Infrastructures to collect, manage, and use data are being deployed worldwide. Major tech companies that use personal information for profit are peeking at all aspects of our lives. While classical cryptography protects the transportation of users' data, more powerful cryptographic tools are needed now since privacy evasion and security vulnerability are often pervasive and unavoidable. My research focuses on pushing the practicality of modern cryptographic protocols to secure real-world systems. This involves making protocols more efficient, secure, capable, and flexible to changing requirements. Most of my projects are mixtures of both theory and systems, where the greatest challenge, and also the fun, is to balance provable security and practical considerations. I approach it from two angles.

1. I explore the foundation of modern cryptographic protocols from a practical perspective. This involves designing optimizations for improved efficiency, innovating new protocol paradigms, and better understanding their security.
2. I bring mature cryptographic tools to various fields within and beyond computer science (e.g., formal methods, machine learning, databases, and health informatics) and, at the same time, bring new challenges from these fields back to the cryptography community. The systems I helped build have attracted huge attention; some are being deployed in the real world.

My long-term research goal is to expand the set of practically deployable cryptography tools and build more systems with real-world impacts.

Accomplishments. My group has made tremendous progress in applied cryptographic protocols and their applications in various fields. As one metric, I have produced over 60 publications (more than 40 since 2019) in peer-reviewed conferences and journals with more than 4500 citations and an H-index of 33. My results have been published in cryptography conferences (2 Crypto, 5 Eurocrypt, 6 Asiacrypt), security conferences (14 CCS, 7 S&P, 4 USENIX Security), ML/AI-related conferences (5 ICLR, ICML, CVPR, AAAI), database conferences (3 VLDB, SIGMOD), etc. These results were awarded a CCS best paper in 2017 (0.6% of all submissions), a CCS best paper runner-up in 2021 (1.6%), a CCS distinguished paper in 2022 (0.5%), an ICLR notable top 5% in 2023, and an ICLR spotlight in 2024 (5%). My research has been funded by NSF, AFRL, DARPA, and the industry (JPMorgan, Google, Meta, Chainlink Labs, and PlatON Network), bringing more than four million dollars of support to Northwestern. Members of my lab have received fellowships from IBM, JPMorgan, and NSF; senior members are joining Arizona State University (Chenkai Weng, as an assistant professor), UIUC (Ning Luo, as an assistant professor), and University of Toronto (Olive Franzese, as a post-doc researcher).

Founding the Deployment of Modern Cryptography

Theoretically, modern cryptographic protocols are highly capable with suitable hardness assumptions. However, when considering practical deployment, many aspects need to be considered beyond merely the feasibility, including efficiency, security, protocol complexity, and details of how they fit particular settings. My research focuses on refining the foundations of modern cryptographic protocols to better prepare them for deployment. We open-source most of our research outcomes as part of the EMP toolkit. As a result, the use of EMP has flourished in the past few years, becoming one of the most prominent tools in the field.

Yet more efficient MPC. Secure multi-party computation (MPC) allows for distributed computation of private inputs and is being pushed for deployment by many companies, where efficiency is still the main obstacle. *My research has been long exploring the frontiers of practically efficient MPC protocols from every aspect.* First, I focus on the efficiency of fundamental tools needed in MPC. My works push the efficiency of essential MPC building blocks, including improved protocols for oblivious transfer [5, 38], distributed-point function [18], and correlated randomness [32]. In particular, the Ferret protocol [38] has been extensively used by the community, including Meta, which used it to deploy private advertisement measurement. Using these tools on top of the authenticated garbling framework [2, 1, 19] that I proposed, I further closed the gap between active and passive security in the context of constant-round MPC with all-but-one corruption [9, 37], one of the harshest settings. As an ongoing effort, I'm leading a team to prepare for a NIST standardization of threshold cryptography that uses many protocols mentioned above. If standardized, it would further accelerate their deployment. I also pay specific attention to designing large-scale MPC protocols for more parties [26] and bigger data [39] to meet practical demands. Finally, I work on MPC-based protocols for specific applications, where there are more opportunities to improve the protocol based on the function being computed and the deployment settings. For example, my recent work that customized MPC to facilitate TLS payload authentication [34] has been built into a browser extension and serviced more than 170,000 users globally.

A new approach to scalable ZKP. Another essential cryptographic tool is zero-knowledge proof (ZKP), which enables a player to prove to other parties about predicates of its secret information without revealing it. Traditionally, most results are in the context of succinct and non-interactive ZKPs, which is suitable for use cases where one person wants to prove to a large set of players asynchronously. These protocols are widely applicable but often hit resource bottlenecks in CPU and memory quickly as the statements become large. As part of the DARPA SIEVE program, *I initiated the development of a new type of interactive ZK with high scalability and concrete efficiency* [32, 4]. Although the communication is often linear to the circuit size, it is most suitable for private auditing and other use cases where the set of verifiers is fixed and small. With a cost of less than 5 US dollars, our recent work, QuickSilver [35], can compute thousands of billions of gates efficiently! It has attracted Chainlink to use our protocol to prove payload data in blockchain applications. More recently, we further extended it for mix-mode statements [31] and private RAM accesses [11], succinctness [33], and non-interactiveness [36]. This line has also attracted many researchers to improve its efficiency and capability, and to use it for proving large statements.

Revisiting the security of modern cryptographic protocols. The urgent need for modern cryptographic protocols has significantly advanced their efficiency, but sometimes too quickly. As a result, some protocols are being used without rigorous security analysis. *My research raises awareness and provides solutions to security issues at all levels of modern cryptography, including the underlying hardness assumption, the low-level symmetric-key primitives, and the high-level protocol design.* First, I study the use of symmetric-key constructions in MPC. In a series of works [14, 13, 15, 16], we found that many building blocks (including garbled circuits, GGM tree, etc.) are proven secure only in the asymptotic sense (i.e., a security loss of $\kappa^c/2^\kappa$ for some constant c) while protocol parameters are selected assuming an optimal concrete bound (i.e., assuming $c = 0$). This could lead to devastating attacks when deployed in practice. Fortunately, we show a solution that leverages modern CPU instructions with almost no slowdown and near-optimal concrete security. In addition, I'm also exploring practically efficient protocols with various notions of adaptive security for extra flexibility in using the protocol. I designed new protocols for adaptively secure oblivious transfer [6] and non-interactive ZKP [7] and, even better, proved that popularly deployed garbled circuit schemes already satisfy more desirable properties and thus can be used in a more flexible way [17]. Finally, I'm also exploring cryptoanalysis of important assumptions like Learning Parity with Noise in the context of its recent popularity in MPC applications [22]. After studying more individual protocols, my ultimate goal is to provide frameworks for analyzing refined security properties more easily (or even automatically).

Long-term vision. My long-term goal is to lead the development of next-generation efficient cryptography protocols. In addition to continuing the next steps in my research portfolio, I would like to understand the efficiency limit of cryptographic protocols, especially with advanced cryptographic assumptions mostly associated with relatively low efficiency. Additionally, I would like to explore how other resources like hardware customization, massive parallelization, and even a "simplistic" quantum device can bring efficiency gains beyond conventional settings.

Cryptography-Enabled Systems

I am actively building cryptography-based systems tailored for various applications by tackling two problems: defining the most relevant and meaningful security requirements for the application and designing customized protocols utilizing the unique features of each application.

Cryptography-Enabled Formal Methods. Formal methods can model complex objects (like programs and configurations) using mathematical tools. *As a pioneer in this intersection, I aim to enable the modeling of private objects for sensitive tasks using cryptography.* In recent works, I've been building efficient ZKP systems for problems beyond NP, which was much unexplored in practical efficiency. Although solving problems beyond NP typically requires high computational costs, they are crucial and used extensively in modeling program safety. In past works, I designed ZKP systems for the unsatisfiability of Boolean formulas [24] and SMT formulas [23]. ZKUNSAT is the first system to prove the safety of real programs (e.g., Linux drivers) privately. Our ongoing work is further exploring ZKP of PSPACE problems and succinct proofs for knowledge beyond NP. On the other hand, I am also exploring other intersections between formal methods and cryptography, including using MPC to enable joint solving of SAT formula [25] and using formal methods for optimized distribution of ZKP parallelization [27].

Cryptography-Enabled Machine Learning. The availability of ample data is the foundation of all machine learning (ML) systems, but high-quality datasets are often sensitive. *My work provides better tradeoffs between data privacy and utility using cryptography* by addressing two challenges: 1) defining rigorous and cryptographically achievable safety properties for ML systems, and 2) efficiently supporting the high-intensity computation in common ML systems in cryptography. First, I'm exploring the use of cryptography to build secure systems for decentralized learning [8]. We integrate cryptography with federated learning to bring in additional properties, including robustness [12] and differential privacy (DP) [10], to the federated learning system without extra trust assumptions. In an ongoing project, we are designing customized schemes that are friendly to the decentralized generation of DP noises even over hundreds of participants. Additionally, I'm exploring using cryptographic proof systems for privacy-preserving auditing of ML systems. In my prior work [31], we constructed a ZKP system that can prove correct inference on an unmodified ResNet-101 model, but the efficiency is still far from practical needs. In my recent works, we use properties of ML models to simplify what needs to be proven cryptographically such that even the fairness [29] and DP properties [28] of a model can be proven efficiently.

Cryptography-Enabled Databases. I have worked extensively at the intersection of cryptography and databases. *My main focus is to use cryptography to provide rigorous security to modern database systems.* I have been working on applying approximation and randomization to accelerate SQL [3] and quantile [20] queries over federated databases, where multiple data providers would like to make queries over the union of every participant's private databases. By using approximation and randomization, one can significantly reduce the amount of computation needed in MPC. I have also worked to use ZKP to build a system that allows a data owner to prove facts represented as SQL queries to other parties without revealing the database itself [21]. Our key insight is that SQL is inherently relational and friendly to set representations; on the other hand, ZKP for set operations can be made highly efficient. This provides a nice interface between SQL operations and ZKP protocols, leading to a system that can prove even TPC-H benchmarking queries efficiently! Additionally, I'm working in a team to pilot the deployment of SQL-over-MPC systems over a set of Chicago hospitals [30] serving tens of millions of patients. In a recent DARPA-funded project, we focused on user-friendly interfaces and smooth deployment so that the details of the cryptographic protocols are completely hidden the curtains from health informatics.

Long-term vision. In addition to the above three directions I will continue pursuing, I would also like to establish new territories where cryptography can have a high social impact. This includes journalism (e.g., combating misinformation using cryptography), law (e.g., understanding how cryptographic protocols fit into existing legal frameworks), and medicine (e.g., private and collaborative diagnosis).

References

[1] Xiao Wang, Samuel Ranellucci, and Jonathan Katz. Global-Scale Secure Multiparty Computation. In *ACM Conference on Computer and Communications Security (CCS)*, 2017.

[2] Xiao Wang, Samuel Ranellucci, and Jonathan Katz. Authenticated Garbling and Efficient Maliciously Secure Two-Party Computation. In *ACM Conference on Computer and Communications Security (CCS)*, 2017, **Best Paper Award**.

[3] Johes Bater, Yongjoo Park, Xi He, Xiao Wang, and Jennie Rogers. SAQE: Practical Privacy-Preserving Approximate Query Processing for Data Federations. In *Proceedings of the VLDB Endowment (PLVDB)*, 2020.

[4] Carsten Baum, Samuel Dittmer, Peter Scholl, and Xiao Wang. SoK: Vector OLE-Based Zero-Knowledge Protocols. In *Designs, Codes and Cryptography*, 2023.

[5] Ran Canetti, Pratik Sarkar, and Xiao Wang. Blazing Fast OT for Three-Round UC OT Extension. In *Public-Key Cryptography (PKC)*, 2020.

[6] Ran Canetti, Pratik Sarkar, and Xiao Wang. Efficient and round-optimal oblivious transfer and commitment with adaptive security. In *Annual International Conference on the Theory and Application of Cryptology and Information Security (Asiacypt)*, 2020.

[7] Ran Canetti, Pratik Sarkar, and Xiao Wang. Triply Adaptive UC NIZK. In *Annual International Conference on the Theory and Application of Cryptology and Information Security (Asiacrypt)*, 2022.

[8] Christopher A. Choquette-Choo, Natalie Dullerud, Adam Dziedzic, Yunxiang Zhang, Somesh Jha, Nicolas Papernot, and Xiao Wang. CaPC Learning: Confidential and Private Collaborative Learning. In *International Conference on Learning Representations (ICLR)*, 2021.

[9] Hongrui Cui, Xiao Wang, Kang Yang, and Yu Yu. Actively Secure Half-Gates with Minimum Overhead under Duplex Networks. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques (Eurocrypt)*, 2023.

[10] Adam Dziedzic, Christopher A. Choquette-Choo, Natalie Dullerud, Vinith Suriyakumar, Ali Shahin Shamsabadi, Muhammad Ahmad Kaleem, Somesh Jha, Nicolas Papernot, and Xiao Wang. Private Multi-Winner Voting for Machine Learning. In *Proceedings on Privacy Enhancing Technologies (PoPETs)*, 2023.

[11] Nicholas Franzese, Jonathan Katz, Steve Lu, Rafail Ostrovsky, Xiao Wang, and Chenkai Weng. Constant-Overhead Zero-Knowledge for RAM Programs. In *ACM Conference on Computer and Communications Security (CCS)*, 2021.

[12] Olive Franzese, Adam Dziedzic, Christopher A. Choquette-Choo, Mark R. Thomas, Muhammad Ahmad Kaleem, Stephan Rabanser, Congyu Fang, Somesh Jha, Nicolas Papernot, and Xiao Wang. Robust and Actively Secure Serverless Collaborative Learning. In *Conference on Neural Information Processing Systems (NeurIPS)*, 2024.

[13] Chun Guo, Jonathan Katz, Xiao Wang, Chenkai Weng, and Yu Yu. Better Concrete Security for Half-Gates Garbling (in the Multi-Instance Setting). In *International Cryptology Conference (CRYPTO)*, 2020.

[14] Chun Guo, Jonathan Katz, Xiao Wang, and Yu Yu. Efficient and Secure Multiparty Computation from Fixed-key Block Ciphers. In *IEEE Symposium on Security and Privacy (S&P)*, 2020.

[15] Chun Guo, Francois-Xavier Standaert, Weijia Wang, Xiao Wang, and Yu Yu. Provable Security of SP Networks with Partial Non-Linear Layers. In *IACR Transactions on Symmetric Cryptology (ToSC)*, 2021.

[16] Chun Guo, Xiao Wang, Xiang Xie, and Yu Yu. The Multi-user Constrained PRF Security of Generalized GGM Trees for MPC and Hierarchical Wallets . In *ACM Transactions on Privacy and Security*, 2023.

[17] Xiaojie Guo, Kang Yang, Xiao Wang, Yu Yu, and Zheli Liu. Unmodified Half-Gates is Adaptively Secure - So is Unmodified Three-Halves. 2023. https://eprint.iacr.org/2023/1528.

[18] Xiaojie Guo, Kang Yang, Xiao Wang, Wenhao Zhang, Xiang Xie, Jiang Zhang, and Zheli Liu. Half-Tree: Halving the Cost of Tree Expansion in COT and DPF. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques (Eurocrypt)*, 2023.

[19] Jonathan Katz, Samuel Ranellucci, Mike Rosulek, and Xiao Wang. Optimizing Authenticated Garbling for Faster Secure Two-Party Computation. In *International Cryptology Conference (CRYPTO)*, 2018.

[20] Xiao Lan, Hongjian Jin, Hui Guo, and Xiao Wang. Efficient and Secure Quantile Aggregation of Private Data Streams. In *IEEE Transactions on Information Forensics and Security (TIFS)*, 2023.

[21] Xiling Li, Chenkai Weng, Yongxin Xu, Xiao Wang, and Jennie Rogers. ZKSQL: Verifiable and Efficient Query Evaluation with Zero-Knowledge Proofs. In *Proceedings of the VLDB Endowment (PLVDB)*, 2023.

[22] Hanlin Liu, Xiao Wang, Kang Yang, , and Yu Yu. The Hardness of LPN over Any Integer Ring and Field for PCG Applications. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques (Eurocrypt)*, 2024.

[23] Daniel Luick, John Kolesar, Timos Antonopoulos, William R. Harris, James Parker, Ruzica Piskac, Eran Tromer, Xiao Wang, and Ning Luo. ZKSMT: A VM for Proving SMT Theorems in Zero Knowledge. In *USENIX Security*, 2024.

[24] Ning Luo, Timos Antonopoulos, William Harris, Ruzica Piskac, Eran Tromer, and Xiao Wang. Proving UNSAT in Zero Knowledge. In *ACM Conference on Computer and Communications Security (CCS)*, 2022, **Distinguished Paper Award**.

[25] Ning Luo, Samuel Judson, Timos Antonopoulos, Ruzica Piskac, and Xiao Wang. ppSAT: Towards Two-Party Private SAT Solving. In *USENIX Security*, 2022.

[26] Radhika, Kang Yang, Jonathan Katz, and Xiao Wang. Scalable Mixed-Mode MPC. In *IEEE Symposium on Security and Privacy (S&P)*, 2024.

[27] Yuyang Sang, Ning Luo, Samuel Judson, Ben Chaimberg, Timos Antonopoulos, Xiao Wang, Ruzica Piskac, and Zhong Shao. Ou: Automating the Parallelization of Zero-Knowledge Protocols. In *ACM Conference on Computer and Communications Security (CCS)*, 2023.

[28] Ali Shahin Shamsabadi, Gefei Tan, Tudor Ioan Cebere, Aurelien Bellet, Hamed Haddadi, Nicolas Papernot, Xiao Wang, and Adrian Weller. Confidential-DPproof: Confidential Proof of Differentially Private Training. In *International Conference on Learning Representations (ICLR)*, 2024, **spotlight**.

[29] Ali Shahin Shamsabadi, Sierra Calanda Wyllie, Nicholas Franzese, Natalie Dullerud, Sebastien Gambs, Nicolas Papernot, Xiao Wang, and Adrian Weller. Confidential-PROFITT: Confidential PROof of FaIr Training of Trees. In *International Conference on Learning Representations (ICLR)*, 2023, **notable top 5%**.

[30] Xiao Wang and Jennie Rogers. VaultDB: Facilitating Secure Analytics over Multiple Private Data Sources. In *Workshop on Privacy Enhancing Technologies for the Homeland Security Enterprise*, 2022.

[31] Chenkai Weng, Kang Yang, Xiang Xie andJonathan Katz, and Xiao Wang. Mystique: Efficient Conversions for Zero-Knowledge Proofs with Applications to Machine Learning. In *USENIX Security*, 2021.

[32] Chenkai Weng, Kang Yang, Jonathan Katz, and Xiao Wang. Wolverine: Fast, Scalable, and Communication-Efficient Zero-Knowledge Proofs for Boolean and Arithmetic Circuits. In *IEEE Symposium on Security and Privacy (S&P)*, 2021.

[33] Chenkai Weng, Kang Yang, Zhaomin Yang, Xiang Xie, and Xiao Wang. AntMan: Interactive Zero-Knowledge Proofs with Sublinear Communication. In *ACM Conference on Computer and Communications Security (CCS)*, 2022.

[34] Xiang Xie, Kang Yang, Xiao Wang, and Yu Yu. Lightweight Authentication of Web Data via Garble-Then-Prove. In *USENIX Security*, 2024.

[35] Kang Yang, Pratik Sarkar, Chenkai Weng, and Xiao Wang. QuickSilver: Efficient and Affordable Zero-Knowledge Proofs for Circuits and Polynomials over Any Field. In *ACM Conference on Computer and Communications Security (CCS)*, 2021, **Best Paper Award Runner-up**.

[36] Kang Yang and Xiao Wang. Non-Interactive Zero-Knowledge Proofs to Multiple Verifiers. In *Annual International Conference on the Theory and Application of Cryptology and Information Security (Asiacrypt)*, 2022.

[37] Kang Yang, Xiao Wang, and Jiang Zhang. More Efficient MPC from Improved Triple Generation and Authenticated Garbling. In *ACM Conference on Computer and Communications Security (CCS)*, 2020.

[38] Kang Yang, Chenkai Weng, Xiao Lan, Jiang Zhang, and Xiao Wang. Ferret: Fast Extension for coRRElated oT with small communication. In *ACM Conference on Computer and Communications Security (CCS)*, 2020.

[39] Wenhao Zhang, Xiaojie Guo, Kang Yang, Ruiyu Zhu, Yu Yu, and Xiao Wang. Efficient Actively Secure DPF and RAM-based 2PC with One-Bit Leakage. In *IEEE Symposium on Security and Privacy (S&P)*, 2024.