Some Open Problems in Information-Theoretic Cryptography

Some Open Problems in Information-Theoretic Cryptography

Vinod Vaikuntanathan

MIT CSAIL, Cambridge, MA, USA
vinodv@csail.mit.edu

Abstract

Information-theoretic cryptography is full of open problems with a communication-complexity flavor. We will describe several such problems that arise in the study of private information retrieval, secure multi-party computation, secret sharing, private simultaneous messages (PSM) and conditional disclosure of secrets (CDS). In all these cases, there is a huge (exponential) gap between the best known upper and lower bounds. We will also describe the connections between these problems, some old and some new.

1998 ACM Subject Classification G.1.0 Mathematics of Computing

Keywords and phrases Cryptography, Information-Theoretic Security, Private Information Retrieval, Secret Sharing, Multiparty Computation

Digital Object Identifier 10.4230/LIPIcs.FSTTCS.2017.5

Category Invited Paper

1 Introduction

Information-theoretic cryptography deals with problems of secure communication and secure computation *against computationally unbounded adversaries.* While much of cryptography relies on unproven computational assumptions (and in particular, provides only conditional security), information-theoretic cryptography provides absolute guarantees that are independent of any computational assumption. As such, in the field of information-theoretic cryptography, one could hope to gain a complete understanding of the landscape of secure communication and computation, namely, classify which tasks are possible and which are not, and precisely quantify the computational and communication cost of security.

Indeed, Shannon's celebrated work [33] gave us such a complete picture for *secure communication* against unbounded adversaries: namely that the one-time pad is essentially the best one can do. While several influential works extended the basic model of secure communication to leverage environmental noise [35] or quantum effects [6] to accomplish information-theoretically secure communication in a larger range of settings, Shannon's characterizations gave us a clean and satisfying answer to the basic question.

The situation in *secure computation* turns out to be very different. Broadly speaking, secure computation [36, 22, 5, 10] deals with settings where two or more parties wish to communicate and compute a joint function f on their private inputs while revealing nothing to each other except the output of the computation. The primary complexity measure of secure computation protocols that we care about is their *communication complexity.*

Secure computation is replete with primitives and settings where there is an exponential gap between the known upper and lower bounds on communication complexity. In the basic setting of information theoretically secure 3-party computation that we describe in more detail below, the best protocols to compute an arbitrary function incur communication cost $2^{O(n)}$ where n is the total bit-length of the inputs of all parties, whereas the best lower bounds are linear in n. This leads us to ask:

What is the true communication overhead of secure computation?

Furthermore, all known approaches for information-theoretically secure computation incur a communication cost that is proportional to the computational cost of the function (in some computational model, say as a Boolean or arithmetic circuit). Thus, these approaches get stuck at the so-called *circuit-size barrier*. Yet another fundamental question in the foundations of information-theoretic cryptography is:

Does the communication cost of secure computation depend on the computational cost?

It is worth noting here that a simple counting argument establishes the *existence* of functions that require exponentially large circuits, but a similar statement is not known for the communication cost. That is, we do not even know whether there *exist* functions that require super-linear communication to securely compute.

In the rest of this extended abstract, we describe a number of objects of interest in information-theoretic cryptography – private information retrieval, secure multiparty computation, private simultaneous messages, conditional disclosure of secrets, and secret sharing – and the relations among them, as well as the open problems associated to these objects.

2 Information-Theoretic Primitives and their Problems

Private Information Retrieval (PIR). PIR is a protocol among one or more non-communicating servers each holding the same database D, thought of as an N-bit string, and a user holding an index $i \in [N]$. The user wishes to retrieve the i-th bit $D[i]$ from the server(s), without revealing any information about i. Clearly, the server(s) can rather inefficiently accomplish this by sending the entire string D to the user. The objective of PIR, then, is to achieve this goal while communicating (significantly) less than N bits. Such PIR schemes are deemed non-trivial.

Chor, Goldreich, Kushilevitz and Sudan [11], who first defined PIR, also showed that non-trivial *single-server* PIR schemes (with communication less than N bits) require computational assumptions. One line of research that resulted from this work showed several constructions of single-server PIR with decreasing communication complexity under various cryptographic assumptions [27, 9, 28, 7, 20, 19, 8], culminating in [8] that achieves the asymptotically optimal communication complexity of $O(\log N + \lambda)$ bits where λ is the cryptographic security parameter.

Chor, Goldreich, Kushilevitz and Sudan also proposed the natural setting of multi-server PIR where two or more *non-communicating* servers each holding the same database D interact with the client holding an index i. The client is guaranteed that its index is private as long as the servers do not collude with each other.

In the spirit of our questions, let us mention here that the best two-server PIR protocols have total communication complexity $2^{\tilde{O}(\sqrt{\log N})}$ [15] while the lower bound is "merely" $(5 - o(1)) \log N$ [31, 34]. We refer the reader to the excellent survey of Yekhanin and the bibliography maintained by Gasarch [37, 17] for pointers to the long line of work on information-theoretically secure multi-server PIR protocols. Curiously, such PIR schemes turn out to be equivalent in a precise sense to locally decodable codes, an object that does not refer to privacy at all [26].

Secure Multiparty Computation (MPC). In the setting of MPC, a collection of k parties wish to collaborate to compute a publicly known function f on their respective inputs $x_1, x_2, \ldots, x_k$ without leaking any information to each other except the output $f(x_1, \ldots, x_k)$. Such a protocol should be secure against collusions of t corrupted parties. It is well-known that 2-party secure computation of even simple functions requires computational assumptions even against a semi-honest adversary, thus for our purposes, the simplest setting to think about is the 3-party setting with a single corrupted party.

Canonical ways of achieving MPC go through some explicit computational representation of the function f either as a circuit [5, 10] or as a branching program [23]. Consequently, there are functions f for which such MPC protocols have communication exponential in the (total) input length simply because, by a counting argument, there are functions f which require exponentially large circuits (resp., branching programs). Must this be the case?

The lower bounds on the communication complexity of MPC are few and far between. To the best of our knowledge, the state of the art is the work of Data, Prabhakaran and Prabhakaran [13] who show a $1.5n - o(n)$ lower bound for 3-party secure computation (in the presence of a single corrupted party) where n is the total input length of all the parties. Since insecure computation can be achieved by communicating just n bits, this shows that *achieving security has its price.* However, here again, there is a large gap in communication between known protocols and lower bounds.

Interestingly, the two problems we just discussed, namely PIR and MPC, turn out to be equivalent to each other. A beautiful result of Ishai and Kushilevitz [24] tells us that any k-server PIR protocol gives us a $(k+1)$-party secure computation protocol tolerating a single corruption with nearly the same communication complexity, and vice versa. In particular, in the case of 3 parties, the PIR protocol of [15] gives us a 3-party computation protocol for arbitrary functions with communication $2^{\tilde{O}(\sqrt{n})}$ where n is the total input size.

Private Simultaneous Messages (PSM). Feige, Kilian and Naor [16] considered multiparty computation in a very structured and clean model called the private simultaneous messages (PSM) model (inspired by the simultaneous messages model of Babai, Kimmel and Lokam [2]). In the two-party PSM setting, Alice has an input x and Bob has input y, and they share a common random string which is unknown to the outside world. They send a single message each to a referee called Charlie. In turn, after receiving these messages, Charlie should be able to learn $f(x, y)$ (for a fixed, publicly known, function f) but nothing else about x or y. Since neither Bob nor Alice receive any additional information in the course of the protocol, they learn nothing about each other's input.

Again, there are large gaps between lower and upper bounds in this model. Feige, Kilian and Naor showed that there are functions that require $1.5n - o(n)$ bits of communication in this model (where n is the total input length of Alice and Bob). In a recent work, Beimel, Ishai, Kumaresan and Kushilevitz [4] showed that every function f has a PSM protocol with communication $2^{n/4}$ where n again is the total input length. Narrowing this gap is an important open problem.

Conditional Disclosure of Secrets (CDS). Two-party conditional disclosure of secrets (CDS) first defined by Gertner, Ishai, Kushilevitz and Malkin [21] is an important special case of PSM: two parties want to disclose a secret to a third party if and only if their respective inputs satisfy some fixed predicate ϕ. Concretely, Alice holds x, Bob holds y and in addition, they both hold a secret $m \in \{0, 1\}$ (along with some additional private randomness w). Charlie knows both x and y but not m; Alice and Bob want to disclose m to Charlie iff

$\phi(x, y) = 1$. How many bits do Alice and Bob need to communicate to Charlie?

This is a very simple and natural model where non-private computation requires very little communication (just a single bit), whereas the best upper bound for private computation is exponential. Indeed, in the non-private setting, Alice or Bob can send m to Charlie, upon which Charlie computes $\phi(x, y)$ and decides whether to output m or $\perp$. This trivial protocol with one-bit communication is not private because Charlie learns m even when the predicate is false. In contrast, in the private setting, we have a big gap between upper and lower-bounds. The best upper bound we have for CDS for general predicates ϕ requires that Alice and Bob each transmits $2^{\tilde{O}(\sqrt{n})}$ bits [29]. This upper bound works by translating a special type of PIR protocol into a CDS scheme. Indeed, the communication complexity of $2^{\tilde{O}(\sqrt{n})}$ is closely related to that of the Dvir-Gopi 2-server PIR scheme.

The best known lower bound is $\Omega(\log n)$ [18, 1] which is a double-exponential factor away from the upper bound! A central open problem is to narrow this gap; a concrete question in this direction is to improve the lower bound to $\Omega(n)$ even for a non-explicit function. On this note, we remark that [18] show an $\Omega(\sqrt{n})$ lower bound for the inner product predicate for special type of CDS protocols where Charlie's reconstruction algorithm is a linear function computed on the messages of Alice and Bob.

One could of course consider multi-party versions of both PSM and CDS. Recently, [30] showed a CDS protocol that achieves the same complexity of $2^{\tilde{O}(\sqrt{n})}$ even for arbitarily many parties.

Secret Sharing. The classical problem of secret sharing [32], more precisely non-threshold secret sharing [25], is closely related to multiparty CDS (see, e.g., [30] for more details on this connection). For general non-threshold secret-sharing schemes, the best upper bounds on the (individual) share size are exponential in the number of parties n, namely $2^{n-o(n)}$, whereas the best lower bounds are nearly linear [12], namely $\Omega(n/\log n)$ (see Beimel's survey [3] for more details).

In summary, there is a rich landscape of problems in information-theoretic cryptography, all closely related to secure computation, where there is a large gap between known upper and lower bounds on their communication complexity. Furthermore, there are non-trivial relations between all these problems. For example, MPC and PIR are equivalent modulo computational considerations [24]; PSM is a special type of MPC with a restricted communication pattern; multi-server PIR protocols of a special form give us CDS protocols with an equivalent communication complexity [29]; and multi-party CDS protocols are closely related to secret sharing. Despite recent progress [4, 14, 1, 29, 30], much remains to be uncovered in this world, eventually leading us to a better understanding of the question: *What is the true communication overhead of secure computation?*

References

1 Benny Applebaum, Barak Arkis, Pavel Raykov, and Prashant Nalini Vasudevan. Conditional disclosure of secrets: Amplification, closure, amortization, lower-bounds, and separations. *IACR Cryptology ePrint Archive*, 2017:164, 2017. URL: `http://eprint.iacr.org/2017/164`.

2 László Babai, Peter G. Kimmel, and Satyanarayana V. Lokam. Simultaneous messages vs. communication. In *STACS*, pages 361–372, 1995.

3 Amos Beimel. Secret-sharing schemes: A survey. In *Coding and Cryptology – Third International Workshop, IWCC 2011, Qingdao, China, May 30-June 3, 2011. Proceedings*, pages 11–46, 2011.

4 Amos Beimel, Yuval Ishai, Ranjit Kumaresan, and Eyal Kushilevitz. On the cryptographic complexity of the worst functions. In *TCC*, pages 317–342, 2014.

5 Michael Ben-Or, Shafi Goldwasser, and Avi Wigderson. Completeness theorems for non-cryptographic fault-tolerant distributed computation (extended abstract). In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, pages 1–10, 1988.

6 C. H. Bennett and G. Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems, and Signal Processing*, page 175, 1984.

7 Dan Boneh, Eu-Jin Goh, and Kobbi Nissim. Evaluating 2-dnf formulas on ciphertexts. In *Theory of Cryptography, Second Theory of Cryptography Conference, TCC 2005, Cambridge, MA, USA, February 10-12, 2005, Proceedings*, pages 325–341, 2005.

8 Zvika Brakerski and Vinod Vaikuntanathan. Efficient fully homomorphic encryption from (standard) LWE. In Rafail Ostrovsky, editor, *IEEE 52nd Annual Symposium on Foundations of Computer Science, FOCS 2011, Palm Springs, CA, USA, October 22-25, 2011*, pages 97–106. IEEE Computer Society, 2011. `doi:10.1109/FOCS.2011.12`.

9 Christian Cachin, Silvio Micali, and Markus Stadler. Computationally private information retrieval with polylogarithmic communication. In Jacques Stern, editor, *Advances in Cryptology – EUROCRYPT'99, International Conference on the Theory and Application of Cryptographic Techniques, Prague, Czech Republic, May 2-6, 1999, Proceeding*, volume 1592 of *Lecture Notes in Computer Science*, pages 402–414. Springer, 1999. `doi:10.1007/3-540-48910-X_28`.

10 David Chaum, Claude Crépeau, and Ivan Damgård. Multiparty unconditionally secure protocols (extended abstract). In *Proceedings of the 20th Annual ACM Symposium on Theory of Computing, May 2-4, 1988, Chicago, Illinois, USA*, pages 11–19, 1988.

11 Benny Chor, Eyal Kushilevitz, Oded Goldreich, and Madhu Sudan. Private information retrieval. *J. ACM*, 45(6):965–981, 1998. `doi:10.1145/293347.293350`.

12 László Csirmaz. The size of a share must be large. *J. Cryptology*, 10(4):223–231, 1997.

13 Deepesh Data, Manoj Prabhakaran, and Vinod M. Prabhakaran. On the communication complexity of secure computation. In Juan A. Garay and Rosario Gennaro, editors, *Advances in Cryptology – CRYPTO 2014 – 34th Annual Cryptology Conference, Santa Barbara, CA, USA, August 17-21, 2014, Proceedings, Part II*, volume 8617 of *Lecture Notes in Computer Science*, pages 199–216. Springer, 2014. `doi:10.1007/978-3-662-44381-1_12`.

14 Deepesh Data, Vinod M. Prabhakaran, and Manoj M. Prabhakaran. Communication and randomness lower bounds for secure computation. *IEEE Trans. Information Theory*, 62(7):3901–3929, 2016. `doi:10.1109/TIT.2016.2568207`.

15 Zeev Dvir and Sivakanth Gopi. 2-server PIR with sub-polynomial communication. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing, STOC 2015, Portland, OR, USA, June 14-17, 2015*, pages 577–584, 2015.

16 Uriel Feige, Joe Kilian, and Moni Naor. A minimal model for secure computation (extended abstract). In *Proceedings of the Twenty-Sixth Annual ACM Symposium on Theory of Computing, 23-25 May 1994, Montréal, Québec, Canada*, pages 554–563, 1994.

17 William Gasarch. Web page on private information retrieval. `http://www.cs.umd.edu/~gasarch/TOPICS/pir/pir.html`.

18 Romain Gay, Iordanis Kerenidis, and Hoeteck Wee. Communication complexity of conditional disclosure of secrets and attribute-based encryption. In *CRYPTO (II)*, pages 485–502, 2015.

19 Craig Gentry. Fully homomorphic encryption using ideal lattices. In Michael Mitzenmacher, editor, *Proceedings of the 41st Annual ACM Symposium on Theory of Computing, STOC*

2009, Bethesda, MD, USA, May 31 – June 2, 2009, pages 169–178. ACM, 2009. doi:10.1145/1536414.1536440.

20 Craig Gentry and Zulfikar Ramzan. Single-database private information retrieval with constant communication rate. In Luís Caires, Giuseppe F. Italiano, Luís Monteiro, Catuscia Palamidessi, and Moti Yung, editors, *Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Lisbon, Portugal, July 11-15, 2005, Proceedings*, volume 3580 of *Lecture Notes in Computer Science*, pages 803–815. Springer, 2005. doi:10.1007/11523468_65.

21 Yael Gertner, Yuval Ishai, Eyal Kushilevitz, and Tal Malkin. Protecting data privacy in private information retrieval schemes. *J. Comput. Syst. Sci.*, 60(3):592–629, 2000. doi:10.1006/jcss.1999.1689.

22 Oded Goldreich, Silvio Micali, and Avi Wigderson. How to play any mental game or A completeness theorem for protocols with honest majority. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing, 1987, New York, New York, USA*, pages 218–229, 1987.

23 Yuval Ishai and Eyal Kushilevitz. Randomizing polynomials: A new representation with applications to round-efficient secure computation. In *41st Annual Symposium on Foundations of Computer Science, FOCS 2000, 12-14 November 2000, Redondo Beach, California, USA*, pages 294–304, 2000.

24 Yuval Ishai and Eyal Kushilevitz. On the hardness of information-theoretic multiparty computation. In *Advances in Cryptology – EUROCRYPT 2004, International Conference on the Theory and Applications of Cryptographic Techniques, Interlaken, Switzerland, May 2-6, 2004, Proceedings*, pages 439–455, 2004.

25 Mitsuru Ito, Akira Saito, and Takao Nishizeki. Secret sharing scheme realizing general access structure. *Electronics and Communications in Japan (Part III: Fundamental Electronic Science)*, 72(9):56–64, 1989.

26 Jonathan Katz and Luca Trevisan. On the efficiency of local decoding procedures for error-correcting codes. In *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, May 21-23, 2000, Portland, OR, USA*, pages 80–86, 2000.

27 Eyal Kushilevitz and Rafail Ostrovsky. Replication is NOT needed: SINGLE database, computationally-private information retrieval. In *38th Annual Symposium on Foundations of Computer Science, FOCS'97, Miami Beach, Florida, USA, October 19-22, 1997*, pages 364–373. IEEE Computer Society, 1997. doi:10.1109/SFCS.1997.646125.

28 Helger Lipmaa. An oblivious transfer protocol with log-squared communication. In Jianying Zhou, Javier Lopez, Robert H. Deng, and Feng Bao, editors, *Information Security, 8th International Conference, ISC 2005, Singapore, September 20-23, 2005, Proceedings*, volume 3650 of *Lecture Notes in Computer Science*, pages 314–328. Springer, 2005. doi:10.1007/11556992_23.

29 Tianren Liu, Vinod Vaikuntanathan, and Hoeteck Wee. Conditional disclosure of secrets via non-linear reconstruction. In *Advances in Cryptology – CRYPTO 2017 – 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part I*, pages 758–790, 2017.

30 Tianren Liu, Vinod Vaikuntanathan, and Hoeteck Wee. Towards breaking the exponential barrier for general secret sharing. Cryptology ePrint Archive, Report 2017/1062, 2017. https://eprint.iacr.org/2017/1062.

31 Eran Mann. Private access to distributed information. In *Master's thesis, Technion – Israel Institute of Technology*, 1998.

32 Adi Shamir. How to share a secret. *Commun. ACM*, 22(11):612–613, 1979. doi:10.1145/359168.359176.

33 C. Shannon. A mathematical theory of communication. *Bell system technical journal*, 27, 1948.

34 Stephanie Wehner and Ronald de Wolf. Improved lower bounds for locally decodable codes and private information retrieval. In *Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Lisbon, Portugal, July 11-15, 2005, Proceedings*, pages 1424–1436, 2005.

35 A. D. Wyner. The wire-tap channel. *The Bell System Technical Journal*, 54(8):1355–1387, Oct 1975. `doi:10.1002/j.1538-7305.1975.tb02040.x`.

36 Andrew Chi-Chih Yao. Protocols for secure computations (extended abstract). In *23rd Annual Symposium on Foundations of Computer Science, Chicago, Illinois, USA, 3-5 November 1982*, pages 160–164, 1982.

37 Sergey Yekhanin. *LDCs and PIRs*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2007. URL: `http://hdl.handle.net/1721.1/42242`.