Study of Two-Hop Message Spreading in DTNs

Antonis Panagakis, Athanasios Vaios, Ioannis Stavrakakis

WiOpt 2007
5th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks
April 16-20, 2007, Limassol, Cyprus

Reference: http://cgi.di.uoa.gr/~ioannis/publications/2007WIOPT.pvs.pdf

Supported by: IST BIONETS, PENED 2003, PYTHAGORAS II

Delay Tolerant Networks

Connectivity established through (low frequency) node encounters

DTN message transport

- Mobility-assisted
- typically, multiple message spreading is needed

Some DTN message transport /routing schemes

Epidemic routing: (unlimited copies)
- At each node encounter, all copies are exchanged
- Minimum message delivery delay but high buffer occupancy and bandwidth utilization

Binary spray-and-wait routing (BSW): (limited copies)
- Every one gives half its copies to a node (with no copy) it encounters
- Faster than other limited copies schemes

2-hop routing: (limited copies)
- The source forwards one of its copies to a node (with no copy) it encounters
- Only the source forwards copies to others than the destination (➔ source controls the message spreading process)

Focus

Consider a DTN environment where the source is in **full control of message spreading: 2-hop relay**

- Determines who to pass a copy to. Intermediate nodes are not allowed to spread a message copy they may have to any node other than the destination. Robustness against intermediate node misbehavior or limitations.
- Limits the max number of copies in the network (overhead concern)
- Limits the lifetime of copies (overhead or QoS concern)

Model description

- *N+1* nodes moving within a square size of L^2.
- Exponential node inter-meeting times (i.e. the time elapsed between two consecutive encounters for a given pair of nodes - fairly accurate in case of $R \ll L$, and random waypoint model)
- Mean rate of encounters λ for a given node (exp parameter):

  $\lambda = cvR/L^2$,

  c: constant depending on the mobility model used

  v: relative speed

  R: communication range

  L: size of the network area

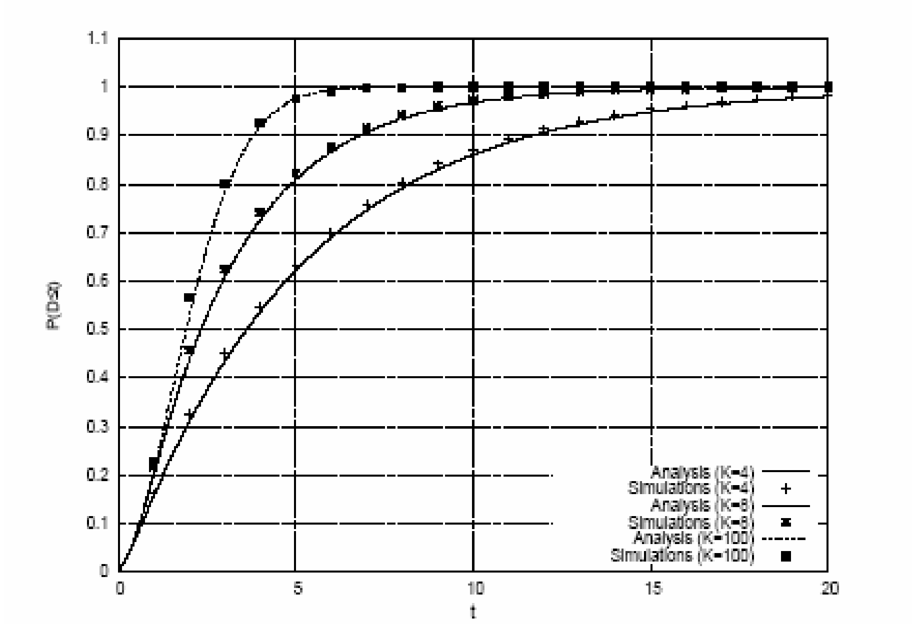
Source – intermediate node differentiation

- Diverse inter-meeting times for the source (λ) and the other nodes (λ_o)
  - May capture diverse transmission range (power) or/and speed or even, indirectly, the cooperation degree
- The parameters capturing differentiation could be considered as:
  - Non-tunable (e.g., misbehaviour)
  - Tunable (e.g., adjustment of transmission range)

Theory and simulations

Exact analytical expressions for the delivery delay cdf

Closeness of simulation with analytical results indicates the validity of the exponential encounter times for nodes moving under the random Waypoint model



Results are shown for: N=100, K=4, 8, 100, λ=0.08 and λo=0.04

Closed form approximation for delay *cdf*

A much simpler expression that approximates fairly accurately the exact one is derived by bounding the *accurate cdf* (for a specific number of copies K ≤ N) by two *cdfs*:

- the *maximum-copy cdf*
- the *zero-spreadtime cdf*

The maximum-copy *cdf*

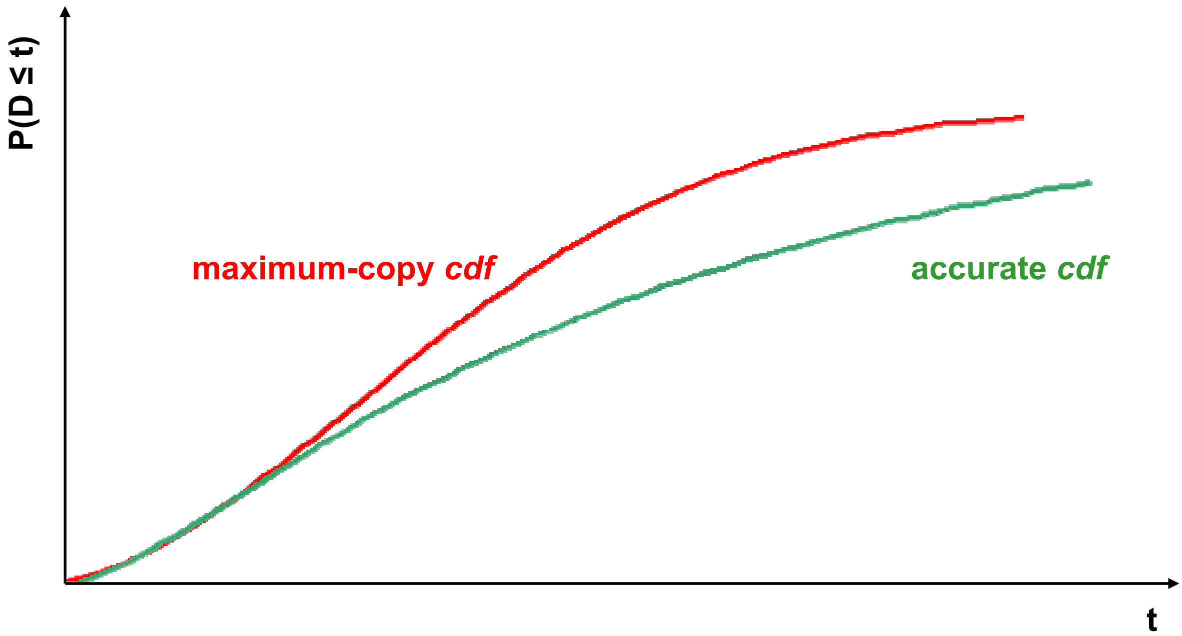
It refers to a modified algorithm where the number of copies employed in the network equals the number of nodes (K=N):

$$Q_N(t) = 1 - e^{-\lambda N t}\left(1 + \frac{\lambda}{\lambda_d}\left(e^{\lambda_d t} - 1\right)\right)^{N-1}$$

The modified algorithm has exactly the same behaviour until the first K copies are spread in the network

afterwards, the performance is enhanced due to the advantage of the surplus copies (N-K)

The maximum-copy *cdf*



The zero-spreadtime *cdf*

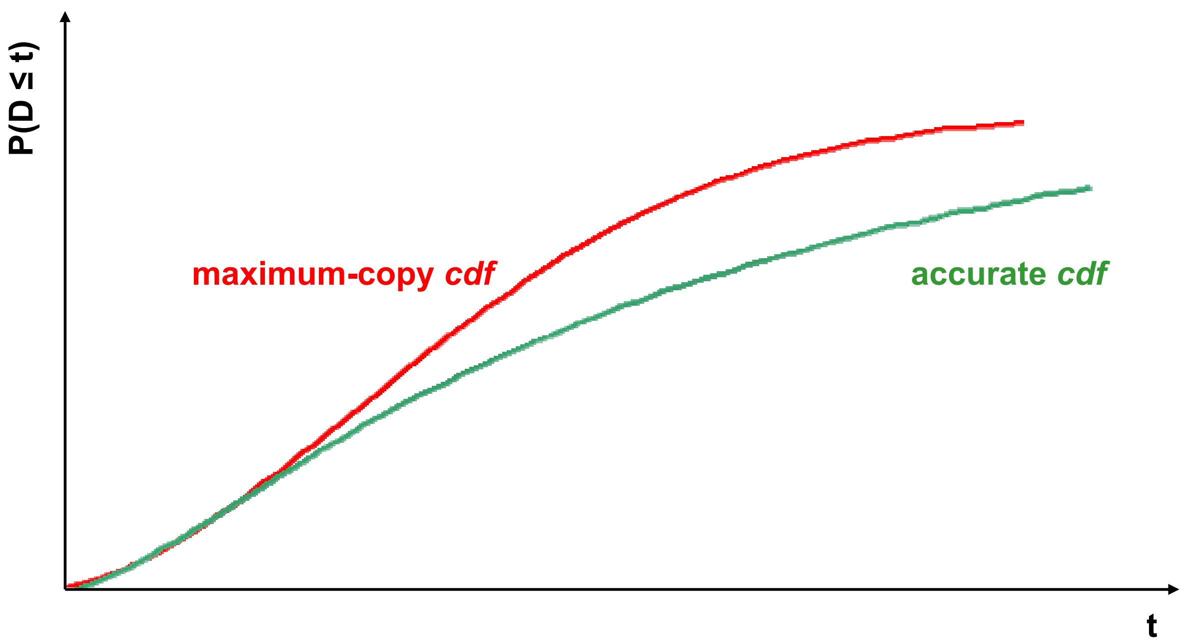
It refers to a modified algorithm where all K copies are instantly spread in the network

$$Q_K(t) = 1 - e^{-(\lambda+\lambda_o(K-1))t}$$

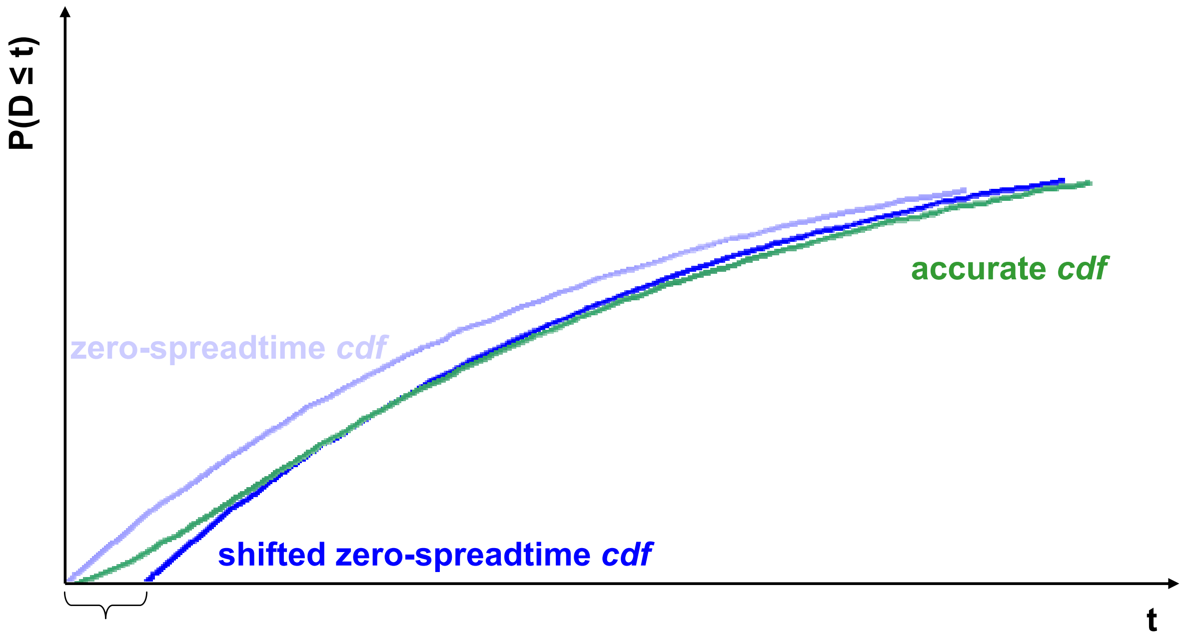
The modified algorithm has identical behavior with the original after the original has spread all K copies. Better performance for small t, converging to the original one for large t.

It is expected and indeed observed that when this *cdf* is shifted by t_o to be tangent at some time t_{cr} to the maximum-copy one, the part of the *cdf* from t_{cr} and afterwards approximates accurately the original *cdf*

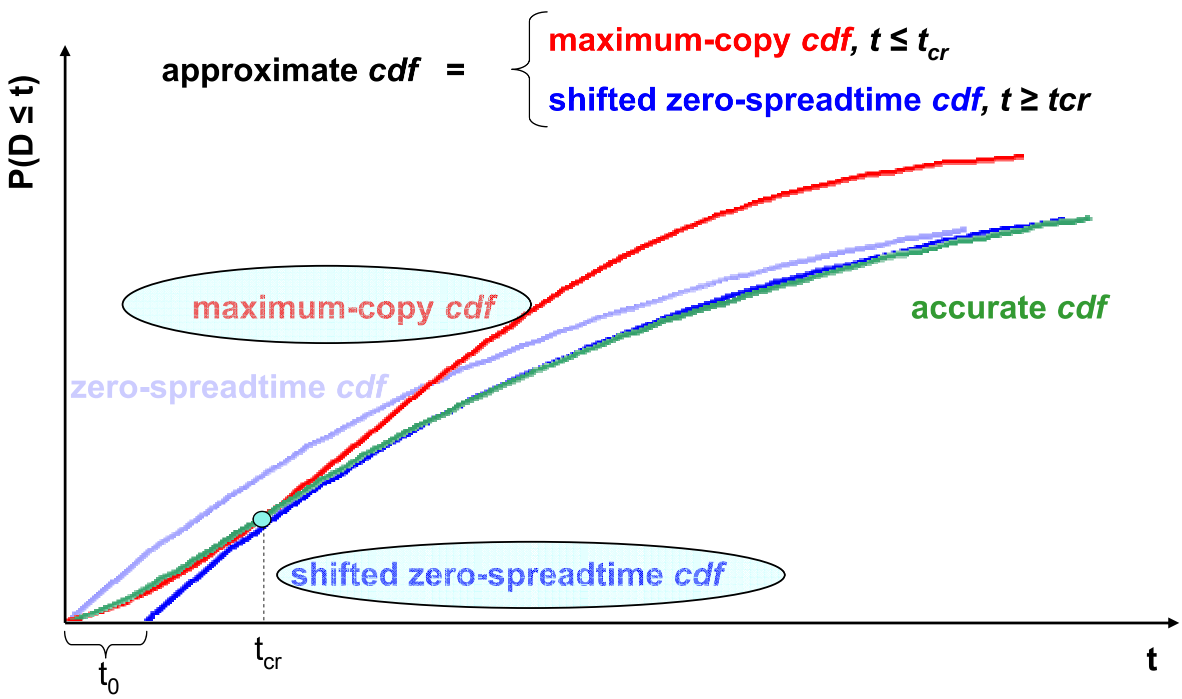
Constructing the approximate *cdf*



Constructing the approximate *cdf*



Constructing the approximate *cdf*



The approximate *cdf*

Thus, the approximate expression may be defined as a two-part function:

- the maximum-copy *cdf* until t_{cr}
- the shifted zero-spreadtime one after t_{cr}

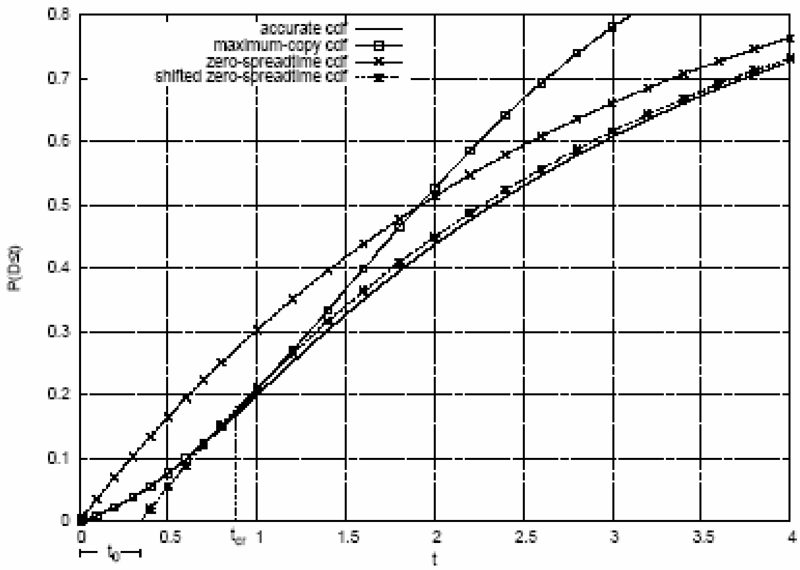
$$\hat{Q}_K(t) = \begin{cases} Q_N(t) = 1 - e^{-\lambda N t}\left(1 + \frac{\lambda}{\lambda_d}\left(e^{\lambda_d t} - 1\right)\right)^{N-1}, \\ 0 \le t \le t_{cr}; \\ Q_{\acute{K}}(t - t_0) = 1 - e^{-(\lambda + \lambda_o (K-1))(t - t_0)}, \\ t \ge t_{cr}, \end{cases}$$

where: $\lambda_d = \lambda - \lambda_o$,

t_o is the time shift of the zero-spreadtime *cdf* needed to be tangent to the maximum-copy one

t_{cr} is the contact point of the above *cdfs*

Results for the approximate *cdf*



For the case of N=100, K=8, λ=0.08 and λ_o=0.04:

ΕΘΝΙΚΟ & ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
NATIONAL & KAPODISTRIAN UNIVERSITY OF ATHENS
WiOpt 2007, April 16-20, 2007, Limassol, Cyprus
DEPT. OF INFORMATICS & TELECOMMUNICATIONS

Solving design problems

The approximation may be used in order to obtain closed form solutions to design problem where the exact analysis allows only for a numerical solution

For instance, **the value of K to achieve a specific delivery ratio Q_d within a specific delay bound *t*** may be estimated:

$$K_{approx} = \frac{-\lambda_d(N + (N-2)(N+1)\lambda t)}{\lambda_o(N - 2\lambda t)} + \frac{\lambda(N + (N(N-1)-1)\lambda t + \ln(1 - Q_d)) \pm \sqrt{C}}{\lambda_o(N - 2\lambda t)},$$

where *C* is a function of the network parameters. The positive value that fulfills the condition $t_o \leq t_{cr}$ should be selected

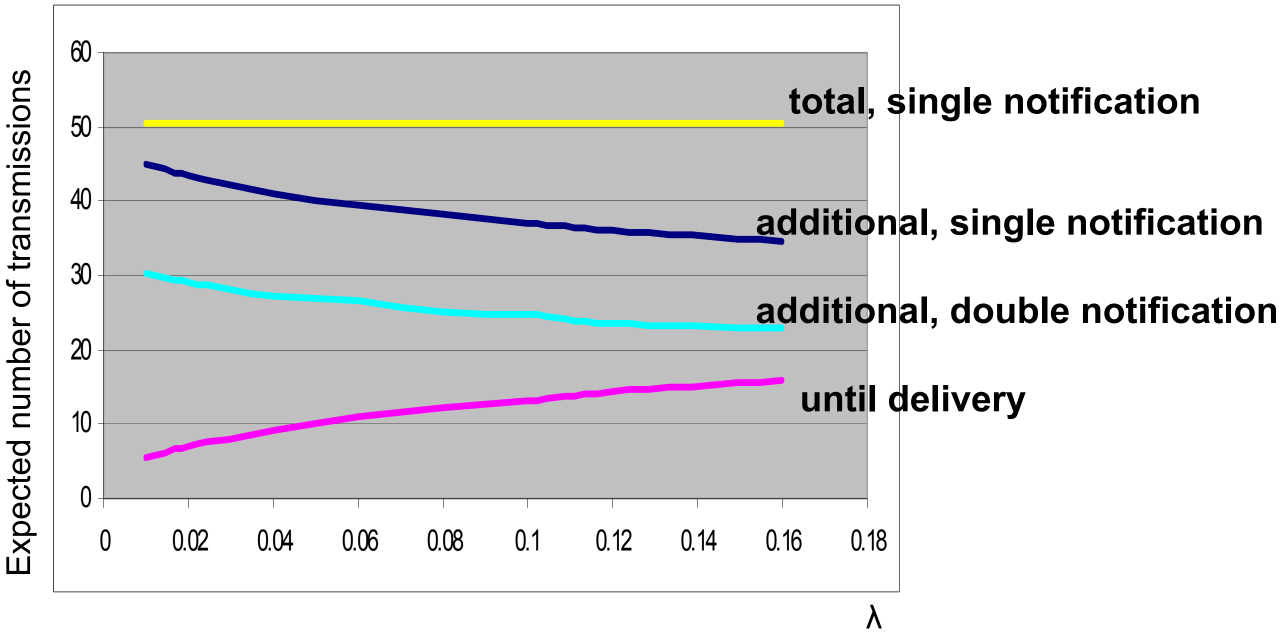
Calculation of the overhead

- In DTNs is typically measured in terms of the number of transmissions induced **until the message delivery** to the destination (or dropped if time constrained)
- **Here**, we also measure the **additional overhead** induced until the spreading process is actually terminated (the source node becomes aware of the message delivery through some notification mechanism)

  Here, the additional overhead is calculated for two distinct cases:
  - Until the source meets the destination (single notification)
  - Until it meets either the destination or the intermediate node that delivered the message to the destination (double notification)
- Besides the number of transmissions, energy overhead considerations are introduced for the case of a network of nodes that employ different transmission powers.

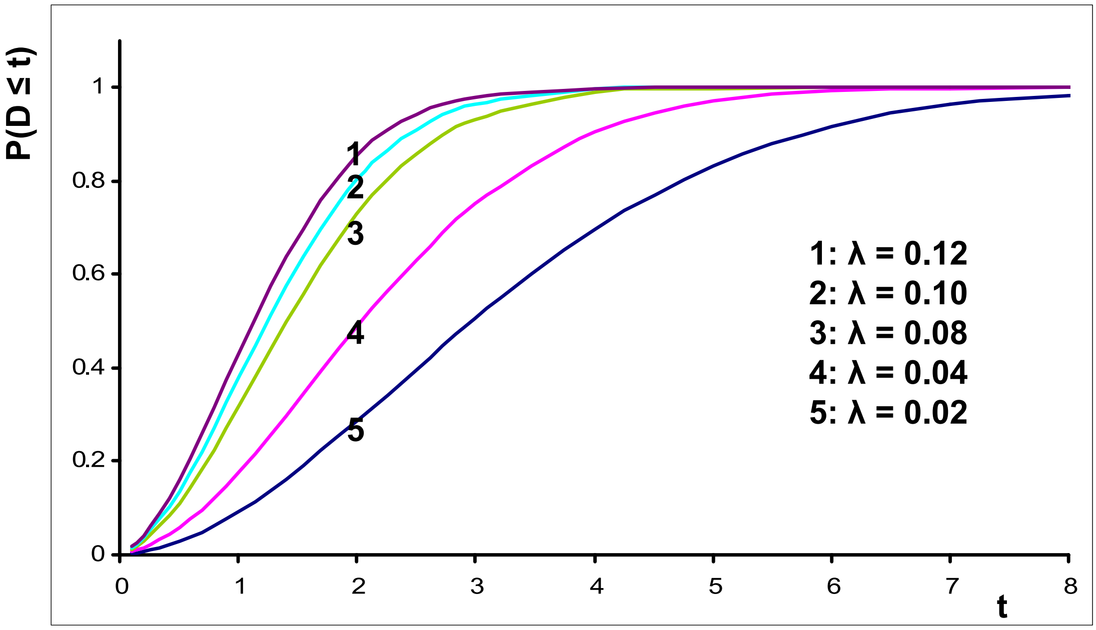
Results

Number of transmissions as a function of λ for the case of $K=N=100$ and λ_o=0.08



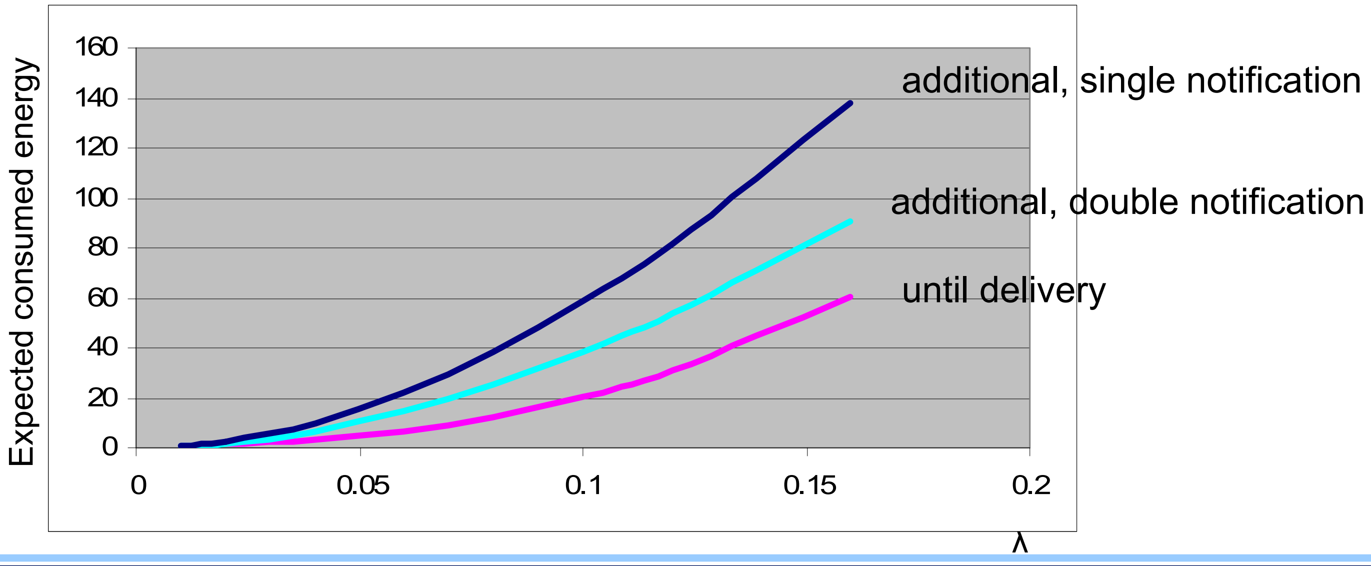
Results

Cdf of the delivery delay for the case of $K=N=100$ and λ_o=0.08

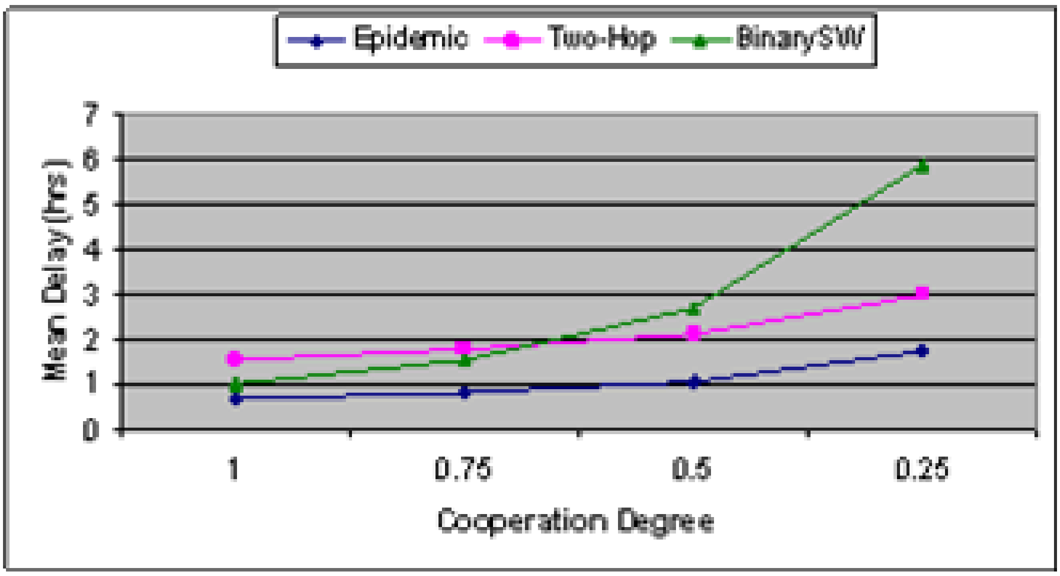


Results

- Energy consumed as a function of λ for the case of *K=N=100* and λ_o=0.08
- The energy consumed for a transmission is assumed to be proportional to the square of λ and equal to 1 for λ=0.08



2-hop relay outperforms BSW for low cooperation



Focus and Contributions

Study analytically the 2-hop relay algorithm (source-controlled spreading)

- cdf of message delivery delay – delivery ratio
- Closed form approximate cdf for delay, allowing for setting design parameters and shape the delivery ratio – overhead trade off.
- Consideration and calculation of the overhead not only until the message delivery but also until the message spreading is actually terminated.
- Differentiation between the source and intermediate nodes, capturing the effects of a more realistic, generally heterogeneous DTN environment, in terms of: transmission power, speed, cooperation degree, etc.
- Energy vs number of transmissions overhead considerations.

National & Kapodistrian University of Athens
Dept. of Informatics & Telecommunications

Additional slides – some details

Derivation of the *cdf*

The *pdf* of the unconditional total delivery delay *D* may be expressed as:

$$f_D(t) = \sum_{i=1}^{K} p_{d,i} f_{D_i}(t)$$

The *pdf* of the conditional delivery delay D_i given that the destination gets the message after *i* copies have been spread

Probability that *i* copies have been spread when the destination gets the message

$$p_{d,i} = \begin{cases} \frac{q(i,A)}{q_i} \prod_{j=1}^{i-1} \left(1 - \frac{q(i,A)}{q_i}\right), & i = 1, \ldots, K-1; \\ 1 - \sum_{j=1}^{K-1} p_{d,j}, & i = K; \\ 0, & \text{otherwise}, \end{cases}$$

where $q_i = q(i, i+1) + q(i, A)$.

Derivation of the *cdf*

Let T_j denote the sojourn time in state *i.* Then the *pdf* of the conditional total delivery delay $\left(D_i = \sum_{j=1}^{i} T_j\right)$ may be expressed as:

$$f_{D_i}(t) = \mathcal{L}^{-1}(F_i(s)), \text{ where } F_i(s) = \prod_{j=1}^{i} \mathcal{L}(f_{T_j}(t))$$

The Laplace transform of $f_D(t)$ may be expressed as:

$$F(s) = \sum_{i=1}^{K-1} B_i F_{1,i}(s) + B_K F_{1,K}(s)$$

$$B_i = \frac{q(i,A)}{q_i}\left(\prod_{j=1}^{i-1}\left(1 - \frac{q(j,A)}{q_j}\right)\right)\prod_{j=1}^{i} q_j$$
$$= (\lambda_d + (\lambda - \lambda_d)i)\lambda^{i-1}\frac{(N-1)!}{(N-i)!}, 1 \le i \le K-1$$

$$B_K = \left(1 - \sum_{i=1}^{K-1}\frac{q(i,A)}{q_i}\prod_{j=1}^{i-1}\left(1 - \frac{q(j,A)}{q_j}\right)\right)\prod_{j=1}^{K} q_j$$
$$= \lambda^{K-1}(K\lambda - (K-1)\lambda_d)\frac{(N-1)!}{(N-K)!},$$

$$\lambda_d = \lambda - \lambda_o$$

Derivation of the *cdf*

The Laplace transform of the *pdf* of *D* may be expressed as:

$$F(s) = \sum_{i=1}^{K-1} B_i F_{1,i}(s) + B_K F_{1,K}(s)$$

$$F_{1,i}(s) = \prod_{j=1}^{i} \frac{1}{q_j + s} = \begin{cases} \prod_{j=1}^{i} \frac{1}{\lambda N - \lambda_d (j-1) + s}, & 1 \le i \le K-1; \\ \frac{1}{\lambda + (\lambda - \lambda_d)(K-1) + s} \prod_{j=1}^{K-1} \frac{1}{\lambda N - \lambda_d (j-1) + s}, & i = K. \end{cases}$$

Derivation of the *cdf*

Since the *cdf* *Q(t)* is obtained by integration of the *cdf:*

$$Q(t) = \mathcal{L}^{-1}\left(\frac{F(s)}{s}\right)$$

Finally it may be concluded that

$$Q(t) = \sum_{i=1}^{K-1} B_i f_{2,i}(t) + B_K f_{2,K}(t)$$

$$f_{2,i}(t) = \begin{cases} \sum_{k=1}^{i} \frac{1}{\lambda_d (k-1) - \lambda N} \frac{1}{\lambda_d^{i-1}} \frac{1}{(k-1)!} \frac{(-1)^{i-k}}{(i-k)!} \left(e^{-\lambda N t} e^{\lambda_d (k-1) t} - 1\right), & 1 \le i \le K-1; \\ \sum_{k=1}^{K-1} \frac{1}{\lambda N - \lambda_d (k-1)} \frac{1}{\lambda_d^{K-2}} \frac{1}{(k-1)!} \frac{(-1)^{K-k-1}}{(K-k-1)!} + \frac{1}{\lambda_d^{K-2}} \sum_{k=1}^{K-1} \frac{1}{(k-1)!} \frac{1}{(K-1-k)!}, & i = K. \end{cases}$$

Deriving t_o

- We solve the equation $Q_N(t) = Q_{\acute{K}}(t - t_0)$
- By setting $y = \lambda t \text{ and } z = \lambda t_0$

$$e^{-y}\left(1+\frac{\lambda}{\lambda_d}\left(e^{\frac{\lambda_d y}{\lambda}}-1\right)\right)^{\frac{N-1}{N}} = e^{-\frac{(\lambda+\lambda_o(K-1))(y-z)}{\lambda N}}$$

- We expand the above equation in a Taylor series keeping the terms up to second order $1 + a_1 y + a_2 y^2 = b_0(1 + b_1 y + \frac{1}{2}(b_1 y)^2)$

where $a_1 = -\frac{1}{N}$, $a_2 = (\frac{\lambda_d}{2\lambda} + \frac{1}{2N^2} + \frac{1}{2N} - \frac{\lambda_d}{2\lambda N} - \frac{1}{2})$, $b_0 = e^{-b_1 z}$, $b_1 = (\frac{\lambda_d K}{\lambda N} - \frac{\lambda_d}{\lambda N} - \frac{K}{N})$

- In order for the polynomial to have a single root (so that the two *cdfs* are tangent to each other) it is required that its discriminant be zero or $(a_1^2 - 2a_1 b_1 + b_1^2) + (2a_1 b_1^2 - 4a_2 b_1)z + O(z^2) = 0$ leading to:

$$t_0 = \frac{\lambda_o(K-1)^2 N}{2\lambda(\lambda K - \lambda_d(K-1))((N-1)N - K + 1)}$$

Deriving t_{cr}

- We solve the equation $Q_N(t) = Q_{\acute{K}}(t - t_0)$
- By setting $y = \lambda t \text{ and } z = \lambda t_0$

$$e^{-y}\left(1+\frac{\lambda}{\lambda_d}\left(e^{\frac{\lambda_d y}{\lambda}}-1\right)\right)^{\frac{N-1}{N}} = e^{-\frac{(\lambda+\lambda_o(K-1))(y-z)}{\lambda N}}$$

- We expand the above equation in a Taylor series keeping the terms up to second order $1 + a_1 y + a_2 y^2 = b_0(1 + b_1 y + \frac{1}{2}(b_1 y)^2)$

t_{cr}

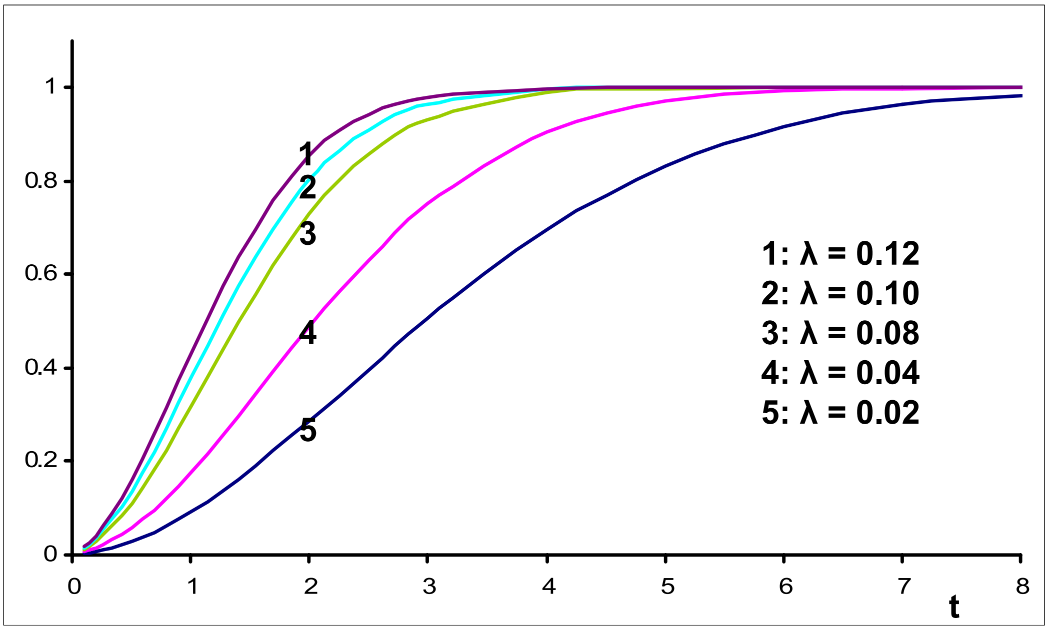
where $a_1 = -\frac{1}{N}$, $a_2 = (\frac{\lambda_d}{2\lambda} + \frac{1}{2N^2} + \frac{1}{2N} - \frac{\lambda_d}{2\lambda N} - \frac{1}{2})$, $b_0 = e^{-b_1 z}$, $b_1 = (\frac{\lambda_d K}{\lambda N} - \frac{\lambda_d}{\lambda N} - \frac{K}{N})$

- In order for the polynomial to have a single root (so that the two *cdfs* are tangent to each other) it is required that its discriminant be zero or $(a_1^2 - 2a_1 b_1 + b_1^2) + (2a_1 b_1^2 - 4a_2 b_1)z + O(z^2) = 0$ leading to:

$$t_0 = \frac{\lambda_o(K-1)^2 N}{2\lambda(\lambda K - \lambda_d(K-1))((N-1)N - K + 1)}$$

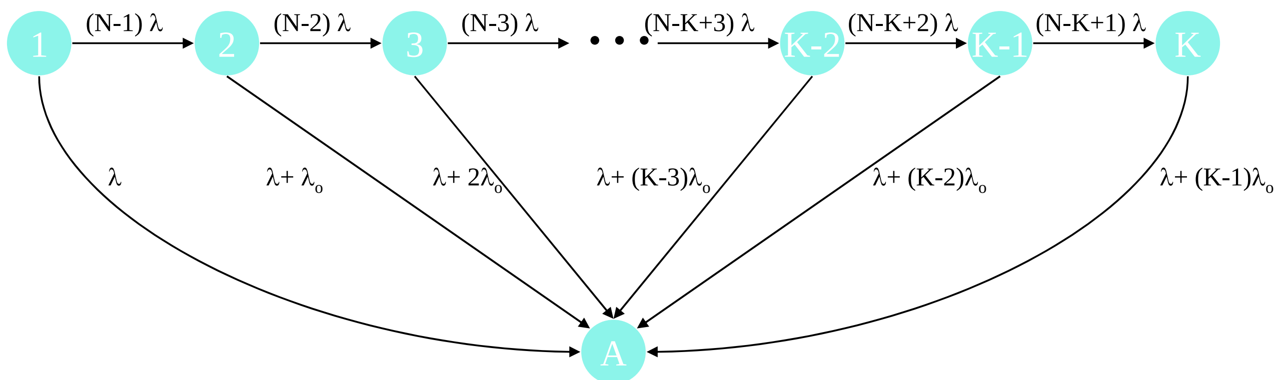
Results

Cdf of the delivery delay for the case of $K=N=100$ and λ_o=0.08



Derivation of the *cdf*

Based on the following Markov chain:



- K states capturing the number of copies spread in the network ($K \leq N$)
- one absorbing state A visited when the message is delivered to the destination

Overhead until delivery or drop (1)

The expected overhead induced until the delivery or drop:

$$\overline{O}_{del|drop} = \sum_{i=1}^{K} \overline{O}_{del|drop,i} P_{del|drop,i}$$

where $\overline{O}_{del|drop,i}$ denotes the expected overhead consumption provided that the message is delivered or dropped when in state *i*, and $P_{del|drop,i}$ denotes the probability that the system is in state *i*, when the destination is reached or the message is dropped

Overhead until delivery or drop (2)

The expected overhead consumption provided that the message is delivered or dropped when the system is in state *i* may be expressed as

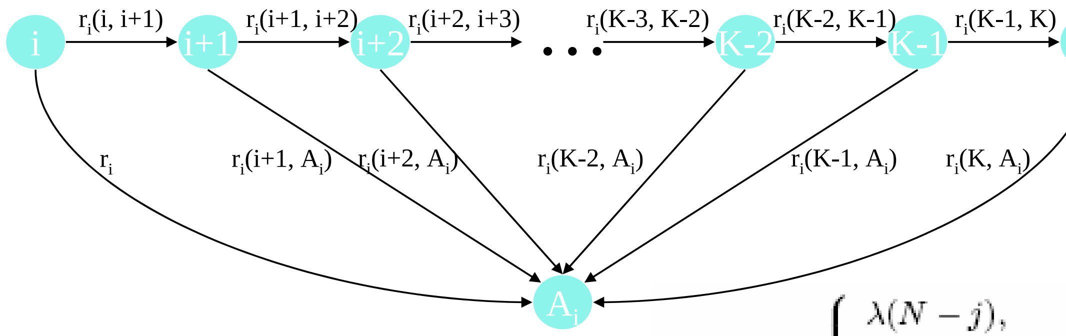
$$\overline{O}_{del|drop,i} = (i-1)E_s + P_{dels,i}E_s + P_{delo,i}E_o$$

where $P_{dels,i}$ ($P_{delo,i}$) denotes the probability that the source (some intermediate node) delivers the message to the destination provided that the message is delivered or dropped when the system is in state *i*

When different power levels for the source and the intermediate nodes are used, the parameters E_s and E_o are used respectively; they are equal to 1 when calculation refers to transmissions

Additional overhead (2)

- The calculation of the additional overhead is based on the following Markov chain, starting from state *i* (for the copies present in the network) up to state *K* and having an absorbing state A_i that corresponds to the case that the source node has been informed of the delivery success



$$r_i(j,k) = \begin{cases} \lambda(N-j), & j = i, \dots, K-1, k = j+1; \\ \lambda_{n,i} + \lambda_{n,d}, & j = 1, \dots, K, k = A_i; \\ 0, \ \text{otherwise}, & \end{cases}$$

Additional overhead (3)

The expected additional overhead may be expressed as

$$\overline{O}_{add} = \sum_{i=2}^{K} P_{del|drop,i} P_{delo,i} \overline{O}_{add,i}$$

where $\overline{O}_{add,i}$ denotes the expected additional overhead provided that the message is delivered by some intermediate node when the system is in state *i* $P_{del|drop,i}$ ($P_{delo,i}$) denotes the probability that the message is delivered or dropped (delivered by an intermediate node) when in state *i*

Additional overhead (4)

- The expected additional overhead provided that the message is delivered by an intermediate node when the system is in state *i* may be expressed as

$$\overline{O}_{add,i} = \sum_{j=i}^{K} \overline{O}_{add,i,j} P_{del|drop,i,j}$$

where $\overline{O}_{add,i,j}$ denotes the expected overhead provided that the source is notified or the message is dropped when the system is in state *j*, and $P_{del|drop,i,j}$ denotes the probability that the system is in state *j* when the source is notified or the message is dropped

Additional overhead (5)

- The term $\overline{O}_{add,i,j}$ may be expressed as

$$\overline{O}_{add,i,j} = (j-i)E_s + P_{nd,i,j}E_{nd} + P_{ni,i,j}E_{ni}$$

where $P_{nd,i,j}$ $(P_{ni,i,j})$ denotes the probability that the source is notified by the destination (the intermediate node that delivered the message) provided that the source is notified or the message is dropped when being in state *j*

- E_{nd} and E_{ni} denote the energy consumed for the transmission of the notification message by the destination node and an intermediate node, respectively

Conclusions

- There is a significant difference between the number of transmissions and the energy consumed in heterogeneous networks where different transmission powers among the nodes may be employed
- It may be concluded that the number of transmissions or consumed energy until the message delivery time are just a small portion of the corresponding totals and cannot be ignored
- The participation of the intermediate node in notifying the source node of the message delivery (double notification) limits noticeably the additional energy spent

Summing up

This work

- Provided a more general and realistic setting for studying the two-hop message spreading in heterogeneous DTNs
- Analytically derived the *cdf* and overhead of the algorithm
- Provided a fairly accurate approximation for the *cdf* that may be used in design problems
- Calculated the overhead under a more realistic framework (as should be employed in every study) by also taking the additional overhead into account
- Proved even a simple notification approach as the one introduced here can be proved to be a valuable mechanism (overhead-limiting)
- Made energy considerations to capture the performance of the algorithm in a network where different transmission power levels among the nodes may be employed