Cong Zhang | Research Statement

(215)-313-8948 • congresearch@gmail.com

Overview

I am broadly interested in theoretical and applied cryptography with particular emphasis on outsourcing encrypted databases, verifiable computation and black-box separations. In particular, my future research interests are shaped by proposing new security notions schemes for outsourcing encryption, building efficient and verifiable computational protocols, and constructing schemes from weak models/assumptions or proving the impossibility results.

My past research focuses on order-revealing encryption (ORE) and indifferentiability, and my dissertation combines my three results on order-revealing encryption, where I propose the first meaningful security notion that is independent of the leaked ciphertexts, called *parameter-hiding* and build an ORE scheme achieving our notion only based on bi-linear map. Moreover, I give two impossibility results to indicate that you cannot build efficient ORE schemes that achieves stronger security notions. Additionally, I also study the work on indifferentiability and I bring it into the realm of public key cryptosystem. Essentially, I propose a new primitive—*ideal PKE*, and build it only from random oracle model and standard model assumptions. In this statement, I briefly summarize the contributions of my prior research, and detail my future research plans.

Prior Research

In this part, I'll roughly illustrate my prior research on order-revealing encryption and indifferentiability, respectively.

• An order-revealing encryption is a symmetric encryption scheme that gives a public procedure by which two ciphertexts can be compared to reveal the order of their underlying plaintexts. ORE enables encrypted range queries in a legacy-compatible form. Moreover, efficient, blockcipher-based constructions are known. ORE was has seen practical deployment as well as application in several research projects, including CryptDB. However, a series of works, starting with Naveed *et al.* (CCS 2015), have shown that when the adversary has a good estimate of the distribution of the data, ORE provides little protection, thus it is an open problem to design efficient ORE with meaningful security.

From the security and privacy perspective, it's natural to start with the strongest notion—ideal leakage, which means that, given the ciphertexts, *only* the order is revealed. Despite this notion provides the best security, the only known constructions of ideal ORE are based on cryptographic multilinear maps and are currently too impractical for real-world applications. In our work, we give evidence that building ideal leakage ORE from weaker tools may be hard. Indeed, we give black-box separations between ideal ORE and most symmetric-key primitives, as well as public key encryption and anything else implied by generic groups in a black-box way.

From the performance efficiency perspective, we'd prefer to build schemes using weak tools. While, due to our impossibility result, we have to make a relaxation on the security notion. Thus, we propose

a new security notion—*parameter-hiding*, which, to the best of my knowledge, is the strongest notion except for the ideal one. In our definition, we consider the case that the database entries are drawn identically and independently from a distribution of known shape, but for which the mean and variance are not(and thus the attacks of Naveed et al. do not apply), and we say an ORE is parameter-hiding, if for any probabilistic and polynomial time adversary, given a sequence of ciphertexts, the mean and variance of the message distribution are hidden. Then we give a construction of ORE that satisfying our new notion from bi-linear maps.

• When designing a cryptographic system, it is difficult to predict how it will be used in practice and what security properties will be required of it. Cryptographers have devised different security models to capture each of the scenarios above and more, each requiring different constructions to satisfy. However, seldom are these different security models considered in tandem, meaning that each application scenario may require a different scheme. Even worse, there are many potential security models that have yet to be considered; after all, it is difficult to predict the various applications devised by software developers that may deviate from the existing provably secure uses. With those issues in mind, our goal is to develop a *single* construction for a given cryptographic concept that simultaneously captures any reasonable security property and can be composed to work in any reasonable larger protocol. As such, only a single instance of the scheme needs to be developed and then deployed in a variety of use cases, even those that have not yet been discovered.

In our work, we define and construct the first ideal public key cryptosystems such as public key encryption. By being ideal, our schemes will immediately satisfy a wide class of security properties, including most studied in the literature. Our proof is based on the indifferentiability framework, under standard model assumptions, in random oracle model.

After that, we give a following-up work, in which, we attempt to build an indifferentiable generic group model from random oracle model and standard model assumptions, via our previous techniques. However, we found a fundamental barrier, and we turn this barrier into a separation result, which means that, it is impossible to build an indifferentiable generic group model from random oracle and any computational assumptions (even non-falsifiable assumptions). Moreover, we extend our result to the separation between generic bi-linear group model and generic group model.

Future Research Plan

I am interested in various theoretical and applied cryptography problems and in the future, I will continue to work on my ongoing research and explore new topics in cryptography and computer security. The following is my research plan:

• Efficient ORE-based protocol against stronger adversary. Due to the attacks on ORE, we know that, even ideal-leakage ORE protects little if the adversary has a good estimate on the message distribution. In my dissertation, I propose a potential solution for it, where in the encryption procedure, the algorithm samples some fake message and encrypt both the real message and the fake messages. The point is that, if the adversary never knows which one is the real one and which one is the fake one, then it cannot sort the ciphertext sequence, as a result, the attacks would fail. However, there are several challenging problems: 1) define a new security notion against such a strong adversary; 2) how to sample the fake points such that our protocol can achieve this new notion. It would be very interesting to have a new security notion for outsourcing encrypted database and build systems that

achieve it.

• Indifferentiability for other primitives. In our work, we bring indifferentiability into the public key setting, where we build an indifferentiable PKE and Signature. One interesting direction is to extend it to other primitives, such as functional encryption or interactive protocols (OT).

• Separation between tri-linear group model and bi-linear group model. In our work, the best separation we can achieve is between bi-linear group model and generic group model, and our technique seems insufficient to show a better separation. On the other hand, there is no good tri-linear map candidate yet, and it would be very exciting if we can give this separation.

• Verifiable protocols. When it comes to cloud computation, one natural question is *how to prevent malicious cloud server without a trusted party*. I aim to design a protocol that can efficiently detect whether the server/non-colluding servers are doing the proper work for some restricted class of computation, such as counting queries.

• Watermarking in neural network. This is a new topic I just start, and I aim to add a watermarking into a trained neural network to provide better efficiency and authenticity.

This plan roughly lists my on-going research and the work I am going to explore recently, and I am interested in various crypto problems and I am flexible to join in other research projects, such as security/privacy in machine learning or cryptocurrency.