Anomalie Erkennung anhand von Flow Informationen in Software Defined Networking

Wie kann Machine Learning die Resilienz kritischer Infrastrukturen optimieren?

Anomaly Detection on the basis of Flow Information in Software Defined Networking

How can Machine Learning optimize critical infrastructures' resilience?

Bachelorarbeit 1
zur Erlangung des akademischen Grades:
Bachelor of Science in Engineering

an der
University of Applied Sciences
Fachhochschule Campus Wien (FHCW)
Favoritenstraße 226, 1100 Wien, Österreich
Fakultät für Informationsechnologien und Telekommunikation

in Kooperation mit der
Austrian Institute of Technology (AIT) GmbH
Center für Digital Safety und Security (DSS)
Abteilung für sichere Kommunikationstechnologien (SCT)
Giefinggasse 4, 1210 Wien, Österreich

eingereicht von:

Julian Timo Magin
julian.timo.magin@stud.fh-campuswien.ac.at julian.magin@ait.ac.at

Betreuer:
FH-Prof. DI Dr. Igor Miladinovic (FHCW)
DI Dr. Oliver Jung (AIT)

Studenten-ID:
1610475122

Spezialisierung:
Telekommunikation

eingereicht am:
19.01.2020

Erklärung:

Ich erkläre, dass die vorliegende Bachelorarbeit von mir selbst verfasst wurde und ich keine anderen als die angeführten Behelfe verwendet bzw. mich auch sonst keiner unerlaubter Hilfe bedient habe.
Ich versichere, dass ich dieses Bachelorarbeitsthema bisher weder im In- noch im Ausland (einer Beurteilerin/einem Beurteiler zur Begutachtung) in irgendeiner Form als Prüfungsarbeit vorgelegt habe.
Weiters versichere ich, dass die von mir eingereichten Exemplare (ausgedruckt und elektronisch) identisch sind.

Datum: Unterschrift:

Declaration of authorship:

First I declare that this Bachelor Thesis has been written by myself. I have not used any other than the listed sources, nor have I received any unauthorized help.

Second I hereby certify that I have not submitted this Bachelor Thesis 1 in any form (to a reviewer for assessment) neither in Austria nor abroad.

Further, I assure that the (printed and electronic) copies I have submitted are identical.

Date: Signature:

Danksagung

Insbesonderer Dank gebührt meinen BA1 Betreuern FH-Prof. DI Dr. Igor Miladinovic von der FHCW und DI Dr. Oliver Jung vom AIT für deren konstruktives Feedback und die Chance im Bereich kritischer Infrastrukturen durch den Einsatz von Machine Learning Methoden aufschlussreiche Lernerfahrungen zu sammeln und meinen Horizont diesbezüglich im Gebiet der Computerwissenschaften auf diese Art und Weise zu erweitern.

An dieser Stelle möchte ich auch die Gelegenheit nutzen, meiner Familie und Angehörigen wie vor allem meiner Freundin Alexandra meine tiefste Dankbarkeit zum Ausdruck zu bringen. Ohne ihre Geduld und ihren zeitlichen Verzicht meiner Anwesenheit hätte diese Arbeit nicht jene Erkenntnisse zu Tage gefördert, die in den nachfolgenden Kapiteln zusammengefasst sind. Des Weiteren gilt mein Dank meinen Kommilitonen Lenhard Reuter (BSc.), Philipp Schunker (Ing., BSc.) und Harun Kocak für Diskussionen über Machine Learning Verfahren allgemein.

Wien, Österreich, Januar 2020
Julian Magin

Kurzfassung

In den letzten Jahren wurden kritische Infrastrukturen, wie sie *Smart Grids (SG)* exemplarisch darstellen, zunehmend Opfer von Advanced Persistent Threats (APTs). Als berühmte Beispiele solcher Angriffe sind unter anderem *MiniDuke* oder *Red October* zu nennen. Der Wohlstand moderner Zivilisationen ist extrem stark abhängig von einer zuverlässigen Energieversorgung. Verschiedene Ansätze werden daher in der Forschung in Betracht gezogen, um einen Weg zu finden, die Resilienz solcher kritischer Infrastrukturen zu erhöhen. Jene Ansätze repräsentieren eine interdisziplinäre Herangehensweise, sodass die Erkenntnisse aus den Wissenschaftsdisziplinen der Mathematik, der Statistik sowie mit Techniken aus dem Bereich des *Machine Learnings* kombiniert werden. Um einerseits sowohl kritische Infrastrukturen vor Schaden zu schützen als auch andererseits Entscheidungen über mögliche Gegenmaßnahmen nach einem erkannten Angriff einleiten zu können, ist es essentiell herauszufinden, inwiefern Netzwerkverkehr ein normales respektive abnormales Verhalten darstellt.

In dieser Arbeit wird ein Konzept für ein *Anomaly Detection System* (ADS) entwickelt, um *Intrusions* zu entdecken. Um wertvolle Informationen diesbezüglich zu gewinnen, werden Entropie Werte von definierten *Features* kalkuliert, um potentielle Anomalien, in einem *Software Defined Network (SDN) Use Case* in Bezug auf SG, zu erkennen. Die Kommunikationsinfrastruktur von Energienetzen hat sich im Laufe der Zeit von einer uni- hin zu einer bi-direktionalen Technologie entwickelt. Im Energienetz beteiligen sich immer mehr dezentral verteilte Produzenten, die sowohl elektrischen Strom einspeißen als auch ebendiesen nutzen. Diese werden *Prosumer* genannt. Somit gab es in den letzten Jahren einen Wandel zu einer zunehmend dezentralisierteren Form der Energieversorgung. Ein SDN Szenario ermöglicht es Ingenieuren*innen, wesentlich dynamischer mit verfügbarem Netzwerkverkehr umzugehen. Da es eine Vielzahl von *Anomaly Detection* Techniken gibt, liegt der Fokus dieser Arbeit darauf, den *Shannon Entropie* Algorithmus näher zu untersuchen. Die Entwicklung eines Konzeptes, welches auf *Machine Learning* Methodiken fußt, birgt jedoch Herausforderungen, welche ebenfalls in dieser Arbeit erörtert werden.

Abstract

In recent years Critical Infrastructures (CIs) such as Smart Grids (SGs) have been the target of Advanced Persistent Threats (APTs). Famous examples of those are known as *MiniDuke* or *Red October*. Because the economic welfare of a society depends strongly on a reliable energy supply, different approaches are combined by the research community to find a way to increase CIs resilience by applying an interdisciplinary approach. Hence, the knowledge from the fields of mathematics, statistics and machine learning are combined. In order to both prevent damage from the CI and to make decisions about possible countermeasures, it is essential to figure out whether some traffic behavior represents a normal or abnormal pattern. An Anomaly Detection System (ADS) is conceptualized to find intrusions. One idea of gaining valuable information is to use Entropy values from defined features to detect potential anomalies in a Software Defined Network (SDN) use case applied in SGs. The communication infrastructure of an energy grid evolved from a uni- to a bi-directional one, in which more and more decentralized energy consumer and producer are involved. Using a SDN scenario enables engineers to handle the available network traffic more dynamically. There are a couple of anomaly detection techniques, the focus of this thesis lays in exploring the *Shannon Entropy* algorithm. Developing a concept based on machine learning one should always keep in mind which potential challenges might occur. These are also discussed as part of this work.

Schlüsselbegriffe

Anomalie Erkennung
Application Programming Interface
Application Plane
Control Plane
Data Plane
IEC 61850
IEC 60870-5-104
Informations-Theoretische Konzepte
Industrial Control Systems
Intrusion Detection System
Entropie
Machine Learning
Unsupervised Learning
Supervised Learning
Reinforcement Learning
Modbus
OpenFlow
Prinicipal Component Analyse
Software Defined Networking
Smart Grids
SCADA
Mustererkennung

Keywords

Anomaly Detection
Application Programming Interface
Application Plane
Control Plane
Data Plane
IEC 61850
IEC 60870-5-104
Information-Theoretic Measures
Industrial Control Systems
Intrusion Detection System
Entropy
Machine Learning
Unsupervised Learning
Supervised Learning
Reinforcement Learning
Modbus
OpenFlow
Prinicipal Component Analysis
Software Defined Networking
Smart Grids
SCADA
Pattern Recognition

Abkürzungsverzeichnis

AD	Anomaly Detection
ADS	Anomaly Detection System
ADU	Application Data Unit
AMI	Advanced Metering Infrastructure
API	Application Programming Interface
ASDU	Application Service Data Unit
BRIC	Acronym for Brasil, Russia, India and China
CI	Critical Infrastructures
DA	Detection Accuracy
DER	Distributed Energy Resources
DDoS	Distributed Denial of Service
DoS	Denial of Service
DSO	Distributed System Operators
ETP	European Smart Grid Technology Platform
FPR	False Positive Rate
HMI	Human Machine Interface
HV	High Voltage
HyRiM	Hybrid Risk Management for Utility Networks Project
ICS	Industrial Controll System
IKT	Informations- und Kommunikationstechnologien
IED	Intelligent Electronic Devices
ID	Intrusion Detection
IDS	Intrusion Detection System
IP	Internet Protocol
ISP	Internet Service Provider
KNN	K-Nearest Neighbor Algorithm
LRM	Linear Regression Model
LV	Low Voltage
ML	Machine Learning
MV	Medium Voltage
NOX	SDN-Controller
NTP	National Technology Platform
OD	Origin Destination
ODL	Open Day Light SDN-Controller
OF	OpenFlow Protocol
ONOS	Open Network Operating System
OCSVM	One-Class Support Vector Machine

Abkürzungsverzeichnis

PDU	Protocol Data Unit
PLC	Programmable Logic Controller
QoS	Quality of Service
RL	Reinforcement Learning
RTU	Remote Terminal Unit
R2L	Remote to Local
REST	Representational state transfer
SCADA	Supervisory Control and Data Acquisition
SDN	Software Defined Networks
SG	Smart Grid
SL	Supervised Learning
TA	Threat Actor
TPR	True Positive Rate
TCP	Transmission Controll Protocol
U2R	User to Root
USL	Unsupervised Learning
VLAN	Virtual Local Area Network
WLAN	Wireless Local Area
xDSL	Digital Subscriber Line

Inhaltsverzeichnis

Kapitel 1

Einleitung

1.1 Aufbau der Arbeit

In diesem Kapitel werden die Zielsetzung und die Methodik der vorliegenden Bachelorarbeit 1 zusammengefasst. Im Anschluss daran werden die Prinzipien von Software Defined Networking (SDN) detailierter in den Subkapiteln über den architektonischen Aufbau von SDN, allgemein Controller in SDN und das OpenFlow (OF) Protokoll beschrieben. Der Hauptteil der Arbeit besteht zudem aus dem Kapitel über intelligente Energienetze, welches in die Subkapitel Supervisory Control und Data Acquisition (SCADA) in Smart Grids (SGs), die verwendeten Protokolle in SGs sowie in die auftretenden Sicherheitsbedrohungen in kritischen Infrastrukturen unterteilt sind. Dies führt den*die Leser*in über zum Kapitel der Anomalie-Erkennung in SDN. Den Anfang bildet hier wiederum das Subkapitel Anomalie-Erkennung mit stark nicht gleich verteilten Daten von Angriffsszenarien, die in einer Laborumgebung generiert wurden. Darauf folgen des weiteren Berechnungen basierend auf dem Grundsatz der *Shannon Entropie*, in Kombination mit linearen Methoden zur Reduktion von Dimensionen. Daran folgend ist das Kapitel Related Work platziert, bevor die Erkenntnisse der Arbeit im Kapitel Zusammenfassung wiedergegeben werden. Die Arbeit schließt mit Ausblicken über mögliche weitere Vorgehensweisen und Ansätze zur Klassifizierung von Anomalien.

1.2 Zielsetzung

Der Energiesektor steckt im Wandel. Die Art und Weise wie elektrischer Strom gewonnen wird, hat sich über die letzten Jahre von einer zentralisierten Form zu einer dezentralen Form, durch das Einbeziehen von erneuerbaren Energieträgern in das Energienetz, entwickelt. Durch die Veränderungen in diesem Bereich sind auch die Steuerungssysteme der Energieanbieter betroffen. Dies bedeutet, dass zunehmend mehr technische Komponenten im Rahmen von mehrstufigen Attacken angegriffen werden, somit als *Advanced Persistence Threats (APT)* auf kritische Infrastrukturen erfolgen. Konkret betrifft das Bauelemente wie *Programmable Logic Controllers (PLCs)* oder *Remote Terminal Units (RTUs)* als Teile von *Industrial Control Systemen (ICS)* respektive dementsprechend auch SCADA Systemen. APTs, wie sie z.B. in Form von Angriffen namens *Flame*, *Stuxnet*, *Nitro*, *Night Dragon Operation* oder *Duqu* aufgetreten sind, stellen für die nationale Sicherheit einer Volkswirtschaft eine Bedrohung dar. Schließlich können sich kaskadierende Effekte von Angriffen auf andere kritische Infrastrukturen, wie die der medizinischen Versorgung oder auf die Einrichtungen der Informationstechnologien- und Telekommunikationsanbieter, auswirken. Die Sicherung des ökonomischen Wohlstands und

somit des sozialen Friedens eines jeden Landes sind daher hochgradig von einer stabilen und zuverlässigen Energieversorgung abhängig. Kritische Infrastrukturen bedürfen daher besonderen Schutzmaßnahmen. Daraus ergibt sich als Anforderung an solche Systeme, Cyberattacken frühzeitig zu detektieren, um gezielt Gegenmaßnahmen einleiten und größere Schäden minimieren zu können. Durch den Einzug von Informationstechnologien und Telekommunikationslösungen (IKT) in die Infrastruktur von ICSs, werden neue Ansätze zur Steuerung von Daten in den zugrunde liegenden ITTK Netzwerken von Energienetzen, auch bekannt als *SGs*, in Betracht gezogen. SDN bildet einen solchen Ansatz, bei welchem *Application-*,*Control-*, und *Data Plane* getrennt und gemäß ihrer Funktionalitäten eingesetzt werden.

Ziel dieser Arbeit ist es, zu untersuchen, wie das Erkennen von Anomalien in Netzwerkverkehr genutzt werden kann, um die Resilienz von kritischen Infrastrukturen zu optimieren. Dies geschieht basierend auf Flow Informationen, die von einem SDN Controller gemanaged werden. Es sollen Entropie Werte berechnet werden, die als Ausgangslage für weitere mathematische Operationen zur Klassifizierung von Datenpunkten dienen sollen. Diese Entropie Werte bilden somit die Grundlage zur Erkennung von unnatürlichem Netzwerkverhalten. Konkret besteht das Ziel dieser wissenschaftlichen Arbeit in der konzeptionellen Erstellung eines Anomalie-Erkennungssystems. Das System besteht aus mehreren Teilen. Diese beinhalten die Berechnung von Entropie Werten aufgrund von SDN Flow Informationen, wie Source IP, Destination IP, Source Port, Destination Port, Packet Count, Average Packet Size und Flow Duration. Die Daten aus den Berechnungen sollen in einem Array abgespeichert werden, das als *Feature Array* benannt wird. Anschließend sollen die Dimensionen des erstellten *Feature Arrays* reduziert werden. Schlussendlich werden die im Datensatz vorhandenen gelabelten Datenpunkte verwendet, um die Performanz des Systems zu ermitteln. Dies bedeutet insbesondere in weiterer Folge den Datensatz in Trainings- und Testdaten zu separieren, um das System im Anschluss trainieren als auch validieren zu können.

1.3 Methodik

Als Ausgangspunkt zur Frage, wie die Verwendung, des aus dem Gebiet der Theoretischen Informatik stammenden Prinzips der Entropie, die Resilienz von kritischen Infrastrukturen optimieren kann, dienen Berechnungen zu Shannon Entropie im Speziellen. Hierbei liegt ein kanadisches Machine Learning Datenset zu Grunde, das *Intrusion Detection Evaluation Dataset (CICIDS2017)*. Dieses wird vom *Canadian Institute for Cybersecurity* der University of New Brunswick öffentlich zur Verfügung gestellt. Es sollen verschiedene Berechnungssalgorithmen kombiniert werden, um letzten Endes neu auftretende Datenpunkte, klassifizieren zu können.

Kapitel 2

Prinzipien von Software Defined Networking

In den folgenden Abschnitten wird nun näher auf den Architektonischen Aufbau von SDN, die in SDN verwendeten Controller und das OF Protokoll eingegangen. Im Anschluss daran werden Intelligente Energienetze und SCADA Systeme beleuchtet, ebenso wie auftretende Sicherheitsbedrohungen in kritischen Infrastrukturen.

2.1 Software Defined Networking Architektur

In SDN sind *Application, Control* und *Data Plane* voneinander getrennt. Der architektonische Aufbau von SDN läßt sich anhand von Abbildung 2.1 illustrieren. Aufgrund der Tatsache, dass der Datenfluss in einem Netzwerk schwierig zu kontrollieren ist, wird dieser Problematik wie folgt begegnet: Die Architektur von SDN wird in kleinere Teile zerlegt. So werden in SDN, im Vergleich zu traditionellen Netzwerken, drei Abstraktionsebenen eingeführt. Ein Vorteil besteht somit in der Vereinfachung des Netzwerkmanagements. Die erste fundamentale Eigenschaft bei der Betrachtung von SDN beinhaltet das Trennen von *Control* und *Data Plane*. Das Weiterleiten von Datenpaketen ist in den *Flow Tables* geregelt. Ein Grund für die Trennung von Verantwortlichkeiten in den *Planes* ist eine angestrebte erhöhte Flexibilität im Netzwerkmanagement. [1, 2]

Eingehende Pakete können sowohl direkt weitergeleitet als auch repliziert und über mehrere *Ports* verschickt werden. In Abhängigkeit von den *Flow Rules* können Pakete auch verworfen werden. Um die Pakete in SDN weiterzuleiten, wird am Switch in der *Flow* Tabelle nachgesehen und darauf hin der zugehörige *Output Port* ermittelt. Der Controller bekommt das Paket zur Analyse, und anhand verschiedener Methoden, die in verschiedenen Applikationen implementiert sind, wird eine neue *Flow Rule* erstellt. Diese neue *Flow Rule* wird der Netzwerkkomponente zur Verfügung gestellt. Diese dient zur weiteren Behandlung der IP Pakete. SDN ändert somit die Art und Weise wie Netzwerke konstruiert und verwaltet werden. [4]

2.2 Software Defined Networking Controller

Der Controller kann als das zentrale Schaltinstrument quasi als *Gehirn* des Netzwerkes betrachtet werden. Es gibt eine Vielzahl an SDN Controllern, welche derzeit am Markt erhältlich sind. Dabei gibt es sowohl *Open Source* als auch kommerzielle Angebote. Diejenigen Controller, die als nicht kommerziell erhältlich sind, haben unterschiedliche Designs, wie un-

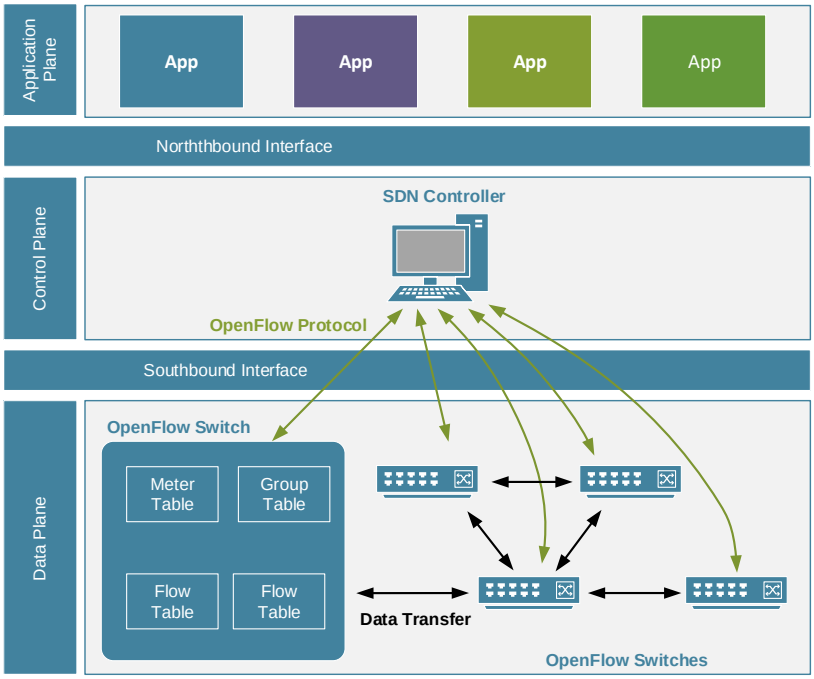

Abbildung 2.1: Konzeptuelle Architektur von SDN und OpenFlow, Quelle: [3]

ter anderem, der auf der Programmiersprache C basierende *NOX* Controller. Dieser kann beispielsweise dynamisch Ende-zu-Ende Pfade konfigurieren. Mit den auf Java basierenden Controller, wie *Beacon* oder *Floodlight* existieren ebenfalls mögliche Lösungsansätze für Testumgebungen, die aber nicht Bestandteil dieser Arbeit sind. Ein weiterer Markteilnehmer ist ein auf der Programmiersprache *Ruby* basierender Controller namens *Trema*. Da ein Controller in einem Netzwerk potentiell einen *Single Point of Failure* darstellen kann, ist die optimale Platzierung von einem oder mehreren Controllern in einem Netzwerk auch von Bedeutung. Die im Rahmen des vertiefenden Wahlfachprojektes erstellte SDN Testumgebung verwendet einen *ONOS* Controller, welcher Java-basiert ist. [1, 5–8]

2.3 OpenFlow Protokoll

Zur Kommunikation zwischen den einzelnen Applikationen und der Data Plane wird das OF Protokoll eingesetzt. Ein OF basierter Switch hat eine oder mehrere Tabellen, mit *Paket-Handling Rules*. Jede dieser Regeln trifft für einen Teil des Netzwerkverkehrs zu, sodass bestimmte Aktionen gesetzt werden. Diese lauten: *Dropping*, *Forwarding* oder *Flooding*. Je nach dem welche Regeln vom Controller vorgegeben sind, kann sich eine Netzwerkkomponente wie ein *Switch*, eine *Firewall* oder wie als ein Netzwerkaddressen *Translator* verhalten. OF ist ein Protokoll, welches die Kommunikation zwischen OF *Switches* und dem Controller ermöglicht. Das Protokoll wurde ursprünglich entwickelt, um es Forschern*innen zu erlauben, Experimente in Netzwerken zu gestalten, um neue Protokolle und Innovationen in Alltagsnetzwerken zu testen. OF basiert auf einem *Ethernet Switch* mit einer internen *Flow* Tabelle und einem standardisierten Interface, um *Flow* Einträge hinzufügen, verändern oder entfernen zu können. Ein Vorteil in der Nutzung dieser Technologie liegt darin, dass das Netzwerkmanagement skalierbarer und flexibler wird. Durch diesen Einsatz können zusätzliche Verzögerungen in der *OF Switch* Architektur in Kauf genommen. [1, 4, 9, 10]

Kapitel 3

Intelligente Energienetze

Das weltweite Verlangen nach elektrischer Energie wird in den nächsten Jahren um mehr als zwei Drittel zunehmen. Diese Betrachtung der steigenden Nachfrage an elektrischer Energie gilt von einem globalen Standpunkt aus. Ein Hauptgrund hierfür ist in dem kontinuierlichen Zuwachs sowohl in der globalen Population als auch im ökonomischen Wachstum zu finden. Dies trifft insbesondere zu in bisher weniger entwickelten Ländern, die einen demographischen Wandel erfahren, sowie in der wachsenden Weltwirtschaft. Daher sind Überlegungen zur Konzeption und Erweiterung von Infrastrukturen der Energienetzbetreiber bedeutsam, um dem stetig wachsenden Energiehunger mit technologischen Mitteln Herr zu werden. Das Interesse an solchen Themen wie denen der Steuerung einer vernetzten und dezentralisierten Energieversorgung spiegelt sich auch in der zunehmenden Anzahl wissenschaftlicher Veröffentlichungen in diesem Bereich wider. Folglich sorgt die höhere Komplexität von Energienetzen auch dafür, dass Anforderungen an Systeme steigen, indem sie nach mehr Zuverlässigkeit im Sinne von höherer Ausfallsicherheit verlangen. Ebenfalls erfordert dies auch strengere Maßnahmen in Bezug auf Sicherheit gegen cyber-kriminelle Aktivitäten. In künftigen Energienetzen wird eine Konstellation bestimmend sein, die sowohl für Produzenten wie auch für Konsumenten eine bereichernde Situation darstellen kann. Konzepte wie *Smart Pricing*, welche die Fluktuationen von preislichen Abweichungen für Kunden in transparenter Art und Weise gestalten, helfen Ressourcen effizienter zu allozieren. Das heißt Produzenten wie auch Konsumenten. Energiekonzerne werden immer mehr Daten über den Energiekonsum ihrer Kunden durch den weiter verbreiteten Einsatz von *Smart Metern* sammeln, um so präzisere Vorhersagen über künftige Nachfragemuster zu generieren. Einerseits helfen solche Datenerhebungen bei der Planung von Produktionskapazitäten und andererseits die Netzauslastung entsprechend zu gestalten. [11–13]

In aufstrebenden Wirtschaften wie zum Beispiel in Brasilien, in Russland, in Indien und in China auch bekannt als *BRIC* Staaten, sowie auch in Europa, sind neue ingenieurwissenschaftliche Techniklösungen von hohem Bedarf, um der Nachfrage nach mehr Energie nachkommen zu können. Neben dem Anstieg des Energiebedarfs ist ein zusätzlich wichtiger Punkt allerdings das Verhindern von *Blackout* Szenarien, da ein Ausfall von Energienetzen kaskadierende Auswirkungen auf viele andere Sektoren haben kann. Technische Transformationen und Reformen innerhalb der elektrischen Energiesysteme beziehungsweise im Energiesektor allgemein sind eine Grundvoraussetzung, um Ausfallszenarien zu vermeiden. Mögliche großflächige Ausfälle sind bei weitem nicht nur ein Problem von weniger industrialisierten Ländern. Daher wurde in 2005 die *European Technology Platform (ETP)* von der EU gegründet, um sich mit solchen Themen auseinanderzusetzen. Ziel war es, ein flexibles, effizientes und zuverlässiges Übertragungssystem zu gestalten. Involviert sind viele Interessensgruppen von Forschungs-

einrichtungen, Energiezulieferern, wie auch Komponenten Herstellern bis hin zu den Energieanbietern selbst. Die *National Technology Platform (NTP) Smart Grids Austria* wurde ins Leben gerufen, um ebendiese Interessen zu bündeln. Was nun unter einem intelligenten Energienetz zu verstehen ist, ist nicht ganz trivial zu beantworten. Laut Technologieplattform *Smart Grids Austria* kann ein Energienetz, welches eine bidirektionale Kommunikation zwischen den einzelnen Netzkomponenten, Erzeugern, Energiespeichern und Konsumenten ermöglicht, als SG bezeichnet werden. Aufgrund des regen Zuwachses an regenerativen Energieformen in der Energielandschaft entstehen auch hier neue Arten von Herausforderungen an die Gesamtinfrastruktur. [11, 14–18]

Um nun dezentralisierte Erzeugereinheiten effizient managen zu können, bedarf es seitens des SGs eines Kommunikationsnetzwerkes und entsprechender Infrastruktur zum Austausch von Informationen und infolgedessen der Kontrolle von involvierten Komponenten. Das Zunehmen von regenerativen Energieformen birgt neue Herausforderungen in Bezug auf die Lastverteilung im Netz. Bei Produktionseinheiten im *Medium Voltage (MV)*, sowie im *Low Voltage (LV)* Bereich können Überspannungsprobleme auftreten. *Distribution System Operators (DSOs)* können nicht bei auftretenden Umspannungen eingreifen. Von daher ist ein Umdenken erforderlich, wie die Spannungskontrolle im Netzwerk generell gehandhabt werden sollte. Energie, die von *Distributed Energy Resources (DERs)* zur Verfügung gestellt wird, kann zu Spannungsspitzen je nach Auslastung des Gesamtnetzes führen. Dies kann wiederum langfristige Schäden an den elektronischen Komponenten der Verbraucher mangels entsprechender Schutzmechanismen in den Verbraucherhaushalten nach sich ziehen. Aufgrund der Tatsache, dass es einen Bedarf darin gibt, DERs sowohl in LV wie auch in MV zu kontrollieren, werden künftige SGs stärker auf IKT Netzwerken aufsetzen. In den vergangenen Jahren gab es ein zunehmendes Interesse, anhand verschiedener Forschungsstudien unterschiedliche Aspekte von IKT Netzen in Kombination mit SGs zu beleuchten. Hohe Schwankungen in der Erzeugung, aufgrund der volatilen Natur von erneuerbaren Energieformen, können durch das Einbeziehen von verschiedenen Energiespeichersystemen ausgeglichen werden. Alternativ kann ein weiterer Ansatz sein, die Auslastung des Netzes auf unterschiedliche Teilnetze zu verteilen. Um diesen Grad an Flexibilität einerseits erst überhaupt erreichen und andererseits dann auch effizient bewerkstelligen zu können, ist eine Art und Weise der Kommunikation von Nöten, die den Datenaustausch von Kontrolleinheiten und Sensoren ermöglicht. Hier kommt SDN ins Spiel. Durch das SDN Konzept im Netzwerk ist es beispielsweise möglich, einzelne technische Anwendungsbereiche innerhalb der Infrastruktur klar zu trennen und die Sichtbarkeit des Netzwerkes zu erhöhen. [15, 19–21]

3.1 Supervisory Control und Data Acquisition Systeme in Smart Grids

Die *Security* von kritischen Infrastrukturen steht mehr und mehr im Fokus des öffent-lichen Interesses aufgrund neu auftretender Cyberattacken, welche auf SCADA Systeme abzielen. Die Raffinesse der Attacken, die stärkere Vernetzung der verbauten Geräte, sowie der Einsatz von Standardhardware bzw. Software haben unter anderem zu dieser Entwicklung beigetragen. SCADA Systeme bestehen aus einer Vielzahl an Computern, welche Echtzeitdaten von Sensoren sammeln, analysieren und auswerten. Jene kommen zum Einsatz, um den Betrieb von großen Industrieanlagen und Einrichtungen zu unterstützen. Potentielle Einsatzgebiete von SCADAs können entweder in der chemischen Industrie, in der Wasserversorgung oder in Elektrizitätswerken liegen. Insbesondere die spezielle Bedeutung solcher Infrastrukturen zur Gewährleistung einer zuverlässigen Stromversorgung ist der Grund, weshalb deren Sicherung

von solch hohem Wert ist. Schließlich können die Konsequenzen mangelhaft umgesetzter Sicherheitsmaßnahmen in diesem Zusammenhang zu katastrophalen Ausmaßen führen. Auch wenn bisher solche SCADA Systeme als vergleichsweise sicher galten, sind diese in den letzten Jahren vermehrt als Ziele von langfristig geplanten mehrstufigen Attacken auserkoren worden. Die Gründe hierfür sind mannigfaltiger Natur. Dank der erhöhten Vernetzung können sowohl Ingenieure*innen als auch System Administratoren*innen per *Remote* auf die Systeme zugreifen und bieten somit unter anderem neue Optionen bzw. weitere Angriffsflächen für Angreifer*innen als bisher. [22–25]

3.2 Protokolle in Smart Grids

In SCADA Systemen kann zwischen technischen Komponenten wie beispielsweise *Control Center*, RTUs oder Intelligente Elektronische Devices (IED) *Human Machine Interfaces (HMIs)* und *Programmable Logic Controllers (PLCs)* unterschieden werden. Beispiele von CPSs können zum einen Photovoltaikanlagen, oder Windturbinen und auch zum anderen Elektrofahrzeuge als potentielle Energiespeicher sein. Für den Austausch von Daten kommen verschiedene Kommunikationsprotokolle wie beispielsweise IEC61850, IEC 608708-5-104 oder Modbus zum Einsatz. Unter der Verwendung von bi-direktionalen Kommunikationsformen und Kontrollapplikationen soll ein SG der Zukunft der Effizienz und Zuverlässigkeit von elektrischen Energieversorgungssystemen dienen. SCADA Systeme lassen sich unter dem allgemeineren Begriff der *Industrial Control Systems (ICS)* zusammenfassen. Die Hauptaufgabe von ICSs besteht darin, Echtzeitdaten zu sammeln, die bei der Automatisierung von technischen Vorgängen innerhalb der Energie-Infrastruktur behilflich sind. [22, 26, 27]

Das IEC 61850 Protokoll

Während in der Vergangenheit der Austausch von Informationen in *Substations* hauptsächlich dadurch erfolgte, dass binäre oder analoge Signale per Kabel gesendet wurden, sind in *Substations* heutzutage moderne Kommunikationsnetzwerke im Einsatz. Um nun den Austausch von Informationen via jener zu standardisieren, wie auch Interoperabilität zwischen Systemkomponenten zu erzielen, wurde IEC 61850 als Standard entwickelt. Das IEC 61850 Protokoll ist als ein international anerkannter Standard ein solches Kommunikationsprotokoll. Dieses ist weit verbreitet vor allem bei der Automatisierung von*Substations*. Die Aufgabe von IEC 61850 besteht darin, die *Substation* Kommunikationssysteme zu unterteilen. Schließlich geschieht diese Aufteilung auf die Prozess-, Interval- und Station- Layer. Der Prozess-Layer inkludiert hierbei eine Variation an technischem Equipment bestehend aus intelligenten elektronischen Komponenten, welche primär die Hauptfunktionalitäten von smarten Substations repräsentieren. Der Interval-Layer beinhaltet Ausstattungskomponenten, die insbesondere auf Geräte zur *Relay Protection* wie auch auf Kontrollgeräte abzielen. Der Station Controll Layer ist hauptsächlich verantwortlich für Aktivitäten wie dem Monitoring und Controlling einer Smart Station. So kann die gesamte Station kontrolliert werden. Vertraut man auf die Trends der Standardisierungsgemeinschaft, dann soll das IEC 61850 in Zukunft die Grundlage bilden, um Informationen von verschiedenen Teilen entlang der Wertschöpfungskette der Energieversorgung zu sammeln. In elektrischen *Substations* werden sogenannte IEDs eingesetzt, um das Equipment zu schützen. Zu erwähnen ist in diesem Zusammenhang, dass es notwendig ist, Informationen zwischen RTUs und lokalen *Human Machine Interfaces (HMI)* auszutauschen, damit die *Substations* ihre Operationen optimal ausführen können. [22, 28–31].

Das IEC 60870-5-104 Protokoll

Zusätzlich kann das IEC 60870-5-104 Protokoll als internationaler Standard für die Kommunikation zwischen SCADA Systemen und *Substations* angesehen werden. Hierbei werden Variationen des IEC 60870-5 Protokolls vermehrt in europäischen Ländern eingesetzt. Verfolgt man die Entwicklung von Netzwerk Technologien, so ist anzumerken, dass IEC 60870-5-104, das auf TCP/IP fußt, in Bezug auf die Kommunikation zwischen *Remote Terminal Units (RTU)* und Kontrollzentren zunehmend zum Einsatz kommt. Im Zuge von Übertragungen vermittelt der Application Layer dieses Protokolls eine *Application Service Data Unit (ASDU)*. Es werden Protokolle in den Komponenten Protokolle verwendet, die nicht mit dem Internet verbunden sind. Diese weisen jedoch auch *Vulnerabilities* auf, welche von Angreifern*innen potentiell ausgenutzt werden können. Um nun zuverlässige SG Operationen zu garantieren, ist es erforderlich authentische Messdaten zu erhalten. Folglich ist es sinnvoll, diese Kommunikationsstrukturen in sekundären *Substations* gegen *Cyber-Security* Angriffe zu schützen. Zusammenfassend ist zu den Protokollen der 60870 Reihe zu sagen, dass die IEC 60870 Versionen eine Ansammlung von offenen Standards repräsentieren. Diese wurden verfasst von jenen Autoren, die für die *International Electrotechnical Commission (IEC)* im Hinblick auf Kontrollmechanismen für SCADA Systeme tätig sind. Im Kontext dieser Bachelorarbeit 1 erfolgt eine Betrachtung von SCADA Systemen insbesondere im Hinblick auf den konkreten Standard 60870-5-104. Die Version 104 ist allerdings vorrangig in der Kontrolle von IKT Komponenten für Einrichtungen zur Energieumwandlung wie z.B. bei Wasserwerken das Mittel der Wahl. [32–35]

Das Modbus Protkoll

Zur Datenkommunikation ist auch das Modbus Protokoll wichtig. Es ist oft im Einsatz bei der Automatisierung von industriellen Prozessabläufen. Konventionelle industrielle Kontrollsysteme folgen einem International Electrotechnical Commission (IEC) *Master-Slave* Ansatz. Dieser weist eine zentralisierte Kontrollarchitektur auf. Somit übernimmt ein zentraler Controller die Hauptkontrollentscheidungen. Dadurch werden *low-level* Geräte anhand von Punkt-zu-Punkt Verbindungen kontrolliert. Einerseits ist Modbus zwar ein zuverlässiges Protokoll, andererseits benötigten die Implementierungsschritte und Re-Konfigurationen eine lange Zeitpanne. [22, 36]

3.3 Sicherheitsbedrohungen in kritischen Infrastrukturen

Nachdem nun die Architektur und Funktionsweisen von SDN und SCADA Systemen vorgestellt wurden, gilt es Sicherheitsbedrohungen zu betrachten, welchen sowohl SDN Infrastrukturen wie auch intelligenten Energienetzen ausgesetzt sind. Verschiedene APTs wie bereits zu Beginn der Arbeit in Kapitel 1 erwähnt, halten nun seit einiger Zeit durch ihre weitreichenden Auswirkungen auf andere technische Systeme die Verantwortlichen im Energiesektor aufgrund ihrer Raffinesse in Planung und Komplexität der Umsetzungen in Atem.

Security Angelegenheiten in SDN

Neben den Vorteilen, die eine SDN Architektur mit sich bringt, entstehen allerdings auch neue *Security* Herausforderungen. Diese sollten SDN Infrastruktur Anbieter bedenken, um nicht Opfer von zum Beispiel *Man-in-the-middle Attacken*, *DoS-* oder *Saturation* Attacken

zu werden. Aufgrund der architektonischen Aufteilung in SDN gilt es, Sicherheitsproblematiken je *Plane* differenzierter zu betrachten. So bestehen auf *Application Plane* beispielsweise folgende Bedrohungen: [37]

1. Mangel an Authentifizierung und Autorisierung
2. Betrügerische *Flow Rule(s) Insertion*
3. Mangel an Zugriffskontrolle *(Access Controll)* und Verantwortung *(Accountability)*

Die Control Plane hingegen ist tendenziell eher DoS Attacken ausgesetzt, genauso wie Versuchen nicht-authorisierte Zugriffskontrolle zu erlangen. Das bedeutet konkret, dies geschieht, wenn keine Mechanismen vorhanden sind, welche die Zugriffskontrolle von Applikationen überprüfen. Die Zentralisierung der Steuereinheit auf eine Entität kann wiederum problematisch bei der Skalierbarkeit und Verfügbarkeit sein. Die dritte Ebene, die *Data Plane*, kann *Flooding* Attacken unterliegen, da die *Flow* Tabellen der *OF Switches* nur eine limitierte Anzahl von *Flow Rules* speichern kann. Da die *Data Plane* hochgradig von der *Controll Plane* abhängig ist, hat zum Beispiel ein *Hijacking* Angriff auf den Controller auch Auswirkungen auf die beide anderen *SDN Planes*. Im folgenden Abschnitt wird kurz darauf eingegangen wie das Zusammenspiel von diversen Angriffen zu einem kaskadierenden APT Problem werden kann. [37]

Advanced Persistent Threats

Beim *Deployment* von APTs werden eine Vielzahl an verschiedenen Attacken-Arten kombiniert. So vermag ein APT mit einem *Social Engineering* Versuch beginnen, um technische *Exploits* vorzubereiten. Diese zielen darauf ab, Schaden an Teilen der Infrastruktur zu verursachen und organisatorische Sicherheitsmechanismen zu umgehen. Neben *Stuxnet*, welches zum Ziel hatte, nukleare Zentrifugen in einem Kraftwerk im Iran zu sabotieren, sind noch weitere APTs in Erscheinung getreten in den letzten Jahren. Zu diesen zählen auch *MiniDuke*, *Operation Aurora*, *Shady Rat*, *Black Energy*, *Crashoverride* oder auch *Red October*, um nur einige zu nennen. Diese Angriffe repräsentieren eine Form der Spionage heutzutage. Anstatt einzelne Schwachstellen auszunutzen, konzentrieren sich APTs auf eine Verkettung von jenen Schwachstellen in unterschiedlichen Systemen, um innerhalb einer Unternehmensnetzwerkstruktur die Hochsicherheitsareale zu beeinflussen. Nachdem der Zugang zum internen Netz erreicht wurde, kann ein sogenannter *Threat Actor* (TA) zunächst, trotz bestehender Sicherheitsrichtlinien, unentdeckt bleiben. Ein Grund hierfür ist, dass das Monitoring ein intensiver Prozess ist. In *Intrusion Detection Systemen (IDS)* müssen vielerlei humane Ressourcen eingesetzt werden, um Kennzahlen wie z.B. *False Positives* richtig zu analysieren [38–40]. Um nun das Risiko von APTs zu vermindern, gibt es die Möglichkeit, Spieltheorie einzusetzen, indem wie im Projekt *Hybrid Risk Management for Utility Networks (HyRiM)* Defensivstrategien optimiert werden [41]. Mögliche Bedrohungen können wiederum in ihrer Attacken Art und Weise unterschieden werden. Beispielsweise kann durch *Denial of Service (DoS)* Angriffe ein Service gezielt in die Knie gezwungen werden. Des Weiteren kann mittels *Remote to Local (R2L)* oder eben auch durch *User to root (U2R)* Attacken versucht werden, systeminterne Daten zu akquirieren. [42]

Kapitel 4

Anomalie-Erkennung in Software Defined Networking

Ein effektives Anomalie Erkennung System zu entwerfen, beinhaltet das Herausfiltern von relevanten Daten aus einer Vielzahl von mehrdimensionalen Daten. Unterschiedliche Anomalien spiegeln sich auf verschiedene Art und Weise in Netzwerkstatistiken wider. Aus diesem Grund kann es sich schwierig gestalten, ein generelles Modell zu konzipieren. [42]

4.1 Motivation für Anomalie Erkennung

Mit zunehmenden Datenvolumen ist es für Netzwerk Operatoren ein immer schwierigeres Unterfangen, jedes einzelne Paket, in dem zu überwachenden Netzwerk zu inspizieren. Anomalie Erkennungsmodelle in Kommunikationsnetzwerken bieten die Grundlage, um neue Attacken aufzuspühren, Fehlkonfigurationen auf den Grund zugehen oder dem Ausfall von Netzwerkkomponenten nachzugehen. Das Detektieren von Attacken stellt sowohl für Netzwerk Operatoren wie auch für Wissenschaftler*innen insbesondere bei *Zero-Day Attacken* eine Herausforderung dar. Um Anomalien in Echtzeit aufzuspühren, benötigen Netzwerkadministratoren*innen Methoden, um neben *signature-based* Techniken auch durch die Verwendung von *statistical-based* Ansätzen Anomalien zuverlässig zu erkennen. Entropie basierte Methoden spiegeln die Verteilung von Netzwerk *Features* wider. Betrachtet man Kombinationen aus mehreren *Features* als Auftrittswahrscheinlichkeiten von Datenpunkten, so lassen sich anhand dieser auftretenden Konzentrationen an Punkten abnormale Netzwerkverhalten ableiten. [43] Die Grundidee hinter Anomalie-Erkennungsmechanismen besteht darin, mittels historischer Daten zu trainieren. Anhand dieser sind Anomalien, auch oft *Outlier* genannt, herauszukristallisieren, um diese mit bestehenden Mustern zu vergleichen. Anwendungsgebiete von Anomalie Erkennung erstrecken sich über Kreditkartenmissbrauch über Aktienkursabweichungen an der Börse bis hinzu *Intrusion Prevention* in *Cyber-Security* Systemen sowie weiteren Anwendungsbereichen in unterschiedlichen Branchen, in denen statistische Ausreißer von Relevanz sind. Hat man beispielsweise bereits im Rahmen seiner Forschungstätigkeiten tausende Datenpunkte gesammelt und in einem geeignetem Format zur Weiterverarbeitung persistiert, kann man diese zum Anlernen seines Systems nutzen. [44]

Machine Learning Ansätze

In *Machine Learning (ML)* lassen sich daher insbesondere folgende drei Ansätze unterscheiden: *Unsupervised Learning (USL), Supervised Learning (SL) und Reinforcement Learning*

(RL). Mit dem Einsatz von *ML* Methoden ist es möglich, Algorithmen zu entwickeln anhand derer Computerprogramme Muster zu erkennen erlernen. Kombiniert man mathematische Konzepte mit maschinellen Lerntechniken ergeben sich interessante Anwendungsfelder im Bereich der Computerwissenschaften. Diese sind unter anderem *Image Processing, Speech Processing, Natural Language Processing* und die Robotik. Aber nicht nur klassische Informatikzweige sind von dem Mehrwert betroffen, den ML bieten kann, vielmehr können enorme Fortschritte auch in anderen Wissenschaftsbereichen erzielt werden. Darunter fallen die Biologie, medizinische Untersuchungsmethoden, die Astronomie, die Physik und auch Materialwissenschaften. Je nach Umfeld und technischer Problemstellung gilt es abzuwägen, welche Methodik im jeweiligen Szenario im Hinblick auf unterschiedlichen Parameter, die optimiert werden sollen am sinnvollsten einzusetzen ist. [45]

Unsupervised Learning Ansatz

Das Ziel von *USL* besteht allgemein darin, Muster in Daten automatisiert zu erkennen. Beim *USL* Ansatz sind die Daten nicht gelabelt. Labeling wird im Abschnitt unten bei Supervised Learning (SL) beschrieben. Das bedeutet zeitgleich, dass die Klassen *Label* nicht bekannt sind. Daten werden geplottet, um zu sehen, wie sie sich als *Cluster* verhalten. Beim *Clustering* gibt es zwei Arten dieses zu unterscheiden. Entweder man betrachtet hierarchische oder partitionale *Cluster*. Der Begriff des *Labeling* hingegen bedeutet im Bereich des *ML* eine Art Codierung, die einem Datenpunkt eine eindeutige Zuordnung zu einer spezifischen Gruppe von Daten mit gleichen Charakteristiken vergibt. *Unsupervised* Lernalgorithmen wie sie zum Beispiel der k-means Algorithmus darstellt, spezialisieren sich auf das Auffinden von *Clustern*, die quasi Punktwolken, akkumulierter Punkteballungen symbolisieren. Dabei wird der Mittelpunkt eines jeden *Clusters* festgelegt. Voraussetzung für *Unsupervised* Algorithmen ist das Wissen, über die Anzahl der *Cluster* der Datenpunkte. Die Zentren der *Cluster* zu berechnen ist die Hauptaufgabe des k-means. Neben k-means gibt es auch andere Möglichkeiten wie zum Beispiel den Einsatz von einer Hauptachsentransformation, der Principal Component Analysis (PCA), bei der multivariate Daten in einen kleineren Untervektorraum überführt werden. Somit werden Informationen zusammengefasst, und im Anschluss weiterverarbeitet. Diese Technik soll bei dieser wissenschaftlichen Arbeit eingesetzt werden. [46]

Clustering wird auch als *Exploratory Data Analysis* bezeichnet, und kann als wichtiger Schritt angesehen werden, um ein Datenset zu verstehen. Hierbei ist das Ziel, aus unbekanntem Verhalten der Daten durch die Aufteilung eines endlichen Datensets (in kleinere Datensets mit bis dato verborgenen Datenstrukturen), konkrete Muster herzuleiten. Betrachtet man ein Datensatz mit n Punkten in einem zwei-dimensionalen Vektorraum, somit ist das Ziel mit Hilfe von *Clustering* Daten so zu gruppieren, die Gemeinsamkeiten aufweisen. [47]

Supervised Learning Ansatz

Dieser Abschnitt befasst sich nun mit der zweiten Kategorie von ML. SL liefert eine Bandbreite an nützlichen Werkzeugen, um Daten zu klassifizieren und weiterzuverarbeiten, ebenfalls unter der Anwendung von ML Algorithmen. Im Kontrast zu USL werden Daten verwendet, die gelabelt sind, das bedeutet konkret einen Satz an Daten zu analysieren, der schon klassifiziert wurde, um einen Lernprozess anzuregen. Das Datenset wird wiederum verwendet, um als Grundlage für die Vorhersage der Klassifizierung anderer ungelabelter Daten zu dienen. In SL unterscheidet man zudem zwischen linearer Regression und Klassifizierungstechniken. Unter linearer Regression wird das Auffinden von Beziehungen zwischen quantitativen Daten verstanden. Das Modell der linearen Regression ist aus der Statistik bereits seit längerem bekannt

und eine der älteren Methoden, um Vorhersagen zu treffen. Dieses Verfahren kann angewendet werden, um beispielsweise Korrelationen zwischen Marketingbudgets und Verkaufszahlen eines Unternehmens ausfindig zu machen. Andere Anwendungsszenarien betreffen vielmehr medizinische Teilbereiche wie die Radiotechnologie, um insbesondere einen etwaigen Zusammenhang zwischen Strahlentherapie und Tumorgrößen festzustellen. Auch im Bankenwesen sind Einsatzmöglichkeiten gegeben wie beispielsweise die Identifizierung von betrügerischen Kreditkartentransaktionen. Dies geschieht durch das Analysieren von Daten und durch die Wiedererkennung von Mustern. Zu bekannten Klassifizierungen zählen hiermit, die Logistische Regression, die Lineare Diskriminanten Analyse, die *K-Nearest Neighbors* Methoden oder Suchbäume, Neurale Netze ebenso wie Support Vector Machines. [48]

Ein SL Algorithmus verwendet also ein bekanntes Set von *Input* Informationen und ist sich bewusst über die zu erwartenden *Outputs*, um anhand derer das Regressions- beziehungsweise Klassifizierungsmodell zu trainieren in Bezug auf neue Datenpunkte. SL nutzt die oben genannten Techniken, um ein vorhersagbares Modell zu entwickeln. Neben den bisher aufgezählten Vorgehensweisen kann auch ein *Random Forest-* oder ein *Naive Bayes-* Verfahren eingesetzt werden. Klassifizierungsmodelle sortieren die Eingangsdaten in verschiedene Kategorien. Weitere Anwendungen sind zum Beispiel Spracherkennungsverfahren oder die Überprüfung von manuell geschriebenem Text, dem sogenannten *handwritting recognition*, um zu erkennen, ob eine Mail echt ist oder ob sie Spam beinhaltet. Viel versprechende Ergebnisse konnten verschiedene Bilderkennungsverfahren auch im medizinischen Bereich erzielen, etwa bei der Unterscheidung von Bildern mit Tumoren und solche ohne Tumorbildung. [49]

Reinforcement Learning Ansatz

Neben Unsupervised und Supervised Learning Methoden gibt es noch eine dritte Kategorie das sogenannte Reinforcement Learning (RL). Hierbei wird sequentiell gelernt. Dabei werden Entscheidungen schrittweise gefällt, um Probleme zu lösen. Ein Beispiel für einen solchen Ansatz zeigt P. Schunker in seiner Arbeit über einige interessante Möglichkeiten auf, die das Prinzip von RL zur Lösung eines Schachspiels bietet. [50] Des Weiteren werden Markov Decision Prozesse eingesetzt als Stellvertreter von Reinforcment Algorithmen zur Berechnung von optimalem Lernverhalten [51]. Bei RL Verfahren gibt es einen Agent und eine Umwelt von welcher er lernt. Der Agent probiert die *Rewards*, die er von seiner Umwelt bekommt, zu maximieren. Er agiert auf verschiedene Arten, um anhand der Reaktionen seiner Umwelt zu erlernen, welche Konsequenzen seine Handlungen bewirken. Eine der größten Herausforderungen bei der Verwendung von RL ist das Assoziieren von Aktionen mit verspäteten *Rewards*, welche der Agent erhält, lange nachdem die Aktionen stattgefunden haben, die den *Rewards* generiert haben. Daher ist dieser Ansatz bisher stark im Einsatz bei der Erkundung von Gesellschaftsspielen wie TicTacToe oder Go. [52]

RL Algorithmen können aber auch eingesetzt werden, um die Performanz von IDS zu verbessern. Allerdings gilt auch hier, dass die Implementierung einer *Reward* Funktion in Kombination mit einem IDS als schwierig gilt, da es keine automatisierte Art gibt Eindringlinge zu identifizieren. Auf die Vor- und Nachteile dieses Modells für ML wird im Rahmen dieser Arbeit nicht eingegangen, da die Priorität auf der Verwendung von Entropie und PCA, beruhen. Dies stellt eine *USL* Methode dar. [53]

4.2 Berechnungskonzepte der Informationstheorie

Das Konzept der Entropie ist in vielerlei technischen Bereichen im Einsatz. Es unterstützt als statistisches Mittel Ingenieure*innen und Analysten*innen, die auf dem Gebiet der Thermodynamik tätig, wie auch mit Problemen der Informationstheorie per se konfrontiert sind. Boltzmann zeigte die Verwendung von Entropie für die Thermodynamik bereits Ende des 19. Jahrhunderts. [54] Betrachtet man jene Entropie, welche von Claude Shannon [55] im Jahre 1948 auf Informationstheoretische Problemstellungen adaptiert wurde, lassen sich mehrere Subkategorien und besondere Ausprägungsformen unterscheiden. So existieren neben der in Definition 4.2.1 näher erläuterten Shannon Entropie prinzipiell, die weiteren folgenden informationstheoretischen Konzepte: Die konditionale Entropie beziehungsweise die relativ-konditionale Entropie, der Ansatz von *Information Gain* sowie die Umsetzungen zu *Information Cost*. Mithilfe der konditionalen Entropie lassen sich die jeweiligen Datenpunkte zwischen zwei Wahrscheinlichkeitsverteilungen ins Verhältnis setzen und spezifischer betrachten. Dies kann von Interesse sein, wenn man bestrebt ist herauszufinden, welches Maß an Unsicherheit für den Rest der Datenpunkte übrig bleibt, wenn man mehrere Datenmengen vergleichen will. Allerdings soll im Rahmen dieser wissenschaftlichen Arbeit als Grundlage für weitere Berechnungen der Fokus auf der Shannon Entropie liegen. [56]

Konzept der Shannon-Entropie

Eine dieser Entropie-Arten stellt wie bereits oben aufgelistet die Form der Shannon-Entropie dar. Sie ist auch bekannt als Shannon-Wiener Index. Sie ist ein Maßstab für Unsicherheit und hilft somit die Streuung von Zufallsvariablen zusammenzufassen. Für ein bestimmtes Datenset mit Datenpunkten, die auf unterschiedliche Klassen verteilt sind, lässt sich folgendes festhalten:

Definition 4.2.1. Shannon Entropie ist definiert als eine diskrete Variable, für die folgendes gilt: H(Y) ist der Entropie Wert einer Zufallsvariablen Y, siehe hierzu Gleichung (4.1) [57], [44]. Hierbei sind Y_α mögliche Werte also Merkmalsausprägungen und w(y_α) ist die Wahrscheinlichkeit, mit welcher Y den Wert y_α annehmen kann, wie in Formel (4.2) zu sehen ist.

$$H(Y) = \left(\sum_{\alpha=1}^{n} w(y_\alpha) log_2(\frac{1}{w(y_\alpha)})\right) = -\left(\sum_{\alpha=1}^{n} w(y_\alpha) log_2(w(y_\alpha))\right) \tag{4.1}$$

wobei

$$w(y_\alpha) = W(Y = y_\alpha) \tag{4.2}$$

Wenn nun der Entropie Wert 0 beträgt, bedeutet dies, dass alle Datenpunkte gleich sind und wenn der Entropie Wert hingegen bei $log_2 n$ liegt, sind die Datenpunkte verschieden. Dies kann helfen, anhand der Streuung der Daten, Analysen über das Netzwerkverhalten zu schließen. Insbesondere beim Anwendungsfall von Anomalie Erkennung in einem Computernetzwerk bedeutet dies das die Zufallsvariablen, die mehr zerstreut sind, einen höheren Entropie Wert implizieren. Enthalten die Daten weniger Streuung, spiegelt sich das letztlich in einem geringen oder gar 0 als Entropie Wert wider. Vom Entropie Wert lassen sich weitere Rückschlüsse ziehen, vor allem auch auf die Verteilung der betrachteten Klassen, welche unterschiedliche Angriffsszenarien symbolisieren. Dies bedeutet konkret, dass bei einem niedrigen Entropiewert eine Klasse und bei hohem Entropiewert mehrere Klassen vorliegen. Vergleicht man die Werte miteinander lassen sich weitere Rückschlüsse auf Verteilung der Daten ziehen. [44]

4.3 Methoden zur Reduktion von Dimensionen

Unter der *PCA* versteht man eine mathematische Vorgehensweise, mit deren Hilfe man die Dimensionen eines mehrdimensionalen Datensets auf einen Unterraum, mit weniger Dimensionen, reduzieren kann. Theoretisch vermindert PCA somit die Anzahl der Variablen, während zeitgleich die Varianz des originalen Datensets beinahe erhalten bleibt [44]. PCA ist zudem auch bekannt als Karhunen-Loeve Transformation [58].

Grundsatz der Principal Component Analyse

PCA ist ein wirkungsvolles Mittel, welches in vielen Bereichen zum Einsatz kommt, wie unter anderem auch bei *Image Recognition* Verfahren. Es ist oft in Verwendung zum Auffinden von Mustern in hoch dimensionellen Datensätzen. Bei PCA handelt es sich grundsätzlich um eine *Unsupervised* Methode. PCA kann auf mehrere Arten eingesetzt werden. Zum einen wie bereits beschrieben zur Reduktion von Dimensionen, aber zum anderen eben auch als Visualisierungswerkzeug oder zum Herausfiltern von *Noise*. Eine Anwendungsmöglichkeit besteht aber auch zum Zwecke des *Feature Extraction* und *-Engineering*. Aufgrund der vielseitigen Verwendungsmöglichkeiten von PCA, wird es in unterschiedlichen Disziplinen eingesetzt. Um seine Datenpunkte zunächst einmal zu visualisieren, ist es ein geeignetes Mittel, um die Beziehungen der Punkte untereinander zu veranschaulichen. Bei der Verwendung von statistischen Lernmethoden gilt es deren Vor- und Nachteile im jeweiligen Kontext abzuwägen. Daher ist im Zusammenhang mit PCA festzuhalten, dass *Outlier* im Datenset Auswirkungen auf die Analyse der Datenpunkte haben, was potentiell als Schwäche dieser Technik gewertet werden kann. Aus diesem Grund gibt es mehrere Abwandlungen von PCA. Die *Scikit-Learn* Bibliothek, welche in Python 3.7 zur Verfügung steht, beinhaltet einige Varianten von PCA Techniken, insbesondere die *RandomizedPCA* oder die *SparsePCA*. [59]

Der folgende Abschnitt gibt dem*der Leser*in eine Einführung in die Konzepte der Hauptachsentransformation auch PCA genannt, die als Mittel zur Reduktion der Dimensionen der Daten dienen. Zu Beginn werden die mathematischen Konzepte aus den Gebieten der linearen Algebra und Wahrscheinlichkeitstheorie erläutert, die bei PCA verwendet werden. Dazu gehören die *Standard Abweichung*, die *Kovarianz* sowie *Eigenvektoren* und *Eigenwerte*.

Mathematisches Hintergrundwissen für multivariate lineare Regressionen

Dieser Abschnitt gibt einen Überblick über das mathematische Rüstzeug, welches notwendig ist, um die Schritte, die bei einer PCA durchgeführt werden, nachzuvollziehen. Ziel ist es, ein Verständniss zu kreieren, weshalb PCA als Methodik einzusetzen ist und wie die Ergebnisse zu interpretieren sind. Zudem werden statistische Maße kurz vorgestellt, um die Verteilung der Datenpunkte zu erklären. Um das arithmetische Mittel, respektive die Standard Abweichung oder andere statistische Werkzeuge aus einer Menge von Zufallsvariablen per se zu verstehen, ist es sinnvoll, exemplarisch kleineres Datenset zu untersuchen. Das arithmetische Mittel $\overline{y}$ ist dabei definiert als die Summe aller Merkmalsausprägungen y_1 bis y_n dividiert durch die Anzahl n ihrer Vorkommnisse (siehe Formel 4.3) [60]. Die Formeln 4.3 bis 4.8 gelten für unabhängige und identisch verteilte Zufallsvariablen [61]:

$$\overline{y} = \left(\frac{1}{n} \sum_{\alpha=1}^{n} y_\alpha = \frac{y_1 + y_2 + \cdots + y_n}{n} \right) \tag{4.3}$$

Das arithmetische Mittel $\overline{y}$ konvergiert somit stochastisch mit zunehmendem Stichprobenumfang gegen den Erwartungswert der Stichprobenmenge, ergo des Datensets. Dies ließe sich mit

Hilfe der Grenzwertsätze und der Ungleichung von Tschebyscheff beweisen. Dies wiederum würde aber zu tief in die Materie von Teilgebieten der Wissenschaftsdisziplin Mathematik greifen und schließlich am eigentlichen Ziel dieser Arbeit vorbeiführen. Mit der Formel 4.4 kann nun die (auch bekannt als Stichproben-, respektive empirische-) Standard Abweichung σ hergeleitet werden: [61]

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{\alpha=1}^{n}(y_\alpha - \overline{y})^2} \tag{4.4}$$

Quadriert man nun die Standard Abweichung hingegen kann man daraus wiederum die empirische *Varianz* σ^2 (wie in Formel 4.5 zu sehen) ableiten [61]:

$$Var(Y) = \sigma^2 = \frac{1}{n-1}\sum_{\alpha=1}^{n}(y_\alpha - \overline{y})^2 = \frac{1}{n-1}\sum_{\alpha=1}^{n}(y_\alpha - \overline{y})(y_\alpha - \overline{y}) \tag{4.5}$$

daraus ergibt sich die empirische *Kovarianz* (wie in Formel 4.6 zu sehen), um den Zusammenhang zwischen zwei Zufallsvariablen Y und Z zu verdeutlichen [61]:

$$Kov(Y,Z) = \frac{1}{n-1}\sum_{\alpha=1}^{n}(y_\alpha - \overline{y})(z_\alpha - \overline{z}) = \frac{1}{n-1}\sum_{\alpha=1}^{n}(z_\alpha - \overline{z})(y_\alpha - \overline{y}) \tag{4.6}$$

Ein potentieller Weg alle in Frage kommenden *Kovarianz* Werte zwischen allen Dimensionen zusammenzufassen, besteht darin, diese Werte zu berechnen und anschließend in einer *Kovarianz* Matrix abzulegen. Dabei ist die *Kovarianz* Matrix für ein Datenset mit n Dimensionen wie in Gleichung 4.7 definiert. Eine Matrix der Form $Kov^{(n\times n)}$ symbolisiert hierbei mit n Zeilen und n Spalten eine Zusammenfassung aller zu berechnenden Kovarianzen. Dies bedeutet, dass jeder Eintrag einen konkreten *Kovarianz* Wert darstellt, der zwischen zwei Zufallsvariablen ermittelt wurde. [60]

$$Kov^{(n\times n)} = (k_{\alpha,\beta}, k_{\alpha,\beta}) = kov(Dim_\alpha, Dim_\beta) \tag{4.7}$$

Eine beispielhafte Kovarianz-Matrix ist unten in 4.8 zu sehen. Sie ist zu ihrer Hauptdiagonalen symmetrisch und repräsentiert die Kovarianzen der Zufallsvariablen U,W,Y sowie Z. Diese kann letztlich auf zusätzliche Dimensionen erweitert werden [60]:

$$Kov = \begin{pmatrix} Kov(U,U) & Kov(U,W) & \dots & Kov(U,Y) & Kov(U,Z) \\ Kov(W,U) & Kov(W,W) & \dots & Kov(W,Y) & Kov(W,Z) \\ \vdots & \vdots & \dots & \vdots & \vdots \\ Kov(Y,U) & Kov(Y,W) & \dots & Kov(Y,Y) & Kov(Y,Z) \\ Kov(Z,U) & Kov(Z,W) & \dots & Kov(Z,Y) & Kov(Z,Z) \end{pmatrix} \tag{4.8}$$

Eine PCA wird also verwendet, um die Anzahl der Variablen ergo die auftretenden Dimensionen zu reduzieren, jedoch immer darauf achtend, dass die Varianzen des originalen Datensets annähernd erhalten bleiben. Zunächst einmal müssen die gesamten Datenpunkte verfügbar sein, um aus diesen die Eigenvektoren und Eigenwerte der Kovarianz-Matrix zu bestimmen. Diese Art der PCA kann als ineffizient angesehen werden. Dies gilt insbesondere für Applikationen, deren Aufgaben das Streamen von Daten ist, da jedes mal die Werte neu ermittelt werden müssen, die mit den Formeln 4.3 bis 4.8 berechnet werden [62]. Jolliffe zeigt in seinem Buch über PCA, in dem er potentielle Arten von PCA ausführlicher diskutiert, dass PCA auch zum Einsatz kommt, um *Features* zu extrahieren, anstatt sich nur auf die Reduktion von Dimensionen zu fokussieren [63].

Abbildung 4.1 von C. Bishop, zu finden in [58], zeigt exemplarisch eine Illustration einer orthogonal Projektion im Rahmen einer PCA.

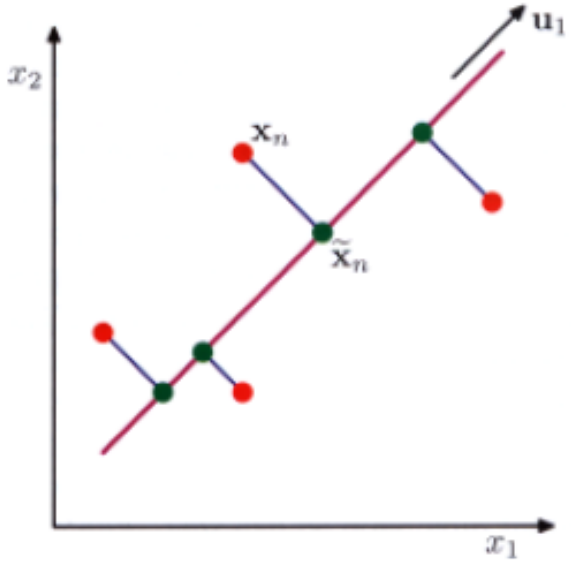

Abbildung 4.1: Illustration einer PCA orthogonalen Projektion, Quelle: [58]

Die Abbildung 4.1 zeigt vor allem, dass die roten Punkte auf die grünen Punkte projiziert werden. Dies geschieht durch die Maximierung der Varianz der grünen Punkte. Hierbei betrachtet man ein Datenset mit Merkmalsausprägungen $[x_n]$ wobei gilt n=1, ..., N und x_n symbolisiert eine euklidische Variable in euklidischen Vektorraum mit der Dimension D, und $\tilde{x_n}$ die Projektion im gleichen. Ziel ist diesbezüglich das Überführen der Daten in einen Unterraum der Dimension M $<$ D. Dabei steht das M stellvertretend für die Formulierung der maximalen Varianz [58]. Danach beschreibt unten stehende Auflistung die einzelnen Schritte, die bei einer PCA sukzessive angewendet werden [60]:

1. vollständige Daten erhalten, aus einem geeigneten Datenset$_{prePCA}$ akquirieren,
2. arithmetisches Mittel der Datenpunkte berechnen,
3. Kovarianzen berechnen und in Kovarianz-Matrix zusammenfassen,
4. Eigenvektoren dieser Kovarianz-Matrix ermitteln,
5. Eigenwerte dieser Kovarianz-Matrix ermitteln,
6. *Principle Components (PCs)* auswählen, anhand der höchsten Eigenwerte priorisieren
7. und im Anschluß das Datenset$_{postPCA}$ mit PCs Werten generieren.

Neben den Werkzeugen aus dem Bereich der Statistik sind für eine PCA noch Konzepte aus der Matrizen Algebra relevant. Hauptsächlich von Interesse für die Betrachtung von PCA sind hierbei die Eigenvektoren und Eigenwerte der Kovarianz-Matrix. Als Eigenschaften von Eigenvektoren ist festzuhalten, dass beispielsweise für eine $Kov^{(n\times n)}$ Matrix, genau n Eigenvektoren existieren, welche orthogonal zueinander sind. Der Eigenvektor mit den größten Eigenwerten stellt nun die erste *Principle Component* dar. Stellt man sich die Frage, weshalb

PCA sinnvoll einzusetzen ist, so findet sich darauf folgende Antwort: Zum Einen hilft es bei der Mustererkennung also auch um Gemeinsamkeiten und Unterschiede in den Datenpunkten aufzuzeigen. Zum Anderen wird es schwierig bei mehr als 3-dimensionalen Daten, diese graphisch letztlich zu veranschaulichen. Nachdem bestimmte Muster erkannt und die Anzahl der Dimensionen reduziert wurden, bleiben dennoch viele Informationen erhalten. [60]

Nicht lineare Projektionen der Daten

Neben der Möglichkeit, Daten von einem zwei dimensionalen Raum auf eine einzige Dimension zu vermindern, dank der Verwendung von linearer Regressionsmethoden, gibt es noch andere Techniken dies zu tun. Als Beispiele hierfür sind *Principal Component Surfaces (PCSs)* oder *Principal Curves (PCu)* zu nennen. Im folgenden zeigen die Abbildungen 4.2 und 4.3 jeweils eine 2-dimensionale PCS. Diese können unter Umständen präzisere Ergebnisse zu Tage fördern als der Einsatz von Regressionsgeraden. [64]

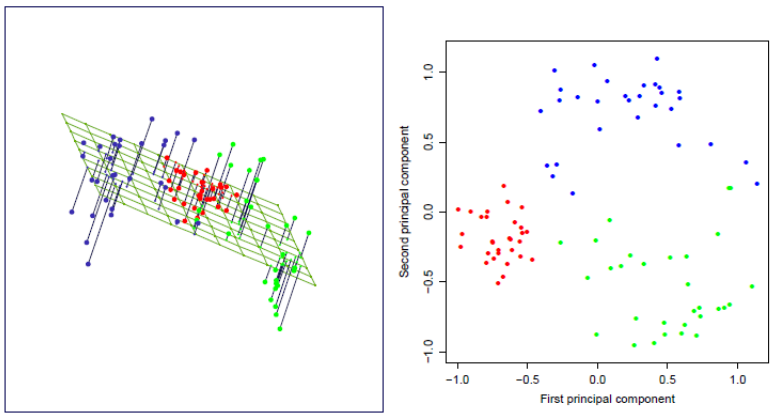

Abbildung 4.2: Illustration einer PCA Projektion auf eine Ebene, Quelle: [64]

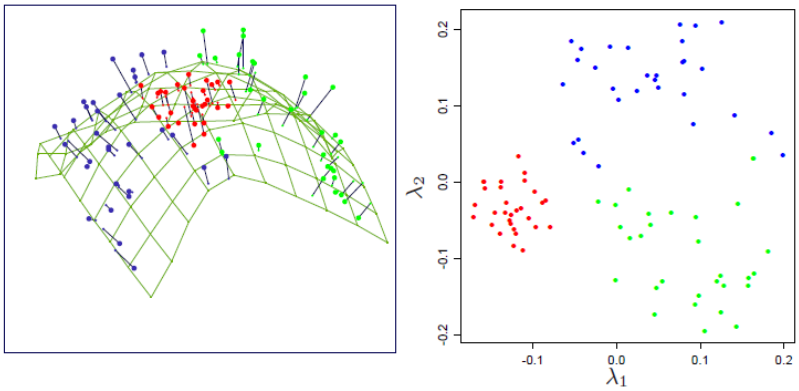

Abbildung 4.3: Illustration einer PCA Projektion auf eine gekrümmte Fläche, Quelle: [64]

Abbildung 4.2 zeigt auf der linken Seite eine zwei dimensionale PCS angepasst auf die *Half-Sphere* Daten, die von Hastie et al. verwendet wurden. Die rechte Seite von Abbildung 4.2 zeigt hingegen die Projektion der Datenpunkte auf die ersten beiden PCs. Dabei symbolisiert die Abszisse die erste PC und die Ordinate die zweite PC des vorliegenden Koordinatensystems. Dieses Verfahren zeigt eine klare Trennung der einzelnen *Cluster*. Abbildung 4.3 zeigt wiederum das Ergebnis einer PCS, allerdings wurde in diesem Fall ein *Scatterplot Smoother*

eingesetzt. Anstatt einer euklidischen Ebene wurde zudem ein *Grid* aus Rechtecken in Form einer Halbkugel verwendet. Dabei zeigt die rechte Seite wieder die Projektion der Daten auf einen Unterraum. Im Rahmen dieser Arbeit wird PCA verwendet um die Dimensionen der vorab berechneten Entropie Werte zu reduzieren, dies ist wiederum notwendig, um Klassifizierungsalgorithmen letzten Endes in weiteren Schritten anwenden zu können. [64]

Kapitel 5

Related Work

Im Bereich der Anomalie Erkennungsmethoden gibt es eine Vielzahl weiterer Ansätze. Da das Gebiet der künstlichen Intelligenz in der Informatik sehr breit gefächert ist, gibt es auch hier unterschiedliche Konzepte und Ansätze, um Muster zu erkennen und aus gesammelten Daten analytische Schlüsse zur Weiterverarbeitung von Informationen zu ziehen. Machine Learning Ansätze bilden neben statistischen Lernverfahren ein solches Teilgebiet künstlicher Intelligenz. Da Mutter Natur in vielen Bereichen der Technik beziehungsweise technologischer ingenieurs-wissenschaftlicher Problemstellungen als beispielhaftes Vorbild zu Lösungen dient, behelfen sich Wissenschaftler*innen im Bereich der Forschung über Lernverhalten auch hier einer Analogie aus der Biologie. Hierbei zu nennen sind beispielsweise die Adaptierung des Lotusblatteffekts oder der Grundsatz von Klettverschlüssen. So wird versucht, das menschliche Gehirn samt seiner Aktivitäten nachzubauen. Dies geschieht in Form von Neuronalen Netzen.

Dies bedeutet konkret aus technischer Sicht, dass eine Vielzahl an miteinander verbundenen berechenbaren Knoten ein Netzwerk bilden. Diese Knoten werden auch als Neuronen bezeichnet. Diese Neuronen sind in einer bestimmten Anzahl an *Layern* platziert. Allgemeiner formuliert bedeutet dies für die Architektur eines Neuronalen Netzes von einer *Highlevel* Perspektive, dass es einen *Input Layer*, einen oder mehrere *Hidden Layer* und einen *Output Layer* gibt. Knoten, die sich auf dem gleichen *Layer* befinden funktionieren in paralleler Mannier. Jedes Neuron hingegen berechnet seine zugehörige *Output* Aktivierung, welche beispielsweise eine Sigmoid oder Tangenshyperbolikus Funktion, also nicht affin lineare Funktionen sein können, anhand der Summe seines *Inputs*. Der daraus generierte Aktivierungswert dient als Berechnungsgrundlage für weitere Neuronen auf nachfolgenden *Layern*. Wichtig zu erwähnen ist in diesem Zusammenhang, dass alle Neuronen eines bestimmten *Layers* mit all den Neuronen der vorangegangenen *Layern* verbunden sind. Um nun das Netzwerk zu befähigen selbstständig zu Erlernen, werden den einzelnen Neuronen Gewichtungen zugewiesen. Diese werden innerhalb der Trainingsphase errechnet. Ein bestimmtes *Input* Muster, welches dem ersten *Layer* übergeben wird, wird durch das gesamte Neuronale Netz weitergegeben. Eine detailiertere Beschreibung der schematischen Struktur des dort angewandten Klassifizierungsmodells sowie des eingesetzten *Deep Autoencoders* ist in L. Reuter et al. [65] zu finden. Dort ist auch der Algorithmus zur Bestimmung einer *Threshold-basierten* Modifizierung der Daten ausführlich erläutert. In der Domäne der intelligenten Energienetze argumentieren L. Reuter et al. in [65] unter anderem über die Möglichkeiten Anomalien anhand von Neuronalen Netzen zu detektieren. Neuronale Netze sind demnach bereits weit verbreitet sowohl beim Identifizieren als auch beim Klassifizieren von *Cyber* Attacken in Computernetzwerken. Insbesondere der Einsatz solcher Technologien für die zuverlässige Erkennung von Angriffen auf

die Infrastruktur von Energieanbietern ist von großem Interesse seitens von Sicherheitsforschern auf diesem Fachgebiet. Die Besonderheit ihrer Idee besteht in der Kombination eines *Classifiers* mit einem *Autoencoder*, welche auf Neuronalen Netzen basieren, um einerseits eine hohe Detektionsrate zu erzielen, aber auch gleichzeitig mit dem Anspruch einer geringeren Fehlerrate. Auch bei diesen Experimenten stehen die Daten eines SDNs im Mittelpunkt der Untersuchungen. Aufgrund der zunehmenden Digitalisierung, die auch vor den technischen Komponenten der Energieanbieter nicht halt macht, bedarf es bei der Anbindung dieser Einheiten im Energienetz an einer Vielzahl von fortgeschrittenen Informations-technologischen Kommunikationsstrukturen. [65]

Zusätzlich gibt es noch weitere bestehende Forschungsbemühnugen wie beispielsweise von Muna et al., welche die Modellierung von *Deep Learning* Techniken als Versuch in industriellen Netzwerken zeigen, um schadhaftes Verhalten aufzuspüren. Die Wissenschaftler vertrauten auf einen Autoencoder innerhalb eines ICS um auch hier Anomalien zu erkennen. Im Gegensatz zu Reuter et al., verwendeten diese allerdings ein Deep Feed Neural Network, um den Klassifizierungsprozess zu ermöglichen. Wenn nun bisher unbekannte Attacken auftreten kann ihr System von einer Ressourcen effektiven Lernphase profitieren. Aufgrund der Tatsache, dass andere Machine Learning Szenarien in hohen False Positive Raten endeten, wurde dieser Ansatz von Muna et al. gewählt, um mit ungelabelten Daten umgehen zu können [66]. Die Forscher validierten ihr Model allerdings mit einem anderen Datenset als Reuter et al. [65]. Vielmehr bestand ihr Datenset aus einer Kombination des NSL-KDD [67] und UNSW-NB15 Datensets [68]. Dadurch erzielten sie eine Erkennungsrate von 99 Prozent [66].

Kapitel 6

Zusammenfassung und Ausblick

6.1 Zusammenfassung

Im Rahmen dieser Bachelorarbeit 1 wurden Möglichkeiten, die Anomalie-Erkennungsalgorithmen unter Verwendung von Flow Informationen in SDN bieten, untersucht. Zunächst wurde ein Überblick über den architektonischen Aufbau von SDN verschafft. Darauffolgend wurden Steuerungsmechanismen in ICS unter die Lupe genommen. Solche technische Komponenten wie SCADA Systeme können jedoch nicht mehr als isolierte Einheiten angesehen werden, sondern sind Bestandteil zunehmend vernetzter Infrastrukturen. Mit steigendem Grad an dezentralen Energieversorgern im gesamten Energienetz nimmt auch die Anzahl an angreifbaren RTUs oder PCLs wesentlich zu. Im Rahmen von APTs werden *Substations* von Energienetzen immer öfter auserkoren, um von *Cyber-Security* Experten als Angriffsziele genutzt zu werden. Diese Form der Industriespionage und möglicher daraus resultierender Sabotageakte führt dazu, dass Betreiber von Energienetzen auf die Kunst des *Machine Learning* vertrauen, um mittels statistischer Verfahren, Anomalien im Netzwerkverkehr festzuhalten. In Kapitel 1 wurde zu Beginn die Zielsetzung und Methodik dieser wissenschaftlichen Arbeit vorgestellt. Die weiteren Kapitel 2 bis 4 zeigten, wie theoretisch in Bezug auf SDN Informationen der Einsatz von Shannon Entropie Werten in Kombination mit linearen Regressionen in Form einer PCA dazu beitragen können, die Resilienz kritischer Infrastrukturen zu verbessern, indem *Cyber* Angriffe systematisch erkannt werden können. Danach wurde in Kapitel 5 aufgezeigt, welche anderen ML Techniken zur Detektion von Anomalien zum Einsatz kommen.

6.2 Ausblick

Nachdem die Entropie Werte aus den Datenpunkten berechnet und die Dimensionen der Features mittels PCA reduziert worden sind, ist es sinnvoll, Überlegungen zur Klassifizierung der Daten anzustellen. Hierbei kommen insbesondere *K-Nearest Neighbor* Methoden wie auch *K-means* und *Support Vector Machines* in Betracht. In diesem Kontext sind auch die Grenzen der einzelnen Techniken zu bedenken. Diese betreffen sowohl die Komplexität der Laufzeit der Berechnungen als auch wie die Ergebnisse *Machine Learning* Algorithmen zu interpretieren sind. Zur Veranschaulichung der Resultate kann der Einsatz von *Confusion-Matrizen* nützlich sein. Besondere Aufmerksamkeit sollte man in diesem Zusammenhang der Interpretation einzelner ML Metriken, wie zum Beispiel der *Accuracy* und dem *F1-Score* schenken. Des Weiteren kann es besonders wichtig sein, Maßnahmen zu ergreifen, die sich mit der Verteilung der Daten innerhalb des Datensets befassen, um etwaigen ungleichen Verhältnissen in den

Verteilungen der Datenpunkte entgegenzuwirken. Außerdem wird von Bedeutung sein, welche Schritte in einer Infrastruktur notwendig sind, nachdem eine Anomalie als Attacke erkannt worden ist.

Anhang A

Anhang

Abbildungsverzeichnis

Literaturverzeichnis

[1] P. Goransson. *Software defined networks : a comprehensive approach*. Morgan Kaufmann, Cambridge, MA, 2017. 3, 4

[2] D. Kreutz, F. M. V. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, and S. Uhlig. Software-defined networking: A comprehensive survey. *Proceedings of the IEEE*, 103(1):14–76, Jan 2015. 3

[3] Julian Magin, Lenhard Reuter, Oliver Jung, and Paul Smith. Anomaly detection in smart grids based on software defined networks. In *Proceedings of the 8th International Conference on Smart Cities and Green ICT Systems - Volume 1: SMARTGREENS,*, pages 157–164. INSTICC, SciTePress, 2019. 4, 24

[4] N. Feamster, J. Rexford, and E. Zegura. The road to sdn: An intellectual history of programmable networks. *SIGCOMM Comput. Commun. Rev.*, 44(2):87–98, April 2014. 3, 4

[5] M. Ramasamy and S. Pawar. Pareto-optimal multi-controller placement in software defined network. In *2018 3rd International Conference for Convergence in Technology (I2CT)*, pages 1–7, April 2018. 4

[6] M. Casado, M. J. Freedman, J. Pettit, J. Luo, N. McKeown, and S. Shenker. Ethane: Taking control of the enterprise. In *Proceedings of the 2007 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, SIGCOMM '07, pages 1–12, New York, NY, USA, 2007. ACM. 4

[7] D. B. Hoang and M. Pham. On software-defined networking and the design of sdn controllers. In *2015 6th International Conference on the Network of the Future (NOF)*, pages 1–3, Sept 2015. 4

[8] X. Zhang, K. Wei, L. Guo, W. Hou, and J. Wu. Sdn-based resilience solutions for smart grids. In *2016 International Conference on Software Networking (ICSN)*, pages 1–5, May 2016.

[9] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner. Openflow: Enabling innovation in campus networks. *SIGCOMM Comput. Commun. Rev.*, 38(2):69–74, March 2008. 4

[10] Y. Ozcevik, M. Erel, and B. Canberk. Noc based banyan openflow switch for software defined networks. In *2015 23nd Signal Processing and Communications Applications Conference (SIU)*, pages 1433–1436, May 2015. 4

[11] M. Brandstetter, A. Schirrer, M. Miletic, S. Henein, M. Kozek, and F. Kupzog. Hierarchical predictive load control in smart grids. *IEEE Transactions on Smart Grid*, 8(1):190–199, Jan 2017. 5, 6

[12] P. Samadi, H. Mohsenian-Rad, R. Schober, and V. W. S. Wong. Advanced demand side management for the future smart grid using mechanism design. *IEEE Transactions on Smart Grid*, 3(3):1170–1180, Sept 2012. 5

[13] K. Moslehi and R. Kumar. A reliability perspective of the smart grid. *IEEE Transactions on Smart Grid*, 1(1):57–64, June 2010. 5

[14] Jacob Theilgaard Madsen, Mislav Findrik, Domagoj Drenjanac, and Hans-Peter Schwefel. Investigating wind farm control over different communication network technologies. In *Energy Informatics*, pages 129–140. Springer International Publishing, 2015. 6

[15] M. Findrik, R. Pedersen, E. Hasenleithner, C. Sloth, and H. Schwefel. Test-bed assessment of communication technologies for a power-balancing controller. In *2016 IEEE International Energy Conference (ENERGYCON)*, pages 1–6, April 2016. 6

[16] O. Valgaev, F. Kupzog, and H. Schmeck. Designing k-nearest neighbors model for low voltage load forecasting. In *2017 IEEE Power Energy Society General Meeting*, pages 1–5, July 2017. 6

[17] A. Lugmaier, H. Fechner, W. Pruggler, and F. Kupzog. National technology platform - smart grids austria. In *CIRED 2009 - 20th International Conference and Exhibition on Electricity Distribution - Part 1*, pages 1–3, June 2009. 6

[18] X. Yao and J. Liu. The potential of economic growth and technology advancement in the brics. In *2011 International Conference on Machine Learning and Cybernetics*, volume 3, pages 1067–1071, July 2011. 6

[19] M. Findrik, J. Groenbaek, and R. L. Olsen. Scheduling data access in smart grid networks utilizing context information. In *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 302–307, Nov 2014. 6

[20] G. Lauss, F. Andrén, M. Stifter, R. Bründlinger, T. Strasser, K. Knöbl, and H. Fechner. Smart grid research infrastructures in austria: Examples of available laboratories and their possibilities. In *2015 IEEE 13th International Conference on Industrial Informatics (INDIN)*, pages 1539–1545, July 2015. 6

[21] R. Pedersen, C. Sloth, G. B. Andresen, and R. Wisniewski. Disc: A simulation framework for distribution system voltage control. In *2015 European Control Conference (ECC)*, pages 1056–1063, July 2015. 6

[22] E. Pleijsier. Towards anomaly detection in scada networks using connection patterns. 2013. 7, 8

[23] Rafael Ramos Regis Barbosa. Anomaly detection in scada systems a network based approach. 2014. 7

[24] Iñaki Garitano, Roberto Uribeetxeberria, and Urko Zurutuza. A review of scada anomaly detection systems. In *SOCO*, 2011. 7

[25] Jesse G. Wales. Analysis of a scada system anomaly detection model based on information entropy. 2014. 7

[26] Niv Goldenberg and Avishai Wool. Accurate modeling of modbus/tcp for intrusion detection in scada systems. *IJCIP*, 6:63–75, 2013. 7

[27] M. Findrik, P. Smith, J. H. Kazmi, M. Faschang, and F. Kupzog. Towards secure and resilient networked power distribution grids: Process and tool adoption. In *2016 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 435–440, Nov 2016. 7

[28] M. N. Noran and Z. Shukri. Adaptive breaker failure protection scheme for double busbar substation using iec 61850 goose message communication. In *2015 International Symposium on Technology Management and Emerging Technologies (ISTMET)*, pages 207–212, Aug 2015. 7

[29] Zhu Yongli, Wang Dewen, Wang Yan, and Zhao Wenqing. Study on interoperable exchange of iec 61850 data model. In *2009 4th IEEE Conference on Industrial Electronics and Applications*, pages 2724–2728, May 2009. 7

[30] C. Yang, J. Xu, and V. Vyatkin. Towards implementation of iec 61850 goose messaging in event-driven iec 61499 environment. In *Proceedings of the 2014 IEEE Emerging Technology and Factory Automation (ETFA)*, pages 1–4, Sept 2014. 7

[31] C. Brunner. Iec 61850 for power system communication. In *2008 IEEE/PES Transmission and Distribution Conference and Exposition*, pages 1–6, April 2008. 7

[32] P. Jafary, S. Repo, and H. Koivisto. Secure communication of smart metering data in the smart grid secondary substation. In *2015 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*, pages 1–6, Nov 2015. 8

[33] A. Shahzad, S. Musa, A. Aborujilah, and M. Irfan. Secure cryptography testbed implementation for scada protocols security. In *2013 International Conference on Advanced Computer Science Applications and Technologies*, pages 315–320, Dec 2013. 8

[34] P. Gama, A. Gomes Varela, and W. Freudenberg. Iec 60870-5-104 as a driver to evolution of substation and distribution automation at edp. In *CIRED 2009 - 20th International Conference and Exhibition on Electricity Distribution - Part 1*, pages 1–4, June 2009. 8

[35] Peter Maynard, Kieran McLaughlin, and Berthold Haberler. Towards understanding man-in-the-middle attacks on IEC 60870-5-104 SCADA networks. In *2nd International Symposium for ICS & SCADA Cyber Security Research 2014*. BCS Learning & Development, sep 2014. 8

[36] E. Joelianto and Hosana. Performance of an industrial data communication protocol on ethernet network. In *2008 5th IFIP International Conference on Wireless and Optical Communications Networks (WOCN '08)*, pages 1–5, May 2008. 8

[37] I. Ahmad, S. Namal, M. Ylianttila, and A. Gurtov. Security in software defined networks: A survey. *IEEE Communications Surveys Tutorials*, 17(4):2317–2346, Fourthquarter 2015. 9

[38] Luca Didaci, Giorgio Giacinto, and Fabio Roli. Ensemble learning for intrusion detection in computer networks. 08 2002. 9

[39] J. McHugh, A. Christie, and J. Allen. Defending yourself: The role of intrusion detection systems. *IEEE Software*, 17(5):42–51, Sept 2000. 9

[40] Paul E. Proctor. *Practical Intrusion Detection Handbook*. Prentice Hall PTR, Upper Saddle River, NJ, USA, 1st edition, 2000. 9

[41] A. Gouglidis, S. König, B. Green, K. Rossegger, and D. Hutchison. *Protecting Water Utility Networks from Advanced Persistent Threats: A Case Study.* In: Rass S., Schauer S, (eds) Game Theory for Security and Risk Management. Static & Dynamic Game Theory, 2018. 9

[42] Félix Iglesias and Tanja Zseby. Analysis of network traffic features for anomaly detection. *Machine Learning*, 101(1-3):59–84, dec 2014. 9, 10

[43] Narmeen Zakaria Bawany, Jawwad A. Shamsi, and Khaled Salah. DDoS attack detection and mitigation using SDN: Methods, practices, and solutions. *Arabian Journal for Science and Engineering*, 42(2):425–441, February 2017. 10

[44] Daocheng Hong, Deshan Zhao, and Yanchun Zhang. The entropy and PCA based anomaly prediction in data streams. *Procedia Computer Science*, 96:139–146, 2016. 10, 13, 14

[45] *Introduction to Statistical Machine Learning.* Elsevier, 2016. 11

[46] Geoff Dougherty. Unsupervised learning. In *Pattern Recognition and Classification*, pages 143–155. Springer New York, September 2012. 11

[47] Shashi Shekhar, Hui Xiong, and Xun Zhou, editors. *Encyclopedia of GIS.* Springer International Publishing, 2017. 11

[48] Tabalis. *Information Security Analytics.* Elsevier, 2015. 12

[49] Shanta Rangaswamy Gangadhar Shobha. *Handbook of Statistics, Computational Analysis and Understanding of Natural Languages: Principles, Methods and Applications.* Elsevier, 1st edition, 2018. 12

[50] Philipp Schunker. *Pfadsuche basierend auf Reinforcement Learning und Monte-Carlo Tree Search.* Bachelorarbeit 1 an der Fachhochschule Campus Wien, 17.01.2019. 12

[51] Martijn van Otterlo and Marco Wiering. Reinforcement learning and markov decision processes. In *Adaptation, Learning, and Optimization*, pages 3–42. Springer Berlin Heidelberg, 2012. 12

[52] Shaked Zychlinski. The complete reinforcement learning dictionary. 12

[53] Manuel Lopez-Martin, Belen Carro, and Antonio Sanchez-Esguevillas. Application of deep reinforcement learning to intrusion detection for supervised problems. *Expert Systems with Applications*, 141:112963, March 2020. 12

[54] E. T. Jaynes. Gibbs vs b̈oltzmann ëntropies. *American Journal of Physics*, 33(5):391–398, May 1965. 13

[55] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27(3):379–423, July 1948. 13

[56] Wenke Lee and Dong Xiang. Information-theoretic measures for anomaly detection. In *Proceedings 2001 IEEE Symposium on Security and Privacy. S P 2001*, pages 130–143, May 2001. 13

[57] Przemyslaw Berezinski, Bartosz Jasiul, and Marcin Szpyrka. An entropy-based network anomaly detection method. *Entropy*, 17:2367–2408, 2015. 13

[58] C. Bishop. *Pattern recognition and machine learning.* Springer, New York, 2006. 14, 16, 24

[59] Jake VanderPlas. *Python Data Science Handbook: Essential Tools for Working with Data.* O'Reilly Media, Inc., 1st edition, 2016. 14

[60] Lindsay I Smith. A tutorial on principal components analysis. 14, 15, 16, 17

[61] Gerald Teschl and Susanne Teschl. *Mathematik für Informatiker, Band 2: Analysis und Statisitk.* 14, 15

[62] The entropy and pca based anomaly prediction in data streams. 15

[63] Jolliffe I. *Principle Component Analysis 2nd Edition, Springer.* 15

[64] Robert Tibdhirani Trevor Hastie and Jerome Friedman. *The elements of Statistical Learning, Data Mining, Inference, and Prediction.* 17, 18, 24

[65] Lenhard Reuter, Oliver Jung, and Julian Magin. Neural network based anomaly detection for SCADA systems. In *2020 23rd Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN) (ICIN 2020)*, Paris, France, February 2020. 19, 20

[66] Muna Al-Hawawreh, Nour Moustafa, and Elena Sitnikova. Identification of malicious activities in industrial internet of things based on deep learning models. *Journal of Information Security and Applications*, 41, 05 2018. 20

[67] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani. A detailed analysis of the kdd cup 99 data set. In *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, pages 1–6, July 2009. 20

[68] Nour Moustafa and Jill Slay. Unsw-nb15: a comprehensive data set for network intrusion detection systems (unsw-nb15 network data set). 11 2015. 20