CREATE-NET
via Solteri 38, Trento
Italy

“Pervasive communication environments: the BIONETS perspective - 2nd part”

F. De Pellegrini

fdepellegrini@create-net.org

Autonomous Computing
in Smart Environments

University of Fribourg, 23 - 24 November 2006

Bionets overview

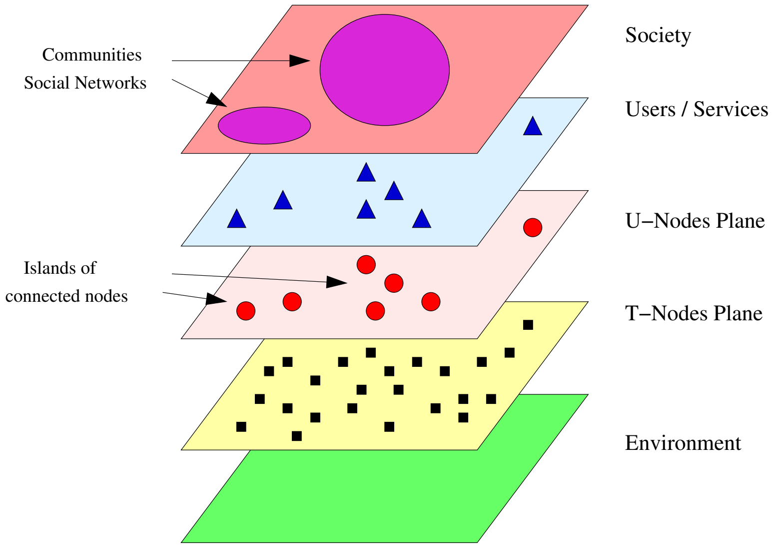
FET EU project funded within the Sixth Framework Programme (FP6)

- BIOlogically inspired NETwork and Services = BIONETS
- Funding Agency: European Commission - Future and Emerging Technologies
- Area: Situated and Autonomic Communications
- Instrument: Integrated Project
- Project reference: FP6-027748
- Duration: 48 months
- Web: http://www.bionets.org/

Bionets overview

- Pervasive computing and communication environments: *extremely large number of networked embedded devices possessing sensing/identifying capabilities: user-situated services to interface directly with the surrounding environment*
- Target [WS]: *“The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.”*
- Traditional communication approaches are ineffective in this context
  - wide *heterogeneity* in node capabilities and service requirements
  - huge number of nodes with consequent *scalability* issues
  - possibly high node mobility and the management *complexity*.
- BIONETS aims at solving such problems through: *network architecture* and *service definition*

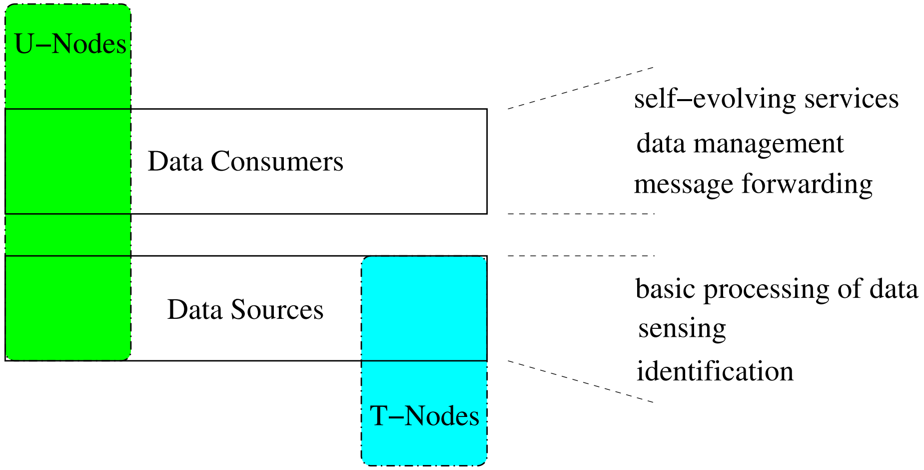
Bionets overview



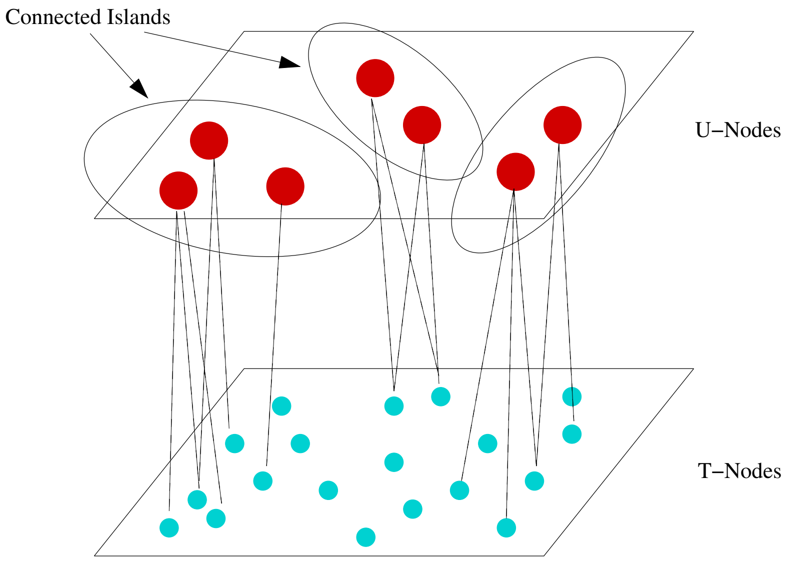
- First principle: information diffusion and filtering
  - Aim: replacing the conventional end-to-end Internet approaches with localized service-driven communications
  - Meta: relaxing connectivity requirements in order to improve scalability
- Second principle: bio-inspired platform, centered around the concept of *evolution*
  - Aim: support of autonomic services
  - Meta: solve complexity in the system management

Bionets overview: network architecture

- Elements of BIONETS Networks: 3 main actors
- Aim: minimal requirements in processing/storage/communications
- **T-Nodes** simple, inexpensive devices with sensing/identifying capabilities: interface with the environment. T-Nodes do not communicate among themselves; just "read" by U-Nodes passing by.
- **U-Nodes** complex, powerful devices (PDAs, laptops and smartphones) carried around by users and run services.
- **Access Points** gateways with the IP world



Bionets overview: network architecture



- 2 tier network architecture
- Lower tier of T-Nodes cheap, tiny devices (=sensors, tags and RFIDs). No store-and-forward operations.
- Upper tier of U-Nodes (= devices running services)
  - computing/communication capabilities
  - exempt from energy consumption issues
  - mobile in nature
  - U-Nodes communicate with U-nodes and T-Nodes
  - U-Nodes are “sinks” for sensor-gathered data and sources/relays/sinks for communication among U-Nodes

Inherent tradeoff: delay versus memory requirements

Bionets overview: evolution

- Complexity: centralized network management is not scalable with the size foreseen with pervasive environments
- Also: we already gave up connectivity, so that even service management should be localized
- Ex. Guidance system: cannot report to the central unit; but somehow should be able to adapt to *local conditions*
- BIO-inspired solution: services which *adapt* and eventually *evolve*
- Seed for the BIONETs project: evolutionary services are modeled after *living organisms* which are *user-situated*; hosted on U-Nodes and move through the physical movement of the users through the network, which is the *habitat* of such moving organisms
- Nice idea, but ... how do you implement that? How do you *engineer* this stuff?
- The final prototype still not there, but we have some idea of how to get there ...

Diffusion

Diffusion

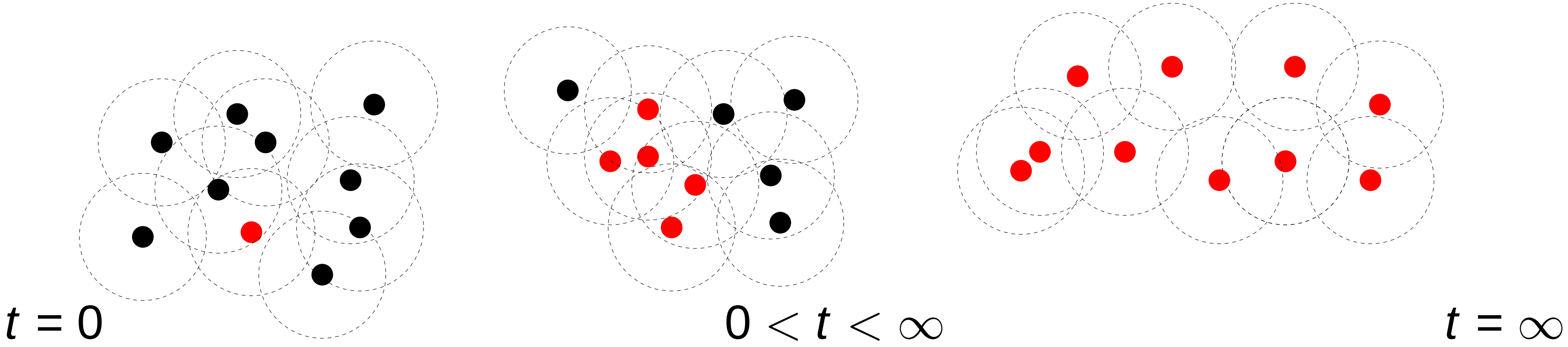
- Diffusion at the U-nodes layer uses *opportunistic* and *disconnected* operations
- Exactly at the opposite end of QoS guaranteed traffic
- Delay Tolerant Networks or Disruption Tolerant Networks (DTNs) show similar features
- Connectivity is not there all the time *intermittent connectivity*
- Examples: Interplanetary Internet, Message ferrying systems [ZeguraMobihoc06], Systems of personal devices on a metropolitan scale [leguaychants06]
  - Interplanetary Internet/ Satellite links [burleighcommag03]: orbits drive the presence/absence of line-of-sight
  - Use of buses and other public transportation means for carrying and disseminating information [burgessinfocom06,leemobeyes06]
- Applications: diffuse information about traffic situation, parking availability, special events (conferences, fairs etc.), local advertisement, video-surveillance ...

Diffusion

- When connectivity cannot be taken for granted, need of architecture able to handle disconnected operations [fallsigcomm03,dtnnetarch]
- V. Cerf et al. "Delay-Tolerant Network Architecture", IETF Internet draft, work in progress, Mar. 2006.
- Several experiments in DTNs carried out recently
  - DieselNet DTN network: 30 buses from UmassTransit equipped with the DieselNet devices IEEE 802.11b access points, one for use by the passengers and one to provide connectivity with other buses [burgessinfocom06]
  - MIT Reality Mining: experiment carried out a.y. 2004/05, approximately 100 people (faculty and students) at MIT; software for Nokia Symbian Series 60 Phone, Bluetooth interface [realitymining]
  - Intel-Cambridge Traces iMotes, Bluetooth radio interface, devices collecting the time epoch of contacts (or sightings) with other Bluetooth devices [Diot_infocom2006]
- traces from the CRAWDAD repository [crawdad]

Diffusion

- The *faster the better* ...
- Pure epidemic spreading [Khelil]
- Once infected, a node remains in the infected state
    - population of size N
    - at time 0 one “infected” node shows up, i.e., reading of a sensor
    - and $(N-1)$ “susceptibles” (i.e., which have not got the message)
    - engine: mobility and opportunistic data exchange
    - question: how long does it take to advertise the new data to all nodes?



Diffusion

- A simple model to describe the infection process [TVT]
- Message generated at a source sensor
- $Z(t)$ being the number of infected nodes, i.e. the number of copies of the message
- $\{Z(t), t \geq 0\}$ is a semi-Markov process over the state space $\{1, \ldots, N\}$.
- The sojourn time depends on the state
- State N is adsorbing: message broadcasted successfully
- Need to solve the transient of the semi-Markov process: need to assess how many copies of the message we have in the system

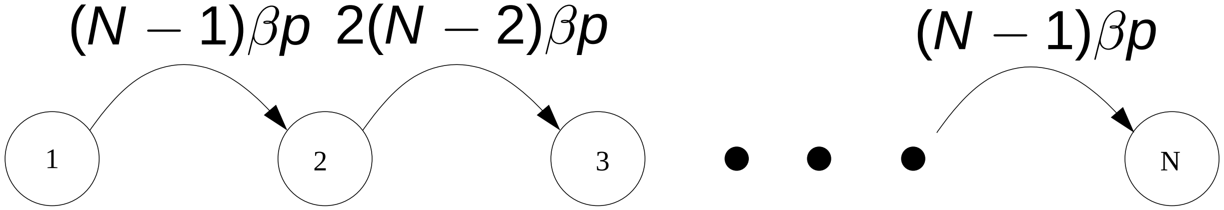
Diffusion

- General sojourn times in each state
- Assume to know: $\psi_{i,i+1}$ the time it takes to pass from state i to state $i+1$
- $\Psi^*_{i,i+1}(s)$ the Laplace-Stieltjes Transform (LST) of its pdf
- θ_k time to reach state k
- $\Theta^*_k(s) = \prod_{i=1}^{k-1} \Psi^*(s)_{i,i+1}$
- Transient probabilities that at time t there exist k copies of the message

$$\begin{array}{rcl} p_N(t) & = & \mathbb{P}[\theta_N \leq t], \\ p_{N-1}(t) & = & \mathbb{P}[\theta_{N-1} \leq t] - \mathbb{P}[\theta_N \leq t], \\ \vdots & & \\ p_2(t) & = & \mathbb{P}[\theta_2 \leq t] - \mathbb{P}[\theta_3 \leq t], \\ p_1(t) & = & 1 - \mathbb{P}[\theta_2 \leq t]. \end{array} \tag{1}$$

- $P^*_k(s) = \frac{\Theta^*_k(s) - \Theta^*_{k+1}(s)}{s}$
- Numerical inversion may then by applied to obtain $p_k(t)$

Diffusion

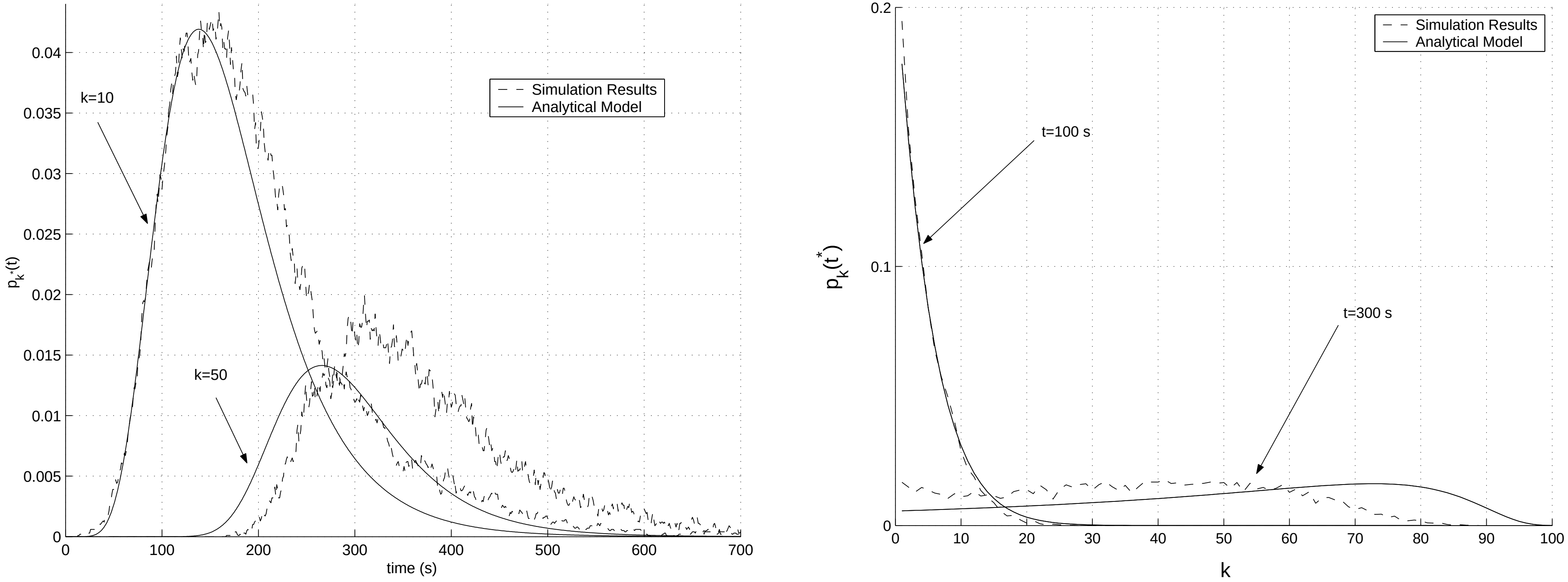


- Case of mobility: user nodes move at a constant speed v, square playground of size $L \times L$, with $L \gg R$, R communication range
- Meeting times of nodes i and j, $\{T_n, \sigma_n = (i,j)\}$: Poisson point process of intensity β [groen:sigm05]
- Function: $\beta \approx 2R \frac{4V}{\pi} \frac{1.3683}{L^2}$
- Packet errors: sampled process with intensity βp
- Exponential form: $\Theta^*_{k+1}(s) = \Theta^*_k(s) \cdot \frac{(N-k)k\beta p}{(N-k)k\beta p+s}$,
- Case k=1,2,…, N-1: $P^*_k(s) = \frac{1}{(N-k)k\beta p+s} \cdot \prod_{i=1}^{k-1} \frac{(N-i)i\beta p}{(N-i)i\beta p+s}$
- Case k=N: $P^*_N(s) = \frac{\Theta^*_N(s)}{s} = \frac{1}{s} \cdot \prod_{i=1}^{N-1} \frac{(N-i)i\beta p}{[(N-i)i\beta p+s]}$

Diffusion

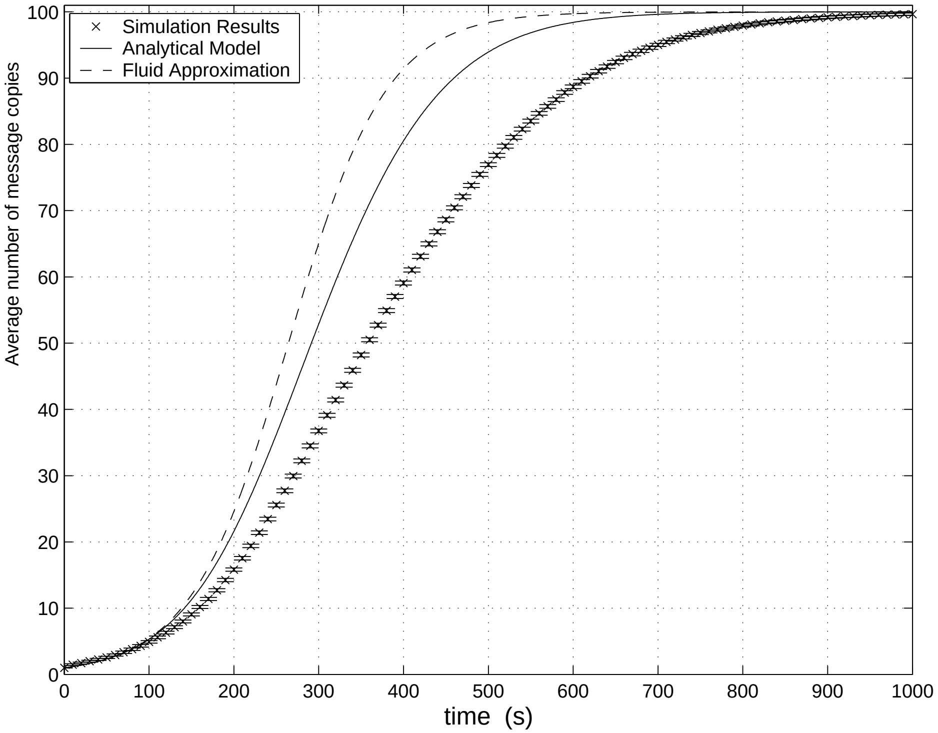
- Inversion for RWP mobility: case N even
- Case $1 \leq k \leq \frac{N}{2}$:
  - $p_k(t) = \sum_{i=1}^{k} r_{k,i} e^{-[(N-i)i\beta p]t} \mathbb{U}(t)$
  - simple poles: $r_{k,i} = \dfrac{\prod_{j=1}^{k-1} j(N-j)}{\prod_{j=1, j\neq i}^{k} j(N-j)-i(N-i)}$
- Case $\frac{N}{2}+1 \leq k \leq N-1$:
  - $p_k(t) = \left\{ r_{k,\frac{N}{2}} e^{-\left[(\frac{N}{2})^2 \beta p\right]t} + \sum_{i=1}^{N-1-k} r_{k,i} e^{-[(N-i)i\beta p]t} + \sum_{i=\frac{N}{2}+1}^{k} \left[r'_{k,i} e^{-[(N-i)i\beta p]t} + r''_{k,i} t e^{-[(N-i)i\beta p]t} \right] \right\} \mathbb{U}(t)$
  - $r''_{k,i} = \beta p \dfrac{\prod_{j=1}^{k-1} j(N-j)}{\prod_{j=1, j\neq i, j\neq N-i}^{k} (N-j)j-(N-i)i}, \; r'_{k,i} = -\frac{r''_{k,i}}{\beta p} \cdot \sum_{\substack{j=1 \\ j\neq i, j\neq N-i}}^{k} \frac{1}{(N-j)j-(N-i)i}$

Diffusion



Match with experimental data RWP “perfect simulations” with 100 nodes, speed 4 m/s, v = 4 m/s, R = 50

Diffusion



- Fluid model: a reasonably simple tool
- Speed of diffusion is paid in terms of overhead
- Dependence on the mobility model

K-relaying

- [fluid_models] The speed of diffusion depends on the number of copies allowed *per node*

- The more the copies, the larger the overhead: if we limit the number of copies?

- K-relaying protocol: each node is allowed to make up to K copies of the message and release them to neighbors within radio range

- Marked point process $\{Z_n\}_{n\in\mathbb{Z}} = \{T_n, \xi_n\}_{n\in\mathbb{Z}}$, where:
  - The points $\{T_n\}$ represent the meeting times, i.e., the times at which two nodes get within mutual communication range;
  - The marks $\{\xi_n\}$ represent the IDs of the nodes getting in contact.

- $\hat{T}$ is the mean intermeeting time among (pairs of) nodes in the system

- $\lambda = \frac{1}{\hat{T}}$ is the mean rate at which (pairs of) nodes meet

K-relaying

- Reference scenario: $S(t)$ infected nodes, $N - S(t)$ susceptible nodes
- Case $K = N - 1$:

$$\frac{\partial S(t)}{\partial t} = \lambda\pi S(t) \cdot [N - S(t)] \qquad \rightarrow \qquad S(t) = \frac{Ne^{N\lambda t}}{N - 1 + e^{N\lambda t}}$$

- Case $K = 1$:

$$\frac{\partial S(t)}{\partial t} = \lambda \cdot [N - S(t)] \qquad \rightarrow \qquad S(t) = \lambda\pi\,[N - S(t)]$$

- Case $K = 2, \dots, N - 2$ (uniform approx. on the number of copies transmitted)

$$\frac{\partial S(t)}{\partial t} = \lambda \frac{K}{K + 1} S(t) \cdot [N - S(t)] \qquad \rightarrow \qquad S(t) = \frac{Ne^{N\lambda \frac{K}{K+1} t}}{N - 1 + e^{N\lambda \frac{K}{K+1} t}}$$

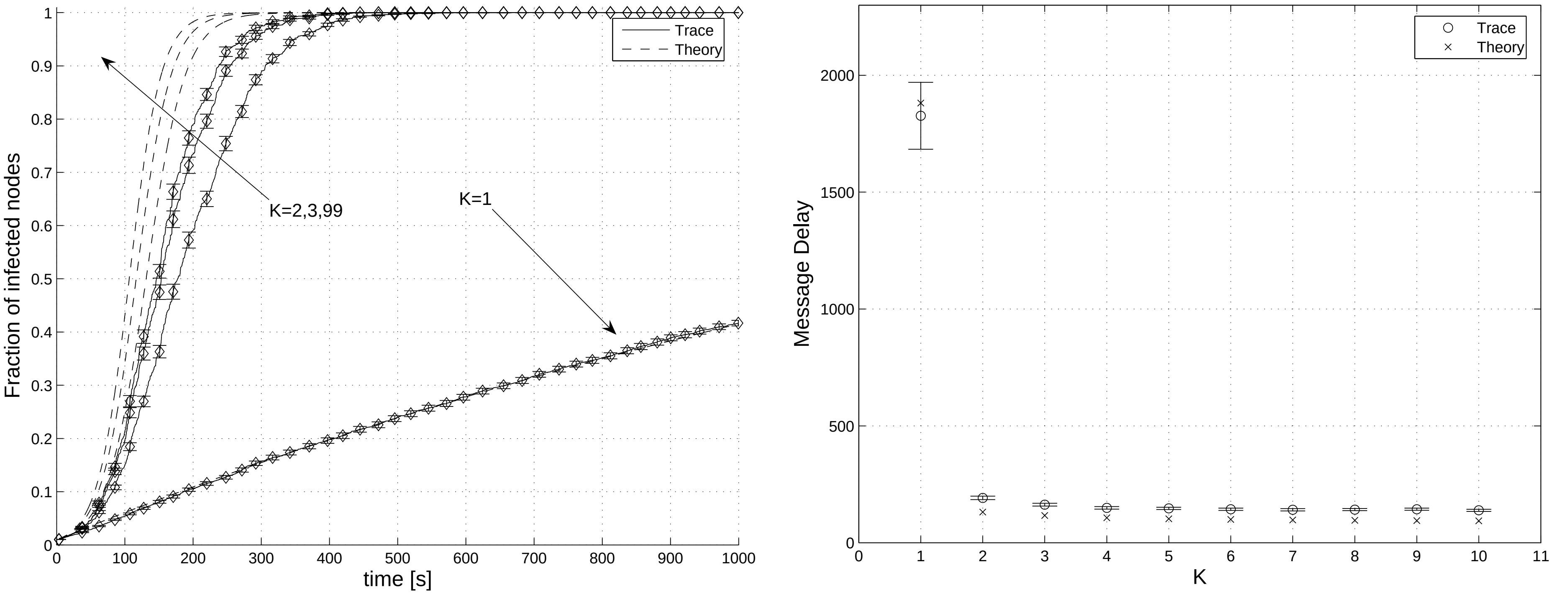
K-relaying

- Delay for the message to reach a uniformly chosen destination
- Under uniformity assumptions

$$\mathbb{E}[D] = \int_0^{+\infty} \mathrm{d}t\,[1 - F_D(t)] = \int_0^{+\infty} \mathrm{d}t \left[1 - \frac{S(t)}{N}\right]$$

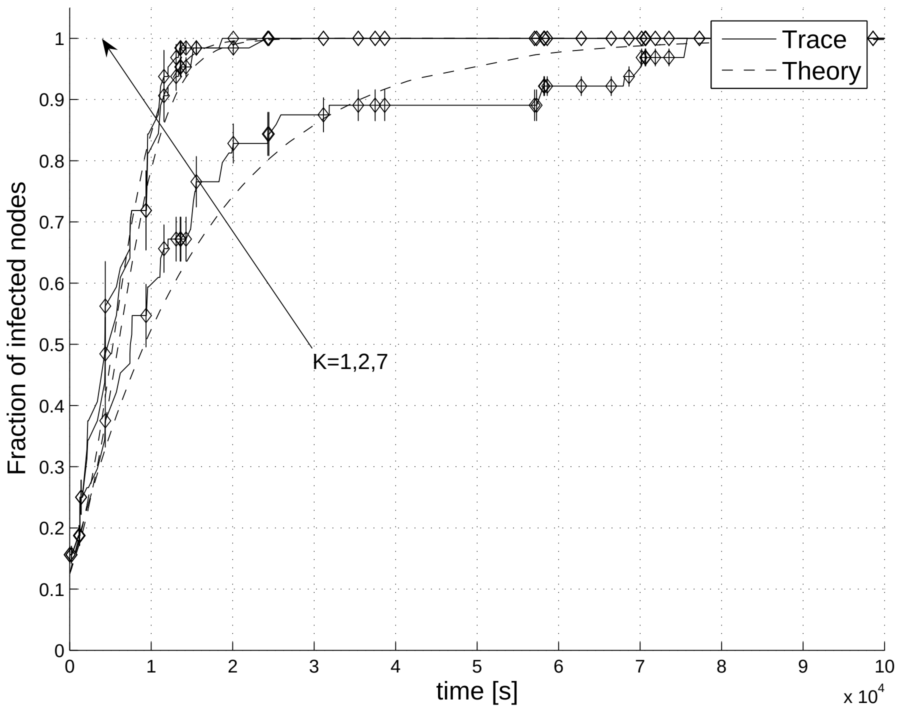
- $K = 2, \ldots, N-2$: $\mathbb{E}[D] = \frac{\log N}{\lambda \frac{K}{K+1}(N-1)}$
- $K = N-1$: $\mathbb{E}[D] = \frac{\log N}{\lambda(N-1)}$
- $K = 1$: $\mathbb{E}[D] = \frac{1}{\lambda}$
- Scaling in the *dense* case: speed $v = cost$, typical distance $d = \hat{T} = \Theta(1/\sqrt{N})$
  - $K = 1$ $\mathbb{E}[D] = \Theta(1)$, $K > 1$ $\mathbb{E}[D] \to 0$
- Scaling in the *extended* case: speed $v = cost$, typical distance $d = \hat{T} = \Theta(\sqrt{N})$
  - $K = 1$ $\mathbb{E}[D] \to \infty$, $K > 1$ $\mathbb{E}[D] \to 0$

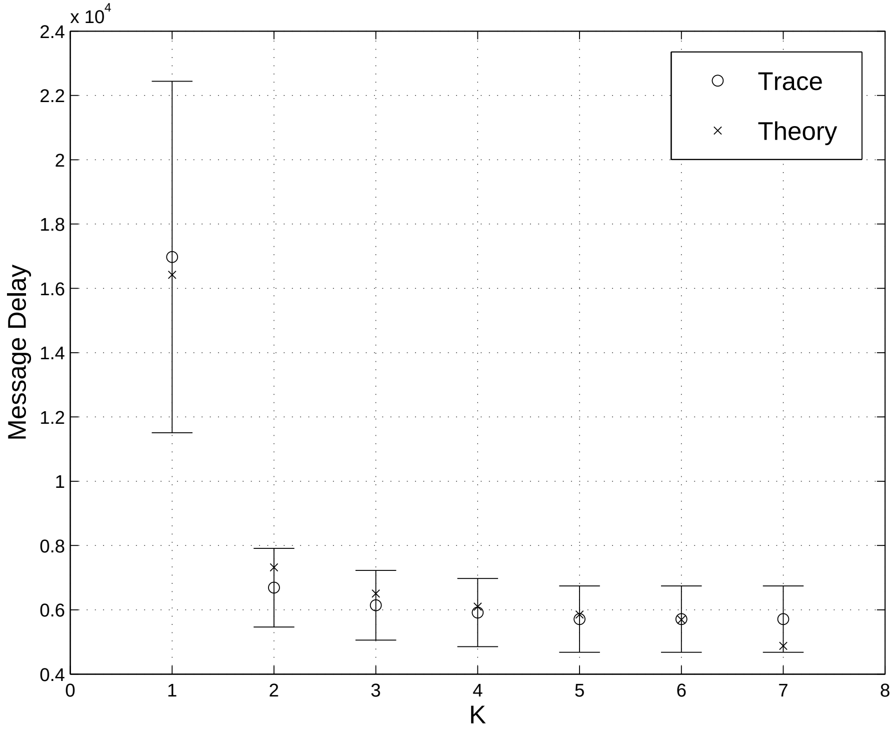
RWP match



100 nodes, speed 5 m/s, square playground of area 1×1 km^2

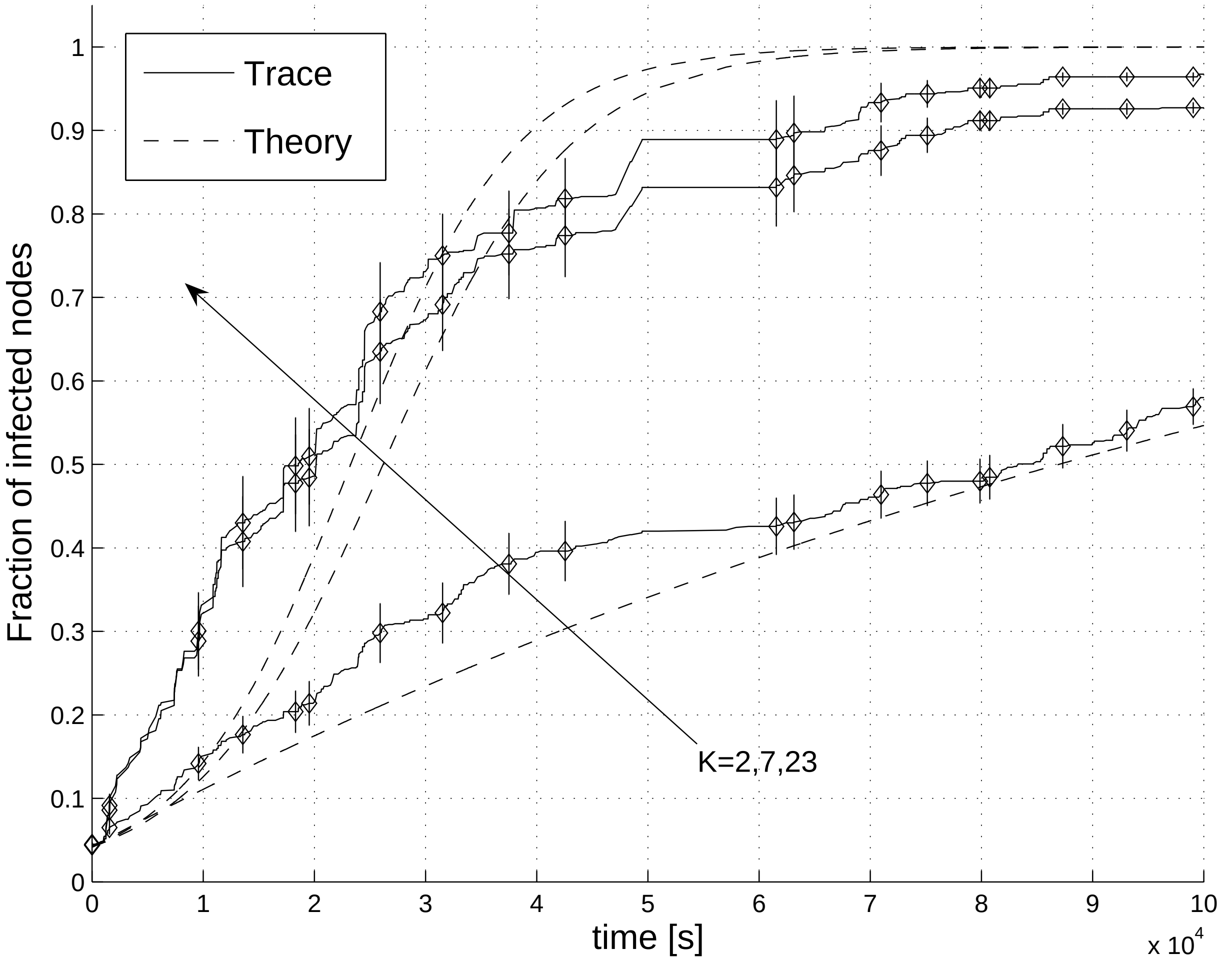
Real-world match





Filtered trace of UMassDieselNet of 20/4/2005, at least 60 contacts

Real-world match



Complete trace of UMassDieselNet of 20/4/2005, $K = 1, 2, 7$.

Dependence on the mobility

- A general model for mobility: M^2 mobility framework
- Assumptions for the mobility model:

  (A1) Communications among nodes take place only at meeting instants.

  (A2) The duration of a contact is long enough to transmit all the data carried by the nodes involved in the meeting.

  (A3) All nodes have infinite buffers, so that no message is dropped due to buffer overflow.

- Meeting instants: marked point process $Z_n = (T_n, \sigma_n)$
- T_n represents the n-th meeting time
- $\sigma_n = (i, j)$, $i, j \in \mathcal{S}$ mark of the IDs of the nodes involved in the meeting
- Sequence of intrameeting times, $\{Y_n\}_{n \in \mathbb{Z}}$, where $Y_n = T_n - T_{n-1}$:

  (i) The marks $\{\sigma_n\}_{n \in \mathbb{Z}}$ form an i.i.d. sequence.

  (ii) The intrameeting times $\{Y_n\}_{n \in \mathbb{Z}}$ form an i.i.d. sequence.

Dependence on the mobility

A precise framework provides sharper results:

Lem. 1 *If* $\pi(i,j) > 0\ \forall i,j \in \mathcal{S}, i \neq j$, *then the mean intermeeting time for any* (i,j) *pair is finite with probability one.*

Th. 1 *If* $\pi(i,j) > 0\ \forall i,j \in \mathcal{S}, i \neq j$, *and the message generation process is independent of the meeting process, a necessary and sufficient condition for the mean message delay to be finite is* $\mathbb{E}\left[Y_0^2\right] < +\infty$, *irrespective of the value of* K.

Th. 2 *If* $\pi(i,j) > 0\ \forall i,j \in \mathcal{S}, i \neq j$, *the message generation process is independent of the meeting process and* $\mathbb{E}\left[Y_0^2\right] < +\infty$, *a necessary and sufficient condition for the variance of the message delay to be finite is* $\mathbb{E}\left[Y_0^3\right] < +\infty$, *irrespective of the value of* K.

Note: scaling is such that the intermeeting times vanish: we know also that D in such case tends to either a constant or zero

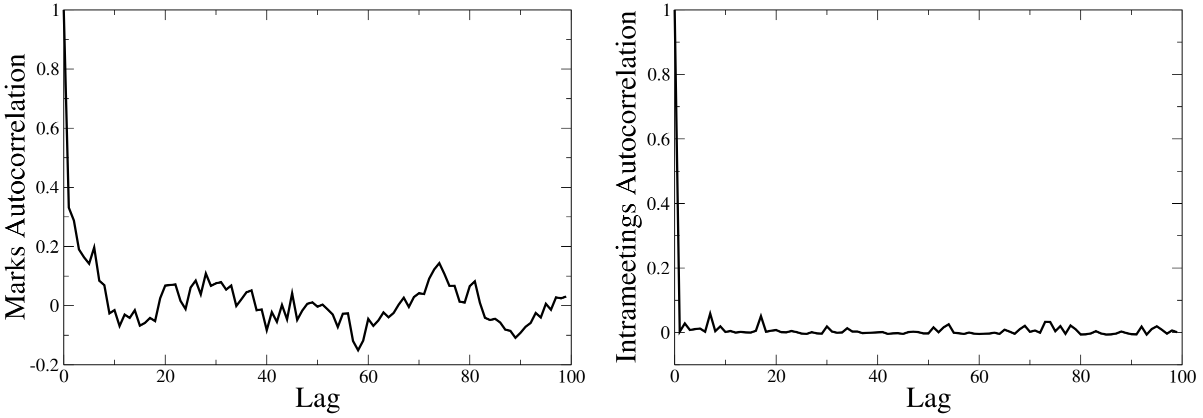
Dependence on the mobility

- [Mobiquitous]: M^2 framework is able to accommodate some of the most used synthetic mobility models,
- Good fit: random waypoint and random direction models
- Poor fit: Brownian Motion
- Real traces? Test on the correlation of *marks* and of *intrameetings*
- Sample normalized correlation function: $r_k = \dfrac{\sum_{n=1}^{L-k}(x_n-\mu)(x_{n+k}-\mu)}{\sigma_x^2}$
- Lesson learnt: the intrameeting times sequence turns out to be *practically uncorrelated*, while a *significant amount of correlation exists in the marks sequence*.
- Reasons: (i) the low number of devices in the system (ii) the periodicity of bus journeys
- The "topology" has a structure ...

Dependence on the mobility

Meetings autocorrelation peak for the DieselNet experiments

Experiment	**Marks Autocorrelation Peak**	**Intrameeting Times Autocorrelation Peak**
3012005	0.3364	0.0645
1282005	0.2476	0.0509
5222005	0.1432	0.0115
1272005	0.2174	0.0434
2012005	0.2598	0.0391
3022005	0.2210	0.0398



(a) Marks.

(b) Intrameeting times.

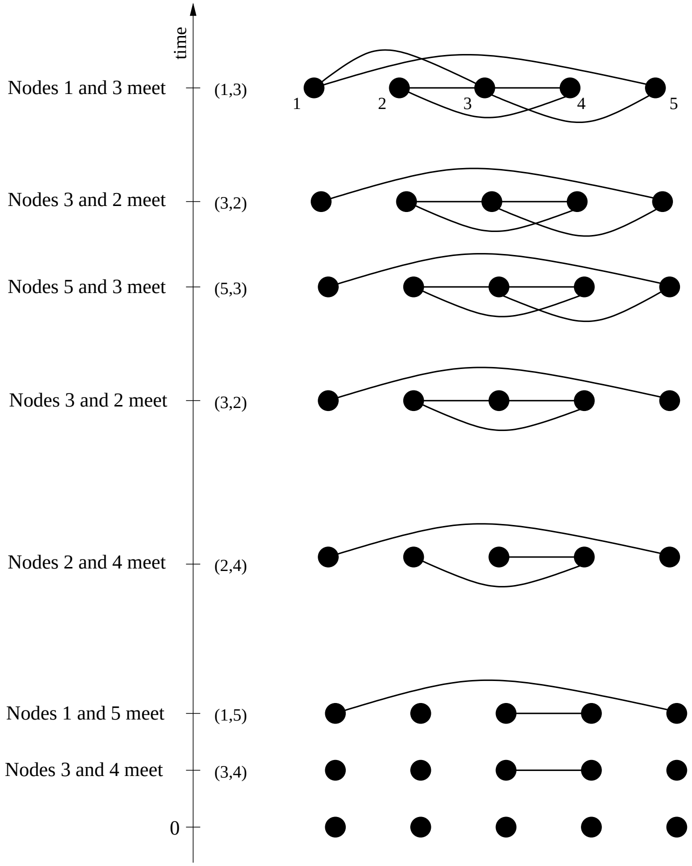
Connectivity graph

- How do we define a topology when everything is mostly disconnected?
- Common notion of a connectivity graph well established in ad hoc network literature
- proposals exist in literature
    - *DTN multigraph* [fallsigcomm03] : several edges may exist between pairs of nodes, each weighted with the capacity and delay functions
    - *space-time graphs* [zegurarouting]: perfect information of connectivity at each point in time.
    - *time-expanded graph* construction proposed in [fordfulkerson]
    - *evolving graph* sequence of graphs constructed from the presence schedule of all nodes and links over a set of intervals
    - Notion of *journey* corresponds to the concept of a path
    - building a minimum cost spanning tree in such domain (evolving spanning tree) is $\mathcal{NP}$-hard [ferreirawiopt04]

Connectivity graph

[infocom07] DTN connectivity graph: given a DTN and the associated *marked point process* $\{Z_n\}_n$, the connectivity graph $G^{(t_0,D)} = (V,E)$ is a graph where the set of vertices is $V = \{1,2,\ldots,n\}$ and the set of edges E is characterized as

$$(i,j) \in E \quad \text{if and only if } \exists\ T_n \text{ such that } \ t_0 \le T_n \le t_o + D, \sigma_n = (i,j)$$



Connectivity graph

- If $\pi(i,j) > 0$ for $\forall i,j$ the connectivity graph will converge to the complete graph as $D \to \infty$.
- Stationary assumption: $\frac{\partial}{\partial D} p_{ij}(D) \geq 0$.
- All nodes move according: intrameeting times will have a uniform distribution, which means that $\pi(i,j) = 1/C$, where $C = \binom{n}{2}$ is the maximum number of possible edges
- The connectivity graph will be an Erdös-Renyi graph [bollobas_rnd]
- Two scaling properties of ER graphs
    - Almost all ER graphs are connected when the average degree of a node scales faster than $\log(n)$, and disconnected below such threshold
    - Above the threshold, the diameter of random graphs is almost surely *concentrated* on a few values around $\frac{\log n}{\log np}$

Connectivity graph

In a uniform scenario of mobility the critical window size scales well:

Th. 3 *In a DTN of n mobile nodes, satisfying a Marks Memoryless mobility model, it holds*

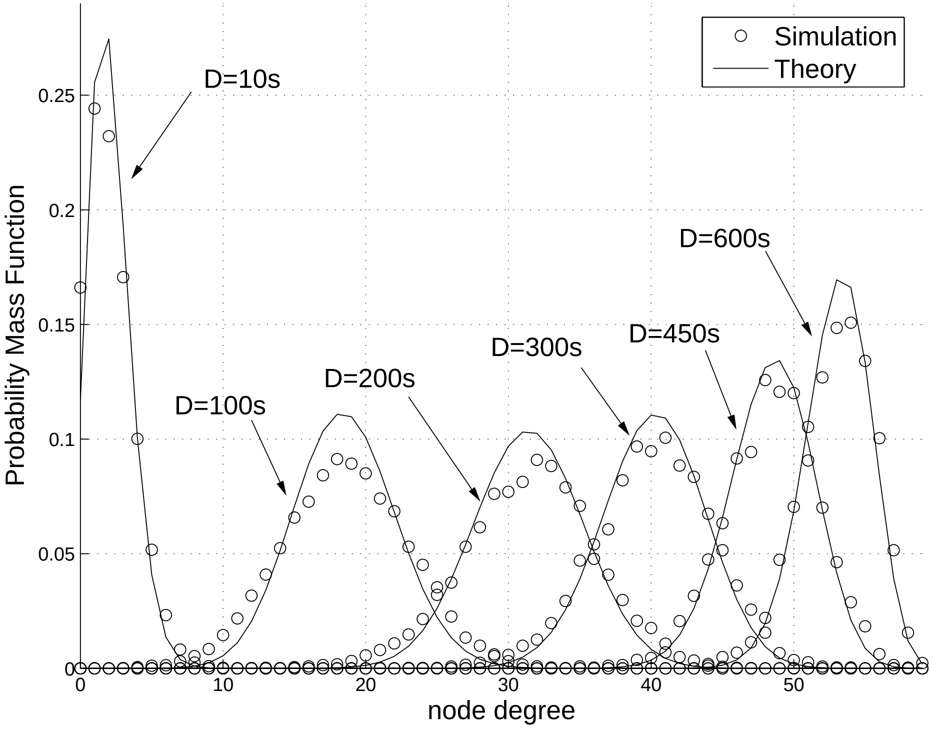
$$D^* = \Theta(n \log(n)). \tag{2}$$

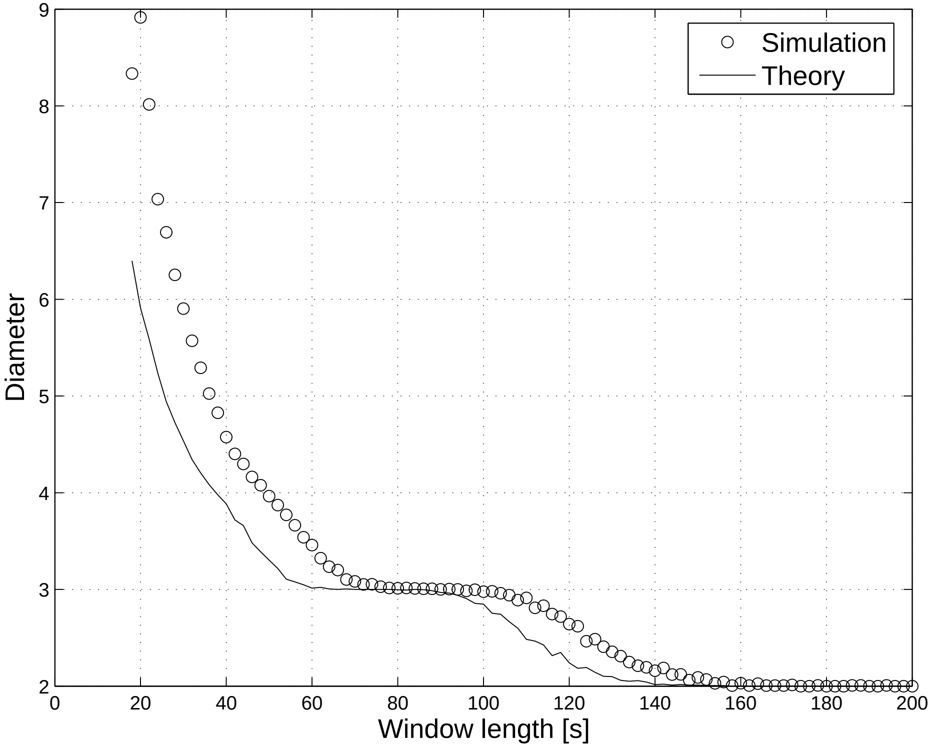
Finally, the delay sounds reasonably scalable:

Th. 4 *In a DTN of n mobile nodes, satisfying a Marks Memoryless mobility model, it holds*

$$\mathcal{D}_{ij} = O\left(\frac{n \log^2(n)}{\log(\log(n))}\right). \tag{3}$$

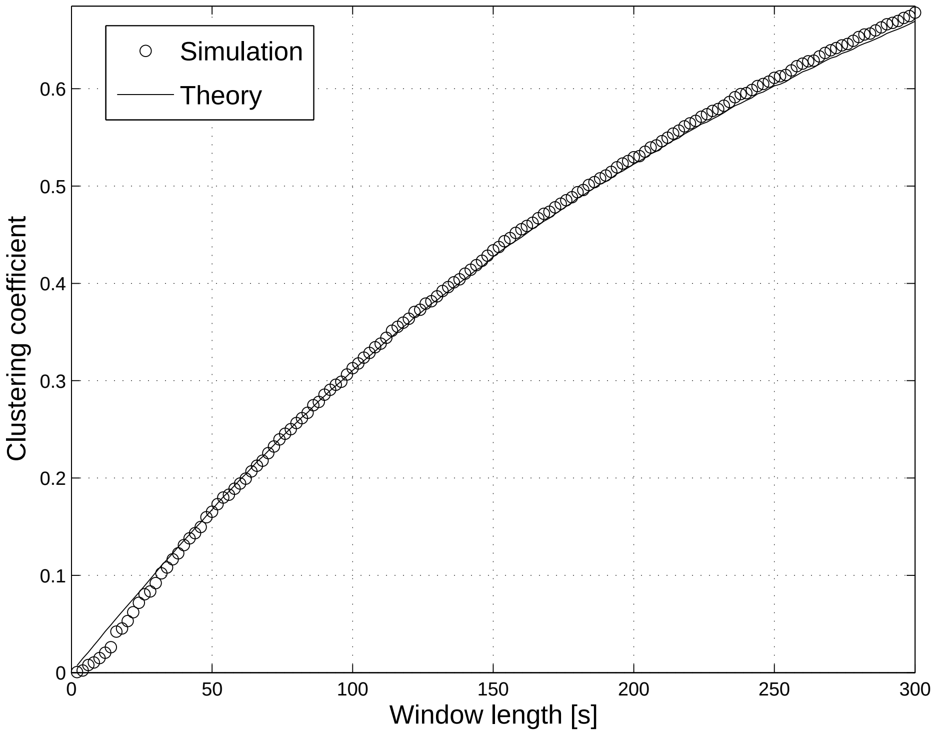
Connectivity graph





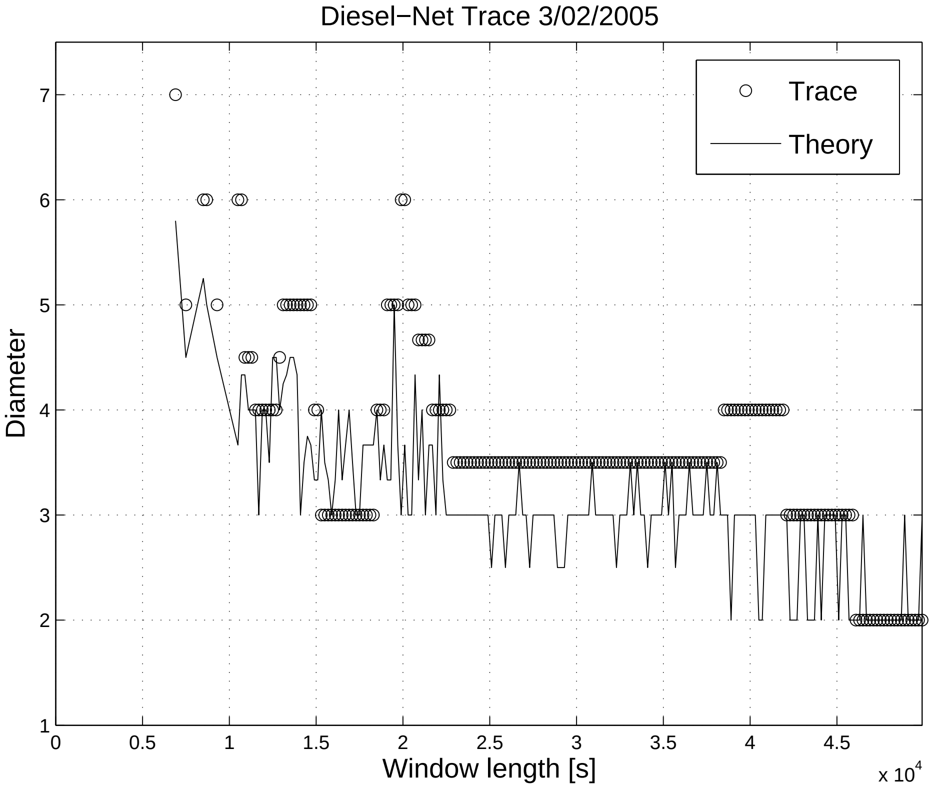
RWP confirms ER modeling

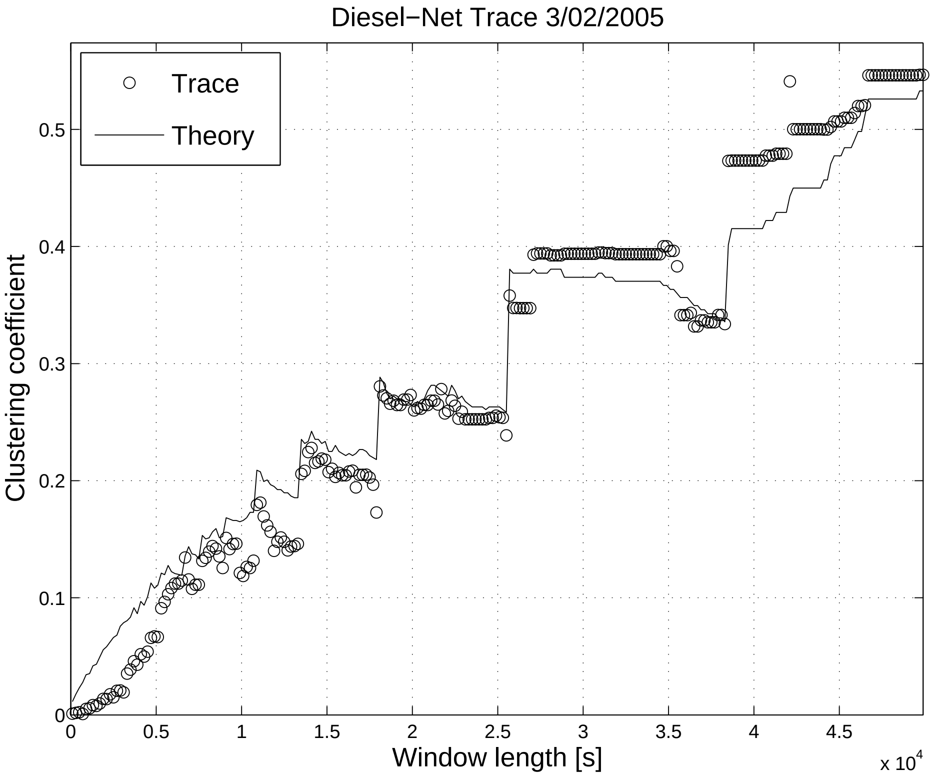
Connectivity graph



RWP confirms ER modeling: clustering

Connectivity graph





Diesel net: stationarity does not hold, p is not monotonically increasing

Information Filtering

Information Filtering

- K-copy 1-hop relaying to cope scalability issues at the U-Nodes level
- Pervasive computing/communication environments: number of sensors is expected to grow some order of magnitude higher than user nodes
- Limiting the diffusion of the data generated by T-Nodes: primary need to avoid network congestion/collapse [TVT]
- The concept itself of somehow trimmering the amount of data packets flowing in the network is not new
    - Simplest example: TTL field in an IP packet
    - TTL decremented every time a router handles the packet
    - TTL reaches zero: the packet is discarded
    - If tight on network diameter: prevents loops and flooding
- Good points: decentralized, very simple and easy to implement.
- Any more general formulation of the problem?

Information Filtering

- A U-Node issues a query at time t from position (x, y, z), concerning the value of a given random field X
- A nearby sensor answers with the data measured at location (x_0, y_0, z_0) at time t_0
- Sensor located at (x_0, y_0, z_0): samples field X at time t_0
- Value $X(x_0, y_0, z_0, t_0)$ is then coded, packetized and transmitted.
- Minimum amount of bits necessary to code the data [TC]

$$H(X) = -\sum_{x \in \mathcal{S}} \pi(x) \log \pi(x)$$

- Do we need this much bits independent of the distance and the time?
- Sensor: the information content should decay over time and space (e.g. temperature collected 2 months ago 100 Km far away)
- User: “sphere of interest” in time and space

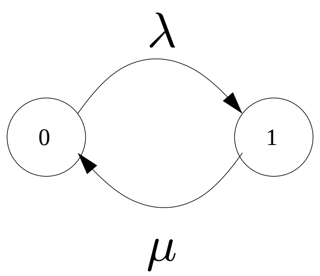
Information Filtering

- Message containing $X(x_0, y_0, z_0, t_0)$: received at (x, y, z, t)
- User needs different information: status of the field at $X(x, y, z, t)$
- Actual information content of the data message:

$$I(x, y, z, t; x_0, y_0, z_0, t_0) = H(X) - \xi, \tag{4}$$

- Equivocation: $\xi = H\left(X(x, y, z, t) | X(x_0, y_0, z_0, t_0)\right)$
- Notice: equivocation in this case is given by a pure distortion term, due to displacement in space and time
- As time and space shift increases: $\xi \rightarrow H(X)$, then $I \rightarrow 0$.
- Optimal coding scheme: adjust the data matching its information content
- Fano's bound [TC]: $H(P_e) + P_e \log\left(|\mathcal{S}| - 1\right) \geq \varepsilon H(X)$

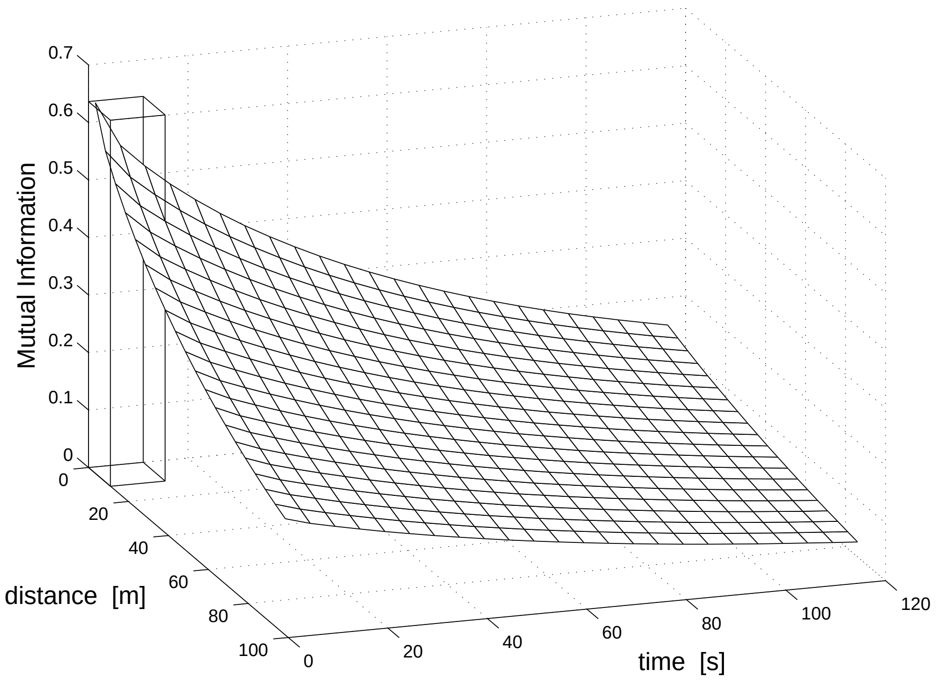
Information Filtering



- Rather crude example: a two-states Markovian model for a sensor placed at (x, y, z)
- Readings are Poisson distributed
- Models a physical quantity exceeding a given threshold: surveillance, environmental monitoring, resources availability or domotic applications
- Correlation function:

$$R_Y(t) = \left(\frac{\lambda}{\lambda + \mu} + \frac{\mu e^{-(\lambda+\mu)t}}{\lambda + \mu} \right) \cdot \frac{\lambda}{\lambda + \mu}$$

Information Filtering



$\lambda = 0.0033$, $\mu = 0.0017$, $\gamma = 0.0050$, $\varepsilon = 0.2$

- Binary process normalized covariance

$$\rho_1(t) = \frac{R_Y(t) - \pi_1^2}{\pi_1 - \pi_1^2} = e^{-(\lambda+\mu)t}$$

- Spatial correlation: $\rho_2(d) = e^{-\gamma|d|}$
- If separable: $\rho_X(d,t) = \rho_1(t) \cdot \rho_2(d) = e^{-(\lambda+\mu)t-\gamma|d|}$

$$I(x+x_0, y_0, z_0, t+t_0; x_0, y_0, z_0, t_0) = \pi_0\pi_1\rho_X(d,t)\log\frac{\left[1+\frac{\rho_X(d,t)\pi_0}{\pi_1}\right]\cdot\left[1+\frac{\rho_X(d,t)\pi_1}{\pi_0}\right]}{(1-\rho_X(d,t))^2} +$$

$$+\pi_1^2\log\left(1+\frac{\rho_X(d,t)\pi_0}{\pi_1}\right)+\pi_0^2\log\left(1+\frac{\rho_X(d,t)\pi_1}{\pi_0}\right)+2\pi_0\pi_1\log(1-\rho_X(d,t))\,.$$

Information Filtering - Packet Discarding

- Assume a model for the T-node sensing process is available [sensfus]
- $P_e(t) = P\{X(t+t_0) = 0|X(t_0) = 1\}P\{X(t_0) = 1\} + P\{X(t+T_0) = 1|X(t_0) = 0\}P\{X(t_0) = 0\}$
- $P_e(t) = 2\frac{\lambda\mu}{(\lambda+\mu)^2}\left(1 - e^{-(\lambda+\mu)t}\right)$
- Example: $\mu = \frac{1}{3}\gamma$ and $\lambda = \frac{2}{3}\gamma$, where $1/\gamma = 900$ s (15 minutes)
- In state 0: 45 minutes; in state 1: 90 minutes (on average)
- $P_e(t) = \frac{4}{9}(1 - e^{-\gamma t})$
- Assume that a U-node needs to know the status of the T-node with an error probability smaller than 5%
- The corresponding time threshold is

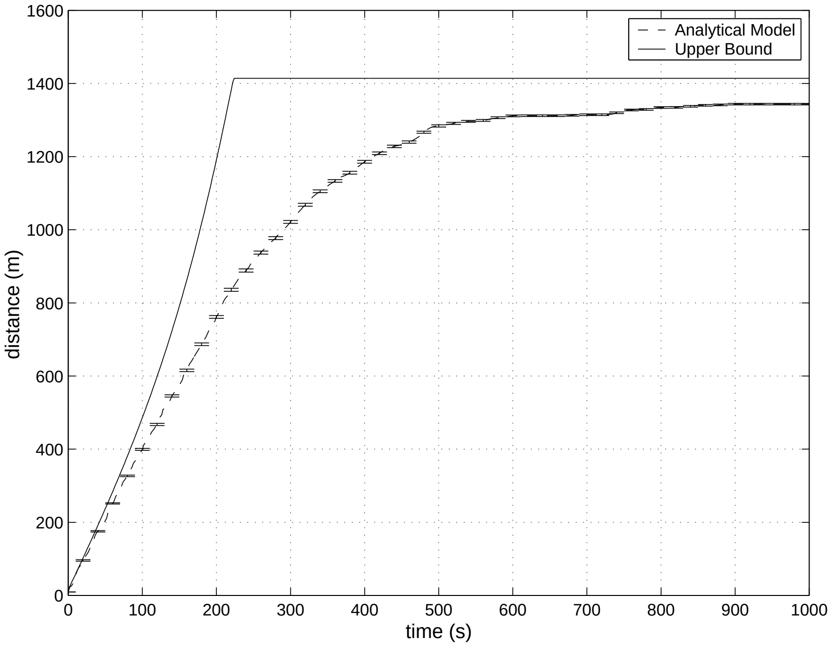
$$t_e = -\frac{1}{\gamma}\ln(1 - 0.05\frac{9}{4})s = \frac{1}{\gamma}0.119346s \tag{5}$$

  which brings $t_e = 107.4114$ s.
- Policy: discard all messages older than t_e.

Information Filtering

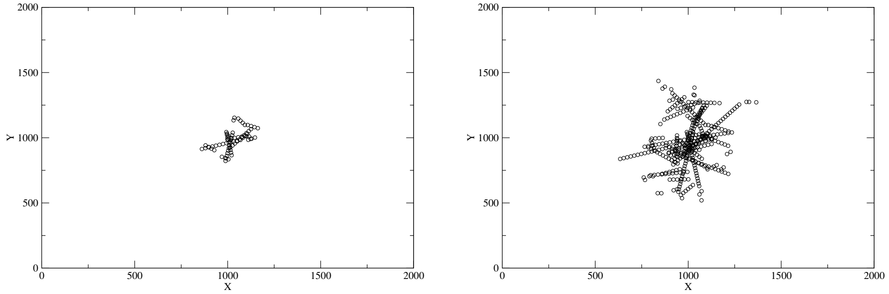
- Assume long persistence of data from sensors, but local scope
- Using the fluid model [TVT]:

$$D(t) = V \cdot t + \frac{R}{\pi} \cdot \frac{Ne^{N\beta pt}}{N + e^{N\beta pt}}. \qquad (6)$$

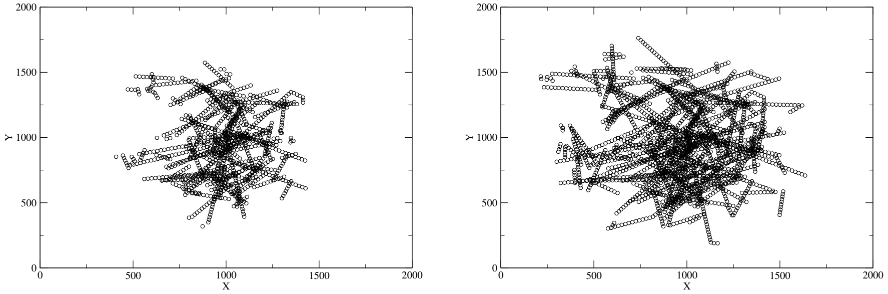


$N = 100$ mobile users, RWP, $v = 4$ m/s, $R = 50$ m

Information Filtering



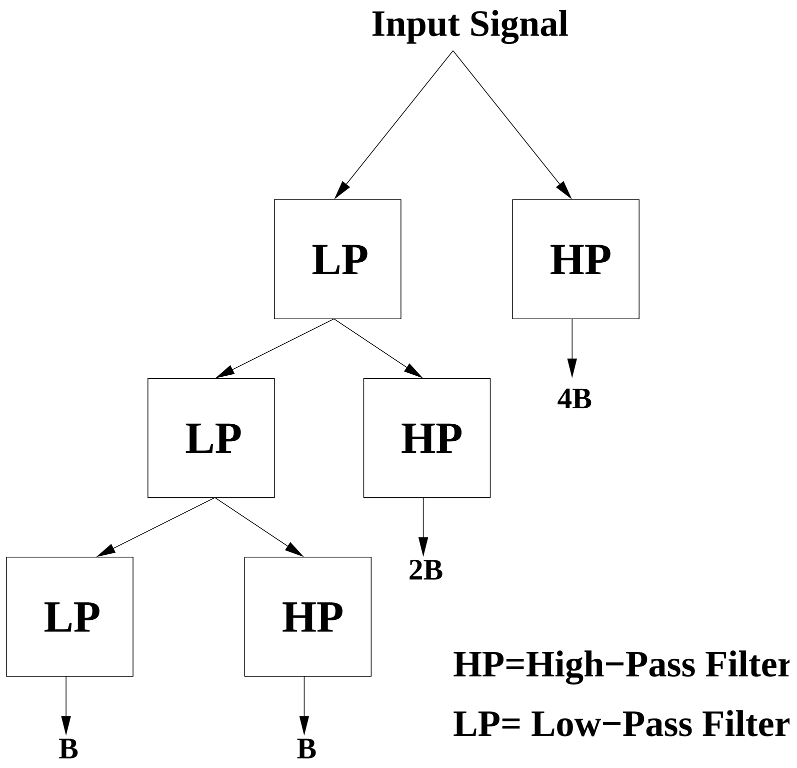
(c) 60 s (d) 120 s



(e) 180 s (f) 240 s

Footprint: $N = 200$, speed of 4 m/s, RWP

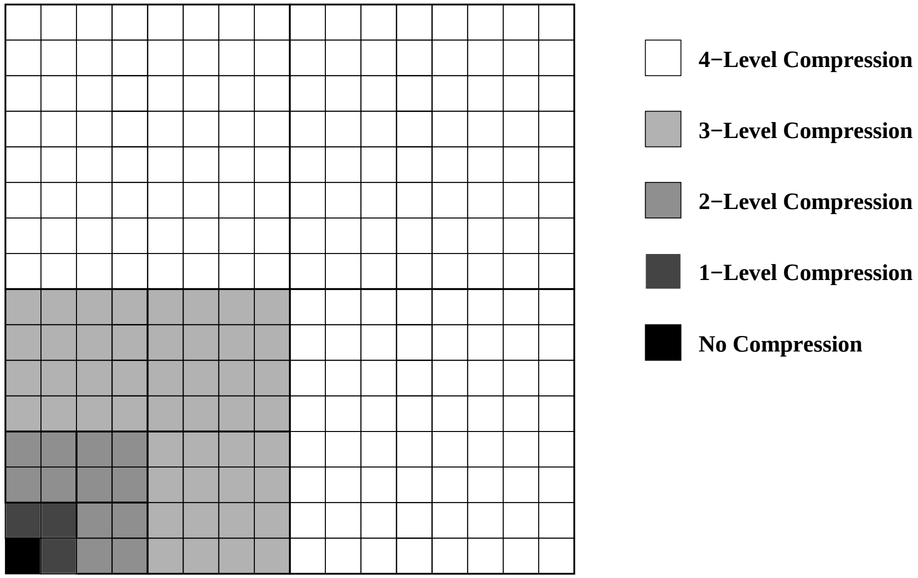
Information Filtering



Iterated Filter Bank implementation of the Discrete Wavelet Transform: components have bandwidth $B, B, 2B$ and $4B$ respectively, signal has bandwidth $8B$

- Instead of *dropping* an entire packet, try to compress it
- *High compression ratios:* wavelet decomposition achieve high compression ratios (image/video compression)
- *Low-complexity:* availability of low-complexity of implementations (by means of filter banks)
- *Multi-resolution:* capability of carrying joint time-frequency information on the considered signal

Information Filtering



- N^2 sensing devices uniformly deployed over a $L \times L$ square playground
- playground logically divided into $T \times T$ square tiles
- M mobile nodes are moving over the playground; reading sensor nodes *in each tile* when in mutual communication range
- the gathered information is then stored in the internal memory of the U-node

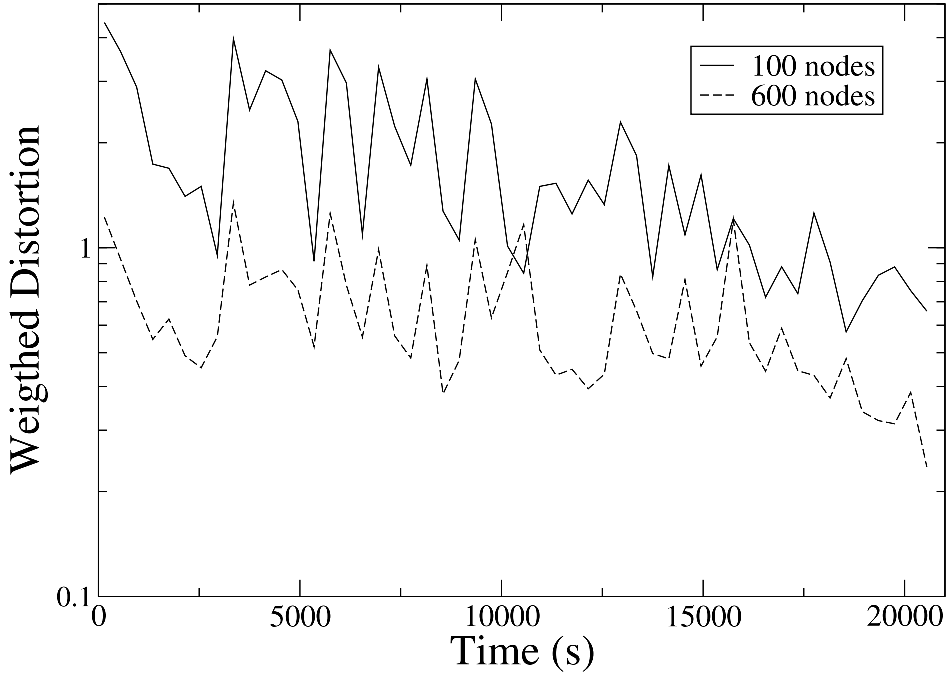
- Every tile compressed separately, the overall map is stored in the sensor memory togheter with meta-information (timestamp and number of sensor readings)
- If two nodes encouter: check for possible update
- Euristic: the more recent and more complete, the better

Information Filtering

- Random field $RF = C + sin(2\pi T_f) + X$
- X is a Gaussian Random Field with given *spatial* covariance
- $T_f = 1$day and $C = 4$
- Metric: weighted distortion $D = \frac{1}{MN^2} \sum_{i=1}^{M} \sum_{j=1}^{N^2} w_{i,j}(R_{i,j} - S_j)^2$
- $R_{j,i}$ is the value stored by node i about sensor j
- $w_{i,j} = \exp(-\gamma d_{i,j})$

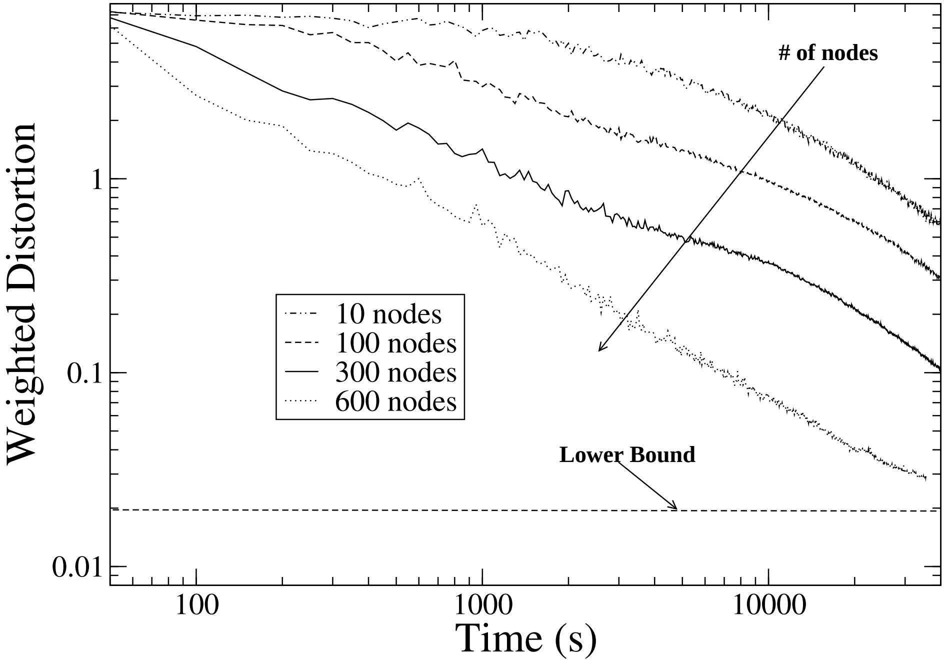
Simulation area	1280×1280 m^2
Sensors communication range	10 m
Mobile nodes communication range	50 m
Mobile nodes speed	4 m/s
Sensor nodes number	4096
U-Nodes PHY/MAC Protocol	$CSMA/CA$
Tiles T	8×8
Sensors per tile	8×8

Information Filtering



Th. 5 *The* Memory Usage (MU) *deriving from the multi-resolution storage of a Sensor Map composed by* N^2 *sensors (with* $N = 2^k, k \in N$*) and subdivided into* T^2 *tiles, is* $MU = \frac{N^2}{T^2}\left[1 + \frac{3}{8}\log_2 N\right]$. *The asymptotic behaviour of* MU *is* $\Theta\left(N \log_2 N\right)$.

Information Filtering



Accuracy depends on how many readings can be collected per tile

Service Evolution

Service Evolution

What we mean for service: a *service is a program running on top of a user device, which is leveraged by user(s), and exploits both computing and communication tasks to obtain the final result requested by the user* (e.g. navigation system)

(i) *Self-optimisation of parameters at run-time*: find the optimum combination of parameter value that is best suited to a given environment (convergence properties)

(ii) *Automatic composition of service/protocol building blocks*: the system seeks the optimal combination of existing building blocks to obtain the desired service or protocol

(iii) *Automatic generation of inter-component glue logic*: complementary to step (ii), allows for the interfaces to evolve as well.

(iv) *Automated code generation*: the most ambitious step, full building blocks could be generated from scratch

Service Evolution

- Identify the building blocks: self-evolving programs, we can distinguish between two “levels” at which evolution may take place
- Evolution at the *micro* scale: evolution taking place at the individual service/program level
  - *micro* scale: program transformations, from self-tuning of some running parameters up to self-generation of code
- Evolution at the *macro* scale comes as a consequence of evolution at the micro level, coupled with interactions among micro-level components
  - *macro* scale: successfully services replicate and propagate throughout the network, competing services share resources, cooperation is built in some regions
- Different tools for each one of these two levels

Service Evolution

- Tools at the *micro-level*:
  - Genetic Algorithms
  - Evolutionary Algorithms
  - (Evolutionary Programming, Evolution Strategy)
  - Problem: *on-line implementation*, the space of solution should be narrowed to non-disruptive ones
- Tools at the *macro-level*
  - Evolutionary games
  - Stochastic processes
  - System biology

Service Evolution

- Evolutionary game theory as a tool for tracing the dynamics of a population when multiple populations are interacting
- Two main concepts
  - equilibrium: proportions of populations are maintained and populations are immune from invasions
  - trajectories and boundedness of trajectories provide informations about the population dynamics
- Adaptation to the environment is an additional feature here
- Important: we are given the chance to engineer trajectories and equilibria, i.e., find the suitable rules to make then range within a certain subset

Service Evolution

- There is already a lot of work done on evolutionary games ...
- Mixed strategies with 2 rules $\{s_1, s_2\}$depending one parameter: with prob. $p > 0$ play s_1 and with prob. $1 - p$ play s_2
- Normal play q^*, mutant play p
- Reward of a play dictates the growth rate of a population: payoff $J(p, q)$ is the reward for having played p against all the population playing q
- In order to obtain immunity from mutants: $J(q^*, q^*) > J(p, q^*)$
- Sufficient condition is q^* strictly dominates all $p \neq q^*$ if it dominates all pure strategies $J(q^*, q^*) > J(s_i, q^*)$
- Weaker domination is also possible: in such cases q^* is a n evolutionary stable equilibrium
- This gives some “computational hope” since it reduces the number of comparisons to a finite set!

Service Evolution

- atomic services exchanging parameters in a GA-like fashion: with crossover and mutation, selection comes is a consequence of interactions [carreraswac05]
- mating procedure: data/code/parameters exchanged with other users
- overall effect: *distributed evolution process*
- standard metric: the *fitness* represents the driving parameter of the mating process
- evolution process: service is confined into a black box, looking at its output only in terms of fitness
- target: understanding how the number of nodes, the nodes speed, the mobility model affect the evolution process.

Service Evolution

- Oversimplifying: service of user i represented as a binary vector

$$\underline{v}_i = [v_i(1), \ldots, v_i(T)]$$

- $v_i(l) \in \{0, 1\}, l = 1, \ldots, T$
- Fitness: $I_i = \frac{d_H(v_i, v^*)}{T}$
- $I_i(t)$ the fitness level of service i at time t, taken to be in the interval $[0, 1]$
- v^* is an unknown optimal service
- $d_H(\cdot, \cdot)$ usual Hamming distance.

Def. 1 *A service mating policy is stable if it leads to convergence of* $X(t)$ *with unitary probability to some reference value*

Def. 2 *A service mating policy is called optimal if it leads to convergence of* $X(t)$ *to* 1 *with unitary probability*

Service Evolution

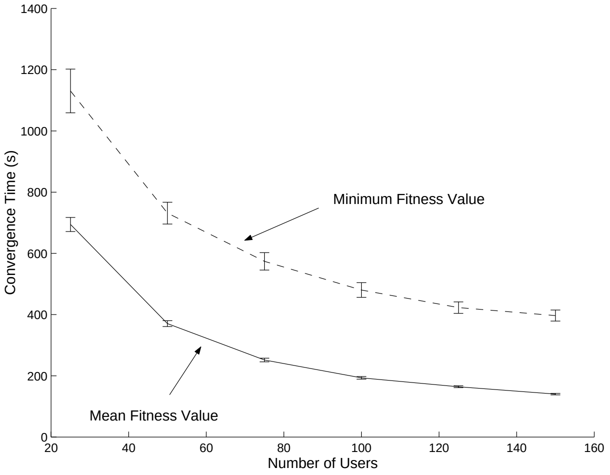
$$\phi[x,y] = (x,x)\,, \qquad \textit{clonation mating policy,}$$
$$\phi[x,y] = (x,x+\xi)\,, \qquad \textit{clone-and-mutate mating policy,}$$
$$\phi[x,y] = (x,\psi\cdot x+(1-\psi)\cdot y+\xi))\,, \qquad \textit{combine-and-mutate mating policy,}$$

Under mild assumptions on the mobility pattern, it follows:

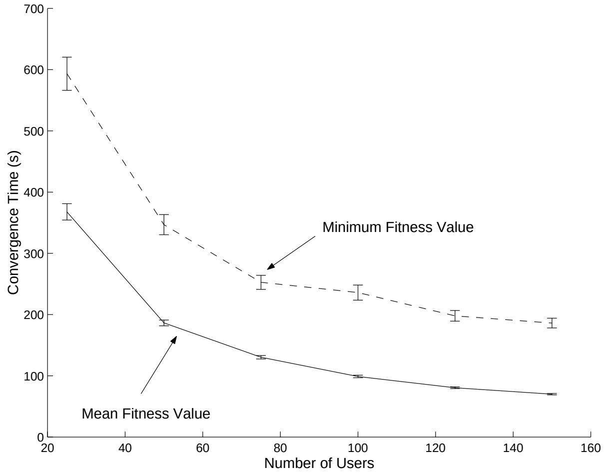
- The *clonation/clone-and-mutate/combine-and-mutate mating policies* are stable.
- The *clonation mating policy* is not optimal.
- The *clone-and-mutate mating policy* is optimal.
- The *combine-and-mutate mating policy* is optimal.

Note: stochastic perturbation induced by the mutation operator is useful to favor convergence; still open question how fast we converge [WAC05]

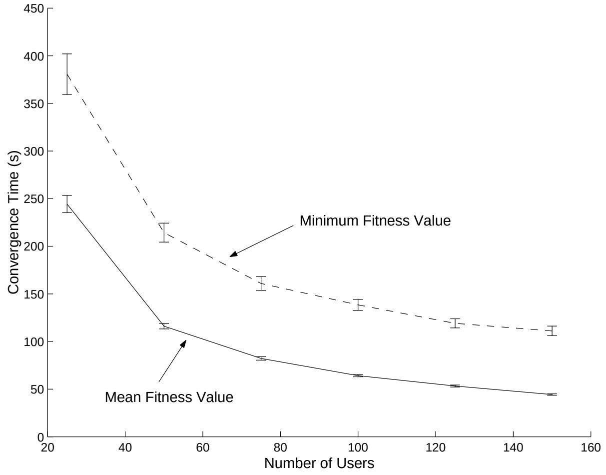
Service Evolution



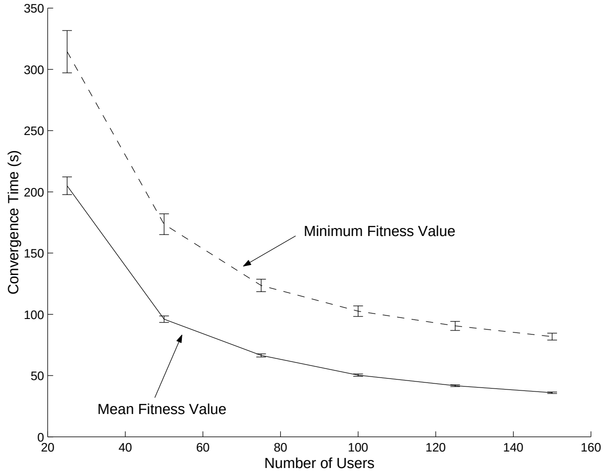
(g) Speed 2 m/s



(h) Speed 5 m/s



(i) Speed 10 m/s

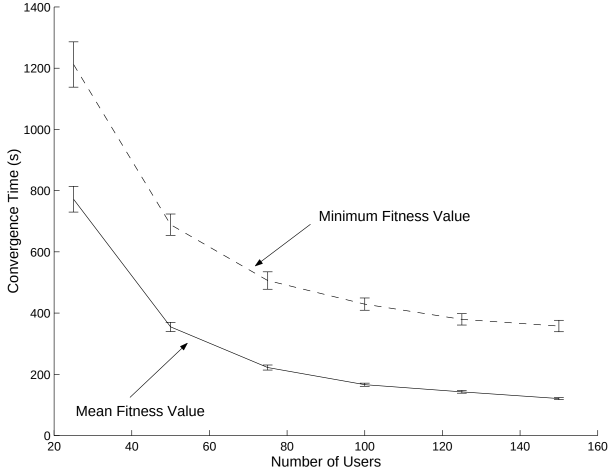


(j) Speed 15 m/s

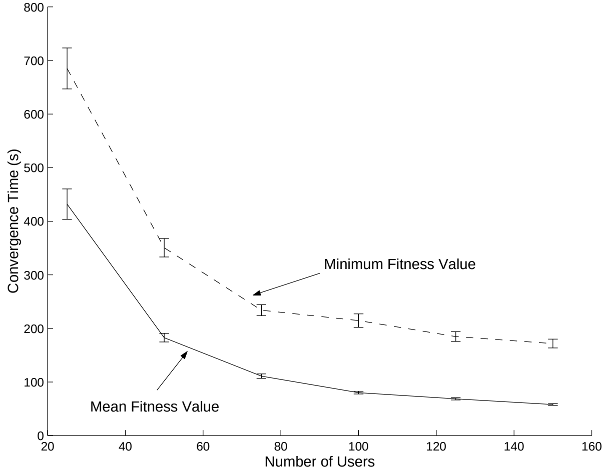
Clone and Mutate

Service Evolution

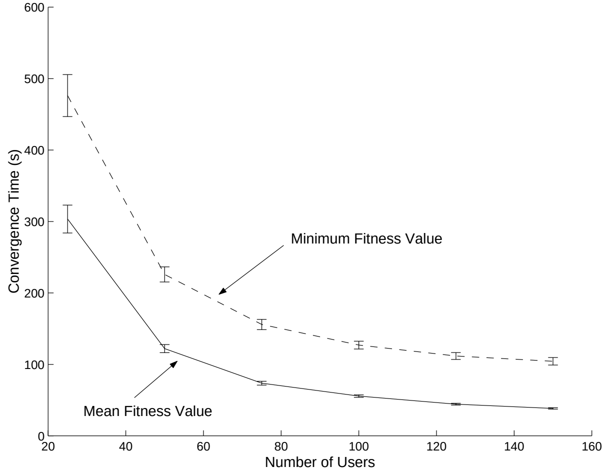
Combine and Mutate



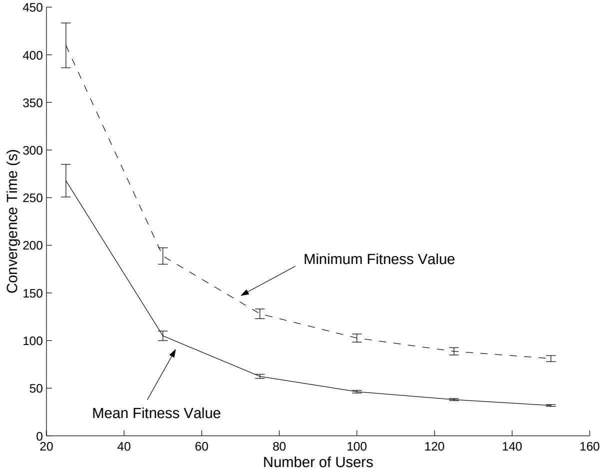
(k) Speed 2 m/s



(l) Speed 5 m/s



(m) Speed 10 m/s



(n) Speed 15 m/s

Bibliography

- [Weiser] M Weiser, *“The computer for the 21st century”*, ACM Mob. Comput. Commun. Rev., Vol. 3, no. 3, 1999, pp. 3-11
- [TC] T. Cover and J. Thomas *Elements of Information Theory*, John Wiley & Sons, New York, NY, 1991.
- [TVT] I. Carreras, I. Chlamtac, F. De Pellegrini and D. Miorandi,*“BIONETS: Bio-Inspired Networking for Pervasive Communication Environments”*, IEEE Transactions on Vehicular Technology, to appear.
- [mobiquitous] I. Carreras, D. Miorandi and I. Chlamtac, “A Framework for Opportunistic Forwarding in Disconnected Networks”, in Proc. of MOBIQUITOUS, San Jose, CA, July 06.
- [sensfus] I. Carreras, F. De Pellegrini, D. Miorandi and I.Chlamtac, *"Information Diffusion in a 2-tier Wireless Sensor Network”*, in the proceedings of SENSORFUSION 2005.
- [infocom07] F. De Pellegrini, D. Miorandi, I. Carreras and I. Chlamtac, “A Graph-Based Model for Disconnected Ad Hoc Networks”, accepted for publication in the proceedings of IEEE INFOCOM 2007
- [fluid_models] I. Carreras, F. De Pellegrini, D. Miorandi and I. Chlamtac, "Fluid Models for Epidemic Spreading of Information in Disruption-Tolerant Networks", work in progress.
- [WAC05] I. Carreras, F. De Pellegrini, D. Miorandi and H. Woesner, "Service Evolution in a Nomadic Wireless Environment", in Proc. of WAC, Athens, Oct. 05.
- [WavTech] I. Carreras, D. Tacconi, D. Miorandi, F. De Pellegrini, "A Multi-Resolution Data Management Scheme for Opportunistic Information Diffusion", CREATE-NET Technichal Report N. 200600012, November 2006.

- [Zegura_Mobihoc06] M. M. Bin Tariq, M. Ammar and E. Zegura, *“ Message ferry route design for sparse ad hoc networks with mobile nodes”* , in the proceedings of ACM MobiHoc, 2006, Florence, Italy.
- [burgess_infocom06] J. Burgess, B. Gallagher, D. Jensen and B. N. Levine, *“MaxProp: routing for vehicle-based disruption-tolerant networks”*, Proc. of IEEE INFOCOM, 2006, Barcelona, ES.
- [burleigh_commag03] S. Burleigh, L. Torgerson, K. Fall, V. Cerf, B. Durst, K. Scott and H. Weiss , *“ Delay-Tolerant Networking: an approach to interplanetary Internet”*, IEEE Comm. Mag., vol 41, no. 6, pp. 128-136, June, 2003.
- [lee_mobeyes06] U. Lee, E. Magistretti, B. Zhou, M. Gerla, P. Bellavista and A. Corradi , *“Efficient Data Harvesting in Mobile Sensor Platforms”*, Second IEEE International Workshop on Sensor Networks and Systems for Pervasive Computing (PerSeNS 2006), Pisa, Italy, March, 2006.
- [fallsigcomm03] K. Fall, *“A Delay-Tolerant Network Architecture for Challenged Internets”*, in Proc. of ACM SIGCOMMm 2003, Jarlsruhe, DE.
- [crawdad] *“CRAWDAD, the Community Resource for Archiving Wireless Data At Dartmouth”*, url: `crawdad.cs.dartmouth.edu/`.
- [umassdiesel]*“The Disruption Tolerant Networking Project at UMass ”*, url: `prisms.cs.umass.edu/diesel/`.
- [realitymining] *“Machine Perception and Learning of Complex Social Systems”*, url: `reality.media.mit.edu/`.
- [khelil_mswim02] A. Khelil and C. Becker and J. Tian and K. Rothermel, *“An epidemic model for information diffusion in MANETs”*, Proc. of ACM MSWiM, 28 Sept., 2002, Atlanta, Georgia, pp. 54-60.

- [Diot_infocom2006] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, Richard Gass and James Scott , *“Impact of human mobility on the design of opportunistic forwarding algorithms”*, Proc. of INFOCOM, 23-29 Apr., 2006, Barcelona, Spain.
- [groensigm05]R. Groenevelt and P. Nain , *“Message Delay in MANETs”*, Proc. of ACM SIGMETRICS, pp. 412-413, 2005; also: R. Groenevelt, Stochastic Models for Mobile Ad Hoc Networks. PhD thesis, University Nice-Sophia Antipolis, April 2005.
- [ferreira_wiopt04] A. Ferreira and A. Jarry , *“Complexity of Minimum Spanning Tree in Evolving Graphs and the Minimum-Energy Broadcast Routing Problem”*, Proc. of WiOpt’04, Cambridge, UK,March 2004.
- [zegura_routing] S. Merugu and M. Ammar and E. Zegura , *“Routing in Space and Time in Networks with Predictable Mobility”*, GIT-CC-04-07, Georgia Tech., College of Computing, 2004.
- [fordfulkerson] L. R. Ford and D. R. Fulkerson , *“Constructing Maximal Dynamic Flows from Static Flows”*, Operations Research, Vol. 6, no. 3, pp. 419-433, May-June, 1958(!).
- [bollobas_rnd] Bela Bollobas , *“Random Graphs”*, Cambridge Univ. Press, 2001, Cambridge.