Private Set Intersection from Pseudorandom Correlation Generators

Submitted for Journée C2 Crypto 2022

Submitter: Dung Bui[1] **Advisor:** Geoffroy Couteau[2]

[1] IRIF, Université de Paris
bui@irif.fr
[2] CNRS, IRIF, Université de Paris
couteau@irif.fr

*Private Set Intersection (*PSI*)* is a special, but important case of secure multi-party computation (MPC). It enjoys numerous real-life applications; for example, Google relies on PSI with third-party data providers to find target audiences for advertising and marketing campaigns [IKN+19] or contact discovery [CLR17]. Private set intersection allows parties to jointly compute the set of all common elements between the data sets of all parties. At the end of PSI protocol, one or both parties should get the correct intersection and will get nothing about the other's data sets outside the set intersection. Several techniques have been proposed that realize the PSI functionality, such as using public-key cryptography, constructing circuit-based generic techniques for secure computation, combining only oblivious transfer OT and the efficiency of symmetric cryptographic primitives. Recently, there has been an active line of work on efficient secure PSI protocol with fast implementations [PSZ14,KKRT16, RR17, KRTW19, PSWW18, PRTY19, PRTY20, CM20, RS21, RT21] that can process millions of items in seconds and makes PSI more practical.

*Pseudorandom Correlation Generator (*PCG*)*, is a primitive introduced in the work of Boyle *et al.* [BCG+19b, BCGI18, SGRR19, BCG+19a, CIK+20]. The goal of PCG is to compress long sources of correlated randomness without violating security. More concretely, the sender and receiver in a (two-party) PCG scheme hold pair of short correlated keys, and then they can locally expand these keys without interactions to obtain a pair of long correlated strings. In a recent line of work [BCGI18,BCG+19b,BCG+19a,CIK+20], Boyle *et al.* have designed multiple concretely efficient PCGs for specific correlations, such as vector oblivious linear evaluation (VOLE) or batch oblivious linear evaluation (BOLE). These primitives are at the heart of modern secure computation protocols with low communication overhead. The VOLE functionality allows a receiver to learn a secret linear combination of two vectors held by a sender and it was constructed (with sublinear communication) under variants of the syndrome decoding assumption.

In this work, we explore how the use of syndrome-decoding-based PCG can speed up private set intersection protocols. We devise two new protocols, aiming either at minimizing computation or communication, in both the semi-honest and the malicious setting. In our first protocol, we use cuckoo-hashing techniques to reduce a PSI on size-n datasets to a single VOLE on length-$1.2n$ vectors, and show how additional efficiency improvements can be achieved by relying instead on a variant of VOLE called *subfield* VOLE. Our protocol achieves the smallest communication accross all protocols that have linear computation, and we estimate that it achieves the best overall efficiency in a low latency setting. In our second protocol, we show how a PSI between size-n datasets can be reduced to a single OLE over the polynomial ring $\mathcal{R}_\mathsf{p} = \mathbb{F}_\mathsf{p}[\mathsf{x}]/\mathsf{F}(\mathsf{x})$, for an irreducible degree-$2n$ polynomial $F(x)$. It turns out that a PCG for this functionality was recently constructed by Boyle *et al.* (as a mean to achieve PCG for the BOLE functionality), under a ring variant of the syndrome decoding assumption. Our protocol achieves the smallest communication across all known PSI protocols for the database range size $2^{16} - 2^{20}$, with comparable computation as previous protocol, making it especially suitable in high latency settings. Another contribution should be constructing a leakage semi-honest PSI protocol which discloses negligible information of input set but we believe it achieves better efficiency in both communication and computation.

References

BCG+19a. E. Boyle, G. Couteau, N. Gilboa, Y. Ishai, L. Kohl, P. Rindal, and P. Scholl. Efficient two-round OT extension and silent non-interactive secure computation. In *ACM CCS 2019*, pages 291–308. ACM Press, November 2019.

BCG+19b. E. Boyle, G. Couteau, N. Gilboa, Y. Ishai, L. Kohl, and P. Scholl. Efficient pseudorandom correlation generators: Silent OT extension and more. In *CRYPTO 2019, Part III*, *LNCS* 11694, pages 489–518. Springer, Heidelberg, August 2019.

BCGI18. E. Boyle, G. Couteau, N. Gilboa, and Y. Ishai. Compressing vector OLE. In *ACM CCS 2018*, pages 896–912. ACM Press, October 2018.

CIK+20. G. Couteau, Y. Ishai, L. Kohl, E. Boyle, P. Scholl, and N. Gilboa. Efficient pseudorandom correlation generators from ring-lpn. Springer-Verlag, 2020.

CLR17. H. Chen, K. Laine, and P. Rindal. Fast private set intersection from homomorphic encryption. In *ACM CCS 2017*, pages 1243–1255. ACM Press, October / November 2017.

CM20. M. Chase and P. Miao. Private set intersection in the internet setting from lightweight oblivious PRF. In *CRYPTO 2020, Part III*, *LNCS* 12172, pages 34–63. Springer, Heidelberg, August 2020.

IKN+19. M. Ion, B. Kreuter, A. E. Nergiz, S. Patel, M. Raykova, S. Saxena, K. Seth, D. Shanahan, and M. Yung. On deploying secure computing: Private intersection-sum-with-cardinality. *Journal of Cryptology*, 2019. `https://eprint.iacr.org/2019/723`.

KKRT16. V. Kolesnikov, R. Kumaresan, M. Rosulek, and N. Trieu. Efficient batched oblivious PRF with applications to private set intersection. In *ACM CCS 2016*, pages 818–829. ACM Press, October 2016.

KRTW19. V. Kolesnikov, M. Rosulek, N. Trieu, and X. Wang. Scalable private set union from symmetric-key techniques. In *ASIACRYPT 2019, Part II*, *LNCS* 11922, pages 636–666. Springer, Heidelberg, December 2019.

PRTY19. B. Pinkas, M. Rosulek, N. Trieu, and A. Yanai. SpOT-light: Lightweight private set intersection from sparse OT extension. In *CRYPTO 2019, Part III*, *LNCS* 11694, pages 401–431. Springer, Heidelberg, August 2019.

PRTY20. B. Pinkas, M. Rosulek, N. Trieu, and A. Yanai. PSI from PaXoS: Fast, malicious private set intersection. In *EUROCRYPT 2020, Part II*, *LNCS* 12106, pages 739–767. Springer, Heidelberg, May 2020.

PSWW18. B. Pinkas, T. Schneider, C. Weinert, and U. Wieder. Efficient circuit-based PSI via cuckoo hashing. In *EUROCRYPT 2018, Part III*, *LNCS* 10822, pages 125–157. Springer, Heidelberg, April / May 2018.

PSZ14. B. Pinkas, T. Schneider, and M. Zohner. Faster private set intersection based on OT extension. In *USENIX Security 2014*, pages 797–812. USENIX Association, August 2014.

RR17. P. Rindal and M. Rosulek. Malicious-secure private set intersection via dual execution. In *ACM CCS 2017*, pages 1229–1242. ACM Press, October / November 2017.

RS21. P. Rindal and P. Schoppmann. Vole-psi: Fast oprf and circuit-psi from vector-ole. *IACR Cryptol. ePrint Arch.*, 2021:266, 2021.

RT21. M. Rosulek and N. Trieu. Compact and malicious private set intersection for small sets. Cryptology ePrint Archive, Report 2021/1159, 2021. `https://ia.cr/2021/1159`.

SGRR19. P. Schoppmann, A. Gascón, L. Reichert, and M. Raykova. Distributed vector-OLE: Improved constructions and implementation. In *ACM CCS 2019*, pages 1055–1072. ACM Press, November 2019.