Multi-party Private Set Intersection (PSI) Using Multi-Key Fully Homomorphic Encryption (FHE)

Supervisor: Nesrine Kaaniche and Akram Bendoukha

Context: Multiple organizations often need to collaborate on sensitive data without exposing it to each other. A classic example is healthcare where hospitals may want to identify common patients for collaborative research while ensuring compliance with privacy regulations such as the GDPR [10]. Similarly, in finance, banks or insurance companies may want to detect fraudulent transactions that span multiple institutions [8].

Private Set Intersection (PSI) is a cryptographic primitive that enables a group of entities to compute the intersection of their private datasets while ensuring that each party learns *only* the intersection and nothing else about the others' data. This makes PSI a main building block of privacy-preserving collaborative analytics. An interesting application is Vertical Federated Learning (VFL), where organizations hold different features about the same users. PSI is then a necessary preprocessing step to align records across datasets before jointly training machine learning models [6].

However, existing PSI protocols face two major limitations: First, most protocols are optimized for *two-party* settings. When extending them to *n-party* scenarios, the computational and communication overhead grows rapidly, making them impractical for large datasets and multiple stakeholders. Second, many multi-party PSI protocols rely on MPC or partially trusted servers. This creates potential trust issues: if one participant or the server is compromised, private data could be exposed.

Recent advances in Fully Homomorphic Encryption (FHE) have introduced new opportunities for PSI, mainly in terms of expressiveness. FHE allows computations to be performed directly on encrypted data, meaning that sensitive records never need to be decrypted during processing [1]. However, like PSI, they are mostly limited to two-party scenarios [9,5]. They require heavy polynomial evaluations and frequent bootstrapping, leading to very high computation and communication costs.

To overcome these limitations, this project focuses on multi-key FHE, a paradigm where each participant encrypts their data under their own key, yet joint computations can be performed across all ciphertexts [3]. The result is decrypted collaboratively, ensuring that no single participant can decrypt data alone. Despite its potential, research on multi-key FHE-based PSI is very limited. For instance, Chen et al. [3] demonstrated that multi-key FHE for federated analytics, but did not address the specific requirements of set intersection, such as efficient matching, hashing, and polynomial encoding.

This project aims at:

- extending FHE PSI protocols to *n-party* scenarios using multi-key FHE,
- introducing efficient set representations and batching techniques for scalability,
- and achieving an optimal balance between computation and communication costs.

Deliverables:

- A formal specification of the designed multi-key -FHE-based PSI protocol
- A prototype implementation using libraries such as SEAL, PALISADE, or Lattigo

References:

1. Brakerski, Z., Gentry, C. and Vaikuntanathan, V., 2014. Leveled fully homomorphic encryption without bootstrapping. Innovations in Theoretical Computer Science (ITCS).
2. Braun, J., Hanzlik, L., Loss, J., and Pietrzak, K., 2022. Universally verifiable threshold MPC via class groups. Advances in Cryptology – EUROCRYPT 2022.
3. Chen, H., Laine, K., and Lauter, K., 2019. Multi-key homomorphic encryption for privacy-preserving federated analytics. IEEE Symposium on Security and Privacy (S&P) 2019.
4. Gong, J., Huang, L., Wang, R. and Wu, J., 2025. Authenticated Private Set Intersection: A Merkle Tree-Based Approach for Enhancing Data Integrity.
5. Halevi, S., Ishai, Y., Kushilevitz, E., Lindell, Y., Pinkas, B., 2022. Secure two-party set intersection with sublinear communication.
6. Hardy, S., Chen, B., and Smith, P., 2017. Private federated learning on vertically partitioned data via entity resolution and federated model training.
7. Kolesnikov, V., Kumaresan, R., Malkin, T. and Rosulek, M., 2016. Practical multi-party private set intersection from function secret sharing. ACM CCS 2016.
8. Nikolaenko, V., Weinsberg, U., Ioannidis, S., Joye, M., Boneh, D. and Taft, N., 2013. Privacy-preserving private set intersection for fraud detection. USENIX Security Symposium 2013.
9. Pinkas, B., Schneider, T., Zohner, M., 2019. Scalable private set intersection based on OT extension, EUROCRYPT 2019.
10. Shi, S., Chen, X., and Li, J., 2023. Privacy-preserving medical record linkage using secure set intersection protocols.
11. Urban, J. and Rambaud, M., 2024. Robust multiparty computation from threshold encryption based on RLWE.