Cooperative Content Offloading Through WiFi and Mobile Device-to-Device Networks

Guoqiang Mao[1,2] and Xiaofeng Tao[2]

Abstract— **This paper investigates the use of WiFi and mobile device-to-device networks, with vehicular ad hoc networks being a typical example, as a complementary means to offload and reduce the traffic load of cellular networks. A novel cooperative content dissemination strategy is proposed for a heterogeneous network consisting of different types of devices with different levels of mobility, ranging from static WiFi access points to mobile devices such as vehicles. The proposed strategy offloads a significant proportion of data traffic from cellular networks to WiFi or device-to-device networks. Detailed analysis is provided for the content dissemination process in heterogeneous networks adopting the strategy. On that basis, the optimal parameter settings for the content dissemination strategy are discussed. Simulation and numerical studies show that the proposed strategy effectively reduces data traffic for cellular networks while guaranteeing successful content delivery.**

I. INTRODUCTION

Recent years have seen an exponential annual increase in mobile data traffic [1], [2]. Content dissemination and offloading, where a portion of data traffic is offloaded from cellular networks to high-capacity and low-cost complementary networks, including WiFi, vehicular and device-to-device networks, to ease the burden of cellular networks, becomes an increasingly important and challenging task [3]–[9].

In addition to complementary networks, the mobility of users can also be exploited to assist the content dissemination. When users carry mobile devices physically while walking around university campus, shopping centres or travelling by taxis, buses or private vehicles, the content in their mobile devices also move with them without consuming any bandwidth. Together with device-to-device communication technologies such as vehicle-to-vehicle communications, mobility of users offers an alternative way to transport delay-tolerant content efficiently and cost effectively [1].

A large proportion of the content delivered over mobile networks is delay-tolerant content, like videos, newspapers, weather reports and vehicular info-entertainment. For these types of content, content offloading reduces the traffic load of cellular networks (and boosts the capacity of cellular networks in the sense that more users can be served) and also provides a higher data rate for the users. The expense is that content offloading will incur higher delay compared with direct transmission using cellular networks.

This paper proposes a novel cooperative content dissemination and offloading strategy for a heterogeneous network consisting of different types of devices with different levels of mobility, ranging from static WiFi access points (APs) to mobile devices such as vehicles, to reduce the traffic load of cellular networks while guaranteeing the successful delivery of content. The strategy is particularly suited for the delivery of delay tolerant content. More specifically, the following contributions are made in the paper:

1) a cooperative content dissemination strategy is proposed, which exploits all three data dissemination methods, i.e. cellular networks, complementary networks and device mobility, to disseminate content;
2) erasure codes, to be introduced in Section III, are employed to further reduce the data traffic load;
3) analytical results are presented to stochastically characterise the content dissemination process, taking into account the heterogeneity in the devices, in terms of mobility and transmission capability. In particular, the reduction in the amount of data traffic in cellular networks is calculated;
4) based on the above results, optimal parameter settings of the content dissemination strategy are discussed, which minimise the data traffic load of the cellular network while guaranteeing the successful delivery of all content.

The rest of this paper is organised as follows: Section II reviews related work. Section III introduces the system model, including the content dissemination strategy. Section IV presents the analysis of the content dissemination process, whose optimal design is studied in Section V. Section VI validates the analysis using simulations. Finally Section VII concludes this paper and discusses possible future work.

II. RELATED WORK

In this section, we give a brief review of the work most related to this paper.

Recent research [9] has shown that WiFi networks have already carried and offloaded a large amount of mobile data traffic. When a device enters into a WiFi covered area, it can switch its data traffic from cellular networks to WiFi networks [10] to reduce the traffic load of cellular networks. A major issue in WiFi offloading is the optimum deployment of WiFi APs. Bulut et al. [10] compared different methods of deploying WiFi APs for efficient offloading of mobile data traffic. They also proposed a greedy approach that can achieve a high offloading efficiency.

In addition to using WiFi APs, recent technology has also allowed mobile devices to become a virtual WiFi access points (a.k.a. WiFi Tethering), so that devices can communicate with one another via ad hoc connections without relying on infrastructure. In this way, devices can *cooperate* with one another to disseminate content.

Using a typical cooperative content dissemination strategy [11], the service providers first deliver the content to only a small group of users, then these users can further disseminate the content to other subscribed users when their mobile devices are in the proximity and can communicate using WiFi tethering or Bluetooth technology. It is obvious that such opportunistic content dissemination cannot guarantee the delivery of content. This paper proposes a mechanism that can provide guarantee on the delivery of content.

Furthermore, some work in cooperative mobile data offloading (e.g. [1], [12]) investigated the design and selection of helpers, viz. some special mobile devices selected to help the content provider to deliver messages to other mobile devices using ad hoc connections.

As vehicular networks form an important class of mobile networks, there is significant research on content dissemination in vehicular networks [8]. Wireless access through vehicle-to-roadside communications can be used in public transport vehicles for streaming applications, e.g., videos and interactive advertisements. As pointed out in [13], it is challenging to develop an efficient wireless access scheme that minimises the cost of wireless connectivity for downloading data. There are also some empirical studies on content dissemination methods for vehicular networks [14].

III. SYSTEM MODEL AND CONTENT DISSEMINATION STRATEGY

A. Network model

We consider a network of N nodes. These N nodes can be classified into H types according to their mobility, e.g. static WiFi APs, mobile devices carried by pedestrians and vehicles, and social characteristics, e.g. students in the same class. Let N_h be the number of nodes of the hth type: $\sum_{h=1}^{H} N_h = N$.

Suppose that at some initial time instant $t = 0$, the N nodes are independently, randomly and uniformly distributed on a torus $(0, L]^2$ [15]. It follows that the nodes' density in this network is $\lambda = N/L^2$. Then these nodes start to move independently according to some mobility models [16]. We assume that the nodes' mobility is such that the spatial distribution of nodes is stationary and ergodic with stationary uniform distribution on the torus. As shown in [17], a number of mobility models have this property.

The use of a toroidal rather than planar region as a tool in analysing network properties is well known. The use of torus allows nodes located near the boundary to have the same number of connections probabilistically as a node located near the centre. Therefore the consideration of a torus implies that there is no need to consider special cases occurring near the boundary of the region. This often simplifies the analysis without obscuring the relationship between main performance-impacting parameters.

B. Wireless communication model

We consider two types of wireless networks: the cellular network and the complementary network. It is assumed that every node is directly connected to at least one cellular base station (BS) at any time. In the complementary network, devices communicate with one another via ad hoc connections using device-to-device communication technologies such as Bluetooth, WiFi or DSRC [18]. These ad hoc connections are usually of high capacity, and can be exploited by mobile devices to cooperatively share content of common interest.

Due to the limited communication range, the ad hoc connection between two nodes only emerges opportunistically when they move close to each other. Considering a commonly used model, called the *unit disk connection model* (UDM), two nodes are directly connected iff their Euclidean distance is not larger than the radio range r_0. Adopting this commonly used connection model, we say that two nodes *meet* each other when they move into the radio range of each other. Consequently define the *inter-meeting time* of two nodes as the time interval between two consecutive meetings of two nodes (a more rigorous definition is given later in Section IV-A).

There have been a number of studies on the distribution of the inter-meeting time. In particular, Cai and Eun [19] analytically proved that for two nodes moving in a finite area (with reflective or wrapping boundary) under random waypoint or random walk mobility models, their inter-meeting time has an exponential distribution, whereas the inter-meeting time of nodes moving in an infinite area follows a power-law distribution.

This paper also considers that the inter-meeting time follows an exponential distribution. The analysis presented in this paper however is directly applicable to other inter-meeting time distributions as shown in Section IV-A.

C. Cooperative content dissemination strategy

This subsection describes the cooperative content dissemination and offloading strategy.

Consider that a content provider has M messages to deliver to the N nodes. The content dissemination process has three phases: initial phase, sharing phase and complement phase.

Phase 1 Initially at time $t = 0$, the cellular BSs transmit β packets to β different nodes via cellular networks. Note that the content in these packets depends on the coding scheme to be described in Section III-C.1.

Phase 2 Then the network enters into the *sharing phase*, where the nodes broadcast their received packets via ad hoc connections using the Susceptible-Infected-Recovered (SIR) epidemic scheme, to be described in Section III-C.2.

Phase 3 At time $\mathcal{T}_{end}$ (called deadline), the sharing phase stops and then the network enters into the *complement phase*, where every node requests the remaining packets required to decode all M messages from BSs, where the *remaining packets* required to

decode the messages are determined by the coding scheme.

The main objective of the content dissemination strategy is to minimise the total number of packets requiring to be transmitted through cellular networks, which include the packets transmitted in the initial phase and in the complement phase, while guaranteeing the successful delivery of the content.

1) Coding scheme: We consider the use of a simple erasure coding scheme. Specifically, the M messages are encoded into β *coded packets* [20]. We assume that the Galois field [20], [21] used in the encoding process is large enough so that the content provider can generate β linear-independently coded packets. It follows that each node needs to acquire M *distinct* coded packets to reconstruct all the M messages.

Consequently, if a node receives $B < M$ distinct coded packets by the end of the sharing phase, then it needs to request $M - B$ coded packets in the complement phase via cellular networks in order to decode all M messages.

For benchmarking, we also consider the case that no coding technique is employed. In this case, the BSs transmit $\hat{\beta}_m \in \{0, 1, \ldots, \beta\}$ copies of the mth message in the first phase for $m = \{1, 2, \ldots, M\}$ and $\sum_{m=1}^{M} \hat{\beta}_m = \beta$. Then each node needs to receive at least one copy of each message.

2) Epidemic sharing scheme: In the sharing phase, the β packets are shared among N nodes using a Susceptible-Infected-Recovered (SIR) epidemic sharing scheme.

Without loss of generality, consider the epidemic sharing of an arbitrary packet, say packet j. Using a classic SIR broadcast scheme, a node in the network can be in any of the following three states: the node that has never received the packet j is in the state of susceptible (S_j). A susceptible node goes into the state of infected and infectious (I_j) immediately after receiving the packet j. The node in state I_j keeps transmitting the packet j to every node it meets for a certain time period, which is referred to as the *active period*. Denote by τ_h the length of the active period of a type-h node. Note that τ_h is a pre-determined value, which is the same for all nodes of the same type. After the active period, the node recovers and enters into state R_j. A recovered node stops transmitting and receiving the packet j. The nodes that have received the packet j are referred to as the *informed nodes of the packet j*.

Note that the value of τ_h may be different for different types of nodes. For example, a pre-installed WiFi AP can have a significantly larger value of τ_h compared with other mobile devices that are powered by battery. Further, the value of τ_h for a mobile device can be tuned by introducing some incentives (e.g. a lower subscription fee or some rewards) to the mobile users [12], so that the mobile users are willing to share more packets with other users.

When the length of the sharing phase $\mathcal{T}_{end}$ is long, the epidemic sharing process stops naturally (i.e. reaches the *steady state*) when, for all the packets, there is no infectious node. We are particularly interested in the case where the $\mathcal{T}_{end}$ is long because it is suitable for delay-tolerant content

disseminations and minimises the traffic load of the cellular network by fully utilising the complementary networks to share content.

Note that this paper considers a large network where $N \gg M$. Furthermore, in the next section, when we consider an *asymptotic* network with $N \to \infty$, we increase the network area $L \to \infty$ while keeping the density of every type of nodes unchanged. That is, a well-known extended network model is considered. The analytical results obtained are therefore applicable for a large network only.

IV. Analysis of the content dissemination process

The main challenge in the analysis of the content dissemination process is the characterisation of the packet propagation process in the sharing phase. This section first analyses the propagation process of a single packet, say the packet j, and then generalises to multiple packets.

A. Characterising the ad hoc connections

Denote by $T_{h,k}$ the inter-meeting time between a randomly-chosen type-h node and a randomly-chosen type-k node, for $h, k \in \{1, 2, \ldots, H\}$. Assume that $T_{h,k}$ follows an exponential distribution with mean $\lambda_{h,k}$. The pdf of the inter-meeting time is then given by $\Pr(T_{h,k} = t) = \lambda_{h,k} \exp(-\lambda_{h,k}t)$. It follows that the probability that the type-h node meets the type-k node during the active period τ_h of the type-h node is

$$\gamma_{h,k} = \int_0^{\tau_h} \Pr(T_{h,k} = t)dt = 1 - \exp(-\lambda_{h,k}\tau_h). \quad (1)$$

Note that given other probability distribution of inter-meeting time, one can use a similar method to calculate $\gamma_{h,k}$. Thus our analysis does not critically depend on the assumption of the exponential inter-meeting time distribution.

B. Extinct probability

In this subsection, we construct a multi-type branching process [22], [23] to study the number of informed nodes of a typical packet, say packet j, where the 0th generation of the branching process includes the nodes that receive the packet j at time 0 (i.e. in the first phase). Further, the number of type-k children generated by a type-h node is denoted by a random variable $\hat{Q}_{h,k}$. Because nodes move independently of one another, it is evident that $\hat{Q}_{h,k}$ follows a Binomial distribution $\text{Bin}(N_k, \gamma_{h,k})$, where $\gamma_{h,k}$ is given in Eq. 1. Denote by $\mathcal{X}_\alpha^h$ the number of type-h nodes in the αth generation. The branching process modelling the number of informed nodes for a typical packet becomes *extinct* if there exists an integer value of α such that $\sum_{h=1}^{H} \mathcal{X}_\alpha^h = 0$.

The following result is required in the later analysis:

Lemma 1 (Threshold phenomenon)**.** *Define matrix* $\hat{\Xi} \triangleq \{\mathbb{E}[\hat{Q}_{h,k}] = N_k\gamma_{h,k}; h, k = 1, 2, \ldots, H\}$. *Let* R_q *be the largest eigenvalue of* $\hat{\Xi}$. *Then the branching process will become extinct with probability 1 if and only if* $R_q \leq 1$.

This result can be readily obtained by applying Theorem 2 in [22, Chapter V]. Hence the proof is omitted.

When $R_q > 1$, there is a positive probability that the branching process does not become extinct, i.e. the packet

can be disseminated to a significant fraction of nodes as the network size becomes asymptotically large.

Note that in a heterogeneous network, the type of source nodes that a branching process is rooted at can have a significant impact on the probability that the branching process becomes extinct. Denote by w_h the *extinct probability of type-h source*, which is defined as the probability that a branching process rooted at a type-h node becomes extinct. The following theorem characterises w_h.

Theorem 1 (Extinct probability). *Consider an asymptotic network with $L \to \infty$ while keeping nodes' density unchanged and the node's communication range varying with L such that $N_k\gamma_{h,k}$ is a finite constant [1]. The extinct probabilities w_h for $h = 1, 2, \ldots, H$ are the solutions to the following system of equations:*

$$w_h = \exp\left(\sum_{k=1}^{H} N_k\gamma_{h,k}(w_k - 1)\right), \text{ for } h = 1, 2, \ldots, H. \quad (2)$$

Proof. Firstly, note that in the construction of the branching process, we consider that every node that a type-h node meets is a susceptible node. As the packet propagates, the probability that a type-h node meets an informed node increases. This may reduce the expected number of newly informed nodes generated by an infectious node. However, there is no need to consider the impact of this effect in the analysis of the extinct probability because when analysing the extinct probability, we are only interested in the case that the fraction of recipients is vanishingly small (i.e. becomes extinct) as $L \to \infty$ and $N \to \infty$. Accordingly, the probability that a type-h node meets an informed node is vanishingly small and hence negligible.

Because $N_k\gamma_{h,k}$ is a finite constant as the network size increases, the distribution of $\hat{Q}_{h,k}$, i.e. a Binomial distribution $\text{Bin}(N_k, \gamma_{h,k})$, approaches a Poisson distribution with an expected value $N_k\gamma_{h,k}$ [24]. The difference between the $\hat{Q}_{h,k}$ and its Poisson distribution counterpart (denoted by $Q_{h,k}$) diminishes as $N_k \to \infty$ and $\gamma_{h,k} \to 0$, where the convergence rate is given in [24].

Denote by $G_{h,k}(s)$ the probability generating function of $Q_{h,k}$:

$$G_{h,k}(s) = \mathbb{E}[s^{Q_{h,k}}] = \exp(N_k\gamma_{h,k}(s - 1)). \quad (3)$$

Further, define the multi-variate probability generating function $\hat{G}_h(\mathbf{s}) \triangleq \mathbb{E}[s_1^{Q_{h1}} s_2^{Q_{h2}} \ldots s_H^{Q_{hH}}]$, where $\mathbf{s} \triangleq \{s_1, s_2, \ldots, s_H\}$ is a row vector. It can be shown that

$$\hat{G}_h(\mathbf{s}) = G_{h,1}(s_1)G_{h,2}(s_2)\ldots G_{h,H}(s_H). \quad (4)$$

Denote the extinct probabilities by a row vector $\mathbf{w} \triangleq [w_1, w_2, \ldots, w_H]$. Then according to Theorem 2 in [22, Chapter V], the extinct probabilities satisfy $\mathbf{w} = \hat{\mathbf{G}}(\mathbf{w})$, where $\hat{\mathbf{G}}(\mathbf{s})$ is a row vector $[\hat{G}_1(\mathbf{s}), \hat{G}_2(\mathbf{s}), \ldots, \hat{G}_H(\mathbf{s})]$. The conclusion follows that the extinct probabilities w_h for $h = \{1, 2, \ldots, H\}$

[1]This condition is required to avoid triviality in the analysis

are the solutions to the following system of equations:

$$w_1 = \prod_{k=1}^{H} \exp(N_k\gamma_{1,k}(w_k - 1)), \quad (5)$$
$$\cdots$$
$$w_H = \prod_{k=1}^{H} \exp(N_k\gamma_{H,k}(w_k - 1)),$$

where N_k is the number of type-k nodes in the network and $\gamma_{h,k}$ is given by Eq. 1. $\qquad \square$

Using Theorem 1, we can further obtain the extinct probability for several special-case of networks.

Corollary 2. *[Extinct probability for homogeneous networks] Consider the special case of a network with only $H = 1$ type of nodes, the extinct probability is $w_1 = -\frac{\mathcal{W}(-N\gamma_{1,1}\exp(-N\gamma_{1,1}))}{N\gamma_{1,1}}$, where $\mathcal{W}(.)$ is the Lambert-W function.*

Corollary 3. *[Extinct probability with multiple source nodes] Suppose that a packet is initially broadcast from $\beta = \sum_{h=1}^{H} \beta_h$ source nodes at the beginning of phase 2, where β_h is the number of type-h source nodes. Then the extinct probability for this packet is $\prod_{h=1}^{H} w_h^{\beta_h}$.*

The proof of the above two corollaries is straightforward and hence omitted.

Note that when the branching process does *not* become extinct, the packet is disseminated to a significant number of nodes and we say that *the packet spreads out*. The next sub-section quantifies the number of recipients of a packet when it spreads out.

C. Expected fraction of recipients

Denote by $\hat{z}_h$ the expected fraction of type-h nodes which receive the packet j in the steady state, where the packet propagation starts from a randomly-chosen source node *and* the packet spreads out. Now we further investigate $\hat{z}_h$.

Theorem 2 (Fraction of recipients when a packet spreads out). *In a large network with $N \to \infty$, suppose that the packet j is broadcast from a randomly chosen source node, no matter which type the source node belongs to. Given that the packet is spread out, the expected fractions of recipients $\hat{z}_h$ for $h = 1, 2, \ldots, H$ are the solutions to the following system of equations*

$$1 - \hat{z}_h = \exp\left(-\sum_{k=1}^{H} N_k\gamma_{k,h}\hat{z}_k\right), \text{ for } h = 1, 2, \ldots, H. \quad (6)$$

This theorem can be readily obtained from the analysis of epidemics [25], [26] and the proof is omitted.

Next we consider the case that there are more than one source node of the packet j.

Corollary 4 (Fraction of recipients with multiple source nodes). *Suppose that initially at time 0, there are $\sum_{h=1}^{H} \beta_h = \beta$ nodes that have received the packet j, where β_h is the number of type-h source nodes. Denote by $z(\beta_1, \ldots, \beta_H)$ the expected fraction of nodes, out of the total N nodes, which receive the*

packet in the steady state. In a large network with $N \to \infty$, there holds

$$z(\beta_1, \ldots, \beta_H) = \left(\sum_{h=1}^{H} \frac{N_h \hat{z}_h}{N} \right) \left(1 - \prod_{h=1}^{H} w_h^{\beta_h} \right). \qquad (7)$$

Proof. From Corollary 3, the probability that the packet spreads out is $\left(1 - \prod_{h=1}^{H} w_h^{\beta_h} \right)$. Note that if the packet is not spread out, the fraction of recipients goes to 0 as $N \to \infty$. If the packet spreads out, the expected number of type-h recipients is $N_h \hat{z}_h$, where $\hat{z}_h$ is given by Theorem 2. The conclusion follows. □

For the special case that there is only $H = 1$ type of node in a network, a closed form expression can be obtained.

Corollary 5 (Fraction of recipients for homogeneous networks). *Suppose that there is only $H = 1$ type of node in a network and a packet is sent to β different nodes in the first phase. Then in the steady state, the expected fraction of recipients of this packet is*

$$
\begin{aligned}
z(\beta) \quad = \quad & \left(1 + \frac{\mathcal{W}(-N\gamma_{1,1} \exp(-N\gamma_{1,1}))}{N\gamma_{1,1}} \right) \\
& \times \left(1 - \left(-\frac{\mathcal{W}(-N\gamma_{1,1} \exp(-N\gamma_{1,1}))}{N\gamma_{1,1}} \right)^{\beta} \right). \quad (8)
\end{aligned}
$$

V. Minimising the traffic load of cellular networks

Based on the above characterisation of the content dissemination process, this section investigates the optimal content dissemination strategy that minimises the total traffic load of cellular networks.

Recall that in the initial phase, the BSs transmit β packets to β different nodes through the cellular network. Denote by Y the expected number of packets that BSs need to transmit in the complement phase.

Definition 1 (Cellular traffic load). *The cellular traffic load is the expected number of packets transmitted by BSs through the cellular network, which consist of the packets transmitted in the first and the third phases, i.e. $\beta + Y$, in order to transmit M messages to N nodes.*

Note that the value of β determines the value of Y, which is calculated later in this section. Then the problem of minimising the cellular traffic load can be formulated as follows:

$$
\begin{aligned}
& \underset{\beta}{\text{Minimise}} \quad \beta + Y \\
& \text{Subject to} \quad \beta \in \{1, 2, \ldots, N\}.
\end{aligned}
\qquad (9)
$$

Using the erasure coding technique introduced in Section III-C.1, the BSs push β coded packets to β different nodes in the initial phase. The following lemma gives the optimum strategy to choose the source nodes to disseminate the packets.

Lemma 6 (Sharing maximisation strategy). *Label all nodes in the network in the descending order of their values of w_h, which is given by Theorem 1. If more than one node have the same value of w_h, their order can be arbitrarily assigned. Suppose that BSs push β encoded packets to β different nodes*

in the initial phase. Then the optimal strategy that minimises the cellular traffic load is to push β different *packets to the first β nodes in the above order.*

Proof. First, it can be readily shown that the strategy that minimises the cellular traffic load is to push β *different* coded packets in the initial phase rather than pushing multiple copies of any coded packet.

Because different encoded packets are shared independently of one another, we next consider a randomly-chosen packet, say packet j. It is obvious that to minimise the cellular traffic load, one needs to maximise the number of recipients of packet j. According to Corollary 4, the expected fraction of nodes, out of the total N nodes, which receive packet j in the steady state is

$$z(\beta_1, \ldots, \beta_H) = \left(\sum_{h=1}^{H} \frac{N_h \hat{z}_h}{N} \right) \left(1 - \prod_{h=1}^{H} w_h^{\beta_h} \right), \qquad (10)$$

where $\beta_1 + \beta_2 + \cdots + \beta_H = 1$ because we are now considering the propagation of single packet - packet j. In other words, only one value among $\beta_1, \beta_2, \ldots, \beta_H$ is equal to one and the other values are all equal to 0. It is obvious that to maximise $z(\beta_1, \ldots, \beta_H)$, one should assign the values of $\beta_1, \beta_2, \ldots, \beta_H$ in a way that minimises $\prod_{h=1}^{H} w_h^{\beta_h}$, i.e. let the only non-zero value $\beta_k = 1$ for the kth type of nodes that have the smallest value of the extinct probability w_k among all w_h for $h \in \{1, 2, \ldots, H\}$.

It can be shown that to maximise the sharing performance, the β coded packets need to be pushed to β *different* nodes. Therefore, when the number of type-k nodes N_k is less than the total number of packets β, some packets need to be pushed to the nodes that have the second (and if needed, the third, forth, etc.) smallest value of the extinct probability w_i among all w_h for $h \in \{1, 2, \ldots, H\}$. □

Next we focus on determining the optimum value of β for the special case of a homogeneous network with $H = 1$. The optimum value of β for the more general case of a heterogeneous network with $H > 1$ can be determined analogously albeit with greater complexity.

Denote by random variable B the number of packets received by a randomly-chosen node at the end of the sharing phase. Then B follows a Binomial distribution, i.e. the probability that a node receives $B = b$ packets is $\Pr(B = b) = \binom{\beta}{b}(z(1))^b (1 - z(1))^{\beta - b}$, where $\binom{\beta}{b} = \frac{\beta!}{b!(\beta - b)!}$.

Then in the complement phase, the number of packets that need to be transmitted to a randomly-chosen node is $(M - B)^+$, where $(x)^+ = \max\{0, x\}$.

Finally, the expected number of packets that the BSs need to transmit in the complement phases is

$$
\begin{aligned}
Y \quad = \quad & N\mathbb{E}[(M - B)^+] \qquad\qquad\qquad (11) \\
= \quad & N \sum_{b=0}^{M} (M - b)\binom{\beta}{b}(z(1))^b (1 - z(1))^{\beta - b}. \quad (12)
\end{aligned}
$$

Then the optimisation problem in Eq. 9 becomes

$$\text{Minimise} \quad \beta + N \sum_{b=0}^{M}(M-b)\binom{\beta}{b}z_1^b(1-z_1)^{\beta-b} \quad (13)$$
$$\text{Subject to} \quad \beta \in \{1, 2, \dots, N\}.$$

This optimisation problem can be readily solved numerically using Matlab, where the results are presented in the next section.

VI. SIMULATION AND DISCUSSION

This section reports on simulations to verify the accuracy of the analysis presented in the previous sections. The simulations are conducted using a mobile network simulator written in C++. Specifically, $N = 960$ nodes are uniformly deployed on a square $(0, 8000]^2$. Consequently the nodes' density equals to 15 nodes/km^2, which is equal to the density of cabs in New York metropolitan area [27]. After initial deployment of the nodes, they start to move according to the random direction mobility model. When the nodes hit the boundary of the square, they may appear from the other side of the boundary. It can be shown that when nodes move according to the random direction mobility model, the inter-meeting time follows an exponential distribution. The node's speed is $V = 10m/s$ (typical vehicle moving speed). The radio range $r_0 = 20$m or 250m (typical radio ranges using Wi-Fi Tethering or DSRC [18]). Every point shown in the simulation result is the average value from 500 simulations.

Consider two types of nodes moving according to the random direction model with speeds $V_1 = 10m/s$ and $V_2 = 0$, i.e. mobile and static nodes, and there are equal number of nodes in each type. Fig. 1 shows the probability that a packet spreads out and the expected fraction of recipients of a single packet. The analytical result of the probability that a packet spreads out when the source node is of type-1 (resp. type-2) is given by w_1 (resp. w_2) from Theorem 1. The analytical result of the expected fraction of recipients of a packet when the source node is of type-1 (resp. type-2) is given by $z(1, 0)$ (resp. $z(0, 1)$) from Corollary 4. It is interesting to note that the probability that a packet spreads out and the expected fraction of recipients can be significantly affected by the type of source node. More specifically, other things being equal, it can be seen that a mobile source node can spread the packet to more recipients than a static source node.

H=2, V_1=10, V_2=0, λ=0.000015, τ_1=τ_2=τ=1~50, β=1, Source Node Type (ST)=1 or 2
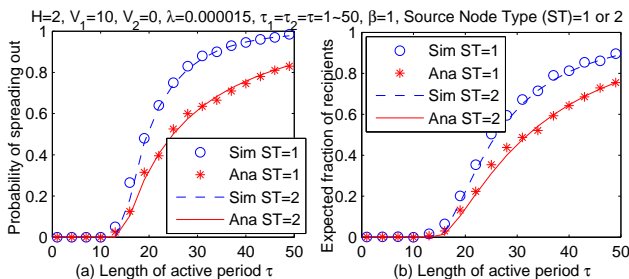
Fig. 1.   Simulation and analytical results of (a) The probability that a packet spreads out and (b) the expected fraction of recipients of a packet.

Fig. 2 shows the results of another interesting case where the complementary network consists of a set of fixed-location base stations (e.g. WiFi APs). In the second phase, the message is disseminated from these WiFi APs to mobile users (i.e. type-1 nodes). Specifically, we consider that a small number of WiFi APs are randomly and uniformly deployed in a given area, which is a widely-used setting for AP deployment [2]. Because the WiFi APs are usually connected to the Internet via wired connections, we set $\gamma_{2,2} = 1$. Other parameters are the same as those in the previous sub-section. A message is transmitted to all the WiFi APs at time 0. Then the WiFi APs keep transmitting the message for a given time period τ_2. In Fig. 2, we let the active period of type-2 nodes be $\tau_2 = 500, 1500$ while varying the active period of type-1 nodes. Note that $\tau_1 = 0$ corresponds to the traditional case [9], [10] where nodes do not cooperatively share received packets and they solely rely on WiFi APs to offload data traffic from cellular networks. It can be seen in Fig. 2 that a longer active period of mobile nodes τ_1 leads to a larger expected fraction of recipients. It is obvious that packet sharing using ad hoc connections between mobile nodes can significantly increase the number of recipients of a packet, hence reducing the number of transmissions required by BSs.

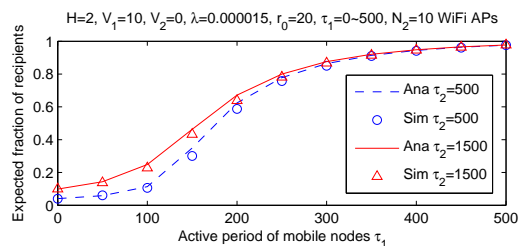H=2, V_1=10, V_2=0, λ=0.000015, r_0=20, τ_1=0~500, N_2=10 WiFi APs

Fig. 2.   The expected fraction of recipients of a packet in a network with 10 fixed WiFi APs and some mobile nodes.

The following results further evaluate the cellular traffic loads.

Fig. 3 shows the expected cellular data traffic load $\beta + Y$ with different values of β, viz. the number of packets sent in the first phase, using different content dissemination strategies. To study the impact of coding on the cellular traffic load, we consider two networks with the same setting except that one network employs the erasure coding technique (c.f. Section III-C) but the other network does not. The performance of the network without employing network coding can be readily obtained using the same technique adopted for analysing network employing the erasure coding technique.

Several interesting trends can be observed in Fig. 3. Firstly, when β is small, two networks have a similar and relatively high cellular traffic load. This is because only a limited number of nodes receive the packets through the complementary network, hence most packets are directly transmitted to the users via cellular networks. As β increases, the cellular traffic load first decreases rapidly, due to a rapid increase in the expected fraction of recipients in the sharing phase. Then after a certain point, the cellular traffic load starts

to gradually increases as β further increases. This is because the expected fraction of recipients has limited increase when β increases further; on the other hand the increase in β causes more cellular data traffic. It is interesting to note that sending out more packets in the initial phase is not always beneficial.

Furthermore, it is interesting to note that above a certain value of β, e.g. $\beta = 17$ in Fig. 3(a), the traffic load of cellular networks employing coding is significantly smaller than that of networks without coding, due to the following reason. Recall that when $\beta > M$, each node only needs M different coded packets to decode all M messages when coding is employed. On the other hand, in a network without coding, there is a non-zero probability that two packets received at a node contain the same message. Therefore a node may need more than M packets.

Note that without employing the cooperative content dissemination strategy, the nodes need to request all packets from the cellular network and the traffic load of cellular networks is 960. Compared with the values of $\beta + Y$ in Fig. 3, it is evident that the cooperative content dissemination strategy can significantly reduce the traffic load of cellular networks.
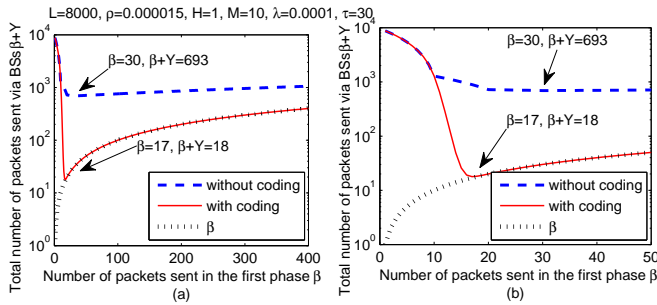

Fig. 3. Comparison between the expected cellular data traffic loads $\beta + Y$ in networks with and without coding. Note that subplot (b) shows the range $\beta \in [1, 50]$ of subplot (a).

VII. Conclusion and future work

This paper investigated a cooperative content dissemination strategy for heterogeneous networks consisting of different types of devices. The content dissemination strategy can effectively offload a significant amount of data traffic from cellular networks to complementary networks such as WiFi and device-to-device networks using ad hoc connections that emerge when devices move and meet one another. Theoretical analysis for the content dissemination process was presented. On that basis, the optimal design of the content dissemination strategy was discussed, which maximally reduces traffic load of cellular networks while guaranteeing the successful delivery of all content.

In our paper, we consider that the duration of the sharing phase $\mathcal{T}_{end}$ is sufficiently long such that the epidemic sharing process is able to reach its steady state. In the future, one may consider the case where only a short time period is allowed for the sharing phase, causing the epidemic sharing process to terminate before it reaches the steady state. In this case, non-trivial analysis is required to calculate the number of recipients of a packet at an arbitrary time instant. Furthermore, it is an interesting extension of our work to consider different probability distributions for the inter-meeting time of nodes, which can be affected by the nodes' mobility and network area as described in Section III.

References

[1] Y. Li, M. Qian, D. Jin, P. Hui, Z. Wang, and S. Chen, "Multiple mobile data offloading through disruption tolerant networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 7, pp. 1579 - 1596, 2014.

[2] J. G. Andrews, "Seven ways that HetNet are a cellular paradigm shift," *IEEE Communications Magazine*, vol. 51, no. 3, pp. 136–144, 2013.

[3] A. Kannan, B. Fidan and G. Mao, "Robust Distributed Sensor Network Localization Based on Analysis of Flip Ambiguities," *IEEE Globecom*, pp. 1–6, 2008.

[4] X. Ge, S. Tu, T. Han, Q. Li and G. Mao, "Energy Efficiency of Small Cell Backhaul Networks Based on Gauss-Markov Mobile Models," *IET Networks*, vol. 4, no. 2, pp. 158–167, 2015.

[5] R. Mao and G. Mao, "Road Traffic Density Estimation in Vehicular Networks," *IEEE WCNC*, pp. 4700–4705, 2013.

[6] G. Mao and B. D. O. Anderson, "Graph Theoretic Models and Tools for the Analysis of Dynamic Wireless Multihop Networks," *IEEE WCNC*, pp. 1–6, 2009.

[7] G. Mao, B. D. O. Anderson and B. Fidan, "Online calibration of path loss exponent in wireless sensor networks," *IEEE Globecom*, pp. 1–6, 2006.

[8] Z. Da and Y. Chai Kiat, "Enabling efficient wifi-based vehicular content distribution," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 3, pp. 479–492, 2013.

[9] K. Lee, J. Lee, Y. Yi, I. Rhee, and S. Chong, "Mobile data offloading: How much can wifi deliver?" *IEEE/ACM Transactions on Networking*, vol. 21, no. 2, pp. 536–550, 2013.

[10] E. Bulut and B. K. Szymanski, "Wifi access point deployment for efficient mobile data offloading," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 17, no. 1, pp. 71–78, 2013.

[11] A. Aijaz, H. Aghvami, and M. Amani, "A survey on mobile data offloading: technical and business perspectives," *IEEE Wireless Communications*, vol. 20, no. 2, pp. 104–112, 2013.

[12] H. Bo, H. Pan, V. S. A. Kumar, M. V. Marathe, S. Jianhua, and A. Srinivasan, "Mobile data offloading through opportunistic communications and social participation," *IEEE Transactions on Mobile Computing*, vol. 11, no. 5, pp. 821–834, 2012.

[13] D. Niyato and E. Hossain, "A unified framework for optimal wireless access for data streaming over vehicle-to-roadside communications," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 6, pp. 3025–3035, 2010.

[14] S. Panichpapiboon and W. Pattara-Atikom, "A review of information dissemination protocols for vehicular ad hoc networks," *IEEE Communications Surveys and Tutorials*, vol. 14, no. 3, pp. 784–798, 2012.

[15] M. Franceschetti and R. Meester, *Random networks for communication: From Statistical Physics to Information Systems*, Cambridge University Press, 2007.

[16] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communications and Mobile Computing*, vol. 2, no. 5, pp. 483 – 502, 2002.

[17] P. Nain, D. Towsley, B. Liu, and Z. Liu, "Properties of random direction models," in *Proceedings IEEE INFOCOM*, vol. 3, 2005, pp. 1897– 1907.

[18] J. Rezgui and S. Cherkaoui, "About deterministic and non-deterministic vehicular communications over DSRC/802.11p," *Wireless Communications and Mobile Computing*, vol. 14, no. 15, pp. 1435-1449, 2012.

[19] H. Cai and D. Y. Eun, "Crossing over the bounded domain: From exponential to power-law intermeeting time in mobile ad hoc networks," *IEEE/ACM Transactions on Networking*, vol. 17, no. 5, pp. 1578–1591, 2009.

[20] M. Sathiamoorthy, A. Dimakis, B. Krishnamachari, and F. Bai, "Distributed storage codes reduce latency in vehicular networks," *IEEE Transactions on Mobile Computing*, vol. 13, no. 9, pp. 2016-2027, 2013.

[21] Y. Li, D. Jin, Z. Wang, L. Zeng, and S. Chen, "Coding or not: Optimal mobile data offloading in opportunistic vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 1, pp. 318-333, 2013.

[22] K. B. Athreya and P. E. Ney, *Branching Processes*, Springer-Verlag, 1972.

[23] T. E. Harris, *The Theory of Branching Processes*, Dover Publications, 2002.

[24] D. Pollard, *Poisson Approximation*, 1997.

[25] F. Ball and D. Clancy, "The final size and severity of a generalised stochastic multitype epidemic model," *Advances in Applied Probability*, vol. 25, no. 4, pp. 721–736, 1993.

[26] T. Britton, "Stochastic epidemic models: A survey," *Mathematical Biosciences*, vol. 225, no. 1, pp. 24–35, 2010.

[27] MKThink, "Unsustainable city: density, transportation, and san franciscos sustainability," 2005. [Online]. Available: http://www.mkthink.com/archives/2470