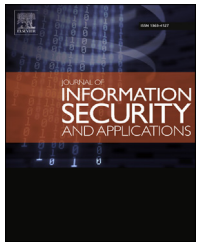
JOURNAL OF INFORMATION SECURITY AND APPLICATIONS ■■ (2016) ■■–■■

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/jisa



# Toward single-server private information retrieval protocol via learning with errors ☆

*Zengpeng Li* [a,b], *Chunguang Ma* [a,b,*], *Ding Wang* [c], *Gang Du* [a,b]

[a] *College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China*
[b] *Fujian Provincial Key Laboratory of Network Security and Cryptology, Fujian Normal University, Fuzhou 350007, China*
[c] *School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China*

ARTICLE INFO

*Article history:*
Available online

*Keywords:*
Private information retrieval
Homomorphic encryption
Multi-bit symmetric encryption
Learning with errors

ABSTRACT

At FOCS2011 Brakerski and Vaikuntanathan proposed a single-server LWE-based private information retrieval (abbreviated as PIR) protocol with a security reduction to hard standard lattice problems and nearly optimal communication complexity. However, Brakerski just described a generic PIR protocol that utilized a somewhat homomorphic encryption and an arbitrary symmetric encryption as building blocks, he did not instantiate the generic construction. In this work, we first modify Brakerski's construction without the evaluating key and construct a new PIR model. Moreover, we instantiate our new model via matrix FHE first proposed by Ryo et al. at PKC2015 and vector symmetric encryption scheme proposed in this work as building block. Then we optimize the Response operations and several other aspects of the scheme.



## 1. Introduction

The advent of cloud computing has greatly changed the IT landscape over the recent years, while security concerns have been long cited as barriers to wider adoption of cloud services. Specifically, on the one hand, if the sensitive data and operation are outsourced to a no-trust third party, maintaining confidentiality, integrity and availability is a challenge. On the other hand, the user does not want to reveal his private information when people search it on Google or Bing. Hence, in this paper, we force on how to keep user's privacy. In fact, a Private Information Retrieval (PIR) Protocol allows a user to retrieve an element from a server in possession of a database without revealing any information about which element is retrieved. An example: Bob is in possession of a database $x = \{0;1\}^n$ from which Alice wants to read the $i$-th bit $x_i$ without Bob knowing the index $i$ in which Alice is interested in. A trivial solution to this problem would be for Alice to download the entire database, assuming that the server would know which elements were not retrieved and therefore gain information about which element was retrieved. This implies that the trivial approach is also the optimal solution to private information retrieval from a single server.

2010 MSC: 00-01, 99-00.

☆ Fully documented templates are available in the elsarticle package on CTAN.

* *Corresponding author.* College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China. Fax: 0451-82588389.

*E-mail addresses:* lizengpeng@hrbeu.edu.cn (Z. Li), machunguang@hrbeu.edu.cn (C. Ma), wangdingg@pku.edu.cn (D. Wang), dugang@hrbeu.edu.cn (G. Du).

http://dx.doi.org/10.1016/j.jisa.2016.11.003
2214-2126/

Please cite this article in press as: Zengpeng Li, Chunguang Ma, Ding Wang, Gang Du, Toward single-server private information retrieval protocol via learning with errors ☆, Journal of Information Security and Applications (2016), doi: 10.1016/j.jisa.2016.11.003

Computationally Private Information Retrieval (CPIR) is a notable branch of privacy protecting protocol research. The definition of CPIR is based on PIR and the assumption that the sever is polynomially bounded. Contrary to Information Theoretic Private Information Retrieval, CPIR allows protocols to involve only one single server. Therefore, the user can only receive one piece of information he queried from the only one server. This variation is called Symmetric private information retrieval (SPIR) (Gertner et al., 1998), where the privacy of the data, as well as the privacy of the user, is guaranteed. That is, in every invocation of a SPIR protocol, the user learns only a single physical bit of $x$ and no other information about the data.

One of the requirements for a PIR protocol is that communication must be sublinear. Therefore, every SPIR protocol is always a 1-out-of-$n$ oblivious transfer protocol as well. Moreover, one can arbitrarily transform single database CPIR protocol into SPIR protocol (Di Crescenzo et al., 2000). This implies that every single database CPIR protocol is also a 1-out-of-$n$ oblivious transfer protocol.

Definition 1. *In a 1-out-of-n oblivious transfer protocol the party called sender has n messages, but wishes to share only one of those. The other party, receiver, wants to receive the message corresponding to an index i, which should remain hidden from the sender.*

PIR is still an active field of theoretical research as even current schemes struggle to be incorporated into application. After (Kushilevitz and Ostrovsky, 1997) proving that CPIR schemes do not need to rely on non-co-operating replications, many different CPIR protocols based on various other homomorphic crypto-systems have been discovered subsequently. Most notably is the result by Lipmaa (2005) which was based on the Damgard–Jurik homomorphic cryptosystem. Homomorphic encryption techniques are very important and are natural methods to protect user's privacy by encrypting algorithm or privacy preserving protocols, such as oblivious transfer and private information protocol. The most recent schemes based on NTRU-FHE (Doröz et al., 2014), (somewhat) additively homomorphic encryption (Lepoint and Tibouchi, 2015), have been published. Dai et al. proposed an Accelerating SWHE based PIR using GPU scheme (Dai et al., 2015). This has been proved to be the best scheme which actually is fast enough to be put into actual application.

Especially motivated by recent breakthrough in FHE, hence, in this paper, we proposed a new single-database PIR protocol via matrix-FHE scheme and multi-bit symmetric scheme. The rest of this paper is organized as follows: In Section 2, we reviewed Brakerski's scheme, including some of its constructions and provided some background on fully homomorphic encryption. In Section 3, we described our two building blocks. In Section 4, we described our PIR protocol. Finally, we concluded our work and talked future research in Section 5.

## 2. Preliminaries

In this section we introduce some notations and briefly review the representation of private information retrieval protocol. The details are as follows:

### 2.1. Notation

For ease of presentation, we employ some initial notations listed in Table 1 and will follow the notations in Brakerski et al.'s scheme as closely as possible.

### 2.2. Review of Brakerski et al.'s private information retrieval

Formally, a single-database PIR protocol is a two-party protocol, consisting of four algorithms as in Cachin et al. (1999) and Gentry and Ramzan (2005), where a user retrieves the $i$-th bit from a $K$-bit database $DB = b_1 b_2 \cdots b_K$, without revealing the value of $i$ to the database server.

- Setup algorithm (*Setup*) phase, $(pp, ss) \leftarrow PIR.Setup(K, i, 1^\lambda)$: The protocol begins in an off-line setup phase that does not depend on the index to be queried nor on the contents of the database. The receiver runs the setup algorithm, taking as input a security parameter $\lambda$, the size $K$ of database, and the index $i$ of a bit in the database. It thus obtains a public set of parameter $pp$ (the public key) that is sent to the sender, and a secret state $ss$ that is kept private. Once the setup phase is completed, the receiver and sender can run the remainder of the protocol an unbounded number of times.
- Query Generation (*QGen*) phase, $(Q, qs) \leftarrow PIR.QGen(1^\lambda, K, i, ss)$: When the receiver wishes to receive the $i$-th element in the database $DB[i]$, it takes as input a security parameter $\lambda$, the size $K$ of database, the index $i$ of a bit in the database, and outputs a query $Q$ and a query secret $qs$.
- Response Generation (*RGen*) phase, $resp \leftarrow PIR.RGen(1^\lambda, pp, DB[K], Q)$: The sender accesses to a database $DB \in \{0,1\}^K$, upon receiving the query message query $Q$ from the receiver.
- Decode Response (*DRes*) phase, $x \leftarrow PIR.DRes(pp, ss, resp, qs)$: Upon receiving $resp$, the receiver decodes the response algorithm and output $x \in \{0, 1\}$ is the output of the protocol.

We note that while in general a multi-round interactive protocol is required for each database query, the protocols we present are of the simple form of a query message followed by a response message. Hence, we choose to present the simple syntax above. The communication complexity of the protocol is defined to be $|query| + |resp|$. Namely, the number of bits has been exchanged to transfer a single database element (excluding the setup phase). We sometime analyze the query length and the response length separately.

**Table 1 – Notations.**

| Notation | Meaning |
|---|---|
| $\lambda$ | Security parameter |
| $K$ | Size of database |
| $i$ | Index |
| $pp$ | Public parameter |
| $ss$ | Secret state |
| $Q$ | Query |
| $resp$ | Response |

Please cite this article in press as: Zengpeng Li, Chunguang Ma, Ding Wang, Gang Du, Toward single-server private information retrieval protocol via learning with errors ☆, Journal of Information Security and Applications (2016), doi: 10.1016/j.jisa.2016.11.003