622.755 – Introduction to Cybersecurity Winter Term, 2020/21

Project Write Up

Denis D'Ambrosi 12245845

1 Outline

This report investigates the implementation of Yao's protocol within the field of financial data analysis from a social, legal and ethical perspective. The paper is structured as follows: section 2 introduces the protocol itself and gives an overview of my implementation of the exchange, along with a relevant use case. Section 3 investigates the deployment of Yao's protocol in some areas of financial data analysis, while section 4 analyzes some of social, ethical and legal questions introduced with such technology. Finally, section 5 briefly summarizes the contents of this report.

2 An overview of Yao's protocol

Yao's protocol is a two-party secure computation scheme firstly introduced in 1982 [1] and then re-elaborated in 1986 [2] by Andrew Yao that enables two parties to jointly compute a function f on their private inputs x and y without disclosing any information about those data points to each other. The protocol relies on the principle of *garbled circuits*, an algorithmic technique for encoding logical circuits while mantaining the inputs secret: to execute the exchange, one party (the garbler) encrypts the circuit that computes the desired function f along with its input x and sends it to an evaluator, which uses the encoded data received to locally compute f with his input y. This *Secure Multi Party Computation* (SMPC) scheme exploits *oblivious transfer*, a cryptographic mechanism introduced by Rabin in an earlier paper [3], that ensures that the two parties are able to evaluate $f(x, y)$ without disclosing to the each their personal input.

Yao's protocol offers strong privacy assurances, as neither party gains any knowledge of the other's data nor intermediate values computed during the protocol. This makes it an ideal communication primitive for applications where confidentiality is essential (see figure 1).

3 Yao's protocol in the wild: SMPC within the financial sector

In the implementation part of this project, I have developed an application based on garbled circuits that allows two clients to securely compute a joint sum between their secret inputs. In financial data analysis, such functionality could be clearly beneficial to organizations that need to evaluate collective statistics without disclosing individual

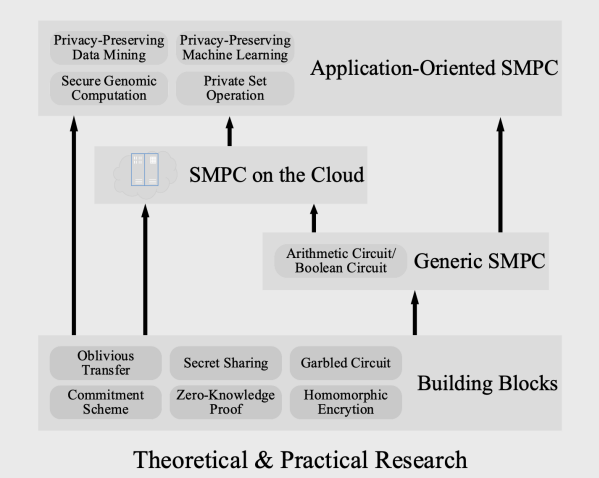

Figure 1: An overview of SMPC and its applications. Image taken from page 2 of [4].

values. This can be especially helpful when financial institutions have to comply with strict regulatory requirements about the data involved in the computation or when the parties taking part to the exchange do not want to share secrets because of economic interests. In the remaining part of this section I will introduce some examples of actual implementations of Yao's protocol within the financial data analysis sector: needless to say, actual real-world applications calculate various functions between multiple parties, but many of them can be easily traced back to the computation of a joint sum (for example, a joint average is just a cumulative sum divided by the number of parties, which is obviously an information established in advance since the circuit must have a fixed size and shape known by all the partecipants).

3.1 Secure Financial Benchmarking

The process of comparing an organization's performance against its competitors, industry standards or historical data to identify areas for improvement and set performance targets is called *financial benchmarking*. Since sharing sensitive financial data with competitors and third-party benchmarking providers may clearly be the cause of several privacy concerns, by implementing a SMPC-based exchange organizations can easily compute performance metrics while easing the additional burden of ensuring data privacy. In particular, a study by Damgard et al. [5] proposes a practical implementation of a secure multi-party computation scheme (implementing Yao's protocol) that allows banks to perform secure benchmarking for various financial metrics, like loan default rates or operational costs, without exposing sensitive information. Such protocols enable organizations to identify best practices and enhance their operations while keeping the risk of data leakage at a minimum.

3.2 Secure Credit Scoring

Another recurrent process heavily (sensitive) data-dependent is *credit scoring*: banks and other financial institutions clearly require lots of private details to assess potential borrowers before lending money away. If they exploited SMPC-based techniques for this process, they would be able to compute very accurate credit evaluation of their customers without requiring the user to disclose personal information.

To showcase the benefits of such innovation within this field, the authors of [6] have proposed a secure scoring system that allows lenders to compute credit statistics without directly accessing borrowers' private information through the use of SMPC. By garbling the lender's private data within the input of Yao's circuit, the user is able to evaluate his credit score locally, without sharing personal details. In this way, neither the financial institution's, nor the borrower's information are disclosed. Such approach obviously is able to guarantee customers' privacy, while alleviating concerns regarding the misuse of personal information in the process. A further upgrade to this procedure has been proposed in [7]: the authors actually show that it is possible to enhance secure credit scoring using privacy-preserving data mining techniques (for further information on this topic's state of the art refer to [8]) to provide alternative input data sources to the credit assessment process while still respecting privacy regulations.

These secure procedures have the potential to clearly benefit all the parties involved: organizations are able to access supplementary information to compute more accurate estimates, while customers do not need to disclose private information.

4 S.E.L. considerations about SMPC

Yao's protocol can clearly offer financial data analysis many advantages, but there are still some ethical, legal, and social concerns that must be taken into account before blindly embracing such technology. In this section, I will briefly introduce some of the possible issues that may arise from an unresponsible use of SMPC.

4.1 Ethical Considerations

Even though Yao's protocol can help protecting sensitive data, it is essential that people do not let their guard down when adopting this technique: SMPC allows for the secure computation of a shared value, but it clearly does not ensure that the choice of function, nor the inputs provided are fair and balanced. A bank could easily garble into the circuit biased input values to influence a customer's credit score and inflate his interest rates and the user would not be able to determine whether the result was tampered or not. More generally, such problems can occur when unbalanced data is input into the process: if flawled datapoints are provided by the parties to the process, the computation could easily lead to a final result that may induce discriminatory outcomes. For example, if

the input data is biased towards certain demographic groups, the resulting analysis could result in investment decisions that unfairly benefit (or harm) said groups.

Additionally, parties with greater resources or computing power could potentially manipulate the protocol's outcome to their benefit, leading to unequal advantages for those in power (as explained in theorems 5 and 7 of [9] and section 1.2 of [10] Yao's scheme is secure with respect to a semi-honest threat model where the attacker does not deviate from the protocol, but more complex exchanges must be developed to embank active adversaries).

A potential risk associated to the deployment of Yao's protocol is thus clearly the false sense of security associated with it: this SMPC procedure can be considered truly safe and fair only when implemented along with a set of preventive measures established to control its execution.

4.2 Legal Considerations

Yao's protocol must also abide by the relevant data protection and privacy regulations, such as the General Data Protection Regulation (GDPR) in the European Union or California Consumer Privacy Act (CCPA) in the United States: even though the garbled-circuit exchange is inherently privacy-oriented, it is imperative to always ensure that sensitive data is managed with extreme caution. These regulations may impose specific requirements on how personal information should be handled even when encrypted to protect users' privacy and thus financial institutions must guarantee adherence to all applicable laws and regulations. For instance, a fallacious deployment of an instance of Yao's protocol could lead to the unintentional sharing of sensitive personal information such as income, investment strategies and financial goals: if customers do not give their explicit consent or the information is inadequately protected, data sharing could potentially violate said regulations.

To reduce privacy risks associated with the deployment of Yao's protocol in financial data analysis, any institution that implements it must ensure all personal data is anonymized and encrypted and appropriate consent and data security measures are in place.

4.3 Social Considerations

While Yao's protocol and other SMPC techniques can contribute to a societal shift towards increased data privacy and security, it is still essential to consider the potential impact on trust and transparency between the parties involved. Using SMPC in any financial process may raise concerns about data manipulation or cheating: any deviation from the agreed-upon protocols could easily lead to inaccurate or manipulated results. For example, within Yao's exchange the evaluator could easily communicate a result different from the one he locally computed and the garbler would not be able to discriminate it from an authentic outcome without additional measures into place.

Financial institutions may thus need to establish appropriate governance structures, cryptographic verification techniques, secure computing infrastructure, and auditing procedures as well as collaborate with trusted third parties for integrity assurance in the secure computation process. By implementing these measures, one can guarantee the integrity of computation and protect against cheating by any party involved during its execution.

5 Conclusion

After briefly introducing the dynamics of Yao's protocol, I have taken a look at some of its potential applications in financial data analysis that can allow organizations to collaborate on sensitive tasks without exposing private information. Examples such as secure financial benchmarking and credit scoring can give us a taste of its potential advantages; however, caution must still be exercised prior to implementation to ensure that deploying such secure exchanges does not introduce systematical discrimination, privacy violating mechanisms or opaqueness in the process.

References

[1] Yao, Andrew C.: Protocols for secure computations. In: *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)*, 1982, S. 160–164

[2] Yao, Andrew C.: How to generate and exchange secrets. In: *Proceedings of the 27th Annual Symposium on Foundations of Computer Science* (1986), S. 162–167

[3] Rabin, Michael: How To Exchange Secrets with Oblivious Transfer. In: *IACR Cryptology ePrint Archive* (1981), S. 187

[4] Zhao, Chuan ; Zhao, Shengnan ; Zhao, Minghao ; Chen, Zhenxiang ; Gao, Chong-Zhi ; Li, Hongwei ; Tan, Yu an: Secure Multi-Party Computation: Theory, practice and applications. In: *Information Sciences* 476 (2019), S. 357–372

[5] Damgard, Ivan ; Damgard, Kasper ; Nielsen, Kurt ; Nordholt, Peter S. ; Toft, Tomas: Confidential Benchmarking Based on Multiparty Computation. In: *Financial Cryptography and Data Security*, 2017, S. 169–187

[6] He, Haoran ; Wang, Zhao ; Jain, Hemant ; Jiang, Cuiqing ; Yang, Shanlin: A privacy-preserving decentralized credit scoring method based on multi-party information. In: *Decision Support Systems* 166 (2023), S. 113910

[7] Djeundje, Viani B. ; Crook, Jonathan ; Calabrese, Raffaella ; Hamid, Mona: Enhancing credit scoring with alternative data. In: *Expert Systems with Applications* 163 (2021)

[8] Aldeen, Yousra Abdul Alsahib S. ; Salleh, Mazleena ; Razzaque, Mohammad A.: A comprehensive review on privacy preserving data mining. In: *SpringerPlus* 4 (2015), Nr. 1, S. 694

[9] Lindell, Yehuda ; Pinkas, Benny: A Proof of Security of Yao's Protocol for Two-Party Computation. In: *Journal of Cryptology* 22 (2009), Nr. 2, S. 161–188

[10] Huang, Yan ; Evans, David ; Katz, Jonathan ; Malka, Lior: Faster Secure Two-Party Computation Using Garbled Circuits. In: *Proceedings of the 20th USENIX Conference on Security*, 2011, S. 35