

Kurzbericht CRYPTTECS Verbundprojekt - Universität Stuttgart

Teilvorhaben „Effizientes MPC und Integration in eine Privacy-Preserving Cloud-Umgebung“

Aufgabenstellung und wissenschaftlicher technischer Stand zu Projektbeginn

Das Ziel der Universität Stuttgart innerhalb des Verbundprojektes CRYPTTECS war es, Multi-Party-Computation (MPC) für den industriellen Einsatz, insbesondere im Umfeld von Anwendungen des Maschinellen Lernens (ML), zu verbessern und verfügbar zu machen. Dabei sollte ein besonderer Fokus auf die Effizienzsteigerung von MPC selbst sowie die Kombination von MPC mit weiteren Technologien für Privacy-Preserving Computing (PPC-Technologien) gelegt werden.

MPC-Protokolle wurden von Yao [Yao86, doi:[10.1109/SFCS.1986.25](https://doi.org/10.1109/SFCS.1986.25)] für den Spezialfall von zwei Parteien eingeführt und in den Folgejahren schrittweise erweitert und verbessert. Mit SPDZ [DPSZ12, <https://eprint.iacr.org/2011/535>] wurden erstmals effiziente MPC-Berechnungen beliebig vieler sich gegenseitig misstrauender Parteien möglich. Diese Protokolle wurden schrittweise verbessert, sodass mit Projektstart [KPR18, <https://eprint.iacr.org/2017/1230>] (mit verbesserten ZKPs aus [BCS19, <https://eprint.iacr.org/2019/035>]) die beste Effizienz für aktiv-sichere Berechnungen weniger Parteien erreichte. Berechnungen mit wenigen Parteien entsprechen der im Verbundprojekt CRYPTTECS angestrebten Cloudinfrastruktur. Auf diesen Vorarbeiten aufbauend, hat die Universität Stuttgart neue effizientere MPC-Protokolle entwickelt und evaluiert.

Ablauf des Vorhabens und wesentliche Ergebnisse

Entsprechend des geplanten und in der Teilvorhabensbeschreibung ausgeführten Arbeitsplans lag der Fokus in der ersten Projektphase auf der Verbesserung SPDZ-basierter MPC-Protokolle für Anwendungen des maschinellen Lernens (ML). Die Effizienz von ML-Anwendungen hängt dabei vorrangig von der effizienten Umsetzung verwendeter Teiloperationen, wie Matrizenmultiplikation, Tensorfaltungen und der Evaluation von Aktivierungsfunktionen, z. B. polynomialer Funktionen oder Vergleichsoperationen, ab. Diese Bestandteile von ML-Algorithmen wurden von uns jeweils verbessert und die neuen Protokolle als Erweiterung von MP-SPDZ implementiert. Insbesondere sind unsere Implementierungen damit verwendbar mit der von unserem Projektpartner bereitgestellten MPC-Cloudinfrastruktur Carbyne Stack (<https://carbynestack.io/>).

Sichere Matrixmultiplikation. Die aktive Berechnung in MPC-Protokollen wie SPDZ nutzt speziell konstruierte strukturierte Zufallszahlen, wie Beavertripel. Diese Zufallszahlen hängen nicht von sensiblen Eingabedaten der Parteien ab, ermöglichen aber die schnelle Berechnung auf sensiblen Daten (in der sogenannten Onlinephase). Beavertripel können vor der eigentlichen Berechnung (in einer Offlinephase) bereits generiert werden, sodass die zu ihrer Konstruktion benötigte Zeit wenig Einfluss auf die Laufzeit einer MPC-Berechnung in der Onlinephase hat. Während Beavertripel für einfache Multiplikationen sehr gut geeignet sind, verwenden wir in unseren neuen Protokollen zur sicheren Matrizenrechnung kürzlich aufgekommene neuartige strukturierte Zufallszahlen, sogenannte Matrixtripel.

Unser Projekt hat zunächst die Generierung von Matrixtripeln erforscht und hierbei über den Antrag hinausgehende Fortschritte erreicht, die es ermöglichen neben Matrixtripel auch klassische Beavertripel schneller zu erzeugen. Dazu wurde für die linear homomorphe Offlinephase ein neuer Sicherheitsansatz entwickelt, der sowohl die Laufzeit als auch die Menge an Daten verringert, die zwischen den Parteien ausgetauscht werden muss. Zusätzlich wurden entsprechend dem Antrag für die Auswertung von Skalarprodukten und das Quadrieren von Matrizen weitere Optimierungen gefunden. Unsere Evaluation zeigt eine Verbesserung um 28% - 74% in der Laufzeit und 57% - 66% in der Bandbreite für verschiedenen ML-Anwendungen jeweils im Vergleich zu den besten bekannten Protokollen.

Die Ergebnisse wurden bei der Top-Konferenz „ASIA Conference on Computer and Communications Security“ (AsiaCCS 2023) präsentiert und veröffentlicht (siehe <https://eprint.iacr.org/2023/462>).

Sichere Tensorfaltung. Ebenso wie für Matrizenoperationen, können auch Tensorfaltungen mit speziellen Tensortripeln effizienter berechnet werden. Entsprechend des Arbeitsplans wurde die effiziente Generierung dieser Tupel untersucht. Dabei konnte ein neuartiger Ansatz zur Generierung mittels „packing“ Techniken gefunden werden, der es erlaubt, ganze Tensoren in einzelne gitterbasierte Chiffretexte zu schreiben und die Multiplikation dieser Chiffretexte zur Tensorfaltung zu nutzen. Sowohl die theoretische Analyse als auch der

Laufzeitvergleich zeigen einen klaren Vorteil unserer neuen Konstruktion. Beispielhaft erhalten wir damit eine um den Faktor 40 schnellere Onlinephase für den ML-Algorithmus ResNet-50 im Vergleich mit [KRP18].

Unsere Ergebnisse wurde auf der Top-Konferenz „Privacy Enhancing Technologies Symposium“ (PETS 2023) präsentiert und veröffentlicht (siehe <https://eprint.iacr.org/2023/359>). **Darüber hinaus erhielt unser Artikel einen zweiten Preis des Andreas Pfitzmann Best Student Awards.**

Sichere Polynomevaluationen und Fixpunktoperationen. In MPC-Protokollen mit additivem Sharing stellen Multiplikationen eine besondere Herausforderung dar, da jede Multiplikation üblicherweise den Austausch von Daten zwischen den Parteien erfordert. Entsprechend unseres Arbeitsplans wurden eine neue Art von strukturierten Zufallszahlen (Polynomielle Tupel) erforscht und gefunden, die viele Multiplikationen (wie bei der Evaluation von Polynomen) in einer Onlinekommunikationsrunde ermöglicht. Unsere polynomiellen Tupel verbessern lange bestehende Komplexitätsergebnisse für derartige Rechnungen signifikant. Darüber hinaus konnten wir zeigen, dass sich auch in ML übliche vergleichsbasierte Aktivierungsfunktionen wie ReLU/Argmax damit verbessern lassen.

Unsere Ergebnisse wurden auf der Top-Konferenz Asiacrypt 2024 präsentiert und veröffentlicht (siehe <https://eprint.iacr.org/2024/1435>). Auch im Bereich der Fixpunktoperationen wurden die theoretischen Arbeiten des Arbeitsplans erfolgreich ausgeführt und ein technischer Bericht erstellt - die Veröffentlichung der Ergebnisse ist in Vorbereitung. Über den ursprünglichen Projektplan hinaus haben wir eine Verbesserung oben genannter Techniken für andersartige Grundringe erreicht, die eine bessere Kompatibilität mit klassischer Rechnerinfrastruktur bietet und ML-Berechnungen weiter beschleunigt (erschieden in PETS 2024, siehe <https://eprint.iacr.org/2023/1932>).

Integration von MPC mit anderen PPT-Technologien. Im Laufe des Projektes wurde die Integration von MPC und lokaler/globaler Differential Privacy als am vielversprechendsten identifiziert. Hier wurden zusammen mit den französischen Partnern INRIA und Orange Anwendungsszenarien in der Telekommunikation identifiziert und dafür eine neuartige Lösung entwickelt, die eine Kombination aus globaler und lokaler Differential Privacy für viele Parteien liefert. Entsprechend dem Antrag wurde ein neuer Sicherheitsbegriff entwickelt. Die Ergebnisse sind unter Begutachtung bei der Top-Konferenz „Computer Security Foundations Symposium (IEEE CSF)“.

Kombinationen von MPC und homomorpher Verschlüsselung wurden im Rahmen der Offlinephasen von MPC-Protokollen untersucht. Forschungsergebnisse, wie neue “Packing”-Methoden in homomorphe Chiffretexte sind in oben genannte Arbeiten zu Tensorfaltungen und Matrizenmultiplikationen eingegangen. Die Kombination von MPC und Trusted Execution Environments (TEEs) wiederum wurde durch Bachelorarbeiten bei der Robert Bosch GmbH und an der Universität Stuttgart [17] untersucht. Die Forschung in diesem Bereich ist noch nicht abgeschlossen und wird im Rahmen einer Masterarbeit an der Universität Stuttgart fortgesetzt.

Neben der ursprünglich geplanten PPT-Technologien hat sich im Laufe des Projekts Federated Learning als vielversprechendes ML-Schema herauskristallisiert, in das sich PPT-Technologien wie MPC effizient integrieren lassen. An einem Anwendungsbeispiel aus der Gesichtserkennung konnte gezeigt werden, dass MPC dafür geeignet ist akkurate ML-Modelle zu trainieren und gleichzeitig die im Trainingsprozess verwendeten Gesichtsbilder privat zu halten. Unsere Ergebnisse wurden im Rahmen des „Symposium on Eye Tracking Research & Applications“ (HCI 2024, <https://arxiv.org/abs/2402.18970>) veröffentlicht. Weiterhin haben wir die Notwendigkeit unserer privatheitsschützenden Protokolle bewiesen, indem wir die Gefahr durch Replayattacken analysiert haben, die Privatheit in E-Votingsystemen und voraussichtlich auch Federated Learning Schemata brechen können (CSF 2022, siehe <https://eprint.iacr.org/2022/743>).

Für eine detaillierte Beschreibung obengenannter Ergebnisse, sowie weiterer über den ursprünglichen Projektplan hinausgehender Veröffentlichungen im Rahmen von CRYPTTECS verweisen wir auf den ausführlichen Abschlussbericht und <https://www.cryptecs.eu>.

Zusammenfassend wurden die im Antrag formulierten Forschungsziele größtenteils erreicht und teilweise übertroffen. Die Veröffentlichung noch ausstehender Ergebnisse wie zu Fixpunkten und zur Kombination von MPC und Differential Privacy ist in Vorbereitung. Über den Antrag hinausgehend wurden weitere für die Integration von MPC in industrielle Anwendungen relevante Forschungsergebnisse erzielt. Das CRYPTTECS-Verbundprojekt und insbesondere das Teilprojekt der Universität Stuttgart liefert damit sicheres MPC-basiertes maschinelles Lernen, das zukünftig von Industrieunternehmen genutzt werden kann und schon jetzt bei unseren Industriepartnern im Einsatz ist.