Der vorliegende Beitrag wurde beim Deutschen Studienpreis 2023 mit dem 2. Preis in der Sektion Natur- und Technikwissenschaften ausgezeichnet. Er beruht auf der 2022 an der Rheinisch-Westfälischen Technischen Hochschule Aachen eingereichten Dissertation „Designing Trustworthy Hardware with Logic Locking“ von Dr. Dominik Šišejkovic.

Entwicklung vertrauenswürdiger Mikroelektronik: Von der Theorie zur Praxis

1 Die Halbleiterindustrie: Die existenzielle Abhängigkeit von nicht vertrauenswürdigen Geschäftspartnern

Die Hardware (HW) stellt zweifellos die kritischste Sicherheitsschicht in modernen mikroelektronischen Systemen dar. Selbst eine geringfügige und sorgfältig konzipierte Veränderung der HW kann jeden Sicherheitsmechanismus in den oberen Schichten des Systems außer Kraft setzen. Aber was ermöglicht solche Manipulationen? Die Halbleiterindustrie ist heute in hohem Maße auf das geistige Eigentum Dritter (engl. IP), die Einbindung externer Designhäuser und die Auslagerung der Herstellung an externe Produktionsstätten angewiesen. Dieses auf externe Parteien angewiesene Geschäftsmodell trägt der immer größer werdenden Notwendigkeit Rechnung, die Produktionskosten zu senken und die Design- und Fertigungszyklen zu verkürzen. Leider ermöglicht die Beteiligung von externen und nicht vertrauenswürdigen Parteien aber auch eine Reihe schwerwiegender Sicherheitsprobleme, und das aus einem einfachen Grund – externe Designhäuser und Halbleiterfertigungsstätten erhalten uneingeschränkten und unkontrollierten Zugang zur Hardware. Die fehlende Möglichkeit, *Vertrauen* in externe Parteien zu gewährleisten, stellt daher die Konkurrenzfähigkeit der seriösen HW-Designer in Widerspruch zu einem sicheren Design- und Fertigungszyklus. Vor allem die Einbindung bösartiger Designmodifikationen, die als Hardware-Trojaner bezeichnet werden, gilt als ernst zu nehmende Bedrohung. Diese unauffälligen Modifizierungen können eingesetzt werden, um mikroelektronische Systeme durch geschickt eingefügte Hintertüren auszunutzen, zu manipulieren und zu steuern. Dies kann potenziell zu einer Vielzahl von Angriffen führen, beispielsweise zu Informationsverlusten, Denial-of-Service (englisch für „Verweigerung des Dienstes”) und einer Beeinträchtigung der Hardwarezuverlässigkeit, um nur einige zu benennen. Da Hardwarechips in allen Bereichen des modernen Zeitalters eine entscheidende Rolle spielen, können Hardware-Trojaner in den unterschiedlichsten Bereichen wie

Telekommunikation, Fahrzeugelektronik, Medizintechnik, Finanzinfrastrukturen und Militärsysteme erheblichen Schaden anrichten.

Diese äußerst schwierige Herausforderung hat im akademischen und industriellen Sektor großes Interesse an der Erforschung neuartiger vertrauenswürdiger HW-Designmethoden geweckt. Die U.S. Defense Advanced Research Project Agency (DARPA) hat mehrere Förderprogramme initiiert, die sich mit der Frage der vertrauenswürdigen Elektronik befassen, darunter das TRUST, IRIS, und SHIELD Programm, um nur einige zu nennen. Die Schwere des Problems ist auch in Deutschland erkannt worden. Das deutsche Bundesministerium für Bildung und Forschung (BMBF) hat ein Rahmenprogramm für die Jahre 2021–2024 aufgelegt, um die Herausforderungen einer vertrauenswürdigen und nachhaltigen Mikroelektronik für Deutschland und Europa anzugehen. Darüber hinaus wurden kürzlich seitens der deutschen Agentur für Innovation in der Cybersicherheit mehrere Vorstudien gestartet, die kritische Forschungslücken im Bereich nachweisbar sicherer und vertrauenswürdiger Informationstechnologie für die digitale Souveränität Deutschlands aufdecken und bearbeiten sollen. Hinzu kommt, dass Hardware-Trojaner über ein bloßes Forschungsproblem hinausgehen. Führungskräfte des US-Militärs und der Geheimdienste haben Hardware-Trojaner zu den größten Bedrohungen eingestuft, denen die Nation im Falle eines Krieges ausgesetzt sein könnte. Daher bleibt es eindeutig: Das Streben nach vertrauenswürdigen mikroelektronischen Systemen ist ein zentraler Punkt der Forschung und von grundlegender Bedeutung für eine digitale und sichere Zukunft.

2 Die Anatomie des Hardwareangriffs

Wie bereits erwähnt, wird das Einschleusen von Hardware-Trojanern dadurch ermöglicht, dass Teile des Hardwareentwurfs und der Herstellung an nicht vertrauenswürdige Drittanbieter ausgelagert werden, die schlichtweg vollen Zugriff auf die Hardwarebeschreibung erhalten müssen, um diese zu verarbeiten. Hierbei ist es wichtig zu verstehen, dass diese Parteien immer noch die Möglichkeit haben, das Hardwaredesign zu ändern, bevor dessen Funktionalität in Silizium „eingebrannt" und somit unveränderbar wird. Aber was benötigt ein Angreifer, um einen Trojaner einzufügen? Hier müssen wir zuerst zwischen zwei Trojaner-Klassen unterscheiden: designunabhängige und designabhängige Hardware-Trojaner. Designunabhängige Trojaner ähneln zufälligen Modifikationen der HW, daher benötigt ein Angreifer für das Einfügen *keine Kenntnisse* über die tatsächliche Funktionalität des Mikrochips. Dies hat eine wichtige Folge: Wir können uns nicht gegen zufällige Veränderungen schützen, da kein Wissen erforderlich ist, um diese durchzuführen. Die gute Nachricht ist: Es ist unwahrscheinlich, dass zufällige und

funktionalitätsunabhängige Änderungen zu gezielten und kontrollierbaren Angriffen führen. Anders sieht es bei designabhängigen Trojanern aus. Damit diese eingeschleust werden können, muss ein Angreifer die Design-Spezifikationen vollständig verstehen und zurückentwickeln (engl. „Reverse Engineering"). Nur dann kann der Trojaner speziell darauf ausgelegt werden, die ursprüngliche Funktionalität des Hardwaredesigns kontrolliert zu verändern. Dieses designspezifische Angriffsmodell führt daher zu einem verheerenden Schaden, da es eine kontrollierbare Angriffsumgebung ermöglicht.

3 Der Weg zu sicherer Mikroelektronik

Wir haben gelernt, dass designabhängige Trojaner der Kern des Problems sind. Doch eine Frage bleibt offen: Wie kann man die Hardware gegen designabhängige Hardware-Trojaner schützen? In den letzten Jahrzehnten wurden viele Sicherheitsmechanismen vorgeschlagen, um die Vertrauenswürdigkeit von HW zu gewährleisten. Insbesondere die Logikverschlüsselung, eine HW-Obfuskationstechnik, wurde als eine der führenden Methoden zum Schutz der Integrität von HW-Designs identifiziert, da sie vor Angreifern schützen kann, die sich überall in der HW-Lieferkette befinden. Die Funktionsweise der Logikverschlüsselung lässt sich im Prinzip wie folgt beschreiben. Der legitime Eigentümer der Hardware-IP ändert (logikverschlüsselt) das Design, bevor der Chip an externe Parteien zur weiteren Bearbeitung übergeben wird. Der korrekte Schlüssel wird erst nach der Produktion auf dem Chip abgelegt, wenn keine Änderungen mehr möglich sind. Auf diese¨ Weise bleibt das Design geschützt, während es von externen Parteien bearbeitet wird.

Die Verschlüsselung äußert sich in Form von *funktionalen und strukturellen Veränderungen*, die von einem geheimen Aktivierungsschlüssel abhängig sind. Damit muss ein potenzieller Angreifer erheblich mehr Aufwand betreiben, um zunächst das Design zu entschlüsseln (bzw. den Aktivierungsschlüssel zu finden) und dann das Reverse Engineering und ggf. die Trojaner-Einschleusung durchzuführen.

Obwohl die Wissenschaft zahlreiche Anstrengungen unternommen hat, um resiliente Logikverschlüsselungsverfahren zu entwickeln, beschränken sich die meisten Lösungen auf theoretische Konzepte, fernab einer praktischen Anwendung. Daher konzentriert sich diese Dissertation auf das Schließen der Lücke zwischen theoretischen Modellen und industrietauglichen Lösungen, um eine sichere Grundlage für die Vertrauenswürdigkeit zukünftiger digitaler Systeme zu schaffen.

Um einen ganzheitlichen Ansatz für die sichere Hardwareentwicklung zu bieten, besteht diese Arbeit aus vier Bausteinen: dem Entwurf von Sicherheitsmetriken, der Implementierung einer industriekompatiblen Softwareinfrastruktur, dem

Entwurf von sicheren Logikverschlüsselungsverfahren für komplexe Hardwaredesigns und der Implementierung einer automatischen Schwachstellenanalyse. Darüber hinaus wurde der Entwurf von Logikverschlüsselungsverfahren und Angriffen auf diese mit den Fähigkeiten moderner maschineller Lerntechniken verschmolzen, um Schwachstellen aufzudecken, die bisher übersehen wurden. Weitere Einzelheiten werden im Folgenden erläutert.

3.1 Wie wird Hardwaresicherheit gemessen?

Die Sicherheitsmetriken sind für Datenverschlüsselung gut definiert, doch für die Hardwaresicherheit ergibt sich ein ganz anderes Bild. Die Herausforderung besteht darin, dass die Logikverschlüsselung eine Vielzahl von Veränderungen an einem HW-Design bewirkt und dadurch potenziell eine Reihe von disjunkter Angriffsvektoren ermöglicht. Es ist zum Beispiel möglich, einen korrekten Aktivierungsschlüssel zu finden, indem man die Verschlüsselungsstrukturen oder das Funktionsverhalten des Chips für fehlerhafte Schlüssel analysiert, da es oft eine Korrelation gibt, die den Schlüsselwert mit den beobachteten Strukturen und dem Verhalten des Chips verknüpft. Daher können oft voneinander unabhängige Angriffsvektoren eine Verschlüsselung knacken. Dies macht es ausgesprochen schwierig, alle sicherheitsrelevanten Merkmale in einer einzigen, einheitlichen Metrik zusammenzufassen. Folglich waren unsere ersten Bemühungen darauf ausgerichtet, eine der ersten in der Literatur verfügbaren verallgemeinerten Metriken für die HW-Sicherheit zu entwerfen. Die Metrik besteht aus einer Aufteilung der wichtigsten Sicherheitsaspekte in eine mehrdimensionale Reihe von Eigenschaften, die sich über die verschiedenen Merkmale der bestehenden Logikverschlüsselungsverfahren erstrecken. Betrachten wir einige der definierten Sicherheitsmerkmale, um besser zu verstehen, was die vorgeschlagene Metrik aussagt. Auf einer hohen Ebene müssen wir zwischen strukturellen und funktionalen Merkmalen der Verschlüsselung unterscheiden. Zu den relevanten Merkmalen im Sinne der Funktionalität gehört zum Beispiel die Fehlerhaftigkeit der Ausgabe für inkorrekte Aktivierungsschlüssel. Ein sicheres Design sollte inkorrekte Outputs für inkorrekte Schlüssel generieren, wobei der Grad der Fehlerhaftigkeit bestimmte Sicherheitsmerkmale aufweisen muss, um heuristisch gesteuerte Angriffe zu verhindern. Ein strukturelles Sicherheitsmerkmal ist hingegen, wie intensiv sich das Design durch die Verschlüsselung topologisch verändert hat und wie diese Veränderung im Design verteilt ist. Dabei ist es wichtig, beide Merkmale gleichzeitig zu betrachten. Eine unsichere Verschlüsselung könnte zum Beispiel nur sehr geringe Veränderungen bewirken, die über die gesamte Schaltung verteilt sind, oder aber viele Veränderungen, die sich nur auf einen Punkt konzentrieren. So kann ein Merkmal ein

hohes Maß an Sicherheit signalisieren, während das Design aufgrund eines anderen Merkmals unsicher bleibt. Diese und andere sicherheitsrelevante Merkmale wurden identifiziert und ein Bewertungssystem für sie vorgeschlagen. Dabei ist jedes Merkmal erweiterbar, um der sich schnell verändernden Landschaft der Logikverschlüsselungsverfahren Rechnung zu tragen. Darüber hinaus wurden erste Kenntnisse über die Auswirkung der Verschlüsselungskosten (in Bezug auf Chipfläche, Stromverbrauch und Leistungseinbußen) auf das Sicherheitsniveau gewonnen. Mit dieser Arbeit haben wir den ersten Schritt zum grundlegenden Verständnis der Bedeutung von Hardwaresicherheit und ihrer Messung unternommen.

3.2 Wie entwirft man eine industrietaugliche Softwareinfrastruktur?

Eine anwenderfreundliche Softwareinfrastruktur ist die Voraussetzung für die Applikation, aber auch für das Design von sicheren Verschlüsselungsverfahren in einem industriellen Umfeld. Daher konzentrierte sich der nächste Baustein unserer Arbeit auf den Entwurf und die Entwicklung einer modernen Software-Toolchain für Logikverschlüsselung.

Die Entwicklung der Software wurde von mehreren Anforderungen bestimmt, darunter Modularität, sicherheitsspezifische HW-Designprozesse und Standardschnittstellen. Die Modularität wurde durch modernes Softwaredesign umgesetzt, welches das Hinzufügen und Evaluieren neuer Verschlüsselungsverfahren mit minimalem Aufwand und maximaler Wiederverwendung des Quellcodes ermöglicht – eine kritische Anforderung in einer sich rapide entwickelnden Verschlüsselungslandschaft. Darüber hinaus ist die Softwareinfrastruktur in konzeptionellen Schritten aufgebaut, die es dem HW-Designer ermöglichen, HW-Komponenten individuell zu bearbeiten. Auf diese Weise kann jede Komponente des potenziell sehr komplexen HW-Designs entsprechend ihrer spezifischen Sicherheitsanforderungen verschlüsselt werden. Schließlich ist ein Software-Tool nur dann nützlich, wenn es sich nahtlos in die herkömmliche Softwarelandschaft einfügen lässt. Die Eingaben und die generierten Ausgaben der Softwareinfrastruktur arbeiten daher mit herkömmlichen Hardwarebeschreibungssprachen, was eine direkte Umsetzung in einer industriellen Umgebung ermöglicht.

3.3 Integritätsschutzmethoden für komplexe Hardwaredesigns

Während der gesamten Forschung haben wir große Mängel in bestehenden Arbeiten festgestellt: die Beschränkung auf isolierte und kleine Schaltkreiskomponenten, unklare Sicherheitsmetriken und eine fehlende Verbindung zu realen Angriffsszenarien. Diese Beobachtungen beeinträchtigen die *Praxistauglichkeit* der

Logikverschlüsselung erheblich. Aus diesem Grund haben wir die beschriebene Softwareinfrastruktur um einen automatischen Ansatz zur Skalierung von Logikverschlüsselungsverfahren für komplexe HW-Designs erweitert. Die automatische Skalierung gewährleistet eine wichtige Funktion – die Sicherheitsabhängigkeit zwischen nicht miteinander verbundenen Designkomponenten wird geschaffen. Das hat eine wichtige Konsequenz: Die Aufteilung des Hardwaredesigns in kleinere Komponenten, die einzeln angegriffen werden können, wird erschwert. Interessanterweise war dies der erste Ansatz, der sich gegen den „Divide and Conquer"-Angriff richtete. Spätere Untersuchungen haben gezeigt, dass gerade diese Art von Angriff ein fundamentaler Schritt von Reverse Engineering ist.

Zusätzlich haben wir einen weiteren Angriffsvektor erforscht: Reicht die Identifikation eines einzigen kritischen Prozessorsignals aus, um einen softwaregesteuerten Hardware-Trojaner einzubauen? Die kurze Antwort lautet: leider ja! Aus praktischer Sicht ermöglicht diese Schwachstelle ein äußerst gefährliches und breites Angriffsspektrum, das sowohl HW- als auch SW-Ressourcen ausnutzt. Daher haben wir die erste Methodik zur Sicherung von Kontrollsignalen zwischen unabhängigen Prozessormodulen vorgestellt, die eine Einbindung von Trojanern verhindert, selbst wenn die Signale von einem Angreifer entdeckt werden.

Mit dem vorgestellten Forschungsbaustein haben wir die ersten greifbaren Lösungen zum Schutz komplexer HW-Designs entwickelt und dabei die Grenzen eines realistischen Angriffsvektors deutlich offengelegt.

3.4 Maschinelles Lernen für Hardwaresicherheit: Ein Blick über den Tellerrand

Mit der Verbreitung effizienter und benutzerfreundlicher Modelle für maschinelles Lernen (ML) haben ML-basierte Methoden langsam Einzug in den Bereich der Logikverschlüsselung gefunden. Die jüngste Forschung hat gezeigt, dass der Einsatz moderner ML-Techniken ein großes Potenzial hat, um die Sicherheit von Logikverschlüsselungsverfahren zu bewerten. Dennoch gab es erhebliche Lücken in der Theorie – was ist die Quelle von ML-ausnutzbaren Schwachstellen, und wie werden ML-resistente Logikverschlüsselungsverfahren entwickelt? In diesem Zusammenhang haben wir uns auf drei wesentliche Herausforderungen konzentriert.

Zunächst haben wir analysiert, welche Schwachstellen der Logikverschlüsselung durch ML oder, allgemeiner gesagt, sogar durch eine menschliche Analyse identifizierbar sind. Diese Arbeit führte uns zur Entwicklung der weltweit ersten theoretischen Grundlage der *Lernresilienz* im Bereich der HW-Obfuskation.

Basierend auf diesen Konzepten, haben wir ein theoretisches Testsystem entwickelt, das in der Lage ist, ML-ausnutzbare Schwachstellen schon während der Entwurfsphase von Logikverschlüsselungsalgorithmen zu identifizieren. Der Test wurde erfolgreich eingesetzt, um strukturelle Fehler verschiedener Logikverschlüsselungsverfahren zu identifizieren, die über ein Jahrzehnt übersehen wurden.

Nachdem wir die theoretische Grundlage für die ML-Lernresilienz gelegt hatten, haben wir einen neuen, auf ML basierenden Angriff entwickelt, der als Erster seiner Art in der Lage ist, durch „nur einen Blick" auf das abgesicherte HW-Design die Logikverschlüsselung effizient zu brechen. Der „One Look"-Angriff basiert auf einer neu entwickelten Darstellungsform von HW-Designs, die verschlüsselungsrelevante Informationen

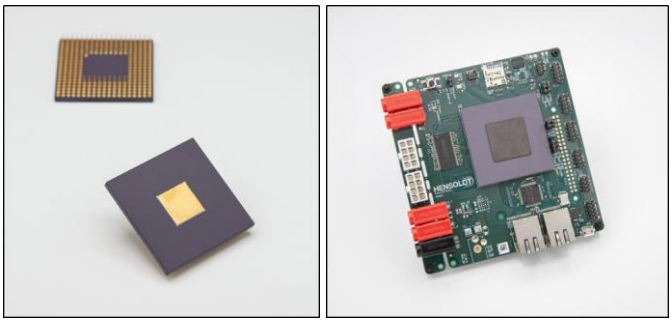

Abbildung 1: Der erste logikverschlüsselte Prozessor „Made in Germany" (MiG-V), basierend auf der Open-Source RISC-V Befehlssatzarchitektur

in der HW erfasst und in ein ML-Modell übertragen kann. Dadurch kann der Angriff ein Design beliebiger Komplexität verarbeiten und die Sicherheit einer beliebigen Verschlüsselungsstrategie bewerten. Dies garantiert die allgemeine Anwendbarkeit des Angriffs als Werkzeug zur Sicherheitsbewertung.

Schließlich ermöglichten uns der entwickelte theoretische Test sowie der ML-basierte Angriff, die erste ML-resistente Logikverschlüsselungsstrategie zu entwickeln. Der Verschlüsselungsalgorithmus basiert auf der Einbindung von schlüsselgesteuerten Logikstrukturen, die keine strukturellen oder funktionalen Schwachstellen beinhalten, die möglicherweise durch einen ML-basierten Angriff aufgedeckt werden könnten. Dieses erste ML-resistente Verschlüsselungsverfahren bahnte den Weg für die Entwicklung sicherer Verschlüsselungsstrategien der nächsten Generation in einer zunehmend ML-getriebenen Welt.

Die oben genannten Ergebnisse der interdisziplinären Forschung im Bereich HW-Sicherheit und maschinelles Lernen haben uns ein neues Verständnis der Sicherheitsgrundlagen moderner Logikverschlüsselungsverfahren ermöglicht, was einen *neuen Entwicklungs- und Forschungskurs* in diesem Bereich ausgelöst hat.

4 Ein Meilenstein für sichere Hardware

Das Framework wurde anhand praktischer Fallstudien evaluiert, bei denen mehrere industrietaugliche Prozessorkerne umfassend verschlüsselt wurden. Dieses ermöglichte eine einfache Kompromissanalyse zwischen Sicherheit und Kosten. Doch wir haben hier nicht aufgehört! Die Forschungsergebnisse sowie die entwickelte Softwareinfrastruktur für den Entwurf und die Anwendung neuartiger Logikverschlüsselungsverfahren wurden durch die Hensoldt Cyber GmbH, ein bayerisches Start-up-Unternehmen, das sich auf die Entwicklung hochsicherer eingebetteter Systeme konzentriert, erfolgreich in die Industrie übertragen. Durch die Zusammenarbeit mit der Industrie hatten wir die Möglichkeit, den „ganzen Weg" zu gehen – von der Grundlagenforschung bis zum Siliziumchip –, indem wir zur Entwicklung des weltweit ersten logikverschlüsselten „Made in Germany"-Prozessors für sicherheitskritische Anwendungen beigetragen haben (Abb 1). Somit wurde ein wichtiger Meilenstein für die Entwicklung von sicherer Hardware geschaffen, der direkt in einem industriellen Umfeld anwendbar ist. Zusammengefasst hat diese Arbeit dazu beigetragen, das Potenzial der Logikverschlüsselung von theoretischen Konstrukten auf praktische Anwendungen zu übertragen und die Konzepte, Werkzeuge und Metriken für den Aufbau vertrauenswürdiger Hardware bereitzustellen.