Efficient Ancestry, Paternity, and Genomic Compatibility Tests Using Server-Aided Secure Two-Party Function Evaluation

Fattaneh Bayatbabolghani

Department of Computer Science and Engineering, University of Notre Dame

Motivations

- ❑ Genomic data is used in medical domain and other domains.
- ❑ Genomic data is highly sensitive; it discloses information about data owner, as well as owner's relatives.

Goals

- ❑To protect genomic data during the computations of genomic tests.
- ❑To develop secure evaluation techniques for malicious and semi-honest users in the server-aided setting.

Preliminaries of Genomic Testing

- ❑ Genomes represent complete hereditary information of an individual.
- ❑ Information extracted from one's genome is often represented as SNPs and STRs.
  - ➢ Each SNP is referenced by a specific index and its value is 0, 1, or 2.
  - ➢ Each STR consists of fixed number of pairs which are integer numbers.
- ❑ Three common genomic tests:
  - ✓ **Ancestry**: Given two SNP sequences belonging to two individuals, the test determines the number of SNPs they have in common.
  - ✓ **Paternity with a Single Parent**: Given two STR profiles $S = \langle\{x_{1,i}, x_{2,i}\}\rangle$ and $S' = \langle\{x'_{1,i}, x'_{2,i}\}\rangle$, the test compute: $\wedge_{i=1}^{N}[\{x_{1,i}, x_{2,i}\} \cap \{x'_{1,i}, x'_{2,i}\}] = True$
  - ✓ **Genetic Compatibility**: Given markers (multiple location of SNPs) of potential partners, the test compare them to evaluate the possibility of their children inheriting special diseases.

Garbled Circuits

- ❑ There is two-party traditional garbled circuits technique which allows two parties P_1 and P_2 to securely evaluate a Boolean circuit of their choice.

Desired function should be converted to a Boolean circuit and then:

- ➢ One party acts as a circuit generator and creates a garbled representation of the circuit by associating both values of each binary wire with random labels.
- ➢ Two parties perform oblivious transfer (OT) to transfer circuit evaluator's input labels.
  - ➢ In 1-out-of 2 OT, sender has x_0 and x_1, and receiver obtains only x_b without revealing any information about b for sender.
- ➢ Another party acts as a circuit evaluator and evaluates the circuit in its garbled representation without knowing the meaning of the labels.

The output labels can be mapped to their meaning and revealed to either or both parties.

❖ **We do not use the traditional technique, we add a server to make the computation faster.**

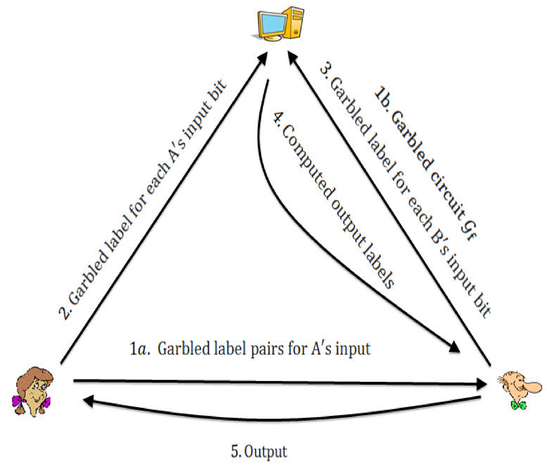
Designed Protocols and Applications

❑ **Ancestry**: This test would often be invoked when two parties already know to be related or have reasons to believe to be related. Thus, they are unlikely to try to cheat each other. For that reason, we use a solution with semi-honest parties to perform this test.

Designed Protocols and Applications

Semi-honest parties:

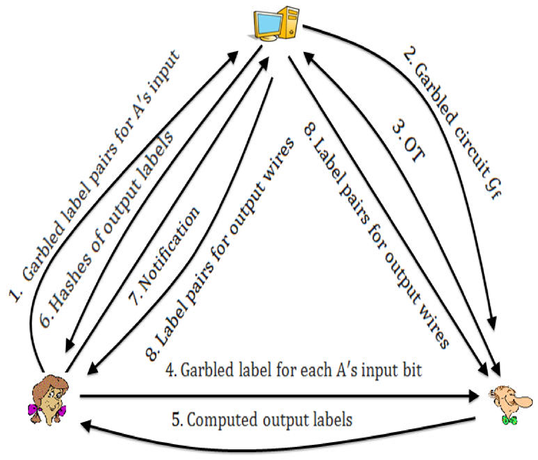
1- Both parties jointly create a garbled representation of the circuits.
2- They send the appropriate label for their input bits to the server for evaluation.
3- The server returns the output to the parties.
4- The parties can determine the result by mapping the output to its actual meaning.



❑ **Paternity**: We assessed that a security setting with malicious users is the most suitable for running paternity tests because the participants may be inclined to tamper with the computation to influence the result.

Malicious parties:

1- Alice and the server jointly create a garbled representation of the circuits.
2- Alice sends Bob her input bits.
3- Bob is engaged by the server in oblivious transfers to find his input bits. Then
4- Bob acts as the evaluator of garbled circuits.
5- Bob sends the computed labels to Alice.
6- With the help of the server, Alice verifies the output labels without recognizing their meanings.
7- Once the output labels are verified, Alice notifies the server which sends the label pairs to Alice and Bob.
8- Both parties can interpret and learn the result.



❑ **Genetic Compatibility**: We added input certification to the security setting with malicious users because one party may modify the input to learn other party's genetic information.

Malicious parties with input certification:

The basic structure of the protocol remains the same as previous protocol, but we extend it with a novel mechanism for obliviously verifying correctness of the inputs.

Conclusions

- ❑ We developed general secure solutions in our server-aided framework with several applications in genomic testing and other domains.
- ❑ Running time of all of our experiments indicate execution efficiency of the protocols.

This work was done under supervision of Professor Marina Blanton and supported by grant CNS-1223699 from the National Science Foundation.