Secure Distributive Storage of Decentralized Source Data: Can Interaction Help?

Salim El Rouayheb  
University of California, Berkeley  
Email: salim@eecs.berkeley.edu

Vinod Prabhakaran  
University of Illinois, Urbana-Champaign  
Email: vinodmp@uiuc.edu

Kannan Ramchandran  
University of California, Berkeley  
Email: Kannanr@eecs.berkeley.edu

Abstract—We consider the problem of securing a distributed storage system with decentralized data, where some of the nodes are compromised by an eavesdropper. The system is formed of \( n \) storage nodes among which \( k \) nodes (\( k < n \)) have information sources. The system is required to have the “MDS property”, i.e., to allow any user to recover all the sources by contacting any \( k \) nodes. To achieve this goal, the source nodes need to disseminate their data to the other nodes in the system while revealing no information to the eavesdropper. We investigate the role of interaction between the sources in reducing the total required bandwidth. When the sources are independent, we show that interaction does not help and that there always exists an optimal non-interactive scheme.

I. INTRODUCTION

Recent and ongoing impressive advances in the different technologies related to data storage are making it an inexpensive and ubiquitous resource. This has resulted in a surge of applications, such as email service and online storage, that require storing and maintaining large amounts of data. For reliability and scalability reasons, this data is distributed on multiple storage nodes interconnected by a network, forming, thus, a distributed storage system \([1]–[4]\). A user of this system should be able to download his or her data by contacting a targeted minimum number of the storage nodes. It is likely, however, for some of these nodes to become compromised by a passive eavesdropper who can observe their stored data, and listen to their incoming and outgoing communications. This would be the case, for instance, when the system is connected to the Internet which makes it vulnerable to different kinds of “phishing” activities.

This paper investigates strategies to make distributed storage systems secure against eavesdropping for a general scenario where the data sources exist in a decentralized fashion at different nodes in the system. Consider for example the storage system depicted in Figure 1 consisting of three nodes \( v_1, v_2, \) and \( v_3 \) having large storage capacities. Two equal-length binary files, \( X_1 \) and \( X_2 \), are stored respectively on nodes \( v_1 \) and \( v_2 \). Node \( v_3 \) joins the network with no initial data. Each node \( v_i \) has available to him an independent random key \( K_i \). The figure also depicts a data collector (DC) wanting \( X_1 \) and \( X_2 \) who is connected to nodes \( v_2 \) and \( v_3 \).

Assuming that each node \( v_i \) has access to an independent random key \( K_i \), Table I proposes such protocol that consists of four rounds. At the beginning (round 0), the nodes start by having their private keys stored, in addition to any original data, \( X_1 \) or \( X_2 \), that they possess. At any subsequent round, the transmitter node is indicated by a star “*” symbol, while the destination node is labeled by the message it receives and stores. At the end of the last round, the column below each node

This research was funded in part by an AFOSR grant (FA9550-09-1-0120), a DTRA grant (HDTRA1-09-1-0032), and an NSF grant (CCF-0830788).
optimal, but non-interactive, scheme would consist of for the system depicted in Figure 1, an alternative each source encodes its data separately. For example, the minimum bandwidth. We show that there always ex-

<table>
<thead>
<tr>
<th>round</th>
<th>v1</th>
<th>v2</th>
<th>v3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>X1,K1</td>
<td></td>
<td>K3</td>
</tr>
<tr>
<td>1</td>
<td>*</td>
<td>K1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>*</td>
<td>X1 + K1</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>X2 + K1 + K2</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>*</td>
<td>K1 + K2</td>
<td></td>
</tr>
</tbody>
</table>

TABLE I
AN OPTIMAL INTERACTIVE PROTOCOL FOR THE DISTRIBUTED STORAGE SYSTEM OF FIGURE 1. THE STAR SYMBOL "*" INDICATES THE TRANSMITTING NODE IN EACH ROUND.

indicates its stored content. It is easy to check that the protocol proposed in Table I will allow the system to satisfy the demands of any data collector while keeping the information leaked to the eavesdropper limited to the sources on the nodes he observes.

We will show later that this protocol is optimal, and that at least four units of information need to be transmitted in the system. Note that this protocol is interactive since all the nodes cooperate together in order to achieve the desired system requirements. For instance, the messages sent by v2 in the third and fourth rounds depend on the messages it has previously received from v1 in the first round. Two important questions arise here. First, what is the minimum amount of information that should be exchanged in a distributed storage system in order to satisfy the demand of the data collectors and meet the secrecy constraint? Second, is interaction necessary for achieving this optimum?

In this paper, we answer these two questions by demonstrating that interaction is not needed to achieve the minimum bandwidth. We show that there always exists a non-interactive bandwidth-optimal scheme where each source encodes its data separately. For example, for the system depicted in Figure 1, an alternative optimal, but non-interactive, scheme would consist of v2 transmitting X2 + K2 and K2 to v1 and v3 respectively, in the last two rounds. Subsequently, analyzing the non-

This paper is organized as follows. In Section II, we describe the system model and define the security problem. In Section III, we propose a non-interactive protocol and analyze its performance. In Section IV, we show that this non-interactive protocol is optimal for a special family of distributed storage systems. We generalize this result in Section V. Finally, we conclude in Section VI.

II. Model

Consider a distributed storage system comprised of n nodes v1, v2, . . . , vn. Each of these nodes is assumed to have a large storage capacity1. At each of the first k nodes vi, i = 1, . . . , k, is located an independent information source represented by a random variable Xi of unit entropy (H(Xi) = 1) and uniformly distributed over a large finite alphabet A. For ease of reference, we call the first k nodes source nodes, and the remaining ones storage nodes. We assume that each node vi has access to an independent random number generator, also called a key, represented by the random variable Ki, taking values from a large finite alphabet.

The nodes v1, v2, . . . , vn are fully connected in a mesh network (complete graph) and communicate among themselves in rounds. Initially, at round 0, the storage content of each node consists of its corresponding information source, if any, and key. At round r, a single node vt can be chosen to transmit a message Yr to a single destination node vd, d = v, through an error-free link. Node vd then stores the received message Yr. Due to the decentralized nature of the problem, the message Yr should satisfy what we call the local causality constraint, i.e., Yr should be a function of only the content stored at node vt, which includes the key Kr, the messages received by vt in the previous r − 1 rounds, in addition to the source Xr, if vt is a source node.

Ultimately, the distributed storage system needs to service users, also called data collectors, interested in knowing X1, . . . , Xk and which can contact any set of k out of the n nodes. We assume that a number b < k of nodes, whose identity is not known, has been compromised by an eavesdropper who can observe their stored data, as well as their incoming and outgoing messages. We are interested in finding the minimum amount of information that needs to be communicated among the nodes in the system in order to achieve the following two conditions: (i) MDS property: any data collector contacting k nodes should be able to recover all the sources X1, . . . , Xk with no errors, (ii) perfect secrecy: any eavesdropper observing b nodes should gain no information about the sources at the other nodes in the system.

Let V = {v1, v2, . . . , vn} be the set of storage nodes and X = {X1, X2, . . . , Xk} the set of information sources. For any positive integer n, we define the set [n] := {1, 2, . . . , n}. For any subset S ⊆ [n], we define VS := {vi : i ∈ S}, KS := {Ki : i ∈ S} and XS := {Xi : i ∈ S ∩ [k]}. Suppose the system

1Upper bounds on the required storage capacity of the nodes will be derived later.
implements a protocol consisting of $N$ rounds, where in round $r$ the message $Y_r$ is transmitted by node $v_r$. Let $T$ and $W$ be subsets of $[n]$, we define $Y_{T,W}(r)$ to be the set of messages transmitted from the nodes in the set $V_T$ to the nodes in the set $V_W$ during the first $r$ rounds. If $T$ and $W$ intersect non-trivially, $Y_{T,W}^N$ includes the messages exchanged among the nodes in $V_T \cap W$ during the first $r$ rounds. We will usually drop the index $N$ from $Y_{T,W}^N$ and write $Y_{T,W}$. For $r = 1, \ldots, N$, the local causality constraint can be now written as

$$Y_r = \begin{cases} f_r(X_{t_r}, K_{t_r}, Y_{V,t_r}^{r-1}) & \text{if } t_r \in [k] \\ f_r(K_{t_r}, Y_{V,t_r}^{r-1}) & \text{otherwise,} \end{cases} \tag{1}$$

where $f_r(.)$ is a deterministic function, $Y_{T,W}^T = \emptyset$, $\forall T, W \subseteq [n]$ and $Y_{V,t_r} = Y_{V,(t_r)}$. When the functions $f_r$ do not depend on $Y_{V,t_r}^{r-1}$ for all $r$, we say that the scheme is non-interactive. Denoting by $D$ the set of nodes that a data collector can contact, the MDS property implies

$$H(X_D|X_D, K_D, Y_{V,D}) = 0, \forall D \subset [n], |D| = k. \tag{2}$$

Let $E$ denote the set of compromised nodes, the perfect secrecy condition can then be written as

$$H(X_E|X_E, K_E, Y_{V,E}) = H(X_E), \forall E \subset [n], |E| = b, \tag{3}$$

where $D$ and $E$ are respectively the complements of $D$ and $E$ in $[n]$.

The total amount of information transmitted in the system is $R := \sum_{r=1}^{N} H(Y_r)$. Our objective is to find an optimal communication scheme that will lead to a minimum value of $R$, denoted by $R^*(n, k, b)$, while satisfying the three conditions (1), (2), and (3).

### III. The Non-Interactive Upper Bound

In this section, we derive an upper bound on the optimum bandwidth $R^*(n, k, b)$ by analyzing a non-interactive protocol where each source encodes its information independently before distributing it to the other nodes. In this case, from the standpoint of any single source, the distributive storage system can be looked at as a combination network [5, Chap. 4], and the problem becomes a special case of designing multicast network codes [6]–[8]. Figure 2 illustrates the system as a combination network from the vantage point of source $X_1$ for the case of $n = 5, k = 3$ and $b = 1$.

![Fig. 2. The encoding of source $X_1$ in a non-interactive protocol for a distributed storage system with $n = 5, k = 3$ and $b = 1$. The other sources $X_2$ and $X_3$ located at nodes $v_2$ and $v_3$ use a similar code. Only the relevant data collectors, those who are not directly connected to $v_1$, are depicted here.](Image)

It was shown in [7] and [8] that the coset codes proposed by Ozarow and Wyner [9] for the wiretap channel of type II can be generalized for multicast networks with a single source to achieve perfect secrecy, provided that the alphabet of the source is large enough. We apply this scheme here to each source $X_i$ separately by taking $X_i = (X_{i1}, \ldots, X_{ik-b}) \in \mathbb{F}_q^{k-b}$ and the key $K_i = (K_{i1}, \ldots, K_{ik}) \in \mathbb{F}_q^k$ for a large prime power $q$. A special coset code based on a linear MDS code then takes $X_i$ and $K_i$ as inputs and outputs the codeword $(C_i^1, \ldots, C_i^{n-1}) \in \mathbb{F}_q^{n-1}$. Node $v_i$ then transmits a different symbol $C_j^i$ to each of the $(n-1)$ remaining nodes in the system. Such code is shown in Figure 2 for the source $X_1 \in \mathbb{F}_2^{k}$ which is split into $k-b = 2$ bits $X_{11}, X_{12} \in \mathbb{F}_2$ and encoded, using the key $K_1 \in \mathbb{F}_2$ into $(C_1^1, C_1^2, C_1^3, C_1^4) = (X_{11} + K_1, X_{12} + K_1, K_1, X_{11} + X_{12} + K_1)$. The other sources use a similar code. The linear MDS code implies $H(C_j^i) = \frac{n-1}{q-1}$. Therefore, using this scheme, each source node $X_i$ transmits during $n-1$ rounds at a total rate of $\sum_{j=1}^{n-1} H(C_j^i) = \frac{n-1}{q-1}$. Summing over all the sources, we obtain the following upper bound on $R^*$:

$$\text{Lemma 1: } R^*(n, k, b) \leq \frac{k(n-1)}{k-b}.$$  

We want to show that this upper bound is tight and that the non-interactive scheme described above is optimal. This result is summarized in Theorem 2.

**Theorem 2:** For the distributed storage system described above any protocol that simultaneously achieves the MDS and the perfect secrecy conditions will need to use a bandwidth of at least

$$R^*(n, k, b) = \frac{k(n-1)}{k-b}, \tag{4}$$

and the non-interactive scheme described above achieves this bound.

To prove this theorem, we first show that it is true for the special case of $n = k+1$. Then, we reason by induction on the total number of nodes $n$ to prove it for the general case.
IV. A SPECIAL CASE: \( n = k + 1 \)

We consider in this section the simplest non-trivial distributed storage systems which are the ones comprising a single storage node. We want to show that the non-interactive scheme of the previous section is optimal in this case. To that end, we first note that any feasible protocol should satisfy the following Markovian property which is a direct consequence of the local causality condition and the independence of the sources.

**Lemma 3:** For every \( i \in [k] \) and \( D = [n] \setminus \{i\} \), \( X_i - (Y_{i,D}, Y_{D,i}) - (X_D, K_D) \), is a Markov chain.

We can now write, for all \( i \in [n] \) and \( D = [n] \setminus \{i\} \)

\[
H(X_i|Y_{i,D}, Y_{D,i}) = H(X_i|X_D, K_D, Y_{i,D}, Y_{D,i})
\]

\[
H(X_i|X_D, K_D, Y_{i,D}, Y_{D,i}) = H(X_i|X_D, K_D, Y_V, D_D)
\]

\[
H(X_i|X_D, K_D, Y_V, D_D) = 0.
\]

(a) follows from Lemma 3, (b) by the local causality condition which implies that \( Y_{DD} \) is a deterministic function of \( X_D, K_D \) and \( Y_{i,D} \), (c) follows from the MDS property of Eq. (2). Equation (5) implies that the information of any source \( X_i \) should be completely recoverable from all the messages transmitted and received by node \( v_i \). Using the chain rule, we can generalize it to any set \( S \subseteq [k] \) of source nodes to get

\[
H(X_S|Y_{S,S}, Y_{S,S}, Y_{S,S}) = 0.
\]

The equivocation rate of the eavesdropper for any set \( E \subseteq [n] \) corresponding to \( b \) compromised nodes can be now written as

\[
H(X_E|X_E, K_E, Y_{VE}) = H(X_E|X_E, K_E, Y_{VE}, Y_{EE})
\]

\[
H(X_E|X_E, K_E, Y_{VE}, Y_{EE}) = H(X_E|X_E, K_E, Y_{VE}, Y_{EE})
\]

\[
H(X_E|X_E, K_E, Y_{VE}, Y_{EE}) \leq H(Y_{EE})
\]

where \( E \) is any subset of \([n]\) of cardinality \( b \). Summing Eq. (8) over all possible choices of \( E \subset [k] \) and \(|E| = b\), we get

\[
(k - 2) \sum_{i,j \in [k]} H(Y_{ij}) + (k - 1) \sum_{i=1}^{k} (H(Y_{ni}) + H(Y_{mi})) \geq \binom{k}{b} (k - b).
\]

Similarly, summing over all possible choices of \( E \) such that \( n \in E \), we get

\[
(k - 2) \sum_{i,j \in [k]} H(Y_{ij}) \geq \binom{k}{b - 1} (k - b + 1).
\]

Adding Eqs. (10) and (11) and using the submodularity of the entropy function, we get

\[
\sum_{i=1}^{N} H(Y_{i}) \geq \sum_{i,j \in [n]} H(Y_{ij}) \geq \frac{k^2}{k - b}.
\]

The previous equality gives a lower bound on the optimal bandwidth \( R_* \) which matches the upper bound of Lemma 1 for the special case of \( n = k + 1 \).

**Lemma 5:** For the case of \( n = k + 1 \), the non-interactive scheme is optimal and the minimum required bandwidth is

\[
R_* (k + 1, k, b) = \frac{k^2}{k - b}.
\]

Next, we show that the non-interactive optimal scheme also minimizes the rate \( R_{sto} \) of messages transmitted to the storage node \( v_n \). This result will constitute an important step in proving Theorem 2 for the general case. Define \( R_{sto} := \sum_{i,j \in [n]} H(Y_{ij}) \), and let \( R_{sto}^* \) be the minimum value of \( R_{sto} \) over all bandwidth-optimal protocols.

**Lemma 6:** \( R_{sto}^* = \frac{k^2}{k - b} \).

**Proof:** Consider a relaxed version of this problem with the same setting, but where the message sent by a node can be a function of all the messages that has been previously transmitted in the system, and not only the ones stored at the transmitting node. Let \( R_{rel}^* \) and \( R_{sto,rel}^* \) be the minimum bandwidth used in the system, and the minimum rate of the messages sent to \( v_n \) for the relaxed problem. Since any protocol for the original problem will also be feasible for the relaxed version, we have \( R_{rel}^* \leq R_*^* \) and \( R_{sto,rel}^* \leq R_{sto}^* \). It can be seen that Lemmas 3, 4, and 5 still hold for the relaxed version. Therefore, we have \( R_{rel}^* = R_*^* \).
In the relaxed problem, there is always an optimal protocol where node \( v_n \) does not transmit any message. Indeed, since \( v_n \) does not have any information source, any other node can take its place in transmitting in the relaxed version of the problem. Therefore, we can set 
\[
H(Y_{ni}) = 0, \quad i = 1, \ldots, k
\]
in Eq. (10) with no loss of generality. Then, taking equalities and subtracting Eq. (11) from Eq. (10), we get 
\[
R_{Sto,rel}^* \geq \frac{k}{k-b}
\]
and those transmitted by the source nodes to the first node, \( v_0 \), and those transmitted by the source nodes to the last node, \( v_{n0+1} \), are identical to those transmitted by \( v_n \) during round \( r \). Again, without loss of generality, we can assume that only the source nodes transmit messages, i.e., \( t_r \in [k] \) for \( r = 1, \ldots, N \). Therefore, the transmitted messages in the system can be partitioned into those transmitted by the source nodes to the first \( n_0 \) nodes, and those transmitted by the source nodes to the last node, \( v_{n0+1} \). Therefore, we can write:
\[
\sum_{r=1}^{N} H(Y_{r}^*) \geq \sum_{r:t_r \in [k]} H(Y_{r}^*) + \sum_{r:t_r \in [n_0]} H(Y_{r}^*) + \sum_{r:t_r \in [n_0+1]} H(Y_{r}^*). \tag{13}
\]
Note that at the end of this protocol, the system formed by the first \( n_0 \) nodes, \( v_1, \ldots, v_{n_0} \), satisfies the MDS and the perfect secrecy conditions. Since these nodes did not receive any transmissions from the last node \( v_{n0+1} \), we can use the induction hypothesis to get:
\[
\sum_{r:t_r \in [k], d_r \in [n]} H(Y_{r}^*) \geq R_{rel}^* (n_0, k, b) = \frac{k(n_0 - 1)}{k-b}. \tag{14}
\]
Similarly, the system formed by the \( k+1 \) nodes \( v_1, \ldots, v_k \) in addition to the last node \( v_{n0+1} \) satisfies the MDS and the perfect secrecy conditions. Therefore, by Lemma 6 we have
\[
\sum_{r:t_r \in [k], d_r \in [n_0+1]} H(Y_{r}^*) \geq R_{Sto,rel}^* = \frac{k}{k-b}. \tag{15}
\]
Adding Equations (14) and (15), we get 
\[
R_{rel}^* (n_0 + 1, k, b) \geq \frac{k(n_0 + 1)}{k-b}. \]

VI. CONCLUSION

We studied in this paper the problem of securing a distributed storage system with decentralized sources, i.e., sources that are located on different nodes, against an eavesdropper that can observe the content of a fixed number of the system nodes not known in advance. Our objective was to protect the data from being leaked to the eavesdropper while minimizing the total bandwidth used in the system. A relevant question that emerges here is whether interaction among the nodes can help reduce the bandwidth necessary for achieving the security requirement. We answered this question in the negative and showed that interaction is not needed when the sources are independent and all the nodes have random keys. To that end, we demonstrated that, in this case, it is always optimal to decouple the problem into achieving security for each source independently. For further research, it is interesting to see how this result would extend to the case of correlated sources or the case where not all the nodes can generate random keys.

REFERENCES