On the Effect of Cooperation in Wireless Content Distribution

Ólafur Ragnar Helgason and Gunnar Karlsson
ACCESS Linnaeus Center and School of Electrical Engineering
KTH, Royal Institute of Technology
Stockholm, Sweden
{olafurr,gk}@ee.kth.se

***Abstract*—In this paper we propose continuous time Markov chain models that capture the dynamics of content spreading in a disruption tolerant wireless content distribution system. We use our models to study the effect of cooperation among the mobile nodes and how limited node resources, such as battery lifetime and confined storage, affect the content distribution process. Based on our models and numerical results we deduce that limiting the number of times a node shares each content entry is a good method to conserve energy while at the same time only slightly reducing system performance. Our study also suggests that the effect of assisting nodes is greatest for content channels with few subscribers. For promoting fairness in distributing channels and giving new channels a chance to spread, assisting nodes should therefore solicit and help in spreading less popular channels.**

I. Introduction

In recent years there has been an enormous increase in the proliferation of small mobile devices carried by humans, vehicles or animals. As a part of this trend, handheld multimedia players have become immensely popular and these are increasingly being equipped with wireless radio such as WiFi or Bluetooth. Mobile phones today are also commonly multimedia capable. Content is traditionally delivered to these devices via docking stations or through contact with a fixed infrastructure such as a WiFi/Bluetooth Access Points or over GSM/3G.

Extending the content distribution for these devices into the ad-hoc domain has been proposed previously[1] [2]. When the mobile nodes are not connected to a fixed-infrastructure network or a docking station they operate in disruption tolerant mode. In this mode they utilize node-to-node contact opportunities which arise because of the node movements, to solicit content in a peer-to-peer manner. This way the nodes may collaborate in distributing content in areas that are not covered by an infrastructure network. The nodes are however generally limited in their resources, particularly battery power and storage, and therefore this collaboration comes at some cost.

In this paper we stochastically model different ways of cooperating in a wireless disruption tolerant content distribution system. We identify three types of cooperation schemes, each differing in the degree of how much the nodes cooperate.

- With type I cooperation nodes can only obtain content from fixed or dedicated infrastructure.
- In type II cooperation nodes share content they are privately interested in during the peer contacts
- With type III cooperation nodes solicit and share content on behalf of others as well as sharing their private content.

For the cooperation schemes we give stochastic models of the content spreading process based on continuous time Markov chains (CTMC). The Markov models capture the dynamics of how content is spread through the opportunistic contacts that arise because of node mobility. We extend our basic models to capture how the performance of the content spreading process is affected by resource limitations of the nodes. In particular we model how content can have limited lifetime at the nodes due to storage restrictions and how nodes can limit their sharing to cut down on the energy needed for transmitting shared content.

Our models are inspired by the mathematical theory of epidemic modelling which studies the spreading of infectious diseases among individuals [3], [4]. In particular it studies the dynamics of how healthy (susceptible) individuals become infected through contacts with infected individuals and how immunization or cure affects the spreading process. The analogy in wireless content distribution is the spreading of content from one or more source (infected) nodes to a group of content receivers (susceptibles) when nodes of each type are in communication range. Because of storage or energy constraints, nodes may cease to spread the content of interest thus becoming immune, which in this context is undesirable.

The contributions of this paper are as follows. We have developed continuous time Markov chains for studying cooperation in a wireless content distribution system inspired by the theory of epidemic modelling. We give basic models for each cooperation type and extend these to capture both storage and energy limitations at the mobile nodes. We report on the performance results for the content distribution by computing the mean time to absorption for the Markov chain models. Our results indicate that limiting the number of times a node shares each content item is a good method to conserve energy needed for sharing data while at the same time only slightly reducing the performance of the content distribution. Our study suggests that the effect of assisting nodes (as in type III cooperation) is greatest for content channels with few subscribers. For promoting fairness in distributing channels and giving new channels a chance to spread, assisting nodes should therefore solicit, and help in spreading, less popular channels.

The rest of the paper is structured as follows. Section II discusses related work. In section III we present our models and assumptions for the cooperation levels identified. In section IV we give performance results for the models in section III by analysis of the mean time to absorption for the Markov chains. In section V we conclude and discuss future directions for this work.

II. Related work

A wireless content distribution system that utilizes opportunistic contacts to exchange content between nodes has been proposed in [1] [2]. Also, [5] describes a system where information propagates on a contact basis between mobile users although it is targeted at seeking information in contrast to broadcasting information and relies heavily on infrastructure. In [6] the authors study cooperation in the mobile infostation system which, like the system we model, uses opportunistic node contacts to spread content. In the mobile infostation system, however, nodes cooperate through *social contracts*: content is not exchanged unless each node gets something it wants from the exchange. We do not consider social contracts in this work and assume that nodes act according to a tamper-proof protocol.

Performance of the 7DS system has been studied in [2] for different collaboration schemes and a power conservation strategy that consists of turning off the nodes wireless interfaces periodically. In the collaboration schemes of [2], nodes cooperate by sharing data, rebroadcasting and caching popular entries. The cooperation types that we identify in this work include both data sharing (type II) and caching popular entries (type III). In the system that we model there is however no abstraction for sending a message to a particular node, it is a pure receiver driven broadcast system and therefore we do not consider rebroadcasting of queries. Also, in contrast to turning off the wireless interface, the power conservation scheme that we consider consists of limiting the number of times a certain content entry is shared. Moreover, our models also capture storage limitations of the nodes. Finally, the study in [2] is mostly based on simulations although an initial model based on diffusion processes is introduced. Here we study the content distribution analytically with a stochastic Markovian model.

The mathematical field of epidemic modelling has a long history where both stochastic and deterministic models are used to study the spreading of infectious diseases [3], [4]. Epidemic modelling has received considerable attention from the networking research community, as there are many scenarios that arise which are analogous to the spreading of epidemics. Some examples are distributed computing and peer-to-peer networks [7] [8], database updates and maintenance [9], ad-hoc networks[10], multicast [11], [12] and more. The base Markov chain models for type I and II cooperation in this work are adapted from epidemic modelling. Then we extend and give variations of these models to capture limited node resources and to model type III cooperation.

In delay and disruption tolerant networks (DTN), models based on epidemic theory and/or Markov chains have been used to study the performance of different routing approaches [13], [14], [15]. In these works the models are mainly aimed at evaluating the tradeoff between message delivery delay and resource consumption, since it is desirable to limit the number of times a packet is copied/cached while still achieving acceptable transmission delay. In contrast to this we do not study routing of a message from a single source to a designated receiver. In the content distribution system that we model there is no explicit routing. Content is only solicited on a single-hop basis between two peering nodes.

A Markov model is applied in [13] to evaluate the tradeoffs in the two-hop multicopy and unrestricted multicopy DTN routing protocols. Moreover the authors show that assumption of independent and exponentially distributed inter-contact times, commonly used in the literature [16] [17], [14], is a good approximation for some common mobility models, such as the random waypoint and random direction models.

In [16] the authors develop a model based on ordinary differential equations (ODE) to study the performance of epidemic routing. The ODE models appear as fluid limits of Markovian models under appropriate scaling as the number of nodes grow. On the one hand, solutions of Markovian models become impractical when the number of nodes is large. On the other hand, Markov models more accurately capture the behavior of small systems and can be used to derive full distributions while the ODE approach evaluates the moments of the distributions.

Recently there have been empirical evidence presented suggesting that inter-contact times are not exponentially distributed but exhibit a heavy tail such as that of a power law distribution [18]. In [19] it is shown that the exponential inter-contact times exhibited by common mobility models is due to the finite domain on which they are defined and by removing the finite boundary assumptions power-law decay arises. Another recent work [20] suggests that inter-contact times feature a dichotomy. Up to a characteristic time they follow a power-law distribution but after that they decay exponentially and therefore the power-law decay suggested by [18] may be overly pessimistic. Finally, [21] advocates, by studying three reference DTN data sets, that pair-wise inter-contact times (as opposed to the aggregate considered in previous works) follow a log-normal distribution with both finite mean and variance, which is positive for the performance of DTN. As a conclusion to this discussion we are aware that the exponential assumption may in some cases provide overly optimistic results. It is however significantly more difficult to model the wireless content distribution without this approximation and in section IV we show by simulation that assuming exponentially distributed inter-contact times captures the qualitative behavior of the system.

III. Stochastic cooperation model

In this section we identify and model analytically three basic types of cooperation schemes for content distribution in a disruption tolerant network. We are interested in analyzing and comparing the performance of the content spreading process under each of the cooperative schemes and to evaluate the effect of limited battery lifetime and storage at the nodes.

We consider identical nodes moving in a closed area of size A. The transmission range of the nodes is r and we assume that r^2 is small in comparison to the area size A. We assume that content is arranged into channels and each channel consists of channel entries which are the actual data elements. Whenever two nodes are in communication range a channel entry can be transferred in zero time if one of the nodes is carrying a channel entry that the other one is interested in obtaining. This is a reasonable approximation if the content in the system is structured such that it consists of small units that can presumably be transmitted in a peer contact, such as the channel entries in [1] [22].

We identify three basic types of cooperation schemes. Each scheme differs in the degree of node-cooperation.

- Type I No cooperation in spreading the content among the nodes. Nodes can only obtain content when connected with a fixed infrastructure or in docking mode.
- Type II Nodes are willing to share private content that they currently have.
- Type III In addition to sharing content that the nodes are privately interested in, nodes are willing to solicit and carry content that is not of direct interest to themselves.

We define the *inter-contact time* as the elapsed time between two successive contacts of the same pair of nodes. In particular we denote by the random variable $\tau_{i,j}(n)$ the n-th inter-contact time between nodes i and j. The inter-contact times are determined by the mobility process of the nodes and in this work we make the following assumption:

> The sequence of random variables $\{\tau_{i,j}(n)\}_{i,j,n}$ are identical, independent and exponentially distributed with rate λ. (1)

This assumption allows us to model the content distribution process using continuous time Markov chains. If the transmission range r of the nodes is small in comparison to the area size A it has been shown that exponential inter-contact times are a good approximation for some common mobility models, such as the random waypoint and random direction models [13]. In section IV we compare the results from our analytical model with simulation results for non-exponentially distributed inter-contact times.

In the following subsections we model content distribution under each of the schemes and extend the models to indirectly capture the effects of battery power or limited storage at the mobile nodes. In the models that follow we study the content distribution process for a single channel entry that is brought into the area at time $t = 0$. A fundamental state variable in all our models is the number of nodes that have obtained the channel entry at time t which we denote by $X(t)$. In particular we are interested in the time it takes for a newly published channel entry to reach all the nodes in the area that are interested in it. This time is a measure of how fast the content is distributed and we denote it by T_c and refer to it as *content spreading time*. We assume that channel interests of the node population do not change under the timescales we consider. Then we have for a given channel that T_c is an identical and equally distributed random variable for all entries of the channel.

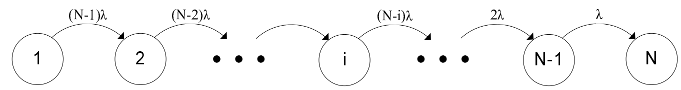


Fig. 1. Markov chain for type I cooperation.

Borrowing terms from epidemic modelling, the nodes in our model can be classified as *susceptible*, *infected* or *recovered*. In epidemic modelling a susceptible individual will become infected by the epidemic disease when it meets an infected individual. After some time an infective may recover from the disease (or die) and become immune and thus no longer infect other susceptibles. Analogously in our model, whenever a node that has a channel (infected) makes a radio contact with an interested node that does not have it (susceptible) the latter node obtains the channel. Because of limited energy or storage a node may stop distributing a channel, thus moving from an infected state into a recovered.

A. Type I: No sharing

Under this scheme there is no collaboration between the individual nodes in distributing the channel. Nodes can only obtain the channel contents when docking, when connected to infrastructure or to a dedicated content provider. This non-cooperative model does thus not capture the relaying effect of the content spreading process. It however serves as a worst-case baseline in performance comparison with the cooperative models.

We assume that there are N nodes in the area that are interested in a particular channel. At time $t = 0$ a single node obtains a new channel entry offline. We assume that this node is a dedicated content provider in the sense that it is willing to provide the channel entry to other nodes is meets. Apart from that it is identical to the other nodes and in particular, assumption (1) is valid. Since the nodes are not cooperating the only way to obtain the content for the $N - 1$ susceptible nodes is by infection from the content provider node.

We denote by the random variable $X(t)$ the number of infected subscribers at time t. The process $\{X(t); t \geq 0\}$ is a pure birth process with rate $\lambda_i = (N - i)\lambda$ for all states $i = 1, ..., N-1$ as shown in Fig.1. The CTMC for this process is absorbing with $i = N$ as the absorbing state and all other states transient.

The state sojourn time in a CTMC is exponentially distributed with rate equal to the sum of the rates going out of the state. Therefore the mean time to absorption from the initial state $i = 1$ is easily calculated in this case as the sum of the time spent in each of the transients states. We have that

$$E[T_c] = \sum_{i=1}^{N-1} \frac{1}{\lambda_i} = \frac{1}{\lambda} \sum_{i=1}^{N-1} \frac{1}{(N-i)} = \frac{1}{\lambda} \sum_{j=1}^{N-1} \frac{1}{j} = \frac{1}{\lambda} H_{N-1} \tag{2}$$

where $H_n = \sum_{i=1}^{n} \frac{1}{i}$ is the n-th harmonic number. It is well known that the harmonic series does not converge when $n \to \infty$. An asymptotic expansion for the harmonic numbers is $H_n = \gamma + ln(n) + \mathcal{O}(\frac{1}{n})$ where γ is Euler's constant. Thus we have

$$E[T_c] = \frac{1}{\lambda}(\gamma + ln(N-1) + \mathcal{O}(\frac{1}{N-1})) \tag{3}$$

from which we deduce that $E[T_c] \in \mathcal{O}(ln(N))$.

For this simple case we can get the probability distribution for $X(t)$ for a general inter-contact time distribution as long as we assume that the inter-contact times are i.i.d. I.e. the distribution of $X(t)$ can be deduced in a closed form without assuming that the inter-contact times are exponentially distributed.

If we have initially that $X(0) = m$ there are m nodes infected and thus $N - m$ susceptible nodes interested in obtaining the channel. Since the $N-m$ susceptible nodes can only obtain the channel from one of the m initially infected nodes the number of infected nodes is given by the binomial distribution, namely

$$P\{X(t) = j+m\} = \binom{N-m}{j} q(t)^j (1-q(t))^{(N-m)-j} \tag{4}$$

where $q(t)$ is the probability that, in time t, a susceptible node obtains the channel from at least one of the m infected nodes. For a a given susceptible node $i \in \{m+1, m+2, ..., N\}$ we thus have

$$q(t) = P\{min(\{\tau_{i,k}\}_{k=1...m}) \leq t\} \tag{5}$$

(For clarity and without loss of generality we have assumed that the m initially infected nodes are labelled from $1..m$ and the susceptible node as $m{+}1, m{+}2, ..., N$). In epidemic theory (4) is known as the Reed-Frost model [4].

If we now return to assumption (1), namely that the inter-contact times are exponentially distributed with rate λ, we have

$$q(t) = F_\tau(t) = 1 - e^{-m\lambda t}$$

Thus the average number of infected nodes at time t is given by

$$E[X(t)] = m + (N-m)(1 - e^{-m\lambda t}) \tag{6}$$

and when there is initially a single node that has the channel entry of interest we have $m = 1$ or

$$E[X(t)] = 1 + (N-1)(1 - e^{-\lambda t})$$

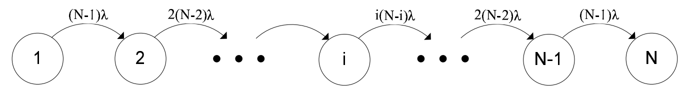


Fig. 2. Markov chain for type II cooperation with infinite storage.

B. *Type II: Cooperative sharing*

With this degree of cooperation nodes share the channels that they are privately subscribed to. The mobile devices we expect to be involved in the content distribution are small, resource limited devices such as mobile phones, small audio/video players, PDA's etc. Therefore we presume that the resource limitations of these devices will restrict the capability of the nodes to cooperate. Limited battery lifetime and storage will in particular discourage cooperative transmissions and caching of data.

In this subsection we describe three different models of which the latter two capture the nodes resource limitations. The *infinite storage* model is a baseline model that assumes that once a node has obtained a channel entry it will keep and share it forever. In the *limited storage* model nodes may delete and cease sharing a channel entry after some time due to shortage of storage. In the *limited upload* model nodes will only share each channel entry k times therefore limiting the energy needed for the transmission of the shared channels. We present this model for the case $k = 2$.

Infinite storage: As before we assume that there are N nodes in the area interested in the channel. Denote by $X(t)$ the number of *infectives* at time t and thus $N - X(t)$ is the number of *susceptibles* for $t > 0$. We assume that at $t = 0$ a single node obtains the content offline and thus $X(0) = 1$.

The process $\{X(t); t \geq 0\}$ is a pure-birth Markov process with positive transition rates

$$\begin{array}{ccc} \text{From} & \text{To} & \text{Rate} \\ i & i+1 & i(N-i)\lambda \end{array} \tag{7}$$

The transition diagram of the Markov chain is shown in Fig.2. State $i = N$ is an absorbing state of the Markov chain and in this state all the susceptible nodes have become infected with the content. We are interested in the mean content distribution time $E[T_c]$. This is the same as the mean time to absorption for the Markov chain, starting from state $i = 1$.

Let $R_{i,i+1}$ denote the time that it takes for the process, starting from state i, to reach state $i+1$, $i \geq 1$. $R_{i,i+1}$ is exponential with rate $(N-i)\lambda$ and thus

$$E[R_{i,i+1}] = \frac{1}{\lambda_i} = \frac{1}{i(N-i)\lambda}$$

The time it takes all the nodes to obtain the content is $T_c = R_{1,N}$ and its expected value is $E[R_{1,N}] = E[R_{1,2}] + E[R_{2,3}] + \cdots + E[R_{N-1,N}]$. We thus have

$$E[T_c] = E[R_{1,N}] = \sum_{i=1}^{N-1} E[R_{i,i+1}] = \frac{1}{\lambda} \sum_{i=1}^{N-1} \frac{1}{i(N-i)} \tag{8}$$

and since

$$\sum_{i=1}^{N-1} \frac{1}{i(N-i)} = \frac{1}{N}(\sum_{i=1}^{N-1} \frac{1}{i} + \sum_{i=1}^{N-1} \frac{1}{N-i}) = \frac{2}{N} H_{N-1} \quad (9)$$

we have the following asymptotic expansion for $E[T_c]$

$$E[T_c] = \frac{2}{\lambda N}(\gamma + ln(N-1) + \mathcal{O}(\frac{1}{N-1})) \quad (10)$$

where γ is Euler's constant. Since $\lim_{x\to\infty} \frac{k_1 ln(x-1)}{x} = 0$ and $\lim_{x\to\infty} \frac{k_2}{x} = 0$, where k_1 and k_2 are constants, we have that $E[T_c] \in \mathcal{O}(1)$.

Thus when the nodes utilize the peer contacts to cooperatively share content the mean content distribution time is bounded by a constant and does not grow to infinity when $N \to \infty$ as is the case when there is no peer cooperation (3). This confirms what has previously been established in [17]. The Markov chain in (7) is a variation of the *simple stochastic Markovian epidemic* [3].

Limited storage: Now we relax the infinite storage assumption of the previous subsection. We denote by $Y(t)$ the number infected nodes at time t. Assume that when a node becomes infected and obtains the channel entry it will keep it for an exponentially distributed time with rate μ before deleting it. When it has deleted the channel the node can not infect other nodes anymore and thus becomes immune. Denote by $X(t)$ the number of infected plus immune nodes at time t.

We have the initial conditions $X(0) = 1$ and $Y(0) = 1$. The process $\{(X(t), Y(t)); t \geq 0\}$ is a two dimensional Markov chain and its positive state transition rates

From	To	Rate	
(i, j)	$(i+1, j+1)$	$j(N-i)\lambda$	(11)
	$(i, j-1)$	$(j-1)\mu$	

This Markov chain is absorbing and the absorbing states are $\{(X(t), Y(t)) = (N, j); \forall j\}$. We point out that we do not allow the content to vanish from the area since we assume that $1 \leq Y(t) \leq X(t)$ for all t. In other words, we assume that the node that brought the content in initially never deletes it.

Limited upload: When a node shares a channel entry with a peering node some energy will be spent on the data upload. It is expected that many of the nodes in the content distribution system will be battery powered and thus have limited energy resources. Therefore it seems natural for a node to limit its uploading to others to save its own energy. One way to accomplish this is to restrict the uploading such that the node will share each channel entry only k times.

Let $X(t)$ be the number of infected and immune nodes at time t, and $Y_m(t)$ the number of infected nodes who have to share the channel m times before becoming immune, where $m = 1,..,k$. Thus right after a susceptible node obtains the channel it will become infected and since it has never shared its channel entry Y_k will increase by one. Also, after an infected node that has shared its channel n times ($n = 1...k-1$) meets a susceptible node Y_{k-n} will decrease by one and Y_{k-n-1} will increase by one.

When $k = 1$ each node will only share the channel once. Thus at each time there is only a single node who is willing to share the content and the process can be modelled by the one-dimensional Markov chain $\{X(t); t \geq 0\}$. This is evident from the fact that when content is shared the sharing node will become immune and the receiver will be infected until it meets one of the remaining susceptible nodes. The absorbing state and the transition rates for the Markov chain are thus the same as for the type I system with no cooperation in Fig.1). It should also be clear that when $k \to \infty$ a node will share a channel whenever it is asked for and then this scenario is the same as the type II infinite storage system described previously in (7).

For a general $k > 1$ the Markov process for the system will consist of $k+1$ random variables since we have to count the number of nodes for each of the sequence of variables $\{Y_m(t); t \geq 0\}_m$. Here we give the model for the case $k = 2$. Initial conditions are $X(0) = Y_2(0) = 1$, $Y_1(0) = 0$. The positive state transition rates for the process $\{(X(t), Y_1(t), Y_2(t)); t \geq 0\}$ are

From	To	Rate	
(i, j, k)	$(i+1, j+1, k)$	$k(N-i)\lambda$	(12)
	$(i+1, j-1, k+1)$	$j(N-i)\lambda$	

C. Type III: Generous sharing

In this section we model the scheme where node cooperation is generous in the sense that nodes are willing to assist in distributing other channels than those they are privately interested in. Thus in addition to sharing the *private channels* a node is interested in it is willing to solicit and cache *public channels* for the benefit of others.

Assume that we have N nodes in the area that are privately subscribed to a channel, herein referred to as *subscribers*. Assume moreover that there are M other nodes in the area which are willing to assist in spreading the channel. Since these M nodes do not have private interest in the channel we refer to them as *assistants*. We model this system as a CTMC where the following events cause state transitions

(a) Infected subscriber meets susceptible subscriber
(b) Infected subscriber meets susceptible assistant
(c) Infected assistant meets susceptible subscriber

Note that assistants never infect each other, they can only be infected in a contact with an infected subscriber. This is similar to *two-hop multicopy forwarding* [13] [17]. More aggressive spreading such as *unrestrictive multicopy forwarding* [13] would presumably result in faster spreading but at the expense of requiring more resources from the participating nodes.

As in type II cooperation the assistants may have resource limitations which restrict their content distribution capabilities. Here we provide models for *infinite assistant storage* and *limited assistant upload*. In both cases we assume infinite storage at the subscribers but extending the models to capture limited resources at the subscribers is a straightforward extension of the models in the previous subsection. Limiting storage for the assistants is a straightforward extension of model 11.

We denote by $X(t)$ the number of infected and immune subscribers. Denote by $Y(t)$ the number of infected and immune assistants at time t and by $Z(t)$ the number of infected assistants at time t.

Infinite assistant storage: In this baseline model each assistant has infinite storage. Thus after it once obtains a public channel it will redistribute it forever. With this assumption we have $Z(t) = Y(t)$ for all t and thus the Markov chain for this model is two dimensional.

At time $t = 0$ a single subscriber obtains the channel offline, giving initial conditions $X(0) = 1$, $Y(0) = 0$. The positive transition rates for the process $\{(X(t), Y(t)); t \geq 0\}$ are

From	To	Rate	Event
(i, j)	$(i+1, j)$	$i(N-i)\lambda + j(N-i)\lambda$	(a) and (c)
	$(i, j+1)$	$i(M-j)\lambda$	(b)

(13)

When $M = 0$ this model is equal to the type II cooperation infinite storage model in (7).

Limited assistant upload: As discussed for the type II limited upload case, nodes may want to limit the number of times they share a channel to conserve energy and storage. Here we give a model for the case when each assistant is willing to share a public channel exactly once. Initial conditions for this case are $X(0) = 1$, $Y(0) = 0$ and $Z(0) = 0$ and the positive state transition rates for the process $\{(X(t), Y(t), Z(t)); t \geq 0\}$

From	To	Rate	Event
(i, j, k)	$(i+1, j, k)$	$i(N-i)\lambda$	(a)
	$(i, j+1, k+1)$	$i(M-j)\lambda$	(b)
	$(i+1, j, k-1)$	$k(N-i)\lambda$	(c)

(14)

IV. Numerical results

In this section we compare and contrast the cooperation models from the previous section. All the CTMC's in the previous section are absorbing and they all reach an absorbing state when the number of infected plus immune subscribers reaches N, i.e. when all subscribers have obtained the channel of interest. Thus the time it takes for the Markov chain to reach an absorbing state from the given initial state is a performance measure of how fast the content spreads for a given cooperation scenario and this is the same as the *content spreading time* previously defined.

For the most simple models we can solve for content spreading time explicitly as in equations (3) and (10). When finding a closed form solution is not feasible we resort to calculating the mean time to absorption numerically.

Absorption analysis of continuous time Markov chains

We denote by $\mathbf{Q}$ the transition-rate matrix of a CTMC. For an absorbing continuous time Markov chain with a single absorbing state the matrix $\mathbf{Q}$ can be decomposed as

$$\mathbf{Q} = \begin{pmatrix} \mathbf{T} & \mathbf{A} \\ \mathbf{0} & 0 \end{pmatrix} \tag{15}$$

where $\mathbf{T}$ is the square matrix of transient-to-transient state transition-rates and the column vector $\mathbf{A}$ are the rates from the transient states to the single absorbing state. If the initial state distribution among the transient states is given by the vector $\mathbf{p_0}$, the mean time to absorption for the chain is

$$E[T_{absorbtion}] = \mathbf{p_0}(-\mathbf{T})^{-1}\mathbf{1} \tag{16}$$

where $\mathbf{1}$ is a column vector of ones. For a more thorough discussion see for example [23].

Some of the Markov chains in the previous section contain more than a single absorbing state since all the states where $X(t) = N$ are absorbing. Equation (16) assumes that there is a single absorbing state and therefore we collapse all the absorbing states into a single state before extracting the matrix $\mathbf{T}$ from $\mathbf{Q}$. Numerically solving (16) consists of the following steps

- If the Markov chain is more than one-dimensional we renumber the states with a 1-dimensional linear index.
- Collapse absorbing states into a single state.
- Renumber states such that $\mathbf{Q}$ is on the form (15).
- Extract the matrix $\mathbf{T}$ from $\mathbf{Q}$ and solve (16).

For the two and three dimensional Markov chains in the previous section the state space grows quite fast as the number of nodes increases. For example, in the type III limited assistant upload model in (14) the size of the $\mathbf{Q}$ matrix grows as $O(N \cdot M^2)$. Therefore we can only solve (16) when the number of nodes N and assistants M is on the order of tens. However, with systems of this size the performance differences between the individual cooperation schemes are very evident. Here we also point out that there might of course be more than $N + M$ nodes in the area. But since we assume that the content distribution for the channel is only affected by the subscribers and assistants for the channel we do not need to consider other nodes in our models.

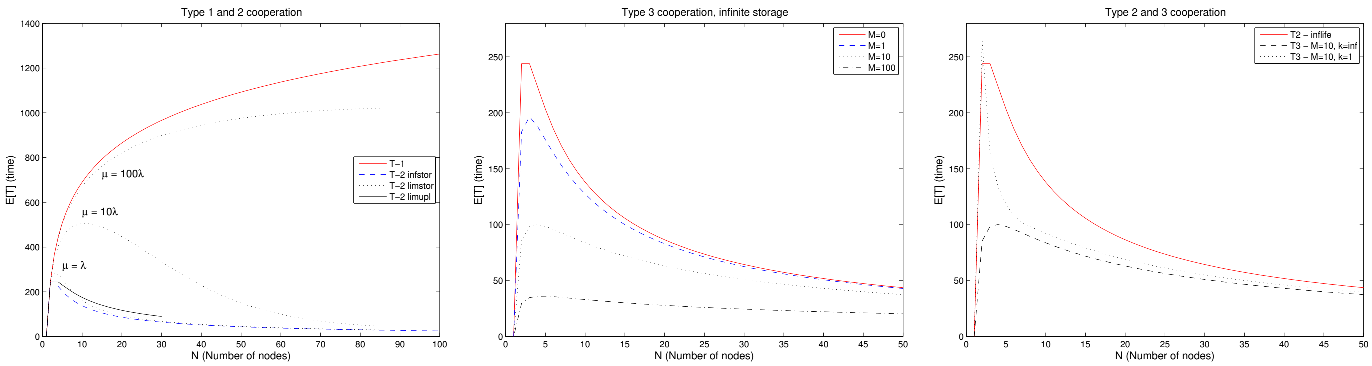
Results

In the following discussion we assume that the nodes in the area are moving according to the random waypoint mobility model [24] and that the average relative speed between two nodes is $E[V^*] = 1.5$ m/s. We assume that the area is a square with sides of length $L = 100$ meters and that the communication range of the nodes is $r = 10$ meters. This could for example represent a square in a city where pedestrians are carrying a device equipped with a short-range radio such as Bluetooth. In [13] the authors show that for the random waypoint and random direction mobility models the inter-contact rate λ is approximated by

$$\lambda \approx \frac{2\omega r E[V^*]}{L^2} \tag{17}$$

where ω is constant specific to the mobility model and $\omega \approx 1.3683$ for the random waypoint model. Thus for the parameters above we have $\lambda = 0.0041 \ s^{-}1$ and then the average inter-contact time is $1/\lambda = 243.9$ s or roughly 4 minutes.

In Fig.3(a) we have plotted the mean content distribution time for type I cooperation and for type II cooperation with infinite storage, limited storage and limited uploading. For the



(a) (b) (c)

Fig. 3. Content distribution time for different cooperation types and different resource constraints. 3(a): Type I and Type II cooperation with infinite storage (infstor), limited storage (limstor) and limited upload (limupl). 3(b): Type III with infinite storage for different number of assistants (M). 3(c): Type III cooperation with infinite assistant storage (k=inf) and limited upload (k=1).

time-limited storage case we have plotted for $\mu = \lambda, 10\lambda$ and 100λ. As expected from (3) and (10) we clearly see that if the nodes are willing to cooperate by sharing the content through the peer contacts the content spreading time is significantly reduced. Type I cooperation gives the worst-case performance for all cases considered and from the closed-form solution for $E[T_c]$ (3) we know that is does not converge as $N \to \infty$. For the type II cooperation with limited memory (11) we see that when $\mu \to \infty$ the mean lifetime of the channels in storage tends to zero and the system tends to one of type I with no cooperation between the nodes. When $\mu \to 0$ the system tends to a type II system with infinite storage. Due to the reasons mentioned previously we are only able to numerically solve the type II limited upload model in (12) for a maximum of $N = 30$ nodes. It is however interesting to see for these values of N, that if each node is willing to share the content only twice, the mean content distribution time is only slightly longer then what it is for the infinite storage case. A benefit with the limited upload approach is also that nodes can determine the maximum amount of energy required for sharing content. If we say that it costs one energy unit to share a channel entry the maximum upload cost in energy units equals the number of channel entries it is carrying that have never been shared. This maximum upload cost is thus independent of node density and contact rate which is not the case for the limited storage models.

In Fig.4 we study the effect of the exponentially distributed inter-contact times assumption. We compare results of our model for type I and II cooperation with simulation results where the inter-contact times follow a log-normal distribution with the same mean and variance of $1/\lambda$ and $1/\lambda^2$ respectively. It has recently been showed that log-normally distributed inter-contact times are a good match for three reference DTN contact traces[21]. A power-law (Pareto) inter-contact distribution with the same mean of $243.9s$ has infinite variance and, as expected, our simulation results for this distribution give highly varying results, differing by orders of magnitude. Therefore we do not include these results here. It is not surprising that the numerical results show some differences and that system performance is worse for the Log-Normal case. However, the assumption of exponentially distributed inter-contact times captures the qualitative behavior of the system, in particular when comparing the difference between cooperation types.

Fig.3(b) shows the effect of assisting nodes on the content distribution process under the assumption that they have infinite storage as in (13). Clearly, the assisting nodes have the largest effect when there are few subscribers. When the systems is already large, assisting nodes have only marginal effect. As an example we see that when there are 10 subscribers for a given channel and 10 other nodes are willing to assist, the content spreading time is approximately 61% of what it is when there are no assistants. When we have 50 subscribers and 10 assistants the difference is not as large or the content distribution time is 86% of of what it is when there are no assistants. This suggests that for promoting fairness and giving new content channels a chance to spread the assistants should help in spreading less popular content. This has been suggested from simulation results in [22] but here we give an argument based on an analytical model.

In Fig.3(c) we compare type II infinite assistant storage cooperation in (13) with the type III limited upload case in (14). As a base comparison we also plot the type II cooperation with infinite storage (7). For both the type III models we assume that the number of assistants is $M = 10$. It is interesting that only when the system is really small (approximately 5 nodes or less) there is significant difference in the content distribution time for the infinite assistant storage versus the limited upload case. Otherwise the content distribution time for the limited assistant upload case is only slightly longer than for the infinite assistant memory case.

V. Conclusions

In this work we have addressed the issue of cooperation in a wireless opportunistic content distribution system. We have identified three basic types of cooperation where the degree of node collaboration and generosity differs. For each of the

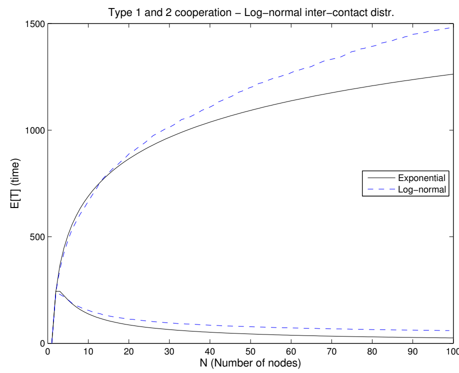


Fig. 4. Content distribution time in Type I and II cooperation for exponential- and Log-Normal distributed inter-contact times.

cooperation types identified we give a Markov chain model that captures the dynamics of content spreading. We have also addressed and captured in our models how limited resources at the mobile nodes can affect the content spreading. In particular we study the effects of limited content lifetime at the nodes due to shortage of storage and limited number of uploads from a node to reduce the power consumption required for data transmissions. The performance of the content distribution for the different cooperation types and node limitations is studied by computing the mean time to absorption for the Markov chains.

The main conclusions from our models and numerical results are the following:

- Performance of the content distribution is highly dependent on the cooperation degree of nodes. When there is high collaboration (type II and type III) the content spreading time is much shorter than when cooperation is limited or none (type I). We have seen that the mean content distribution time grows with the number of nodes N as $\mathcal{O}(ln(N))$ when nodes are non-cooperative. When nodes cooperate by sharing content, the mean content distribution time is bounded by a constant and is $\mathcal{O}(1)$.
- Limiting the number of times a node or assistant shares a channel gives only a slightly worse performance than when nodes have infinite storage. Moreover, by limiting upload each node has a bound on the maximum energy required for uploading the content it currently shares.
- The effect of assisting nodes is most significant for small systems (or equally for non-popular channels) while for popular content with many subscribers their effect is marginal. This suggests that for promoting fairness in distributing channels and giving new channels a chance to spread, assisting nodes should also solicit and help spreading less popular channels.

As part of future work we intend to study the effect of different solicitation and caching strategies for type III cooperation. This includes examining multiple channels and the effect of channel popularity. Finally we intend to relax the boundary assumptions of our models and study the content distribution in an open system.

References

[1] G. Karlsson, V. Lenders, and M. May, "Delay-tolerant broadcasting," *IEEE Transactions on Broadcasting*, vol. 53, no. 2, pp. 369 – 381, Mar. 2007.

[2] M. Papadopouli and H. Schulzrinne, "Effects of power conservation, wireless coverage and cooperation on data dissemination among mobile devices," in *Proc. ACM MobiHoc*, Long Beach, CA, USA, 2001, pp. 117–127.

[3] D. Daley and J. Gani, *Epidemic Modelling: An Introduction*. Cambridge, United Kingdom: Cambridge University Press, 1999.

[4] H. Andersson and T. Britton, *Stochastic Epidemic Models and Their Statistical Analysis*. New York, USA: Springer-Verlag, 1999.

[5] M. Motani, V. Srinivasan, and P. S. Nuggehalli, "Peoplenet: engineering a wireless virtual social network," in *Proc. ACM MobiCom*, Cologne, Germany, 2005, pp. 243–257.

[6] W. Yuen, R. Yates, and S. Mau, "Exploiting data diversity and multiuser diversity in noncooperative mobile infostation networks," in *Proc. IEEE Infocom*, San Francisco, USA, 2003.

[7] P. T. Eugster, R. Guerraoui, A.-M. Kermarrec, and L. Massoulie, "From epidemics to distributed computing," *IEEE Computer*, vol. 37, no. 5, pp. 60–67, 2004.

[8] S. Voulgaris and M. van Steen, "An epidemic protocol for managing routing tables in very large peer-to-peer networks," in *Proc. of IFIP/IEEE DSOM*. Berlin, Germany: Springer-Verlag, 2003.

[9] A. Demers, D. Greene, C. Hauser, W. Irish, J. Larson, S. Shenker, H. Sturgis, D. Swinehart, and D. Terry, "Epidemic algorithms for replicated database maintenance," in *Proc ACM PODC*, Vancouver, British Columbia, Canada, 1987, pp. 1–12.

[10] A. Khelil, C. Becker, J. Tian, and K. Rothermel, "An epidemic model for information diffusion in manets," in *Proc. ACM MSWiM*, Atlanta, Georgia, USA, 2002, pp. 54–60.

[11] A.-M. Kermarrec and A. J. Ganesh, "Efficient and adaptive epidemic-style protocols for reliable and scalable multicast," *IEEE Trans. Parallel Distrib. Syst.*, vol. 17, no. 7, pp. 593–605, 2006.

[12] K. P. Birman, M. Hayden, O. Ozkasap, Z. Xiao, M. Budiu, and Y. Minsky, "Bimodal multicast," *ACM Trans. Comput. Syst.*, vol. 17, no. 2, pp. 41–88, 1999.

[13] R. Groenevelt, P. Nain, and G. Koole, "The message delay in mobile ad hoc networks," *Perform. Eval.*, vol. 62, no. 1-4, pp. 210–228, 2005.

[14] T. Small and Z. J. Haas, "The shared wireless infostation model: a new ad hoc networking paradigm (or where there is a whale, there is a way)," in *Proc. ACM MobiHoc*, Annapolis, Maryland, USA, 2003, pp. 233–244.

[15] T. Small and Z. J. Haas, "Resource and performance tradeoffs in delay-tolerant wireless networks," in *Proc. WDTN'05 ACM SIGCOMM workshop on Delay-tolerant networking*, Philadelphia, Pennsylvania, USA, 2005, pp. 260–267.

[16] X. Zhang, G. Neglia, J. Kurose, and D. Towsley, "Performance modeling of epidemic routing," *Comput. Networks*, vol. 51, no. 10, pp. 2867–2891, 2007.

[17] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, no. 4, pp. 477–486, 2002.

[18] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on the design of opportunistic forwarding algorithms," in *Proc. IEEE Infocom*, Barcelona, Spain, 2006.

[19] H. Cai and D. Y. Eun, "Crossing over the bounded domain: from exponential to power-law inter-meeting time in manet," in *Proc. MobiCom*, Montreal, Quebec, Canada, 2007, pp. 159–170.

[20] T. Karagiannis, J.-Y. L. Boudec, and M. Vojnovic, "Power law and exponential decay of inter contact times between mobile devices," in *Proc. MobiCom*, Montreal, Quebec, Canada, 2007, pp. 183–194.

[21] V. Conan, J. Leguay, and T. Friedman, "Characterizing pairwise inter-contact patterns in delay tolerant networks," in *Proc ACM Autonomics*, Rome, Italy, Oct. 2007.

[22] V. Lenders, G. Karlsson, and M. May, "Wireless ad hoc podcasting," in *Proc. IEEE SECON*, June 2007.

[23] P. Brémaud, *Markov chains : Gibbs fields, Monte Carlo simulation, and queues*. New York, USA: Springer-Verlag, 1998.

[24] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad hoc wireless networks," in *Mobile Computing*, Imielinski and Korth, Eds. Kluwer Academic Publishers, 1996, vol. 353.