Dynamic Human Behaviour Based Epidemic Content Dissemination In Mobile Social Networks

Jie Hu, Lie-Liang Yang, Lajos Hanzo

School of Electronics and Computer Science, University of Southampton,

Southampton, UK, SO17 1BJ,

Tel: +44-23-8059 3125, Fax: +44-23-8059 4508

Email: {jh10g11, lly, lh}@ecs.soton.ac.uk, http://www-mobile.ecs.soton.ac.uk

Abstract

Social networks are penetrating our daily life, connecting people across the globe. As a combination of social science and mobile communication networks, mobile social networks (MSNs) are attracting an increasing attention across the research community. In this paper, based on the common interests of a specific community, the epidemic content dissemination across a MSN is studied as a powerful supplement to the conventional centralized-infrastructure (CI) based communication for the sake of conserving precious radio resources, enhancing coverage and reducing power-dissipation. The Factor of Altruism (FA) concept is introduced for quantifying the willingness of the MSN subscribers to share their content. We model the epidemic content dissemination by a pure-birth based Markov chain and evaluate the statistical properties of the content dissemination delay and

The financial support of the RC-UK's India-UK Advanced Technology Centre (IU-ATC), that of the EU's concerto project and of the China Scholarship Council (CSC) is gratefully acknowledged. L. Hanzo would also like to acknowledge the fiscal support of the European Research Council under its Advanced Fellow Award.

the delay of a specific MSN subscriber receiving the desired content. We also approximate the tail distribution function (TDF) of these two delays by the Gamma distribution. Simulation results are provided for supporting our analysis, which show the difference with respect to the conventional CI systems, demonstrating that the delay is substantially reduced upon increasing the number of MSN subscribers, especially when the MSN subscribers are altruistic.

Index Terms

Mobile social network, epidemic content dissemination, Markov chain, factor of altruistic

I. Introduction

Given the rapid development of social networks, an increasing number of people who share similar interests are connected to each other. They create a virtual social community, and they regularly communicate with each other based on their common interests. Social networks were deemed to contribute to the tremendous success of the London Olympic Games, which inherited the fond connotation of "Twitter Olympics" [1]. As the combination of mobile networks and social science, mobile social networks (MSNs) [2] have attracted substantial research attention. Developing techniques for exploiting the social relations among mobile users for improving the wireless communication and networking services deserves further research.

The operators are now facing two important wireless networking problems. The first one is networking in densely populated areas. The conventional solution is to install more picocell or femtocell base stations (BSs) [3] [4]. However, people sometimes get together only temporarily for large events, such as the 2012 Olympic Game in London, which would not justify investment into a permanent infrastructure. For example, during the Olympic Games, a large fraction of the audience in the Wembley Arena, who were watching the men's final in Badminton, were also interested in the result of Andy Murray's tennis-final against Roger Federer. Based on the audience's common interests, we may find an alternative way of satisfying their appetite for data traffic without a permanent infrastructure. The BSs select some users and transmit Murray's winning stroke to them. Then this content may be further disseminated spontaneously to the others in the Wembley Arena. In this manner, we may provide a service at a low energy consumption and enhance the coverage by supporting all the users in densely populated areas.

The second problem is faced by the network in the areas having a low infrastructure density, as exemplified by emergency communication. After a natural disaster, such as the Tsunami in Japan, most of the infrastructure may be destroyed, hence it is hard for victims to send SOS messages in conventional ways. The wireless terminals may create a self-organized ad hoc network, and SOS messages may be sent to specific nodes, which are capable of communicating with the outside world through a surviving BS. The fundamental goal of emergency communication is to find the best route for sending the required information [5]. However, in reality every user's phone is unaware of which particular node is capable of communicating with the outside world. In our solution, after generating an SOS message, the victims flood the network with it and when the specific node having a link to the outside world via a BS receives it, it forwards the information to the outside 'universe'.

Epidemic spreading in scale-free networks was studied in [6], where a birth-death Markov chain was invoked whose states represent the number of nodes infected in the network. A two-dimensional continuous time Markov chain (CTMC) model relying on a so-called absorption state is proposed in [7] for evaluating the performance of a heterogeneous delay-tolerant network, where the state-transition rate is defined by the inter-encounter rate of mobile users. An important application of content dissemination is constituted by facilitating the spectrum sensing decisions in cognitive radio (CR) networks [8], where the CR users are grouped for propagating their spectrum preference[1] to each other, so that the best decision may be made.

[1]Spectrum preference of a cognitive user indicates which is the best spectrum band for this user to achieve the optimal throughput when it is released by primary users

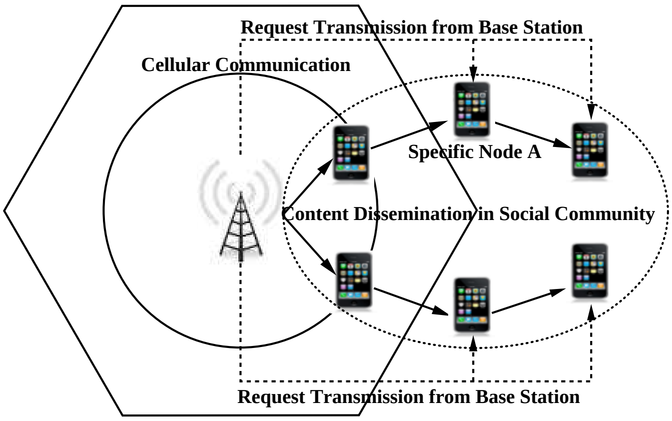


Fig. 1: System Model

The authors of [9] investigate how the content providers and network operators interact with each other for disseminating the content across a MSN.

Against this background, our major contributions are listed as follows:

(1) A hybrid MSN architecture is studied, which includes conventional centralized-infrastructure based communication as the first stage, followed by content dissemination across the MSN during the second stage.

(2) In order to investigate the associated delay, content dissemination across the MSN is modelled by a pure-birth Markov chain having an absorption state.

(3) The Factor of Altruism (FA) concept is introduced for reflecting the willingness of users sharing the content with others.

(4) The statistical characteristics of the content dissemination delay and the delay of a specific user receiving the content are also derived.

Our paper is organized as follows. Our system overview is provided in Section II, including the description of the Network, MAC and PHY layer. The content dissemination delay and the delay of a specific user receiving the content are analysed in Section III, while our simulation and analytical results are provided in Section IV. Finally, we conclude in Section V.

II. System Overview

A. Network Layer

We assume that there are N subscribers in the MSN, who are divided into active subscribers (ASs) as well as passive subscribers (PSs) and are able to communicate with each other. The ASs either receive the content from the CIs or actively generate an SOS messages themselves. By contrast, the PSs receive the content either from other subscribers or from the CIs. As shown in Fig.1, our system may support two different sessions. In the scenario of high downlink traffic load, U ($0 \leq U < N$) ASs receive their desired content from the CIs, such as a BS or WiFi hot-spot, while in the scenario of emergency communication, they themselves generate the SOS information. Secondly, the ASs disseminate the content to all the others in the MSN based on their common interests. After receiving the content, regardless of where it originates from, the subscribers will further disseminate the content to the others who do not have it, until all the MSN subscribers receive this content.

Naturally, each MSN subscriber acts both as a receiver and as a transmitter. However, when acting as a transmitter, the subscribers' terminals assist in the content dissemination by dissipating their own energy, which will substantially shorten their battery recharge-period, potentially without any payback. Consequently, some of the subscribers may be reluctant to help the content dissemination in order to save their own battery charge. We define the probability of a user becoming a transmitter as q ($0 \leq q \leq 1$), which is termed as the FA. Clearly, a lower q value represents a more selfish inclination, while a higher q represents more altruistic behaviour. Naturally, a subscriber's preference as to whether to become a transmitter may vary from time to time. Without loss of generality, we assume that the subscribers, who already have the content, will make their own decision upon receiving the content during the content dissemination.

Depending on the specific value of the FA, at any moment there might not be any subscribers willing to broadcast the content. If so, the subscribers, who are keen on downloading the content, have to request it from a CI, although dissemination from the CI is of lower efficiency owing to its higher path loss (PL), than those between the MSN subscribers.

B. MAC Layer

In the MAC Layer we assume that the content is only transmitted successfully in a time slot (TS), when the near-instantaneous received signal to noise ratio (SNR) of the link from a source to a target is higher than a predefined threshold. We define this successful reception probability of the links spanning from the CI to the MSN subscribers as μ_b and its counterpart amongst the MSN subscribers as μ_s. By jointly considering the PHY Layer model, we calculate these two probabilities in Section II-B.

Automatic repeat request (ARQ) [10] relying on an unlimited number of re-transmissions is adopted for ensuring that no contents are lost. Furthermore, we assume that the acknowledgement sent from a target is always successfully received by a source, while the associated delay is ignored. As a result, we can formulate the following theorem.

Theorem 1: Given the successful reception probability μ of a single content packet during a specific TS in the ARQ-aided system, the continuous-valued time required for the successful transmission of a packet obeys the exponential distribution with the mean of $1/\mu$.

Proof: The proof of *Theorem 1* is provided in Appendix A. Here μ can also be referred to as the average service rate of a link. ■

C. PHY Layer

In the PHY layer, radio propagation between any pair of nodes is assumed to experience uncorrelated stationary Rayleigh flat-fading channels $|h_i(t)|$ associated with $E[|h_i(t)|^2] = 1$, while $|h_i(t)|^2$ obeys an exponential distribution with a unity mean, whose tail distribution function (TDF) is $P[|h_i(t)|^2 > x] = \exp(-x)$. The average PL of a link is denoted by Ω_i. We assume that the transmit power is P_i, and the noise power is $N_0 W_i$, where W_i is the available

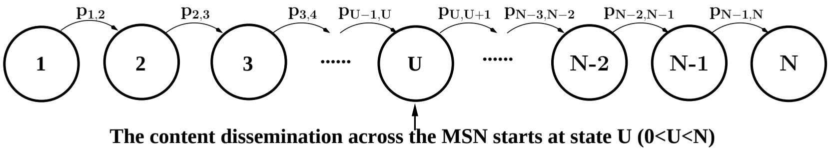


Fig. 2: A pure birth process having a single absorption state

bandwidth, while the subscript i represents 'b' for the links connecting the CIs and the MSN subscribers and 's' for the links amongst the MSN subscribers. We define the successful content reception SNR threshold at a subscriber as β. As a result, we can formulate the relevant expressions for μ_b and μ_s as follows:

$$\mu_b = P(P_b|h_i(t)|^2/\Omega_b > \beta) = \exp(-\beta \cdot \Omega_b N_0 W_b / P_b), \tag{1}$$

$$\mu_s = P(P_s|h_i(t)|^2/\Omega_s > \beta) = \exp(-\beta \cdot \Omega_s N_0 W_s / P_s). \tag{2}$$

III. Content Dissemination Delay

In this section, we will study the statistical properties of the epidemic content dissemination delay and the delay of a specific user receiving the content with the aid of a pure birth Markov chain.

A. Pure Birth Markov Chain

In the MSN, epidemic content dissemination is invoked for ensuring that every subscriber receives the content of common interest. We define a state-machine having states given by the number of subscribers who have already received the content. Clearly, the dissemination is completed when all the N subscribers of the MSN received the content. Hence, we can model the content dissemination as a pure-birth based Markov chain with an absorption state N, as shown in Fig.2.

Let us first consider the adjacent-state transition rate from state k to state $(k + 1)$. Again, depending on the FA, not necessarily all of the k subscribers are willing to act as transmitters. We assume that there are n_k $(0 \leq n_k \leq k)$ subscribers willing to further disseminate the content.

Since $(N-k)$ subscribers in the MSN are waiting for the content, there are $n_k \cdot (N-k)$ possible links in this state, each of which has an average service rate of μ_s. Hence, given a sufficiently short time unit Δt, in which a single transmission takes place, each of the links successfully sends the content with the probability of $\mu_s \Delta t$. Then provided that n_k subscribers are willing to disseminate the content, the adjacent-state transition probability in state k may be expressed as

$$\begin{aligned} p_{k,k+1|n_k} &= \binom{n_k(N-k)}{1} \mu_s \Delta t (1-\mu_s \Delta t)^{n_k(N-k)-1} \\ &= n_k(N-k)\mu_s \Delta t \left[1 + \sum_{i=1}^{n_k(N-k)-1} \binom{n_k(N-k)-1}{i} (-\mu_s \Delta t)^i \right] \\ &\approx n_k(N-k)\mu_s \Delta t, \end{aligned} \tag{3}$$

where the approximation is due to the fact that the time unit Δt was assumed to be sufficiently short for encountering a single successful reception.

Secondly, we demonstrate that the probability of a two-step transition $p_{k,k+2|n_k}$ is negligibly low. Each receiver is connected to the transmitters by n_k links. Provided that one of these links successfully transmits, this receiver acquires the content. In this case, only two out of $(N-k)$ receivers acquire the content and the resultant conditional probability $p_{k,k+2}$ associated with n_k users willing to disseminate the content may be formulated as

$$\begin{aligned} p_{k,k+2|n_k} &= \binom{N-k}{2} \left[1 - (1-\mu \Delta t)^{n_k} \right]^2 (1-\mu \Delta t)^{n_k(N-k-2)} \\ &= \binom{N-k}{2} \left[\sum_{i=0}^{n_k} (-\mu \Delta t)^i \right]^2 (1-\mu \Delta t)^{n_k(N-k-2)}. \end{aligned}$$

Therefore, if Δt is sufficiently short, which results in $p_{k,k+2|n_k} \ll p_{k,k+1|n_k}$, we have a reasonable approximation of $p_{k,k+2|n_k} \approx 0$.

B. Content Dissemination Delay Across the MSN

In state k, we express the number of subscribers willing to share the content as n_k ($0 \leq n_k \leq k$). Since each subscriber has a probability q of sharing his/her version of the content, n_k follows a Binomial distribution having a pair of parameters k and q, whose probability mass function (PMF) is given by [11]

$$p(n_k) = \binom{k}{n_k} q^{n_k}(1-q)^{k-n_k}. \tag{4}$$

In case of $n_k \geq 1$, which indicates that at least one subscriber would like to assist in the content dissemination process, we have $n_k(N-k)$ possible links for dissemination. According to *Theorem 1* and Equation (3), we may claim that the delay from state k to state $(k+1)$ obeys an exponential distribution having a parameter of $n_k(N-k)\mu_s = n_k\mu_k$, where we define $\mu_k = (N-k)\mu_s$. Given that n_k subscribers are willing to share the content, the conditional probability density function (PDF), the mean and the second moment of the random delay T_k from state k to state $(k+1)$ are

$$f_{T_k|n_k}(t_k) = n_k\mu_k \cdot \exp\Big(-n_k\mu_k t_k\Big),\ \ t_k \leq 0 \tag{5}$$

$$E\Big[T_k\Big|n_k\Big] = \int_0^{\infty} t_k f_{T_k|n_k}(t_k)dt_k = \frac{1}{n_k\mu_k}, \tag{6}$$

$$E\Big[T_k^2\Big|n_k\Big] = \int_0^{\infty} t_k^2 f_{T_k|n_k}(t_k)dt_k = \frac{2}{(n_k\mu_k)^2}. \tag{7}$$

In case of $n_k = 0$, which indicates that no subscribers would like to assist in the content dissemination process, some of the $(N-k)$ MSN subscribers, who are still waiting for the content, have to request its transmission from the CI, although these transmissions may experience more substantial degradation. As a result, $(N-k)$ links are established from the CI in total. According to *Theorem 1*, we may claim that the delay from state k to state $(k+1)$ obeys an exponential distribution with a parameter of $(N-k)\mu_b$. Hence, the conditional PDF, the mean and the second order moment of the random delay T_k are

$$f_{T_k|n_k=0}(t_k) = (N-k)\mu_b \cdot \exp\Big(-(N-k)\mu_b t_k\Big),\ t_k \le 0 \tag{8}$$

$$E\Big[T_k\Big|n_k = 0\Big] = \int_0^{\infty} t_k f_{T_k|n_k=0}(t_k)dt_k = \frac{1}{(N-k)\mu_b}, \tag{9}$$

$$E\Big[T_k^2\Big|n_k = 0\Big] = \int_0^{\infty} t_k^2 f_{T_k|n_k=0}(t_k)dt_k = \frac{2}{((N-k)\mu_b)^2}. \tag{10}$$

According to Bayes' theorem, the PDF of the random delay T_k from state k to state $(k+1)$ may be expressed as

$$\begin{aligned} f_{T_k}(t_k) =\ & f_{T_k|n_k=0}(t_k) \cdot p(n_k = 0) + \sum_{n_k=1}^{k} f_{T_k|n_k}(t_k) \cdot p(n_k) \\ =\ & (N-k)\mu_b \exp(-(N-k)\mu_b t_k)(1-q)^k \\ &+ \sum_{n_k=1}^{k} n_k\mu_k \exp\Big(-n_k\mu_k t_k\Big) \cdot \binom{k}{n_k} q^{n_k}(1-q)^{k-n_k}. \end{aligned} \tag{11}$$

Given the PDF of (11), the mean and the second moment of the random delay T_k are

$$\begin{aligned} E\Big[T_k\Big] &= E\Big[T_k\Big|n_k = 0\Big]p(n_k = 0) + \sum_{n_k=1}^{k} E\Big[T_k\Big|n_k\Big]p(n_k) \\ &= \frac{(1-q)^k}{(N-k)\mu_b} + \sum_{n_k=1}^{k} \binom{k}{n_k} \frac{q^{n_k}(1-q)^{k-n_k}}{n_k\mu_k}, \end{aligned} \tag{12}$$

$$\begin{aligned} E\Big[T_k^2\Big] &= E\Big[T_k^2\Big|n_k = 0\Big]p(n_k = 0) + \sum_{n_k=1}^{k} E\Big[T_k^2\Big|n_k\Big]p(n_k) \\ &= \frac{2(1-q)^k}{((N-k)\mu_b)^2} + \sum_{n_k=1}^{k} \binom{k}{n_k} \frac{2q^{n_k}(1-q)^{k-n_k}}{(n_k\mu_k)^2}. \end{aligned} \tag{13}$$

Hence, we can also derive the variance of T_k using the formula of $Var[T_k] = E[T_k^2] - \{E[T_k]\}^2$.

As shown in Fig.2, the epidemic content dissemination is initiated in state U and completed in state N. Therefore, the total random delay T is given by $T = \sum_{k=U}^{N-1} T_k$. Since the delay from each state to its adjacent one is independent of each other, the mean and variance of T

are derived as:

$$E[T] = \sum_{k=U}^{N-1} \frac{(1-q)^k}{(N-k)\mu_b} + \sum_{k=U}^{N-1} \sum_{n_k=1}^{k} \binom{k}{n_k} \frac{q_k^n (1-q)^{k-n_k}}{n_k \mu_k}, \tag{14}$$

$$Var[T] = \sum_{k=U}^{N-1} Var[T_k]. \tag{15}$$

Unfortunately, it is impossible to derive the exact closed-form PDF of the random delay T. But given the mean and variance, we may approximate it by the Gamma distribution, which is more accurate than its Gaussian counterpart [12], when non-negative component random variables are concerned. According to the Gamma distribution [12], we have the shape parameter $m = \{E[T]\}^2/Var[T]$ and the scale parameter $\Theta = Var[T]/E[T]$. Then, given a delay threshold D_{th}, we may express the approximate probability of the content dissemination delay exceeding D_{th} as

$$P(T > D_{th}) = \frac{\Gamma\left(m, \frac{D_{th}}{\Theta}\right)}{\Gamma(m)} = \frac{\Gamma\left(\frac{\{E[T]\}^2}{Var[T]}, \frac{D_{th}E[T]}{Var[T]}\right)}{\Gamma\left(\frac{\{E[T]\}^2}{Var[T]}\right)}. \tag{16}$$

This approximation will be characterized by our simulation results in Section IV.

C. Delay of a Specific Subscriber Receiving the Content

Apart from the delay of content dissemination across the entire MSN, the delay of a specific subscriber receiving the content is also of high importance, for example in the scenario of emergency communication and for content dissemination across multiple MSNs. A MSN is defined based on the common interests of its subscribers. A subscriber may however belong to several MSNs. In the scenario of multiple MSNs, the subscribers shared by these MSNs are of particular significance for the content dissemination amongst MSNs. The delay of these subscribers receiving the content substantially affects the total delay of content dissemination across these MSNs. Hence, we study the delay of a specific subscriber $\mathcal{A}$ receiving the content under the assumption that $\mathcal{A}$ does not belong to ASs (defined in Section II-A), since if it does, its delay is zero.

Under the above assumptions, subscriber $\mathcal{A}$ may receive the content during the state transition from $(U+1)$ to N. When considering the state transition from k to $(k+1)$, $(N-k)$ subscribers, who at this stage have not as yet received the content, have an equal probability of $1/(N-k)$ of receiving the content and similarly an equal probability of $(N-k-1)/(N-k)$ of not receiving it. The probability of $\mathcal{A}$ receiving the content in state $(k+1)$, which naturally implies that it has not received it during any of the previous states, may be expressed as

$$\begin{aligned} p_{k+1} &= \frac{N-U-1}{N-U}\frac{N-(U+1)-1}{N-(U+1)}\cdots\frac{N-(k-1)-1}{N-(k-1)}\frac{1}{N-k} \\ &= \frac{1}{N-U}. \end{aligned} \tag{17}$$

Given the event that $\mathcal{A}$ receives the content in state $(k+1)$, the random delay is denoted by $T_{\mathcal{A}} = \sum_{j=U}^{k} T_k$ and the conditional PDF of $T_{\mathcal{A}}$ is denoted by $f_{T_{\mathcal{A}}|k+1}(t_{\mathcal{A}}) = f_{T_U+\cdots+T_k}(t_{\mathcal{A}})$. According to the Bayes' theorem, we arrive at the PDF of $T_{\mathcal{A}}$ in the form of:

$$f_{T_{\mathcal{A}}}(t_{\mathcal{A}}) = \sum_{k=U}^{N-1} f_{T_{\mathcal{A}}|k+1}(t_{\mathcal{A}})\cdot p_{k+1} = \sum_{k=U}^{N-1}\frac{f_{T_U+\cdots+T_k}(t_{\mathcal{A}})}{N-U}. \tag{18}$$

With the aid of (18), we arrive at the average value of the random delay $T_{\mathcal{A}}$ as

$$\begin{aligned} E[T_{\mathcal{A}}] &= \int_0^{\infty} t_{\mathcal{A}}\sum_{k=U}^{N-1}\frac{f_{T_U+\cdots+T_k}(t_{\mathcal{A}})}{N-U}dt_{\mathcal{A}} \\ &= \sum_{k=U}^{N-1}\frac{\int_0^{\infty} t_{\mathcal{A}} f_{T_U+\cdots+T_k}(t_{\mathcal{A}})dt_{\mathcal{A}}}{N-U} \\ &= \sum_{k=U}^{N-1}\frac{1}{N-U}E\left[T_U+T_{U+1}+\cdots+T_k\right] \\ &= \frac{1}{N-U}E\left[T_U+(T_U+T_{U+1})+(T_U+T_{U+1}+T_{U+2})+\cdots+(T_U+T_{U+1}+\cdots+T_{N-1})\right] \\ &= \sum_{k=U}^{N-1}\frac{N-k}{N-U}E[T_k], \end{aligned} \tag{19}$$

where $E[T_k]$ is given by (12). In order to complete the statistical characterization of the delay of $\mathcal{A}$ receiving the content, its delay variance is also needed. Before deriving this variance, we have to determine its second moment, which may be derived as

$$\begin{aligned}
E[T_{\mathcal{A}}^2] &= \int_0^{\infty} t_{\mathcal{A}}^2 \sum_{k=U}^{N-1} \frac{f_{T_U+\cdots+T_k}(t_{\mathcal{A}})}{N-U} dt_{\mathcal{A}} \\
&= \sum_{k=U}^{N-1} \frac{\int_0^{\infty} t_{\mathcal{A}}^2 f_{T_U+\cdots+T_k}(t_{\mathcal{A}}) dt_{\mathcal{A}}}{N-U} \\
&= \sum_{k=U}^{N-1} \frac{1}{N-U} E\Big[(T_U + T_{U+1} + \cdots + T_k)^2\Big] \\
&= \frac{1}{N-U} E\Big[T_U^2 + (T_U + T_{U+1})^2 + (T_U + T_{U+1} + T_{U+2})^2 + \cdots + (T_U + T_{U+1} + \cdots + T_{N-1})^2\Big] \\
&= \frac{1}{N-U} \sum_{k=U}^{N-1} \sum_{i,j=U}^{k} E\Big[T_i T_j\Big] \\
&= \frac{1}{N-U} \Bigg[\sum_{k=U}^{N-1} (N-k) E\Big[T_k^2\Big] + \sum_{k=U+1}^{N-1} \boldsymbol{\xi_k} \Big[\mathbf{H}_{k-U+1} - \mathbf{I}_{k-U+1}\Big] \boldsymbol{\xi_k}^T \Bigg],
\end{aligned} \tag{20}$$

where $\boldsymbol{\xi_k} = (E[T_U], E[T_{U+1}], \cdots, E[T_k])$, $\mathbf{H}_{k-U+1}$ is a $(k-U+1)\times(k-U+1)$-element matrix having elements all of which are unity, and $\mathbf{I}_{k-U+1}$ is a $(k-U+1)\times(k-U+1)$-element identity matrix. Consequently, the variance of $T_{\mathcal{A}}$ can be expressed as $Var(T_{\mathcal{A}}) = E[T_{\mathcal{A}}^2] - \{E[T_{\mathcal{A}}]\}^2$. In a similar manner as shown in (16), we may obtain the approximate probability of $T_{\mathcal{A}}$ exceeding the threshold D_{th}.

Here a special case is considered. When the FA is $q = 0$, no MSN subscribers are willing to share the content with each other. The content has to be broadcast by the CI all the time to all the unserved MSN subscribers. In state k, there are $(N-k)$ unserved users, which implies that $(N-k)$ links are established for broadcasting by the CI. According to the analysis in Section III-A, the transition rate from state k to state $(k+1)$ is $(N-k)\mu_b$ and the average delay is $E[T_k|q=0] = 1/[(N-k)\mu_b]$. Substituting this expression into (19), we have

$$E[T_{\mathcal{A}}\big|q=0] = \sum_{k=U}^{N-1} \frac{N-k}{N-U} \frac{1}{(N-k)\mu_b} = \frac{1}{\mu_b}. \tag{21}$$

It is however a bit of a challenge to derive the second moment of the delay, when $\mathcal{A}$ receives the content in this special case. According to (20), we have

$$E[T_{\mathcal{A}}^2|q=0] = \frac{1}{N-U}\sum_{k=U}^{N-1}(N-k)E[T_k^2|q=0] + \frac{1}{N-U}\sum_{k=U+1}^{N-1}\sum_{i,j=U,i\neq j}^{k} E[T_iT_j|q=0]$$

$$= \frac{1}{N-U}\Bigg\{\sum_{k=U}^{N-1}(N-k)E[T_k^2|q=0] + E[(2T_UT_{U+1}(N-U-1) + 2T_UT_{U+2}(N-U-2) + \cdots$$

$$+ 2T_UT_{N-1}) + (2T_{U+1}T_{U+2}(N-U-2) + \cdots + 2T_{U+1}T_{N-1}) + \cdots + 2T_{N-2}T_{N-1}|q=0]\Bigg\}. \quad (22)$$

Since we have $E[T_k^2|q=0] = 2/[(N-k)^2\mu_b^2]$, (22) may be reformulated as

$$E[T_{\mathcal{A}}^2|q=0] = \frac{1}{N-U}\left\{\sum_{k=U}^{N-1}\frac{2}{(N-k)\mu_b^2} + \left[\frac{2(N-U-1)}{(N-U)\mu_b^2} + \frac{2(N-U-2)}{(N-U-1)\mu_b^2} + \cdots + \frac{2}{2\mu_b^2}\right]\right\}$$

$$= \frac{1}{N-U}\left\{\sum_{k=U}^{N-1}\frac{2}{(N-k)\mu_b^2} + \sum_{k=U+1}^{N-1}\frac{2}{\mu_b^2}\frac{N-k}{N-k+1}\right\}$$

$$= \frac{2}{(N-U)\mu_b^2}\left\{\frac{1}{N-U} + \sum_{k=U+1}^{N-1}\left[\frac{1}{N-k} + \frac{N-k}{N-k+1}\right]\right\}$$

$$= \frac{2}{(N-U)\mu_b^2}\left\{\frac{1}{N-U} + (N-U-1) + \sum_{k=U+1}^{N-1}\left[\frac{1}{N-k} - \frac{1}{N-k+1}\right]\right\} = \frac{2}{\mu_b^2}. \quad (23)$$

Consequently, the variance of $T_{\mathcal{A}}$ is $Var[T_{\mathcal{A}}|q=0] = E[T_{\mathcal{A}^2}|q=0] - \{E[T_{\mathcal{A}}|q=0]\}^2 = 1/\mu_b^2$. According to the above analysis, we can see that when the FA is $q=0$, the delay of $\mathcal{A}$ receiving the content is only determined by the quality of the link spanning from the CI to the subscriber $\mathcal{A}$, but it is unrelated to any of the MSN's properties.

IV. Numerical Results

In this section, we investigate the average delay and the tail distribution of both the content dissemination delay and the delay of a specific user receiving the content. Both analytical and simulation results are provided for quantifying the impact of the MSN's properties, such as the number of subscribers in the MSN and the FA q on these performances. The parameters are set as follows.

In the PHY layer, we firstly characterize the links spanning from the CI to the MSN subscribers. We invoke a macro BS as part of the CI whose transmit power is 17 dBm. The thermal noise floor for 1 Hz bandwidth at room temperature (20°C) is −174 dBm. In line with the LTE-Advanced standard, we assume that the transmissions from the BS to the MSN subscribers take place in a bandwidth of 100 MHz. As a result, the noise power in this bandwidth is −94 dBm. When the average distance from the BS to MSN subscribers is 1.5 km and the carrier frequency is 1.8 GHz[2], based on the 20 dB/decade free-space PL model [16], the PL is $\Omega_b = 101$ dB.

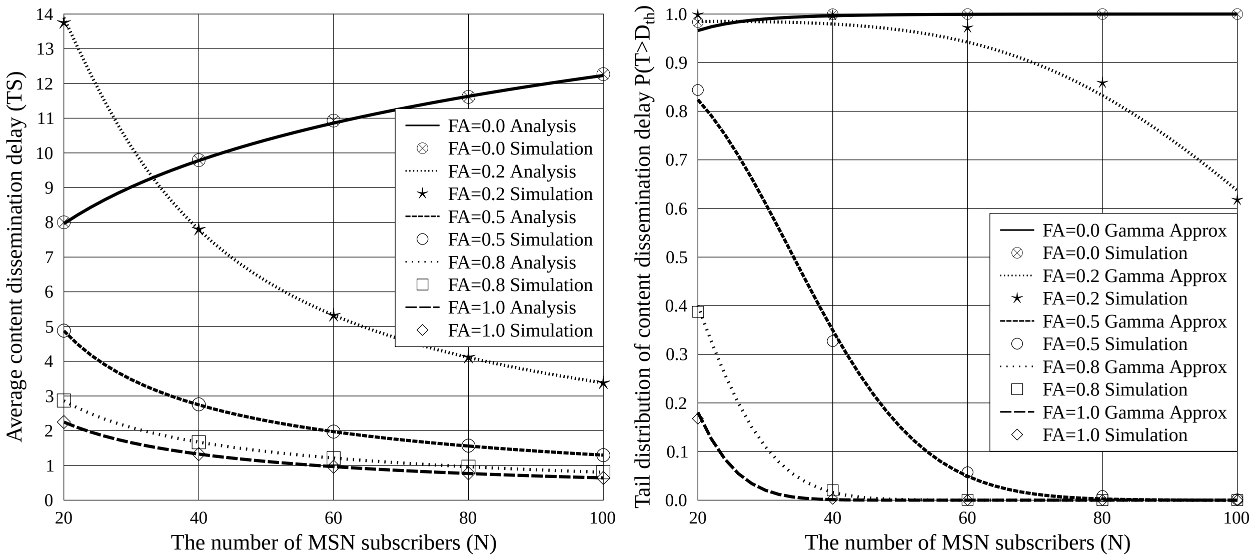
Let us now characterize the links amongst the MSN users. The short-range communication used for the content dissemination across the MSN is assumed to be operated in the vicinity of 3 GHz, while the bandwidth available for a MSN subscriber's broadcast is 10 MHz. As a result, the noise power in this bandwidth is −104 dBm, while the transmit power is assumed to be −30 dBm. If the average distance between the MSN subscribers is 20 m, based on the 20 dB/decade free space PL model [16], the average PL is $\Omega_s = 68$ dB. We assume that the SNR threshold of MSN subscribers required for successful content-reception is $\beta = 10$ dB.

In the Network layer, we assume that half of the MSN subscribers receive the content from the BS or generate it themselves. We investigate the attainable performance mentioned at the beginning of this section upon increasing the number of MSN subscribers for different values of the FA.

In order to arrive at a more reliable statistical characterization of the delay, we conducted simulations repeated 10 000 times and set the time-interval of the content dissemination system to 0.001 TS.

In Fig.3(a), we plot the number of MSN subscribers versus the average content dissemination delay for different FA values. When the FA is non-zero, the average content dissemination

[2]Since LTE-Advanced has not been put into business, no specific frequency band is allocated to it. However, as an technology evolved from LTE, it is reasonable to assume that LTE-Advanced is operated in the same frequency band as defined for LTE [13]. Although in [14] the authors assume 2 GHz carrier frequency for LTE-Advanced, we assume 1.8 GHz carrier frequency in line with the LTE networks operated by British company EE [15]

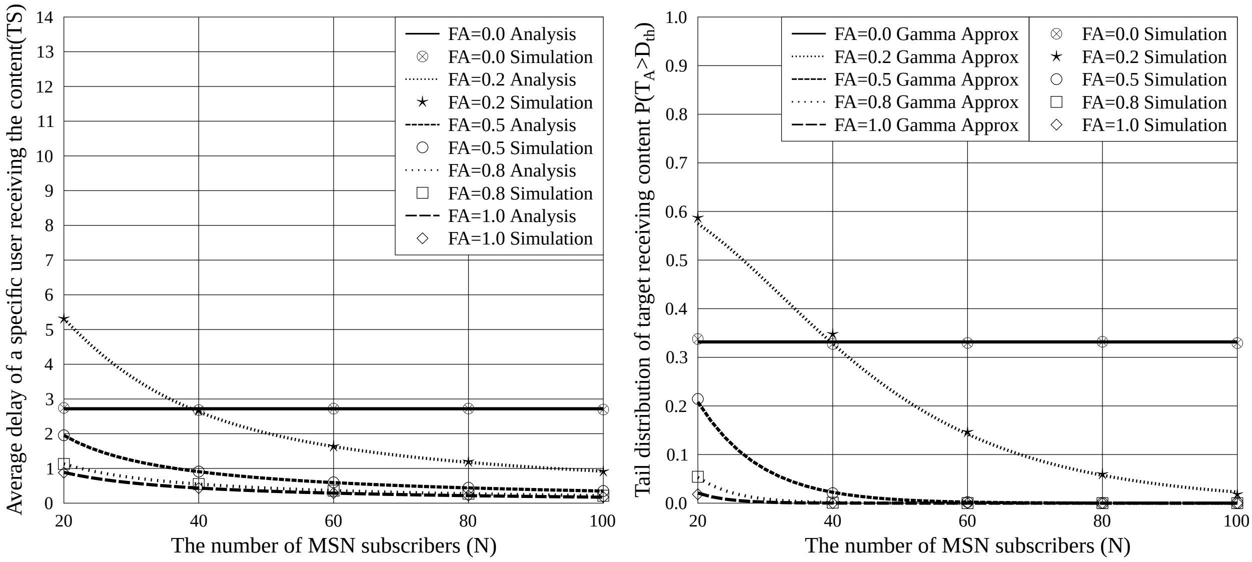


(a) Average delay (b) Tail behaviour

Fig. 3: The delay of the content dissemination: (a) Average delay (b) Tail distribution

delay reduces steadily, as the number of MSN subscribers increases. Having more subscribers in the MSN implies having more potential transmitters for content dissemination, which results in a substantial reduction of the content dissemination delay. For example, for FA = 0.2, the delay is reduced from 13.8 TS to 3.4 TS when the number of subscribers is increased from $N = 20$ to 100. Furthermore, a higher FA produces a lower delay, since more subscribers are willing to share the content, which consequently accelerates the content dissemination across the MSN. Conversely, for FA = 0 the MSN subscribers have to turn to the BS for the content. Since only the BS acts as a transmitter, if the number of receivers in the MSN increases, the delay of delivering the content to all the subscribers from the BS will also be increased. Observe in Fig.3(a) that our analysis perfectly matches the simulation results.

In Fig.3(b), we investigate the probability of the content dissemination delay exceeding a threshold of $D_{th} = 3$ TS. We plot the curves both for the TDF of the Gamma approximation used in (16) and for our simulation results. As shown in Fig.3(b), the Gamma approximation matches the simulation results better for FA > 0.2. For FA $\neq 0$, the probability of the delay

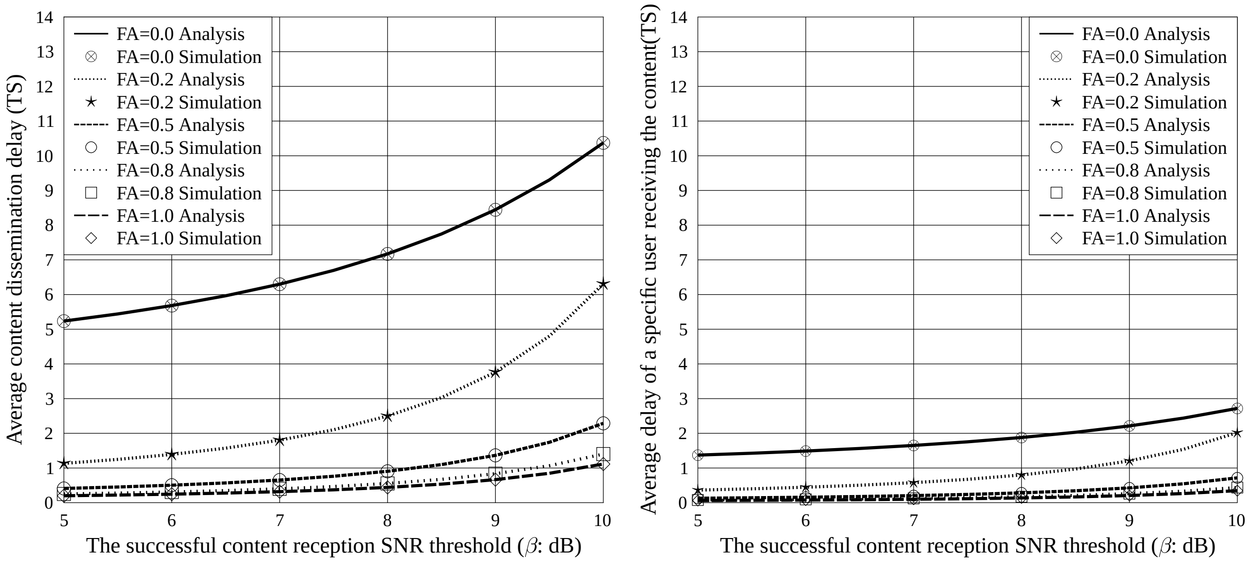


(a) Average delay (b) Tail behaviour

Fig. 4: The Delay of $\mathcal{A}$ receiving the content: (a) Average delay (b) Tail behaviour

exceeding D_{th} = 3 TS falls rapidly as the number of subscribers increases. In the special case of FA = 0, the probability of the delay exceeding D_{th} = 3 TS increases, upon increasing the number of MSN subscribers when the BS is the one and only source of the content.

In Fig.4(a), we investigate the average delay of a specific subscriber receiving the content. As expected, this delay is lower than its counterpart, when all the subscribers in the MSN receive the content. The average delay reduces, as the number of MSN subscribers increases. For example, for FA = 0.2, the delay falls from 5.2 TS to nearly 1.0 TS, when the number of subscribers is increased from N = 20 to 100. Nevertheless, in the special case of FA = 0 we observe a different trend. Although the number of subscribers increases, the average delay, which only relies on the link between this subscriber and the BS, remains near-constant at 3.8 TS. This is in line with our analysis provided in Section III-C. According to our parameter setting provided at the beginning of this section, the quality of the links connecting the BS and the subscribers is statistically speaking better than the links connecting one subscriber to another. Consequently, for less than N = 40 subscribers, the average delay of FA = 0.2 is higher than that of FA = 0.0. However, when the number of MSN subscribers continues



(a) Average delay of content dissemination across the MSN

(b) Average delay of a specific user receiving the content

Fig. 5: Average delay affected by the successful content reception SNR threshold (β)

to increase, the average delay of FA = 0.2 becomes lower than that of FA = 0.0. Another observation is that our analysis perfectly matches the simulation results of Fig.4(a).

In Fig.4(b), we investigate the probability of the delay exceeding a threshold of D_{th} = 3 TS, when a specific user $\mathcal{A}$ receives the content. When increasing FA = 0.2 to FA = 1.0, the probability of the delay being above D_{th} = 3 TS tends to zero, as the number of MSN subscribers increases. A higher FA usually leads to a reduced probability of exceeding the delay threshold of D_{th}, except for FA = 0 or for the scenario, when the link between the BS and subscribers is better than that between one subscriber and another. In the case of FA = 0, the probability of a delay in excess of D_{th} = 3 TS remains at 0.33, which indicates that the delay is essentially determined by the link between the BS and subscriber $\mathcal{A}$, as also indicated by Equations (21) and (23).

Let us now focus on the impact of the successful content reception SNR threshold (β) on the delay of content dissemination across the MSN (as shown in Fig.5(a)), as well as on the delay of a specific user receiving the content (as shown in Fig.5(b)), where we vary the value of β from 5 dB to 10 dB for different values of the FA, while fixing the number of MSN subscribers at $N = 50$.

Observe in Fig.5(a) that the average delay of content dissemination across the entire MSN increases, as the successful content reception SNR threshold β increases. For example, for FA = 0.2, the average delay is increased from about 1 TS to 6.3 TS upon increasing the value of β from 5 dB to 10 dB. Furthermore, recall from Fig.3 and Fig.4 that an increased FA implies that more subscribers are willing to share the content which reduces the content dissemination delay. In Fig.5(b) we observe a similar trend to Fig.5(a), which is recorded in Fig.5(b) for the average delay of a specific user receiving the content as a function of the successful content reception SNR threshold β.

V. Conclusions and Future Work

In this paper, a hybrid MSN architecture was studied, which relies on both CI based communication and on spontaneous content dissemination across the MSN based on the users' common interests. In the first stage, the ASs of the MSN receive the content of common interest directly from the CI in the scenario of a high traffic load, or they generate the content themselves in the scenario of emergency communication. In the second stage, regardless of how they acquired the content initially, they will disseminate it amongst all the MSN subscribers. Having received the content, a subscriber further disseminates it until all of them received it. During the second stage, we introduce the FA for reflecting the probability of a MSN subscriber willing to disseminate the content. In order to smoothly carry out the dissemination, the CI is also used for disseminating the content, when no subscribers are willing to share it.

The statistical properties of both the content dissemination delay across the entire MSN and the delay of a specific subscriber receiving the content were derived. We approximated the TDF of both delays by the Gamma distribution.

Our simulation results accurately matched our theoretical analysis. We concluded that both the delays mentioned in the previous paragraph are substantially improved, when the number of MSN subscribers increased compared to conventional CI-based communication. Furthermore, both the delays increase as the successful content reception SNR threshold increases. Given a higher FA, which indicates that the subscribers are more willing to share the content, the delay was reduced. Furthermore, our Gamma-distribution-based approximation used for the tail distribution of the content dissemination delay also matches the simulation results.

Our work mainly focused on the performance analysis of a single MSN. In the future, we are going to study the interaction between different MSNs and conceive MSNs-based resource allocation for the sake of improving the overall system performance.

Appendix A

The Proof of Theorem 1

We assume a continuous exponentially distributed random variable T in the unit of TS having a mean of $1/\mu$ TS, whose CDF is $P(T \le t) = 1 - \exp(-\mu t)$. Then T can be discretized to generate a new random variable $N = \lceil T/\Delta t \rceil$, where Δt is an interval, which is significantly shorter than one TS. Consequently, the probability mass function (PMF) of N may be expressed as

$$
\begin{aligned}
P(N = n) &= P(n - 1 < \frac{T}{\Delta t} \le n) = P[(n-1)\Delta t < T \le n\Delta t] \\
&= \left[1 - \left(1 - \exp\left(-\mu\Delta t\right)\right)\right]^{n-1}\left[1 - \exp\left(-\mu\Delta t\right)\right].
\end{aligned}
\tag{24}
$$

Assuming that Δt is sufficiently low and μ is in the interval of $[0, 1]$, we may rewrite Equation (24) with the aid of exploiting that $1 - \exp(-x) \approx x$ when $x \to 0$ as

$$
P(N = n) \approx (1 - \mu\Delta t)^{n-1} \cdot \mu\Delta t. \tag{25}
$$

Explicitly, Equation (25) is the PMF of the geometrically distributed discrete random variable N associated with the parameter of $\mu\Delta t$, which represents the successful packet reception probability within the duration Δt, while n in Equation (25) indicates that the packet is first successfully received by the target within the nth Δt duration. As a result, the total time dedicated to transmitting this content is $n \cdot \Delta t$. We can clearly see that if Δt is small enough, which makes the discrete random variable near-continuous, we can model the time spent in disseminating the content by an exponential distribution. Hence *Theorem 1* is proven.

References

[1] J. D. Sutter, "Welcome to the twitter olympics," CNN, http://edition.cnn.com/2012/08/01/tech/social-media/welcome-twitter-olympics/index.html, Tech. Rep., August 2012.

[2] N. Kayastha, D. Niyato, P. Wang, and E. Hossain, "Applications, architectures, and protocol design issues for mobile social networks: A survey," *Proceedings of the IEEE*, vol. 99, no. 12, pp. 2130 –2158, December 2011.

[3] J. Zhang, R. Zhang, G. Li, and L. Hanzo, "Remote coalition network elements for base station cooperation aided multicell processing," *IEEE Transactions on Vehicular Technology*, vol. 61, no. 3, pp. 1406 –1415, March 2012.

[4] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1380 –1408, December 2010.

[5] C. Dong, L.-L. Yang, and L. Hanzo, "Performance analysis of multihop-diversity-aided multihop links," *IEEE Transactions on Vehicular Technology,*, vol. 61, no. 6, pp. 2504 –2516, July 2012.

[6] P. Van Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *IEEE/ACM Transactions on Networking*, vol. 17, no. 1, pp. 1 –14, February 2009.

[7] Y.-K. Ip, W.-C. Lau, and O.-C. Yue, "Performance modeling of epidemic routing with heterogeneous node types," in *IEEE ICC 2008*, May 2008, pp. 219 –224.

[8] H. Li, C.-F. Chen, and L. Lai, "Propagation of spectrum preference in cognitive radio networks: A social network approach," in *IEEE ICC 2011*, June 2011, pp. 1 –5.

[9] D. Niyato, P. Wang, W. Saad, and A. Hjϕandrungnes, "Controlled coalitional games for cooperative mobile social networks," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 4, pp. 1812 –1824, May 2011.

[10] L. Cao and P.-Y. Kam, "On the performance of packet ARQ schemes in rayleigh fading: The role of receiver channel state information and its accuracy," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 704 –709, February 2011.

[11] P. V. Mieghem, *Performance Analysis of Communications Networks and Systems*. New York, NY, USA: Cambridge University Press, 2005.

[12] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, 7th ed. Elsevier/Academic Press, Amsterdam, 2007.

[13] J. Wannstrom, "Lte-advanced," May 2012. [Online]. Available: http://www.3gpp.org/lte-advanced

[14] G. Yuan, X. Zhang, W. Wang, and Y. Yang, "Carrier aggregation for lte-advanced mobile communication systems," *IEEE Communications Magazine*, vol. 48, no. 2, pp. 88 –93, February 2010.

[15] "Notice of proposed variation of Everything Everywhere's 1800 MHz spectrum licences to allow use of LTE and WiMAX technologies," *Ofcom*, March 2012. [Online]. Available: http://stakeholders.ofcom.org.uk/consultations/variation-1800mhz-lte-wimax/summary

[16] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 2001.