622.755 – Introduction to Cybersecurity	Summer Term, 2022/23
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 protocol firstly introduced by Andrew Yao in 1982 [Yao82] and then re-elaborated in 1986 [Yao86] that enables two parties to jointly compute a function f on their private inputs x and y without disclosing any information about those inputs to each other. The protocol relies on the principle of *garbled circuits*, an algorithmic technique that encodes logical circuits while mantaining their inputs secret: to execute the protocol, one party (the garbler) encrypts the circuit that computes the desired function f along with its input x and sends it to an evaluator who uses the encoded data to locally compute f with his data y. This *Secure Multi Party Computation* (SMPC) exchange exploits *oblivious transfer*, a cryptographic exchange introduced by Rabin in an earlier paper [Rab81] to ensure that the two parties are able to evaluate $f(x, y)$ without knowing the each other's 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 Real-World Applications

In the implementation part of this project, I have realized a garbled circuit application that allows two parties 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 values. This can be especially helpful when financial institutions have to comply with strict regulatory
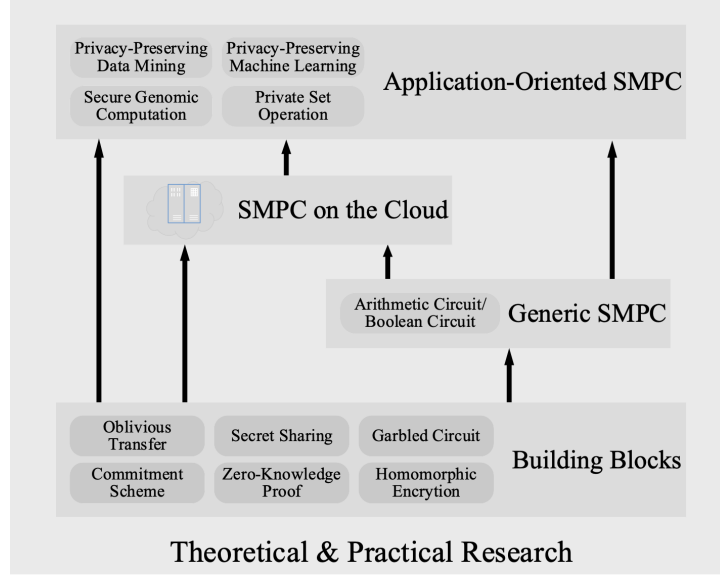
Figure 1: An overview of SMPC and its applications. Image taken from page 2 of [ZZZ+19].

requirements about the data involved in the computation or when the parties taking part to the exchange do not want to share secrets beacause of economic interests. In the remaining part of this section I will introduce some examples of actual implementations of Yao's protocol in the financial data analysis sector: of course, actual real-world applications compute 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 known in advance since the circuit must have a fixed size).

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 organisations can easily compute performance metrics while easing the additional burden of ensuring data privacy. For example, a study by Brickell and Shmatikov [BS08] proposes a practical implementation of secure multi-party computation (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 organisations to identify best practices and enhance their operations without risking sensitive information exposure implied by standard exchanges.

3.2 Secure Credit Scoring

Similarly, Yao's protocol can be profitably employed in financial data analysis to provide *secure credit scoring*: since SMPC-based protocols allow institutions to assess the risk associated with borrowing money to a particular customer without disclosing private information, Yao's protocol can clearly play a key role in streamlining the lending process.

Researchers have recently proposed the deployment of SMPC exchanges to create a secure scoring system that allows lenders to compute credit scores without directly accessing the borrowers' private information [HWJ+23]. Within this protocol, both parties (the borrower and the lender) jointly compute the result of a shared function using Yao's protocol, without sharing knowledge of private data. This approach helps safeguarding borrowers' privacy and alleviates concerns regarding the misuse of personal information in the process. A further upgrade to this procedure has been proposed in [DCCH21]: the authors actually show that it is possible to enhance secure credit scoring using privacy-preserving data mining [ASR15] 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: organisations are able to access supplementary information to compute more accurate estimates, while customers do not need to disclose private information.

4 Ethical, Legal and Social Considerations

Yao's protocol can clearly offer financial data analysis many advantages, but there are several ethical, legal, and social concerns that must be taken into account before blindly adopting such technology.

4.1 Ethical Considerations

Even though Yao's protocol can help protecting sensitive data, it is essential to guarantee that the deployment of this SMPC procedure does not cause unintended consequences for the end users such as discrimination or unfair treatment. A practical instance of such issue in secure credit scoring can occour when providing private data points as input leads to biased scores or prevent some borrowers from accessing credit. In particular, such problems can occur when unbalanced data is input into the process: if flawed datapoints provided by the parties could lead to a final result that produces 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 those groups.

Additionally, parties with greater resources or computing power could manipulate the protocol's outcome to their benefit, leading to unequal advantages for those in power. To address these concerns, it is essential to guarantee that the inputs provided to the protocol are representative and unbiased, with appropriate safeguards in place to prevent manipulation of its outcome.

A potential risk associated to the deployment of Yao's protocol is the false sense of security associated with it: this SMPC procedure can be considered truly safe only when implemented along with a set of preventive measures 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 data is inadequately protected, data sharing could easily 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. For example, using secure computation in financial benchmarking may raise concerns about data manipulation or cheating: any deviation from the agreed-upon inputs could easily lead to inaccurate or manipulated results. For instance, one party could provide false input data in order to influence the outcome of a computation to their advantage.

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, we 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

[ASR15]   Yousra Abdul Alsahib S. Aldeen, Mazleena Salleh, and Mohammad Abdur Razzaque. A comprehensive review on privacy preserving data mining. *SpringerPlus*, 4(1):694, 2015.

[BS08]   Justin Brickell and Vitaly Shmatikov. The cost of privacy: Destruction of data-mining utility in anonymized data publishing. In *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 70–78, 2008.

[BTW12]   Dan Bogdanov, Riivo Talviste, and Jan Willemson. Deploying secure multi-party computation for financial data analysis. In Angelos D. Keromytis, editor, *Financial Cryptography and Data Security*, pages 57–64, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg.

[DCCH21]   Viani B. Djeundje, Jonathan Crook, Raffaella Calabrese, and Mona Hamid. Enhancing credit scoring with alternative data. *Expert Systems with Applications*, 163:113766, 2021.

[HWJ+23]   Haoran He, Zhao Wang, Hemant Jain, Cuiqing Jiang, and Shanlin Yang. A privacy-preserving decentralized credit scoring method based on multi-party information. *Decision Support Systems*, 166:113910, 2023.

[Rab81]   Michael Rabin. How to exchange secrets with oblivious transfer. *IACR Cryptology ePrint Archive*, page 187, 01 1981.

[Yao82]   Andrew C. Yao. Protocols for secure computations. In *23rd Annual Symposium on Foundations of Computer Science (sfcs 1982)*, pages 160–164, 1982.

[Yao86]   Andrew Chi-Chih Yao. How to generate and exchange secrets. *Proceedings of the 27th Annual Symposium on Foundations of Computer Science*, pages 162–167, 1986.

[ZZZ+19]   Chuan Zhao, Shengnan Zhao, Minghao Zhao, Zhenxiang Chen, Chong-Zhi Gao, Hongwei Li, and Yu an Tan. Secure multi-party computation: Theory, practice and applications. *Information Sciences*, 476:357–372, 2019.