Performance Modeling of Network Coding in Epidemic Routing

Yunfeng Lin, Ben Liang, and Baochun Li
Department of Electrical and Computer Engineering
University of Toronto
Toronto, Ontario, Canada
ylin@eecg.toronto.edu, liang@comm.toronto.edu, bli@eecg.toronto.edu

ABSTRACT

Epidemic routing has been proposed to reduce the data transmission delay in opportunistic networks, in which data can be either replicated or network coded along the opportunistic multiple paths. In this paper, we introduce an analytical framework to study the performance of network coding based epidemic routing, in comparison with replication based epidemic routing. With extensive simulations, we show that our model successfully characterizes these two protocols and demonstrates the superiority of network coding in opportunistic networks when bandwidth and node buffers are limited. We then propose a priority variant of the network coding based protocol, which has the salient feature that the destination can decode a high priority subset of the data much earlier than it can decode *any* data without the priority scheme. Our analytical results provide insights into how network coding based epidemic routing with priority can reduce the data transmission delay while inducing low overhead.

1. INTRODUCTION

Opportunistic networks or disruption tolerant networks (DTN) represent a class of networks where nodes do not have contemporaneous connections, but intermittent connections. Such networks usually have sparse node density, and each node has short radio range. The connection between nodes may be disrupted due to node movements, node power-saving sleep schedules, and harsh environment changes. The examples of opportunistic networks include networks in an undeveloped area without Internet connections, sensor networks monitoring nature and military fields, or mobile opportunistic networks composed of moving vehicles and pedestrians.

For a mobile opportunistic network, an *opportunistic link* can be setup when a pair of nodes move into the radio range of each other such that they can communicate directly. A possible data propagation path from the source to the destination, referred to as an *opportunistic path*, is composed of multiple opportunistic links. Clearly, multiple opportunistic paths exist by node movements. Epidemic routing has been proposed to utilize such multiple opportunistic paths to reduce data delivery delay by replicating packets whenever two nodes meets. In essence, epidemic routing replicates data along the multiple opportunistic paths from the source to the destination. The delay in delivering a packet is hence the time to propagate a packet in the shortest opportunistic path.

Network coding [1], along with its randomized distributed implementation [12, 5], allows intermediate nodes perform coding operations besides replication and forwarding. Using the paradigm of network coding in epidemic routing, a node may transmit a coded packet, a random linear combination of data packets, to another node during a transmission opportunity. In contrast to such *network coding based epidemic routing*, the traditional epidemic routing is referred to as *replication based epidemic routing* in this paper.

In this paper, we focus on studying epidemic routing in realistic network environments with limited bandwidth and node buffers. In such environments, if using the replication based protocol, when a transmission opportunity arrives, ideally, a node should transmit the packet with the minimal number of replicas in the network to reduce its transmission delay, since it is the packet with the longest expected delivery delay. However, a node has no such precise knowledge in opportunity networks. Therefore, it is difficult to select the *best* packet for transmission. On the other hand, in the network coding based protocol, a node can transmit any coded packets since all of them can contribute the same to the eventual delivery of all data packets to the destination with high probability. Similarly, the network coding based protocol has the advantage in utilizing limited buffer resource since dropping any coded packet has the same effect.

In this paper, we propose an analytical framework to characterize network coding based and replication based epidemic routing protocols. Our analytical model demonstrates that the network coding based protocol delivers data with shorter delay when bandwidth is limited and such advantage is more significant when the buffer sizes are constrained.

However, in network coding based epidemic routing, one has to pay the price that any useful data can be decoded only after the destination receives a sufficient number of coded packets and can decode all data altogether. That is, the destination may wait too long before any useful data can be decoded. Hence, we propose a simple priority coding protocol that decodes high priority data much earlier than the original network coding based protocol can decode

MobiOpp'07, June 11, 2007, San Juan, Puerto Rico, USA.

any data. Utilizing our analytical model, we show that the priority protocol achieves such a goal with low overhead.

The remainder of the paper is organized as follows. We compare our work with related work in Sec. 2 and describe the network model in Sec. 3. We present the analytical models for network coding based and replication based epidemic routing protocols in Sec. 4 and Sec. 5, respectively. The analytical results are verified by experiments in Sec. 6. Sec. 7 introduces our simple priority coding protocol and investigates the tradeoff in the protocol design, using our analytical framework. We conclude the paper in Sec. 8.

2. RELATED WORK

There have been multiple efforts proposing different variants of DTN routing protocols based on different assumptions of the underlying DTN networks. Some protocols assume prior knowledge on connectivity patterns, *e.g.* [14], or that the past mobility patterns can be used to predict future node movements and message delivery probabilities [2], others assume control over node movements [27]. The purpose of this paper is to understand the performance of the network coding based epidemic routing protocol and its variants with no knowledge of network connectivities and no control on node movements.

Previous studies have proposed to use erasure coding to combat network failures on opportunistic networks with no information of node mobility patterns [23] or DTN networks with prior knowledge of network topology [13]. Chen *et al.* [4] further demonstrate a hybrid approach combining erasure coding and replication. Unlike network coding, in such source-based erasure coding approaches, different upstream nodes may transmit duplicate coded data to the same node and waste bandwidth in a multi-hop (opportunistic) network.

It has been shown that network coding can save data transmissions for both unicast [15] and broadcast applications [8] by exploring the broadcast nature of the wireless medium. However, in the sparse DTN environment considered in this paper, a node seldom has more than one neighbors and such wireless coding opportunities rarely occur.

Deb *et al.* [6] demonstrated that a gossip protocol based on network coding can broadcast multiple messages among nodes with a shorter period of time, than a gossip protocol without network coding, by a logarithmic factor. With the same spirit, the benefit of network coding on wireless network on broadcast applications has been investigated in [9, 24]. In contrast to their work, we show that network coding can efficiently utilize multiple opportunistic paths in unicast applications.

Several research efforts [11, 20, 10, 26, 18, 21, 22] have analytically study replication based epidemic routing. However, none of them has studied the performance of network coding based epidemic routing. Chou *et al.* [5] consider priority encoding in network coding on networks with known topologies. In contrast, our performance modeling and the priority coding protocol are for opportunistic networks without topology information.

Zhang *et al.* [25] studied the benefit of network coding for unicast applications on opportunistic networks, which is, to the best of our knowledge, the closest to our work. However, they use only simulations in their investigation. Our work differs from theirs in three folds. First, we propose an analytical framework, which can be used to study the tradeoff in designing new protocol variants. Second, we have introduced a priority coding protocol to combat the disadvantage of decoding delay in the coding based protocol. Furthermore, we have utilized our analytical framework to show how the proposed simple priority coding protocol is effective and induces low overhead.

3. NETWORK MODEL

Our model of opportunistic network consists of N relay nodes nodes, one source, and one destination node, moving within a constrained area. The source has K packets to be transmitted to the destination. Two nodes *meet* and can transmit data packets to each other when they are within the transmission range of each other. Throughout the paper, we assume there is no background traffic in the network during the transmission of the K packets. Such network model is realistic for mobile sensor networks detecting infrequent events. We leave the analysis for the protocol performance when there is background traffic in the network as futurework.

In this paper, we study the performance of epidemic routing in network environments where the bandwidth and node buffers are limited. In particular, to simplify the analysis but still capture the essence of protocol details, when node i and node j meet, we assume that the bandwidth is only sufficient to transmit one packet from one node to the other and vice versa. It is straightforward to extend the model to the general case where the bandwidth during node meeting is sufficient to deliver an arbitrary number of packets. We further assume that the source node and the destination nodes have sufficient buffer space to hold all K packets. However, the relay buffers on all other nodes have size B, where $1 \leq B \leq K$. Finally, we assume the relay buffers can be cleaned by either an ACK from the destination after it receives all K packets or the expiration of a global timer.

We notice most analytical work in opportunistic networks either explicitly assume that the pairwise meeting time between nodes is exponentially distributed [10, 26, 18] or implicitly assume the Markov property of underlying mobility model while using measured meeting rate in simulations as a parameter in the mobility model [11, 20]. In addition, Groenevelt *el at.* [10] have shown that the inter meeting time between any pair of nodes is almost exponentially distributed if the following three conditions hold. First, nodes move according to the common mobility modes such as the random way point or random direction model. Second, node transmission range is small compared to the area of the node moving region. Third, the speed of nodes is sufficiently high. Although there are measurement evidences (*e.g.*, [3]) that the node meeting time may be distributed in heavy tail in some applications, we believe the insight gained from the analytical result on the performance difference of various protocols based on simple and tractable mobility models is a good indication of their performance difference based on more realistic mobility models. Therefore, throughout the paper, we assume that the node meeting time is exponentially distributed and let λ denote the pairwise meeting rate.

4. NETWORK CODING BASED EPIDEMIC ROUTING

In this section, we develop the analytical model for the network coding based epidemic routing.

4.1 Protocol

We first describe the protocol details. When two nodes meet, they transmit coded packets to each other. Let node a and node b denote the two meeting nodes. A coded packet x is a linear combination of the K source packets $E_1, \ldots, E_K$ in the form: $x = \sum_{i=1}^{K} \alpha_i E_i$, where α_i are coding coefficients. Suppose node a holds m coded packets in its buffer, where $1 \leq m \leq B$. Node a encodes all coded packets in its buffer, namely $x_1, \ldots, x_m$, to generate a coded packet x_a by combining them together:

$$x_a = \sum_{i=1}^{m} \beta_i x_i, \tag{1}$$

where β_i is randomly chosen from a Galois field. It is easy to see that x is also the linear combination of the K source packets with different coding coefficients. Node a then transmits x_a along with its coding coefficients to node b. When node b receives x_a, it inserts x_a into its buffer if there is free space. Otherwise, node b encodes x with each packet in its buffer as follows:

$$x'_i = x'_i + \alpha x_a, \tag{2}$$

where x'_i represents the ith coded packet in the buffer of node b, and α is randomly chosen from a Galois field.

The destination obtains a coded packet when it meets another node, and attempts to decode the K source packets as long as K coded packets have been collected. Because the coding coefficients and the coded packet are known, each coded packet represents a linear equation with the K source packets as the unknown variables. Decoding the K source packets is equivalent to solving the linear system composed of the K coded packets. The *decoding matrix* represents the coefficient matrix of such linear system. When the rank of the decoding matrix is K, the linear system can be solved and the K source packets are decoded. Otherwise, there is linear dependence among the K coded packets, and the node will continue to obtain more coded packets until the K source packets can be decoded.

4.2 Analytical Model

We proceed to describe the analytical model. Our ultimate goal is to compute the delivery delay of all K packets from the source to the destination. If there are more nodes with coded packets in their buffers, the destination has higher opportunity to get a useful coded packet from a contact with another node and proceeds towards the decoding of all K packets. Hence, to compute the delivery delay of all K packets from the source to the destination, we first compute the network state, defined here as the packet distribution on the relay nodes. Let B denote the maximal relay buffer size. We classify the relay nodes in the network by three types: the nodes with no coded packets, the nodes with 1 to $B-1$ coded packets, and the nodes with B coded packets, denoted by v_O, v_M, and v_B, respectively. We then use a 2-tuple $\{X_M(t), X_B(t)\}$ to represent the network state at time t, where $X_M(t)$ and $X_B(t)$ denote the number of v_M and v_B in the network at time t, respectively. We further use $X_O(t)$ to represent the number of v_O. Obviously, we have $X_O(t) = N - X_M(t) - X_B(t)$.

We examine the transmission opportunity when two nodes meet. We say that one node can transmit a *novel* coded packet to another node, if the coded packet it transmits can increase the rank of the decoding matrix on the other node. Clearly, either v_M, v_B or the source can transmit a novel coded packet to v_O. We make the following important assumption in the analysis: *v_M or v_B can transmit a novel coded packet to another v_M with high probability.* In the case of abundant buffers, Deb *et al.* [6] have shown that the probability that a coded packet is useful to another node is $1 - 1/q$, where q is the size of the Galois Field to generate random coding coefficients. In practice, q is usually sufficiently large such that $1 - 1/q$ is very close to 1. Although the relay buffer is limited in our protocol, we will see that the numerical analysis based on such assumption is still very close to the simulation result in Sec. 6.

Let $D_O(t)$, $D_M(t)$, and $D_B(t)$ denote the receiving rate of v_O, v_M, and v_B, *i.e.*, the expected number of novel coded packets received in unit time interval for v_O, v_M, and v_B. Since v_O and v_M can receive a novel coded packet from any relay node with at least one coded packet, namely v_M, v_B, and the source node, with probability 1, as discussed previously, we have

$$\begin{aligned} D_O(t) &= \lambda(X_M(t) + X_B(t) + 1), \\ D_M(t) &= \lambda(X_M(t) + X_B(t) + 1), \\ D_B(t) &= 0, \end{aligned} \tag{3}$$

where the last equation holds since the relay buffer size is B and all packets in the relay buffer are linearly independent with high probability.

Next, we consider the changing rate of $X_M(t)$, which is composed of two parts. First, $D_O(t)X_O(t)$ number of v_O becomes v_M since they obtain one novel coded packet. Second, $D_M(t)X_M(t)/(B-1)$ number of v_M becomes v_B because $D_M(t)X_M(t)$ number of v_M obtain one novel packet within a short time interval, but only $1/(B-1)$ of them become v_B, assuming the fraction of nodes with different number of coded packets are approximately identical. Similarly, the changing rate of $X_B(t)$ is $D_M(t)X_M(t)/(B-1)$. Therefore, we can use the following Ordinary Differential Equations (ODEs) to compute $X_M(t)$ and $X_B(t)$:

$$\begin{aligned} \frac{dX_M}{dt} &= D_O(t)X_O(t) - D_M(t)X_M(t)/(B-1), \\ \frac{dX_B}{dt} &= D_M(t)X_M(t)/(B-1), \end{aligned} \tag{4}$$

with the initial values $X_M(0) = 0$, $X_B(0) = 0$, and $X_O(t) = N - X_M(t) - X_B(t)$.

We proceed to compute the distribution of the delivery delay from the time that the source begins transmitting data to the time that the destination decodes all K packets. We use the random variable T_M and T_K to denote the time that the destination obtains 1 and K coded packets, respectively. Let $F_M(t)$ and $F_K(t)$ be the Cumulative Distribution Function (CDF) of T_M and T_K, respectively.

We derive $F_K(t)$ with ODEs by computing the derivative of $F_K(t)$ over t. In particular, to derive $F_K(t)$, *i.e.* $\Pr(T_K < t)$ with ODEs, we compute the value change of $\Pr(T_K > t)$ within a small time interval $[t, t + \delta t]$. Hence, we can compute the CDF $F_K(t)$ of the delivery delay of K packets by solving the following ODEs, where the derivation details are presented in [16] due to space constraint:

$$\begin{aligned} \frac{dF_M}{dt} &= D_O(t)(1 - F_M(t)), \\ \frac{dF_K}{dt} &= D_M(t)(F_M(t) - F_K(t))/(K-1). \end{aligned} \tag{5}$$

The initial values for the above ODEs are $F_M(0) = 0$ and $F_K(0) = 0$. $D_O(t)$ and $D_M(t)$ are given in (3) by solving (4).

Evidently, when $B = 1$ or $K = 1$, (4) and (5) are no longer valid. However, the analysis of such cases can be trivially extended from the above model, which is presented in [16].

5. REPLICATION BASED EPIDEMIC ROUTING

In this section, to compare with the coding based protocol, we analyze the replication based protocol.

5.1 Protocol

We first describe the protocol details. When two nodes, *e.g.* node a and b, meet, we assume through the exchanging of packet identifiers, node a knows the set of packets in node b and vice versa. Let S_a and S_b denote the set of packets on node a and b, respectively. In the following, we describe the protocol for only node a since the protocol for node b is identical. Node a chooses one packet in the set $S_a - S_b$ to transmit to node b such that the packet transmitted to node b is always new to node b. If $S_a - S_b$ is empty, node a will miss this transmission opportunity.

We examine three policies in selecting which packet from $S_a - S_b$ to be transmitted. First, in the *random* policy, node a chooses a packet with the same probability for each packet in $S_a - S_b$. Second, in the *local rarest* policy, node a uses a counter for each packet in the buffer to record how many times that each packet has been transmitted and chooses the packet with the smallest counter. Third, in the *global rarest* policy, we assume that an oracle maintains the global counters for K packets, the number of copies of each packet in the network. Node a chooses the packet with the smallest counter to transmit. It is clear that the last two policies try to maintain an even distribution of the copies of the K different packets in the network. Although the global rarest policy is impractical, by comparing it with the other two policies and the analytical result, we have clearer understanding on the assumption made in the modeling and the difference between the simulation and analysis results as we will show in Sec. 6.

Upon receiving a packet P_a from node a, node b inserts P_a into its buffer if the buffer is not full. If the buffer is already full, node b uses P_a to replace a random packet in its buffer in the random policy. In the local or global rarest policy, node b compares the local or global counter of P_a with the counter of P_b, the packet that has the largest counter among all packets in the buffer of node b, and drops the one with the larger counter.

5.2 Analytical Model

We proceed to study the delivery delay of the above replication based protocol. If there are more nodes with K packets in their buffers, the destination has higher opportunity to get a new packet from a contact with another node. Hence, to compute the delivery delay of all K packets from the source to the destination, we first compute the network state, the packet distribution on the relay nodes. Let $B \le K$ denote the size of the relay buffer on all relay nodes. We classify the relay nodes in the network by $B+1$ types: the nodes with i packets, denoted by v_i, where $0 \le i \le B$. We then use a B-tuple $\{X_1(t), \ldots, X_B(t)\}$ to represent the network state at time t, where $X_i(t)$ denotes the number of v_i in the network. We further use $X_0(t)$ to represent the number of v_0 and its value is $N - \sum_{i=1}^{B} X_i(t)$.

We make the following assumption in analysis: *the i packets on v_i are uniformly distributed among the K original packets.* This assumption is reasonable if the global rarest policy are employed since it maintains close to even proportion of K packets in the network. We will show the accuracy of this assumption on all three policies in Sec. 6.1. We then examine the probability $\Pr(i, j)$ that v_i obtains a new packet from v_j under such assumption. First, it is easy to see that, if $i < j$, v_i can always obtain a new packet from v_j. Second, if $i \ge j$, v_i cannot obtain a new packet from v_j only if v_i contains all packets on v_j, which has the probability $\binom{i}{j}/\binom{K}{j}$ under the assumption of uniform packet distribution. Hence, we have $\Pr(i, j) = 1 - \binom{i}{j}/\binom{K}{j}$ in such case. In summary, we have

$$\Pr(i, j) = \begin{cases} 1 & \text{if } i < j, \\ 1 - \binom{i}{j}/\binom{K}{j} & \text{if } i \ge j. \end{cases} \tag{6}$$

We notice that similar analysis has been applied in BitTorrent like P2P file sharing systems such as in [7].

Let $D_i(t)$ denote the receiving rate, the expected number of new packets received in unit time interval, of v_i, for $1 \le i \le B$. We further use $D_{B+1}(t) \ldots D_K(t)$ to denote the receiving rate of the destination, when it has obtained $B+1, \ldots, K$ packets, respectively. For v_0, it can receive new packet from any relay node with at least one packet, namely v_j where $1 \le j \le B$, and the source node with probability 1. For v_i, it can receive new packets from v_j with probability $\Pr(i, j)$ and the source node with probability 1. Similar arguments also apply to the receiving rates of the destination. Hence, we have

$$\begin{aligned} D_0(t) &= \lambda\left(\sum_{j=1}^{B} X_j(t) + 1\right), \\ D_i(t) &= \lambda\left(\sum_{j=1}^{B} X_j(t)\Pr(i, j) + 1\right), \\ &\quad \text{for} \quad i = 1, \ldots, K-1, \\ D_K(t) &= 0, \end{aligned} \tag{7}$$

where $\Pr(i, j)$ is computed in (6). $D_0(t), \ldots, D_B(t)$ are useful for both relay nodes and the destination, whereas $D_{B+1}(t), \ldots, D_K(t)$ are useful for only the destination since relay nodes can hold at most B packets.

Next, we consider the changing rate of $X_i(t)$ within a short time interval, which is composed of two parts. First, $D_{i-1}(t)X_{i-1}(t)$ number of v_{i-1} becomes v_i since they obtain one new packet. Second, $D_i(t)X_i(t)$ number of v_i becomes v_{i+1} since they also obtain one new packet. Therefore, we can use the following ODEs to compute $X_i(t)$:

$$\begin{aligned} \frac{dX_i}{dt} &= D_{i-1}(t)X_{i-1}(t) - D_i(t)X_i(t), \\ &\quad \text{for} \quad i = 1, \ldots, B-1, \\ \frac{dX_B}{dt} &= D_{B-1}(t)X_{B-1}(t), \end{aligned} \tag{8}$$

where $D_i(t)$ is computed in (7) as a function of $X_i(t)$. The above ODEs can be solved with the initial value $X_i(t) = 0$ for $i = 1, \ldots, K$.

We proceed to compute the distribution of the delivery delay from the time that the source begins transmitting data to the time that the destination obtains all K packets. We use the random variable T_i to denote the time that the destination obtains i packets. Hence, the delivery delay for all K packets is T_K. We derive the distribution of T_i similar to the derivation of (5):

$$\begin{aligned} \frac{dF_1}{dt} &= D_0(t)(1 - F_1(t)), \\ \frac{dF_i}{dt} &= D_{i-1}(t)(F_{i-1}(t) - F_i(t)), \\ &\quad \text{for} \quad i = 2, \ldots, K. \end{aligned} \tag{9}$$

The initial values of the above ODEs are $F_i(0) = 0$, for $i = 1, \ldots, K$, and $D_i(t)$ is given in (7) by solving (8).

6. MODEL VALIDATIONS

In this section, we use experiments to verify the accuracy of our ODE models. We also show that our analytical result can demonstrate the the advantage of the network coding based protocol over the replication based protocol when bandwidth and buffer are limited. We have developed a discrete-event simulator with the implementation of epidemic routing and network coding. To mitigate randomness in simulations, we show, for each data point in all figures, the average and the 95% confidence intervals from 100 independent experiments. We set the node meeting rate λ to 0.005 and the number of packets K to 10 in most experiments unless explicitly pointed out. We use GF(2^8) as the Galois fields where network coding is operated in all simulations.

6.1 The Case for Limited Bandwidth

We first study the impact of the number of relay nodes on the delivery delay of $K = 10$ packets. The relay buffer size is set to 10 in this experiment such that the buffer is sufficient to hold all K packets on each relay node. From Fig. 1, we observe that the analytical result is close to the simulation result for the global rarest policy in the replication based protocol. The delivery delay of the random policy is longer than the delivery delay of the global rarest policy since the assumption that the packets on a node are uniformly distributed among all K packets is less accurate. The delivery delay of local rarest is much longer than random and global rarest policy. This shows that local counters are not an accurate estimation of the proportion of packets in the entire network. One may imagine if the nodes use the average of the local packet counters of the last several nodes it meets and its own counters as a more accurate estimation. In the following, we omit the experimental result for the local rarest policy.

There is a gap between the numerical result and the simulation result because in the replication based protocols, the packets distribution on buffer are not exactly uniformly distributed. Furthermore, for the case of network coding based protocol, we have ignored linear dependence among coded packets. Nevertheless, such approximation simplifies the analysis while captures the difference between protocols.

Fig. 1 also shows that the analytical result of the replication and network coding based protocols are almost identical. This illustrates that, theoretically, network coding can achieve even distribution of all packets without exchanging packet identifiers as in the replication based protocol.

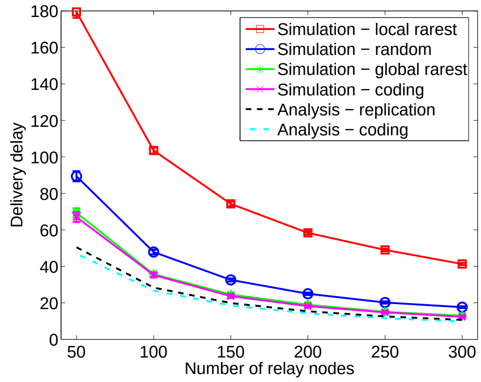

Figure 1: Delivery delay under different numbers of relay nodes.

Furthermore, in practice, this conclusion is also correct because network coding has the same performance as the idealized global rarest policy as shown in Fig. 1. We emphasize Fig. 1 shows that the practical replication based protocols, *i.e.*, random and local rarest policy, both have significantly longer delivery delay than the network coding based protocol.

Finally, we notice that Fig. 1 shows that the delivery delay decreases as the number of relay nodes increases. This is because given the same node meeting rate, more relay nodes can aid more transmissions from the source to the destination.

6.2 The Case for Both Limited Bandwidth and Buffer

We proceed to study the impact of the relay buffer size on the delivery delay. We set the number of relay nodes to 100 in this set of experiments and adjust the relay buffer size from 1 to 10. Fig. 2 shows that our analysis agrees with the simulation result for the network coding based protocol and the replication based protocol with global rarest policy.

In addition, we note that both analytical and simulation result demonstrate the benefit of network coding under limited buffer: the delivery delay of network coding based protocol is not influenced by the buffer size, whereas the delivery delay of the replication based protocols increases significantly when the buffer size decreases. Such performance degradation of the replication based protocols is due to the coupon collector effect [17]. If we consider the extreme case that each buffer can store only one packet, assuming that the packet in a buffer is uniformly randomly chosen from the K packets, the coupon collector effect dictates that the destination node needs to collect $O(K \ln K)$ packets in order to obtain all K packets. On the other hand, under the same setting, the destination in the network coding based protocol can decode all K source packets from K coded packets with high probability.

Finally, we observe that the delivery delay of the practical replication based protocol with random policy increases much more significantly than the global rarest policy when the buffer size decreases. This is because under the random policy, the packet distribution in node buffers is not the uniform distribution, but a biased distribution. If the node buffer size is K, such bias does not have as much im-

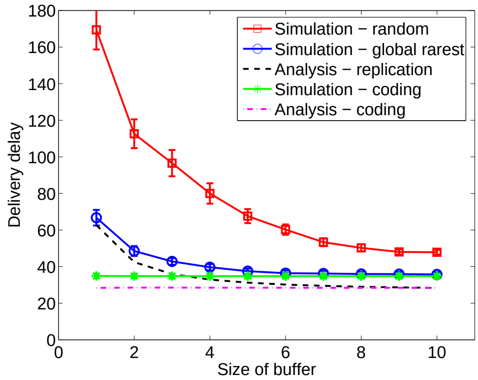

Figure 2: Delivery delay under different buffer sizes.

pact after most nodes collect all packets. However, if the buffer size is very small, such bias has influence throughout the delivery process and degrades the protocol performance.

7. PRIORITY CODING PROTOCOL

As described in Sec. 1, the destination has to wait for a sufficient number of coded packets before decoding any useful data in the network coding based protocol, despite its superiority over practical replication based protocols under limited bandwidth and buffer. In this section, we first introduce a simple priority coding protocol such that a subset of data, *i.e.*, the high priority data, can be decoded much earlier than the time to decode all data. We then use our analytical framework to study the trade-off in designing such a protocol.

7.1 Protocol and Analysis

We assume the K packets in the source can be classified into M different priority levels in descending levels of urgency — the packets in the ith level are more preferable and are decoded before the packets in the jth level, if $i < j$. The number of packets in the ith level is denoted by K_i, where $1 \leq i \leq M$. We further assume through layered coding [19] or particular application semantics, the number of packets K_i in each level can be adjusted to improve the utility of the application under our priority coding protocol. To make the analysis independent of application details, we assume the sum of the number of packets in all priority levels keeps constant after adjusting K_i.

Next, we describe our priority protocol. *First*, the source transmits the data in the 1st level through the network using the network coding based protocol as described in Sec. 4.1. *Second*, after the destination decodes all data in the 1st level, the destination propagates an ACK towards the source by replicating the ACK whenever two nodes meet. *Third*, upon receiving the ACK, the source starts to transmit the data in the 2nd level with the same protocol as used in transmitting the data in the 1st level. Since the data in the 1st level has arrived at the destination, a node drops the data in the 1st level whenever the buffer is full and new data in the 2nd level arrive. *Finally*, such process continues until the destination decodes the data in all priority levels.

We proceed to investigate the effectiveness and overhead of the above priority coding protocol by our analytical framework proposed in Sec. 4.2. It is easy to see in the priority protocol, the transmission process of the data in a priority level is identical to the network coding based protocol described in Sec. 4.1. Therefore, we can use our analytical framework to compute the expected delivery delay of the data within any priority level. In particular, the delivery delay distribution of K_i packets, $F_{K_i}(t)$, can be computed with (5) by replacing K with K_i, and the expected delay $E[T_i]$ for the data in the ith level can be computed from $F_{K_i}(t)$. Next, we notice that the expected delay $E[T_{ACK}]$ in transmitting an ACK is equivalent to transmiting a packet, under the condition of infinite bandwidth, infinite buffer, and the replication based epidemic routing, which has been derived in [26]. Hence, we have

$$E[T_{ACK}] = \ln(N+1)/(\lambda N), \tag{10}$$

where N is the total number of relay nodes, and λ is the inter meeting rate of any pair of nodes in the network. Because the delivery process is composed of the data transmissions for M priority levels and the $M-1$ ACK transmissions interleaved among them, we can compute the total expected delay $E[T]$ to deliver all data as follows:

$$E[T] = \sum_{i=1}^{M} E[T_i] + (M-1)E[T_{ACK}], \tag{11}$$

where $E[T_{ACK}]$ is given in (10), and $E[T_i]$ is given in (5) by replacing K with K_i.

7.2 Priority Coding Advantage

In the following, we conduct numerical analysis on the performance of the above priority protocol. We study the simplest case, where only two priority levels exist. We set the total number of relay nodes N to 200, *i.e.*, the total number of nodes is 202, including the source and the destination. We further set the total number of packets to be transmitted to 100. We perform a set of numerical analysis by adjusting the number of packets in the high priority level from 1 to 99, and compare the delivery delay of the priority coding protocol with the original network coding based protocol, where all 100 packets are sent through the network in one priority level altogether. In all experiments, we set the relay buffer size to 10, and the node inter meeting rate λ to 0.005.

Fig. 3 shows that our protocol is effective. For example, if the high priority level has 10 packets, the network delivers them with delay 14.9473, which is much smaller than the total data delivery delay, 104.3826, in the original network coding based protocol. Furthermore, the total delivery delay, 114.6457, in the priority coding protocol is only 10.26% longer than the data delivery delay in the original protocol. Hence, our priority coding protocol brings low overhead. We explain the overhead in the priority coding protocol with more details in the following.

The overhead of the priority coding protocol consists of two parts: the ACK propagation delay and the delivery time of the first packet, where the former is obvious, and the latter is explained in the following. After examining Fig. 3 more carefully, we observe that the delivery delay of high priority data is almost in linear relation with the number of packets in the high priority. Such observation shows that the delivery delay in the network coding based protocol is composed by two types of components: the delivery delay of the first packet (5.1928 in Fig. 3) and the delivery delay of

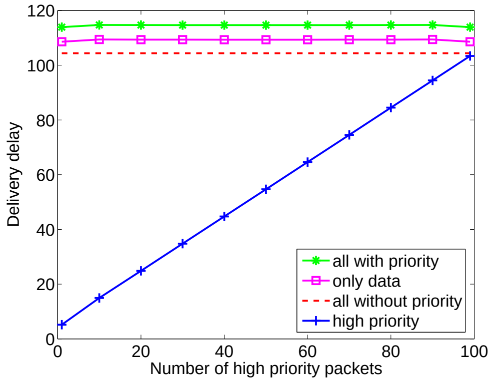

Figure 3: Delivery delay under different numbers of packets in the high priority level. The plot labeled "only data" represents the sum of the delivery delay in two priority levels without the ACK packet.

the remaining packets, where each packet delivery delay is almost identical (about 0.9945) and much shorter than the delay of the first packet. This is because the transmission of the first packet incurs a delay with approximately the length of the shortest opportunistic path. Afterwards, the delivery delay of each packet is around the expected time $E[T_m]$ in which the destination meets another node because the destination can obtain a novel coded packet from each contact with another node, a relay node or the source node, with high probability. We further confirm this by noting that $E[T_m] = \frac{1}{\lambda} \cdot \frac{1}{N+1} = 1/(0.005 * 201) = 0.995$ (agreeing with the value observed in Fig. 3), since $\frac{1}{\lambda}$ is the expected delay that two nodes meet. Because the delivery delay of each packet (excluding the first packet) is identical for both the priority protocol and the original network coding based protocol, it is easy to see that transmitting data in two priority levels separately will induce a delay overhead as the delivery delay of the first packet.

Therefore, the overhead with our priority protocol is low when there are two priority levels, because the ACK propagation delay 5.3033 and the delivery delay of the first packet 5.1928 are much shorter than the delivery delay of the all packets 104.3826. It can be expected that when we increase the number of priority levels, the overhead of our priority protocol increases. The quantitative relation of the protocol overhead and the number of priority levels can be easily estimated by our analytical framework. We omit such analysis in the paper due to space constraint.

8. CONCLUSION

In this paper, we introduce an analytical framework to study the performance of network coding and replication based epidemic routing protocols. Our models capture the dynamics of these protocols on opportunistic networks, and show the superiority of network coding based protocol under limited bandwidth and node buffer. Our analytical models are sufficiently accurate to be used to examine the tradeoff involved in new protocol design. Furthermore, we propose a simple priority coding protocol, which can decode emergent data with much shorter delay than the original network coding based protocol. Through our analytical model, we show that the priority coding protocol is effective and induces only low overhead.

In our future work, we would like to extend our basic analytical model to explore the trade-off between energy and packet delivery delay, using similar energy-saving ideas in non-coding based protocols, *e.g.*, "spray and wait" [22]. Furthermore, we will extend our model to study the protocol performance when multiple flows compete for limited bandwidth and buffer in opportunistic networks. Finally, we would like to investigate the case under more realistic mobility models.

9. REFERENCES

[1] R. Ahlswede, N. Cai, S. R. Li, and R. W. Yeung. Network Information Flow. *IEEE Transactions on Information Theory*, 46(4):1204–1216, July 2000.

[2] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine. MaxProp: Routing for Vehicle-Based Disruption-Tolerant Networks. In *Prof. of IEEE INFOCOM*, 2006.

[3] 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. of IEEE INFOCOM*, Barcelona, Spain, 2006.

[4] L.-J. Chen, C.-H. Yu, T. Sun, Y.-C. Chen, and H. hua Chu. A Hybrid Routing Approach for Opportunistic Networks. In *Proc. of ACM SIGCOMM Workshop on Challenged Networks*, 2006.

[5] P. A. Chou, Y. Wu, and K. Jain. Practical Network Coding. In *Proc. of 41th Annual Allerton Conference on Communication, Control and Computing*, October 2003.

[6] S. Deb, M. Médard, and C. Choute. Algebraic Gossip: A Network Coding Approach to Optimal Multiple Rumor Mongering. *IEEE Transactions on Information Theory*, 52(6):2486–2507, June 2006.

[7] B. Fan, D.-M. Chiu, and J. C. Lui. Stochastic Analysis and File Availability Enhancement for BT-like File Sharing Systems. In *Fourteenth IEEE International Workshop on Quality of Service(IWQoS)*, Yale University, New Haven, CT, USA, 2006.

[8] C. Fragouli, J. Widmer, and J.-Y. L. Boudec. A Network Coding Approach to Energy Efficient Broadcasting: from Theory to Practice. In *Proc. of IEEE INFOCOM*, 2006.

[9] C. Fragouli, J. Widmer, and J.-Y. L. Boudec. On the Benefits of Network Coding for Wireless Applications. In *Second Workshop on Network Coding, Theory, and Applications (NetCod)*, 2006.

[10] R. Groenevelt, P. Nain, and G. Koole. Message delay in mobile ad hoc networks. In *Performance*, October 2005.

[11] Z. J. Haas and T. Small. A New Networking Model for Biological Applications of Ad Hoc Sensor Networks. *IEEE/ACM Transactions on Networking*, 14(1):27–40, February 2006.

[12] T. Ho, R. Koetter, M. Medard, D. R. Karger, and M. Effros. The Benefits of Coding over Routing in a Randomized Setting. In *Proc. of IEEE International Symposium on Information Theory*, 2003.

[13] S. Jain, M. Demmer, R. Patra, and K. Fall. Using Redundancy to Cope with Failures in a Delay

Tolerant Network. In *Proc. of ACM SIGCOMM*, Philadelphia, Pennsylvania, USA, 2005.

[14] S. Jain, K. Fall, and R. Patra. Routing in a Delay Tolerant Network. In *Proc. of ACM SIGCOMM*, 2004.

[15] S. Katti, H. Rahul, W. Hu, D. Katabi, M. Medard, and J. Crowcroft. XORs in The Air: Practical Wireless Network Coding. In *Proc. of ACM SIGCOMM*, 2006.

[16] Y. Lin, B. Liang, and B. Li. Performance Modeling of Network Coding in Epidemic Routing. Technical report, http://iqua.ece.toronto.edu/papers/netcodmodel.pdf, ECE, University of Toronto, 2007.

[17] M. Mitzenmacher and E. Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis*. Cambridge University Press, 2005.

[18] G. Neglia and X. Zhang. Optimal Delay-Power Tradeoff in Sparse Delay Tolerant Networks: a preliminary study. In *Proc. of ACM SIGCOMM Workshop on Challenged Networks*, 2006.

[19] K. Sayood. *Introduction to Data Compression*. Morgan Kaufmann, third edition, 2006.

[20] T. Small and Z. J. Haas. Quality of Service and Capacity in Constrained Intermittent-Connectivity Networks. *to appear in IEEE Transactions on Mobile Computing*.

[21] T. Spyropoulos, K. Psounis, and C. Raghavendra. Efficient Routing in Intermittently Connected Mobile Networks: The Multi-copy Case. *to appear in ACM/IEEE Transaction on Networking*, 2007.

[22] T. Spyropoulos, K. Psounis, and C. Raghavendra. Efficient Routing in Intermittently Connected Mobile Networks: The Single-copy Case. *to appear in ACM/IEEE Transaction on Networking*, 2007.

[23] Y. Wang, S. Jain, M. M. tonosi, and K. Fall. Erasure-Coding Based Routing for Opportunistic Networks. In *Proc. of ACM SIGCOMM Workshop on Delay Tolerant Networking and Related Topics (WDTN)*, Philadelphia, PA, USA, 2005.

[24] J. Widmer and J.-Y. L. Boudec. Network Coding for Efficient Communication in Extreme Networks. In *Proc. of ACM SIGCOMM Workshop on Delay Tolerant Networking and Related Topics (WDTN)*, Philadelphia, PA, USA, 2005.

[25] X. Zhang, G. Neglia, J. Kurose, and D. Towsley. On the Benefits of Random Linear Coding for Unicast Applications in Disruption Tolerant Networks. In *Second Workshop on Network Coding, Theory, and Applications (NetCod)*, 2006.

[26] X. Zhang, G. Neglia, J. Kurose, and D. Towsley. Performance Modeling of Epidemic Routing. *Elsevier Computer Networks journal*, 2007.

[27] W. Zhao, M. Ammar, and E. Zegura. A Message Ferrying Approach for Data Delivery in Sparse Mobile Ad Hoc Networks. In *Proc. of the Fifth ACM International Symposium on Mobile Ad Hoc Networking and Computing(MobiHoc)*, 2004.