Energieforschungsprogramm

Publizierbarer Endbericht

Programmsteuerung:
Klima- und Energiefonds

Programmabwicklung:
Österreichische Forschungsförderungsgesellschaft mbH (FFG)

Endbericht
erstellt am
10/01/2019

Projekttitel:

Ausfallssicherheit digitalisierter Verteilungsnetze für elektrische Energie

Projektnummer: 853660

Ausschreibung	2. Ausschreibung Energieforschungsprogramm
Projektstart	01/04/2016
Projektende	31/10/2018
Gesamtprojektdauer (in Monaten)	31 Monate
ProjektnehmerIn (Institution)	FH St. Pölten Forschungs GmbH
AnsprechpartnerIn	Paul Tavolato
Postadresse	Matthias-Corvinus-Straße15, 3100 St. Pölten
Telefon	+43 2742/313 228-200
Fax	+43 2742/313 228-339
E-mail	paul.tavolato@fhstp.ac.at
Website	www.fhstp.ac.at

Ausfallssicherheit digitalisierter Verteilungsnetze für elektrische Energie

Substation Security

AutorInnen:
Institut für IT Sicherheitsforschung, FH St. Pölten Forschungs GmbH:
Paul Tavolato
Philipp Kreimel
Henri Ruotsalainen
Ersi Hodo
Wels Strom GmbH:
Markus Aigner
Siemens AG Österreich:
Alexander Sadofsky
Walter Wutzl

1 Inhaltsverzeichnis

2 Einleitung

2.1 Problemstellung

Der Begriff „Smart Grid“ wird von der österreichischen Plattform Smart Grids Austria in Übereinstimmung mit internationalen Definitionen folgendermaßen definiert: *„Smart Grids sind Stromnetze, welche durch ein abgestimmtes Management mittels zeitnaher und bidirektionaler Kommunikation zwischen Netzkomponenten, Erzeugern, Speichern und Verbrauchern einen energie- und kosteneffizienten Systembetrieb für zukünftige Anforderungen unterstützen“* [40].

Das US Department of Energy identifiziert vier Technologiebereiche, die für die Entwicklung eines Smart Grid wesentlich sind (zitiert nach [41]):

- Integrierte, automatische Kommunikation zwischen Komponenten des Elektrizitätsnetzes
- Sensor- und Metrologie-Technologien
- Automatische Kontrolle von Verteilung und Wartung
- Verbesserte Management Informationssysteme

Gemeinsam ist diesen Definitionen die Integration von Informations- und Kommunikations-technologien (IKT) in die Erzeugung, die Verteilung und den Verbrauch von elektrischer Energie.

Diese Integration hat allerdings zur Folge, dass Cyber-Sicherheit in den Bereichen Erzeugung, Verteilung und Verbrauch von elektrischer Energie plötzlich als neues Thema auftaucht. Durch die Möglichkeit von Cyber-Angriffen auf die Stromversorgungsinfrastruktur entsteht ein nicht zu unterschätzendes Gefahrenpotential, dem mit entsprechenden Maßnahmen und technischen Vorkehrungen begegnet werden muss.

Die Anzahl der potentiellen Angriffspunkte auf das Elektrizitätsversorgungsnetz wird stark erhöht. Durch die Vernetzung des Systems ist außerdem ein lokaler Angriff nicht mehr auf einen Angriffspunkt beschränkt, sondern kann sich über das gesamte Netz ausbreiten. Wie sich in der Vergangenheit bereits gezeigt hat, sind stark vernetzte Systeme aus dem Bereich kritischer Infrastrukturen und insbesondere aus dem Bereich der Energieerzeugung und Energieverteilung bereits Ziel von Cyber-Angriffen geworden (siehe dazu die halbjährlichen Berichte des im US Department of Homeland Security angesiedelten ICS-CERT – Industrial Control Systems Cyber Emergency Response Teams [39]). Sowohl für Gruppen mit kriminellen, auf finanziellen Erfolg abzielenden Absichten als auch für Gruppen mit terroristischen Motiven stellen solche Systeme lohnende Ziele dar. Die Erforschung von Technologien und Verfahren zur Absicherung aller Komponenten des Stromnetzes ist daher unumgänglich.

Wesentliche Komponenten bei der Energieverteilung sind die Umspannwerke und die Ortsnetzstationen (Trafo-Stationen). Hier führt die Umsetzung der Smart Grid Anforderungen

zu einem erhöhten Automatisierungsgrad, der nur mit dem Einsatz von IKT bewerkstelligt werden kann. Das eröffnet eine Reihe neuer bisher nicht vorhandener Sicherheitsprobleme. In Umspannwerken und Ortsnetzstationen hat sich heute eine Reihe moderner Technologien weitgehend durchgesetzt, die unter anderem folgendes umfasst: Mikroprozessor basierte Intelligent Electronic Devices (IEDs), Standard Protokolle wie TCP/IP, Ethernet und Anbindung an Wide Area Networks (WAN) unter Verwendung von Internet Technologien. Der Fernzugriff auf IEDs oder auf die interne Benutzerschnittstelle des Kommunikationsnetzes im Zuge von Fernwartungstätigkeiten ist mittlerweile durchwegs üblich. Genau daraus ergibt sich ein nicht unbeträchtliches Gefahrenpotential: Durch die Verwendung von unsicheren Standardprotokollen, von aus der Ferne steuerbare IEDs oder durch nicht legitimierten Zugriff können schwerwiegende Folgen für die gesamte Stromversorgung entstehen.

Natürlich werden bereits jetzt bekannte Sicherheitstechnologien wie Firewalls und Verschlüsselung auch für die in Umspannwerken und Ortsnetzstationen verwendeten Informations- und Kommunikationskomponenten eingesetzt. Trotzdem können die Gefahren dadurch nur in beschränktem Maße abgemildert werden. Der Grund dafür besteht im Wesentlichen darin, dass diese Technologien generell für ein anderes Einsatzgebiet (konventionelle IT Systeme) entwickelt worden und daher nicht auf die besonderen Erfordernisse der Hard- und Softwarearchitektur eines Umspannwerks bzw. einer Ortsnetzstation ausgerichtet sind. Vorhandene Technologien sind zum Beispiel nicht dafür geeignet, die Netzwerk Pakete in diesen Kommunikationsnetzen zu interpretieren, da diese auf Industrieprotokollen basieren, die sonst nicht zum Einsatz kommen (im Wesentlichen handelt es sich dabei um Protokolle Familien IEC 60870 und IEC 61850).

Ein weiteres Charakteristikum der Sicherheitsanforderungen von Kommunikationsnetzen im Bereich der Energieverteilung besteht darin, dass es sich bei diesen Systemen um eingebettete Systeme handelt, in denen ausschließlich Komponenten mit beschränkten Ressourcen (in Bezug auf Speicherkapazität, Prozessorgeschwindigkeit) verwendet werden. Jede Lösung muss diese Tatsache berücksichtigen und entsprechend dimensioniert sein.

Da die Sicherheit beim Betrieb eines Umspannwerks bzw. einer Ortsnetzstation ein unabdingbarer Aspekt bei der Verwirklichung eines Smart Grids ist, ist die Forschung im Bereich der Methoden und Technologien zur Überwachung der entsprechenden Kommunikationsnetze im Umspannwerk eine wesentliche Voraussetzung für den sicheren Betrieb eines Smart Grids. Allerdings befindet sich der Forschungsbereich, der sich mit Sicherheitsaspekten bei der Integration von IKT in Energieverteilungsnetzen und speziell in Umspannwerke und Ortsnetzstationen befasst, noch in seinen Anfängen. [38]

Die IKT-Architektur der Netzwerke in Umspannwerken und Ortsnetzstationen kann in zwei Teilbereiche unterteilt werden:

das Schutznetz, das im Wesentlichen IEDs zur der Absicherung vor Auswirkungen von Fehlern im Energieübertragungsnetz (Kurzschluss, Erdschluss, …) enthält, und

das Automatisierungsnetz, das IEDs und RTUs für verschiedene Aufgaben im Zusammenhang mit dem automatisierten Betrieb des Umspannwerkes enthält.

Diese beiden Teilbereiche arbeiten weitgehend unabhängig voneinander, sind jedoch physisch im Kommunikationsnetz miteinander verbunden. Zusätzlich ist im Allgemeinen noch ein Arbeitsplatz für manuelle Bedienung (egal ob vor Ort oder aus der Ferne) ins Netz integriert. In Ortsnetzstationen ist im Allgemeinen nur das Automatisierungsnetz vorhanden.

In beiden Netzen werden unterschiedliche Varianten der Protokolle IEC 60870 und IEC 61850 eingesetzt: meistens IEC 61850-GOOSE, IEC 61850-MMS und IEC 60870-5-103 für das Schutznetz und IEC 60870-5-104 für das Automatisierungsnetz. Die Protokolle dieser Protokollfamilien verwenden als Träger TCP/IP und/oder Ethernet. Durch die Verwendung dieser weitverbreiteten Standardprotokolle (TCP/IP, Ethernet) und durch den Anschluss der Netze an potentiell unsichere Übertragungsnetze (z.B. Internet) ändert sich die Sicherheitssituation grundlegend: waren die Netze vorher in sich abgeschlossen und auf proprietäre Bussysteme basiert, sind sie nun quasi offen und arbeiten überdies mit bekannten Protokollen. Das ermöglicht einen Angriff auf die Kommunikationsnetze dieser Komponenten und in weiterer Folge auf das Elektrizitätsversorgungsnetz. Neue Technologien und Maßnahmen, die den laufenden Betrieb überwachen, Abweichungen vom Normalverhalten entdecken und entsprechend darauf reagieren, sind daher erforderlich. Nur dadurch kann die Sicherheit der Netze weiterhin gewährleistet werden.

Im Projekt Substation Security steht insbesondere das Automatisierungsnetz im Fokus. Dadurch, dass gerade das Automatisierungsnetz von außen erreichbar sein muss (Stichwort Fernwartung), ist es eine besonders kritische Komponente der IKT-Architektur eines Umspannwerks und einer Ortsnetzstation, das im Umspannwerk außerdem noch mit dem Schutznetz verbunden ist. Es wurden im Projekt Substation Security daher Technologien und Methoden erforscht, die eine Anomalie-Erkennung in diesen Automatisierungsnetzen ermöglichen, damit frühzeitig Angriffe und/oder Fehlfunktionen erkannt und die entsprechenden Gegenmaßnahmen rechtzeitig eingeleitet werden können. Die praktische Umsetzbarkeit der ermittelten Methoden wurde durch die Entwicklung einer Proof-of-Concept-Implementierung des Überwachungssystems nachgewiesen.

2.2 Schwerpunkte des Projektes

Die Schwerpunkte des Projekts Substation Security waren wie folgt definiert:

- Konzept eines Intrusion Detection Systems (IDS) für Automatisierungsnetze in Umspannwerken und Ortsnetzstationen
- Formales Modell zur Beschreibung derartiger Systeme
- Definition des Normalverhaltens in Form von formalen Mustern zum Zwecke der laufenden Überprüfung des Systemverhaltens in Echtzeit
- Proof-of-Concept Implementierung des Intrusion Detection Systems

Wobei folgende Teilziele definiert wurden:

- Definition der Sicherheitsanforderungen
- Auswahl eines geeigneten formalen Beschreibungsverfahrens
- Modellierung der Kommunikationsinfrastruktur
- Definition von Verhaltensmustern
- Entwicklung eines Analyseverfahrens zur Mustererkennung
- Proof-of-Concept Implementierung des Analyseverfahrens
- Integration der Proof-of-Concept Implementierung in eine den realen Bedingungen möglichst nahekommende Testumgebung

2.3 Einordnung in das Programm

In Bezug auf die gegenständliche 2. Ausschreibung des Energieforschungsprogramms zielte das vorliegende Projekt Substation Security insbesondere auf das Themenfeld 4 – Intelligente Netze und das darunterliegende Themenfeld 4.1 – Stromnetze ab. Es adressiert vor allem die Sicherheit des Kommunikationsnetzes in wichtigen Komponenten des Smart Grids, dem Umspannwerk und der Ortsnetzstation.

Es konnten im Wesentlichen die zwei folgenden strategischen Ziele des Programms erfüllt werden:

Aufbau und Absicherung der Technologieführerschaft bzw. Stärkung der internationalen Wettbewerbsfähigkeit österreichischer Unternehmen und Forschungsinstitute. Es wird auch der Zugang der Industrie zu relevanter Forschungskompetenz an Forschungseinrichtungen weiter unterstützt und es erfolgt ein gezielter Ausbau von Forschungskompetenz in Forschungseinrichtungen. Durch die Stärkung der Technologiekompetenz und Wettbewerbsfähigkeit wird der Wirtschafts- und Innovationsstandort Österreich gestärkt. (Ziel 3)

1. Beitrag zur Erfüllung der energie-, klima und technologiepolitischen Vorgaben der österreichischen Bundesregierung. (Ziel 1)

2.4 Verwendete Methoden

Eine erste Sichtung der Literatur zum engeren Thema – Sicherheit in Automatisierungsnetzen in Umspannwerken und Ortsnetzstationen – ergab, dass es nur sehr wenig wissenschaftliche Literatur für dieses spezifische Thema gibt. Grund dafür ist im Wesentlichen die Tatsache, dass der Cyber-Sicherheit in derartigen Anlagen bis vor kurzem nur sehr beschränkte Aufmerksamkeit gewidmet wurde. So lange die Systeme proprietär und in sich abgeschlossen waren, bestand nur geringe Gefahr für (erfolgreiche) Cyber-Angriffe. Diese Situation hat sich allerdings in den letzten Jahren drastisch geändert: Durch Ausbau

des Energienetzes in Richtung Smart Grid und die damit einhergehende Öffnung der Systeme nach außen wurde das Problem virulent.
Die spezifische Literatur in diesem Zusammenhang findet sich in den folgenden Arbeiten (siehe Literaturverzeichnis):
[5], [6], [7], [8], [9], [10], [11], [12], [13],

Zur Erreichung der Ziele des Projekts wurden folgende Methoden eingesetzt:

Definition der Sicherheitsanforderungen

Die Definition der Sicherheitsanforderungen erfolgte auf Basis einer Sammlung von Informationen aus folgenden Quellen
Recherche der einschlägigen wissenschaftlichen Literatur
Recherche von Vorfällen in andern Anwendungsbereichen
Diskussionen und Workshops mit den Projektpartnern

Auswahl eines geeigneten formalen Beschreibungsverfahrens

Es wurden drei formale Verfahren als mögliche Kandidaten gewählt und im Rahmen von Workshops auf Ihre Eignung für das Projekt untersucht. Dabei wurde zunächst ein Kriterienkatalog erstellt, mit dessen Hilfe die Bewertung vorgenommen wurde. Bei den drei formalen Verfahren handelte es sich um:

- Formale Grammatiken (im Hinblick auf syntaktische Mustererkennungsalgorithmen)
- Numerische Verfahren (im Hinblick auf statistische Mustererkennungsalgorithmen)
- Simulationen (im Hinblick auf Datengenerierung)

Modellierung der Kommunikationsinfrastruktur

Zur Modellierung der Kommunikationsstruktur wurde zunächst eine Simulation eingesetzt, die die Wiederholung unterschiedlicher Kommunikationsszenarien in einem Umspannwerk bzw. einer Ortsnetzstation ermöglichte. Um auch anormalen Netzwerkverkehr durch die Simulation generieren zu können, wurden auch verschiedene Angriffsszenarien simuliert. Dabei wurden Szenarien für drei Arten von Angriffsvektoren durchgespielt, die wesentliche Bestandteile fast aller Angriffe sind:

1. Passive Man-in-the-Middle Angriffe (Lauschangriffe)
2. Aktive Man-in-the-Middle Angriffe (Replay und Reprogramming Angriffe)
3. Denial-of-Service Angriffe.

Definition von Verhaltensmustern

Die Verhaltensmuster selbst wurden durch Methoden des maschinellen Lernens ermittelt. Zur Verringerung der Dimensionalität der Daten wurde Feature Extraction angewandt und

basierend auf den extrahierten Features wurde ein Modell des Normalverhaltens durch einen Klassifikator trainiert. Es wurden mehrere Klassifikatoren getestet, um denjenigen mit dem besten Ergebnis ermitteln zu können. Die Bewertung der Klassifikatoren erfolgte mit Hilfe

einer Kreuzvalidierung, wobei die besten Ergebnisse mit einem neuronalen Netz erzielt wurden.

Entwicklung eines Analyseverfahrens zur Mustererkennung
Aufbauend auf den Ergebnissen des maschinellen Lernens wurde ein relativ einfaches Analyseverfahren entwickelt, das die – ebenfalls durch maschinelles Lernen ermittelten – Konfidenzintervallen für die gewählten Features heranzieht. Liegt ein gemessener Feature-Wert außerhalb dieser Konfidenzintervalle, wird eine Anomalie diagnostiziert.

Proof-of-Concept Implementierung des Analyseverfahrens
Das Verfahren einschließlich der Trainingsphasen wurde als Proof-of-Concept implementiert, wobei methodisch eine Reihe von verfügbaren Tools, allen voran RapidMiner, herangezogen wurden.

Integration der Proof-of-Concept Implementierung in eine den realen Bedingungen möglichst nahekommende Testumgebung
Der Proof-of-Concept wurde in 2 Testumgebungen installiert und die Ergebnisse der Tests ausgewertet. Die eine Testumgebung war ein vom Projektpartner Siemens zur Verfügung gestelltes Test-Setup mit vier RTUs und entsprechender Software zur Simulation unterschiedlicher Betriebsszenarien. Gegen diese Testanlage wurden auch Cyber-Angriffe direkt durchgeführt, um entsprechende Situationen realistisch nachbilden zu können und die Wirksamkeit des Überwachungssystems überprüfen zu können. Als zweite Testumgebung wurde das System direkt im Umspannwerk Wels Mitte des Projektpartners Wels Strom installiert, um die Funktionsweise des Überwachungssystems auch im Realbetrieb testen zu können. In diesem Fall konnten natürlich keine Cyber-Angriffe durchgeführt werden. Anormale Situationen wurden durch Betriebshandlungen, die für den realen Netzbetrieb nicht gefährlich waren, simuliert.

2.5 Aufbau der Arbeit

Im Kapitel 3 werden die inhaltlichen Aspekte des Projekts behandelt. Die Beschreibung lehnt sich an die definierten Arbeitspakete an und gliedert sich in folgende Bereiche:

Zunächst wird die Analyse der Sicherheitsanforderungen, die zu Beginn des Projekts erstellt wurde, beschrieben. Das detaillierte Ergebnis dieser Analyse ist in einem eigenen Dokument zu finden: „Sicherheitsanalyse und Bedrohungsszenarien“.

Anschließend wird die Auswahl des formalen Beschreibungsverfahrens erläutert. Mit Hilfe des gewählten Verfahrens wird dann die formale Modellierung des Kommunikationsnetzes vorgenommen und ist im folgenden Abschnitt beschrieben. Danach wird beschrieben, wie das Modell im Bereich der Automatisierungsnetze in Umspannwerken und Ortsnetzstationen angewendet wird. Darauf folgt die Definition der Muster des Normalverhaltens des Netzwerkverkehrs mit Hilfe von Methoden des maschinellen Lernens sowie die Spezifikation des Analyseverfahrens zur Erkennung von Anomalien im Netzwerkverkehr. Um die praktische Umsetzbarkeit des entwickelten Verfahrens zu zeigen wurde eine Proof-of-

Concept Implementierung durchgeführt, die im Kapitel 3.6. genau beschrieben ist. Weiters folgt eine Beschreibung der Evaluation dieser Implementierung durch Anwendung sowohl in einer Testanlage als auch in einem realen Umspannwerk.

Kapitel 4 fasst die Ergebnisse zusammen und Kapitel 5 gibt einen Ausblick auf die mögliche Verwertung der Ergebnisse und auf weitere interessante Forschungsthemen in diesem Zusammenhang.

3 Inhaltliche Darstellung

3.1 Analyse der Sicherheitsanforderungen

Am Beginn des Projekts (Arbeitspaket 1) wurden relevante Vorarbeiten bezüglich Anomalie-Erkennung beim Netzwerkverkehr in Umspannwerken geleistet. Diese Arbeiten bestanden aus detaillierten Analysen unterschiedlicher Aspekte mit einem Fokus auf Sicherheitsanforderungen. Im ersten Schritt wurde die gesamte Kommunikationskette von der Prozessebene bis zum HMI (Human-Machine Interface) untersucht, wobei der Schwerpunkt bei den verwendeten Kommunikationsprotokollen (61850, 60870-5-104) lag. Die Analyseergebnisse bezogen sich im Wesentlichen auf die Arten von in den Protokollen erlaubten Netzwerk-Paketen und Paket-Sequenzen. Die wichtigste Informationsquelle dafür waren Netzwerk Protokolldaten, die vom Projektpartner Wels Strom GmbH erfasst und bereitgestellt wurden. In weiterer Folgen war das Augenmerk auf die Infrastruktur von Umspannwerken und Ortsnetzstationen und das Verhalten des Netzwerkverkehrs in diesen Komponenten gerichtet. Information über Netzwerktopologie, Netzwerk-Auslastung und die asynchrone Kommunikation wurden als relevante Vorbereitungen für die späteren Projektschritte erhoben. Für die Auswahl der Methode der Anomalie Erkennung beziehungsweise für die Struktur des IDS sind das grundlegende Parameter. Dazu wurden auch Details über elektrotechnische Eigenschaften eines Umspannwerks bzw. einer Ortsnetzstation und deren Bezug zu den Protokolldaten sowie die Funktionalität der eingesetzten Intelligent Electronic Devices auf der Prozessebene analysiert.

Der Hauptteil der Sicherheitsanalyse bestand in der Erfassung möglicher Bedrohungsszenarien. Es wurden auf Basis der Literatur und auch aufbauend auf Ergebnisse anderer Forschungsprojekte unterschiedliche Angriffsszenarien definiert und detailliert beschrieben. Als Vorgangsweise wurde die von Bruce Schneier vorgeschlagene Methodik der Attack Trees verwendet.

Die Ergebnisse liegen in einem eigenen Dokument vor: „Sicherheitsanalyse und Bedrohungsszenarien“.

3.2 Auswahl einer formalen Methode

Es wurden verschiedene formale Beschreibungsverfahren auf ihre Eignung in Bezug auf die Projektziele untersucht. Dazu wurde ein Kriterienkatalog erstellt, anhand dessen die Bewertung der Beschreibungsverfahren vorgenommen wurde. Folgende Kriterien wurden für eine Bewertung der Verfahren herangezogen:

1. Semantische Eignung – Möglichkeiten der möglichst realistischen Darstellung des Netzwerkverkehrs.
2. Eignung zur Verifizierung der letztendlich entwickelten Anomalie-Erkennungsalgorithmen.
3. Komplexität – das Modell soll eine gewisse Komplexität nicht übersteigen, da bei der Umsetzung auf die eingeschränkten Kapazitäten, wie sie in einem industriellen Kontrollsystem gegeben sind, Rücksicht genommen werden muss.
4. Skalierbarkeit – das Modell muss für Anlagen unterschiedlicher Größenordnung geeignet sein (allein wegen der Anwendung sowohl auf Umspannwerke als auch auf Ortsnetzstationen).
5. Verfügbarkeit von Tools – dies erleichtert die tatsächliche Anwendung beträchtlich. Überdies wäre die Entwicklung eigener Tools für die Modellierung außerhalb des Projektumfangs.

Es wurden drei formale Beschreibungsverfahren gewählt und hinsichtlich ihrer Eignung für die Ziele des Projekts bewertet. Konkret wurden folgende Verfahren untersucht:

- Formale Grammatiken (im Hinblick auf syntaktische Mustererkennungsalgorithmen) siehe [14], [15]
- Numerische Verfahren (im Hinblick auf statistische Mustererkennungsalgorithmen) siehe [16], [17], [18]
- Simulationen (im Hinblick auf Datengenerierung) siehe [19],[20]

Die Bewertung wurde in zwei zeitlich versetzten Workshops vorgenommen; im ersten Workshop wurde jedes Verfahren von einer/m ProjektmitarbeiterIn vorgestellt und dann im Sinne der Kriterien diskutiert; anschließend wurden Punkte festgestellt, die bei den einzelnen Verfahren noch genauer zu untersuchen wären, um eine endgültige Bewertung vornehmen zu können. Im zweiten Workshop wurden die Ergebnisse der genannten Untersuchungen diskutiert und schließlich eine endgültige Bewertung vorgenommen.

Um zu einer fundierten Beurteilung zu kommen, war es auch erforderlich, die Daten des Netzwerkverkehrs, die vom Projektpartner Wels Strom geliefert wurden, genauer zu analysieren (siehe weiter unten).

Es wurde folgende Skala für die Bewertung verwendet:

1 → sehr gut geeignet
2 → geeignet
3 → weniger geeignet
4 → nicht geeignet

In der folgenden Tabelle sind die Ergebnisse der Bewertung aufgeführt.

	Semantische Eignung	Eignung zur Verifizierung	Komplexität	Skalierbarkeit	Verfügbarkeit von Tools
Formale Grammatik	2	2	2	3	1
Numerische Verfahren	2	2	1	1	1
Simulation	1	2	2	1	1

Wie aus der Bewertung ersichtlich wurden formale Grammatiken als Beschreibungsverfahren ausgeschlossen. Auf Grund der Ergebnisse wurde beschlossen, folgendermaßen vorzugehen:

Das Kommunikationsnetz bzw. der Netzwerkverkehr wurden als Simulation beschrieben, um eine formale Basis für die weitere Arbeit zu haben. Die Definition der Muster normalen Verhaltens wurde dann mit Hilfe numerischer, d.h. im konkreten Fall statistischer Methoden vorgenommen.

Zur Beurteilung der Verfahren wurden einzelne Experimente mit den Daten unternommen, um zu einer fundierten Bewertung zu kommen. Dabei kamen verschiedene Softwarewerkzeuge einerseits für die Kommunikationssimulation und andererseits für die Datenanalyse zum Einsatz. Zum Zwecke der Netzwerkverkehr-Simulation wurde diverse open source Simulationssoftware sowie Open IEC 61850 evaluiert. Diese Softwarewerkzeuge ermöglichen eine einfache Leistungsbeurteilung von Intrusion Detection Systemen, aber auch die Simulation von Angriffen. Weitere Aktivitäten umfassten die Erstellung eines Parsers für die aufgenommenen 60870-5-104 Protokolldaten aus den Umspannwerken des Projektpartners Wels Strom. Implementiert wurde der Parser mit Matlab und dem Protokollanalyseprogramm tshark. Mit Hilfe des Parsers ist es möglich, verschiedene Informationen aus der Prozessebene (z.B. die Zahl der Messpunkte), über das Netzwerkverhalten (Datendurchlauf und Zeitinformationen), aber auch Daten aus mehreren Protokollebenen zu extrahieren. Schlussendlich wurden auch erste statistische und visuelle Analysen von 104-Protokolldaten durchgeführt.

3.3 Erstellung eines formalen Modells des Kommunikationsnetzes

Um eine genaue Definition des Normalverhaltens sowie eine effiziente Leistungsbeurteilung des Anomalie-Erkennungsverfahrens zu gewährleisten, wurde das Kommunikationsnetz mit Hilfe einer Simulation modelliert. Es wurde eine Simulation entwickelt, die die Wiederholung unterschiedlicher Kommunikationsszenarien in einem Umspannwerk bzw. einer Ortsnetzstation ermöglicht. Das Werkzeug besteht aus zwei Raspberry Pi 3 Rechnern und geeigneter Software, die für die TCP/IP Kommunikation zuständig ist. Damit war es auch

möglich, auf Basis der aufgenommenen echten 60870-5-104 Protokolldaten die Kommunikation zwischen den Intelligenten Elektronischen Einheiten (IED – Intelligent Electronic Devices) und den Routern auf der Prozessebene genau zu reproduzieren. Die Simulation erlaubt auch die Sammlung von Zeit-Informationen über einzelne 104-Protokoll Pakete, was für die Beschreibung des normalen Netzwerkverhaltens erforderlich ist.

Der zweite wichtige Punkt in diesem Zusammenhang ist die Generierung von anormalem Netzwerkverkehr durch Simulation verschiedener Angriffsszenarien. Auf Basis der am Beginn des Projekts vorgenommenen Sicherheitsanalyse wurde beschlossen, dass sich die Szenarien auf drei Arten von Angriffsvektoren konzentrieren, die wesentliche Bestandteile fast aller Angriffe sind:

1. Passive Man-in-the-Middle Angriffe (Lauschangriffe)
2. Aktive Man-in-the-Middle Angriffe (Replay und Reprogramming Angriffe)
3. Denial-of-Service Angriffe.

Diese Angriffe wurden unter Verwendung der Software ettercap und hping3 realisiert. Letztendlich wurden anormale Protokolldaten, wie sie bei Angriffen entstehen, mit dem Simulationswerkzeug generiert.

3.4 Definition von Mustern des Normalverhaltens

Im nächsten Schritt wurden potentielle Muster normalen Netzwerkverhaltens mit Hilfe von Algorithmen des maschinellen Lernens definiert. Da es sich jedoch bei den aufgezeichneten Rohdateien um Netzwerkpakete handelt und diese Form keine einfache Verarbeitung bzw. Training eines Algorithmus erlaubt, wurden Merkmale aus den übertragenen Informationen extrahiert. Die Merkmalsextraktion (Feature Extraction) ist in Abbildung 1 dargestellt. Aus den Rohdaten wurden bestimmte Features, nicht redundante, informative und vergleichbare Merkmale, extrahiert. Zusätzlich wird dabei die Dimensionalität der Daten verringert – unter Beibehaltung aller wesentlicher Informationen.

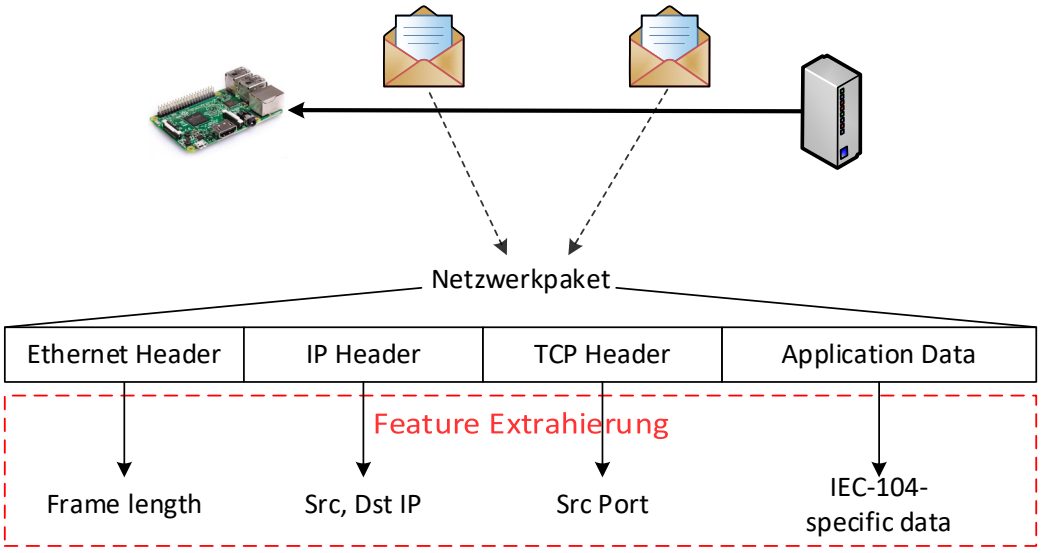

Abbildung 1: Extrahierung von Features aus einem Netzwerkpaket

Als Features wurden statistische Werte (Durchschnitt, Standardabweichung etc.) relevanter Netzwerkinformationen (round-trip-time, Paketlänge, Messwerte u.ä.) berechnet. Die Feature Extraction ist für beide Übertragungsprotokolle strukturell gleich. Die Auswahl der Features stellt eine der wichtigsten, sowie schwierigsten Aufgaben bei der Anomalie-Erkennung dar. Werden zu wenige Features gewählt, können Informationen verloren gehen, bei zu vielen besteht die Gefahr des Overfitting. Im Projektverlauf wurden verschiedene Features auf ihre Eignung zur Beschreibung des Normalverhaltens getestet. Die besten Ergebnisse wurden mit dem in Tabelle 1 angegebenen Set erzielt.

Tabelle 1: Set von extrahierten Features

label	stdev(ioa89)	stdev(rtt)	stdev(length)	stdev(wsize)	avg(ioa89)	avg(rtt)	avg(length)	avg(wsize)	packets/s
rtu20_valid	4,26066	215,1461	6,363961	0	399,4765	317,6	78,9	457	3,148615
rtu20_valid	5,008532	317,681	6,893475	0	400,8722	332,42	78,48	457	3,008243
rtu20_valid	5,110769	296,1252	8,060713	0	399,6217	318,74	80,38	457	3,137353
rtu20_valid	5,620426	274,5979	7,807087	0	399,4502	340,64	79,22	457	2,935651
rtu20_valid	5,440132	325,0333	7,300433	0	398,8247	378,14	79,64	457	2,644523
rtu20_valid	4,539931	220,9048	6,893475	0	399,9002	318,86	78,48	457	3,136173

Basierend auf diesen extrahierten Features wurde ein Modell des Normalverhaltens durch einen Klassifikator (z.B. *k*-Nearest-Neighbour-Algorithmus, Naive Bayes, Support Vector Machines, neuronales Netzwerk) trainiert.

3.5 Entwicklung eines Analyseverfahrens für die Muster

Der erste Schritt für die Erstellung eines Modells des Normalverhaltens ist das Training von validen Daten (normalem, unverändertem Verhalten) um das Systemverhalten des Netzwerkverkehrs im Normalzustand zu beschreiben. In den Trainingsdaten müssen Klassen, sogenannte Labels, zugewiesen sein, damit ein Klassifikator die Daten nach diesen Labels unterscheiden kann. Als klassifiziertes Datenset wurde das Normalverhalten von zwei RTUs, aufgezeichnet und die jeweiligen Features extrahiert und entsprechende Labels zugewiesen.

Für die Anomalie-Erkennung wurden verschiedene Ansätze getestet. Hierbei war der asynchrone Aufbau der beiden im Einsatz befindlichen Übertragungsprotokolle (IEC 60870-5-104 und IEC 61850) eine Herausforderung:

- Neue Messwerte werden nur gesendet, wenn sich Werte ändern
- Pakete können nicht einzeln klassifiziert werden

Bei der Analyse des Netzwerkverkehrs wurde erkannt, dass über einen längeren Zeitraum gesehen, bestimmte Übertragungen periodisch durchgeführt werden und somit Klassen (z.B. RTUs) eindeutig identifiziert werden können. Daraus wurde eine mögliche Lösung der zuvor genannten Probleme abgeleitet: Eine Überprüfung/Klassifizierung per definierter Anzahl von Paketen. Beispiel: Für jeweils *k* (z.B. 100) Pakete werden Durchschnittswerte der Features berechnet und klassifiziert. Dadurch werden periodische Muster in den Aufzeichnungsdaten sichtbar und können durch einen Klassifikator trainiert werden. Eine Visualisierung des klassifizierten (labeled) Datensets ist in Abbildung 2 dargestellt. Ein Klassifikator versucht die beiden Klassen (hier zwei RTUs) möglichst genau zu trennen. Die Visualisierung zeigt ein Streudiagramm der Features der beiden RTUs und zeigt ihre Abhängigkeitsstruktur auf.

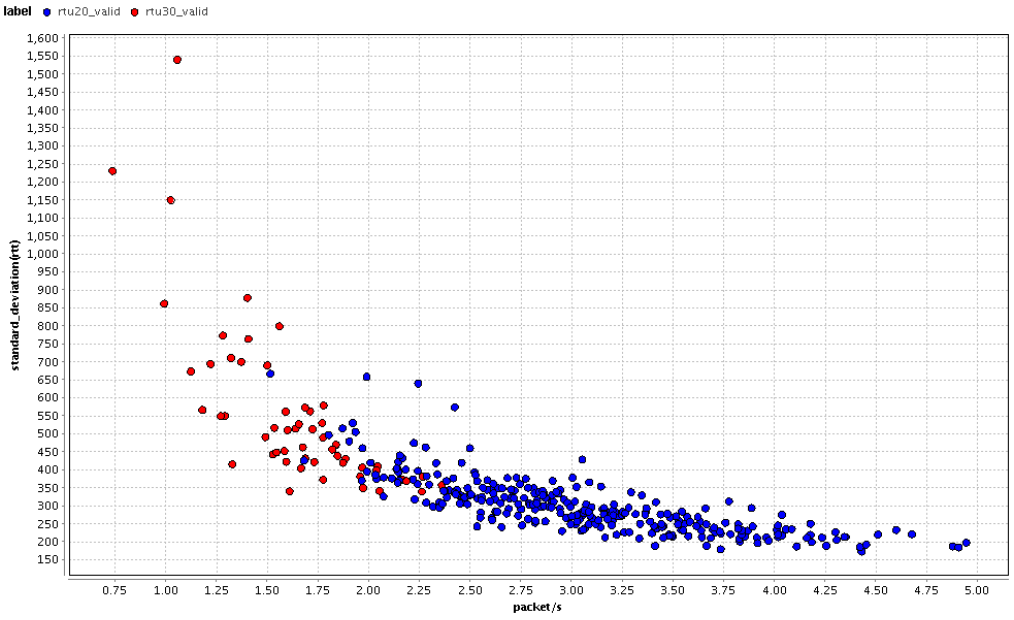

Abbildung 2: Visualisierung der klassifizierten Datensets

Basierend auf dem Datenset wurde mit Hilfe des Software-Werkzeugs RapidMiner, einer Software Plattform für maschinelles Lernen und Data-Mining, ein Modell mit verschiedenen Klassifikatoren trainiert. Um die Genauigkeit der Klassifikatoren zu überprüfen, wurde jeweils eine Kreuzvalidierung vorgenommen. Kreuzvalidierung ist eine Technik zur Bewertung der Leistung eines Modells beim Machine Learning. Mit neuen Datensätzen, welche nicht in der Trainingsphase genutzt wurden, wird die Güte der Vorhersage geprüft. Dies geschieht durch die Partitionierung eines Datensatzes in Teilmengen für das Training und das Testen des Algorithmus.

Jeder Durchlauf der Kreuzvalidierung umfasst eine zufällige Partitionierung des Originaldatensatzes in einen Trainingssatz und einen Testsatz. Der Trainingssatz wird verwendet, um einen Algorithmus für überwachtes Machine Learning zu trainieren und der Testsatz wird zur Bewertung der Leistung des Trainingsergebnisses verwendet. Dieser Vorgang wird mehrmals wiederholt und der mittlere Kreuzvalidierungsfehler als Leistungsindikator verwendet.

Es wurde mit verschiedenen Klassifikatoren experimentiert. Dabei wurden die Genauigkeit, sowie die Performance (Rechenaufwand und -dauer) berücksichtigt. Die höchste Genauigkeit, siehe Tabelle 2, wurde mit einem neuronalen Netz erreicht. Das durch diesen Klassifikator erzeugte Modell stellt die Basis für die Klassifikation unbekannter Daten dar.

Tabelle 2: Ergebnis der Kreuzvalidierung mit neuronalem Netz

Accuracy: 96,53%	**true rtu20_valid**	**true rtu30_valid**	**class precision**
pred. rtu20_valid	284	4	98,61%
pred. rtu30_valid	8	52	86,67%
class recall	**97,26%**	**92,86%**	

3.6 Proof-of-Concept Implementierung

Das gesamte Anomalie-Erkennungskonzept wurde als Proof-of-Concept für einen Raspberry Pi 3 entwickelt und in einer durch den Projektpartner Siemens zur Verfügung gestellten Testanlage mit mehreren RTUs implementiert. Durch die Testanlage konnte eine realitätsnahe Implementierung in einer realen Netzwerkumgebung vorgenommen werden.

Der Aufbau des IDS Prototyps ist in Abbildung 3 dargestellt. Der Prototyp führt folgende Abläufe automatisiert durch:

- *Data Acquisition*: Netzwerkpakete von den definierten Geräten werden aufgezeichnet.
- *Feature Extraction*: Merkmale (Features) werden aus den Netzwerkpaketen extrahiert bzw. berechnet.
- *Modelling Neural Network*: In der Trainingsphase wird mit den extrahierten Features ein

Neuronales Netz trainiert.

- *Anomalie-Erkennung*: Im Betrieb werden neue, unbekannte Daten gegen das Modell getestet. Ein automatisiertes Update mit den neuen Daten („Weiter-lernen“) kann konfiguriert werden.
- *Output*: Die Klassifizierungsergebnisse der aktuellen Netzwerkpakete werden ausgegeben.

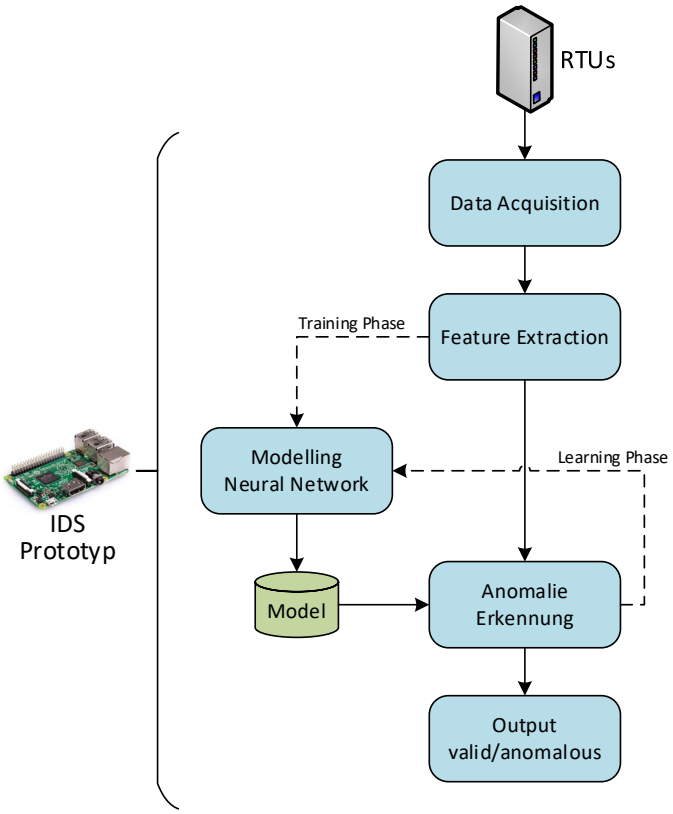

Abbildung 3: Konzept des Substation Anomalie-Erkennungssystems

3.6.1 Paketaufzeichnung

In der Testanlage sind alle RTUs über eine Fortinet Firewall angeschlossen. Um die Pakete am Raspberry Pi zu übertragen, wurde ein Mirror Port konfiguriert, welcher alle Netzwerkpakete der RTUs spiegelt. Dadurch können Informationen aus den Netzwerkpaketen extrahiert werden. Die Aufzeichnung der Pakete wurde mittels der *pcap4j* library, welche Paketmanipulation erlaubt, durchgeführt. Zusätzlich wurde ein Filter definiert, welcher Pakete von Nicht-Industrieprotokollen (die für den geplanten Zweck irrelevant sind) verwirft (z.B. SNMP). Folgende Filter wurden eingesetzt:

- TCP Port 2404 (IEC-104)
- TCP/UPD Port 102 (IEC-61850)
- IP Adressen der konfigurierten RTUs

Das daraus resultierende Datenset enthält die Messwerte des jeweiligen konfigurierten Protokolls (IEC 60870-5-104 oder IEC 61850), sowie relevante Netzwerkpaketinformationen.

3.6.2 Feature Extrahierung

Die Feature Extrahierung (siehe Kapitel 3.4) wurde für die Test-Anlage konfiguriert. Dabei werden die Features von je 50 Netzwerkpaketen per RTU automatisiert extrahiert. Dadurch konnten periodische Muster der RTUs in der Testanlage erkannt werden. Das Normalverhalten der RTUs wurde aufgezeichnet und das Datenset diente als Basis für die Anomalie-Erkennung.

3.6.3 Anomalie-Erkennung

In der Trainingsphase des Prototyps wurde das aufgezeichnete Normalverhalten mit einem neuronalen Netz trainiert. Um das Modell zu prüfen, wurden verschiedene Netzwerkangriffe supervised durchgeführt, um die Erkennungsrate zu testen.

Folgende spezielle Angriffe gegen das IEC 60870-5-104 Protokoll eines Umspannwerks bzw. einer Ortsnetzstation wurden durchgeführt und die Netzwerkpakete während dieser Angriffe aufgezeichnet:

- Man-in-the-Middle Filter Angriff, bei dem übertragene Messwerte konstant überschrieben werden.
- Man-in-the-Middle Increment Angriff, bei dem die übertragenen Messwerte geringfügig (Wert + 0,1-1,0) verändert werden. Dadurch können unbekannte und unerwartete Systemzustände erreicht werden.
- Man-in-the-Middle Drop Angriff, bei dem Pakete, die ein bestimmtes Signal senden, verworfen werden.

Mit dem dabei gewonnenen Datenset wurde das neuronale Netz erneut trainiert, um die Angriffe auch klassifizieren zu können. Für die Berechnung der Güte und Performance des Klassifikators wurde wiederrum eine Kreuzvalidierung durchgeführt. Die Ergebnisse dieser Kreuzvalidierung sind in Tabelle 3 aufgelistet. Das Resultat ist durchaus beachtlich. Einerseits konnten die RTUs untereinander gut unterschieden werden, andererseits konnten sowohl der Filter als auch der Drop Angriff mit hoher Genauigkeit klassifiziert werden. Nur der Increment Angriff wurde bisweilen falsch klassifiziert, da die Änderungen teilweise sehr gering und somit schwer erkennbar waren.

Tabelle 3: Ergebnisse der Anomalie-Erkennung

	true rtu20_valid	**true rtu30_valid**	**true rtu30_mitm-filter**	**true rtu30_mitm-incr**	**true rtu30_mitm-drop**
pred. rtu20_valid	381	16	0	4	2
pred. rtu30_valid	17	365	0	44	1

pred. rtu30_mitm-filter	0	0	136	0	0
pred. rtu30_mitm-incr	2	15	0	31	0
pred. rtu30_mitm-drop	0	4	0	1	65
class recall	**95,25%**	**91,25%**	**100,00%**	**38,75%**	**95,59%**

Das Modell mit Hilfe des neuronalen Netzes trainierte Modell wird in weiterer Folge zur Klassifikation von neuen, unbekannten Daten herangezogen. Dabei berechnet der Klassifikator „confidences“ (Konfidenzen) für eine Klassenzugehörigkeit. Um die Anwendbarkeit des Modells zu prüfen, wurden mehrere unbekannte Datensätze aufgezeichnet, darunter Normalverhalten und diverse Angriffe, und gegen das Modell getestet. Die Ergebnisse, ersichtlich in Tabelle 4, sind durchaus zufriedenstellend. Bis auf eine Fehlklassifikation wurden alle Instanzen korrekt klassifiziert. Zusätzlich sind die Konfidenzen ausreichend groß und lassen auf eindeutige Ergebnisse schließen.

Tabelle 4: Klassifizierung von unbekannten Daten

filename	conf rtu20_valid	conf rtu30_valid	conf rtu30_mitm-filter	conf rtu30_mitm-incr	conf rtu30_mitm-drop	prediction label
rtu-172-16-1-129-valid	**0,981**	0,000	0,000	0,018	0,000	rtu20_valid
rtu-172-16-1-129-valid	**0,989**	0,000	0,000	0,011	0,000	rtu20_valid
rtu-172-16-1-129-valid	**0,998**	0,000	0,000	0,002	0,000	rtu20_valid
rtu-172-16-2-1-mitm-drop	0,000	0,020	0,000	0,000	**0,979**	rtu30_mitm-drop
rtu-172-16-2-1-mitm-drop	0,000	0,013	0,005	0,000	**0,981**	rtu30_mitm-drop
rtu-172-16-2-1-mitm-drop	0,001	0,009	0,000	0,000	**0,987**	rtu30_mitm-drop
rtu-172-16-2-1-mitm-incr	0,000	0,001	0,001	**0,997**	0,000	rtu30_mitm-incr
rtu-172-16-2-1-mitm-incr	0,000	0,001	0,007	**0,992**	0,000	rtu30_mitm-incr
rtu-172-16-2-1-mitm-incr	0,000	0,383	0,000	**0,617**	0,000	rtu30_mitm-incr
rtu-172-16-2-1-mitm-incr	0,000	0,306	0,000	**0,693**	0,000	rtu30_mitm-incr
rtu-172-16-2-1-mitm-incr	**0,549**	0,269	0,000	0,181	0,000	rtu20_valid

3.7 Integration in Testumgebung

In weiterer Folge wurde die Proof-of-Concept Implementierung des Intrusion Detection Systems in eine realitätsnahe Testumgebung integriert. Vom Projektpartner Wels Strom wurde eine direkte Integration in das Umspannwerk Wels Mitte ermöglicht, um reale Bedingungen zu gewährleisten. Es konnte der Prototyp direkt im Umspannwerk Wels Mitte in eingebaut und in Betrieb genommen werden. Dazu wurde der Raspberry Pi mittels Port Mirror am Verteilerswitch angeschlossen. Somit konnte der Netzwerkverkehr aller relevanten Endpunkte aufgezeichnet werden. Bei der Erstkonfiguration wurde eine Liste von Geräten ausgewählt, für welche Trainingsdaten aufgezeichnet werden sollen. Für diese Geräte wurden Netzwerkpakete über einen Zeitraum von 30 Minuten aufgezeichnet und

anschließend durch das neuronale Netzwerk trainiert. Die Genauigkeit des Modells wurde mittels einer Kreuzvalidierung überprüft. Eine Gesamtgenauigkeit von 92,41 %, siehe Tabelle 5, wurde erreicht. Durch eine längere Aufzeichnungsphase, sowie spezifischere Konfiguration kann dieser Wert weiter erhöht werden.

Tabelle 5: Klassifizierungsergebnisse der Prototyp Integration

	true 100_valid	**true 102_valid**	**true 103_valid**	**true 106_valid**	**true 107_valid**	**true 109_valid**	**class precision**
pred, 100_valid	72	2	0	0	0	0	97,30%
pred, 102_valid	4	14	0	0	0	0	77,78%
pred, 103_valid	0	0	216	0	0	0	100,00%
pred, 106_valid	0	0	0	46	12	0	79,31%
pred, 107_valid	0	0	0	26	56	0	68,29%
pred, 109_valid	0	0	0	0	0	132	100,00%
class recall	**94,74%**	**87,50%**	**100,00%**	**63,89%**	**82,35%**	**100,00%**	

Zusätzlich zur Klassifikation von validen Daten wurden im Umspannwerk bestimmte Anomalien (soweit dies ohne Beeinträchtigung des Realbetriebs möglich war) durchgeführt und aufgezeichnet. Es handelte sich dabei im Wesentliche um manuelle Umschaltungen von Transformatoren. Ausgehend vom Modell der validen Daten wurde ein Outlier-Erkennungsverfahren, basierend auf dem *k*-nearest-neighbours Algorithmus, entwickelt und ein Outlier-Wert für die Daten während der Anomalie berechnet. Die Ergebnisse sind in Tabelle 6 aufgelistet. Teilweise sind die Outlier-Werte höher als der Durchschnittswert bei Normalverhalten. Dies ist auf diverse Änderungen in den Netzwerkpaketen (höhere Paketdichte, größere Payload etc.) zurückzuführen. Zusätzlich könnte die Genauigkeit des Verfahrens durch eine spezifischere Systemkonfiguration sowie eine Gewichtung bestimmter Attribute optimiert werden.

Tabelle 6: Outlier Berechnung einer Anomalie

label	**outlier**
109_valid	5,13
109_valid	5,46
109_valid	3,33
109_valid	2,90
109_valid	5,19
109_valid	4,89
109_valid	5,85
109_valid	3,98
rtu-10-90-10-109	3,68

rtu-10-90-10-109	10,03
rtu-10-90-10-109	7,79
rtu-10-90-10-109	6,71
rtu-10-90-10-109	4,48
rtu-10-90-10-109	10,09
rtu-10-90-10-109	20,03

4 Ergebnisse und Schlussfolgerungen

Die Ziele des Projekts Substation Security waren wie folgt definiert:

- Konzept eines Intrusion Detection Systems (IDS) für Automatisierungsnetze in Umspannwerken und Ortsnetzstationen
- Formales Modell zur Beschreibung derartiger Systeme
- Definition des Normalverhaltens in Form von formalen Mustern zum Zwecke der laufenden Überprüfung des Systemverhaltens in Echtzeit
- Proof-of-Concept Implementierung des Intrusion Detection Systems

4.1 Konzept des Intrusion Detection Systems

Das erste Ziel des Projekts – Konzept eines Intrusion Detection Systems (IDS) für Automatisierungsnetze in Umspannwerken und Ortsnetzstationen – wurde vollinhaltlich erreicht. Es wurde ein IDS speziell für Umspannwerke und Ortsnetzstationen konzipiert, das anormales Verhalten im Netzwerkverkehrs des Automatisierungsnetzes in Echtzeit erkennen kann und entsprechend eine Alarmmeldung auslöst.

Das wesentliche Charakteristikum des Ansatzes besteht darin, dass es sich um ein Anomalie-Erkennungssystem handelt („white-list“-Ansatz): Es wird das typische Normalverhalten des Systems definiert und alle Abweichungen werden als potentiell gefährlich bzw. als ein potentieller Angriff eingestuft. Das steht im Gegensatz zu den üblicherweise in Intrusion Detection Systemen eingesetzten Verfahren, die eine Misuse-Erkennungsstrategie verfolgen („black-list“-Ansatz): Dabei werden anormale Verhaltens-muster (Angriffe) definiert und es wird versucht, diese Muster im Netzwerkverkehr zu erkennen. Im konkreten Anwendungsfall (Umspannwerk und Ortsnetzstation) ist die Variabilität im Normalbetrieb wesentlich geringer als in einem allgemeinen Computernetz-werk (in dem ja – zulässigerweise – jede erdenkliche Software laufen kann). Daher ist im gegebenen Fall der Normalbetrieb einfacher zu modellieren als die vergleichsweise große Anzahl an möglichen Angriffen. Überdies kann eine Misuse-Strategie nur bereits bekannte und durch entsprechende Regeln definierte Angriffe erkennen. Ein Anomalie-Erkennungsverfahren hat diese Einschränkung nicht. Sein Nachteil besteht in der Gefahr, eine zu große Menge an False-Positives (Fehlalarmen) zu erzeugen. Auf diesen Aspekt

wurde im Projekt besonderes Augenmerk gelegt. Demgegenüber ist der Nachteil einer Misuse-Strategie, nämlich dass nur bereits bekannte Angriffe erkannt werden können, ein prinzipieller, der auch mit verbesserten Verfahren nicht wettgemacht werden kann. Insbesondere im Hinblick auf die Zunahme von zielgerichteten Angriffen (targeted attacks oder advanced persistent threats) im Bereich kritischer Infrastrukturen kommt diesem Punkt eine große Bedeutung zu.

Im Projekt wurde das Modell des Normalverhaltens mit Hilfe von Methoden des maschinellen Lernens konstruiert. Das System muss zuerst an die spezielle Konfiguration des Umspannwerks bzw. der Ortsnetzstation angepasst werden. Anschließend wird das System während des Normalbetriebs „trainiert“, d.h. dass der Netzwerkverkehr einen bestimmten Zeitraum lang mitgeschnitten und analysiert wird. Dabei werden vordefinierte Features aus den Daten extrahiert. Aus diesen extrahierten Daten werden dann mit Hilfe eines Klassifikators entsprechende Outlier-Werte (die spezifisch für die jeweilige Anlage sind) gebildet. Die Dauer diese Trainingsphase hängt von der Größe der Anlage bzw. ihrer Heterogenität ab. In den Tests haben bereits kurze Trainingsphasen im Bereiche von 30 bis 60 Minuten zu hinreichend genauen Outlier-Werten geführt. Etwas längere Trainingsphasen sind allerdings empfehlenswert und werden im Allgemeinen zu besseren Werten führen. Mit Hilfe der auf diese Weise gewonnenen Outlier-Werte kann dann der operationale Netzwerkverkehr überwacht werden. Dazu müssen die Netzwerkdaten analysiert und die definierten Features in Echtzeit berechnet werden. Der anschließende Vergleich mit den trainierten Outlier-Werten entscheidet dann, ob ein normales Verhalten der Anlage vorliegt oder ob es sich um eine Anomalie handelt. Im letzteren Fall wird ein entsprechender Alarm ausgelöst.

4.2 Formales Modell

Als Voraussetzung zur Erreichung des in 4.1 genannten Ziels, der Konzeption eines Intrusion Detection Systems, musste ein formales Modell der Kommunikationsinfrastruktur des Automatisierungsnetzwerks in einem Umspannwerk bzw. in einer Ortsnetzstation erforscht werden. Das Modell musste so allgemein gehalten werden, dass es auf alle diesbezüglichen Situationen mindestens im österreichischen – so weit möglich auch im europäischen – Kontext anwendbar ist. Es muss grundsätzlich formal definiert werden, damit ein automatisches Überwachungssystem eindeutig daraus abgeleitet werden kann. Das Modell muss dafür geeignet sein, Muster des Normalverhaltens hinreichend flexibel definieren zu können.

Zunächst wurden verschiedene Möglichkeiten der Anomalie-Erkennung auf ihre Tauglichkeit für die gegebene Problemstellung hin untersucht. Prinzipiell kommen zwei verschiedene Verfahren zur Definition des Normalverhaltens eines Systems in Frage:

- Statistische Verfahren
- Regelbasierte Verfahren

- Simulationen

Am weitesten verbreitet sind die statistischen Verfahren. Dabei werden Ereignisse definiert und quantifiziert. Für die Parameter werden Schwellwerte definiert und die Echtdaten werden im Hinblick auf diese Schwellwerte überwacht. In komplexeren Verfahren können auch Abhängigkeiten der Parameter untereinander definiert werden. Auch die Zeit kann als bestimmender Wert miteinbezogen werden. Die entsprechenden Muster zur Beschreibung des Normalverhaltens des Netzwerks (also die Schwellwerte bzw. ihre Abhängigkeiten und Korrelationen untereinander) werden üblicherweise mit Hilfe von Methoden des Machine Learnings gewonnen (supervised oder unsupervised learning). Der Erkennungsalgorithmus leitet sich aus den definierten Parameterwerten ab; je komplexer die Abhängigkeiten der Parameter untereinander sind, desto aufwändiger wird der Erkennungsalgorithmus.

Bei regelbasierten Verfahren definiert man die Muster mit Hilfe von Regeln. Das hat den Vorteil, dass der Erkennungsalgorithmus durch einen Parser implementiert werden kann, der im Allgemeinen aus den Regeln automatisch generiert werden kann. Die Muster selber können entweder manuell von ExpertInnen erstellt werden oder es können Algorithmen entwickelt werden, die die Regeln aus empirischen Daten (halb-)automatisch ableiten. Zu den regelbasierten Systemen zählt man auch die neuronalen Netze. Die Regeln selber können das zeitliche Verhalten des Systems abbilden (time based inductive systems) oder logische Zusammenhänge darstellen.

Simulationen stellen einen Sonderfall dar, da sie an sich nicht direkt zur Anomalie-Erkennung eingesetzt werden können. Sie können jedoch ein wichtiges Hilfsmittel für die Erforschung des Netzwerkverhaltens sein.

Die verschiedenen Möglichkeiten wurden im Projekt Substation Security auf ihre Tauglichkeit für den gegebenen Anwendungsbereich analysiert und es wurde folgende Vorgangweise gewählt:

Zunächst wurde eine Simulation des Netzwerkverkehrs durchgeführt, um eine Basis für die Beurteilung der Situation und der Auswirkungen von Angriffen zu schaffen.

Für die eigentliche Anomalie-Erkennung wurden dann statistische Methoden eingesetzt. Dabei wurden Features bestimmt, für die dann mit Methoden des maschinellen Lernens ein Anomalie-Erkennungsverfahren entwickelt wurde.

4.3 Muster des Normalverhaltens

Die mit Hilfe eines neuronalen Netzwerks gewonnenen Outlier-Werte und Klassifikatoren definieren das Normalverhalten des Netzwerkverkehrs im Automatisierungsnetz eines Umspannwerks bzw. einer Ortsnetzstation hinreichen genau, um eine zuverlässige Anomalie-Erkennung zu ermöglichen.

Dazu wurden Daten aus Netzwerken im Echtbetrieb (die von den beiden Projektpartnern Wels Strom GmbH und Siemens AG Österreich bereitgestellt wurden) herangezogen. Zusätzlich wurden Daten mit Hilfe der Simulation gewonnen.

In einem praktischen Anwendungsfall muss der Lern- bzw. Trainingsprozess jeweils neu für das Zielsystem (das Ziel-Umspannwerk bzw. die Ortsnetzstation) durchgeführt werden. Die Muster des Normalverhaltens des Systems sind

4.4 Proof-of-Concept Implementierung und Test in Real-Umgebung

Als letztes Teilziel wurde das gesamte definierte Anomalie-Erkennungskonzept als Proof-of-Concept Implementierung des Intrusion Detection Systems umgesetzt. Das umfasst auch eine Integration in eine möglichst nahe an die Realität heranreichende Umgebung mit beschränkten Ressourcen. Ziel ist ein System, das das Verfahren soweit praktisch abbilden kann, damit es als Grundlage für spätere Produktentwicklungen herangezogen werden kann.

Ein Prototyp wurde für einen Raspberry Pi 3 entwickelt und auf einer vom Projektpartner Siemens zur Verfügung gestellten Testanlage implementiert. Durch die Testanlage konnte eine realitätsnahe Implementierung, sowie eine reale Netzwerkumgebung dargestellt werden. Basierend auf den Ergebnissen der vorangegangenen Arbeitspakete implementiert der Prototyp folgende Abläufe automatisiert: Datenaufzeichnung, Merkmalsgenerierung, Anomalie-Erkennung und Ausgabe der Ergebnisse. Um die Anwendbarkeit des Prototyps zu testen, wurde ein vollständiger Anomalie-Erkennungsprozess an der Testanlage durchgeführt: Zuerst wurde das Normalverhalten des Systems aufgezeichnet und mittels des Klassifikators trainiert. Um die Erkennungsrate des daraus resultierenden Modells zu prüfen, wurden verschiedene Netzwerkangriffe durchgeführt. Die dabei aufgezeichneten Daten wurden anhand des Modells klassifiziert. Die Ergebnisse waren vielsprechend. Die meisten Angriffe konnten mit sehr hoher Wahrscheinlichkeit erkannt werden.

In weiterer Folge wurde die Proof-of-Concept Implementierung im Umspannwerk Wels Mitte des Projektpartners Wels Strom probeweise installiert, um die Anwendbarkeit der Lösung auch in einer realen Umgebung zeigen zu können. Es konnte der Prototyp direkt im Umspannwerk Wels Mitte integriert werden. Dazu wurde der Raspberry Pi mittels Port Mirror am Verteilerswitch angeschlossen. Somit konnte der Netzwerkverkehr aller relevanten Endpunkte aufgezeichnet werden und mit den Daten das System spezifisch für das Umspannwerk trainiert werden. Es wurde eine Gesamtgenauigkeit von 92,41 % erreicht. Durch eine längere Aufzeichnungs- bzw. Trainingsphase kann dieser Wert sicher noch weiter verbessert werden.

Zusätzlich zu der Klassifikation von validen Daten wurden im Umspannwerk bestimmte Anomalien (soweit dies ohne Beeinträchtigung des Realbetriebs möglich war) durchgeführt und aufgezeichnet. Diese wurden mit hoher Validität vom System erkannt.

5 Ausblick und Empfehlungen

Die positiven Ergebnisse des Projekts, insbesondere wie sie auch in den Tests unter möglichst realen Bedingungen zu Tage getreten sind, ermöglichen die folgende Einschätzung: Die im Projekt entwickelte Modellierung des Netzwerkverkehrs in einem Umspannwerk respektive in einer Ortsnetzstation erfüllt die gestellten Anforderungen und kann als Basis für die Entwicklung eines Anomalie-Erkennungssystems herangezogen werden. Methoden des maschinellen Lernens sind sehr gut geeignet, den Netzwerkverkehr im Automatisierungsnetz der genannten Komponenten des Energieverteilungssystems zu überwachen, um auftretende Anomalien – seien sie durch Fehlbedienung oder durch Cyber-Angriffe induziert – in Echtzeit zu erkennen. Eine Echtzeiterkennung ist die Voraussetzung dafür, dass gegebenenfalls entsprechende Maßnahmen gesetzt werden können. Diese Einschätzung beruht auch auf den Ergebnissen, die beim Test der Proof-of-Concept Implementierung unter realen Bedingungen erzielt wurden.

Auf Grund dieser Einschätzung erscheint es möglich und realistisch, auf der Basis des Prototyps ein Produkt zu entwickeln, das für die Überwachung des Netzwerkverkehrs in einem Umspannwerk bzw. in einer Ortsnetzstation in der Praxis eingesetzt werden kann. **Der Projektpartner Siemens hat auch angekündigt, eine solche Produktentwicklung zu beabsichtigen**. Ein derartiges Produkt wäre ein wesentlicher Baustein, um wichtige Komponenten der Verteilung elektrischer Energie gegenüber Cyber-Angriffen abzusichern. Das hat umso mehr praktische Bedeutung, als durch das neue NIS-Gesetz (Bundesgesetz zur Gewährleistung eines hohen Sicherheitsniveaus von Netz- und Informationssystemen (Netz- und Informationssystemsicherheitsgesetz – NISG) derartige technische Maßnahmen zur Erhöhung der Sicherheit von Anlagen der kritischen Infrastruktur – wozu die Netze zur Verteilung elektrischer Energie zweifelsohne gehören – vorgeschrieben werden. Die Ergebnisse des Projekts können daher einen wesentlichen Beitrag zur Erhöhung der IT-Sicherheit des Energiesystems leisten. Außerdem konnte durch das Projekt das erforderliche Wissen in Österreich erarbeitet werden und verringert somit die Abhängigkeit Österreichs vom Ausland in einem sehr kritischen Bereich – was insbesondere in Krisensituationen von großer Bedeutung ist.

6 Literaturverzeichnis

[1] Smart Grids Austria, die österreichische Technologieplattform zum Thema Smart Grids des FEEI – Fachverband der Elektro- und Elektronikindustrie und Oesterreichs Energie; http://www.smartgrids.at/smart-grids/

[2] http://whatis.techtarget.com/reference/Smart-Grid-Glossary

[3] https://ics-cert.us-cert.gov/#monitornewsletters

[4] J. Hong, A. Stefanov, C. Liu, M. Govindarasu: Cyber-Physical Security in a Substation; Power and Energy Society General Meeting, 2012 IEEE

[5] U. Premaratne, J. Samarabandu, T. Sidhu, R. Beresh und J.-C. Tan, „Security Analysis and Auditing of IEC61850-Based Automated Substations," IEEE Transactions on Power Delivery, Bd. 4, Nr. 25, pp. 2346 - 2355 , 2010.

[6] M. Rashid, S. Yussof, Y. Yusoff und R. Ismail, „A review of security attacks on IEC61850 substation automation system network," in 2014 International Conference on Information Technology and Multimedia (ICIMU), Putrajaya , 2014.

[7] U. K. Premaratne, J. Samarabandu, T. S. Sidhu, R. Beresh und J.-C. Tan, „An Intrusion Detection System for IEC61850 Automated Substations," *IEEE Transactions on Power Delivery,* pp. 2376-2383, Oktober 2010.

[8] U. Premaratne, J. Samarabandu, T. Sidhu, B. Beresh und J.-C. Tan, „Evidence Theory based Decision Fusion for Masquerade Detection in IEC61850 Automated Substations," in *Information and Automation for Sustainability, ICIAFS* , Colombo, 2008.

[9] C.-W. Ten, J. Hong und C.-C. Liu, „Anomaly Detection for Cybersecurity of the Substations," IEEE Transactions on Smart Grid, Bd. 2, Nr. 4, pp. 865-873, 2011.

[10] Y. Yang, K. McLaughlin, S. Sezer, Y. Yuan und W. Huang, „Stateful intrusion detection for IEC 60870-5-104 SCADA security," PES General Meeting | Conference Exposition, 2014 IEEE, pp. 1-5, 2014.

[11] Y. Yang, K. McLaughlin, T. Littler, S. Sezer, B. Pranggono und H. Wang, „Intrusion Detection System for IEC 60870-5-104 based SCADA networks," Power and Energy Society General Meeting (PES), 2013 IEEE, pp. 1-5, 2013.

[12] P. Oman, E. Schweitzer und D. Frincke, „Concerns about intrusions into remotely accessible substation controllers and SCADA systems," Proceedings of the Twenty-Seventh Annual Western Protective Relay Conference, Nr. 160, 2000.

[13] H. Yoo und T. Shon, „Novel Approach for Detecting Network Anomalies for Substation Automation based on IEC 61850," Multimedia Tools and Applications, Nr. 74, pp. 303-318, 2015.

[14] Gonzales, Thomason: Syntactic Pattern Recognition, Addison Wesley

[15] King Sun Fu: Syntactic Pattern Recognition and Applications, Prentice Hall

[16] Hastie, Dipshirani, Friedman: The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer

[17] Duda, Hart, Stork: Pattern Classification, Wiley

[18] Witten, Frank: Data Mining, Fourth Edition: Practical Machine Learning Tools and Techniques, Morgan-Kaufmann

[19] Borshchev: The Big Book of Simulation Modeling, AnyLogic

[20] Law: Simulation Modeling and Analysis. McGraw-Hill

7 Kontaktdaten

Projektleitung:
Paul Tavolato

Institut für IT Sicherheitsforschung, Department Informatik und Security, FH St. Pölten
A-3100 St. Pölten, Matthias-Corvinus-Straße 15
Tel.: +43 2742 313228-0
Fax: +43 2742 313228-339
paul.tavolato@fhstp.ac.at
https://www.fhstp.ac.at

Weitere Projektpartner:
Siemens AG Österreich
Wels Strom GmbH