

**Massachusetts Institute of Technology
Department of Electrical Engineering
and Computer Science**

**Proposal for Thesis Research in Partial
Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy**

TITLE: Relationships Between Functionality, Security, and Privacy
SUBMITTED BY: Rio LaVigne
32 Vassar Street, G636
Cambridge, MA 02139

(SIGNATURE OF AUTHOR)
DATE OF SUBMISSION: December 12, 2019
EXPECTED DATE OF COMPLETION: June 2020
LABORATORY: Theory of Computation

BRIEF STATEMENT OF THE PROBLEM:

One of the core goals of cryptography is to be able to offer security and privacy without sacrificing functionality. This thesis describes three different functionalities and notions of security and privacy and how they interplay. First, we delve into Topology-Hiding Computation, which has a very restrictive definition of privacy making it difficult to construct; nevertheless we provide a two constructions. Second, we invent a new cryptographic concept called Property Preserving Hashes, where we want to compress inputs while securely preserving a certain property between them, e.g. hamming distance. Finally, we explore Fine-Grained Cryptography, and develop a public key cryptosystem. In this notion of cryptography, security takes on a much less restrictive role (e.g. adversaries must run in $O(n^{100})$ time), but the protocols and security reductions must run in “fine-grained” time, e.g. less than $O(n^2)$.

Massachusetts Institute of Technology
Department of Electrical Engineering
and Computer Science
Cambridge, Massachusetts 02139

Doctoral Thesis Supervision Agreement

TO: Department Graduate Committee
FROM: Professor Vinod Vaikuntanathan

The program outlined in the proposal:

TITLE: Relationships Between Functionality, Security, and Privacy
AUTHOR: Rio LaVigne
DATE: December 12, 2019

is adequate for a Doctoral thesis. I believe that appropriate readers for this thesis would be:

OTHER SUPERVISOR: Professor Virginia Vassilevska Williams
READER 1: Doctor Tal Moran

Facilities and support for the research outlined in the proposal are available.
I am willing to supervise the thesis and evaluate the thesis report.

SIGNED: _____
ASSOCIATE PROFESSOR OF ELECTRICAL ENGINEERING
AND COMPUTER SCIENCE AT MIT

DATE: _____

Comments:

Massachusetts Institute of Technology
Department of Electrical Engineering
and Computer Science
Cambridge, Massachusetts 02139

Doctoral Thesis Supervision Agreement

TO: Department Graduate Committee
FROM: Professor Virginia Vassilevska Williams

The program outlined in the proposal:

TITLE: Relationships Between Functionality, Security, and Privacy
AUTHOR: Rio LaVigne
DATE: December 12, 2019

is adequate for a Doctoral thesis. I believe that appropriate readers for this thesis would be:

OTHER SUPERVISOR: Professor Vinod Vaikuntanathan
READER 1: Doctor Tal Moran

Facilities and support for the research outlined in the proposal are available.
I am willing to supervise the thesis and evaluate the thesis report.

SIGNED: _____
ASSOCIATE PROFESSOR OF ELECTRICAL ENGINEERING
AND COMPUTER SCIENCE AT MIT

DATE: _____

Comments:

Massachusetts Institute of Technology
Department of Electrical Engineering
and Computer Science
Cambridge, Massachusetts 02139

Doctoral Thesis Reader Agreement

TO: Department Graduate Committee
FROM: Doctor Tal Moran

The program outlined in the proposal:

TITLE: Relationships Between Functionality, Security, and Privacy
AUTHOR: Rio LaVigne
DATE: December 12, 2019
SUPERVISOR: Professor Vinod Vaikuntanathan
OTHER SUPERVISOR: Professor Virginia Vassilevska Williams

is adequate for a Doctoral thesis. I am willing to aid in guiding the research
and in evaluating the thesis report as a reader.

SIGNED: _____
FACULTY OF SCHOOL OF COMPUTER SCIENCE
AT IDC HERZLIYA

DATE: _____

Comments:

1 Overview

One of the core goals of cryptography is to be able to offer security and privacy without sacrificing functionality. To do this, cryptographers define notions of both correctness and security. We can then build systems that we prove satisfy both of these definitions. Correctness states that the system provides the functionality we want, while security is essentially the notion that we can effectively hide information from an adversary. Depending on the notion of security, this adversary may be all powerful (information-theoretic or computationally unbounded), may run only in polynomial time (computationally bounded), or even bounded by a specific polynomial runtime like $O(n^2)$ (a fine-grained adversary). This thesis will focus on realizing three different functionalities, each one hiding different information, and three different notions of security.

In the first part of this thesis, we discuss results in the area of Topology Hiding Computation (THC). THC is a generalization of secure multiparty computation. In this model, parties are in an incomplete but connected communication network, each party with its own private inputs. The functionality these parties want to realize is evaluating some function (computation) over all of their inputs. On the security and privacy side, these parties want to hide both their private inputs and who their neighbors are. So, the goal is for these parties to run a protocol, communicating only with their neighbors, so that by the end of the protocol, they learn the output of the function on their inputs and *nothing else*, including information about what the communication network looks like other than their own neighborhood.

It turns out that THC is a difficult notion to realize, mired in impossibility results. We circumvent two impossibility results in THC in two different ways. First, there is an impossibility result for THC stating that it is impossible to design THC against adversaries that can “turn off” parties during the protocol [MOR15]. In essence, an adversary that has this power can always learn something about the graph. So, instead of designing a protocol that would attempt to leak no information, we designed a protocol to limit the amount of information leaked as much as possible. We ran into similar impossibility results when considering a model more similar to the real-world: channels between parties now have unknown delays on them, and an adversary may be able to control the delay. We had to find a model that mirrored this real-world complication, and this was also not impossible to achieve.

In the second part of the thesis, we focus on Adversarially Robust Property Preserving Hashes (PPH). PPHs are a generalization of collision resistant hash functions (CRHFs). Recall that a CRHF is a family of hash functions $\mathcal{H} = \{h : \{0, 1\}^n \rightarrow \{0, 1\}^m\}$ that is compressing ($m < n$), and has the property that no Probabilistic Polynomial-Time (PPT) adversary given h can find two inputs $x_1 \neq x_2$ such that $h(x_1) = h(x_2)$ except with negligible probability. Looking at CRHFs from a slightly different perspective, their functionality is to preserve the equality predicate between two inputs while compressing them, preventing an adversary from coming up with inputs where the equality predicate is not preserved (even though these inputs exist, they are hard to find). A PPH is

also a compressing function that preserves some other property. For example, we focus on the property of “gap-hamming” distance: if two inputs are close in hamming distance, then their hashes are also close, and if two inputs are far, their hashes are also far. Note that this example is of a promise predicate since we do not care about inputs that are neither close nor far.

Property preserving hashes were a new cryptographic primitive that we designed, and so it was important to understand what was possible and what was not in our new model. We found strong connections between One-Way Communication Complexity (OWC complexity) and our primitive. Fortunately for us, OWC is a rich, well-studied area of complexity. In many cases, we could not directly port OWC results to our setting, but we could use their proof techniques and get lower bounds. Once we had these lower bounds, we had a much better sense of what was and was not possible, and could construct PPHs more easily.

In the final chapter, for a different perspective, we design a public-key cryptography functionality against weak adversaries. This functionality allows for a party to publish a public key while they keep the secret key, and any other party to use that public key to encrypt a message that only the party with the secret key can decrypt, *hiding* that message from any eavesdroppers. Unlike in standard models for cryptography, however, our eavesdroppers much less powerful: their runtime is bounded by an explicit polynomial instead of “probabilistic polynomial-time”, e.g. can only run in time $O(n^{100})$. This, of course, is only interesting as a cryptographic notion if the adversary has the same or more time to run than the honest parties. So, we get a notion of cryptography we call *fine-grained*: for examples, honest parties might only need to run in $O(n^2)$ time, while any adversary running in time $O(n^{100})$ time still cannot glean any useful information with some probability. Motivating the study of this cryptography is the fact that it does not rely on any of the normal cryptographic assumptions, including $P \neq NP$, $BPP \neq NP$, or even the assumption that one-way functions exist.

However, due to its fine-grained nature many standard cryptographic reductions (like Goldreich-Levin hardcore bits [GL89]) no longer work. And so, in this thesis we demonstrate variations that can work in our setting, including a notion of fine-grained one-way functions, fine-grained hardcore-bits, a fine-grained key exchange, and finally fine-grained public-key cryptography.

2 Results

In this thesis, we explore these three different notions of hiding while preserving a functionality:

1. hide private inputs and network topology while preserving computation,
2. hide as much information as possible while preserving a property between inputs,

3. and hide messages from a weak adversary while preserving the communication and fine-grained runtime.

Topology-Hiding Computation To address the first perspective, we explore the realm of topology-hiding computation (THC). THC is a generalization of secure multiparty computation (MPC), where we hide not only each party’s input, but also the communication graph; parties know who their neighbors are, but learn nothing else about the structure of the graph. In 2017, we showed that THC is possible against a very weak adversary [ALM17], but there were still many open questions surrounding the nature of THC.

- **Fail-stop Adversaries.** Fail-stop adversaries can turn off parties during the protocol. As shown by Moran, Orlav, and Richelson in 2015 [MOR15], it is impossible to achieve perfect topology hiding when giving the adversary this power. The next question is how to get around such an impossibility while still providing meaningful privacy and security. Since the adversary was going to learn something in any protocol we could devise, we decided to use a model that would incorporate this leakage.

Our results here were a definition of security that included access to a “leakage oracle” and a protocol that used this oracle in the security proof. This oracle would spit out yes or no answers to specific queries, leaking one bit of information at a time. In the security reduction of our protocol, a simulator would only need to call this oracle at most once, leaking at most one bit of information to an adversary. Concretely, we could bound the probability that the simulator would have to call the leakage oracle, meaning that we would leak only a polynomial-fraction of a bit. Our protocol ran in $\tilde{O}(\kappa n^4/\rho)$ rounds, where κ is the security parameter, n is the number of parties, and ρ is the probability that we leak a bit.

These results are detailed in our paper at TCC 2018 [LZM⁺18].

- **Probabilistic Delay Model.** This model was the result of first trying to build THC in a purely asynchronous setting, only to prove that it was impossible to do so. So, to get around this without just reducing ourselves to the fail-stop case, we designed a new model that more closely resembled an asynchronous, real-world setting.

More formally, each edge between parties in the Probabilistic Delay Model has its own distribution over delays. Each time a party sends a message to another party, that delay is sampled *independently*. This simulates real world traffic having different delays as it gets sent from one server to another.

Our results also included two new constructions in this model. The first is a construction that works under standard cryptographic assumptions, but only on trees, cycles, or graphs with small circumference. The second is a construction that requires secure hardware boxes/tokens, but works for arbitrary communication graphs. Although this is a very powerful

assumption, there are still many challenges involved since an adversarial party can query this hardware at any point with any input.

These results are currently in submission.

Adversarially Robust Property Preserving Hashes and One-Way Communication Lower Bounds Addressing the second perspective, we look at a new cryptographic primitive: adversarially robust Property Preserving Hashes (PPH). The inspiration here is two-fold. First, collision-resistant hash functions (CRHFs) are a staple in cryptography: no PPT adversary can produce an $x_1 \neq x_2$ such that $h(x_1) = h(x_2)$ except with negligible probability over h sampled from the CRHF family and coins of the adversary, and hence one can preserve the “equality” predicate on compressed inputs even against adversarially chosen inputs. We want to compress inputs while maintaining some property other than equality between them, even in the presence of adversarially chosen inputs. The properties we considered were those that already had non-robust constructions, in particular locality-sensitive hashing (LSH) [IM98]. An LSH family is a hash function family $\mathcal{H} = \{h : \{0,1\}^n \rightarrow \{0,1\}^m\}$ that compresses inputs ($m < n$) and is *mostly* correct. For example, say we were preserving a gap- ℓ_p norm for a gap between r and cr for some constant $c > 1$ (note that in our case, gap-hamming is equivalent to gap- ℓ_0 norm). Then, we have a threshold τ so that for any pair $x_1, x_2 \in \{0,1\}^n$

- if $\|x_1 - x_2\|_p < r$, $\|h(x_1) - h(x_2)\|_p < \tau$, and
- if $\|x_1 - x_2\|_p > cr$ then $\|h(x_1) - h(x_2)\|_p \geq \tau$

with high probability over our choice of h sampled from \mathcal{H} .

We can use almost the same language to define PPHs for some property $P : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}$, except our probability will be negligible and taken over both our sampled hash function and the coins of a PPT adversary. We also generalize the predicate evaluation: instead of using the same predicate on hashed values, we can use an “evaluation function” called **Eval**. So, even in the presence of (PPT) adversarially chosen values of x_1 and x_2 ,

- if $P(x_1, x_2) = 0$, then $\text{Eval}(h(x_1), h(x_2)) = 0$, and
- if $P(x_1, x_2) = 1$, then $\text{Eval}(h(x_1), h(x_2)) = 1$

with all but negligible probability. In other words, for any PPT adversary given h sampled from \mathcal{H} , the probability that the adversary produces x_1 and x_2 such that $P(x_1, x_2) \neq \text{Eval}(h(x_1), h(x_2))$ is negligible.

The link between robust PPH and LSH was very clear, but less obvious was the link between PPH and one-way communication (OWC): in essence, compressing an input as much as possible for a OWC protocol is akin to compressing an input as much as possible while preserving some property of that input as we would want to do for a hash function. These connections led to the following results, which will be included in the thesis.

- OWC is a rich field with a lot of work detailing lower bounds on the complexity of certain predicates. These lower bounds *almost* directly translated to impossibility results for PPHs. However, we needed to be careful because OWC lower bounds dealt with constant error, whereas cryptographers, we care about negligible error. We characterized these OWC lower bounds and how they mapped to PPH lower bounds. We showed that a class of predicates we call “reconstructing predicates” could not have PPHs. This class includes useful predicates like GreaterThan, Index, and ExactHamming.
- With these lower bounds and previous results from LSH in mind, we turned to constructing a PPH for Gap-Hamming: the promise predicate where we can distinguish between pairs of points that are $(1 - \epsilon)d$ *close* in hamming distance or $(1 + \epsilon)d$ *far*. For all pairs within the gap, we “don’t care” how our PPH behaves. We build two constructions for this predicate, each with different drawbacks and benefits. Our first construction relies only on collision-resistant hashing. Our second uses a new assumption related to the hardness of syndrome decoding [AHI⁺17].
- Rounding out this paper and its myriad of definitions, we demonstrate that even for very simple predicates, we require collision resistance, and even if we weaken our main definition of a robust PPH, we still need one-way functions.

These results were published in ITCS 2019 [BLV19].

In unpublished follow-up research with Cyrus Rashtchian, we explored the intricacies of the CRHF method for constructing gap-hamming PPHs. We fleshed out the relationship between d , the “center” of our gap, and the size of the hash function output. We also discussed how to use this kind of construction to get a gap- ℓ_1 -norm PPH and found that with standard methods of transforming ℓ_1 into ℓ_0 , it could not work.

Fine-Grained Cryptography and Barriers Addressing the final perspective, this thesis will discuss our work in Fine-Grained cryptography. With my coauthors Lincoln and Vassilevska Williams, we built the first fine-grained key exchange built from fine-grained assumptions. The premise was to explore what cryptography could be like in “Pessiland,” a world in which there are no cryptographic one-way functions (which also implies no public key cryptography). We looked at this through the lens of fine-grained complexity, using both assumptions and reduction techniques from that field.

While designing cryptography for this setting, we came across multiple barriers. The first was that some fine-grained problems would not lend themselves to build cryptography without refuting NSETH [CGI⁺16]. So, we would need to use a fine-grained problem that was not associated with that barrier. Another form of barrier we ran into was worst-case to average-case reductions for problems that would make for good cryptography. Even now, the only worst-case to average-case reductions for fine-grained problems are for counting versions of

these problems. Unfortunately, counting-style problems do not lend themselves well to building cryptography, as there are no longer small enough witnesses. Finally, many standard cryptographic reductions no longer work in a fine-grained world, since they take a non-trivial polynomial amount of time, and so a challenge we faced was to ensure all of our reductions were fine-grained and build different types of primitives based off of these restricted reductions.

Despite these difficulties, we achieved the following results:

- Assuming that an average-case version of Zero- k -Clique requires $n^{k-o(1)}$ time, we constructed a non-interactive key exchange that required honest parties to run in $O(N)$ time and an adversary to run in time $\tilde{\Omega}(N^{1.5-\epsilon})$, where ϵ decreases with respect to k .¹
- Generalize the properties of Zero- k -Clique to construct this key exchange: plantable, list-hard, and splittable.
- Define and construct fine-grained hard-core bits.

This work was published in CRYPTO 2019 [LLW19].

In unpublished follow-up work, we also show how to use another property of Zero- k -Clique to construct a key exchange where the adversary now needs $\Omega(N^{2-\epsilon})$ time. This N^2 gap is the best we are able to do since we base these constructions on Merkle puzzles [BM09].

3 Timeline

This thesis will have an introduction, conclusion, and three other chapters making up the body of the work. The main tasks left to finishing this thesis are compiling the results for the three main topics, crafting an introduction and conclusion, and compiling acknowledgements. My timeline is as follows:

- A detailed outline and draft of an introduction will be finished by January 2020.
- January and February of 2020 will be devoted to compiling the results in all three topics and finishing the introduction. This includes polishing unpublished results from property preserving hashing and fine-grained cryptography, and producing a rough draft.
- March of 2020 will be used to write the conclusion and assemble acknowledgements.
- April to early May of 2020 will be spent tying the topics together and incorporating my thesis committee’s feedback.
- Defend early May 2020.

¹The tilde in $\tilde{\Omega}$ ignores any *subpolynomial* factors.

References

- [AHI⁺17] Benny Applebaum, Naama Haramaty, Yuval Ishai, Eyal Kushilevitz, and Vinod Vaikuntanathan. Low-complexity cryptographic hash functions. In *8th Innovations in Theoretical Computer Science Conference, ITCS 2017, January 9-11, 2017, Berkeley, CA, USA*, pages 7:1–7:31, 2017.
- [ALM17] Adi Akavia, Rio LaVigne, and Tal Moran. Topology-hiding computation on all graphs. In *Advances in Cryptology - CRYPTO 2017 - 37th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 20-24, 2017, Proceedings, Part I*, pages 447–467, 2017.
- [BGI⁺01] Boaz Barak, Oded Goldreich, Russell Impagliazzo, Steven Rudich, Amit Sahai, Salil P. Vadhan, and Ke Yang. On the (im)possibility of obfuscating programs. In *Proceedings of the 21st Annual International Cryptology Conference on Advances in Cryptology, CRYPTO '01*, pages 1–18, London, UK, UK, 2001. Springer-Verlag.
- [BGR95] Mihir Bellare, Roch Guérin, and Phillip Rogaway. XOR macs: New methods for message authentication using finite pseudorandom functions. In *Advances in Cryptology - CRYPTO '95, 15th Annual International Cryptology Conference, Santa Barbara, California, USA, August 27-31, 1995, Proceedings*, pages 15–28, 1995.
- [BKR94] Mihir Bellare, Joe Kilian, and Phillip Rogaway. The security of cipher block chaining. In *Advances in Cryptology - CRYPTO '94, 14th Annual International Cryptology Conference, Santa Barbara, California, USA, August 21-25, 1994, Proceedings*, pages 341–358, 1994.
- [BLV19] Elette Boyle, Rio LaVigne, and Vinod Vaikuntanathan. Adversarially robust property-preserving hash functions. In *10th Innovations in Theoretical Computer Science Conference, ITCS 2019, January 10-12, 2019, San Diego, California, USA*, pages 16:1–16:20, 2019.
- [BM09] Boaz Barak and Mohammad Mahmoody-Ghidary. Merkle puzzles are optimal - an $O(n^2)$ -query attack on any key exchange from a random oracle. In *Advances in Cryptology - CRYPTO 2009, 29th Annual International Cryptology Conference, Santa Barbara, CA, USA, August 16-20, 2009. Proceedings*, pages 374–390, 2009.
- [CGI⁺16] Marco L. Carmosino, Jiawei Gao, Russell Impagliazzo, Ivan Mihajlin, Ramamohan Paturi, and Stefan Schneider. Nondeterministic extensions of the strong exponential time hypothesis and consequences for non-reducibility. In *Proceedings of the 2016 ACM Conference on Innovations in Theoretical Computer Science, Cambridge, MA, USA, January 14-16, 2016*, pages 261–270, 2016.

- [GL89] O. Goldreich and L. A. Levin. A hard-core predicate for all one-way functions. In *Proceedings of the Twenty-first Annual ACM Symposium on Theory of Computing*, STOC '89, pages 25–32, New York, NY, USA, 1989. ACM.
- [IM98] Piotr Indyk and Rajeev Motwani. Approximate nearest neighbors: Towards removing the curse of dimensionality. In *Proceedings of the Thirtieth Annual ACM Symposium on the Theory of Computing, Dallas, Texas, USA, May 23-26, 1998*, pages 604–613, 1998.
- [LLW19] Rio LaVigne, Andrea Lincoln, and Virginia Vassilevska Williams. Public-key cryptography in the fine-grained setting. *IACR Cryptology ePrint Archive*, 2019:625, 2019.
- [LZM⁺18] Rio LaVigne, Chen-Da Liu Zhang, Ueli Maurer, Tal Moran, Marta Mularczyk, and Daniel Tschudi. Topology-hiding computation beyond semi-honest adversaries. In *Theory of Cryptography - 16th International Conference, TCC 2018, Panaji, India, November 11-14, 2018, Proceedings, Part II*, pages 3–35, 2018.
- [MOR15] Tal Moran, Ilan Orlov, and Silas Richelson. Topology-hiding computation. In *Theory of Cryptography - 12th Theory of Cryptography Conference, TCC 2015, Warsaw, Poland, March 23-25, 2015, Proceedings, Part I*, pages 159–181, 2015.