Colby Rees
Reading & Research
Advisor: Professor Peihan Miao
Fall 2024

In Fall 2023, I enrolled in Professor Peihan Miao's class, CSCI 1510: Introduction to Cryptography and Computer Security. In this class, I was introduced to Private Set Intersection (PSI) which sparked my fascination with the wide range of possibilities surrounding secure communication. Led by the interest to further explore cryptography and better understand the process of research, I began Reading & Research with Peihan in the Spring 2024 semester where we decided to focus on PSI.

Spring 2024
In Spring 2024, our research focus was on PSI when applied to specific real-world problems. The paper that sparked our research question is titled, *Estimating Incidental Collection in Foreign Intelligence Surveillance: Large-Scale Multiparty Private Set Intersection with Union and Sum* by Anunay Kulshrestha and Jonathan Mayer. This paper focuses on the US government Section 702 of the Foreign Intelligence Act. The paper identifies the issue that Section 702 is not only collecting data from foreign adversaries but it is also collecting incidental data of those living in the US. The paper proposes using multiparty PSI to determine the number of incidental messages collected through Section 702. The existing protocol uses Diffie Hellman. For our research, we explored the research question, "Can we create a multiparty PSI with union and sum that is secure against a malicious adversary and more efficient than the one presented in the paper?".

My work consisted of adjusting and developing the protocol with different cryptographic primitives in order to result in a more efficient protocol that could potentially be used more widely than this specific government intelligence use case. I worked independently each week and worked collaboratively with Peihan on a weekly basis. I worked to use and implement cuckoo hashing, binary trees, and OKVS in different protocols in order to improve efficiency. In my research, we found that the binary tree approach did not work within this multiparty setting as there were difficulties when considering updatability when incorporating other parties. We decided to halt this research as we found that the protocols we were developing were not significantly different enough from existing work in order to produce meaningful results. For potential future work, this question can be re-explored in the context of a malicious adversary and the new state of the art PSI protocols like VOLE-PSI that are more efficient than Diffie Hellman.

Fall 2024
In Fall 2024, I continued researching with Peihan and now had the opportunity to conduct research with a fellow Computer Science Master's student, Dagar Rehan. The research question we focused on was, "Can we develop a PSI-Cardinality (PSI-CA) protocol that is as efficient as the state of the art plain PSI protocol?". In plain PSI, the two parties, Alice and Bob, learn all the items in their intersection. In PSI-CA, Alice and Bob only learn the cardinality of their intersection and no items are revealed to either party. Currently, the state of the art PSI protocol is VOLE-PSI. VOLE-PSI is presented in the paper titled, *VOLE-PSI: Fast OPRF and Circuit-PSI from Vector-OLE* by Peter Rindal and Phillipp Schoppman. The research on

VOLE-PSI reveals that the plain PSI protocol is 19 times faster than existing PSI-CA protocols. I worked independently through each week as well as collaboratively in weekly meetings with Dagar and weekly meetings with both Peihan and Dagar, in order to further explore this question.

My work began with a literature review on the most current PSI protocols as well as historical PSI protocols to provide context. This literature review continued throughout the semester as we encountered different complexities. I worked on protocol development which consisted in reworking the existing VOLE-PSI protocol framework to have Alice and Bob only learn the cardinality. This protocol development led to multiple adjustments. We adjusted the VOLE relationship to try multiple different configurations in order to have Alice or Bob need the other party in order to compute the intersection. This did not always work in the protocols because the different VOLE configurations were either not possible or they gave the parties too much information, leading to information being revealed. For future work concerning the VOLE relationship, I believe adjusting it is possible and should be further explored but it is important to keep in mind what elements of the relationship can feasibly be split apart.

We also made the decision we must use symmetric key operations in order to get as close to the state of the art PSI. We did try some variations of protocols that used an offline phase for more expensive operations but we departed from this idea as the final decryption in the online phase cost too much. As used in the VOLE-PSI protocol, we continued to use PaXoS within our protocol. For future work, I believe that using symmetric key operations is necessary for better efficiency.

The other main component of the protocols developed during this research semester was the incorporation of at least one Oblivious Programmable Pseudorandom Function (OPPRF). The purpose of the OPPRF was to have Bob program his values such that when Alice enters hers into the OPPRF, she would get back a set that was either the programed values or random but these would be indistinguishable to her. We tried multiple variations of protocols involving an OPPRF with different equations for how Bob constructs his values but we continuously ran into one of two problems. Either Alice had enough information to identify what items were in the intersection or neither Bob or Alice could determine the cardinality. Thus, PSI-CA was not being achieved.

As we ran into setbacks with incorporating the OPPRF and the potential configurations of the OPPRF, we also explored using linear maps and matrices. This would leverage that PaXoS is a linear system solver so we can maintain linearity. We ran into an issue with this approach because in order to be secure, the matrices would have to be very large which led to the protocol taking a significantly longer time with increased communication costs. We attempted methods to make matrices more efficient but could not find a way so we decided to no longer explore matrices. We also briefly considered function secret sharing but we did not explore incorporating this in depth because of time constraints. Secret sharing would be interesting to consider for future work.

These attempts brought us to the end of the semester. While we were not able to answer our research question of, “Can we develop a PSI-Cardinality (PSI-CA) protocol that is as efficient as the state of the art plain PSI protocol?”, I believe there is still important future work to be done. For future work, I think it would be very interesting to use the knowledge learned on VOLE-PSI and apply it to the multi-party PSI setting. I would be interested in exploring if malicious security can be achieved and how the costs compare to existing protocols.

Papers that influenced and provided guidance for Spring 2024:

Ben-Efraim, A., Nissenbaum, O., Omri, E., & Paskin-Cherniavsky, A. (2022). Psimple: Practical multiparty maliciously-secure private set intersection. *Proceedings of the 2022 ACM on Asia Conference on Computer and Communications Security*, 1098–1112.

Chandran, N., Dasgupta, N., Gupta, D., Obbattu, S. L. B., Sekar, S., & Shah, A. (2021). Efficient linear multiparty PSI and extensions to circuit/quorum PSI. *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, 1182–1204.

Kerschbaum, F., Blass, E.-O., & Mahdavi, R. A. (2022). Faster secure comparisons with offline phase for efficient private set intersection. *arXiv Preprint arXiv:2209. 13913*.

Kolesnikov, V., Kumaresan, R., Rosulek, M., & Trieu, N. (2016). Efficient batched oblivious PRF with applications to private set intersection. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, 818–829.

Kulshrestha, A., & Mayer, J. (2022). Estimating incidental collection in foreign intelligence surveillance: large-scale multiparty private set intersection with union and sum. *31st USENIX Security Symposium (USENIX Security 22)*, 1705–1722.

Nevo, O., Trieu, N., & Yanai, A. (2021). Simple, fast malicious multiparty private set intersection. *Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security*, 1151–1165.

Papers that influenced and provided guidance for Fall 2024:

Bienstock, A., Patel, S., Seo, J. Y., & Yeo, K. (2023). Near-optimal oblivious key-value stores for efficient PSI, PSU and volume-hiding multi-maps. *32nd USENIX Security Symposium (USENIX Security 23)*, 301–318.

Boyle, E., Couteau, G., Gilboa, N., & Ishai, Y. (2018). Compressing vector OLE. *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, 896–912.

Boyle, E., Gilboa, N., & Ishai, Y. (2015). Function secret sharing. *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, 337–367. Springer.

Garimella, G., Pinkas, B., Rosulek, M., Trieu, N., & Yanai, A. (2021). Oblivious key-value stores and amplification for private set intersection. *Advances in Cryptology--CRYPTO 2021: 41st Annual International Cryptology Conference, CRYPTO 2021, Virtual Event, August 16--20, 2021, Proceedings, Part II 41*, 395–425. Springer.

Pinkas, B., Rosulek, M., Trieu, N., & Yanai, A. (2020). PSI from PaXoS: fast, malicious private set intersection. *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, 739–767. Springer.

Raghuraman, S., & Rindal, P. (2022). Blazing fast PSI from improved OKVS and subfield VOLE. *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, 2505–2517.

Rindal, P., & Schoppmann, P. (2021). VOLE-PSI: fast OPRF and circuit-PSI from vector-OLE. *Annual International Conference on the Theory and Applications of Cryptographic Techniques*, 901–930. Springer.