Abschlussbericht
Teil 1: Kurzbericht

Zuwendungsempfänger	Fachgebiet Multimedia Kommunikation (KOM), Technische Universität Darmstadt
Verbundprojekt	**AI-NET-PROTECT**
Förderkennzeichen	**16KIS1295**
Laufzeit/Berichtszeitraum	01.02.2021 – 31.03. 2024
Datum	27.09.2024

1 Ursprüngliche Aufgabenstellung sowie wissenschaftlicher Stand zu Projektbeginn

Die ursprüngliche Aufgabenstellung des Verbundprojektes AI-NET-PROTECT lag zum Zeitpunkt des Projektbeginns darauf, Konzepte, Architekturen und Methoden für automatisierten Netzbetrieb und sichere Netze auf vertrauenswürdigen Geräten für kritische Infrastrukturen und Unternehmen zu entwickeln. AI-NET-PROTECT sollte den Schutz kritischer Daten und hohe Robustheit der Netzinfrastruktur durch eine skalierbare Netz- und Knotenarchitektur, Netztelemetrie und softwaredefiniertes Netzmanagement für die Automatisierung von Ende-zu-Ende-Diensten mit Hilfe von künstlicher Intelligenz (KI), und starke Sicherheit auf Basis von mehrschichtiger Kryptographie, agilen Kryptofunktionen und quantensicheren Algorithmen gewährleisten.

Die TU Darmstadt war im Rahmen ihres Teilvorhabens dafür verantwortlich, für lokale Vorhersage- und Analysemodelle neue Verfahren zu entwickeln, um Informationen von vielen Teilnehmern in Kommunikationsnetzen vorherzusagen und vorzufiltern können. Weiterhin sollten komplexe Applikationen durch Verfahren lokaler Vorhersagen unterstützt und Interferenzen zwischen globalen und lokalen Informationen analysiert und genutzt werden.

Das Projekt AI-NET-PROTECT hat inhaltlich an zahlreiche Vorgängerprojekte angeschlossen. Besonders zu erwähnen ist das CELTIC-Plus SENDATE Projekt, an dem die TU Darmstadt ebenfalls maßgeblich beteiligt war. In diesem Projekt wurden erste vorarbeiten im Bereich des Monitorings von SDN-basierten Netzwerken durchgeführt, welche als Grundlage für das Teilvorhaben der TU Darmstadt innerhalb von AI-NET-PROTECT legte.

2 Ablauf des Vorhabens

Das Projekt wurde im Zeitraum 02/2022 bis 03/2024 an der TU Darmstadt durchgeführt und gegenüber der ursprünglichen Projektlaufzeit um 2 Monate kostenneutral verlängert. Persönlich zu erwähnen ist Frau Pegah Golchin. Sie war fast während der gesamten Projektlaufzeit als wissenschaftlicher Mitarbeiter in dem Projekt tätig und trug maßgeblich zu dem Erfolg des Projekts bei. Die Ergebnisse dieses Projekts finden sich in ihrer Dissertation mit dem Titel „*Machine Learning Models in Network Intrusion Detection Systems - Self-*

Supervised Detection of Malicious Flows and Traffic Patterns Recognition in Programmable Networks“ wieder, die sie am 16. Juli 2024 an der TU Darmstadt erfolgreich verteidigte.

Die inhaltlichen Arbeiten in den Arbeitspaketen AP1 – AP5 wurden in enger Zusammenarbeit mit den projektbeteiligten Partnern durchgeführt und teilweise auch gemeinsam in form von wissenschaftlichen Publikationen veröffentlicht (siehe 3.).

3 Wesentliche Ergebnisse und Forschungskollaboration

Das Fachgebiet Multimedia Kommunikation der TU Darmstadt hat die im Rahmen des Projekts gewonnenen Erkenntnisse erfolgreich in die Lehre und Forschung integriert. Die Inhalte des Projekts wurden in Form von Fragestellungen und Themen für studentische Abschlussarbeiten aufgegriffen und haben somit zur Ausbildung der nächsten Generation von Ingenieuren und Wissenschaftlern im Bereich der Software-Definierten-Computernetze beigetragen. Konkret wurden 1 Bachelorthesis und 4 Masterthesen erfolgreich betreut.

Im Bereich der Forschung wurden die wissenschaftlichen Ergebnisse auf internationalen Konferenzen veröffentlicht. Insgesamt wurden 7 peer-reviewed Publikationen in den Proceedings von IEEE- und ACM-Konferenzen veröffentlicht und auf den entsprechenden Tagungen diskutiert.

Besonders erwähnenswert ist der erarbeitete Ansatz des kollaborativen maschinellen Lernens (CML-IDS) der in verschiedenen Evolutionsstufen mehrfach wissenschaftlich publiziert und auf der IEEE/IFIP NOMS 2023 Konferenz mit dem *Best Student Paper Award* ausgezeichnet wurde.

Zusätzlich wurde das Anschlussprojekt *"Secure SDN"* initiiert, das inhaltliche Anknüpfungspunkte von AI-NET Protect aufgreift. Die in diesem Projekt gewonnenen Erkenntnisse konnten hier für weitere Forschung sowie die Cybersouveränität der Bundesrepublik Deutschland aufgegriffen werden.

Darüber hinaus wurden Synergien im Rahmen des *BMBF Open6GHub Projekts* genutzt. Hierbei fließen die im ursprünglichen Projekt erarbeiteten Lösungen und Ansätze in die Entwicklung von Technologien für zukünftige resiliente 6G-Kommunikationsnetze ein.

Ein weiterer Erfolg des Projekts ist die enge Kooperation mit der *TU Chemnitz*. Diese Partnerschaft führte zu mehreren gemeinsamen wissenschaftlichen Publikationen, die sowohl den Fortschritt des Projekts unterstreichen als auch die zukünftige Zusammenarbeit im Bereich Sicherheit und Maschinelles Lernen für Software-Definierte-Netzwerke weiter stärken.

GEFÖRDERT VOM

Bundesministerium
für Bildung
und Forschung

Abschlussbericht

Teil 2: Eingehende Darstellung

Zuwendungsempfänger	**Fachgebiet Multimedia Kommunikation (KOM), Technische Universität Darmstadt**
Verbundprojekt	**AI-NET-PROTECT**
Förderkennzeichen	**16KIS1295**
Vorhabenbezeichnung	**Automatisierte Resilienz und sichere Netze für kritische Infrastrukturen und Unternehmen - AI-NET-PROTECT** **Teilvorhaben: Automatisierte lokale Netzrepräsentation zur sicheren und flexiblen Netzverwaltung**
Laufzeit/Berichtszeitraum	01.02.2021 – 31.03. 2024
Datum	27.09.2024

1 Zusammenfassung der Projektziele zu Beginn des Projekts

Zum Beginn des Verbundprojekts AI-NET-PROTECT bestand die zentrale Forschungsvision darin, innovative Konzepte, Architekturen und Methoden für die automatisierte Verwaltung und sichere Bereitstellung von Netzwerken auf vertrauenswürdiger Hardware für kritische Infrastrukturen und Unternehmen zu erforschen. Dabei lag der Fokus auf dem Schutz sensibler Daten und der Sicherstellung einer hohen Ausfallsicherheit der Netzwerke durch eine skalierbare Architektur, fortschrittliche Netztelemetrie und ein softwaregesteuertes Managementsystem. Mithilfe von künstlicher Intelligenz sollten Ende-zu-Ende-Dienste automatisiert und durch eine mehrschichtige Kryptografie, flexible Verschlüsselungsprotokolle und quantensichere Algorithmen abgesichert werden.

Im Rahmen des Teilvorhabens „Automatisierte lokale Netzrepräsentation zur sicheren und flexiblen Netzverwaltung“ war die Technische Universität Darmstadt dafür verantwortlich, neuartige Verfahren für die lokale Vorhersage und Analyse von Netzwerkdaten zu entwickeln. Diese sollten die Effizienz und Genauigkeit der Datenverarbeitung in Kommunikationsnetzen verbessern, indem sie die Informationen zahlreicher Netzwerkteilnehmer vorhersagen und filtern. Ein weiterer Schwerpunkt der TU Darmstadt lag auf der Entwicklung von Methoden zur Unterstützung komplexer Anwendungen durch lokale Vorhersagemodelle sowie der Untersuchung und Nutzung von Wechselwirkungen zwischen globalen und lokalen Informationen.

AI-NET-PROTECT knüpfte inhaltlich an mehrere Vorgängerprojekte an. Für die TU Darmstadt ist hierbei insbesondere das CELTIC-Plus-Projekt SENDATE zu nennen. In diesem Vorprojekt wurden grundlegende Arbeiten im Bereich des Monitorings von Software-Definierten Netzwerken durchgeführt, die als Basis für die aktuellen Forschungsarbeiten im Rahmen von AI-NET-PROTECT dienten.

A. Gesamtziel des Vorhabens AI-NET-PROTECT

Das Projekt fokussierte sich darauf, die Digitalisierung in Europa voranzutreiben, wofür eine sichere und skalierbare Kommunikationsinfrastruktur essenziell ist. Diese Infrastruktur soll durch intelligente Netzautomatisierung die Effizienz, Sicherheit und Flexibilität von Kommunikationsnetzen erhöhen. Ein besonderes Augenmerk des Projekts lag auf der Integration von KI zur Optimierung des Netzbetriebs und der Bereitstellung neuer digitaler Dienste und Anwendungen. Das Projekt sollte dazu beitragen, die technologische Souveränität Deutschlands und Europas im Bereich der Kommunikationstechnik zu stärken. Hierfür waren neben akademischen Partnern auch diverse europäische Unternehmen an diesem Vorhaben beteiligt.

B. Wissenschaftliche und Technische Arbeitsziele

AI-NET-PROTECT verfolgt innovative Ansätze zur automatisierten Netzüberwachung und Dienststeuerung. Dabei werden Technologien wie Software-Defined Networking (SDN), Network Functions Virtualization (NFV) und Netztelemetrie mithilfe programmierbarer Netzelemente untersucht, um eine möglichst autonome, agile und intelligente Netzwerkverwaltung zu ermöglichen. Das Projekt ist in verschiedene Teilprojekte gegliedert, die sich jeweils auf spezifische Anwendungsbereiche konzentrieren:

- **Netzwerkautomatisierung und -steuerung**: Entwicklung von KI-basierten Pipelines für die Echtzeit-Datenerfassung und Analyse, um die Netzwerkleistung zu überwachen und darauf aufbauend zu optimieren.
- **Integrierte Sicherheit und Vertrauen**: Implementierung von fortschrittlichen Sicherheitskonzepten zur Erkennung und Abwehr von Cyberangriffen. Hierzu werden u.a. adaptive Überwachungstechniken erforscht.
- **Demonstration und Kommunikation**: Präsentation und Evaluierung der entwickelten Technologien in realen Szenarien, um die Praxistauglichkeit und den Nutzen der Lösungen zu demonstrieren.

2 Darstellung der durchgeführten Arbeiten

Im Rahmen dieses Projekts wurden die operativen und fachlichen Arbeiten auf der Grundlage vorbereitender Tätigkeiten durchgeführt. Die Projektorganisation ermöglichte eine regelmäßige Zusammenarbeit mit den Partnern sowohl durch persönliche als auch virtuelle Treffen. Der dem Projekt zugewiesene wissenschaftliche Mitarbeiter stellte die fristgerechte Erfüllung der Arbeitspakete während der Projektlaufzeit sicher.

Durch die aktive Teilnahme an den monatlichen virtuellen Meetings zu jedem Arbeitspaket konnten wir Fortschritte präsentieren und die Abstimmung mit unseren Projektpartnern gewährleisten.

Während des Projektlaufzeit konzentrierten wir uns in enger Zusammenarbeit mit der TU Chemnitz auf die folgenden Schwerpunktthemen:

- Kollaboratives, maschinelles Lernen zur netzwerkbasierten Erkennung von Angriffen in softwaredefinierten Netzwerken
- Maschinelles Lernen mit Online-Kolaboration zur Angriffserkennung
- Untersuchung von Generativen Adversarial Netzwerken (GANs) und Shift Distribution für C2-Kommunikationsangriffe
- GAN-basierte Generierung von Intrusionsströmen für unausgeglichene Netzwerkverkehrsdaten
- Verbesserung der Generalisierungsfähigkeit von maschinell lernbasierten Netzwerkangriffserkennungssystemen

Das Institut für Multimediale Kommunikation (KOM) ist an verschiedenen Arbeitspaketen beteiligt gewesen, darunter AP0 (2 PM), AP1 (6 PM), AP3 (18 PM), AP4 (3 PM) und AP5 (7 PM) sowie an weiteren Aktivitäten. Die Aktivitäten von KOM in diesen Arbeitspaketen beziehen sich auf die verteilte lokale Überwachung von Kommunikationsnetzen unter Verwendung von KI-Methoden im Rahmen von AI-NET-PROTECT. Die Einzelheiten der Arbeitspakete sind im Nachfolgenden zusammengefasst:

AP0 - Projektkoordination und horizontale Aktivitäten

Das Hauptziel dieses Arbeitspakets bestand darin, die Koordination der Projektaktivitäten effektiv zu gestalten, um die Gesamtziele des Projekts zu erreichen. Zu diesem Zweck nahm mindestens ein wissenschaftlicher Mitarbeiter der TUDa regelmäßig an virtuellen und persönlichen Treffen der verschiedenen Teil-Arbeitspakete sowie des gesamten Projektverbunds teil. Darüber hinaus wurde der Fortschrittsbericht des Projekts alle sechs Monate in einer Plenarsitzung vorgestellt. Zusätzlich wurde jedes Jahr ein Zwischenbericht erstellt, um den Fortschritt des Projekts zu dokumentieren. Durch diese Treffen und Berichte konnte sichergestellt werden, dass alle Aktivitäten in die zielführend verlaufen, um die Projektziele zu erreichen.

AP1 - Anforderungen, Architektur und technisch-ökonomische Analyse kritischer End-to-End-Netzwerke

Das Hauptziel dieses Arbeitspakets war die Entwicklung und Erprobung von Netzwerkarchitekturen, die eine sichere Konnektivität für Unternehmen bieten können. Die TU Darmstadt hat sich auf das Management und die Absicherung programmierbarer Netzwerke unter Verwendung von Kerntechnologien wie Software-Defined-Networking (SDN) und P4 (Programming Protocol-Independent Packet Processors) fokusiert. Der Schwerpunkt dieser Arbeiten lag auf der Anwendung von Techniken des maschinellen Lernens zur Verbesserung des Sicherheitsmanagements in Netzwerken. AP1 bildete dabei die Grundlage für alle weiteren Arbeitspakete.

Um eine ganzheitliche Method für das maschinelle Verarbeiten von Daten(eine so genannte Pipeline) für dieses Projekt zu schaffen, haben wir in diesem Arbeitspaket eine dynamische Vorverarbeitungspipeline entworfen, die in allen anderen Arbeitspaketen des Projekts verwendet werden konnte. In dieser Pipeline werden die Netzwerkverkehrsmerkmale untersucht und die relevantesten Merkmale ausgewählt, die dann in den Modellen des maschinellen Lernens trainiert werden. Um die Richtigkeit der ausgewählten Merkmale sicherzustellen, wurde auch Expertenwissen hinzugezogen. Zu diesem Zweck führten wir mehrere Meetings mit Projektpartnern durch, die Experten im Bereich Netzwerksicherheit sind, und diskutierten, warum jedes dieser Merkmale für die Erkennung von Netzwerkanomalien von Bedeutung sein könnte.

AP3 – KI-gesteuerte Netzwerkkontrolle und Dienstinformationen

Das Hauptziel dieses Arbeitspakets war die Entwicklung und Implementierung automatisierter Netzwerkmanagementsysteme, die auf künstlicher Intelligenz basieren. Ein zentrales Ziel bestand darin, KI-gestützte Lösungen für den Betrieb und das Management der gesamten SDN-Steuerebene nahtlos zu entwickeln. Der besondere Schwerpunkt von TUDa lag im Bereich des Netzwerk-Sicherheitsmanagements, insbesondere bei der Netzwerküberwachung in SDN. Unsere Modelle kamen nicht nur in der SDN-Steuerebene, sondern auch in der programmierbaren Datenebene, konkret in P4-Switches, zum Einsatz.

AK 3.1 - Datenanalyse, ML-Pipelines für Netzwerk- und Dienstmanagement

In dieser Aktivität (AK) des AP 3 haben wir mehrere öffentlich zugängliche Netzwerkverkehrsdatenbanken wie CICIDS2017 [10], CTU-13 [11], CICDoS [12], Botnet [13] und UNSW-NB [14] in unsere Evaluierungsanalyse einbezogen. Diese Betrachtung ermöglichte es uns, Einblicke in unterschiedliche Netzwerkverkehrsmuster zu gewinnen und aufzuzeigen, wie heterogen der Netzwerkverkehr in verschiedenen Anwendungsszenarien sein kann. Unsere explorative Datenanalyse hat zeigt, dass jedes Netzwerkverkehrsmuster erhebliche Ungleichgewichte und spezifische Verteilungen aufweist, die Herausforderungen beim Training von maschinellen Lernmodellen darstellen und deren Erkennungsleistung bei neuen Netzwerkverkehrsmustern beeinträchtigen können. Diese Dynamik kann sich zudem auf die in der Vorverarbeitungspipeline entworfene Extraktion von Flussmerkmalen auswirken.

Daher lag der weitere Fokus in diesem Arbeitspaket auf der Untersuchung der Unterschiede im Netzwerkverkehr und deren Auswirkungen auf das Training von maschinellen Lernmodellen. Ein tiefes Verständnis der Netzwerkverkehrsmuster und ihrer Unterschiede führte uns dazu, unsere Vorverarbeitungspipeline zu erweitern und sie in Bezug auf ihre Generalisierungsfähigkeit zu verbessern.

AK 3.5 – Erweiterte Überwachung und KI-gestützter Netzwerkbetrieb

Das Ziel dieser Aktivität war die Entwicklung von KI-gesteuerten Netzwerkmanagementsystemen, die komplexe Netzwerke verwalten können – eine Notwendigkeit aufgrund der zunehmenden Heterogenität von Anwendungen und Diensten, die in AK 3.1 untersucht wurde. Um dieses Ziel zu erreichen, haben wir verschiedene Aspekte eines KI-basierten Intrusion-Detection-Systems untersucht, darunter die Generalisierungsfähigkeit, Anpassungsfähigkeit und Einsetzbarkeit in neuen Netzwerkarchitekturen wie SDN.

Zur Verbesserung der Generalisierungsfähigkeiten von maschinell lernbasierten Intrusion-Detection-Systemen haben wir ein neuartiges Modell namens SSCL-IDS [6] entwickelt, das einen selbstüberwachten Lernalgorithmus umfasst. Dieser Ansatz kann die Erkennungsleistung des Intrusion-Detection-Systems verbessern, insbesondere bei unausgewogenen Netzwerkverkehrsmustern mit einer Mehrheit an benignen Flüssen und

einer Minderheit an Intrusionsflüssen. Darüber hinaus ermöglicht der Einsatz des selbstüberwachten Ansatzes die Verbesserung der Generalisierung des maschinellen Lernens bei der Erkennung neuer, zuvor unbekannter Netzwerkverkehrsmuster. Außerdem kann dieser Ansatz dem Mangel an annotierten Netzwerkverkehrsproben zur Schulung der ML-Modelle begegnen, was die Erkennungsleistung des Modells beeinträchtigen könnte. Tatsächlich wurde unser vorgeschlagenes SSCL-IDS ausschließlich auf unauffälligen Datenflüssen trainiert, sodass kein Training mit Angriffsdaten erforderlich war, während es dennoch in der Lage war, neue, zuvor unbekannte Netzwerkverkehrsflüsse bzw. deren Muster korrekt zu erkennen.

Neben der Generalisierung des entwickelten maschinell lernbasierten Intrusion-Detection-Systems war unser Ziel, den neuen Ansatz in einer SDN-basierten Netzwerkarchitektur einzusetzen. Dies bringt jedoch mehrere Herausforderungen mit sich. Wenn sie ausschließlich in der SDN-Steuerebene eingesetzt werden, steigt das Risiko von Angriffen auf die Steuerebene, es gibt Verzögerungen bei der Erkennung, und die Latenz nimmt zu. Andererseits ist die Erkennungsleistung, wenn sie ausschließlich in der Datenebene eingesetzt werden, auf einfache ML-Modelle beschränkt. Um diese Herausforderungen zu überwinden, haben wir in diesem Arbeitspaket ein kollaboratives Framework namens CML-IDS [5] entwickelt. Dieses Framework ermöglicht die gleichzeitige Nutzung von ML-Modellen in der Steuerebene und der Datenebene. Basierend auf einem vordefinierten Schwellenwert werden die Ergebnisse als vertrauenswürdig angesehen, wenn das Vertrauen des in der Datenebene implementierten maschinellen Lernmodells hoch genug ist; andernfalls werden die Flussmerkmale an die Steuerebene weitergeleitet, wo ein Ensemble-Ansatz mit höherer Lernkapazität zum Einsatz kommt. Durch die Nutzung dieses Frameworks konnten wir eine hohe Erkennungsleistung erzielen, während die Netzwerk-Latenz reduziert und die Erkennungsgeschwindigkeit erhöht wurde.

Um die Anpassungsfähigkeit des vorgeschlagenen CML-IDS weiter zu verbessern, untersuchen wir die Integration von Online-Lerntechniken, was uns dazu führte, das COML-IDS [8] zu entwickeln. In diesem Ansatz werden die Modelle unter bestimmten Bedingungen dynamisch neu trainiert, um eine kontinuierliche Verbesserung sicherzustellen. Um optimale Bedingungen zu erreichen, haben wir zunächst verschiedene Konzeptdrifts [4] untersucht und anschließend das Online-Lernen in das kollaborative, maschinell lernbasierte Intrusion-Detection-System (CML-IDS) integriert.

WP4 – Integrierte Sicherheit und Vertrauen

Das Arbeitspaket 4 konzentriert sich auf den Schutz kritischer Daten und Infrastrukturen vor Angriffen durch Eindringlinge. In diesem Arbeitspaket lag unser Schwerpunkt darauf, den Bedarf an menschlicher Interaktion, insbesondere bei der Annotierung von Netzwerkverkehr, zu reduzieren und Methoden zur Erkennung von KI-basierten Netzwerkangriffen zu entwickeln. Dieses Arbeitspaket wurde in Zusammenarbeit mit der TU Chemnitz durchgeführt.

Die Ergebnisse dieses Arbeitspakets konzentrierten sich auf die Erhöhung der Automatisierung von Intrusion-Detection-Systemen unter der Nutzung von Maschine-Learning-Methoden sowie auf die Verbesserung ihrer Robustheit gegenüber Angriffen, die durch GAN-Modelle generiert wurden. Durch die entwickelten Ansätze konnte die Notwendigkeit menschlicher Annotation reduziert und gleichzeitig die Fähigkeit zur Erkennung von Angriffen, die durch KI-basierte Methoden erzeugt wurden, gesteigert werden.

Im Fokus des Aktivität 4.2 stand zunächst die automatische Annotation von Netzwerkverkehrsproben, um den Bedarf an menschlicher Unterstützung zu reduzieren und die Erkennungsleistung von Intrusion-Detection-Systemen zu verbessern. Zu diesem Zweck wurde ein aktiv lernbasierter Ansatz entwickelt, der nur die informativsten Datenflüsse zur Annotation durch Menschen auswählt und lernt, andere Datenflüsse auf deren Grundlage zu annotieren. Diese Ergebnisse sind ein Erfolg der sehr erfolgreichen Zusammenarbeit zwischen der TU Darmstadt und der TU Chemnitz.

Darüber hinaus haben wir eine Datenverarbeitunspipeline entwickelt, um neue Netzwerkangriffe unter Verwendung von Generativen Adversarial Netzwerken (GANs) zu erzeugen, um das ML-basierte Intrusion-Detection-System mit diesen neuen Angriffen zu trainieren und die Robustheit des Modells zur Erkennung von KI-basierten Netzwerkangriffen weiter zu verbessern. Insbesondere nutzten wir fortschrittliche Command-and-Control (C2) Kommunikationsangriffe, indem wir diese mit GAN-Modellen generieren und ihre Wirksamkeit durch die Analyse von Korrelationen zwischen Datenflussmerkmalen überprüfen.

In dieser Arbeit konzentrierten wir uns darauf, veränderbare Merkmale zu identifizieren, die von Angreifern manipuliert werden können, und deren Auswirkungen auf andere Paketmerkmale zu bewerten und um die Erzeugung realistischer Netzwerkverkehrspakete sicherzustellen.

AP5: Demonstration & Externe Kommunikation

Die in den technischen Arbeitspaketen (AP1 bis AP4) entwickelten Konzepte und Lösungen wurden auf verschiedenen wissenschaftlichen Tagungen und Konferenzen demonstriert und vorgestellt. Diese Präsentationen erhielten bedeutende Anerkennung, einschließlich Auszeichnungen für die Forschungsneuheit, wie beispielsweise den Best Paper Award auf einer wissenschaftlichen Konferenz. Auch die Vorträge selbst wurden ausgezeichnet. Darüber hinaus arbeitete die TU Darmstadt eng mit anderen AI-NET-Partnern wie der TU Chemnitz und Consistec zusammen, um die Forschung weiter voranzutreiben und den Wissensaustausch zu fördern.

3 Zusammenfassung der erzielten Erkenntnisse und Ergebnisse

In diesem Abschnitt werden die Ergebnisse der Arbeitspakete und darin entwickelten neuen Methoden erläutert:

AP1 - Anforderungen, Architektur und technisch-ökonomische Analyse kritischer End-to-End-Netzwerke

In diesem Arbeitspaket wurde eine allgemeine Vorverarbeitungspipeline entworfen, die Netzwerkverkehr in Flussmerkmale umwandeln kann. Diese Vorverarbeitung ist nicht nur für die anderen Arbeitspakete nützlich, sondern auch für jedes maschinell lernbasierte Intrusion-Detection-System von großem Wert. Die Funktionsweise und Evaluierung dieser Vorverarbeitungspipeline anhand verschiedener verfügbarer Netzwerkverkehrsdaten wurde in einer wissenschaftlichen Publikation [1] beschrieben. Diese Pipeline bildet eine grundlegende Basis für die weiteren Arbeiten in diesem Projekt.

AP3 – KI-gesteuerte Netzwerkkontrolle und Dienstinformationen

In diesem Arbeitspaket lag der Fokus auf der Entwicklung und Gestaltung von maschinell lernbasierten Intrusion-Detection-Systemen, die generalisiert sind. Das bedeutet, dass sie auch bei zuvor unbekannten Netzwerkverkehrsmustern funktionieren und in die programmierbare Datenebene von Netzwerken eingebettet werden können, um die Erkennungsgeschwindigkeit signifikant zu erhöhen und die Gesamtlatenz im Netzwerk gering zu halten.

AK 3.1 - Datenanalyse, ML-Pipelines für Netzwerk- und Dienstmanagement

In diesem Arbeitspaket wurden verschiedene gewöhnliche Netzwerkverkehrsmuster untersucht, um zu verstehen, wie und wann sich diese Muster unterscheiden können. Um eine generalisierte Flussmerkmalextraktion zu entwerfen war es wichtig, den entwickelten Ensemble-Feature-Selection-Ansatz [1] zu erweitern und seine Generalisierungsfähigkeit zu verbessern. Zu diesem Zweck haben wir einen datengetriebenen Ansatz entwickelt, um die relevantesten Merkmale basierend auf verschiedenen Netzwerkverkehrsmustern auszuwählen. Dieser Ansatz, DD-EFS [2], und seine Ergebnisse wurden auf einer wissenschaftlichen Konferenz veröffentlicht und präsentiert.

AK 3.5 – Erweiterte Überwachung und KI-gestützter Netzwerkbetrieb

Im ersten Schritt des AK 3.5 wurde die Generalisierungsfähigkeit der maschinell lernbasierten Intrusion-Detection-Systeme verbessert. Zu diesem Zweck entwarfen wir das SSCL-IDS [6], das ausschließlich mit benignen Proben trainiert wurde, um dem Mangel an annotierten Intrusionsproben zu begegnen. Dank seines neuartigen Trainingsprozesses im selbstüberwachten Modus war es in der Lage, neue, zuvor unbekannte Netzwerkverkehrsflüsse mit hoher Leistung zu erkennen. Durch den Einsatz von SSCL-IDS [6] konnten wir die Generalisierungsfähigkeit unseres zuvor entwickelten ML-basierten Intrusion Detection Systems erheblich steigern. Dieser Ansatz und die entsprechenden Ergebnisse wurden auf der IFIP Networking 2024-Konferenz veröffentlicht.

Der nächste Schritt bestand darin, das entwickelte ML-basierte IDS in einer SDN-basierten Netzwerkarchitektur bereitzustellen, um eine hohe Erkennungsleistung zu erzielen und gleichzeitig die Erkennungsgeschwindigkeit zu erhöhen sowie die Netzwerklatenz zu verringern. Um diese Ziele zu erreichen, entwarfen wir den Ansatz CML-IDS [5], ein kollaboratives ML-basiertes Framework, das eine Zusammenarbeit zwischen einem leichtgewichtigen ML-Modell, das in der programmierbaren Datenebene eingesetzt wird, und einem Ensemble-Modell mit höherer Kapazität, das in der Steuerungsebene implementiert ist, ermöglicht. Unsere Ergebnisse zeigen, dass dieses Framework die Erkennungsleistung im Vergleich zu Szenarien erhöht, in denen lediglich ein ML-Modell in der Datenebene verwendet wird, und zudem die Erkennungsgeschwindigkeit steigert sowie die Netzwerklatenz im Vergleich zu Fällen, in denen ausschließlich ein ML-Modell in der Steuerungsebene eingesetzt wird, reduziert. Diese Arbeit wurde auf der IEEE/IFIP CNSM 2023-Konferenz präsentiert und erhielt den Best Paper Award. Darüber hinaus wurde der Programmcode dieser Arbeit der Öffentlichkeit über ein GitHub-Repository (https://github.com/golchinpg/CML-IDS) zur Verfügung gestellt.

Der nächste Schritt bestand darin, die Anpassungsfähigkeit des CML-IDS zu verbessern, insbesondere im Hinblick auf neue, zuvor unbekannte Netzwerkverkehrsmuster. Um dieses Ziel zu erreichen, untersuchten wir zunächst die möglichen Netzwerkdrifts und deren Auswirkungen auf die Netzwerkverkehrsmuster [4]. Diese Ergebnisse wurden auf dem KuVS-MaLeNe-Workshop 2023, der insbesondere eine Deutsche Community adressiert, veröffentlicht. Basierend auf diesen Ergebnissen integrierten wir später einen Online-Lernansatz in das CML-IDS, der als COML-IDS [8] bezeichnet wird. In diesem Ansatz wird das Modell kontinuierlich basierend auf bestimmten Kriterien aktualisiert. Dieser Ansatz und die Ergebnisse wurden zur Veröffentlichung auf der IEEE NFV-SDN-Konferenz 2024 angenommen.

AP4 – Integrierte Sicherheit und Vertrauen

Die nachfolgenden Ergebnisse beziehen sich auf AK4.2, basieren allerdings auch auf den Ergebnissen aus AP3. Als erster Schritt zur Erhöhung der Automatisierung von maschinell lernbasierten Intrusion-Detection-Systemen lag der Fokus auf der Automatisierung der Annotation von Netzwerkverkehrsproben. Die manuelle Annotation durch Experten ist eine aufwändige Aufgabe, die für einige spezifische Netzwerkverkehrsmuster ohne KI-Unterstützung schwierig oder unmöglich sein kann. Um dieses Ziel zu erreichen, haben wir einen aktiven, lernbasierten Ansatz entwickelt, der die informativsten Flussproben zur Annotation durch Experten auswählt und automatisch andere Proben annotiert, um den Bedarf an menschlicher Unterstützung zu reduzieren. Diese Arbeit war eine erfolgreiche Zusammenarbeit zwischen der TU Darmstadt und der TU Chemnitz. Die Ergebnisse wurden auf der NOMS 2023 Konferenz veröffentlicht.

Darüber hinaus haben wir zur Verbesserung der Robustheit von ML-basierten Intrusion Detection Systemen gegen manipulierte Netzwerkverkehrsflüsse, die durch KI-Methoden erzeugt werden, eine Datenverarbeitungspipeline entwickelt, um valide, veränderte Angriffsproben mithilfe von Generative Adversarial Networks (GANs) für hochentwickelte Command-and-Control (C2) Kommunikationsangriffe zu generieren. In dieser Arbeit haben wir untersucht, welche Flussmerkmale veränderbar sind, während die Gültigkeit der veränderten Probe erhalten bleibt, und wie die Robustheit der maschinell lernbasierten Intrusion-Detection-Systeme erhöht werden kann. Auch diese Arbeit ist eine erfolgreiche Kooperation zwischen der TU Chemnitz und der TU Darmstadt. Die Arbeiten sind abgeschlossen.

4 Erläuterungen zum zahlenmäßigen Nachweis

Der zahlenmäßige Nachweis der Projektausgaben beinhaltet Ausgaben der folgenden Positionen:

- Personalkosten wissenschaftliche Mitarbeiter (0812)
- Personalkosten wissenschaftliche Hilfskräfte (0822)
- Gegenstände bis 410€ (0831)
- Gegenstände ab 410€ (0850)
- Nationale und internationale Reisekosten (0846)
- Sonstige allgemeine Verwaltungskosten (0843)

Personalkosten: Die Arbeiten in diesem Projekt wurden durch einen wissenschaftlichen Mitarbeiter in der Entgeltgruppe E13 durchgeführt. Dieser leitete weitere wissenschaftliche Hilfskräfte an, die beispielsweise einzelne Implementierungsaufgaben übernommen haben.

Gegenstände: Im Rahmen des Projekts wurde ein SDN-fähiger Netzwerkswitch (0850) sowie zugehöriges Zubehör (bspw. Netzwerkkabel) angeschafft. Mithilfe dieses Switches wurde die Control-Data-Plane Problematik in SDN-Netzwerken aus Monitoring-Sicht betrachtet.

Reisen: Zum Zwecke des Austauschs zwischen den Projektpartnern wurden diverse Reisen innerhalb Europas zu AI-NET-PROTECT-Projektmeetings durchgeführt. Aufgrund der COVID-19 Pandemie waren dies jedoch weniger Reisen als zum Zeitpunkt der Antragstellung erwartet. Des Weiteren wurden Reisen zu internationalen Konferenzen durchgeführt, auf denen wissenschaftlicher Austausch praktiziert und Forschungsergebnisse des Projekts vorgestellt wurden.

Verwaltungskosten: Das AI-NET-PROTECT-Projekt wurde durch CELTIC-NEXT operativ betreut. Für diese Betreuung entstanden (gemäß des Projektantrags) laufende Kosten.

5 Notwendigkeit und Angemessenheit der geleisteten Projektarbeiten

Aufgrund des bisher nicht hinreichend erforschten Themengebiets um AI-NET-PROTECT haben sich einige essenzielle Forschungsfragen aufgetan, die im Rahmen dieses Projekts erforscht wurden. Dies war nötig, da das Themenfeld der Netzwerkautomation sowie der Einsatz von KI essenziell für zukünftige Netzwerke sein wird, um den Herausforderungen und Bedrohungen im 21. Jahrhundert gerecht zu werden.

Da es der TUD jedoch nicht möglich war, die Finanzierung des wissenschaftlichen Mitarbeiters und dem darüberhinausgehenden Aufwand selbst zu übernehmen, wurden die im Antrag spezifizierten Drittmittel zur Finanzierung des Vorhabens zu 100% beantragt.

Die Angemessenheit der geleisteten Projektausgaben ergibt sich aus den tariflich bedingten Gehältern von wissenschaftlichem Personal in Deutschland. Die Notwendigkeit dieses Personaleinsatzes ergibt sich aus den geplanten und durchgeführten Arbeitspaketen gemäß des Projektantrags, die während der Projektlaufzeit durchgeführt wurden und zu den hochkarätigen Forschungsergebnissen dieses Vorhabens geführt haben. Neben den beantragten und bewilligten finanziellen Ressourcen waren für die Durchführung dieses Projekts keine weiteren Mittel notwendig.

6 Nutzbarkeit und Verwertbarkeit der Projektergebnisse

Die Ergebnisse des Vorhabens wurden in verschiedenen, wissenschaftlichen wie anwendungs-bezogenen, Publikationen aufbereitet. Analog zum bisherigen Vorgehen bei zurückliegenden Forschungsprojekten dienen die vorgestellten Ergebnisse als Grundlage für zukünftige Forschungsvorhaben und -projekte. Dies ist insbesondere wichtig, um eine wissenschaftliche Kontinuität zu erreichen.

Die Ergebnisse aus diesem Projekt für die KI-basierte Anomalie-Erkennung in SDN-fähigen Netzwerkswitches bietet vielseitiges Potential für weitere Arbeiten sowohl im akademischen als auch industriellen Umfeld. Neben den vielen Erkenntnissen sind aber ebenso viele neue Folgefragestellungen entstanden. Beispielsweise konnte gezeigt werden, wie einfache KI-gelernte Modelle in marktverfügbaren Netzwerkswitches abgebildet werden können. Offen sind jedoch Fragestellungen hinsichtlich der dynamischen Anpassbarkeit sowie der Umsetzung von komplexeren Modellen. Hierfür könnte zukünftig beispielsweise auch der Einsatz von KI-Hardwaretechnologien, wie beispielsweise TPUs, im Netzwerkkontext untersucht werden.

7 Fortschritt auf dem Gebiet des Vorhabens während der Projektlaufzeit

Im Verlauf dieses Projekts gab es verschiedene inkrementelle Neuheiten, sowohl im akademischen als auch im industriellen Kontext. Diese Erkenntnisse haben wir entweder genutzt, um unsere vorgeschlagenen Methoden zu verbessern, oder mit denen wir unsere Arbeit verglichen, um die Verbesserung unserer Ergebnisse zu zeigen.

Ein Beispiel hierfür ist die Entwicklung der allgemeinen Vorverarbeitungspipeline die mithilfe des NFStream-Tools [9], vorgestellt im Jahr 2022, zur Flussmerkmalextraktion entwickelt wurde. In unserer vorgeschlagenen Vorverarbeitungspipeline haben wir zusätzlich eine Methode entwickelt, um die relevantesten Merkmale auszuwählen.

Grundlegende Fortschritte auf dem Gebiet des Vorhabens, die eine fundamentale Neuausrichtung bedürft hätten, gab es nicht.

8 Veröffentlichungen im Rahmen des Vorhabens

Die nachfolgende, chronologisch sortierte, Liste an Publikationen gibt einen Überblick über alle bisher veröffentlichten Ergebnisse auf wissenschaftlichen Konferenzen oder in Form von Dissertationen. Der Bezug zu AI-NET-PROTECT wurde jeweils durch ein Acknowledgement sichergestellt.

[1] Pegah Golchin, Ralf Kundel, Tim Steuer, Rhaban Hark, and Ralf Steinmetz: Improving ddos attack detection leveraging a multi-aspect ensemble feature selection. NOMS 2022 - 2022 IEEE/IFIP Network Operations and Management Symposium, pages 5, 04/2022.

[2] Mehrdad Hajizadeh, Sudip Barua, Pegah Golchin. FSA-IDS: A Flow-based Self-Active Intrusion Detection System. IEEE/IFIP Network Operations and Management Symposium (NOMS), pages 9, IEEE, 05/2023.

[3] Pegah Golchin, Jannis Weil, Ralf Kundel, and Ralf Steinmetz: Dynamic network intrusion detection system in Software-Defined Networking. Proceedings of the MaLeNe workshop “Network Softwarization”, pages 2, KuVS, 09/2023.

[4] Pegah Golchin, Chengbo Zhou., Pratyush Agnihotri, Mehrdad Hajizadeh, Ralf Kundel, Ralf Steinmetz: CML-IDS: Enhancing Intrusion Detection in SDN Through Collaborative Machine Learning. 19th International Conference on Network and Service Management (CNSM), pages 9. IEEE/IFIP, 11/2023.

[5] Pegah Golchin, Nima Rafiee, Mehrdad Hajizadeh, Ahmad Khalil, Ralf Kundel, Ralf Steinmetz: SSCL-IDS: Enhancing Generalization of Intrusion Detection with Self-Supervised Contrastive Learning. IFIP Networking Conference, pages 9, IEEE, 06/2024.

[6] Pegah Golchin, Nima Rafiee, Ralf Kundel: SSCL-IDSA Data-Driven Solution for Improving Transferability of Traffic Flow Feature Selection. IFIP Networking Conference, pages 3, IEEE, 06/2024.

[7] Pegah Golchin: Machine Learning Models in Network Intrusion Detection Systems : Self-Supervised Detection of Malicious Flows and Traffic Patterns Recognition in Programmable Networks. Accepted dissertation at TU Darmstdt, pages 165, TU prints, 08/2024.

[8] Pegah Golchin, Chengbo Zhou, Hengyu Liu, Björn Scheuermann, Ralf Kundel, Tobias Meuser. Integrating Online Learning with Collaborative Machine Learning for Continuous Intrusion Detection in SDN. IEEE Conference on Network Function Virtualization and Software Defined Networks, pages 6, IEEE, 11/2024, *accepted to be published in conference proceedings.*

9 Weitere Referenzen

[9] Aouini, Zied, and Adrian Pekar. "NFStream: A flexible network data analysis framework." *Computer Networks* 204 (2022): 108719.

[10] Sharafaldin, I., Lashkari, A.H. and Ghorbani, A.A., 2018. Toward generating a new intrusion detection dataset and intrusion traffic characterization. ICISSp, 1, pp.108-116.

[11] Garcia, S., Grill, M., Stiborek, J. and Zunino, A., 2014. An empirical comparison of botnet detection methods. computers & security, 45, pp.100-123.

[12] Jazi, H.H., Gonzalez, H., Stakhanova, N. and Ghorbani, A.A., 2017. Detecting HTTP-based application layer DoS attacks on web servers in the presence of sampling. Computer Networks, 121, pp.25-36.

[13] Beigi, E.B., Jazi, H.H., Stakhanova, N. and Ghorbani, A.A., 2014, October. Towards effective feature selection in machine learning-based botnet detection approaches. In 2014 IEEE Conference on Communications and Network Security (pp. 247-255). IEEE.

[14] Moustafa, N. and Slay, J., 2015, November. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). In 2015 military communications and information systems conference (MilCIS) (pp. 1-6). IEEE.