2014 IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)

ASYMPTOTIC ANALYSIS OF FAILED RECOVERY PROBABILITY IN A DISTRIBUTED WIRELESS STORAGE SYSTEM WITH LIMITED SUM STORAGE CAPACITY

Bi Hong and Wan Choi

Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea
Department of Electrical Engineering
E-mail: bi_hong@kaist.ac.kr, wchoi@ee.kaist.ac.kr

ABSTRACT

In distributed wireless storage systems, failed recovery probability depends on not only wireless channel conditions but also storage size of each distributed storage node. For efficient utilization of limited storage capacity, we asymptotically analyze the failed recovery probability of a distributed wireless storage system with a sum storage capacity constraint when signal-to-noise ratio goes to infinity, and find the optimal storage allocation strategy across distributed storage nodes in terms of the asymptotic failed recovery probability. It is also shown that when the number of storage nodes is sufficiently large the storage size required at each node is not so large for high exponential order of the failed recovery probability.

Index Terms— Distributed storage system, wireless storage, maximum distance separable coding, failed recovery.

1. INTRODUCTION

In recent years, the advent of social networks, high-definition (HD) video sharing, and cloud storage require seamless large-scale storage. Distributed storage systems are regarded as a key solution for the demand because distributed storage systems can efficiently utilize limited storage capacity and increase reliability of data storing and recovery.

To improve reliability of data storing and recovery, network coding techniques have been applied to distributed storage systems. The functional repair problem in distributed storage systems was studied in [1] by interpreting the problem as a multicasting problem over an information flow graph. For the exact repair, Rashmi *et al.* [2] showed that the optimal minimum bandwidth regenerating (MBR) code can be found for $d = n - 1$, where d and n are the number of surviving nodes and the number of storage nodes, respectively. Suh and Ramchandran [3] found the exact minimum storage regenerating (MSR) code by interference alignment when $\frac{k}{n} \leq \frac{1}{2}$ and $d \geq 2k - 1$, where k is the minimum number of nodes required for data recovery. Other key issues on network codes for distributed storage are well summarized in [4].

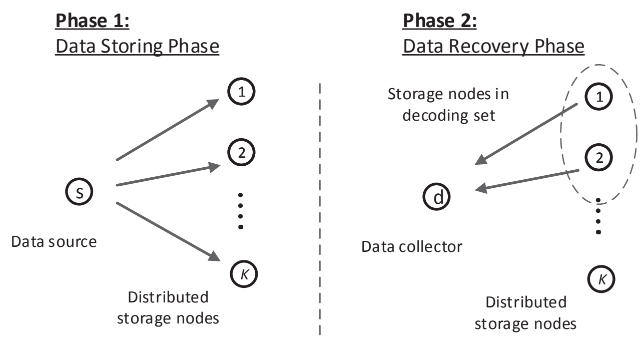

Fig. 1. A distributed wireless storage system model constituted by data storing and recovery phases

For efficient utilization of limited storage capacity, resource allocation in distributed storage systems has been actively researched. Leong *et al.* [5] studied a storage allocation problem under a constraint of total storage capacity. They analyzed recovery probability at the data collector when the link connection probability from each node to the data collector is p, and found the asymptotically optimal allocation policies in terms of recovery probability for large and small total storage budgets, respectively. Ntranos *et al.* [6] extended the result of [5] to a distributed storage system with heterogeneous links where the connection probability from node i to the data collector is p_i. Although the links modeled by Bernoulli random variables partially characterize imperfect channels, they do not exactly address key features of wireless links such as channel fading.

In this context, we consider a distributed wireless storage system where data storing and recovery are carried out through wireless fading channels. We analyze and characterize failed recovery probability in an asymptotic sense, and quantify the effect of limited storage capacity on the asymptotic failed recovery probability. We also find the optimal storage allocation under a constraint of total storage capacity; the optimal allocation strategy is to equally allocate the total storage budget to distributed storage nodes. We also show that the storage size required at each node is not so much for high exponential order of the failed recover probability when the number of storage nodes is sufficiently large.

2. SYSTEM MODEL AND NOTATIONS

2.1. System Model

As shown in Fig. 1, the number of storage nodes is set to K. The source, the collector, and the storage nodes are denoted by s, d and $\{1,\ldots,K\}$, respectively. The collector node recovers the stored data from the storage nodes so that the direct link from the source node to the data collector does not exist. Let s_i be the amount of coded data stored in storage node $i \in \{1,\ldots,K\}$. In our model, each storage node does not have an individual restriction on the storage size but the maximum amount of data stored at each storage node is T because the total storage budget is limited to T such that $\sum_{i=1}^{K} s_i \le T$. For storing data object, the source node broadcasts the data object to the storage nodes during a given time period. We assume that the size of data object is normalized to be unit compared to the total storage capacity. The received signal at the storage node i is given by

$$y_i = h_{si}x_s + z_i \tag{1}$$

where x is the Gaussian distributed data object from the source with an average power constraint $\mathcal{E}[|x_s|^2] \le P$. The channel gain from node i to node j and an additive white Gaussian noise (AWGN) are denoted by $h_{si} \sim \mathcal{CN}(0,1)$ and $z_i \sim \mathcal{CN}(0,N_0)$, respectively. We denote $\rho = P/N_0$ as the average receive signal-to-noise ratio (SNR) of a link.

If a storage node has successfully decoded the broadcast data object from the source, the storage node converts the decoded data object into suitable MDS coded data, and stores it. The amount of MDS coded data stored at node i $(\in \mathcal{D})$ is s_i, where $\mathcal{D}$ is the decoding set whose elements are storage nodes which have successfully decoded the broadcast data object. Successful decoding is possible only when the instantaneous mutual information between the source node and storage node i is greater than or equal to the required rate of the data object, i.e., $\log_2(1+\rho|h_{s,i}|^2) \ge R$. The data object is not properly stored if $\sum_{i\in\mathcal{D}} s_i \le 1$ because the data object cannot be properly recovered from the stored MDS coded data. On the other hand, if $\sum_{i\in\mathcal{D}} s_i > 1$, the collector node has a chance to properly recover the data object from the stored MDS coded data depending on channel conditions from the storage nodes to the collector node. Note that instead of MDS coding, random linear coding over a sufficiently large field can be also used for data object recovery [7].

When the data collector wants to recover the data object, it requests the storage nodes in the decoding set to send the stored MDS coded data. We assume storage node $i \in \mathcal{D}$ transmits during a $t_i = s_i/\sum_{k\in\mathcal{D}} s_k$ fraction of given time period for the recovery stage. That is, the storage node with a larger amount of stored MDS coded data spends more time in the time period for the recovery stage.

2.2. Notations

The exponential equality is denoted as the symbol $\doteq$, i.e., $f(\rho) \doteq \rho^b$, when

$$\lim_{\rho\to\infty} \frac{\log(f(\rho))}{\log(\rho)} = b. \tag{2}$$

In (2), b is called the *exponential order* of $f(\rho)$. Assume that h is a Gaussian random variable with zero mean and unit variance. Then, the asymptotic probability density function (pdf) of the exponential order of $1/|h|^2$ denoted by v could be shown to be $p_v = \lim_{\rho\to\infty} \ln(\rho)\rho^{-v}\exp(-\rho^{-v})$ where $v = -\lim_{\rho\to\infty} \frac{\log(|h|^2)}{\log(\rho)}$. By limiting ρ to infinity, the pdf of v is given by

$$p_v \doteq \begin{cases} \rho^{-\infty} = 0, & \text{for } v < 0, \\ \rho^{-v}, & \text{for } v \ge 0. \end{cases} \tag{3}$$

Thus, for independent random variables $\{v_j\}_{j=1}^{K}$ distributed identically to v, the probability $P_{\mathcal{O}}$ that $(v_1,\ldots,v_K)$ belongs to a non-empty set $\mathcal{O}$ can be characterized by

$$P_{\mathcal{O}} \doteq \rho^{-d_0}, \quad \text{for} \quad d_0 = \inf_{(v_1,\ldots,v_K)\in\mathcal{O}^+} \sum_{j=1}^{K} v_j. \tag{4}$$

3. ASYMPTOTIC ANALYSIS OF THE FAILED RECOVERY PROBABILITY

For the system model described in the previous section, the failed recovery probability (the complimentary recover probability) is hard to obtain in a closed form as noted in [5]. We instead analyze the failed recovery probability for high SNR to understand its asymptotic behavior. That is, we derive the exponential order of the failed recovery probability and find the optimal storage allocation to maximize the exponential order. The exponential order characterizes the decreasing tendency of the failed recovery probability versus SNR, and is interpreted as diversity order if bit error probability or outage probability is considered in traditional wireless communication systems. Contrary to the conventional diversity order, the exponential order of the failed recovery probability is determined by not only the number of independent fading paths for data storing and recovery but also limited total storage capacity.

Theorem 1. *With a resource allocation vector* $\mathbf{s} = \{s_1,\ldots,s_K\}$, *the exponential order of the failed recovery probability is determined as*

$$d(\mathbf{s}) = \min_{\mathcal{D}\subseteq\{1,\ldots,K\}} \left(K - |\mathcal{D}| + \min_{i\in\mathcal{D}} {t_i}^{-1} \cdot \mathbf{1}[\sum_{k\in\mathcal{D}} s_k \ge 1] \right). \tag{5}$$

Proof. By the law of total probability, the failed recovery probability is given by

$$\mathrm{Pr}_f[R] = \sum_{\mathcal{D}\subseteq\{1,\dots,K\}} \mathrm{Pr}_f[R \mid \mathcal{D}]\mathrm{Pr}[\mathcal{D}] \quad (6)$$

where R is the rate of data object (which is equivalent to the data object size). The probability for a decoding set $\mathcal{D}$ is obtained as

$$\begin{aligned}\mathrm{Pr}[\mathcal{D}] &= \prod_{i\in\mathcal{D}} \mathrm{Pr}\left[\log_2(1+\rho|h_{s,i}|^2) > R\right] \\ &\quad\times \prod_{i\in\{1,\dots,K\}\setminus\mathcal{D}} \mathrm{Pr}\left[\log_2(1+\rho|h_{s,i}|^2) < R\right] &(7)\\ &\quad\times \mathrm{Pr}\left[\log_2(1+\rho|h|^2)^k < R\right]^{K-|\mathcal{D}|} &(8)\\ &= \left(e^{-\frac{2^R-1}{\rho}}\right)^{|\mathcal{D}|}\left(1-e^{-\frac{2^R-1}{\rho}}\right)^{K-|\mathcal{D}|} &(9)\\ &\overset{(a)}{\approx} \rho^{-(K-|\mathcal{D}|)} &(10)\end{aligned}$$

where $|\mathcal{D}|$ is the cardinality of the decoding set $\mathcal{D}$ and the approximation of (a) comes from the Taylor's expansion as $\rho \to \infty$. For a given decoding set $\mathcal{D}$, the conditional failed recovery probability with given a decoding is obtained as

$$\begin{aligned}\mathrm{Pr}_f[R \mid \mathcal{D}] =& \mathrm{Pr}[\sum_{i\in\mathcal{D}} t_i \log_2(1+\rho|h_{i,d}|^2) < R]\mathbf{1}[\sum_{k\in\mathcal{D}} s_k \geq 1] \\ &+ \mathbf{1}[\sum_{k\in\mathcal{D}} s_k < 1] &(11)\\ &\overset{(a)}{\doteq} \begin{cases} \rho^{-\min_{i\in\mathcal{D}} t_i^{-1}}, & \text{for } \sum_{k\in\mathcal{D}} s_k \geq 1, \\ 1, & \text{for } \sum_{k\in\mathcal{D}} s_k < 1 \end{cases} &(12)\end{aligned}$$

where $t_i = s_i/\sum_{k\in\mathcal{D}} s_k$ and $\mathbf{1}(\cdot)$ is the indicator function which returns 1 if the argument is true or 0 otherwise. The exponential equality (a) is proved by using the exponential order and (4) such that

$$\begin{aligned}\mathrm{Pr}\left[\sum_{i\in\mathcal{D}} t_i \log(1+|h_{i,d}|^2\rho) < R\right] &\doteq \mathrm{Pr}\left[\sum_{i\in\mathcal{D}} t_i(1-v_i) < 0\right] \\ &\doteq \rho^{-\min_{i\in\mathcal{D}} t_i^{-1}} &(13)\end{aligned}$$

where v_k is the exponential order of $1/|h_{k,d}|^2$. Note that recovery fails if $\sum_{k\in\mathcal{D}} s_k < 1$ due to the property of MDS code and if $\sum_{k\in\mathcal{D}} s_k \geq 1$ the failed recovery probability mainly depends on the cardinality of the decoding set. Combining (10) and (12), the failed recovery probability can be obtained as in Theorem 1. □

From Theorem 1, we can also find the optimal storage allocation to maximize the exponential order. The following corollaries are used for finding the optimal storage allocation.

Corollary 1. *For the non-trivial total storage capacity T (> 1), the optimal storage allocation in terms of exponential order is to symmetrically allocate the sum storage capacity acorss all storage nodes.*

Proof. Our optimization problem is formulated as

$$\begin{aligned}&\max_{\mathbf{s}}\ d(\mathbf{s}) \\ &\text{subject to } s_1 + s_2 + \cdots + s_K = T, &(14)\\ &\quad s_i \geq 0 \qquad \forall i. &(15)\end{aligned}$$

For a given decoding set, the term $\min_{i\in\mathcal{D}} {t_i}^{-1}\cdot\mathbf{1}[\sum_{k\in\mathcal{D}} s_k \geq 1]$ in $d(\mathbf{s})$ has to be maximized. To maximize $\min_{i\in\mathcal{D}} {t_i}^{-1}$, each storage node in the decoding set has the same storage size. For $s_i = \frac{T}{|\mathcal{D}|}$, the term $\mathbf{1}[\sum_{k\in\mathcal{D}} s_k \geq 1]$ obviously becomes 1. Consequently, the term $d(\mathbf{s})$ is maximized when $s_i = \frac{T}{|\mathcal{D}|}$. On the other hand, for all possible $\mathcal{D} \subseteq \{1,\dots,K\}$, the decoding set with full cardinality (i.e., $|\mathcal{D}| = K$) minimizes the exponential order $d(\mathbf{s})$. Therefore, the exponential order $d(\mathbf{s})$ is maximized when $s_1 = \cdots = s_K = \frac{T}{K}$.

□

Corollary 2. *For given K and T, the optimal exponential order of the failed recovery probability is*

$$d^*(K,T) = K - \left\lceil \frac{K}{T} \right\rceil + 1 \quad (16)$$

with the optimal storage allocation policy.

Proof. From Corollary 1 and Theorem 1, we can get the result (16). □

Remark 1. *Corollary 1 is on the same line with the result of [5]. The maximal spreading of the sum storage capacity T (> 1) yields the optimal recovery probability even for distributed wireless storage systems suffering from channel fading.*

.

Remark 2. *The optimal exponential order of the failed recovery probability is bounded by $\left(1-\frac{1}{T}\right)K \leq d^*(K,T) \leq \left(1-\frac{1}{T}\right)K+1$. Thus, the approximated slope of the exponential order is $\left(1-\frac{1}{T}\right)$, which is strictly less than 1, for the sum storage capacity T.*

Although the exponential order well characterizes asymptotic behavior of the failed recovery probability, we also derive a high-SNR approximation of the failed recovery probability for better understanding of recovery success and failure when the sum storage capacity constraint is equally and maximally allocated to each storage node.

Theorem 2. *When SNR is sufficiently high, the failed recovery probability is approximated as*

$$\begin{aligned}&Pr_f^{high}[R] \\ &= \binom{K}{\lceil \frac{K}{T} \rceil - 1}(2^R-1)^{K-\lceil\frac{K}{T}\rceil+1}\rho^{-(K-\lceil\frac{K}{T}\rceil+1)} \quad (17)\end{aligned}$$

Proof. When SNR is sufficiently high, the failed recovery probability is dominated by the probability when $|\mathcal{D}| = \lceil \frac{K}{T} \rceil - 1$ (i.e., the worst exponential order) which is given by

$$\Pr\left[|\mathcal{D}| = \lceil K/T \rceil - 1\right]$$

$$= \binom{K}{\lceil \frac{K}{T} \rceil - 1} \Pr[\log_2(1+\rho|h|^2) > R]^{\lceil \frac{K}{T} \rceil - 1}$$

$$\times \Pr[\log_2(1+\rho|h|^2) < R]^{K - \lceil \frac{K}{T} \rceil + 1} \quad (18)$$

$$\overset{(a)}{\approx} \binom{K}{\lceil \frac{K}{T} \rceil - 1} \left(\frac{2^R - 1}{\rho}\right)^{K - \lceil \frac{K}{T} \rceil + 1} \quad (19)$$

where (a) is due to Taylor's expansion as $\rho \to \infty$. □

4. NUMERICAL RESULTS

Fig. 2 shows that the exponential order of the failed recovery probability increases with the number of storage nodes regardless of sum storage constraints, but the increasing slopes depends on the sum storage constraints. For $T = 2$, the exponential order is limited and the increasing slope of the exponential order is $\frac{1}{2}$ as noted in Remark 2 because the sum storage constraint is small. Fig. 3 shows the effect of the sum storage capacity constraint on the exponential order of the failed recovery probability when the number of storage nodes is fixed. As the number of storage nodes increases, the storage size at each node required for the near optimal exponential order becomes smaller. For example, to obtain 80% of the maximum exponential order under an unlimited sum storage constraint, the required storage sizes at each node (i.e., T/K) are just $s_i = 1/40$ and $s_i = 1/3$ for $K = 200$ and $K = 10$, respectively, as marked as black circles in Fig. 3. The failed recovery probability is also presented in Fig. 4. The figure shows that the asymptotic analysis of the failed recovery probability characterizes the exact failed recovery probability. As shown in the results of exponential order, the failed recovery probability is degraded as the sum storage capacity is smaller. Also, it is also verified that the high-SNR approximation matches well with the simulation results if SNR is greater than 10 dB.

5. CONCLUSIONS

This paper analyzed the failed recovery probability in an asymptotic sense, and showed that the exponential order of the failed recovery probability depends on the number of independent fading paths for storing and recovery but also the sum storage capacity constraint. We proved that the exponential order of the failed recovery probability is maximized when the sum storage capacity is equally allocated across storage nodes. It was also showed that the failed recovery probability can be well approximated in the high-SNR region.

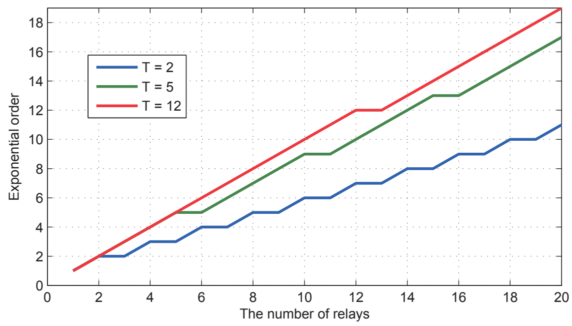

Fig. 2. Exponential order growth for different numbers of storage nodes. Three different sum storage capacity constraints (i.e., $T = 2, 5, 12$) are considered.

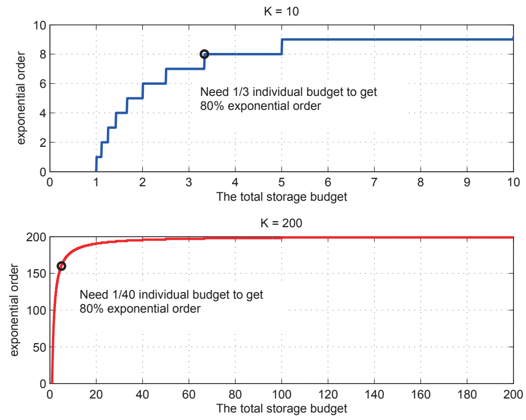

Fig. 3. Exponential order growth for different sum storage capacity T when $K = 10, 200$.

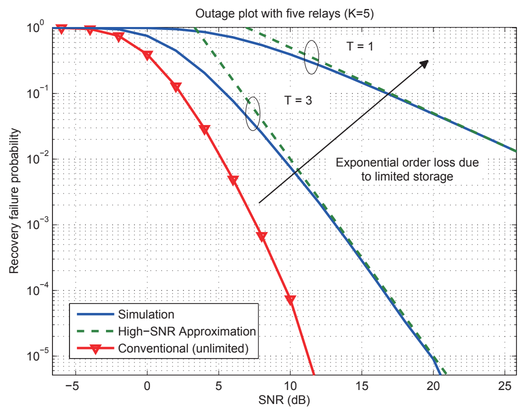

Fig. 4. Failed recovery probability for $K = 5$ and its high-SNR approximations for $T = 1, 3$

6. ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No. 2011-0029329).

7. REFERENCES

[1] A.G. Dimakis, P.B. Godfrey, Y. Wu, M.J. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *Information Theory, IEEE Transactions on*, vol. 56, no. 9, pp. 4539–4551, 2010.

[2] K. V. Rashmi, N.B. Shah, and P.V. Kumar, "Optimal exact-regenerating codes for distributed storage at the msr and mbr points via a product-matrix construction," *Information Theory, IEEE Transactions on*, vol. 57, no. 8, pp. 5227–5239, 2011.

[3] Changho Suh and K. Ramchandran, "Exact-repair mds code construction using interference alignment," *Information Theory, IEEE Transactions on*, vol. 57, no. 3, pp. 1425–1442, 2011.

[4] A.G. Dimakis, K. Ramchandran, Y. Wu, and Changho Suh, "A survey on network codes for distributed storage," *Proceedings of the IEEE*, vol. 99, no. 3, pp. 476–489, 2011.

[5] D. Leong, A.G. Dimakis, and Tracey Ho, "Distributed storage allocations," *Information Theory, IEEE Transactions on*, vol. 58, no. 7, pp. 4733–4752, 2012.

[6] V. Ntranos, G. Caire, and A.G. Dimakis, "Allocations for heterogenous distributed storage," in *Information Theory Proceedings (ISIT), 2012 IEEE International Symposium on*, 2012, pp. 2761–2765.

[7] Tracey Ho, M. Medard, R. Koetter, D.R. Karger, M. Effros, Jun Shi, and B. Leong, "A random linear network coding approach to multicast," *Information Theory, IEEE Transactions on*, vol. 52, no. 10, pp. 4413–4430, 2006.