

Extreme Value FEC for Reliable Broadcasting in Wireless Networks

Weiyao Xiao* and David Starobinski*

* Department of Electrical and Computer Engineering
 Boston University, Boston, MA 02215
 Email: {weiyao, staro}@bu.edu

Abstract—The advent of practical rateless codes enables implementation of highly efficient packet-level forward error correction (FEC) strategies for reliable data broadcasting in loss-prone wireless networks, such as sensor networks. Yet, the critical question of accurately quantifying the proper amount of redundancy has remained largely unsolved. In this paper, we exploit advances in extreme value theory to rigorously address this problem. Under the asymptotic regime of a large number of receivers, we derive a closed-form expression for the cumulative distribution function (CDF) of the completion time of file distribution. We show the existence of a phase transition associated with this CDF and accurately locate the transition point. We derive tight convergence bounds demonstrating the accuracy of the asymptotic estimate for the practical case of a finite number of receivers. Further, we asymptotically characterize the CDF of the completion time under heterogeneous packet loss, by establishing a close relationship between the data broadcasting and multi-set coupon collector problems. We demonstrate the benefits of our approach through simulation and through real experiments on a testbed of 20 Tmote Sky sensors. Specifically, we augment the existing Rateless Deluge software dissemination protocol with an extreme value FEC strategy. The experimental results reveal reduction by a factor of five in retransmission request messages and by a factor of two in total dissemination time, at the cost of a marginally higher number of data packet transmissions in the order of 5%.

Index Terms—Extreme Value Theory, Coupon Collector’s Problem, Forward Error Correction (FEC), Rateless Coding, Over-the-Air Programming.

I. INTRODUCTION

Reliable data broadcasting for wireless networks is an essential service supporting a plethora of applications, including distribution of text and multimedia contents, podcasting, and over-the-air programming (OAP) [1–6].

The lossy nature of wireless channels significantly complicates the task of reliable broadcasting, however. Due to the potentially large number of wireless devices, the so-called “broadcast storm” [7] phenomenon may arise when multiple receivers contend over a shared channel to request retransmissions of lost packets via either acknowledgements (ACKs) or negative acknowledgements (NACKs) messages. Although some mechanisms exist to mitigate the broadcast storm problem, such as ACK and NACK suppression [8], the impact of this problem can still be considerable [2, 5].

Ideally, instead of relying on receivers to notify a source about missing packets, a procedure commonly referred to as automatic repeat request (ARQ), the source should be able to accurately predict the total number of transmissions required and sends out data without the need for acknowledgements. Packet-level forward error correction (FEC) [9] provides a practical approach towards implementing this idea. With the advent of rateless codes, such as random linear codes, LT, and raptor codes [10, 11], FEC can be implemented in a very efficient fashion, whereas a source continuously encodes new packets based on the M original packets of a given file. The source then sends out the encoded packets and as soon as receiver obtains M (or slightly more) distinct packets, it can reconstruct the entire file successfully.

Although FEC broadcasting has been shown to outperform ARQ in many cases [12, 13], the major issue of quantifying the proper amount of redundancy has remained unsolved to a large extent (cf. Section II for related work). While transmitting too many redundant packets wastes bandwidth and energy, too little redundancy leaves many receivers unable to reconstruct the original file, leading to retransmission requests and eventually the same problems as encountered by ARQ schemes.

In this paper, we exploit advances in extreme value theory (EVT) [14] to rigorously address the problem of quantifying FEC redundancy in lossy wireless broadcast networks. Our main contributions are as follows.

First, under the asymptotic regime of a large number of receivers N , we derive a *closed-form* expression for the cumulative distribution function (CDF) of the completion time (i.e., the total number of packets to be sent by a source to ensure file recovery by all the receivers). Our analysis reveals the existence of a *phase transition* property associated with this CDF. Specifically, we show that there exists a threshold on the number of packets to be sent below which the probability that a file can be recovered by all the nodes in the network is close to zero. However, if the number of packets sent is slightly greater than the threshold, then the probability that every node in the network is able to reconstruct the file quickly approaches one. We accurately locate the threshold value and conduct a sensitivity analysis for the case where the packet loss probability and the number of receivers are imperfectly known. Further, we extend the analysis to the case where complete file reception is only required for $N - K$ out of N receivers, where K is a small, fixed number.

Our second contribution is the derivation of tight *con-*

A preliminary version of this paper appeared in the proceedings of the IEEE INFOCOM 2009 conference.

vergence bounds for a finite number of receivers N . These bounds allow us to estimate the error committed by replacing the exact CDF by its limiting form. They also provide a means to compute the amount of redundancy needed for finite values of N . The bounds reveal that the asymptotic formula is remarkably accurate even for small values of N (e.g., 10).

Third, we analyze the *heterogeneous* packet loss case, whereby different receivers experience different packet loss probabilities. To this effect, we establish a relationship between the data broadcasting problem and the multi-set coupon collector's problem [15]. Exploiting this relationship, we provide asymptotically tight bounds on the CDF of the completion time to successfully disseminate a file to a set of receivers with heterogeneous packet loss probabilities.

Last, we conduct *real experiments* on a testbed of 20 Tmote Sky sensors that illustrate practical use of our theoretical findings. Specifically, we embed our extreme-value FEC strategy into the Rateless Deluge OAP protocol [5] and demonstrate potential for significant reduction in retransmission requests (about 80%) and in completion time (about 50%) at the cost of marginally higher data packet transmissions (less than 5%) with respect to the original protocol.

The rest of this paper is organized as follows. We first discuss related research on FEC data broadcasting in Section II. After reviewing basics of extreme value theory in Section III, we present our network model and problem formalization for homogeneous packet loss in Section IV-A. We conduct an asymptotic analysis of FEC broadcasting as $N \rightarrow \infty$ in Section IV-B, derive convergence bounds for finite N in Section IV-C, and perform a sensitivity analysis in Section IV-D. An asymptotic analysis of the completion time under heterogeneous packet loss is carried out in Section V. We present our simulation results and prototype implementation in Sections VI and VII respectively, and conclude the paper in Section VIII. Due to space limitation, proofs of some theorems are omitted. They can be found in [16].

II. RELATED WORK

Reliable data dissemination is a key enabling technology for wireless sensor networks. It provides fundamental services, such as dissemination of a software program from one source to an entire network [1–6]. Thus, the problem considered in this paper is different from that of data aggregation, where multiple sensors send their data to a sink [17, 18].

The concept of exploiting FEC for reliable multicasting/broadcasting¹ has been the subject of considerable amount of work, both in wireline and wireless settings. We survey here only most closely related work. Rubenstein *et al* [19] propose a multicast protocol that requires a source to forward redundant packets in advance. This protocol is shown to achieve a significant decrease in the expected time for reliable delivery of data. Rizzo *et al* propose RMDP [20], another FEC-based reliable multicast protocol, and show that FEC effectively reduces the amount of acknowledgments. However, the problem of quantifying FEC redundancy remains unsolved.

¹The terms multicasting and broadcasting are used interchangeably in this paper.

Huitema [21] and Nonnenmacher *et al* [22] evaluate the performance improvements achieved with different levels of FEC redundancy via numerical computation. Ghaderi *et al* [13] and Mosko *et al* [23] obtain numerical evaluation of the distribution of the completion time. However, no closed form is provided to relate the redundancy needed with the probability of success. Eryilmaz *et al* [24] provide recursive expression for the *average* completion time and Ghaderi *et al* [13] derive an asymptotic expression for it. However, they do not provide results for the CDF. Furthermore, to our knowledge, our work is the first to demonstrate the phase transition associated with this CDF and to derive bounds on the asymptotic error for the case of a finite number of receivers.

While in practice the packet loss probability differs from node to node due to many factors (i.e., link quality, distance to the source, antenna sensitivity) all the following references [4, 10–13, 19–21] assume homogeneous packet loss rates in their analysis. The work in [22, 23] provide analysis for heterogeneous packet loss probability scenarios. However, unlike our paper, the results are only numerical.

III. BACKGROUND

A. Extreme Value Theory

Let X_1, X_2, \dots, X_N be independent, identically distributed (i.i.d.) random variables. Extreme value theory provides tools for characterizing possible limit distributions of sample maxima of the above i.i.d. random variables. Denote by F the CDF of X and by F^N the CDF of the maximum of X_1, X_2, \dots, X_N . Suppose there exists a sequence of constants a_N and b_N , such that $\frac{\max(X_1, X_2, \dots, X_N) - b_N}{a_N}$ has a nondegenerate limit distribution as $N \rightarrow \infty$, then

$$\lim_{N \rightarrow \infty} F^N(a_N x + b_N) = G(x), \quad (1)$$

or equivalently,

$$\lim_{N \rightarrow \infty} N \bar{F}(a_N x + b_N) = -\log G(x), \quad (2)$$

where $G(x)$ is the CDF of one of the three extreme value distributions, namely Fréchet, Gumbel and Weibull [14, p.9].

For a given random variable, various tests exist to determine its *domain of attraction* (i.e., the corresponding extreme value distribution) and its normalization constants. In our paper, all the distributions of interest belong to the domain of attraction of the Gumbel distribution, i.e.,

$$G(x) = \exp(-e^{-x}) \quad \forall x \in \mathbb{R}. \quad (3)$$

Under mild technical assumptions [14, p.77], the domain of attraction conditions imply also moment convergence. Thus, for distributions belonging to the Gumbel's domain of attraction

$$\lim_{N \rightarrow \infty} \frac{\mathbb{E}[\max(X_1, X_2, \dots, X_N)] - b_N}{a_N} = \gamma, \quad (4)$$

where $\gamma \approx 0.5772$ is the Euler's constant.

The above result can be generalized to obtain characterization of the asymptotic distribution of the K -th largest random variable, where $K \geq 1$ is a fixed number. Arrange the random variables X_i in increasing order as following, $X_{1:N} \leq X_{2:N} \leq \dots \leq X_{N:N}$. Thus, $X_{N:N} = \max_{n=1, \dots, N} X_n$

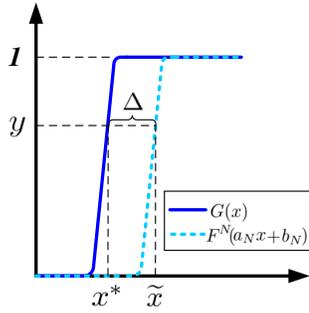


Fig. 1. Illustration of convergence metric.

is the largest random variable, $X_{N-1:N}$ is the second largest random variable, etc. According to Ref. [25, p. 127], if

$$\lim_{N \rightarrow \infty} \Pr\left(\frac{X_{N:N} - b_N}{a_N} \leq x\right) = G(x), \quad (5)$$

then

$$\lim_{N \rightarrow \infty} \Pr\left(\frac{X_{N-K:N} - b_N}{a_N} \leq x\right) = G(x) \sum_{i=0}^{K-1} \frac{1}{i!} \left(\log \frac{1}{G(x)}\right)^i. \quad (6)$$

B. Convergence Metric

As mentioned above, an estimation based on EVT assumes $N \rightarrow \infty$. We provide now a metric to study the quality of convergence when N is finite. Specifically, we fix a value on the y -axis and measure the distance on the x -axis between the points corresponding to the exact distribution F^N and the limit distribution G . Specifically, as shown in Fig. 1, let

$$\begin{cases} G(x^*) = y, \\ F^N(a_N \tilde{x} + b_N) = y, \end{cases} \quad (7)$$

then the convergence metric is set as follows

$$\Delta = |x^* - \tilde{x}|. \quad (8)$$

In the following section, we will derive a bound on Δ that applies uniformly to an entire interval $[y_l, y_h]$, where $0 \leq y_l \leq y_h < 1$. If the desired completion probability y is known in advance, then the values of y_l and y_h can simply be set to y leading to a tighter bound on Δ . Otherwise, one can select a larger interval and the bound will apply to all values of y belonging to that interval.

IV. THE HOMOGENEOUS CASE: LIMIT DISTRIBUTION AND CONVERGENCE BOUNDS

A. Model and Problem Formulation

We consider the problem of broadcasting a file consisting of M packets from a source (e.g., a base station) to N nodes within its transmission range. The time axis is slotted and each packet transmission takes one time slot. In this section, we assume that each node experiences the same packet loss probability p , independent of any other events.

We assume that FEC is implemented using a perfect rateless code, i.e., each node needs to correctly receive M distinct packets to recover a file. Thus, a source transmits new packets

until all the nodes received M different packets. If slightly more packets are needed (say M') because of the imperfection of codes, then one just need to replace M by M' in the following analysis.

Denote by T the random variable representing the completion time, i.e. the number of time slots, needed to disseminate M packets to a cluster of N nodes. Our goal is to characterize the CDF of T , namely $\Pr\{T \leq t\}$, with which one can determine the number of redundant packets needed in FEC. Towards this end, we will use EVT to characterize the limiting form of the CDF of T when $N \rightarrow \infty$ and then derive bounds on the error Δ for finite values of N .

In this paper, we do not enter into the details of how to estimate the network parameters N and p . We refer the interested reader to [5, 26] for possible approaches. Nevertheless, in Section IV-D, we will study robustness of our FEC prediction model to inaccurate estimation of these parameters.

B. Asymptotic Analysis of Completion Time

Denote by T_n^m the number of slots needed for node n to receive its m -th packet, $1 \leq m \leq M$. Clearly, T_n^m follows a geometric distribution with mean $1/(1-p)$, i.e., $\Pr\{T_n^m = i\} = p^{i-1}(1-p)$.

Thus, the time T_n needed for node n to receive M different packets is the sum of M i.i.d geometric random variables with mean $1/(1-p)$, i.e., $T_n = \sum_{m=1}^M T_n^m$ and T_n is said to follow a *negative binomial* or *Pascal* distribution [27, p.164]. Due to the broadcast nature of the channel, the completion time for broadcasting a file to all the nodes is the maximum of N negative binomial random variables, i.e., $T = \max(T_1, T_2, \dots, T_N)$. The following theorem tightly bounds the distribution of T as $N \rightarrow \infty$. Before proceeding, we recall the definition of stochastic ordering [28, p.404].

Definition 1: A random variable X is stochastically larger than a random variable Y , denoted $X \geq_{st} Y$, if

$$\Pr(X > a) \geq \Pr(Y > a), \quad \text{for all } a. \quad (9)$$

Theorem 1: The completion time T to disseminate M packets to N nodes using FEC broadcasting is bounded by random variables belonging to Gumbel's domain of attraction. Namely, there exist T_l and $T_u = T_l + 1$, satisfying

$$T_l \leq_{st} T \leq_{st} T_u, \quad (10)$$

$$\lim_{N \rightarrow \infty} \Pr\left\{\frac{T_l - b_N}{a_N} \leq x\right\} = G(x), \quad (11)$$

$$\text{where, } a_N = 1/\log\left(\frac{1}{p}\right), \quad (12)$$

$$\begin{aligned} b_N &= \log_{\frac{1}{p}}(N) + (M-1) \log_{\frac{1}{p}}(\tau) \\ &\quad + (M-1) \log_{\frac{1}{p}}\left(\frac{1-p}{p}\right) - \log_{\frac{1}{p}}(M-1)!, \end{aligned} \quad (13)$$

$$\tau = \log_{\frac{1}{p}}(N) + (M-1) \left(\log_{\frac{1}{p}}\left(\frac{1-p}{p}\right)\right). \quad (14)$$

Proof: Let $D(t) = \Pr\{T_n \leq t\}$. Since T_n follows a negative binomial distribution, $D(t) = I(1-p; M, t-M+1)$ [27, p.

164], [29, p. 519], where t ($t \geq M$) is an integer and $I(z; a, b)$ is the regularized beta function, defined as following [29, p. 516],

$$I(z; a, b) = 1 - \frac{(1-z)^b}{B(a, b)} \sum_{i=0}^{a-1} (-1)^i \binom{a-1}{i} \frac{(1-z)^i}{b+i}, \quad (15)$$

where a and b are integers, and $B(a, b)$ is the complete beta function [29, p. 594, 597]

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad (16)$$

where Γ is gamma function, i.e., $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$, and $\Gamma(a) = (a-1)\Gamma(a-1)$ for $a > 0$ [27, p. 66].

Create a continuous R.V. T_u^n with CDF $F(x) = I(1-p; M, x-M+1)$, where $x > M-1$. Let $T_l^n = T_u^n - 1$. The probability distribution function for T_l^n is thus $F(x+1)$. According to Eq. (15) and Eq. (16), we have,

$$\bar{F}(x) = 1 - I(1-p; M, x-M+1) \quad (17)$$

$$= \frac{p^{x-M+1}}{B(M, x-M+1)} \sum_{i=0}^{M-1} (-1)^i \binom{M-1}{i} \frac{p^i}{x-M+1+i} \quad (18)$$

$$= \frac{\prod_{j=0}^{M-1} (x-j)}{(M-1)!} p^{x-M+1} \sum_{i=0}^{M-1} \binom{M-1}{i} \frac{(-1)^i p^i}{x-M+1+i}. \quad (19)$$

Let $\bar{D}(x) = 1 - D(x)$. Since $I(x)$ is an increasing function of x , we have

$$\bar{F}(x+1) \leq \bar{D}(\lceil x \rceil) \leq \bar{F}(x). \quad (20)$$

From Eq. (20), we have $T_l^n \leq_{st} T \leq_{st} T_u^n$ according to Definition 1. Let $T_u = \max_{n=1..N} T_u^n$, and $T_l = \max_{n=1..N} T_l^n = T_u - 1$. Since the max operation conserves stochastic ordering, we have $T_l \leq_{st} T \leq_{st} T_u$.

Inserting Eq. (12) and Eq. (13) into Eq. (19) yields

$$\begin{aligned} & \lim_{N \rightarrow \infty} N \bar{F}(a_N x + b_N) \\ &= \lim_{N \rightarrow \infty} N \frac{\prod_{j=0}^{M-1} (a_N x + b_N - j)}{(M-1)!} p^{a_N x + b_N} p^{-(M-1)} \\ & \quad \times \sum_{i=0}^{M-1} \binom{M-1}{i} \frac{(-1)^i p^i}{a_N x + b_N - M + 1 + i}. \end{aligned} \quad (21)$$

According to Eq. (12) and Eq. (13), we have

$$\begin{aligned} p^{a_N x + b_N} &= p^{a_N x} \cdot p^{b_N} \\ &= e^{-x} \frac{1}{N} \left(\frac{1}{\tau}\right)^{M-1} \left(\frac{p}{1-p}\right)^{M-1} (M-1)!. \end{aligned} \quad (22)$$

Therefore, from Eq. (21) and Eq. (23), we obtain

$$\begin{aligned} & \lim_{N \rightarrow \infty} N \bar{F}(a_N x + b_N) \\ &= \lim_{N \rightarrow \infty} \prod_{j=0}^{M-1} (a_N x + b_N - j) e^{-x} \left(\frac{1}{\tau}\right)^{M-1} \left(\frac{1}{1-p}\right)^{M-1} \\ & \quad \times \sum_{i=0}^{M-1} \binom{M-1}{i} \frac{(-1)^i p^i}{a_N x + b_N - M + 1 + i}. \end{aligned} \quad (24)$$

Noting that $b_N \rightarrow \infty$ as $N \rightarrow \infty$, we get

$$\begin{aligned} & \lim_{N \rightarrow \infty} \sum_{i=0}^{M-1} \binom{M-1}{i} \frac{(-1)^i p^i}{a_N x + b_N - M + 1 + i} \\ &= \lim_{N \rightarrow \infty} \sum_{i=0}^{M-1} \binom{M-1}{i} \frac{(-1)^i p^i}{a_N x + b_N} \end{aligned} \quad (25)$$

$$= \lim_{N \rightarrow \infty} \frac{1}{a_N x + b_N} \sum_{i=0}^{M-1} \binom{M-1}{i} (-1)^i p^i \quad (26)$$

$$= \lim_{N \rightarrow \infty} \frac{1}{a_N x + b_N} (1-p)^{M-1}. \quad (27)$$

The last step of the above equations follows from the binomial theorem. Inserting Eq. (27) into Eq. (24), yields

$$\begin{aligned} & \lim_{N \rightarrow \infty} N \bar{F}(a_N x + b_N) \\ &= \lim_{N \rightarrow \infty} \frac{\prod_{j=0}^{M-1} (a_N x + b_N - j)}{\tau^{M-1} (a_N x + b_N)} e^{-x}. \end{aligned} \quad (28)$$

Consider the expression

$$\frac{\prod_{j=0}^{M-1} (a_N x + b_N - j)}{\tau^{M-1} (a_N x + b_N)}. \quad (29)$$

According to Eq. (12), Eq. (13) and Eq. (14), the dominating component of both the numerator and denominator of that expression is $(\log_{\frac{1}{p}}(N))^M$.

Accordingly, from Eq. (28), we obtain

$$\lim_{N \rightarrow \infty} N \bar{F}(a_N x + b_N) = e^{-x} = -\log G(x). \quad (30)$$

Thus, F is in the domain of attraction of G with normalizing constants a_N and b_N [14], namely,

$$\lim_{N \rightarrow \infty} \Pr\left\{\frac{T_l - b_N}{a_N} \leq x\right\} = G(x). \quad (31)$$

Theorem 1 shows that as $N \rightarrow \infty$, the CDF of the completion time converges to a scaled and shifted Gumbel distribution, namely,

$$\Pr\{T \leq t\} \approx G\left(\frac{t - b_N}{a_N}\right). \quad (32)$$

Since the packet loss probability p is usually small, a_N is small as well and the completion time distribution has a sharp phase transition around the point b_N . This will be verified by our numerical results in Section VI.

As a corollary from the theorem, we can also asymptotically characterize the distribution of the completion time when allowing up to K receivers not to recover the entire file. We denote the corresponding random variable $T_{N-K:N}$. From Theorem 1 and Eq. (6), we have

$$\Pr(T_{N-K:N} \leq t) \approx G\left(\frac{t - b_N}{a_N}\right) \sum_{i=0}^K \frac{e^{-\frac{i(t-b_N)}{a_N}}}{i!}. \quad (33)$$

Such a characterization sheds light on the trade-off between the stringency of the requirement for file completion and the amount of FEC redundancy. For instance, according to Eq. (32) and Eq. (33), if the source transmits $t = a_N + b_N$

packets, the probability that at least $N - 1$ nodes receive the file is about 37% larger than the probability that all N nodes receive the file. We will investigate this trade-off further in the numerical results section.

Another corollary from the theorem is that the performance of rateless coding on a single channel is identical to that of plaintext coding over an unlimited number of channels [4].

C. Convergence Bounds

Theorem 1 characterizes the limiting form of the CDF of T as $N \rightarrow \infty$. The following theorem bounds the asymptotic error using the convergence criterion defined in Eq. (8), that is, it bounds the distance between the asymptotic estimate \tilde{x} and the exact value x^* .

Theorem 2: The distance $\Delta = |\tilde{x} - x^*|$ between the exact distribution $\Pr\{T \leq a_N \tilde{x} + b_N\} = y$ and the Gumbel distribution $G(x^*) = y$ is bounded as follows for all probability values y belonging to the interval $[y_l, y_h]$:

$$\Delta \leq |\Delta_l| + |\Delta_h + \log(1 + \frac{1}{N} \log \frac{1}{y})|, \quad (34)$$

$$\text{where } \Delta_l = \log p - \frac{(M-1)^2}{2(a_N G^{-1}(y_l) + b_N) - M + 3} + (M-1) \left(1 - \frac{\tau}{a_N G^{-1}(y_l) + b_N + 1}\right), \quad (35)$$

$$\Delta_h = (M-1) \log\left(\frac{1}{1-p}\right) + (M-1) \frac{a_N G^{-1}(y_h) + b_N + 1 - \tau}{\tau}, \quad (36)$$

and G^{-1} is the inverse function of G .

Theorem 2 provides a means to conservatively implement FEC for finite values of N , that is, if one wants to guarantee a completion probability y , then the source should transmit at least $\lceil a_N(x^* + \Delta) + b_N \rceil$ packets. The bound provided by Eq. (34) also exhibits the desirable property of becoming tighter as the number of recipient nodes N increases and as the completion probability y approaches 1.

D. Sensitivity Analysis

In practice, one seldom has perfect information on the network parameters p and N . We are going next to analyze the effect of imperfect estimation of these parameters on the computation of the FEC redundancy.

Denote by \hat{N} and \hat{p} the estimations of N and p , respectively. Without loss of generality, we can write $\hat{N} = (1 + \epsilon_N)N$ and $\hat{p} = (1 + \epsilon_p)p$. A positive value of ϵ_N means that N is overestimated, while a negative value of ϵ_N means that N is underestimated. A similar relation holds between ϵ_p and p .

Denote by T the number of packet transmissions to achieve a completion probability y , as determined by Theorem 1 using the parameters N and p . Correspondingly, \hat{T}_N and \hat{T}_p denote the number of packet transmissions to achieve a completion probability y , as determined by Theorem 1 using \hat{N} and p in the first case and N and \hat{p} in the second case.

When p and M are fixed, T roughly increases logarithmically with N . Thus,

$$\hat{T}_N \approx T + \log_{\frac{1}{p}}(1 + \epsilon_N). \quad (37)$$

When N and M are fixed, T is approximately a linear function of $1/\log \frac{1}{p}$. Therefore,

$$\frac{\hat{T}_p}{T} \approx \frac{\log \frac{1}{p}}{\log \frac{1}{(1+\epsilon_p)p}} = \frac{\log \frac{1}{p}}{\log \frac{1}{1+\epsilon_p} + \log \frac{1}{p}} = \frac{1}{1 - \frac{\log(1+\epsilon_p)}{\log \frac{1}{p}}}. \quad (38)$$

For a small value of ϵ_p , by Taylor expansion, we have

$$\log(1 + \epsilon_p) \approx \epsilon_p, \quad (39)$$

and

$$\frac{1}{1 - \frac{\epsilon_p}{\log \frac{1}{p}}} \approx 1 + \frac{\epsilon_p}{\log \frac{1}{p}}. \quad (40)$$

From Eq. (38), Eq. (39) and Eq. (40), we get

$$\frac{\hat{T}_p}{T} \approx 1 + \frac{\epsilon_p}{\log \frac{1}{p}}. \quad (41)$$

Eq. (37) and Eq. (41) can be used to conservatively calibrate the amount of FEC redundancy. They show that the completion time is more sensitive to the receiver packet loss probability p than to the number of receivers N . For example, when $p = 5\%$, a 20% overestimate of N , namely $\epsilon_N = 20\%$, will result in sending only 0.06 extra packets on average. A 20% overestimate of p will require the transmission of 6% more packets overall, which is still reasonable.

V. THE HETEROGENEOUS CASE: LIMIT OF DISTRIBUTION AND EXPECTATION

In this section we relax the assumption of homogeneous packet loss probabilities. We consider a model whereby receivers are deployed uniformly at random within a disk of radius of R , with the source at the origin. The signal quality is discretized into L levels based on the distance from the source. Denote by $\vec{\alpha} = [\alpha_1, \alpha_2, \dots, \alpha_L]$ the distance vector (normalized by R), where $0 < \alpha_1 < \dots < \alpha_L = 1$. Next, let $\vec{\omega}_\alpha = [\omega_{\alpha_1}, \omega_{\alpha_2}, \dots, \omega_{\alpha_L}]$ be the corresponding packet loss vector, where $0 < \omega_{\alpha_1} < \dots < \omega_{\alpha_L} < 1$. The packet loss probability for a node is ω_{α_l} if its distance from the source is between $\alpha_{l-1}R$ and $\alpha_l R$ ($\alpha_0 = 0$ by definition). This radio model, which captures spatial correlation of the packet loss, is illustrated in Fig 2. Then, the CDF of the packet loss probability for node n is a multi-step function defined as follows

$$\Pr(p_n \leq x) = \begin{cases} 0 & 0 < x < \omega_{\alpha_1}, \\ \alpha_l^2 & \omega_{\alpha_l} \leq x < \omega_{\alpha_{l+1}}, \quad l = 1, \dots, L-1 \\ 1 & x \geq \omega_{\alpha_L}. \end{cases} \quad (42)$$

Note that the radio model of the previous section is a special case of this model, by setting $L = 1$ and $\omega_{\alpha_L} = p$.

In the remainder of this section, we first establish a relation between the FEC data broadcasting problem in wireless networks and the multi-set coupon collector's problem. This

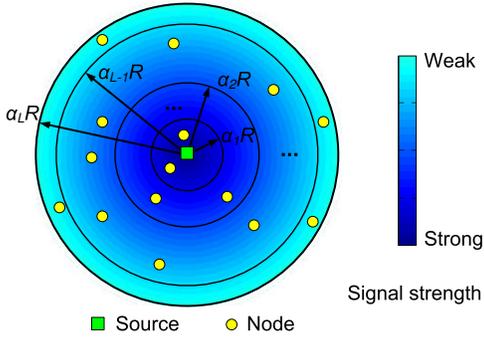


Fig. 2. Radio model for heterogeneous packet loss.

connection enables us in the second part of the section to leverage recent analytical results on the asymptotic behavior of heterogeneous coupon collector systems [15] to analyze FEC data broadcasting with heterogeneous packet loss probabilities.

A. Relation between Coupon Collector and Data Broadcasting Problems

In the multi-set coupon collector's problem, a shopper tries to collect M complete sets of N different coupons in several attempts. Coupon n is associated with a value $q_n > 0$. Each attempt provides the collector with a coupon n with probability $q_n / \sum_{i=1}^N q_i$. Assume that there is unlimited supply of coupons of each kind. Let $\eta_{\{q_n\}}$ be the number of attempts the collector needs to make in order to obtain M complete sets of N coupons, where $\{q_n\}$ represents the set of coupon's values. Asymptotic limits of the CDF of $\eta_{\{q_n\}}$ for large values of N are studied in [15].

Back to our original problem, let $\{p_n\}$ be a set of packet loss probabilities associated with each receiver and $T_{\{p_n\}}$ be the time to transmit M packets to N users with packet loss probabilities p_1, p_2, \dots, p_N using FEC data broadcasting. The following Theorem establishes a relation between the CDF of $T_{\{p_n\}}$ and that of $\eta_{\{q_n\}}$.

Theorem 3: Suppose the packet loss probability at each receiver n , $n = 1, \dots, N$, is an i.i.d. random variable p_n and let $q_n = \log(\frac{1}{p_n})$. Then, as $N \rightarrow \infty$,

$$\Pr\left(\frac{\eta_{\{q_n\}}}{N\mu} + M \leq x\right) \leq \Pr(T_{\{p_n\}} \leq x) \leq \Pr\left(\frac{\eta_{\{q_n\}}}{N\mu} \leq x\right), \quad (43)$$

where μ is the mean of the random variable $\log(\frac{1}{p_n})$.

B. Asymptotic Completion Time

We next derive a closed-form expression for the distribution and the expectation of the completion time as $N \rightarrow \infty$, for the heterogeneous radio model presented at the beginning of the section.

Theorem 4: If the CDF of the packet loss probability of each node n , $n = 1, \dots, N$, satisfies Eq. (42), then, as $N \rightarrow \infty$

$$G\left(x - \frac{M}{a_N}\right) \leq \Pr\left(\frac{T_{\{p_n\}} - b_N}{a_N} \leq x\right) \leq G(x), \quad (44)$$

$$\mathbb{E}[T_{\{p_n\}}] = b_N + \Theta(1), \quad (45)$$

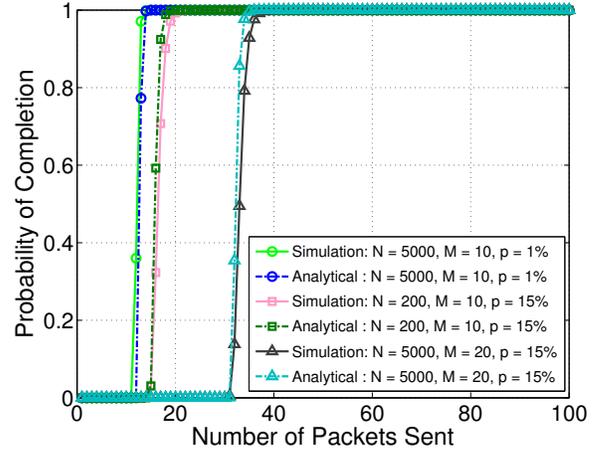


Fig. 3. Accuracy of asymptotic estimate and phase transition demonstration.

where

$$p' = \omega_{\alpha_L}, \quad N' = (\alpha_L^2 - \alpha_{L-1}^2)N, \quad (46)$$

$$a_N = 1 / \log\left(\frac{1}{p'}\right), \quad (47)$$

$$b_N = \log_{\frac{1}{p'}}(N') + (M-1) \log_{\frac{1}{p'}} \tau + (M-1) \log_{\frac{1}{p'}} \frac{1-p'}{p'} - \log_{\frac{1}{p'}}(M-1)!, \quad (48)$$

$$\tau = \log_{\frac{1}{p'}}(N') + (M-1) \left(\log_{\frac{1}{p'}} \frac{1-p'}{p'} \right). \quad (49)$$

This theorem provides the following insight. On average the number of nodes staying within the furthest ring in the disk is $(\alpha_L^2 - \alpha_{L-1}^2)N$. By comparing Eq. (47), Eq. (48) and Eq. (49) with Eq. (12), Eq. (13) and Eq. (14), respectively, we can see that $\bar{T}(\{p_i\})$ is asymptotically identical to the expected time needed to disseminate M packets to $(\alpha_L^2 - \alpha_{L-1}^2)N$ nodes with homogeneous packet loss probability ω_{α_L} . Hence, as $N \rightarrow \infty$, the time needed to disseminate packets to the nodes with the highest packet loss probability dominates.

VI. NUMERICAL RESULTS

In this section, we illustrate the major analytical findings of this paper, namely, (i) the accuracy of the asymptotic estimate of the CDF of the completion time provided by Theorem 1 and the phase transition behavior of this CDF, (ii) the tightness of the upper and lower bounds derived along the proof of Theorem 2, (iii) the tradeoff between redundancy and completion requirement, and (iv) the accuracy of the asymptotic limit on the expected completed time provided by Theorem 4. All the simulation plots are obtained by averaging results over 10000 simulations with identical parameters, but different random seed.

A. Accuracy of Asymptotic Estimate

Fig. 3 compares the CDF estimated by Theorem 1, with the CDF obtained from simulation for various parameters M , N , and p . It is evident from the figure that the limit form provides an accurate estimate of the actual distribution. It is

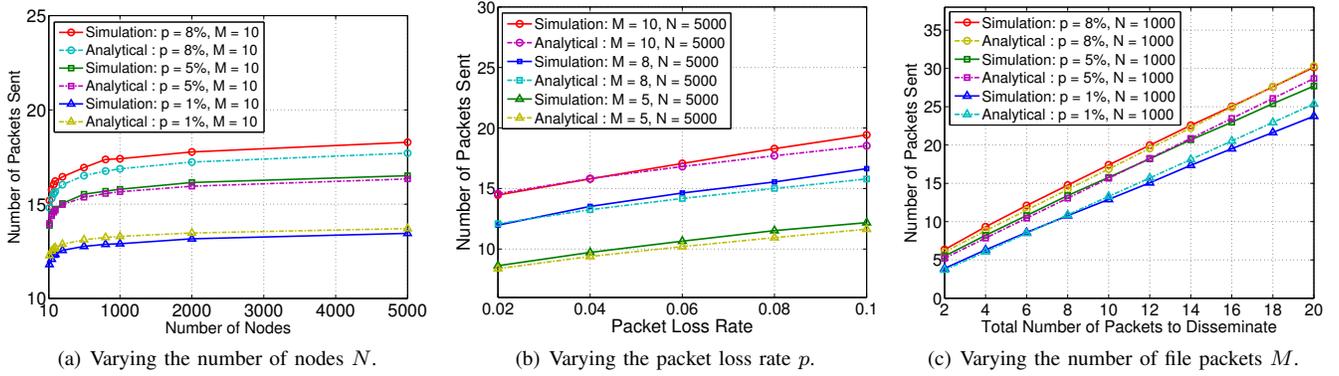


Fig. 4. Number of packets needed to be sent to guarantee completion with probability 99%: Varying different parameters.

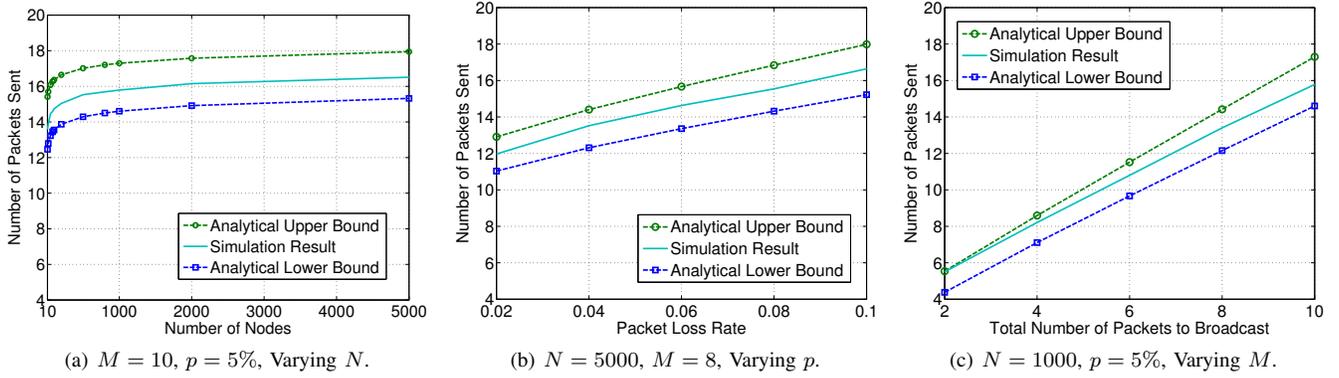


Fig. 5. Number of packets needed to guarantee completion with probability 99%, comparison of simulation and analytical bounds.

interesting to note that even with a large number of receivers and relatively high loss packet probability, we do not need a large number of redundant packets to ensure file reception (with high probability) by all the nodes.

Fig. 3 also clearly demonstrates the phase transition behavior of the CDF. As expected, the CDF shifts to the right as the number of nodes N or the number of file packets M increases, but the sharpness phase transition is not much affected. On the other hand, the packet loss rate p has an effect on both translating and scaling the CDF. A smaller value of p shifts the CDF to the left and also results in a sharper transition.

As discussed in Section IV-B, the phase transition occurs around the point b_N . Using Eq. (13), one can compute the values of b_N for the three cases shown in Fig. 3, which are found to be 12.709, 15.659, 32.020. These values accurately locate the phase transition points.

According to Theorem 1, as $N \rightarrow \infty$, the CDF converges to a Gumbel distribution scaled by $\frac{1}{a_N}$ and translated to the right by $\frac{b_N}{a_N}$. To verify this finding, we closely examine each parameter by fixing the other two. Results are shown in Fig. 4(a), 4(b) and 4(c). We study the phase transition shift of the CDFs as the parameters change by evaluating the case where the completion probability is 99%.

Fig. 4(a) shows the number of packets needed as N increases. As predicted by Eq. (37), when M and p is fixed, the number of packets needed increases logarithmically with N . This is true even when the number of nodes is small, e.g. $N = 10, 20, 50$. This result explains why the redundancy needed is relatively small, even for large values of N .

Fig. 4(b) demonstrates the case where N and M are fixed. From Eq. (41), the number of packets is approximately a linear function of p , which coincides with Fig. 4(b).

Fig. 4(c) shows that, fixing N and p , the phase transition shifts to the right linearly as M increases. This is because the CDF shifts by $\frac{b_N}{a_N}$, which is roughly a linear function of M .

B. Tightness of Bounds

We next compare the analytical upper and lower bounds with simulation results. Each parameter (i.e., N, M, p) is investigated by fixing the other two. Simulation results are compared with analytical bounds, shown in Fig 5(a), Fig 5(b) and Fig 5(c). The analytical bounds are obtained by Theorem 2. We fix $y_l = y_h = 99\%$, that is, a 99% probability of completion. As expected, the curve representing simulation result lies between the analytical lower bound and upper bound. More importantly, the gap between the upper bound and the simulation result is reasonably small for a variety of different parameters. If one is to use the analytical upper bound to estimate the amount of redundant packets, then only one or two more packets than necessary would be transmitted.

C. Tradeoff between Redundancy and Completion Requirement

We next characterize the benefit of loosening the requirement of file recovery by all nodes. Specifically, we allow incomplete file reception at up to K nodes. The analytical estimate in that case is obtained from Eq. (33). Fig. 6 shows

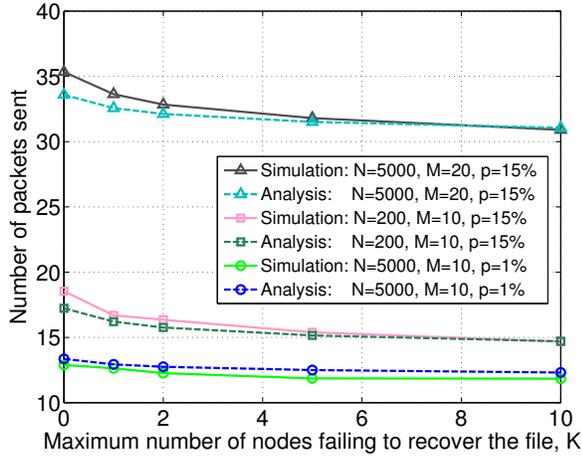


Fig. 6. Tradeoff between redundancy and number of incomplete file reception allowed.

	$\vec{\alpha}$	$\vec{\omega}_{\alpha}$	N	M
Case 1	[0.5, 1]	[0.1, 0.11]	50	20
Case 2	[0.5, 1]	[0.1, 0.15]	50	20
Case 3	[0.9, 1]	[0.0498, 0.1353]	50	20
Case 4	[0.2, 0.4, 0.6, 0.8, 1]	[0.001, 0.008, 0.027, 0.064, 0.125]	50	20
Case 5	[0.2, 0.4, 0.6, 0.8, 1]	[0.001, 0.008, 0.027, 0.064, 0.125]	1000	20

Fig. 7. Different network settings for heterogeneous packet loss simulation.

the number of transmissions required as a function of K for various cases. We set the completion probability to 95%. We note a good match between the analytical estimates and the simulation results. We also observe that sacrificing one or two nodes can significantly reduce FEC redundancy. For example, for the case $N = 5000$, $M = 20$ and $p = 15\%$, allowing incomplete file recovery by one node out of 5000 leads to a 10% reduction in the redundancy amount. However, the marginal gain becomes less significant as K increases.

D. Heterogeneous Packet Loss

In this part we verify the result of Theorem 4, which states that as $N \rightarrow \infty$, the time needed to disseminate packets to the nodes with the highest packet loss probability dominates. We investigate a variety of network settings as shown in the table of Fig. 7. We compare three results: (i) the expected time to disseminate M packets to N nodes under heterogeneous packet loss described by distance vector $\vec{\alpha}$ and packet loss vector $\vec{\omega}_{\alpha}$, based on simulation; (ii) the expected time to disseminate M packets to $(\alpha_L^2 - \alpha_{L-1}^2)N$ nodes all having the same packet loss probability ω_{α_L} , based on simulation; and (iii) the analytical estimate based on Theorem 4.

The results are shown in 8. In cases 1 through 3 we assume, $L = 2$ levels of signal quality: nodes within distance $\alpha_1 R$ of the source have lower packet loss probability ω_{α_1} . Nodes beyond this radius have higher packet loss probability ω_{α_2} . The results show that the presence of nodes with lower packet loss rates has little impact on the completion time for the entire network. This is true even when only a small fraction of nodes suffers from higher packet loss rates. For example, in case 3,

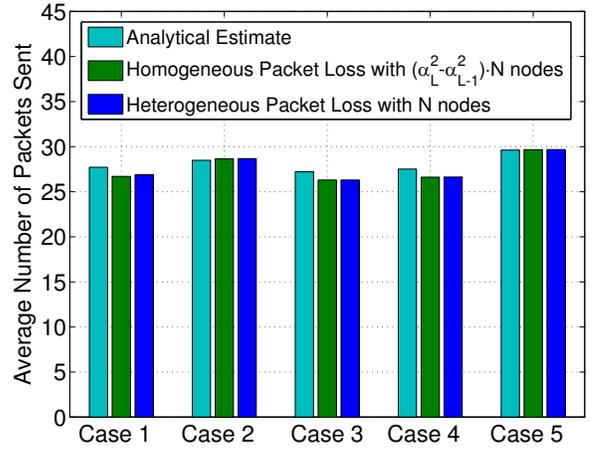


Fig. 8. Average completion time for scenarios with heterogeneous packet loss vs. homogeneous packet loss.

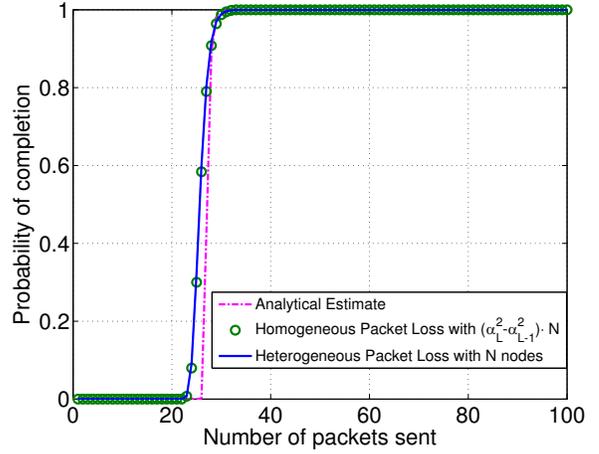


Fig. 9. CDF of completion time heterogeneous packet loss: parameters as described in Case 3 of Fig 7.

on average only $19\% \cdot N$ of nodes have higher packet loss probability e^{-2} , and yet distributing a file only to these nodes takes as much as the time to distribute a file to a network where $81\% \cdot N$ nodes have packet loss probability e^{-3} and $19\% \cdot N$ nodes have packet loss probability e^{-2} . This fact can also be observed from Fig. 9 which depicts the CDF. Cases 4 and 5, in which the signal quality is discretized into $L = 5$ levels, reveal similar behavior. In all cases, Theorem 4 predicts well the simulation results.

VII. PROTOTYPE IMPLEMENTATION

A. Set-up

In this section, we describe practical implementation of extreme value FEC into the Rateless Deluge over-the-air programming protocol [5]. This protocol uses random linear codes for data encoding and enables efficient distribution of a new file program to all the nodes of a sensor network.

The default setting of Rateless Deluge is as follows. A file is divided into pages and each page consists of 20 packets, where each packet contains 23 bytes of data. A sensor sends out a request if it discovers its neighbors have new data. The

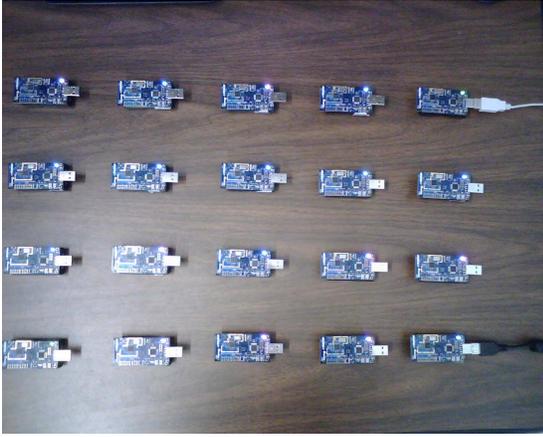


Fig. 10. Experimental testbed with 20 Tmotes.

request message specifies the page number and the number of packets it needs. When a sensor receives enough number of packets (in our case 20), it can decode the page successfully. As in the original Deluge protocol [1], a sensor suppresses its request if it overhears similar requests sent recently.

Here, we augment the original Rateless Deluge with extreme value FEC, and refer to the new protocol as Extreme Value FEC Deluge. Extreme Value FEC Deluge operates the same as Rateless Deluge except that when receiving a request for a new page, the base station broadcasts a redundant amount of packets. The redundancy is set to guarantee with high enough probability that all the receivers recovered the file. In our case, we set the desired completion probability to be 97%, and the redundancy is then computed using Theorem 1.

The performance of Rateless Deluge and Extreme Value FEC Deluge are evaluated on a testbed consisting of 20 Tmote Sky sensors (see Fig. 10). All the sensors are within communication range. Sensors transmit at their highest power setting over short distances to ensure a good link, and packet loss at the receiver is forced by dropping packets uniformly at random. One sensor serves as the base station and 18 others are receivers. The last sensor is used to record network traffic. During each experiment, a new file is injected from a PC into the base station and the base station then disseminates it to the network.

B. Results

In our first experiment, we disseminate a single page, 20-packet file using Rateless Deluge. The packet loss probability is $p = 18\%$. We record the number of data packets sent until every node finishes receiving the file. Based on 200 identical iterations, we plot in Fig. 11 the CDF of the number of packets sent and compare it with the analytical estimate from Theorem 1. We observe that the theory predicts well the experimental results. Further, even though the number of sensors in the network is relatively small, the sharp phase transition is still evident.

Next, we compare the performance of Rateless Deluge and Extreme Value FEC Deluge. We distribute a 20-packet file and take averages over 200 identical experiments. We analyze the

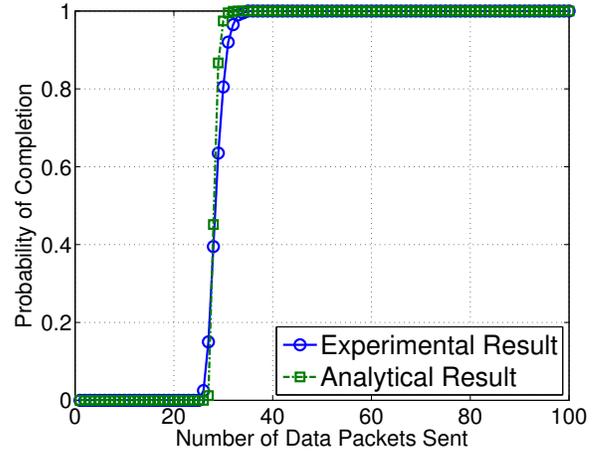


Fig. 11. Real sensor experiments vs. analysis: $N = 18$, $M = 20$, $p = 18\%$.

	Rateless Deluge	EV-FEC Deluge
Average Number of Requests	5,685	1,225
Average Number of Data Packets	29,120	30,270
Completion Time (sec)	2,215	1,140

Fig. 12. Rateless Deluge vs. EV-FEC Deluge: 1 page, $N = 18$, $M = 20$, $p = 18\%$.

network traffic in control plane as well as data plane, namely, we record the number of request messages and data messages sent. We also record the completion time to disseminate the file. The results of the comparison are summarized in Fig. 12.

The results show that Extreme Value FEC Deluge sends out slightly more data messages (less than 5%). However, it drastically reduces the amount of feedback request messages by a factor of about five compared to Rateless Deluge. Note that the minimum possible number of request messages is one since at least one request message must be sent to initiate the dissemination process. With Extreme Value FEC Deluge, the average number of requests is 1.225. Thus, most of the time the entire network finishes receiving enough packets after the base station's first set of transmissions. Thanks to its lower control plane overhead, Extreme Value FEC Deluge effectively reduces the completion time to disseminate a 20-packet file to a 18-node network to 1.14 sec, which is about half of the time needed by Rateless Deluge. We observed similar results when disseminating larger files.

VIII. CONCLUDING REMARKS

In this paper, we developed theoretical foundations and demonstrated practical use of a highly efficient strategy for reliable data broadcasting, called extreme value FEC. This strategy accurately predicts the number of redundant packets to be disseminated by a source so to avoid (with high probability) unnecessary retransmission requests by receivers.

Our analysis, based on extreme value theory, accurately captures characteristics of the completion time of FEC data broadcasting. Not only does it demonstrate the phase transition of the CDF of the completion time, but also accurately pinpoints the location of the phase transition point. The

analysis also reveals that the number of redundant packets required to guarantee file completion by all receivers increases only logarithmically with N . Another major contribution of the paper is in providing convergence bounds for finite N , demonstrating fast convergence of the asymptotic estimate.

By establishing a relation with the coupon collector's problem, we provide asymptotically tight bounds on CDF of the completion time to disseminate a file to receivers with heterogeneous packet loss probabilities. The result points out that, as N gets large, the time needed to disseminate packets to the nodes with the highest packet loss probability dominates. Simulations confirm this finding even when only a small fraction of nodes suffers from high packet loss rates.

Our sensitivity analysis shows that FEC redundancy is robust to imperfect knowledge of the total number of receivers, while uncertainty in the packet loss probability will result in the same order of uncertainty in FEC redundancy. On the other hand, we show that FEC redundancy can be significantly reduced (e.g., on the order of 10%), if we allow incomplete file reception at a single node in the network. However, the marginal gain in allowing incomplete file reception at more nodes quickly diminishes.

Finally, the paper reports a practical implementation of the extreme value FEC strategy in conjunction with the Rateless Deluge OAP protocol. The results show significant performance improvement with respect to control-plane overhead and average data dissemination time, thereby validating the benefits of our approach under real network settings.

The paper leaves many interesting problems for future work. This includes extending the analysis to the case where receivers have temporally correlated packet loss probabilities as well as to multihop network scenarios.

REFERENCES

- [1] J. Hui and D. Culler, "The dynamic behavior of a data dissemination protocol for network programming at scale." in *SenSys'04*, Nov. 2004.
- [2] W. Xiao and D. Starobinski, "Poster abstract: Exploiting multi-channel diversity to speed up over-the-air programming of wireless sensor networks," in *SenSys'05*, San Diego, California, USA, Nov. 2005.
- [3] S. Kulkarni and L. Wang, "MNP: Multihop Network Reprogramming Service for Sensor Networks," in *25th IEEE International Conference on Distributed Computing Systems*, 2005, pp. 7–16.
- [4] D. Starobinski and W. Xiao, "Asymptotically optimal data dissemination in multi-channel wireless sensor networks: Single radios suffice," *IEEE/ACM Transactions on Networking*, to appear.
- [5] A. Hagedorn, D. Starobinski, and A. Trachtenberg, "Rateless deluge: Over-the-air programming of wireless sensor networks using random linear codes," in *IPSN 2008*, Saint Louis, MO, USA, Apr. 2008.
- [6] Chieh-Jan Mike Liang and Răzvan Musăloiu-E. and Andreas Terzis, "Typhoon: A Reliable Data Dissemination Protocol for Wireless Sensor Networks," in *Wireless Sensor Networks*. Springer Berlin / Heidelberg, 2008, pp. 268–285.
- [7] Y.-C. Tseng, S.-Y. Ni, Y.-S. Chen, and J.-P. Sheu, "The broadcast storm problem in a mobile ad hoc network," in *Wireless Networks*, vol. 8, no. 2/3. Kluwer Academic Publishers, 2002, pp. 153–167.
- [8] P. Levis, N. Patel, S. Shenker, and D. Culler, "Trickle: A self-regulating algorithm for code propagation and maintenance in wireless sensor networks," University of California at Berkeley, Tech. Rep., 2004.
- [9] N. Shacham and P. McKenney, "Packet recovery in high-speed networks using coding and buffer management," *INFOCOM '90*, Jun 1990.
- [10] Y. Bartal, J. Byers, M. Luby, and D. Raz, "Feedback-free multicast prefix protocols," *ISCC '98*, pp. 135–141, 1998.
- [11] J. Byers, M. Luby, and M. Mitzenmacher, "A digital fountain approach to asynchronous reliable multicast," *Selected Areas in Communications, IEEE Journal on*, vol. 20, no. 8, pp. 1528–1540, Oct 2002.
- [12] M. Ghaderi, D. Towsley, and J. Kurose, "Network coding performance for reliable multicast," *MILCOM 2007. IEEE*, pp. 1–7, Oct. 2007.
- [13] —, "Reliability gain of network coding in lossy wireless networks," *INFOCOM 2008*, April 2008.
- [14] S. I. Resnick, *Extreme Values, Regular Variation, and Point Processes*. Springer, 1987.
- [15] L. Holst, "Extreme value distributions for random coupon collector and birthday problems," *Extremes*, vol. 4, no. 2, pp. 129–145, 2001.
- [16] W. Xiao, "Reliable data dissemination in dense wireless networks," Ph.D. dissertation, Boston University, 2010.
- [17] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "TAG: a Tiny AGgregation service for ad-hoc sensor networks," *SIGOPS Oper. Syst. Rev.*, vol. 36, no. SI, pp. 131–146, 2002.
- [18] T. He, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "AIDA: Adaptive application-independent data aggregation in wireless sensor networks," *ACM Trans. Embed. Comput. Syst.*, vol. 3, no. 2, pp. 426–457, 2004.
- [19] D. Rubenstein, J. Kurose, and D. Towsley, "Real-time reliable multicast using proactive forward error correction," in *NOSSDAV'98*, 1998.
- [20] L. Rizzo and L. Vicisano, "RMDP: an FEC-based reliable multicast protocol for wireless environments," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 2, no. 2, pp. 23–31, 1998.
- [21] C. Huitema, "The case for packet level FEC," in *Protocols for High-Speed Networks'96*. Chapman & Hall, Ltd., 1996, pp. 109–120.
- [22] J. Nonnenmacher, E. Biersack, and D. Towsley, "Parity-based loss recovery for reliable multicast transmission," in *SIGCOMM '97*, 1997.
- [23] M. Mosko and J. J. Garcia-Luna-Aceves, "An analysis of packet loss correlation in FEC-enhanced multicast trees," in *ICNP '00*, 2000.
- [24] A. Eryilmaz, A. Ozdaglar, and M. Medard, "On delay performance gains from network coding," *CISS*, 2006.
- [25] J. Galambos, *The Asymptotic Theory of Extreme Order Statistics*. Robert Krieger Publishing Company, 1987.
- [26] D. S. J. De Couto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multi-hop wireless routing," in *MobiCom '03*, 2003, pp. 134–146.
- [27] W. Feller, *An Introduction to Probability Theory and Its Applications*. John Wiley & Sons, Inc., 1968, vol. 1.
- [28] S. Ross, *Stochastic Processes*, 1996.
- [29] M. Fogiel and J. R. Ogdin, *Handbook of Mathematical, Scientific, and Engineering Formulas, Tables, Functions, Graphs, Transforms*. Research & Education Assoc., 1984.



Weiyao Xiao received B.E. degree from Harbin Institute of Technology, Harbin, China and M.S. degree from Boston University in 2004 and 2007 respectively. Currently, he is working towards his Ph.D. degree in Electrical and Computer Engineering, also at Boston University. His research interest centers around reliable data dissemination in wireless networks.



David Starobinski received his Ph.D. in Electrical Engineering (1999) from the Technion-Israel Institute of Technology. In 1999-2000 he was a visiting post-doctoral researcher in the EECS department at UC Berkeley. In 2007-2008, he was an invited Professor at EPFL (Switzerland). Since September 2000, he has been with Boston University, where he is now an Associate Professor.

Dr. Starobinski received a CAREER award from the U.S. National Science Foundation and an Early Career Principal Investigator (ECPI) award from the U.S. Department of Energy. He is on the Editorial Board of the *IEEE/ACM Transactions on Networking*. His research interests are in the modeling and performance evaluation of high-speed, wireless, and sensor networks.