On the resistance to Byzantine and unresponsive servers in code-based PIR schemes

Jun Kurihara\(^1, 2, \textit{a)}\) and Toru Nakamura\(^3\)

\(^1\) Graduate School of Applied Informatics, University of Hyogo, 7-1-28 Minatojima-minamimachi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
\(^2\) Zettant Inc., 4-8-16-5F (BcH) Nihonbashi-honcho, Chuo-ku, Tokyo 103-0023, Japan
\(^3\) Advanced Telecommunications Research Institute International (ATR), 2-2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-0288, Japan
\(\textit{a)}\) kurihara@ieee.org

Abstract: In private information retrieval (PIR) from coded storage servers, consider the case where some of servers are Byzantine adversaries and unresponsive. There have been proposed several specialized schemes guaranteeing that the user can correctly obtain the desired message even in the adversarial situation. However, to the best of our knowledge, such resistance to the adversaries in PIR schemes based on arbitrary codes have been not precisely characterized. In this paper, we reveal that the exact resistance to Byzantine and unresponsive servers is expressed in terms of the coset distance of linear codes in linear PIR schemes based on arbitrary storage code.

Keywords: private information retrieval, Byzantine adversaries, coset distance, relative generalized Hamming weight

Classification: Fundamental Theories for Communications

References

[1] B. Chor, O. Goldreich, E. Kushilevitz, and M. Sudan, “Private information retrieval,” Proc. IEEE FOCS 1995, Milwaukee, WI, USA, pp. 41–50, Oct. 1995. DOI: 10.1109/sfcs.1995.492461

[2] R. Tajeddine, O.W. Gnilke, and S. El Rouayheb, “Private information retrieval from MDS coded data in distributed storage systems,” IEEE Trans. Inf. Theory, vol. 64, no. 11, pp. 7081–7093, Nov. 2018. DOI: 10.1109/TIT.2018.2815607

[3] R. Freij-Hollanti, O.W. Gnilke, C. Hollanti, and D.A. Karpuk, “Private information retrieval from coded databases with colluding servers,” SIAM J. Appl. Algebra Geometry, vol. 1, no. 1, pp. 647–664, Jan. 2017. DOI: 10.1137/16M102562

[4] L. Li, M. Militzer, and A. Datta, “rPIR: ramp secret sharing-based communication-efficient private information retrieval,” Int. J. Inf. Security, vol. 16, no. 6, pp. 603–625, Nov. 2017. DOI: 10.1007/s10207-016-0347-8

[5] R. Tajeddine, O.W. Gnilke, D. Karpuk, R. Freij-Hollanti, and C. Hollanti, “Private information retrieval from coded storage systems with colluding, Byzantine, and unresponsive servers,” IEEE Trans. Inf. Theory, vol. 65, no. 6, pp. 3898–3906, June 2019. DOI: 10.1109/TIT.2018.2890285
[6] L. Holzbaur, R. Freij-Hollanti, and C. Hollanti, “Towards the capacity of private information retrieval from coded and colluding servers,” https://arxiv.org/abs/1903.12552, March 2020.

[7] I.M. Duursma and S. Park, “Coset bounds for algebraic geometric codes,” Finite Fields and Their Applications, vol. 16, no. 1, pp. 36–55, Jan. 2010. DOI: 10.1016/j.ffa.2009.11.006

[8] Y. Luo, C. Mitrpant, A.J. Han Vinck, and K. Chen, “Some new characters on the wire-tap channel of type II,” IEEE Trans. Inf. Theory, vol. 51, no. 3, pp. 1222–1229, March 2005. DOI: 10.1109/TIT.2004.842763

1 Introduction

Private Information Retrieval (PIR) [1] is a protocol allowing a user to retrieve a message from a set of $n$ storage (database) servers, without revealing any information about the identity of the message to each or sets of colluded servers. Recently in the community of information theory, PIR schemes for distributed coded storage have been actively investigated, e.g., [2, 3, 4], where every original message is encoded by an $[n, k]$ linear code $C$, and each coordinate of an encoded message is stored at each server. On the other hand, in the context of PIR schemes, several new adversarial settings have also been introduced. Byzantine and unresponsive servers [5, 6] are such ones responding to the user’s query with polluted answers, where a Byzantine server is defined as a storage server maliciously returning an erroneous answer, and an unresponsive server is the one that returns nothing.

In the setting of PIR from coded storage, Tajeddine et al. [5] proposed an explicit scheme in which the user correctly and privately retrieves the desired message when $n > k + t + 2b + r - 1$ for $b$ Byzantine and $r$ unresponsive servers, assuming $t$ colluded servers. However, although the scheme was proven to be optimal in the sense of the download efficiency under a specific setting of $C$ [6], the resistance to $b$ Byzantine and $r$ unresponsive servers in PIR schemes based on arbitrary linear codes have not been characterized yet. In this paper, we reveal that for PIR schemes based on arbitrary linear codes, the exact resistance to Byzantine and unresponsive servers can be expressed in terms of the coset distance [7] (also known as the first relative generalized Hamming weight [8]) of the linear codes.

2 Preliminary

Let $\mathbb{F}$ stand for a finite field, and let $[n] \triangleq \{1, \ldots, n\}$. For a matrix $X \in \mathbb{F}^{m \times n}$, we represent its elements, row vectors and column vectors by subscripts and superscripts as

$$X = \begin{bmatrix} x_1^1 & x_2^1 & \ldots & x_n^1 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^m & x_2^m & \ldots & x_n^m \end{bmatrix} = [x_1, \ldots, x_n].$$

Let $\mathbb{F}^n = \mathbb{F}^{1 \times n}$. For two row vectors, $v = [v_1, \ldots, v_n] \in \mathbb{F}^n$ and $w = [w_1, \ldots, w_n] \in \mathbb{F}^n$, we denote by $\langle v, w \rangle$ their standard inner product, and by $v \circ w \triangleq [v_1w_1, \ldots, v_nw_n]$...
their Hadamard product. We also define a subspace $C \circ D$ from two subspaces $C, D \subseteq \mathbb{F}^n$ as $C \circ D \triangleq \text{span}\{v \circ w : v \in C, w \in D\}$.

**Definition 1** (Minimum Hamming distance and coset distance [7, 8]). For a linear code $C \subseteq \mathbb{F}^n$, the minimum Hamming distance of $C$ is defined as $d_{\text{min}}(C) \triangleq \min \{d(v, w) : v, w \in C, v \neq w\}$, where $d(v, w) = |\{i : v_i \neq w_i\}|$. For two linear codes $C \subseteq \mathbb{F}^n$ and $D \subseteq C$, the coset distance [7] is defined by

$$d_{\text{min}}(C/D) \triangleq \min \{d(x, y) : x \in v + D, y \in w + D : v, w \in C, v - w \not\in D\} = \min \{d(v, 0) : v \in C \setminus D\}, \quad (1)$$

which coincides with the first relative generalized Hamming weight [8].

### 3 Private information retrieval from coded storage

Following [2, 3, 4, 5, 6], first let $C$ be a linear code $C \subseteq \mathbb{F}^n$ of $\dim C = k$, called the storage code. Let $y^i \in C$ be an encoded message for $i \in [m]$ mapped from an original message $x^i \in \mathbb{F}^k$ by a certain bijection $\mathbb{F}^k \rightarrow C$, and we have the following matrix $Y$.

$$Y \triangleq \begin{bmatrix} y^1 \\ \vdots \\ y^m \end{bmatrix} = \begin{bmatrix} y_1^1 & \cdots & y_n^1 \\ \vdots & \ddots & \vdots \\ y_1^m & \cdots & y_n^m \end{bmatrix} = [y_1, \ldots, y_n] \in \mathbb{F}^{m \times n}. \quad (2)$$

Then the $j$-th server ($j \in [n]$) stores the $j$-th column $y_j \in \mathbb{F}^{m \times 1}$ of $Y$. Under this storage setting, the user aims to retrieve a message(s) by a linear PIR scheme. Here we shall introduce the linear PIR scheme that is generalized from existing ones [2, 3, 5, 6] by re-defining the possible set of queries by the direct sum of subspaces as given in the following.

**Definition 2** (Linear PIR scheme). Considering the setting described above, a linear PIR scheme over $\mathbb{F}$ executes the following steps.

1. First fix the inner query code $D_{\text{in}} \subseteq \mathbb{F}^n$, and the outer query code $D_{\text{out}} \supseteq D_{\text{in}}$. Also choose a subspace $E \subseteq \mathbb{F}^n$ such that $D_{\text{out}} = E + D_{\text{in}}$, and choose a vector $e \in E$.

2. For every $i \in [m]$, set a probability space $(Q[i], \mu[i])$ of queries as

$$Q[i] = \left\{ \begin{array}{l} v^j \in \mathbb{F}^{m \times n} : v^j \in D_{\text{in}} \text{ for } j \neq i, v^j \in e + D_{\text{in}} \text{ for } j = i \end{array} \right\}. \quad (2)$$

When the user wishes to download a $x^i$, a query $Q[i] \in Q[i]$ is selected randomly according to the probability measure $\mu[i]$, e.g., typically the uniform probability $\mu[i] = 1/|\mathbb{F}|^m \dim D_{\text{in}}$. Each $Q[i]$ is also defined by a matrix $Q[i] \triangleq [q[i], \ldots, q[i]] \in \mathbb{F}^{m \times 1}$, where $q[i] \in \mathbb{F}^{m \times 1}$ is sent to the $j$-th server.

3. At the $j$-th server, its response $r[i]_j \triangleq \langle q[i], y_j \rangle \in \mathbb{F}$ is computed and transmitted to the user. We then set $r[i] \triangleq [r[i], \ldots, r[i]] \in \mathbb{F}^m$ as the response to $Q[i]$. 


4. The user and servers repeatedly execute steps 1–3 until the user retrieves sufficient data segments.

The scheme of Definition 2 does not guarantee the correctness meaning that the user can correctly retrieve some coordinates of \( y' \), by itself. In order to guarantee it, we need to additionally introduce stronger conditions on \( D_{\text{out}}, D_{\text{in}} \) and \( E \) in step 1 as \([2, 3, 4, 5]\). Here we emphasize that such (specific) conditions are not required to characterize the resistance to Byzantine and unresponsive servers.

**Remark 3.** For the correctness, in \([2, 3, 5]\), \( E \) and its basis \( \{e^1, \ldots, e^{\dim E}\} \) are chosen such that \( \dim E = \dim(C \circ E) \leq d_{\min}(C \circ D_{\text{in}}) - 1 \) by setting its basis with \( d(e^j, 0) = 1 \) for all \( j \in [\dim E] \) and \( e = \sum_{j \in [\dim E]} e^j \). This condition was generalized as the strong linearity in \([6]\). On the other hand, under the setting of a repetition code \( C \), the paper \([4]\) assumes that \( D_{\text{out}} \subseteq \mathbb{F}^{m} \) and \( D_{\text{in}} \subseteq D_{\text{out}} \) are both MDS, and that \( E \cap D_{\text{in}} = \{0\} \) holds.

**Remark 4.** In the scheme of Definition 2, the maximum number of colluded servers, \( t \), capable to guarantee the privacy is characterized by \( t \leq d_{\min}(D_{\text{in}}^c) - 1 \) \([3, \text{Theorem 8}]\).

4 Resistance to Byzantine and unresponsive servers

This section derives the resistance to Byzantine and unresponsive servers in the generalized scheme of Definition 2. To this end, we first give the following analytic lemma characterizing the component of the response \( r[i] \).

**Lemma 5.** Consider the scheme of Definition 2. Then the response \( r[i] \) to the query \( Q[i] \) is represented as an element of a coset in a quotient space \( (C \circ D_{\text{out}})/(C \circ D_{\text{in}}) \), which is given by

\[
r[i] \in z^i + C \circ D_{\text{in}} \in (C \circ D_{\text{out}})/(C \circ D_{\text{in}}),
\]

where \( z^i \in \mathbb{Z} \) is an element of a subspace \( \mathbb{Z} \subseteq C \circ E \) with \( \mathbb{Z} + C \circ D_{\text{in}} = C \circ D_{\text{out}} \) and \( \mathbb{Z} \cap C \circ D_{\text{in}} = \{0\} \).

**Proof.** Recall that for each row \( q[i] \) of the query \( Q[i] \in Q[i] \), we have \( q[i] \in D_{\text{in}} \) for \( j \neq i \), or \( q[i] \in E + D_{\text{in}} = D_{\text{out}} \) for \( j = i \) from Eq. (2). We denote \( d^j \in D_{\text{in}} \) as an element of \( D_{\text{in}} \) chosen to generate \( q[i] \), that is, \( q[i] = d^j \) for \( j \neq i \) and \( q[i] = d^i + e \) for \( j = i \). Since \( D_{\text{in}} \subseteq D_{\text{out}} \), we see \( q[i] \in D_{\text{out}} \), and hence \( y^j \circ q[i] \in C \circ D_{\text{out}} \) for any \( y^j \in C \). From this observation, the response \( r[i] \) to the query \( Q[i] \in Q[i] \) can be rewritten as follows.

\[
r[i] = \sum_{\mu \in [m]} y^\mu \circ q[i] = \sum_{\mu \in [m] \setminus \{i\}} y^\mu \circ d^\mu + y^i \circ (d^i + e) = y^i \circ d^i + y^i \circ e
\]

Here we immediately see that \( y^i \circ e \in C \circ E \) is decomposed as \( y^i \circ e = z^i + w \) for a certain \( w \in C \circ E \cap C \circ D_{\text{in}} \). This completes the lemma. \( \square \)
Remark 6. In \[2, 3, 4, 5, 6\], \(C \circ E = Z\) in Lemma 5 holds by choosing \(E \subseteq Z^n\) as the one satisfying \(C \circ D_{in} \cap C \circ E = \{0\}\). That is, \(z^i = y^i \circ e\) holds.

In the existing PIR schemes \[2, 3, 5, 6\], desired parts in the response are hidden by the randomness, and the user cancels the randomness to elicit the desired parts. In the context of Lemma 5, \(C \circ D_{in}\) is served as the source of randomness, and \(z^i \notin C \circ D_{in}\) is exactly the desired parts in \(r[i]\). To obtain \(z^i\), the existing schemes first identify the randomness given as a codeword of \(C \circ D_{in}\), and subtract the randomness from \(r[i]\). For this operation, Lemma 5 directly means that eliciting \(z^i\) from \(r[i]\) coincides with identifying the coset \(z^i + C \circ D_{in}\). From this perspective, eliciting \(z^i\) from the response polluted by Byzantine and unresponsive servers can be viewed as a game to identify the unique coset \(z^i + C \circ D_{in} \ni r[i]\) from the polluted response. Here we shall define the resistance to Byzantine and unresponsive servers in this sense.

**Definition 7** (The \((b, r)\)-resistance of linear PIR scheme). Suppose that arbitrary \(b\) of \(n\) storage servers are Byzantine servers who maliciously return erroneous responses. Also suppose that arbitrary \(r\) of \(n\) storage servers are unresponsive servers who do not return responses. Then, assume the user has no knowledge on server indices of Byzantine servers, and receives the polluted response \(\hat{r}[i]\) instead of \(r[i]\), in which arbitrary \(b\) coordinates and \(r\) coordinates of \(r[i]\) are alternated and erased, respectively. Under this setting, a PIR scheme is called the one attaining the \((b, r)\)-resistance if the user can correctly identify the original coset \(W \in (C \circ D_{out})/(C \circ D_{in})\), \(r[i] \in W\) from the polluted response \(\hat{r}[i]\).

In the existing schemes \[2, 3, 4, 5, 6\], by introducing additional conditions on \(E, D_{in}\) and \(D_{out}\), it is guaranteed that some coordinates of \(y^i\) (and hence \(x^i\)) is obtained from the identified \(z^i\). In the setting of Definition 2 and Definition 7, since definitions of \(E, D_{in}\) and \(D_{out}\) are more relaxed than those, such decodability on the identified coset \(z^i + C \circ D_{in} \ni r[i]\) is not always guaranteed.

For Definition 7, we eventually present the following theorem that exactly characterize the \((t, b)\)-resistance of PIR schemes based on arbitrary \(E, D_{out}\) and \(D_{in}\) only by the parameter of the quotient space \((C \circ D_{out})/(C \circ D_{in})\).

**Theorem 8.** Consider the PIR scheme in Definition 2. Then, the user can attain the \((b, r)\)-resistance if and only if \(d_{min}((C \circ D_{out})/(C \circ D_{in})) > 2b + r\).

**Proof.** From Lemma 5, the response \(r[i]\) to the query \(Q[i]\) is an element of a coset \(z^i + C \circ D_{in}\), and \(z^i \in Z\) satisfying \(Z + C \circ D_{in} = C \circ D_{out}\) and \(Z \cap C \circ D_{in} = \{0\}\). Thus from Eq. (1) and [7, Lemma 1.1], when at most any \(b\) coordinates of \(r[i]\) are alternated, the coset \(z^i + C \circ D_{in} \ni r[i]\) is identified if and only if \(d_{min}((C \circ D_{out})/(C \circ D_{in})) > 2b\). Considering \(r\) erasures in \(n\) coordinates in addition to \(b\) alternations, we additionally require more than or equal to distance \(r\) to identify the coset. Therefore the theorem holds.

**Example 9.** As an example, we analyze the scheme of [5]. In the scheme, the storage code \(C\) and the inner query code \(D_{in}\) are both generalized Reed-Solomon (GRS) codes. The outer query code \(D_{out}\) is a GRS code chosen such that \(\dim D_{out} = \dim D_{in} + 1\). These imply that \(C \circ D_{in}\) and \(C \circ D_{out}\) are also GRS codes with...
\text{dim}(C \circ D_{\text{in}}) = \text{dim} C + \text{dim} D_{\text{in}} - 1 \text{ and } \text{dim}(C \circ D_{\text{out}}) = \text{dim} C + \text{dim} D_{\text{in}} \text{ from [3, Proposition 3]}, \text{ and that } d_{\text{min}}((C \circ D_{\text{out}})/(C \circ D_{\text{in}})) = n - \text{dim}(C \circ D_{\text{out}}) + 1 \text{ from [8, Corollary 2]}. \text{ We thus have } d_{\text{min}}((C \circ D_{\text{out}})/(C \circ D_{\text{in}})) = n - \text{dim} C - \text{dim} D_{\text{in}} + 1. \text{ Therefore from Theorem 8, the scheme attains the } (b, r)-\text{resistance if } n - \text{dim} C - \text{dim} D_{\text{in}} - 1 > 2b + r \text{ is supposed as shown in [5].}

**Acknowledgments**

The authors would like to thank Dr. Daishi Kondo for valuable discussions.