Weekly Research Progress Report

Jaspal Singh, Ning Luo

March 5, 2018

Research Problems

Both of us were thinking of looking at only the following two research problems closely for the next few weeks.

1. Designing function secret sharing schemes based on pseudo-random functions other than the GGM construction
2. Designing ORAM based SSE schemes for conjunction, disjunction and general boolean queries

For the first question we currently looking at the Naor-Reingold function and its properties and trying to see if it can be used to construct an FSS scheme for some class of function and for more than 2 parties. We don't yet have a concrete correct construction for this problem.

For the second problem we have been reading some recently published literature on SSE (the list of papers discussed are mentioned in the next Section). We believe that there is a good scope to improve these constructions. There is one concrete constructions that we are currently brainstorming on: (An overview of the idea is described below)

Krell et al. [3] describe in their paper on how the leakage in the Blind Seer [2] SSE protocol can be reduced using an ORAM protocol. The setting assumed in [3] is that of Outsourced Symmetric Private Information Retrieval (OSPIR), in which case there are three parties: a client, a server and a data owner. For the case of 2 parties (where the client is the data owner) the authors of [3] also mention an ORAM based construction where each bloom filter can be stored as a block in an ORAM construction. In this construction, to access a path (of length $\log n$, where n is the number of documents) in the bloom filter tree would require us to perform $\log n$ ORAM accesses on the ORAM primitive that stores the bloom filters. We believe that we can optimize the above construction further:

Rather than using a bloom filter, which supports only addition and search functions, we would like to use a hash function h that supports addition, search and deletion functions, along with being space efficient. Given such a hash function, we hope to convert the Blind Seer's bloom filter based tree construction to a similar hash tree construction with the following modifications:

- Each node (bucket) in the hash tree stores a hash table h along with a set of documents
- Whenever a document d present in the hash tree is accessed along the path from the root to the node containing d, we can remove the document d from its location and add it to the root node. We also delete the document d from all the hash function on the path of the hash tree on which document d resides
- To avoid overflow at topmost buckets of the tree, we perform an eviction algorithm along some pre-determined paths after some fixed number of accesses to the tree. Unlike a block in an ORAM tree, which is labeled with a path, a document present is the hash tree is not associated with any leaf node. Hence, no matter which path is chosen for eviction, documents can always be flushed down the path. A major point to note is that if we flush a document d down a path P then we would also need to insert the document d in the hash functions present on the path P. This is necessary to have the SSE search protocol to be similar to that of Blind Seer.

The efficiency of this scheme largely depends on the space efficient hash function h, hence we will further look at candidate hash functions and see what level of efficiency they can provide. Another key optimization parameter we are looking at is the number of rounds, For this we will try looking at if ideas from TWORAM [4] can be used in this scenario.

The references mentioned above correspond to the papers described below.

Research Papers Discussed

1. Cash, David, et al. "Highly-scalable searchable symmetric encryption with support for boolean queries." Advances in Cryptology–CRYPTO 2013. Springer, Berlin, Heidelberg, 2013. 353-373.

    - In this paper,the author present a SSE protocol which support general boolean queries with non-naive way. Moreover, the complexity for conjunctive search is sub-linear in the total number of documents in the database.
    - The basic idea behind this protocol for the conjunctive search $w_1 \wedge w_2 ... \wedge w_q$, is that the result must be a subset of DB(w_1). Hence, in the proposed protocol encrypted indexes contained in DB(w_1) are leaked to the server, who can then efficiently check if each of these documents contain keywords $w_2, \ldots, w_q$.
    - To make server be able to help the client to compute the intersection of DB(w_1),...DB(w_q) without decrypting encrypted index, it introduces a XSET, with which, the tokens given by the client and the precomputed blinded value stored in the EDB, the server can check if the documents in DB(w_1) are also contained in DB(w_i) for all i. The author employ two very basic and simple cryptographic primitives: DDH hard problem and a PRF.
    - A key drawback of this protocol is that the complexity for disjunctive search is linear in the total number of documents in the database.

2. Pappas, Vasilis, et al. "Blind seer: A scalable private DBMS." Security and Privacy (SP), 2014 IEEE Symposium on. IEEE, 2014.

    - In this paper, the authors combine bloom filter and MPC to realize a scalable private DBMS, which keep the queries secret. Further more, this system will support any boolean formula, and the complexity (for both communication and search) is sub-linear in the number of documents.
    - The authors mainly use two cryptographic primitives, bloom filter and MPC. The EDB here is actually a tree with intermediate nodes containing a encrypted bloom filter and leaf nodes containing a index for a document. Client can pick up correct traveling paths via secure multi-party computations with clients.
    - While this system achieve sub-linear complexity for general boolean formula, the disadvantages of this system are obvious : for efficiency concerns, the round complexity is proportional to the query complexity; for security concerns, the leakage cause by access pattern is considerable.

3. Krell, Fernando, et al. "Low-Leakage Secure Search for Boolean Expressions." Cryptographers' Track at the RSA Conference. Springer, Cham, 2017.

    - In this paper the authors employ the ORAM technique to reduce the leakage associated with the Blind Seer construction for searchable encryption.
    - The authors are interested in the OSPIR setting containing three parties: a client, a server and a data owner. In this setting, on each search query, we must not only ensure that the server does not learn anything about the database and the query but we must also ensure that the client does not learn anything more than the query output.
    - The authors use three cryptographic primitives in their protocol: Oblivious PRF, MUL-OPRF and Masked MOPRF, all of which could be constructed from a Hashed Diffie-Hellman PRF.
    - We believe that the search complexity of this scheme compared to the Blind Seer protocol has two multiplicative overheads due to: (i) the ORAM protocol and (ii) the size of the encryption of each bit in the bloom filter.

4. Garg, Sanjam, Payman Mohassel, and Charalampos Papamanthou. "TWORAM: efficient oblivious RAM in two rounds with applications to searchable encryption." Annual Cryptology Conference. Springer, Berlin, Heidelberg, 2016.

    - In this paper the authors design a 2 round ORAM protocol assuming server side computation. The bandwidth of this protocol is a multiplicative factor (the security parameter) more than the Path ORAM construction.
    - The authors also describe how the proposed ORAM construction can be used to develop a 4 round SSE scheme for single keyword searches.

- The key idea we found in the paper is to avoid $O(\log n)$ rounds spent on evaluating the recursive ORAM structure which stores the position map. And the authors achieve this by storing a garbled circuit in each tree node, which encodes the content of that particular bucket. All the garbled circuits evaluated during one ORAM access are replaced with new ones, since non-reusable garbled circuits are assumed.

5. Boyle, Elette. "Recent Advances in Function and Homomorphic Secret Sharing." International Conference in Cryptology in India. Springer, Cham, 2017.

   This is a survey paper providing all recent advances (upto October 2017) on both the HSS and the FSS primitive. The paper also proposes a large number of open problems, of which we are currently trying to address the following while using non-GGM based FSS construction:

   - FSS schemes for CNF/DNF formulas
   - 3-server FSS schemes for DPF (distributed point function) with complexity better than $\mathcal{O}(2^{n/2})$ key size, where n is the length of the input for the point function

Aim for the next week

- Formalize the proposed construction for SSE using ORAM and ideas from Blind Seer. Further analyze the search complexity of the protocol.
- Look at other possible optimizations for SSE constructions allowing boolean queries. For example, we will be looking at combining ideas from TWORAM into proposed Blind Seer based construction.
- Discuss the following recent paper on SSE : "Kamara, Seny, and Tarik Moataz. 'Boolean searchable symmetric encryption with worst-case sublinear complexity.' Annual International Conference on the Theory and Applications of Cryptographic Techniques. Springer, Cham, 2017."
- Aim at coming up with FSS constructions for more than 2 parties and for new function classes.