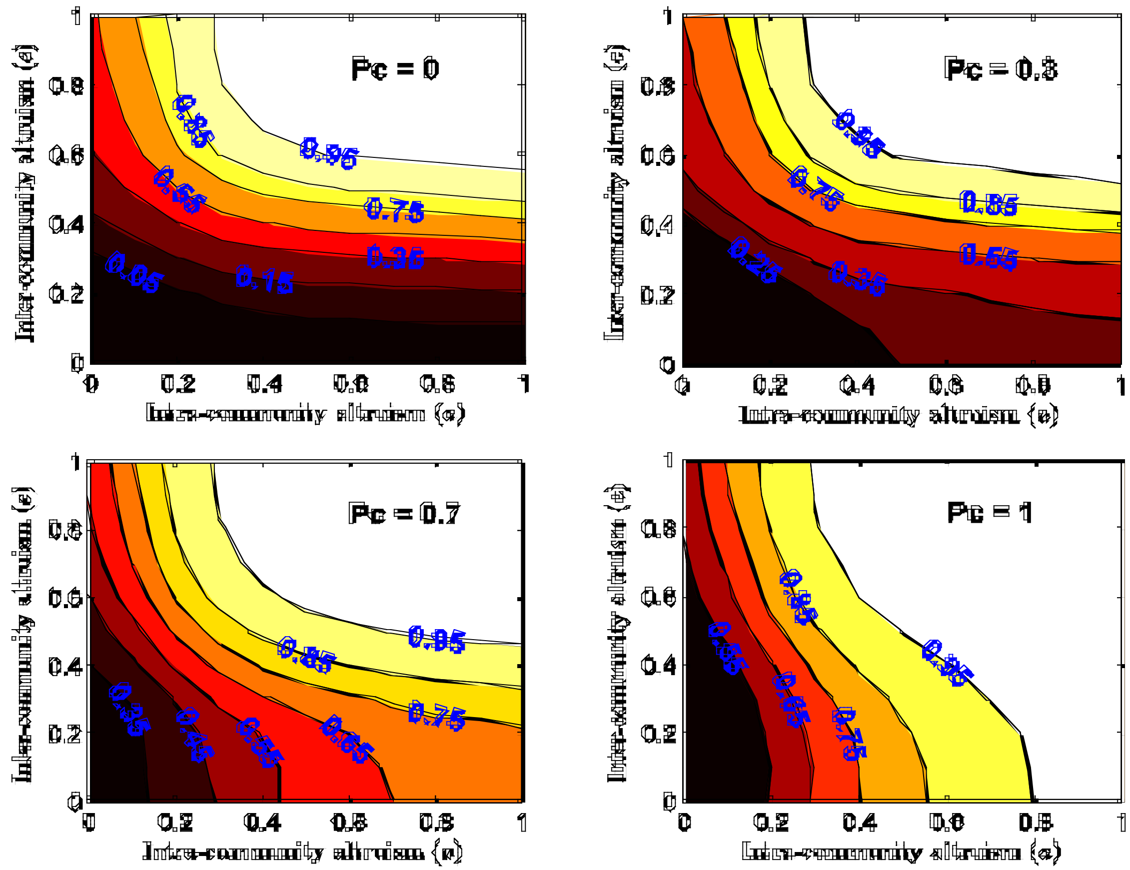
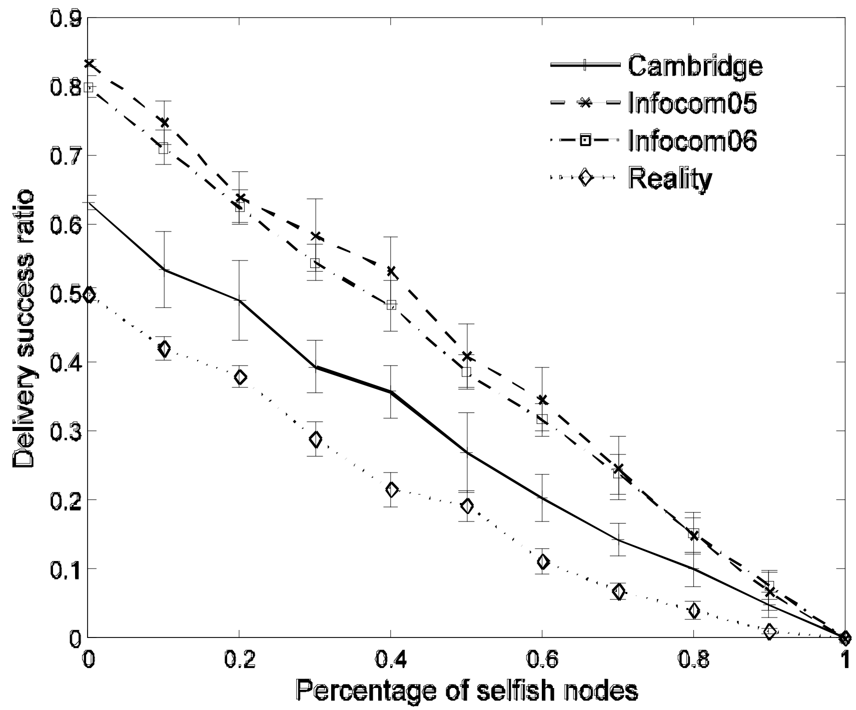


Fig. 4 Delivery ratio of the Caveman model with varying intra- and inter-community altruism

Results and Evaluations (4)



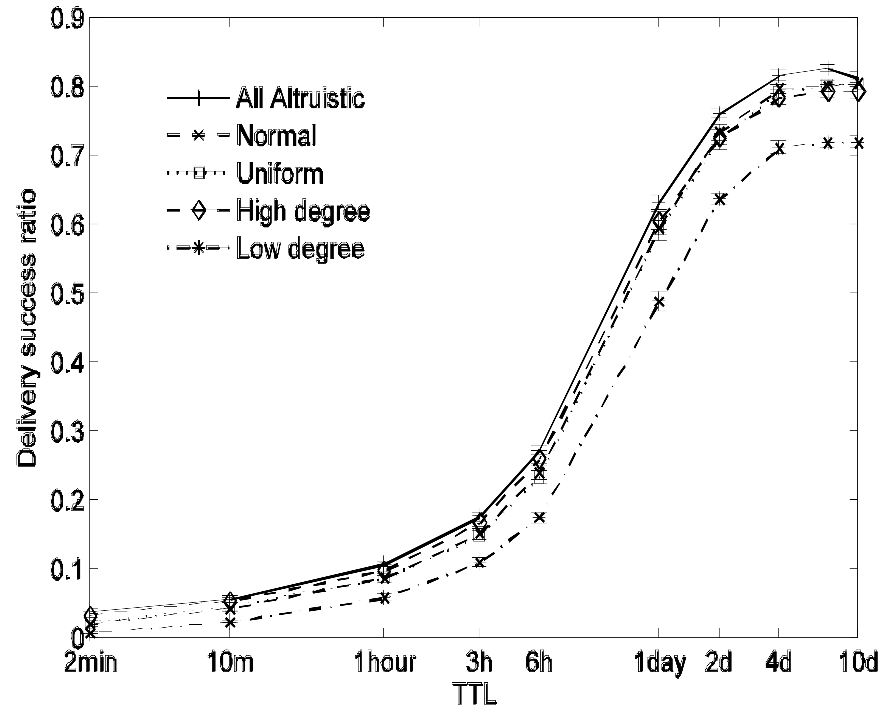


Fig.5. Delivery ratio of four datasets with percentage of selfish nodes

Fig.6. Delivery ratio with uniform traffic on Cambridge data

Conclusions & Future Work

- Opportunistic communication and information dissemination in social networks are very robust toward altruism distribution due to their multiple forwarding paths

- Altruism values resulting from gaming strategies

- Feedbacks from previous delivery histories

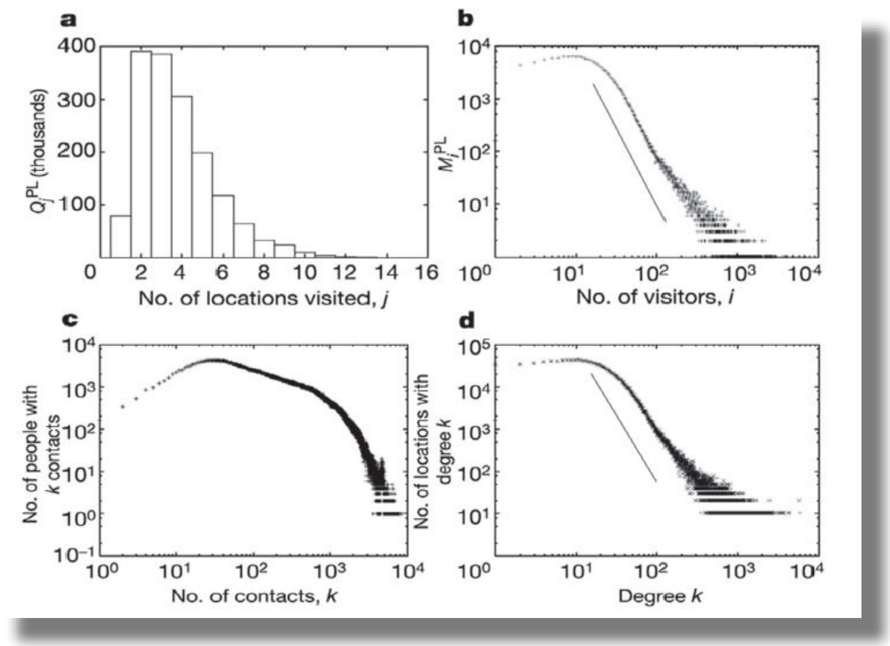
- Power consumption of nodes, and how it is affected by the altruism

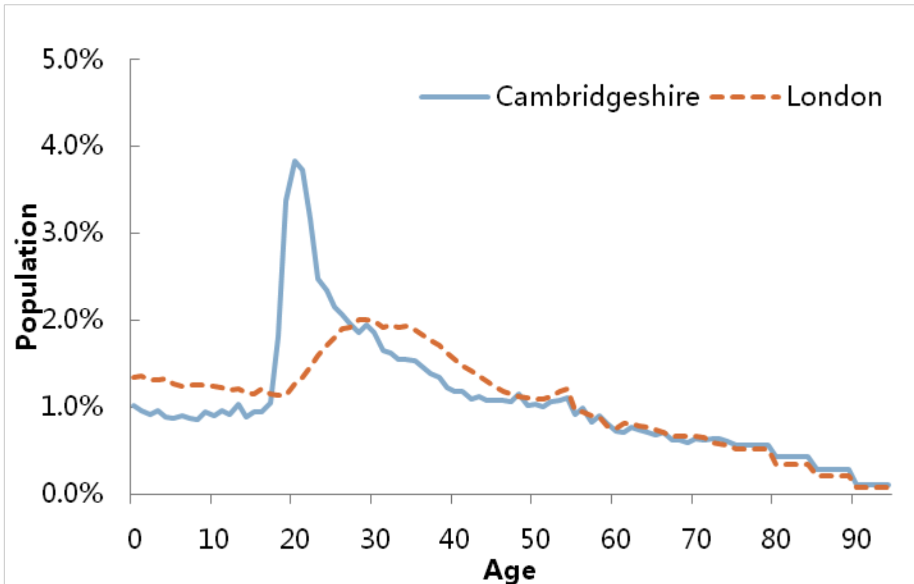
Epidemic Modeling (LNCS ACRI 2008)

- Demographics
- Non-homogeneous population
  - Age Dependency
    - Influenza spread firstly in children aged 3-4 (Browstein)
    - Infectivity/Susceptibility varies with age
  - Location Dependency
- Human Mobility
- Social Containment Strategies
- Vaccination Strategies

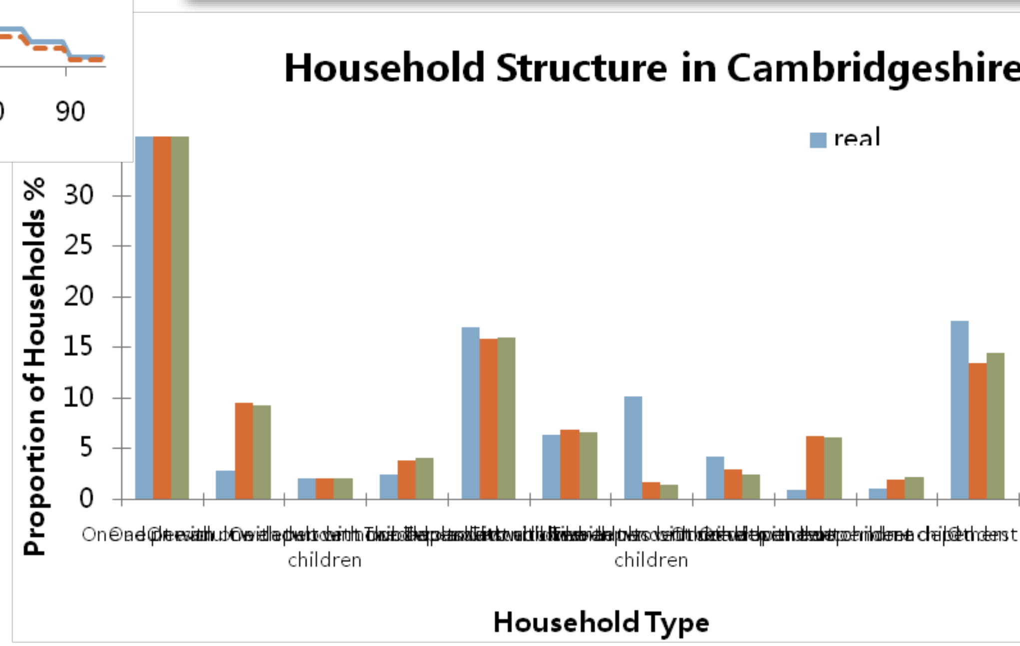
Demogr

*Nature [0028-0836] Eubank (2004)

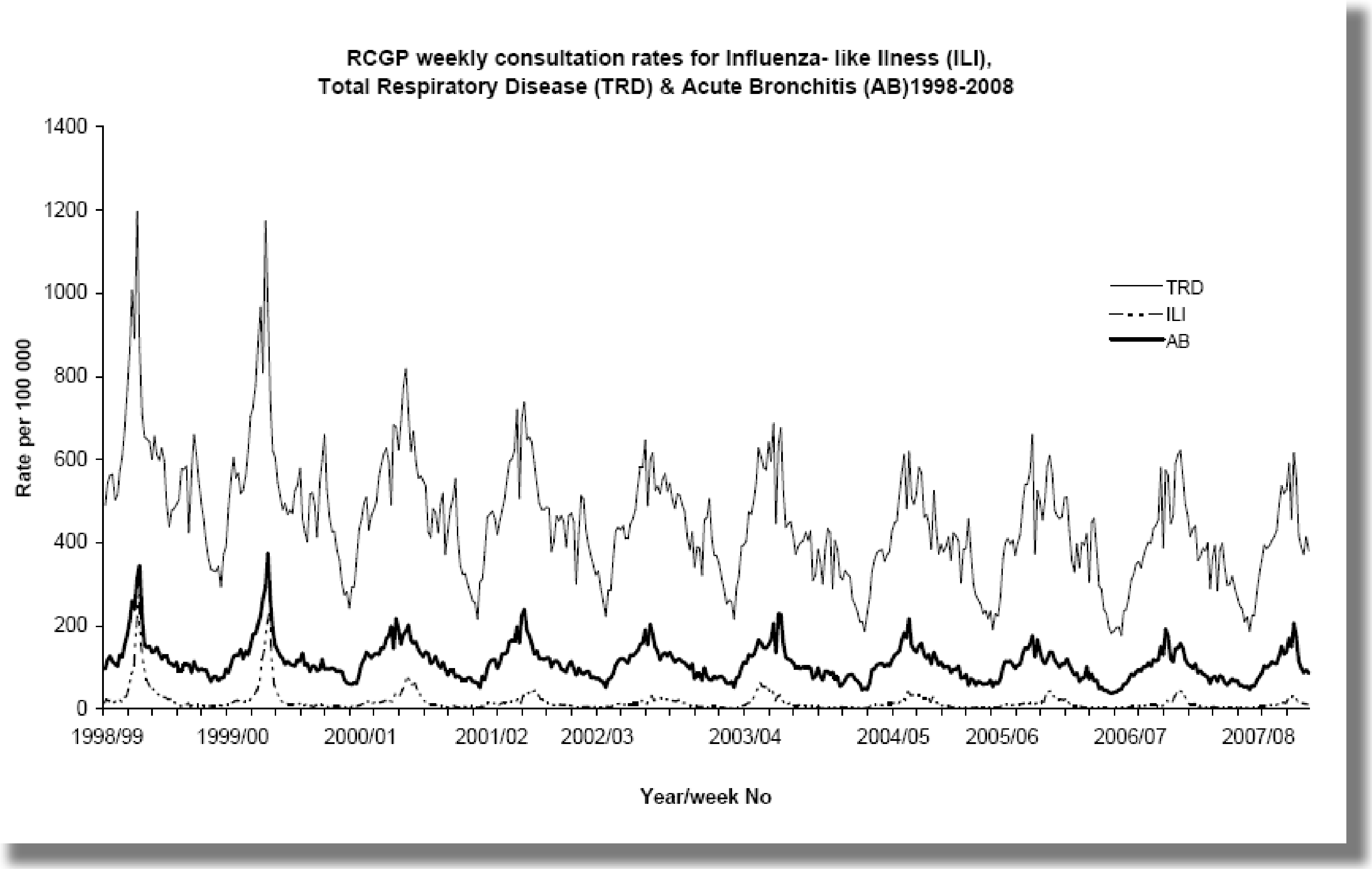




*Statistics.gov.uk



Seasonal Factor



*UK Health Protection Agency (www.hpa.org.uk)

Model

If person i goes to location k, an edge T_{ik} is drawn. We could also extend this definition by weighing the link with the time spent in the location, and considering time coincidence. However, due to the lack of precise data, we just consider the average one-day window, and set $T_{ik} = 1$ if the person is expected to spend more than one hour in that location. The effective social contact network J_{ij} is then obtained as

$$J_{ij}(k) = T_{ik}T_{jk}. \tag{1}$$

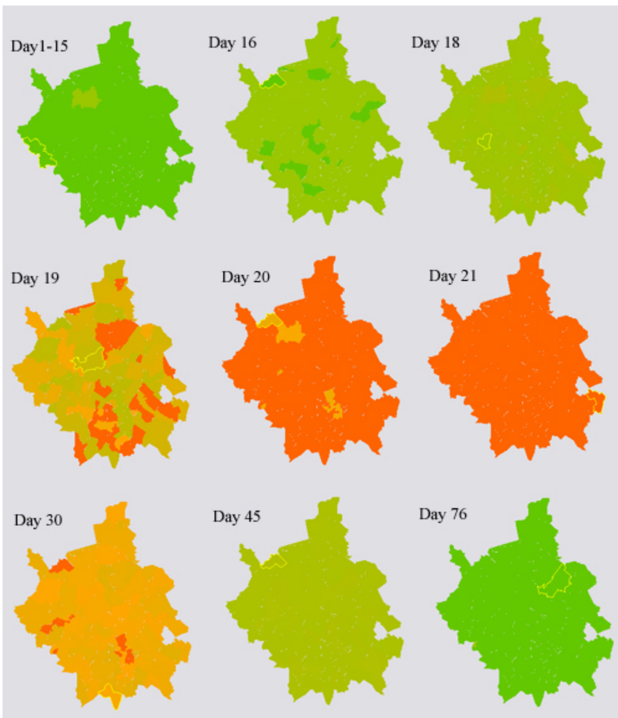
probability of a susceptible becoming exposed, we first define the notion of pairwise propensity of infection $Q(i,s)$ between a susceptible of age s and infective of age i:

$$Q(i,s) = \mathrm{Inf}(i) \cdot \mathrm{Sus}(s) \cdot V \cdot T, \tag{5}$$

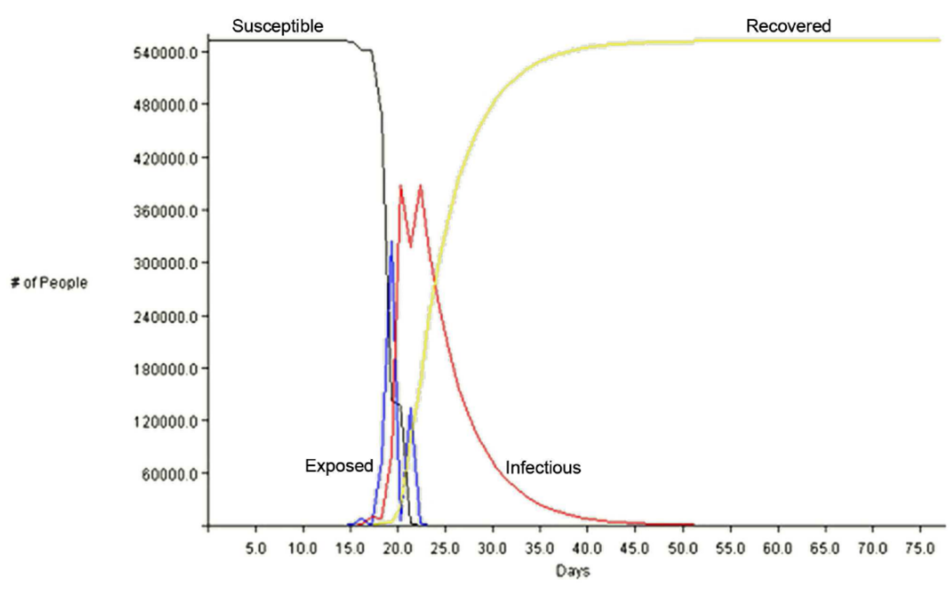
V are assumed to be constant. To estimate the overall probability $P_{E|S}(s,l)$ of susceptible of age s catching the disease in the location l, we take the normalised summation of the above pairwise infection propensity:

$$P_{E|S}(s,l) = \frac{\sum_i J_{is}(l) \cdot Q(i,s)}{\mathrm{Deg}(l)} \tag{6}$$

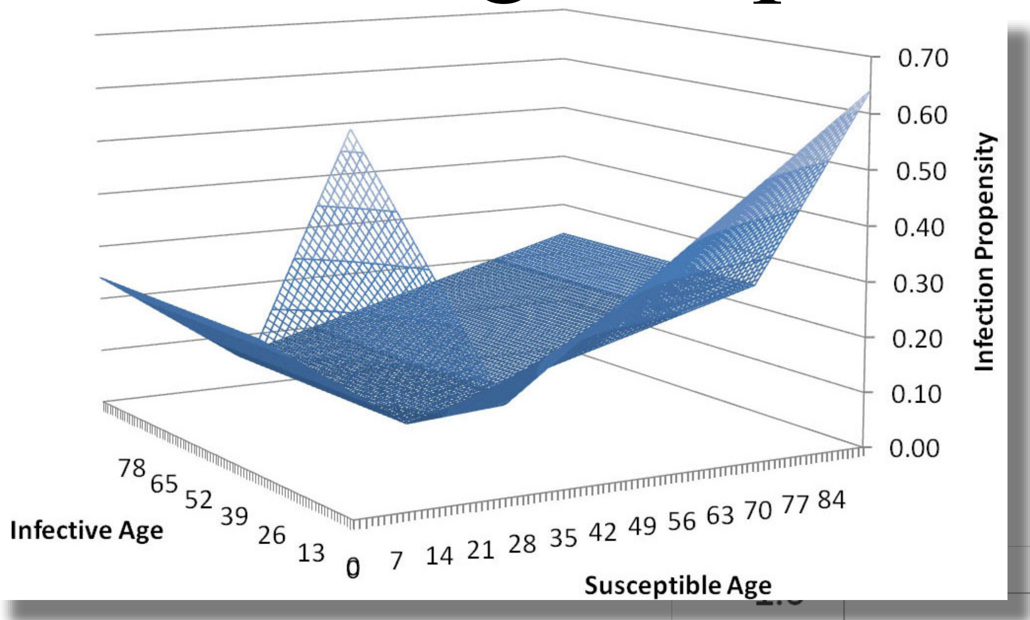
Epidemic Spreading

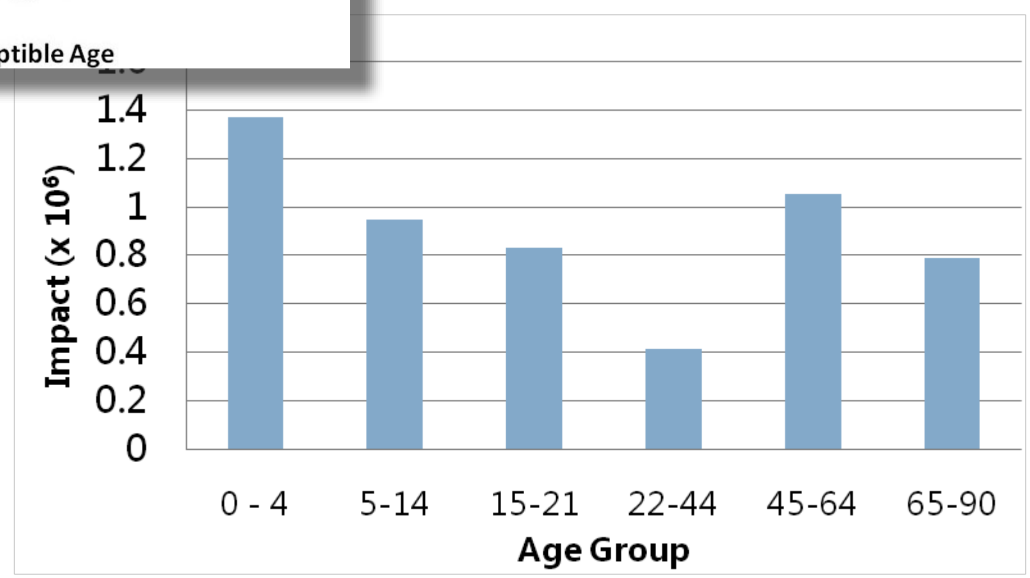


- Parameters:
- External Model based on population density and travel statistics

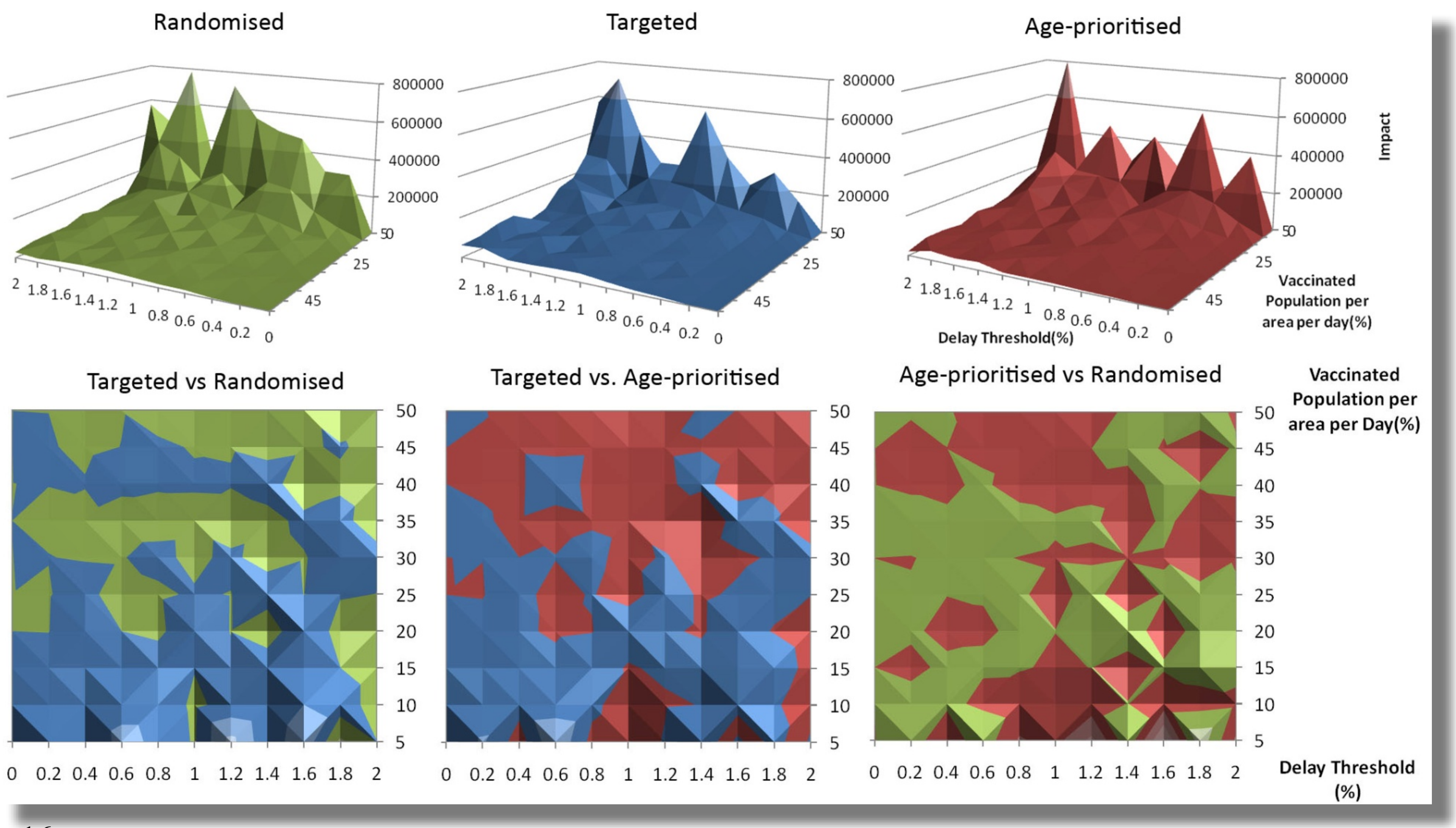


Age Dependency





Vaccination Strategies



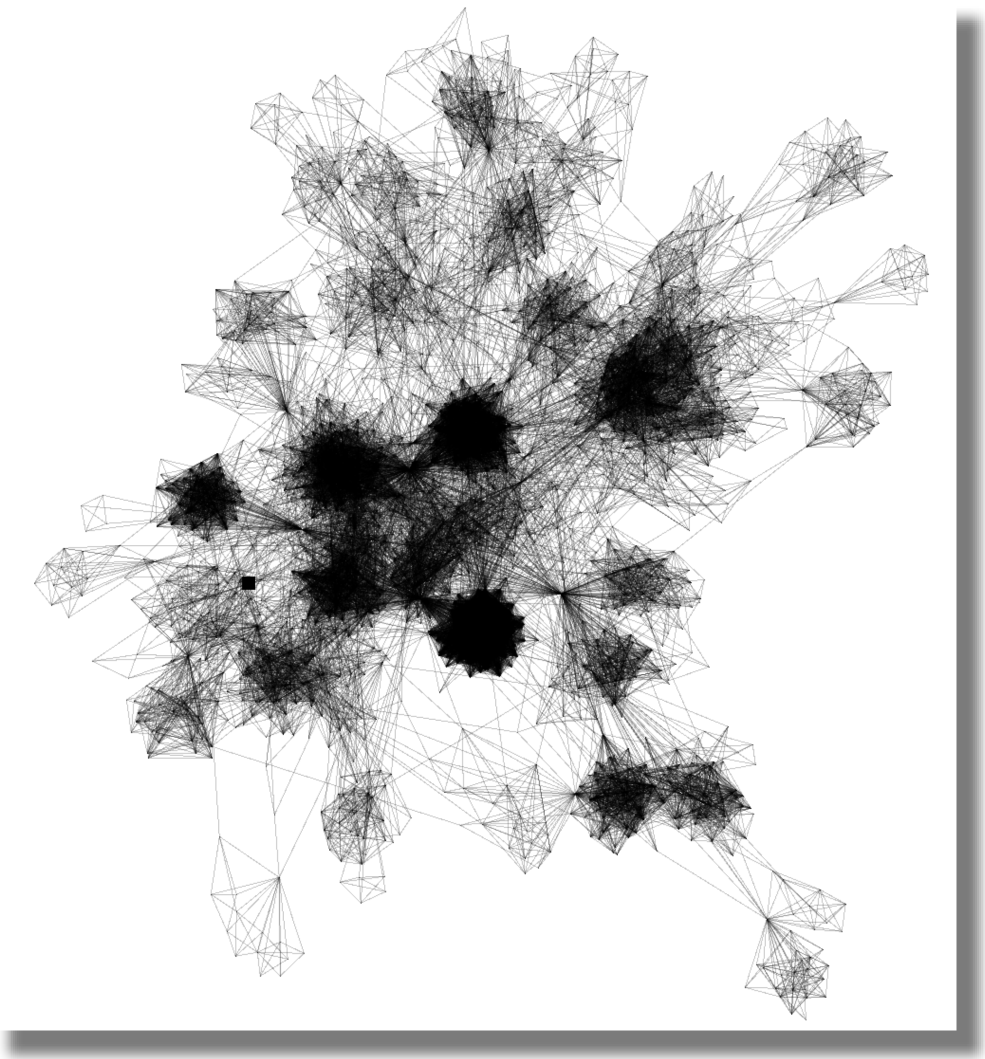
Future Work

- More realistic Contact Network
  - Time/Age dependency
  - Human Dynamics/Mobility Model
  - Probabilities for infection
  - Network Parameters
  - Non Co-location Infection
  - Demographical Information
- Social Containment Policies
- Travel Patterns
- Death/Birth Rate

Complex Networks

- Community Detection in Large Networks
  - Static Social Networks
  - Time Dependent/Dynamic Networks
- Novel Network Characteristics
  - Recurrence Plots
- Novel Network Models
  - Time dependent networks

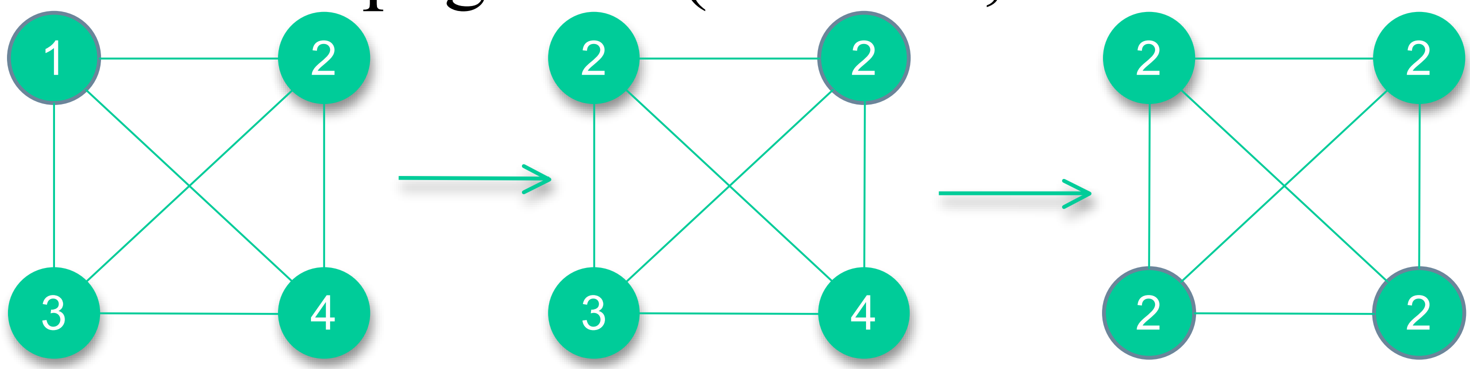
Community Detection in Large Social Networks



- Orkut Online Social Network (3M nodes, 0.2B edges)(Mislove et. Al.)
- Largest available network (1B edges)
- Traditional Modularity-optimization approach is slow
- Aim : Real-time
  - Linear
  - Accurate

Real Time Community Detection (in submission to PRE)

- Label Propagation (Nandini, Albert et. Al.)

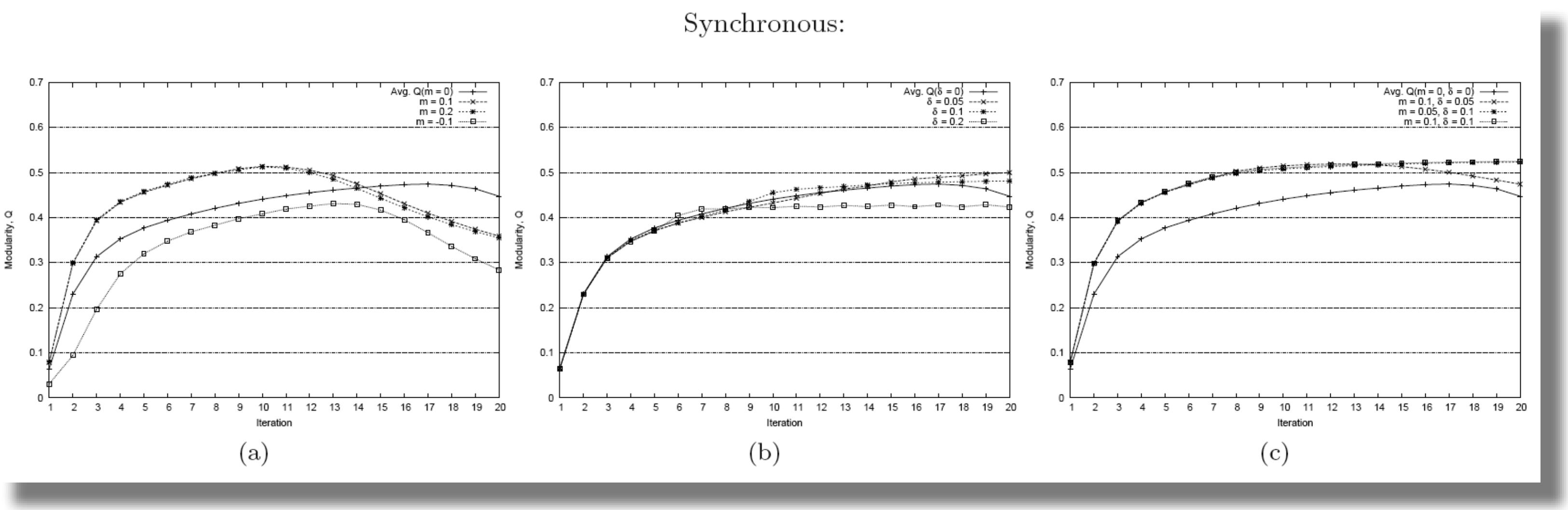


- Gene

$$l'_p = \underset{l}{\operatorname{argmax}} \sum_{p' \in n(p)} s(l_{p'}, p') \cdot f(p')^m \cdot w(p', p)$$

$$s'(l'_p, p) = \left(\max_{p' \in n(p, l'_p)} s(l_{p'}, p') \right) - \delta$$

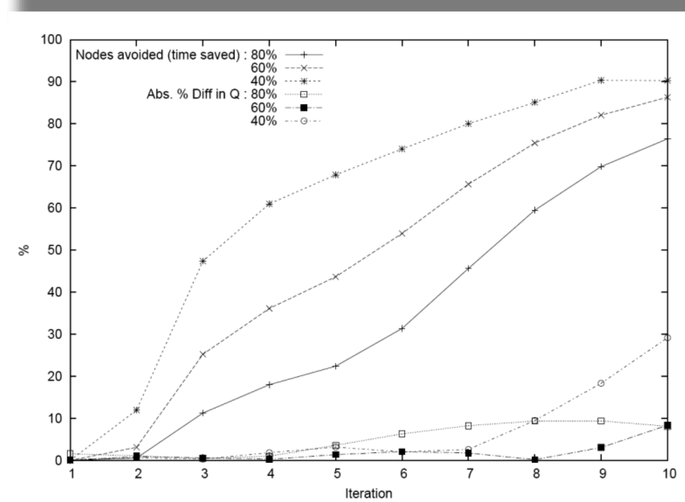
Real Time Community Detection



- Asynchronous v. Synchronous
- m, Label Preference
- ∂, Hop Attenuation

Real Time Communi[redacted]

- Linear Running Time (almost) – $O(|E|)$
- Accurate
  - 5% within state of the art
- Adaptive
- Exploitable:
  - Localised Metric – Easily Parallelised
  - Avoid updating nodes well inside the community



Future Work

- Static Network (OSN, Amazon, Google...relatively speaking) can be boring
- Mostly done offline
- Community Detection in **Dynamic** Networks in **Real-Time**
- Idea: Consider moving spin glasses
- Can apply a similar paradigm as the basis for the detection
- More realistic mobility model

Summary

- Altruism in Opportunistic Communication
- Contact Network Modeling of Epidemics
- Real-Time Community Detection in Large Networks
- ...

Thank You!

Selfish DTN capacity Experimental estimation

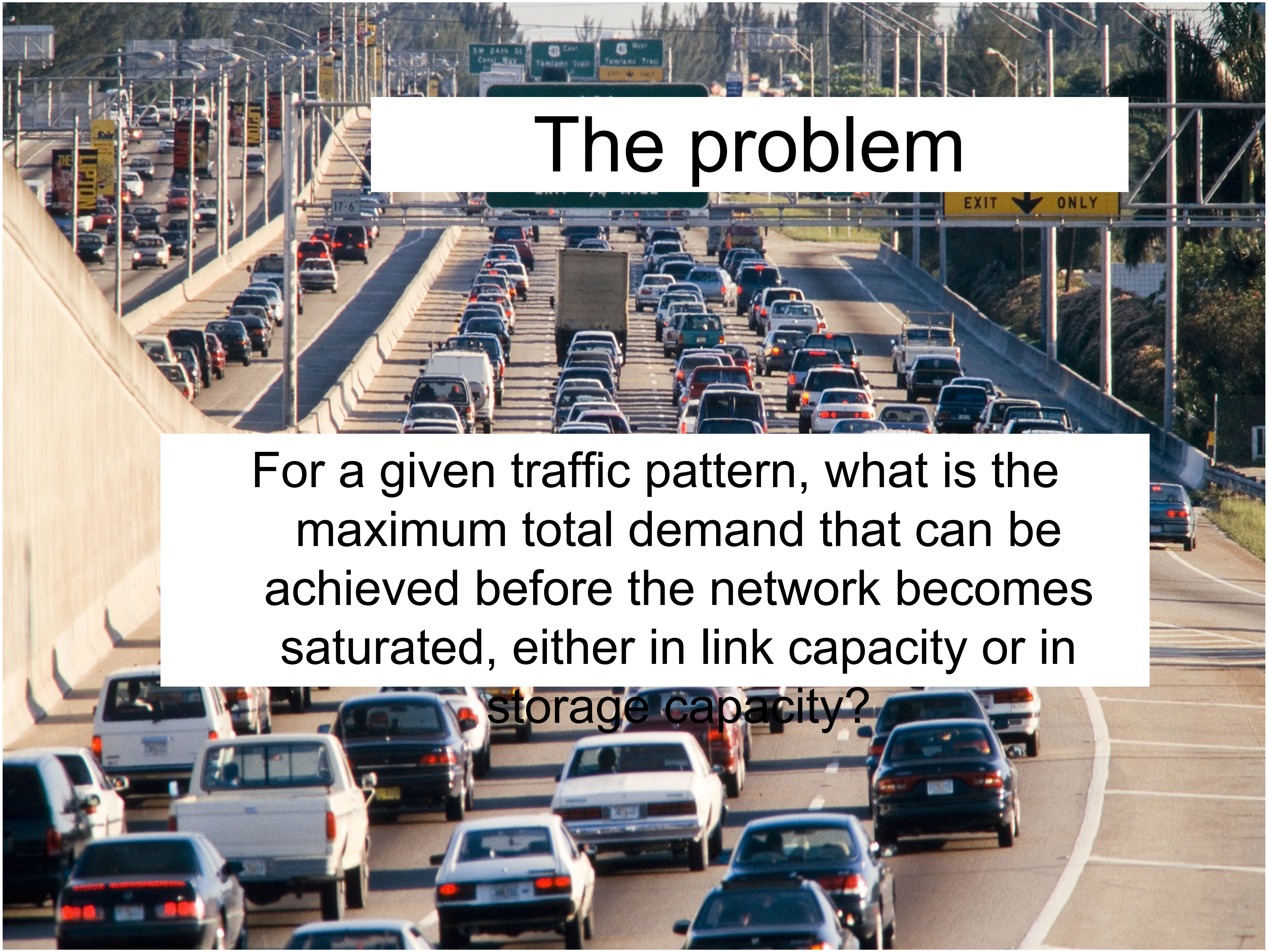
Dagstuhl Seminar DTN II
February 2009

Joint work of
PU Tournoux, V. Conan, J. Crowcroft,
J. Leguay, F. Benbadis, M De Amorim

DTN performance

- Measured in terms of packet delivery ratio, packet delay for non-congested networks
- Performance is limited by:
  - Storage constraints
  - Transmission constraints
  - Node selfishness

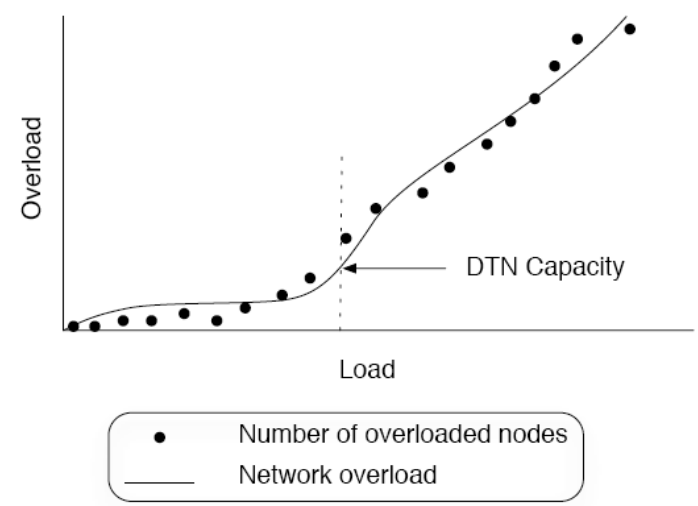
The problem



For a given traffic pattern, what is the maximum total demand that can be achieved before the network becomes saturated, either in link capacity or in storage capacity?

Capacity estimation

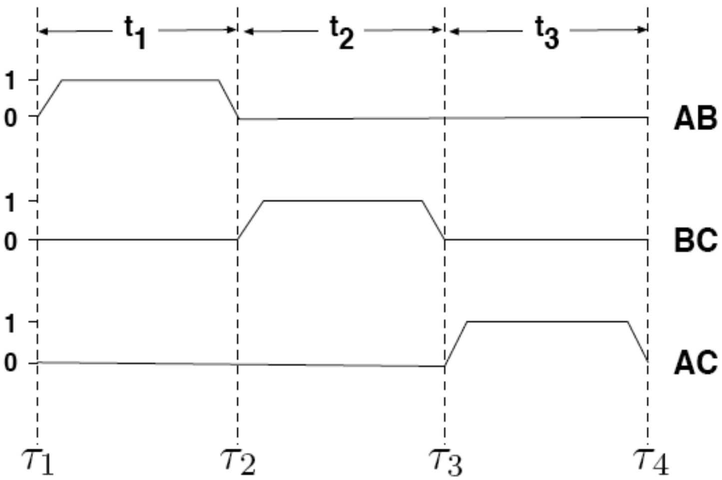
- Use traces of node contacts as input
- Consider fixed demands
  - One demand = transfer x Kbits from node s to node d injected at time t
- Traffic is routed at Wardrop equilibrium
- Increase x for all demands
- Stop when storage/trans mission overload increases faster than load

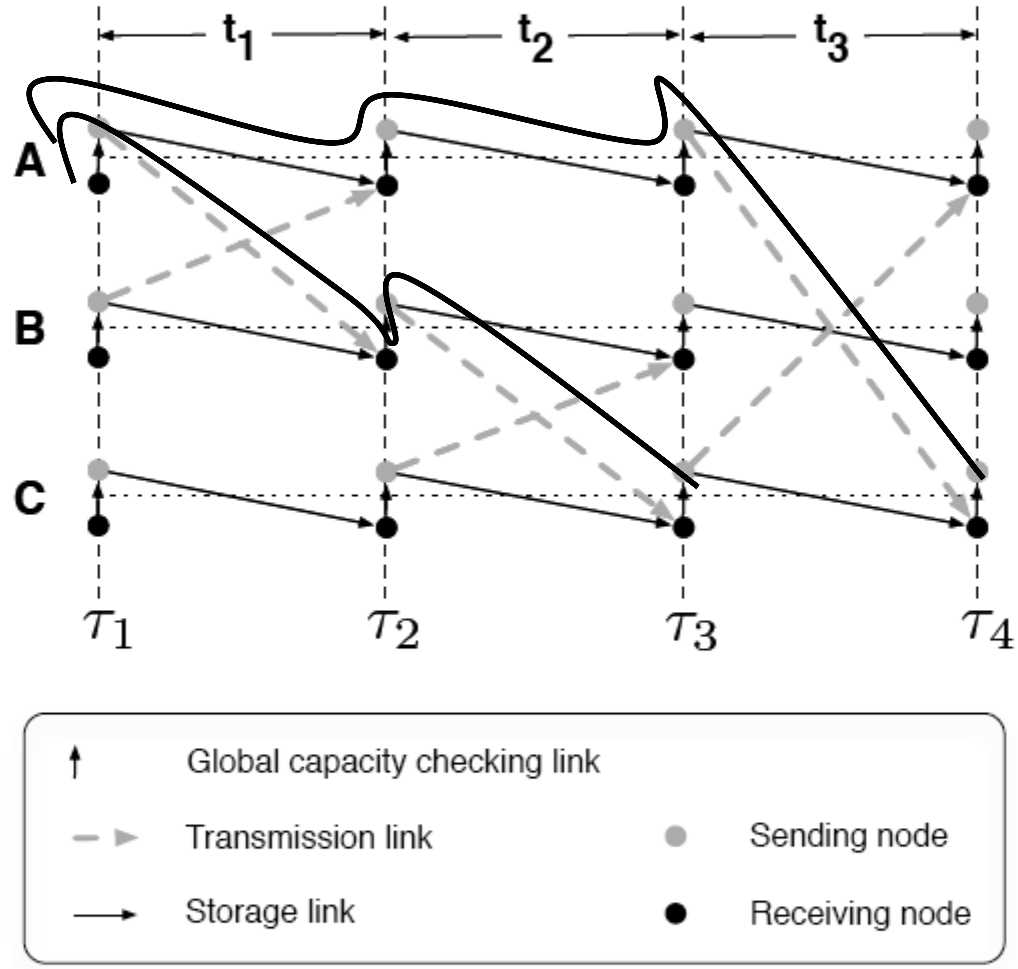


Wardrop equilibrium

- *The journey times (or costs) in all routes actually used are equal and less than those which would be experienced by a single vehicle (or message) on any unused route*.
- For additive and separable positive and continuous cost functions the Wardrop equilibrium is a convex minimization problem. The equilibrium is unique, and may be obtained by running the Frank-Wolfe algorithm

Time Discretized Graph



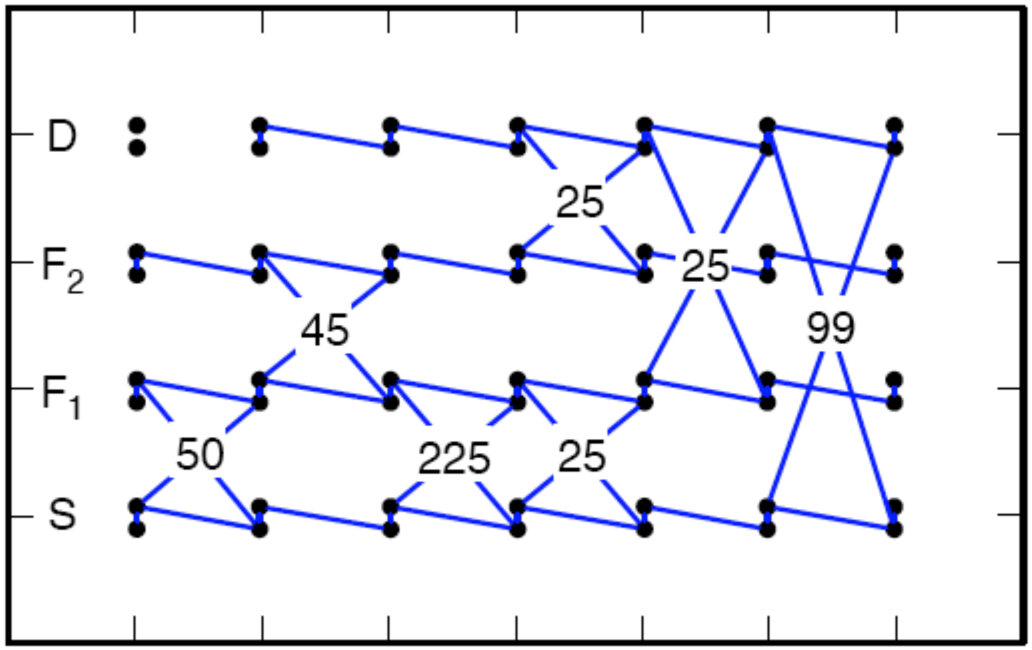


$$C(e \in \mathcal{T}) = \Delta w_\Delta + f(e) w_T + \frac{f(e)^p}{(C_T(e))^{p-1}}, \quad (5)$$

$$C(e \in \mathcal{S}) = \Delta w_\Delta, \quad (6)$$

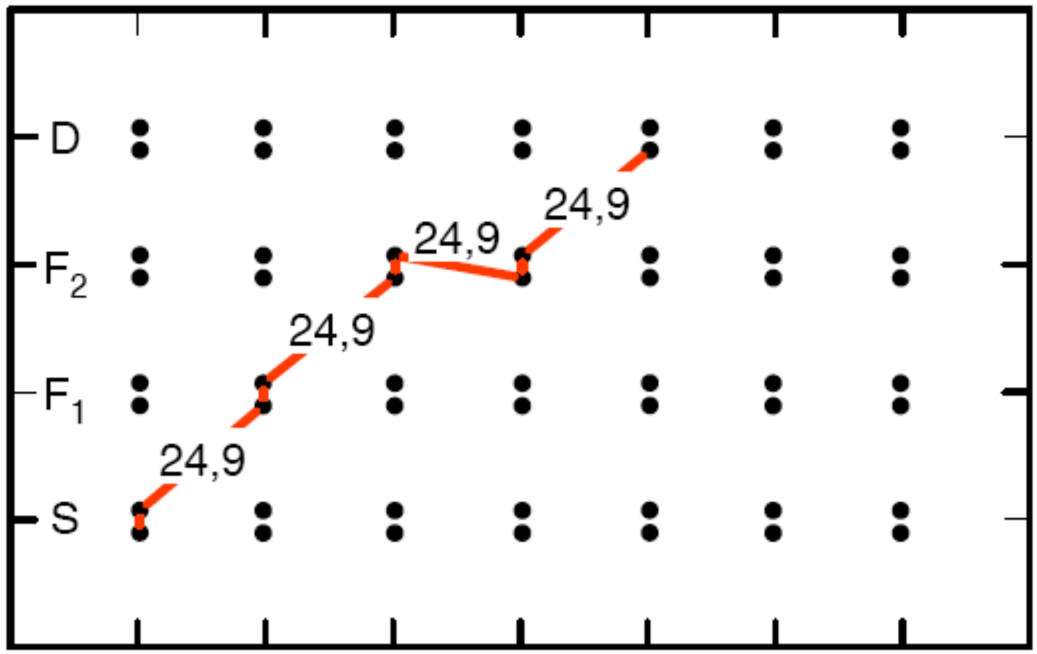
$$C(e \in \mathcal{C}) = \frac{f(e)^p}{(C_S(e))^{p-1}}. \quad (7)$$

Example



Transmission capacities of connections

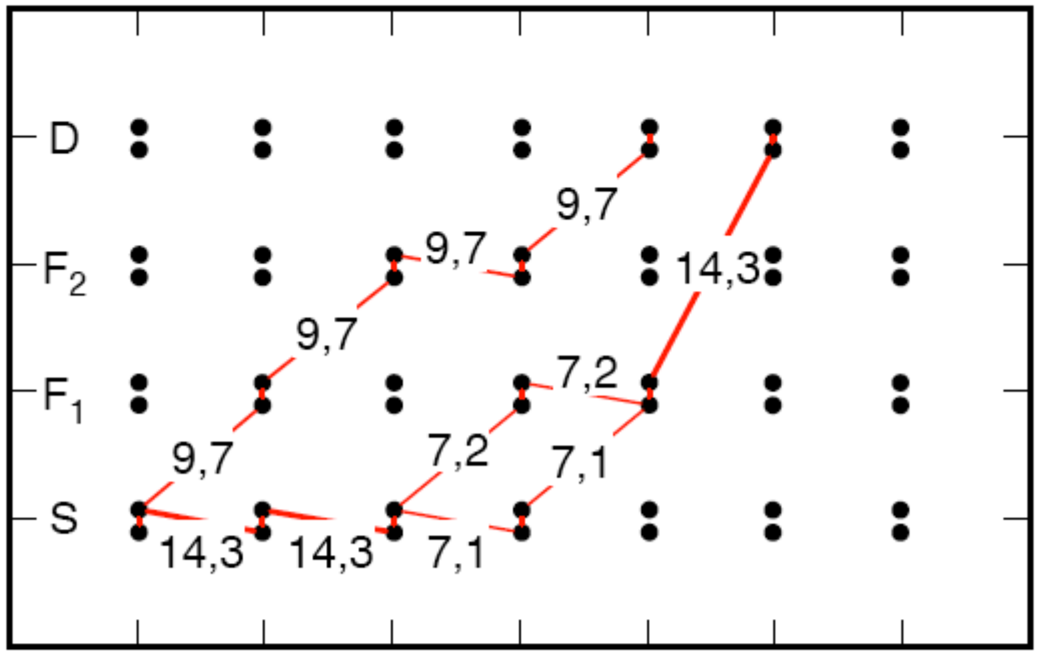
Fastest route



Weight on delay = 1
Weight on transmission cost = 1

- Delay is costly: the fastest route is chosen

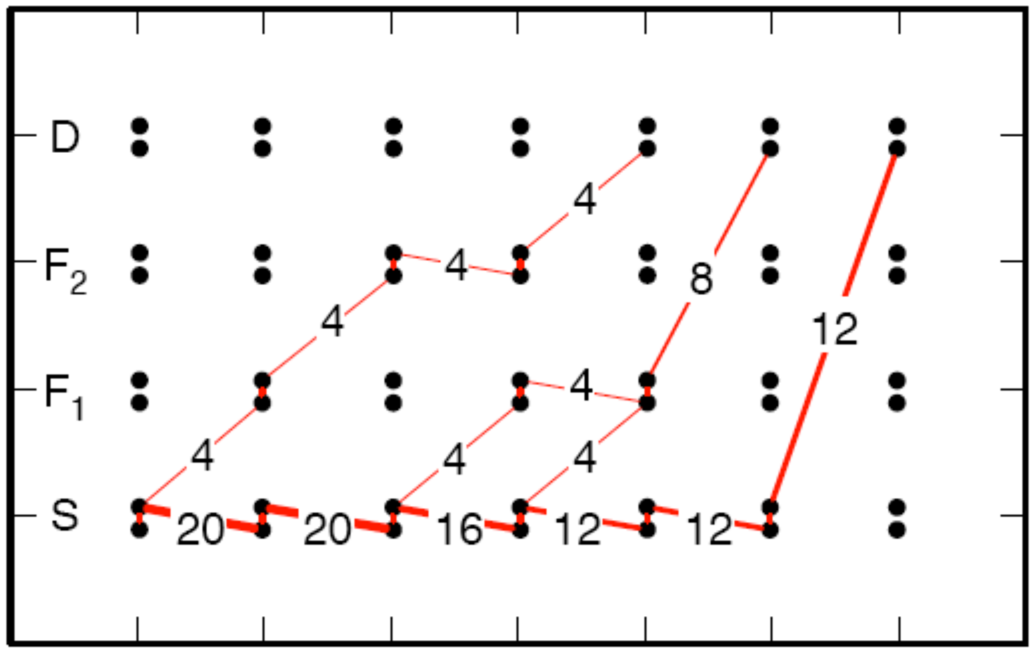
Costly transmissions



Weight on delay = 1
Weight on transmission cost = 30

- Increasing transmission cost forces usage of slower routes

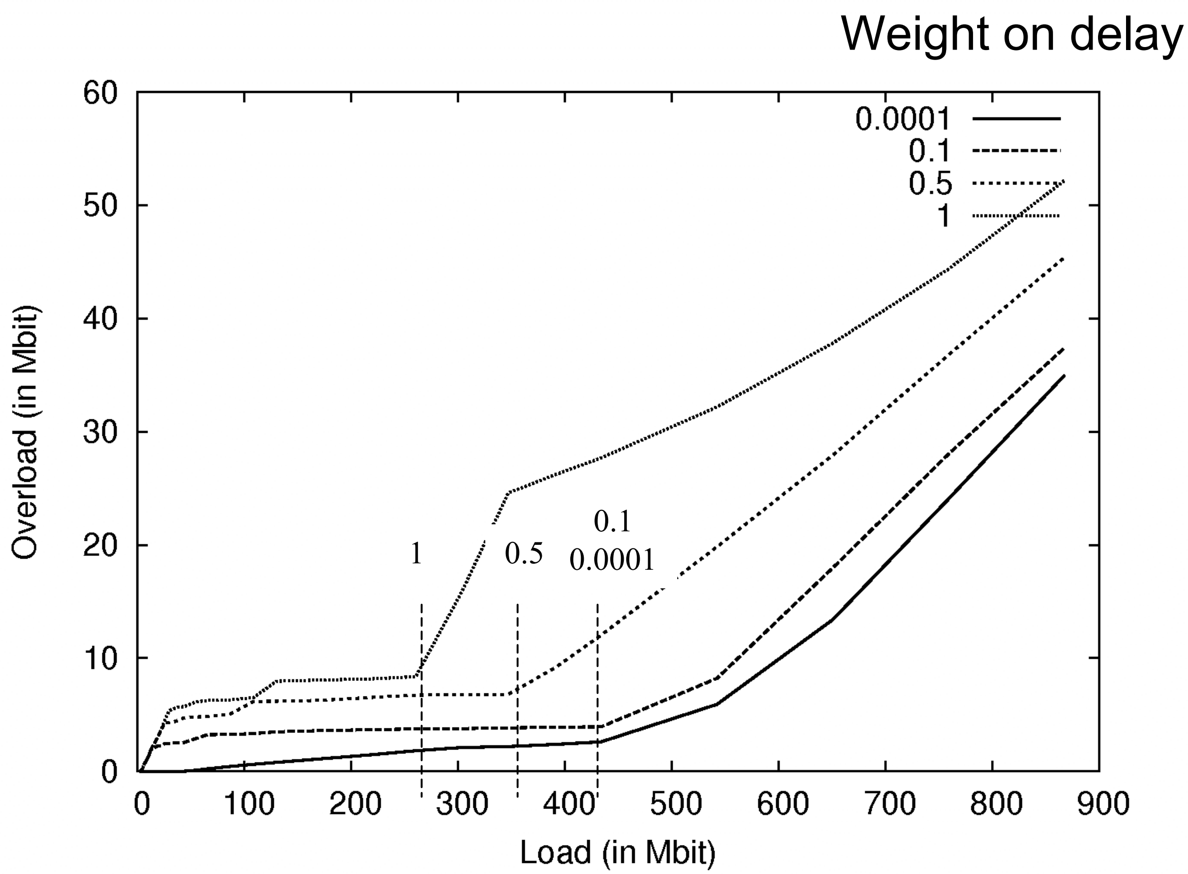
More delay



Weight on delay << 1
Weight on transmission cost = 30

- Allowing for more delay forces usage of even more slow routes

Infocom data set



- Capacity increases with delay tolerance

Conclusions

- Delay-tolerance helps increase capacity.
- Routing protocols should be multi-dimensional.
- Selfishness makes network engineering challenging.