Private Data Broker

No Author Given

No Institute Given

The Original Protocol

The original protocol is designed by De Cristofaro and Tsudik [CT09]. We have public data:

$$\{n, e, H(), H'()\},$$

which consists of the RSA modulus n, the RSA public key e, and two hash functions $H(), H'()$. The client's input to the protocol is: $\mathcal{C} = \{hc_1, \cdots, hc_v\}$, where $hc_i = H(c_i)$. The server's input to the protocol is d, the RSA private key, and $\mathcal{S} = \{hs_1, \cdots, hs_w\}$, where $hs_j = H(s_j)$. In the offline phase of the protocol, the Client computes:

$$\{R_{c:i} \leftarrow \mathbb{Z}_n^* \text{ and } y_i = hc_i \cdot (R_{c:i})^e \bmod n\}_{\forall i},$$

and the Server computes:

$$\{K_{s:j} = (hs_j)^d \bmod n \text{ and } t_j = H'(K_{s:j})\}_{\forall j}.$$

The goal of the protocol is for the client and server to be able to privately compute the set intersection:

$$\mathcal{C} \cap \mathcal{S} = \{hc_1, \cdots, hc_v\} \cap \{hs_1, \cdots, hs_w\}.$$

In Figure 1, we outline the online phase of the protocol. Here, we rename the client to be the *query provider*, and we rename the server to be the *data provider*.

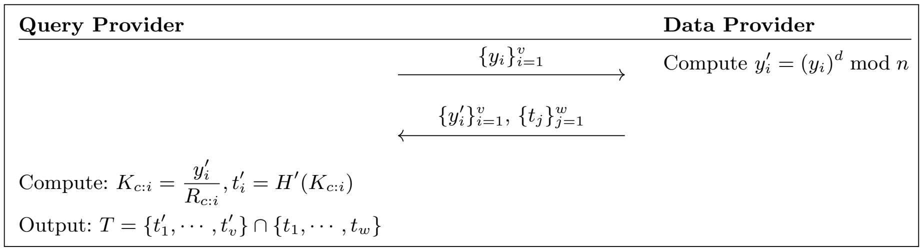
Query Provider
Data Provider
$\{y_i\}_{i=1}^v$
Compute $y'_i = (y_i)^d \bmod n$
$\{y'_i\}_{i=1}^v$, $\{t_j\}_{j=1}^w$
Compute: $K_{c:i} = \dfrac{y'_i}{R_{c:i}}, t'_i = H'(K_{c:i})$
Output: $T = \{t'_1, \cdots, t'_v\} \cap \{t_1, \cdots, t_w\}$

Fig. 1. The on-line phase of the original protocol

Adding a Broker

The updated version of the protocol introduces a third party, referred to as the *broker*, to the process.

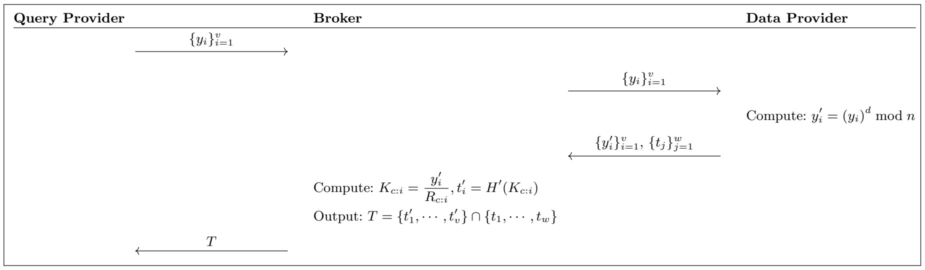
Query Provider
Broker
Data Provider
$\{y_i\}_{i=1}^v$
$\{y_i\}_{i=1}^v$
Compute: $y_i' = (y_i)^d \bmod n$
$\{y_i'\}_{i=1}^v$, $\{t_j\}_{j=1}^w$
Compute: $K_{c:i} = \frac{y_i'}{R_{c:i}}, t_i' = H'(K_{c:i})$
Output: $T = \{t_1', \cdots, t_v'\} \cap \{t_1, \cdots, t_w\}$
T

Fig. 2. The on-line phase of the new protocol

A Comparison of Views

In this section, we briefly compare the view of each of the parties in the two respective protocols.

Party	View
QP	$\{y_i\}, \{y_i^d\}, \{t_j\}, \{t_j'\}, T$
DP	$\{y_i\}, \{y_i^d\}, \{t_j\}$
Broker	`n/a`

Table 1. The view of each of the two parties in the original protocol.

Party	View
QP	$\{y_i\}, T$
DP	$\{y_i\}, \{y_i^d\}, \{t_j\}$
Broker	$\{y_i\}, \{y_i^d\}, \{t_j\}, \{t_j'\}, T$

Table 2. The view of each of the three parties in the new protocol.

So the view of the *data provider* is unchanged, but the query provider gets to see less information about the items that the data provider computes in the original protocol. So, essentially, the query provider only gets to see their query and the result.

Discussion

So what does this mean? Essentially, the new protocol has the following differences:

1. A new party, the *data broker*, is able to view $\{y_i\}, \{y_i^d\}, \{t_j\}, \{t'_j\}, T$, essentially allowing them a full view of the process.
2. The query provider is no longer able to view $\{y_i^d\}, \{t_j\}, \{t'_j\}$.

Let's look at these items one-by-one. The first item $\{y_i^d\}$, corresponds to the query provider's values $\{y_i\}$ to the value of the data provider's private key d. Due to the hardness of the DLP [McC90], we know that the query provider cannot recover the private key d from the values $\{y_i^d\}$. The second item, $\{t_j\}$, corresponds to hash values provided by the server and the third item, $\{t'_j\}$, corresponds to hash values originally provided by the query provider in the first protocol.

Questions: does this mean that the query provider no longer gets to see the size of the data provider's dataset (i.e. the value w in the above two figures)? In the updated protocol, the query provider only sees the size of the resulting intersection.

References

[CT09] Emiliano De Cristofaro and Gene Tsudik. Practical private set intersection protocols with linear computational and bandwidth complexity. Cryptology ePrint Archive, Report 2009/491, 2009. `https://eprint.iacr.org/2009/491`.

[McC90] Kevin S McCurley. The discrete logarithm problem. In *Proc. of Symp. in Applied Math*, volume 42, pages 49–74. USA, 1990.