

IT Innovationen

Band 31

Juni 2023



Grußwort der Fakultät

Liebe Leserinnen und Leser,

als Hochschule für angewandte Wissenschaften befinden wir uns an der Schnittstelle zwischen Forschung und der praktischen Anwendung in Wirtschaft und Industrie. Diese Schnittstelle wird besonders im Rahmen der Abschlussarbeiten der Bachelor- und Master-Absolventinnen und Absolventen gelebt. Dabei werden sowohl bestehende und etablierte Themenbereiche um neue Ansätze und Problemstellungen erweitert als auch neue Problemfelder erschlossen.

Allen Arbeiten ist gemein, dass ihr Kern eine neue Erkenntnis und Innovation darstellt. Eine Abschlussarbeit mutet zu Beginn oft an wie eine Expedition ins Ungewisse. Natürlich bereitet man sich gut vor! Man prüft die Ausrüstung. Statt Eispickel und wetterfester Kleidung erarbeitet man sich die Grundlagen eines Themas und vergegenwärtigt sich nochmals die im Studium erworbenen Kenntnisse und Fähigkeiten. Man plant die Route und bemisst die zur Verfügung stehende Zeit. Die Arbeit bekommt einen roten Faden, der wie eine Reiseroute die Strecke vorzeichnen soll. Statt Proviant zu packen, recherchiert man nach verwandten Arbeiten, von deren Ergebnissen man zehren und auf denen man aufbauen kann. Zu guter Letzt stellt man sich das Expeditionsteam zusammen: Firmenbetreuer:innen, Professor:innen und Freund:innen für die Unterstützung beim Gelesen und für den Rückhalt, wenn man nicht mehr weiterweiß.

Und dann bricht man auf. Und plötzlich ist alles anders. Mit jedem Schritt verändert sich die Route, die auf der Karte so klar aussah. Neue Probleme tauchen auf. Andere Probleme verschwinden plötzlich und erlauben keine eigene Lösung mehr. Auf dem richtigen Weg zu bleiben ist harte Arbeit und manchmal stellt sich erneut die Frage, welcher Weg wohl der Richtige sein mag. Doch sind all diese Unwägbarkeiten bezwungen, steht man am Ziel: einem Stückchen Wissen oder Innovation, dass es zuvor nicht gab. Jeder einzelne Beitrag dieses Bandes trägt diesen Schatz in sich. Gleichzeitig zeigt die Breite der Artikel die Vielfalt der Herausforderungen und Lösungen in der Informatik.

Machen Sie sich ein Bild dieser Expeditionen in diesem Band der IT-Innovationen mit den vorliegenden Kurzbeschreibungen der Abschlussarbeiten des Abschlussjahrgangs des Sommersemesters 2023.

Viel Spaß auf der Reise wünscht Ihnen

A handwritten signature in black ink that reads "Tobias Heer".

Ihr Prof. Dr. Tobias Heer

IMPRESSUM

ERSCHEINUNGSORT

73732 Esslingen am Neckar

HERAUSGEBER

Prof. Jürgen Nonnast
Dekan der Fakultät Informatik und Informationstechnik
der Hochschule Esslingen - University of Applied Sciences

REDAKTIONSANSCHRIFT

Hochschule Esslingen - University of Applied Sciences
Fakultät Informatik und Informationstechnik
Flandernstraße 101
73732 Esslingen am Neckar

Telefon +49(0)711.397-4211
Telefax +49(0)711.397-4214
E-Mail it@hs-esslingen.de
Website www.hs-esslingen.de/it

REDAKTION, LAYOUT UND DESIGN

Prof. Dr.-Ing. Reinhard Schmidt
Hochschule Esslingen - University of Applied Sciences
Fakultät Informatik und Informationstechnik
Flandernstraße 101
73732 Esslingen am Neckar

SATZ, ANZEIGEN und VERLAG

Dipl.-Inf.(FH) Rolf Gassner
Hochschule Esslingen - University of Applied Sciences
Fakultät Informatik und Informationstechnik
Flandernstraße 101
73732 Esslingen am Neckar

ERSCHEINUNGSWEISE

Einmal pro Semester, jeweils Januar und Juni

ISSN 1869-6457

Muhammed Adel	Automatisierung von Dynamischer Malware Instrumentierung	6
Miralem Ademi	Analyse der Nvidia VPI im Vergleich zu generalisierten Image Processing APIs wie OpenCV	9
Raphael Andree	Implementierung des V2X Security Layers für ein POSIX-System	11
Emre Arslan	Standardisierter Projektablaufplan für SAP Applikation Upgrades	14
Betiel Assefaw	Modellierung der approximativen Bestimmung von Nutzenempfinden für serviceorientierte Geschäftsmodelle	17
Jan Lukas Autenrieth	Konzeptentwicklung für die Einführung und Veränderung von Prozessen auf Basis des ITIL Change Managements und des Change of Mind-Ansatzes	20
Rozin Babat	The opportunity of creating a Data Hub for the Mercedes-Benz AG in order to optimize sales steering.	23
Thibaut Becher	Re-Engineering einer Software für Smartphone-Reparatur Management: Usability, Performance und Security	26
Maxim Becht	Semantic Visual SLAM in Outdoor Environments	29
Constantin Blessing	Development, Simulation and Validation Environment for Autonomous Driving Algorithms based on a ROS Architecture	32
Marc Brugger	Robuste Features für die visuelle Lokalisierung mobiler Robotik im Outdoor-Bereich	34
Jeremy Camier	Konzeption und Implementierung einer webbasierten Peer-To-Peer-Infrastruktur für kooperative Virtual Reality Anwendungen	36
Irem Nur Cankaya	Wohin führt uns die künstliche Intelligenz? Eine Analyse der Auswirkung von KI auf Wirtschaft, Gesellschaft und Ethik	39
Lars-Hendrik Crone	Konzeptionelle Entwicklung einer One Service Data Plattform im Bereich der Messanalyse mit Big und Small Data im industriellen Umfeld	42
Dimitar Dimov	Untersuchung von Containern zur Entwicklung einer Build-Umgebung für Embedded Systems	46
Christian Duerr	Data Analytic Tools im Controlling am Beispiel der Prognose	48
Nick Eschen	Konzeptentwicklung und prototypische Entwicklung eines First-Level-Support-Chatbots auf Basis von künstlicher Intelligenz	51
Julius Freudenberger	Vergleich und Evaluation von GraphQL-Implementierungen LowCode Framework gegenüber traditioneller Entwicklung	54
Thimo Froehner	Einsatz von Process Mining zur Optimierung von Geschäftsprozessen	57

Frank Fuchs	Entwicklung von Echtzeit-Chatkomponenten mit den Web-Technologien SignalR und Blazor am Beispiel eines Azure-Cloud basierten 2-Personen Spieles	60
Ismet Gezer	Entwicklung einer Softwarelösung zur Verifizierung und Lizenzierung eines Ladesäulen-Prüfgeräts	64
Maximilian Gittel	Analyse der Anforderungen für den Einsatz eines Systems zur Angriffserkennung für kritische Infrastrukturen	67
Luisa Glass	Migration und Optimierung einer Schwingungsanalyse-Software: Implementierung in Python und Anwendung zur Untersuchung von Abgasanlagen	70
Julian Groshaupt	Verwendung von Software-Bill-of-Materials zur automatischen netzwerkweiten Erfassung von Softwarekomponenten	73
Raphael Gschwind	Data Insights im Connected Car Umfeld: Verbesserung der Entscheidungsfindung durch den Business Intelligence Prozess mittels manuellem Testing	76
Rahim Haidari	Untersuchungen zum Bidirektionalen Laden von batterieelektrischen Fahrzeugen mit ISO15118-20 und Aufbau eines Hardware- und Software-in-the-loop Testers	79
Sami Hazzouri	Entwicklung einer Anwendung zur Erkennung und Darstellung von Sensoren und deren Wirkbereichen in Augmented Reality	84
Johannes Hofmayer	Evaluation und Vergleich verschiedener Methoden zur prädiktiven Instandhaltung von CNC-Maschinen anhand von zeitbasierten Maschinendaten	88
Tolgahan Kandemir	Evaluierung von Einsatzmöglichkeiten für Post-Quanten-Kryptografie-Algorithmen im Automobilumfeld	91
Leon David Kiefer	Disparity Estimation with Stereo Event-Based Cameras and Sparse Coding	93
Florian Korotschenko	Design und Implementierung einer Synchronisationselektronik für digitale Messgeräte im Pikosekunden-Bereich	97
Markus Lang	Erweiterung eines HF Performance Messplatzes mit automatischer Erzeugung normkonformer Reports	101
Steffen Lang	Eine vergleichende Evaluierung von Flutter und React als Alternative zu Qt für die Programmierung der Bedienoberfläche von professionellen Kaffeemaschinen	104
Andreas Lautner	Finding States in Black-Box Fuzzing	107
Dennis Lehmann	Design und Implementierung einer Datenpipeline-Architektur als Grundlage für IIoT-Anwendungen zur Nutzung industrieller Prozessdaten	109

Simon Leibfritz	Analyse zeitbasierter Seitenkanäle und deren Einfluss auf die Sicherheit von Automotive-Software	112
Philipp Lichtmanegger	Analyse des Anforderungsmanagements mit Entwicklung von Lösungen zur Bewertung und Steuerung von Produktprojekten	115
Nico Linder	Vergleich der Webframeworks Spring und Phoenix anhand der Implementierung einer Cloud-nativen Messenger-Anwendung	118
Svenja Luik	Hardware-Tracing von Standard Utility Modules (GM). Eine Laufzeitanalyse auf einem Infineon Aurix System.	121
Christof Meyer	Methodik zur Bewertung und Verwaltung der Datenqualität in einer Powertrain Solution Pipeline	125
Nikolai Miguel Mora Lozano	Integration von PaaS und SaaS zur Verwaltung sowie Verknüpfung verschiedener IoT-Edge-Geräte und Services für intelligente Analyseprozesse	127
Aaron Mueller	Konzeption und Implementierung eines chinesischen V2X-Stacks für ein POSIX-Betriebssystem	130
Robin Mueller	Workflow zur Frühwarnerkennung von gesundheitlichen Veränderungen bei Patienten auf Basis von Web-Technologien	133
Silas Mueller	Virtualisierung von GPUs: Einbindung in SLURM, Accounting und Energieeffizienz	135
Brikena Mushkolaj	Entwicklung und prototypische Implementierung einer Cross-Media-Marketingstrategie für ein IT-Software-Systemhaus	137
Maximilian Nikitsch	Darstellung und Bewertung der Trends der Arbeitswelt	140
Fawad Rasuly	Integration einer datenbasierten Schnittstelle zwischen SAP SuccessFactors und Payroll-Systemen in die bestehende IT-Landschaft der Alfred Kärcher SE & Co. KG - Eine Potenzialanalyse und Handlungsempfehlung	142
Cornelius Rost	Untersuchung einer Schnittstelle zwischen PyScript und Angular	145
Moritz Ruder	Unified UX - Analyse und Optimierung der produktübergreifenden User Experience am Beispiel des Produktportfolios der USU Software AG	148
Julian Saile	Konzeption und Umsetzung von CI/CD-Praktiken im Kontext einer webbasierten Anwendung	152
Artur Schnell	Konzeption und Entwicklung einer Smart Office Management Anwendung für Desk Sharing und Energieeffizienz mit Anbindung an MS-Teams	155
Kevin Schramm	Low-Code Plattformen: Vergleich am Beispiel Monitoring von Code Compliance	159

Katharina Schulz	Evaluierung der IT-Sicherheitsmaßnahmen einer Software-as-a-Service-Lösung zur Umsetzung des Lieferkettensorgfaltspflichtenschutzgesetzes	162
Thomas Sebast	Künstliche Intelligenz in der Musik	165
Arasch Seddiqi	Die Einsatzmöglichkeiten von ChatGPT im Unternehmen und die damit verbundenen Chancen und Risiken	168
Felipe Sibila Valtierre	Exploring and Application of Large Language Models for Rule-based Tasks using the Example of Chess Playing	171
Manuel Staub	UI und UX Design eines Alpaka Herdmanagementsystems	175
Maximilian Stock	Relative Positionsbestimmung von Agrarrobotern mit visueller Odometrie	178
Tim Straetz	Untersuchung und Entwicklung verschiedener Prediction Heads in Convolutional Neural Networks für Fahrzeugerkennung	181
Nick Suess	Evaluierung eines RISC-V Microcontrollers	184
Murat Tekin	Entwicklung eines künstlichen neuronalen Netzes für das Erkennen von Rissen auf Fotos von Mikrochips	187
Panagiotis Theofanidis	V2X Security: Analysis and Implementation	190
Mladen Tramosljika	Konzeption und Realisierung eines vollautomatisierten, hochpräzisen Sensor-Kalibrator mit einem Datenbanksystem	193
Markus Treitz	Einsatz von Bildverarbeitung zur Qualitätskontrolle von Fertigteilebehältern	196
Konstantinos Tsolakidis	Evaluierung von Frameworks zur Erstellung moderner Bots im Web 2.0	198
Joshua Uhl	Leistungsfähigkeit von neuronalen Netzen zur Verbesserung von Bilddaten	201
Viktoria Wagner	Analyse und Konzeptionierung eines SAP-Addons zur automatisierten Verarbeitung eingehender Zahlungsavisen	203
Jonathan Weyl	Konzeption und prototypische Umsetzung einer Migrationsstrategie zur Überführung einer Eclipse RCP-Anwendung in eine webbasierte Lösung	206
Jan von Scholley	Konzeption und Realisierung einer verlustfreien, zeitsynchronen Speicherung von Messdaten mittels moderner Echtzeitmechanismen	211

Automatisierung von Dynamischer Malware Instrumentierung

Muhammed Adel

Reinhard Schmidt

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma adesso SE, Stuttgart Vaihingen

Einleitung

Die rasante technologische Entwicklung und ihre Integration in unseren Alltag haben zu zahlreichen Veränderungen geführt. Heutzutage spielen Computer und andere intelligente Geräte eine immens wichtige Rolle in unserem Leben. Sie unterstützen uns bei der Sicherung und dem Abruf von Daten, erleichtern unsere Aufgaben und ermöglichen die Automatisierung mühseliger Prozesse. Dabei arbeiten sie präziser und schneller als es ein Mensch jemals könnte.

Mit dem technologischen Fortschritt steigt auch der Bedarf an Datensicherung und -verarbeitung. Während früher eine Diskette mit wenigen Kilobytes Speicherplatz ausreichend war, sind heute selbst 100 GB Cloud-Speicher häufig nicht genügend. Wertvolle und persönliche Daten werden nun online und auf verschiedenen Geräten gespeichert, sodass sie von überall auf der Welt zugänglich sind. Idealerweise sind alle Geräte intelligent miteinander verknüpft und können somit von jedem Gerät aus erreicht werden. Durch den Anstieg der im Internet zirkulierenden Daten wächst jedoch auch das Risiko von Datenschutzverletzungen.

In den letzten Jahren wurde ein stetiger Anstieg von Datenschutzverletzungen und Datendiebstählen beobachtet. Eine Datenschutzverletzung liegt vor, wenn eine unbefugte Person oder Organisation Zugang zu Daten erhält, die sie nicht haben sollte. Um die Ausmaße der Datenschutzverletzungen und Diebstähle in Zahlen zu verdeutlichen, zeigt der 2023 SONICWALL CYBER THREAT REPORT, dass es im Jahr 2022 allein 5,5 Milliarden Malware-Angriffe (+2%), 6,3 Billionen (+19%) Einbruchversuche, 139,3 Millionen (+43%) Cryptojacking-Angriffe und 112,3 Millionen (+87%) neu entdeckte IoT-Malware gab, im Vergleich zum Vorjahr. [6] Allein in Deutschland schätzt man den jährlichen gesamten Schaden der durch Cyber-Attacks entsteht auf ca. 203€ Milliarden. [2] Vor diesem Hintergrund gewinnt das Thema Cybersicherheit immer mehr an Bedeutung für Unternehmen und Regierungen, aber auch für Privatpersonen.

Ziel dieser Arbeit

Das primäre Ziel dieser Arbeit ist die Entwicklung einer voll funktionsfähigen Testumgebung, die die automatisierte dynamische Malware-Analyse unterstützt. Am Ende dieses Projekts sollte es jedem Anwender möglich sein, mit minimalen Aufwand, diese Testumgebung lokal einzurichten und Malware automatisch zu testen und zu analysieren.

Die zugrundeliegende Hoffnung besteht darin, dass Forscher, die möglicherweise nicht über das erforderliche Wissen zur Automatisierung und Erstellung einer solchen Testumgebung verfügen, mittels unserer entwickelten Plattform Malware einfach testen können. Dadurch sollen sie befähigt werden, neue Techniken zur Entdeckung und Beseitigung von Malware zu entwickeln.

Malware

Malware, eine Kontraktion aus "malicious software" (böswartige Software), bezeichnet Programme, die entwickelt wurden, um Systeme oder Netzwerke zu schädigen. Sie stellt ein stetig wachsendes Problem dar, dem sich alle Internetnutzer gegenübersehen, da Cyber-Kriminelle zunehmend professioneller und innovativer werden. Das IT-Sicherheitsunternehmen Kaspersky berichtete 2021, dass sie täglich 380.000 neue Malware-Proben identifizierten. Nur ein Jahr später stieg diese Zahl auf 400.000 neue Malware-Proben pro Tag – ein Trend, der sich fortzusetzen scheint. [5]

Laut AV-Atlas gab es seit Anfang 2023 23,4 Millionen neue Malware-Exemplare, wodurch die Gesamtzahl der seit 1984 erfassten Malware auf über 1,254 Milliarden anstieg. [1]

Dynamische Analyse

Die dynamische Analyse zeichnet sich dadurch aus, dass sie den Inhalt der Software nicht direkt analysiert, sondern das Verhalten nach deren Ausführung

beobachtet. Daher wird sie auch als Verhaltensanalyse (engl. "behavioral analysis") bezeichnet. In der dynamischen Analyse wird die Malware in einer geschützten Umgebung, auch als Sandbox bezeichnet, ausgeführt und beobachtet. Es gibt viele verschiedene Aspekte, die man während dieser Beobachtung aufzeichnen kann, um die Malware zu klassifizieren. Zu den verschiedenen Techniken, die bei der Durchführung einer dynamischen Analyse eingesetzt werden können, gehören die Überwachung von Funktionsaufrufen, die Parameteranalyse von Funktionen, das Verfolgen des Informationsflusses, das Erfassen von Befehlsspuren und Autostart-Erweiterungspunkten, das Überprüfen von aufgerufenen API-Aufrufen sowie das Beobachten der Verwendung von verschiedenen Systemaufrufen. [4]

Umsetzung

Der erste Schritt unserer Arbeit bestand darin, eine sichere Testumgebung für die Malware-Analyse zu erstellen. Diese Umgebung wurde mithilfe von Oracle VirtualBox konstruiert und kann durch Automatisierungsskripte eigenständig aufgesetzt werden. Anschließend lag unser Fokus auf der Vorbereitung der virtuellen Maschinen zur Malware-Analyse. Aufgrund des inhärenten Risikos, das mit der Ausführung von Malware verbunden ist, wurden diverse Sicherheitsvorkehrungen getroffen, um zu gewährleisten, dass die Malware keinen Schaden verursacht. Ein wesentlicher und arbeitsintensiver Teil des Projekts befasste sich mit der Datenerfassung und -aufbereitung. Die Suche nach geeigneter Goodware stellte eine Herausforderung dar, war jedoch unerlässlich, um hochwertige Trainingsdaten für das maschinelle Lernmodell bereitzustellen. Dabei dienten die Windows Sysinternals als Hauptreferenz für Goodware. Die Beschaffung von Malware-Daten hingegen erwies sich als einfacher, da zahlreiche öffentlich zugängliche Datenbanken mit Malware-Proben zur Verfügung stehen. Nach der Durchführung der Malware- und Goodware-Analyse auf der Windows-Maschine war es notwendig,

die gewonnenen Daten zu sichern und aufzubereiten, um sie für die weitere Analyse bereitzustellen. Das System wurde so konzipiert, dass es nach jeder Ausführung von Good- oder Malware die Ergebnisse an einen FTP-Server in einem internen Netzwerk sendet, genau wie die Windows-Maschine, die für die Malware-Analyse verwendet wird.

Mit dem erfolgreichen Abschluss dieses Projekts haben wir eine automatisierte Testumgebung für die dynamische Malware-Analyse geschaffen. Sie ermöglicht es Forschern, mit einem einfachen System bis zu 30 Malware-Samples pro Stunde zu testen. Diese Daten können dann weiterhin verwendet werden, um verschiedene Machine-Learning-Algorithmen zu trainieren oder generell um neue Techniken zur Erkennung und Bekämpfung von Malware zu entwickeln.

Ablauf der Automatisierten Analyse

Abbildung 1 stellt den detaillierten Ablauf des Malware-Analyseprozesses dar. Der Prozess beginnt mit dem Start einer virtuellen Ubuntu-Maschine, die einen integrierten FTP-Server enthält und über die VirtualBox CLI und API gesteuert wird. Sobald die virtuelle Windows-Maschine betriebsbereit ist, führt ein Skript auf der Ubuntu-Maschine die Malware-Analyse auf der Windows-Maschine aus. Dieses Skript ist verantwortlich für die Analyse der Malware, die auf der Windows-Maschine bereitgestellt wurde. Nach der Generierung eines Ergebnisses wird dieses an den FTP-Server übertragen. Bei erfolgreichem Datenempfang sendet der FTP-Server ein Signal an die Windows-Maschine, um sie herunterzufahren. Danach wird die Windows-Maschine auf einen Zustand vor der Malware-Ausführung zurückgesetzt, repräsentiert durch einen vorherigen Snapshot, und der Analyseprozess startet wieder von vorn.

Alle Befehle müssen stets von der Ubuntu-Maschine aus erfolgen, um eine Beeinträchtigung der Malware-Analyse zu vermeiden. Da Malware die VirtualBox CLI und API erkennen kann, würde eine direkte Interaktion mit dem Analyseprozess die Ergebnisse verfälschen und den Prozess komplizieren

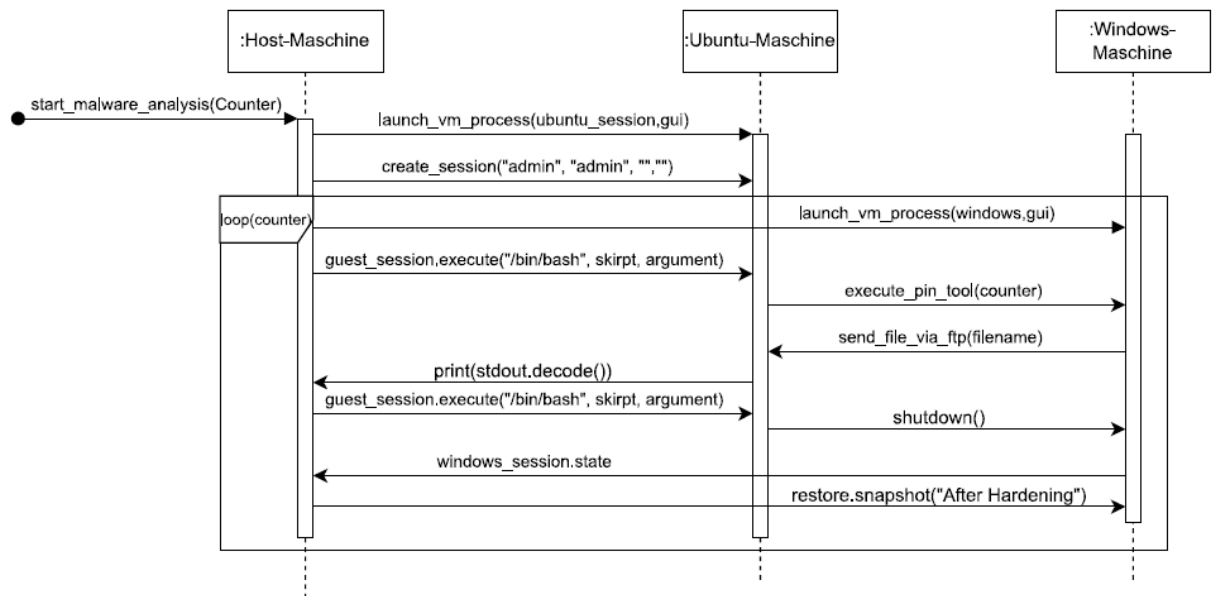


Abb. 1: Sequenzdiagramm des automatisierten Ablaufs der Malware-Analyse [3]

Ausblick

Nach Abschluss der Arbeit und nach einer längeren Beobachtungsphase des Systems sind natürlich neue Ideen und Verbesserungsvorschläge entstanden.

Der wohl wichtigste Aspekt ist die Skalierung dieses Systems. Es ist relativ einfach, im Malware-Analyse-Skript eine zusätzliche Windows-Maschine einzurichten, die parallel zur anderen Windows-Maschine läuft und auf der ebenfalls Malware getestet wird. Die Herausforderung besteht hier jedoch in den begrenzten Ressourcen, da ein gewöhnlicher PC oder Laptop nur über eine bestimmte Anzahl von RAM und CPUs verfügt, die den Windows-Systemen zugewiesen

werden können. Daher wäre die Überlegung, das Ganze auf einem Server einzurichten, der skaliert werden kann. Zwar würde ein Server Kosten verursachen, es gibt jedoch viele Online-Anbieter, bei denen man relativ günstig an einen Server mit umfangreicher Hardware gelangen kann. Das System ist aktuell in der Lage durchschnittlich 30 Malware-Proben pro Stunde analysieren, aber wir hatten in Abschnitt erwähnt, dass täglich fast 400.000 neue Malware-Arten auftreten. Daher wäre dies ein wichtiger Punkt, der angegangen werden könnte. Daher ist es notwendig die Analyse zu skalieren, sodass man schneller auf neue Malware reagieren kann.

Literatur und Abbildungen

- [1] ATLAS AV. Total Amount of Malware and PUA. <https://portal.av-atlas.org/malware>, 2023.
- [2] e.V. Bitkom. Wirtschaftsschutz 2022. *Bitkom*, 2022.
- [3] Eigene Darstellung.
- [4] Ekta Gandotra. Malware analysis and classification: A survey. *Journal of Information Security*, 2014.
- [5] Org. Kaspersky. Cybercriminals attack users with 400,000 new malicious files daily – that is 5% more than in 2021. https://www.kaspersky.com/about/press-releases/2022_cybercriminals-attack-users-with-400000-new-malicious-files-daily—that-is-5-more-than-in-2021, 2022.
- [6] Inc SoincWall. SonicWall Cyber Threat Report 2023. *SonicWall*, 2023.

Analyse der Nvidia VPI im Vergleich zu generalisierten Image Processing APIs wie OpenCV

Miralem Ademi

MarkusENZweiler

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Einleitung

Was wie Zukunftsmusik klingt, wird in nicht mehr weit entfernter Zukunft Standard werden. Die Rede ist von sich autonom bewegenden Fahrzeugen. Dabei ist die Vision von autonomer Mobilität eine, die schon lange besteht. Zwischenzeitlich hat die Technologie jedoch weite Sprünge gemacht. Inzwischen werden nicht Drähte im Boden, sondern komplexe Sensor- und Kamerasysteme mit maßgeschneiderter Software verwendet um Objekte wie Straßenschilder und Verkehrsteilnehmer aller Art, aber auch Gebäude und befahrbare Flächen zu erkennen. Darüber hinaus ist es dem Fahrzeug jeweils möglich selbstständig zu beschleunigen und zu bremsen, den Abstand zum Vorderfahrzeug zu bestimmen und zu regulieren, beim Anfahren an Kreuzungen diese als solche zu erkennen und dementsprechend die Logik dahinter zu verstehen und sich korrekt zu verhalten, Verkehrsschilder erkennen und interpretieren und weiteres.

Problemstellung

Fortschritte in vielen Bereichen der IT ermöglichen das Vernetzen von vielen Systemen. Häufig spielt die Performance dabei eine große Rolle. Diese Fortschritte werden auch von vielen anderen Branchen bemerkt und finden darin Anerkennung und Verwendung. In der Fahrzeugtechnik gibt es schon seit jeher verschiedene Methoden zur Autonomisierung von Fahrzeugen. Beispielsweise gelang es der MAN SE, ehemals MAN, in Kooperation mit der Robert Bosch GmbH bereits 1971 einen MAN-Lastwagen Typ 16.320 mittels eines Leitkabels, welches nicht nur im Boden verbaut wurde, sondern auch oberhalb des Bodens bewegt werden konnte, das Fahrzeug ohne Fahrer aktiv fahren zu lassen [2]. Heute verwenden Fahrzeuge Stereo-Kamerasysteme. Diese sind mit einem Software-Suite ausgestattet, welcher Teile der Bibliotheken von OpenCV, der Open Source Computer Vision Library, verwendet. Im Bereich der Hochschule wird beim Projekt it:movES ebenfalls auf Basis von OpenCV

programmiert. Da Nvidia die letzten Jahre im Bereich der KI immer mehr und häufiger beeindruckende Fortschritte vermeldet und nun eine eigene vielversprechende VPI für Image Processing veröffentlicht hat, liegt es nahe, diese beiden Schnittstellen gegeneinander zu vergleichen und zu erkennen, welche der beiden performanter ist.



Abb. 1: MAN vom Typ 16.320 mit leerer Fahrgastzelle [2]

Robot Operating System

Das Robot Operating System, im folgenden ROS, ist ein Framework zur Erstellung von Software für Roboter. ROS stellt viele Bibliotheken zur Verfügung, welche die Arbeit mit ROS und das Implementieren von anspruchsvollen Softwaremodulen in die Steuerung erleichtern. Der Grundgedanke hinter ROS ist die Zerstückelung des Gesamtsystems in einzelne Module, welche dann gemeinsam wirken können. Das Ziel ist es für die Zwecke des hochschuleigenen Projekts die Module in ROS zu implementieren und sie so in den Modellfahrzeugen verwendbar zu machen.

Lösungsansatz

Nach eingänglichem Studium der Dokumentationen und dem Vergleich wesentlicher Funktionen der beiden Schnittstellen werden Programme zu vergleichbaren Funktionalitäten geschrieben und auf Development Kits geflasht. Ebenfalls wird die Software in Simulationen gegeneinander verglichen. Die Ergebnisse

des Benchmarks werden schließlich gegenübergestellt und miteinander verglichen. Allgemein gilt, sollte der Gain in der Performance nur marginal ausfallen, wird es fraglich werden, ob sich die Verwendung der VPI von Nvidia lohnt. Andererseits entscheidet bei Wettbewerben auf Zeit, wie zum Beispiel dem Carolo Cup oder dem Bosch Future Mobility Cup, ein Tausendstel einer Sekunde über Sieg oder Niederlage.

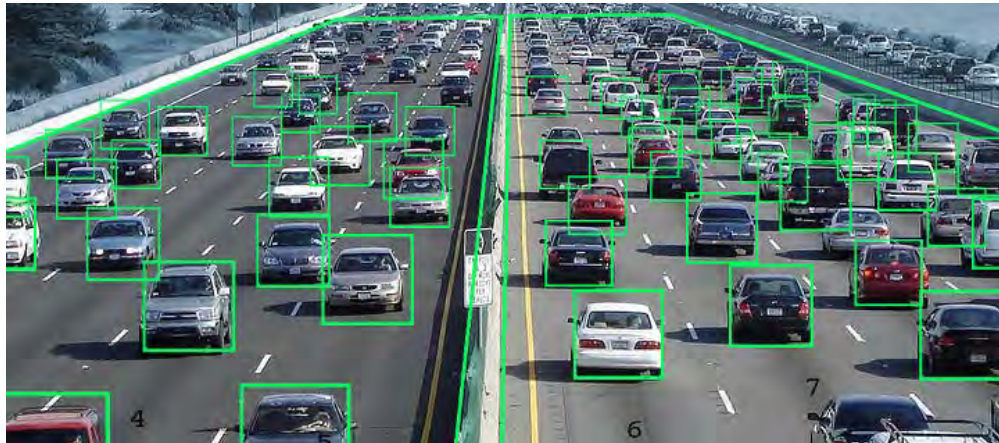


Abb. 2: Instanzierte Autos auf Autobahn via OpenCV [1]

Ausblick

In den kommenden Jahren werden immer mehr Innovationen im Bereich der IT erfolgen, soweit ist das Mooresche Gesetz noch immer zutreffend. Dabei lässt sich das Gesetz nicht nur auf die Menge der Transistoren auf dem Chiplet einer CPU begrenzen, sondern auch beispielsweise auf die Fortschritte in der KI-Forschung. Darüber hinaus haben sich viele Firmen in der IT-Branche das Ziel gesetzt die Geschwindigkeit des Gesetzes nicht nur weiterhin einzuhalten, sondern über die kommenden Jahre auch zu überflügeln

[3]. Diese Fortschritte werden der Treiber für weitere Entwicklungen in allen Bereichen sein, großen Nutzen wird auch die Branche der Fahrzeugtechnik finden. So werden die Stereo-Kamerasysteme noch fortschrittlicher, die Verbunde zwischen Software und Hardware werden noch effizienter zusammenarbeiten. Außerdem ist die Vernetzung der Fahrzeuge mittels IoT noch immer im Gange. Schließlich bleibt die Frage nicht „Ob?“, sondern „Wann?“ die Produkte für den Massenmarkt zugänglich werden und die Straßen sicherer gestalten [4].

Literatur und Abbildungen

- [1] Gary Bradsky and Vadim Pisarevsky. OpenCV Documentation. <https://docs.opencv.org/4.x/d1/dfb/intro.html>, 06 1999.
- [2] Hans Hagen. Von Geisterhand gesteuert. <https://www.mantruckandbus.com/de/momente/1971.html>, 01 2020.
- [3] Luke James. Moore's Law in 2023: What's the status quo? <https://www.power-and-beyond.com/moores-law-in-2022-whats-the-status-quo-a-dc63a87e669b554d4d33d2a5ba73692a/>, 02 2022.
- [4] X. Krasniqi and E. Hajrizi. Use of IoT Technology to Drive the Automotive Industry from Connected to Full Autonomous Vehicles. In -. IFAC-PapersOnline, 2016.

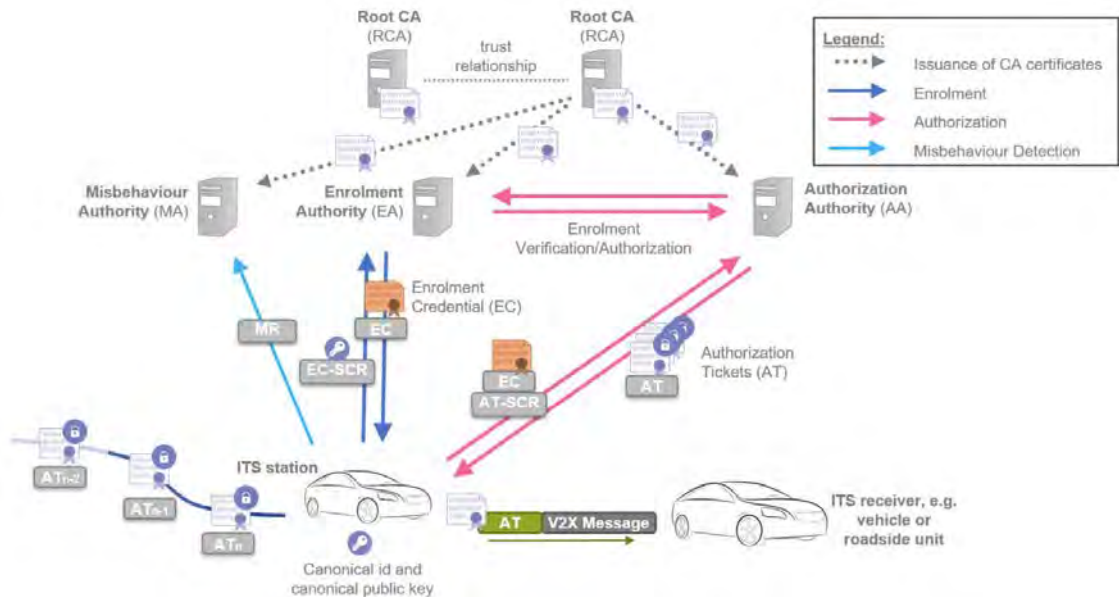


Abb. 2: Die ETSI ITS PKI Architektur [2]

Security in V2X

Um ein ITS vor Manipulation zu schützen, wurden von ETSI Mechanismen definiert, die die Sender von Nachrichten authentifizieren und damit die Verbreitung von gefälschten Nachrichten verhindern sollen. Hierfür wurde von ETSI eine hierarchische PKI-Infrastruktur entwickelt, die der im Web verwendeten PKI-Baumstruktur ähnelt. Diese ist in Abbildung 2 dargestellt.

Innerhalb der PKI ermöglichen die Authorization Authority (AA) und die Enrollment Authority (EA) einem Verkehrsteilnehmer (ITS-Station) Zugang zum ITS. Diese Certificate Authorities sind von der Root CA zertifiziert und stellen ITS-Teilnehmern Zertifikate aus. Ein von der EA signiertes Schlüsselpaar (Enrollment Credential) ermöglicht einem ITS-Teilnehmer den Zugang zum ITS. Mit diesem kann der ITS-Teilnehmer die Signatur von Schlüsselpaaren für Pseudonymzertifikate, sogenannte *Authorization Tickets (AT)* von der AA beantragen. Mit einem AT Private-Key können ausgehende Nachrichten signiert und damit die Authentizität des Absenders durch die ITS-Station nachgewiesen werden [2].

Bei der Nutzung von ATs ist zu berücksichtigen, dass durch den Authentizitätsnachweis in den regelmäßig gesendeten Nachrichten bei gleichbleibendem AT ein Bewegungsprofil des Fahrzeugs erstellt werden kann. Daher sieht der ETSI V2X Standard einen Mechanismus vor, bei dem ein aktives AT zusammen mit den IDs der anderen Layer des Fahrzeugs nach bestimmten Kriterien (z. B. Zeit, zurückgelegte Entfernung) möglichst zufällig ausgetauscht wird [2].

Ziel der Arbeit

In der Bachelorarbeit soll der V2X Security Layer des ETSI Kommunikationsstacks für ITS-Teilnehmer, die ein POSIX-basiertes Betriebssystem nutzen, in C++ prototypisch implementiert werden. Die Implementierung soll dem benachbarten N&T Layer folgende Funktionalität zur Verfügung stellen:

- `sn_encap()`: Nachrichten, die von V2X-Anwendungen über den Facilities Layer an den N&T Layer für das Versenden weitergegeben werden, sollen mit dem Private-Key eines AT signiert und die signierte Nachricht an das N&T Layer zurückgegeben werden.
- `sn_decap()`: Für Nachrichten, die von anderen ITS-Teilnehmern empfangen werden, soll die Authentizität des sendenden Teilnehmers und die Integrität der empfangenen Nachrichten geprüft werden, nicht konforme Daten sollen verworfen werden.
- `sn_certExg()`: Ermöglicht das Auslösen des Austauschs und den koordinierten Austausch von aktiv verwendeten ATs und Identifiern aus anderen Schichten des V2X-Stacks.

Da die genaue interne Implementierung der Funktionalität des Security Layers nicht standardisiert ist, umfasst der Implementierungsprozess zunächst die Zusammenstellung der Anforderungen an das Security-Modul aus Dokumenten der beteiligten Standardisierungsorganisationen. Dazu gehören die Standards der ETSI, Spezifikationen des Car2Car-Communication Consortium, Standards der EU und des BSI. Anschließend

muss eine geeignete generische Architektur für die Implementierung gewählt werden, die Module und zugehörige Funktionalitäten definiert, die die zuvor definierten Anforderungen erfüllen.

Danach wird evaluiert, wie die Architektur spezifisch auf einem POSIX-basierten System implementiert werden kann und welche Werkzeuge und Bibliotheken dafür notwendig sind. Ein kritischer Teil ist dabei das sogenannte *Secure Element*, das für die sichere

Speicherung von kryptographischem Material wie z. B. privaten Schlüsseln verwendet werden soll. Das Secure Element soll als *Hardware Security Module (HSM)* für die prototypische Implementierung emuliert werden, aber durch ein echtes in Hardware implementiertes HSM ersetzt werden können.

Nach der Fertigstellung soll das entwickelte Programmmodul in einer Testumgebung simuliert und die korrekte Funktion durch Tests nachgewiesen werden.

Literatur und Abbildungen

- [1] ETSI ETSI. EN 302 665 - V1.1.1 - Intelligent Transport Systems (ITS); Communications Architecture. https://www.etsi.org/deliver/etsi_en/302600_302699/302665/01.01.01_60/en_302665v010101p.pdf, 2010.
- [2] ETSI ETSI. TS 102 940 - V2.1.1 - Intelligent Transport Systems (ITS); Security; ITS communications security architecture and security management; Release 2. https://www.etsi.org/deliver/etsi_ts/102900_102999/102940/02.01.01_60/ts_102940v020101p.pdf, 2021.
- [3] Charles McLellan. What is V2X communication? Creating connectivity for the autonomous car era. <https://www.zdnet.com/home-and-office/networking/what-is-v2x-communication-creating-connectivity-for-the-autonomous-car-era/>, 2019.

Standardisierter Projektablaufplan für SAP Applikation Upgrades

Emre Arslan

Thomas Rodach

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Robert Bosch GmbH, Stuttgart

Einleitung

Warum sind SAP Upgrades nötig? Durch regelmäßige Upgrades stellt die Robert Bosch GmbH sicher, dass gewisse Qualitätsstandards beibehalten werden. Somit wird stets die Sicherheit gewährleistet und auch neue Funktionen seitens SAP werden den Systemen kontinuierlich hinzugefügt.

Problematik

SAP Upgrades sind nicht immer gleich. Diese können sowohl in der zeitlichen Komplexität als auch im Volumen der betreffenden Maschinen variieren. So kann ein Upgrade drei bis einhundert Maschinen beinhalten und zeitlich zwischen weniger Wochen bis mehrere Monate dauern. Auch gibt SAP vor, dass ein System mindestens alle 24 Monate ein Upgrade erhalten muss, damit es zukünftigen Support seitens SAP erhält. Da die Robert Bosch GmbH zahlreiche SAP-Systeme und Systemlinien mit regelmäßigen Upgrade Projekten versorgen muss, stellt dies eine unternehmensweite organisatorische Hürde dar. Da nicht jedes Upgrade Projekt vom selben Projektleiter betreut wird, variiert die Durchführung eines SAP Upgrades von Projekt zu Projekt stark. Deshalb stehen alle Beteiligten im Laufe eines Upgrades immer wieder von unvorhersehbaren Problemen, welche es mit einem standardisierten Ablaufplan zu beseitigen gilt.

Ziel

Ziel der Bachelorarbeit ist es, einen standardisierten Ablaufplan zu erstellen, welcher unternehmensweit eingesetzt werden soll. Der Ablaufplan soll skalierbar sein und unabhängig von der Größe und Komplexität des Upgrade Projekts anwendbar sein. Durch die Standardisierung werden die zukünftigen SAP Upgrades:

- Einheitlich in der Durchführung
- Effektiver und effizienter in der Durchführung

- Projektleiter hat eine feste Projektstruktur
- Feste Rollen sollen definiert und besetzt werden
- Aus best practices werden künftige Upgrades immer besser



Abb. 1: SAP Upgrade Approaches [1]

Umsetzung

Der erste Schritt wird es sein, Interviews mit Projektleitern der Robert Bosch GmbH durchzuführen, welche bereits SAP-Upgrade Projekte durchgeführt haben. Damit verbunden wird auch eine GAP-Analyse erbracht. Da es zu diesem Zeitpunkt keine vorgegebene standardisierte Durchführung eines SAP-Upgrades gibt, werden die Erfahrungen und Strategien befragter Projektleiter stark voneinander abweichen. Durch die Interviews lassen sich nützliche Kennzahlen dokumentieren wie:

- Bisherige Herangehensweisen
- Probleme bei der Durchführung
- Erwartete Probleme
- Unerwartete Probleme
- Was lief besonders gut
- Welche Rollen sind für SAP-Upgrades zwingend notwendig

- Mit den Infos verbundener Erfolgsgrad des SAP-Upgrades
- Weitere Einschätzungen und Erfahrungen
- Ideen und Anregungen für den standardisierten Ablaufplan

Der nächste Schritt wird es sein die bisherigen Dokumentationen, welche aus früheren SAP-Upgrades entstanden sind, zu sichten. So lassen sich Unterschiede auch in der Dokumentation feststellen. Auch die Dokumentation ist ein wesentlicher Bestandteil eines Ablaufplans und soll später in der Art und Weise, wie sie zu erstellen ist, festgelegt werden. Selbstverständlich ist das Erstellen des Standardisierten Ablaufplans für SAP-Upgrades die Hauptaufgabe und wird im Anschluss in Angriff genommen. Der Ablaufplan soll differenzierte Level an Komplexität beinhalten und unabhängig von der SAP-Version funktionieren. Besonderes Augenmerk besteht hierbei auf der Organisation der nötigen Prozesse und Abläufe, welche für ein solches SAP-Upgrade nötig sind. Eingesetzt wird der Plan von zukünftigen Projektleitern. Sowohl durch die Informationen, welche der Ablaufplan beinhalten und an den Projektleiter weitergeben wird, als auch durch die vorgegebene Struktur, soll es möglich werden jedes SAP-Upgrade einheitlich in seiner Durchführung zu gestalten. Unter anderem soll hierfür eine Dokumentation angefertigt werden, welche den Gesamtprozess in Macro- und

Microsicht genauestens beschreibt. So werden einheitliche Darstellungen erstellt, die genutzt werden, um in zukünftigen SAP-Upgrades folgende Informationen festzuhalten:

- Einheitliche Darstellungen erstellen für zukünftige Projekte (Templates)
- Einheitlicher Projekt timeline-Plan
- Organisation der Rollen/Stakeholder, welche zu besetzen sind
- Einheitliche Meeting und Abstimmungsstruktur erstellen
- Zusammenarbeitsmodell im Projekt festlegen und erstellen
- Steuerungsmanagement System entwickeln
- Einheitliches und modulares Projektmanagementsystem
- Weitere detaillierte Templates
- Unterstützende Maßnahmen für eine erfolgreiche Durchführung des Projekts beschreiben
- Training der Projektteilnehmer
- Gemeinsames Verständnis schaffen

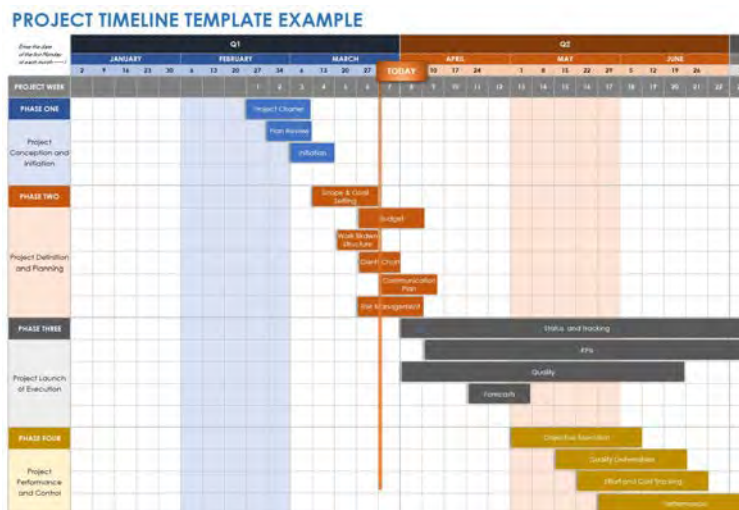


Abb. 2: Beispiel für ein Projekt Timeline Template [2]

Ausblick

Im weiteren Verlauf der wissenschaftlichen Arbeit sollen die beschriebenen Schritte durchgeführt und ein

fertig entwickelter standardisierter Projektablaufplan für SAP-Upgrades in Zusammenarbeit mit der Robert Bosch GmbH erstellt werden.

Literatur und Abbildungen

- [1] Protera Com. Your Guide To Executing An SAP Upgrade. <https://www.protera.com/your-guide-to-executing-an-sap-upgrade/>, 2021.
- [2] Smartsheet Com. Project Timeline. <https://www.smartsheet.com/content/project-timeline-templates>, 2020.

Modellierung der approximativen Bestimmung von Nutzenempfinden für serviceorientierte Geschäftsmodelle

Betiel Assefaw

Thomas Rodach

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt am Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, Stuttgart-Vaihingen

Einleitung

Anbieter von Dienstleistungen, welche die Bedürfnisse ihrer Kunden zielgerichtet befriedigen, können sich langfristig ihren ökonomischen Erfolg am Markt sichern. Aus betriebswirtschaftlicher Sicht ist es somit erstrebenswert, sich auf die Bereitstellung von Dienstleistungen zu konzentrieren, für die ein entsprechender Nutzen gegeben ist. Dabei ist Nutzen ein Maß für die Bedürfnisbefriedigung des Kunden, die durch das Angebot einer Dienstleistung erzielt wird. Der Berücksichtigung subjektiver Bedürfnisse kommt daher für die Nutzenbestimmung eine besondere Relevanz zu. Basierend auf dieser Erkenntnis, lässt sich die Grundhaltung der Serviceorientierung beschreiben, alle Aktivitäten so auszurichten, dass sie die Bedürfnisse der beteiligten Akteure (u.a. Kunden) befriedigen. Zur Überprüfung, ob die angestrebte Dienstleistung die Grundausrichtung der Serviceorientierung erfüllt, bedarf es der Beurteilung bzw. Bewertung der Graderfüllung der Bedürfnisse für den jeweiligen Kunden, mit anderen Worten, die Quantifizierung/Messung von Nutzen. Aufgrund der subjektiven Natur von Nutzen, gestaltet sich das Messen und Bewerten allerdings als schwierig. Zudem lassen sich viele bestehende Ansätze, die sich mit dem Thema der Nutzenmessung beschäftigen, nur mittels aufwendiger empirischer Sozialforschung realisieren. Vor allem bei der Überprüfung und dem Vergleich von neuen serviceorientierten Dienstleistungen ist demnach der Bedarf nach einem Ansatz zur „einfachen“ Quantifizierung von Nutzen und der Modellierung von Nutzenempfinden gegeben. Daraus lässt sich folgende Fragestellung ableiten: Wie lässt sich Nutzen aufwandsarm modellieren, so dass die Bedürfnisse der einzelnen Akteure in die approximative Bestimmung deren Nutzenempfindens miteinfließen können?

Zielsetzung

Ziel der Bachelorarbeit ist es ein ausreichend genaues, aufwandsarmes Modell zu entwickeln, dass die approximative Bestimmung des Nutzenempfindens von an der Dienstleistung beteiligten Akteuren für serviceorientierte Geschäftsmodelle ermöglicht. Im Fokus steht dabei die Systematisierung von Nutzen, um aufwandsarm und ausreichend genau Nutzenempfinden vergleichbar machen zu können und den Wert der Dienstleistung zu bestimmen.

Service Value

Als Erklärungsbeispiel für den „Service Value“ kann ein Besuch bei einem Friseur herangezogen werden. Eine Betrachtungsweise ist, dass der Friseur seine Fähigkeit des Haarschneidens und die dafür eingesetzten Ressourcen als Dienstleistung verkauft. Der Wert der Dienstleistung wird hierbei durch den Preis festgelegt. Die angebotene Dienstleistung hat demnach einen inhärenten Tauschwert (engl. Value in Exchange) (z.B. monetär) [7] [5]. Demgegenüber steht die Service-Logik. Hier steht der Service im Mittelpunkt der Betrachtung. Als Service wird ein Prozess verstanden, in dem sich der Wert einer Dienstleistung oder eines Produkts aus dem Nutzen für den beteiligten Akteur erschließt (Service Value) [2]. Der Service Value ergibt sich somit nicht nur aus dem Preis, sondern auch aus dem Kontext, in welchem der Akteur den Nutzen der Dienstleistung subjektiv wahrnimmt. Daraus resultiert der Gebrauchswert (engl. Value in Use), der für jeden Akteur einzigartig und individuell ist. In dem eingangs genannten Beispiel könnte der Value in Use „ein gutes Aussehen“ sein oder aber „bessere Chancen im Bewerbungsprozess“. Der Value in Use kann zur Ermittlung des Service Values als Messgröße hinzugezogen werden [1] [5].

Elements of Value nach Bain & Company

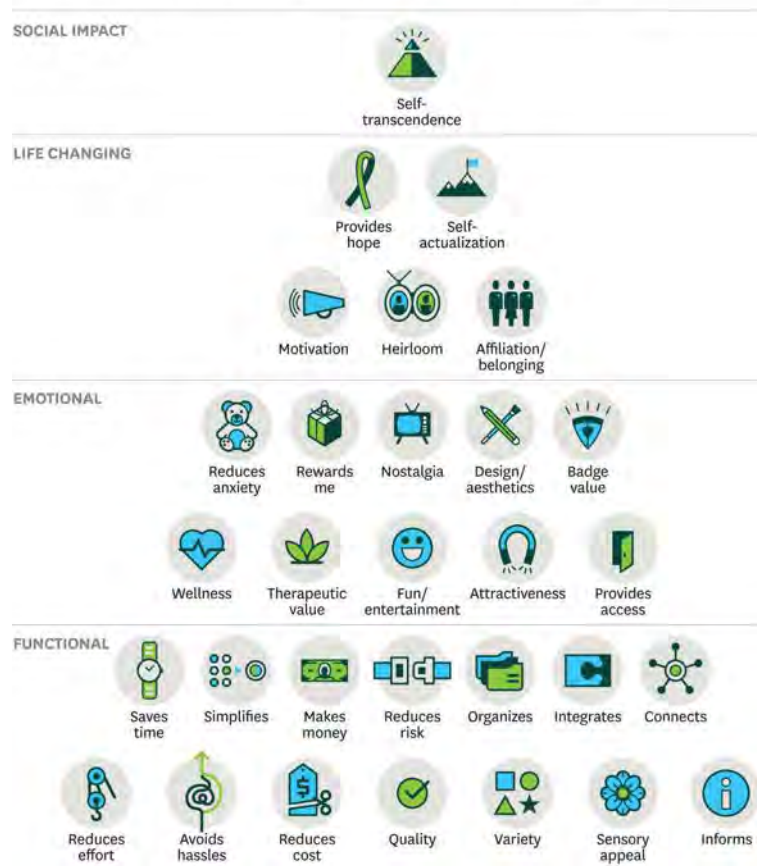
Die Entscheidung für die Beschaffung eines Produkts oder einer Dienstleistung kann sowohl von objektiven Kriterien wie Preis, Servicegrad und/oder Garantieleistung abhängen als auch von subjektiven Kriterien wie Design und Ästhetik. Die Höhe und Art des Werts ergeben sich somit aus der subjektiven Einschätzung des Kunden. Für Unternehmen ist es daher wichtig das ganze Spektrum der rationalen und emotionalen Faktoren zu kennen, die für den Beschaffungsprozess ihrer Kunden von Relevanz sind. Mit dem „Elements of Value“-Modell, ist es Bain & Company gelungen 30 Werteelemente - grundlegende Attribute, die eine Dienstleistung oder ein Produkt aus Kundensicht spezifizieren - zu identifizieren. Das Modell bietet den Unternehmen die Möglichkeit zu verstehen, wie und in welchem Maß Verbraucher eine bestimmte Dienstleistung oder ein bestimmtes Produkt

schätzen, um basierend darauf den Kundenwert zu maximieren. Die sogenannten Verbraucherwerte lassen sich hierarchisch in folgende vier Kategorien einteilen und variieren in ihrer Relevanz je nach Branche, Demografie und Kultur (siehe Abb. 1) [3] [6]:

- Funktionselemente (engl. functional): Was macht das Produkt oder die Dienstleistung?
- Emotionale Elemente (engl. emotional): Wie fühle ich mich mit dem Produkt oder der Dienstleistung?
- Lebensverändernde Elemente (engl. life changing): Wie verändert es oder sie mein Leben?
- Elemente der sozialen Wirkung (engl. social impact): Welche Auswirkung hat es oder sie auf die Welt?

The Elements of Value Pyramid

Products and services deliver fundamental elements of value that address four kinds of needs: functional, emotional, life changing, and social impact. In general, the more elements provided, the greater customers' loyalty and the higher the company's sustained revenue growth.



SOURCE: © 2015 BAIN & COMPANY INC. FROM "THE ELEMENTS OF VALUE," SEPTEMBER 2016

© FOUR_ORG

Abb. 1: The Elements of Value Pyramid [3]

Das Modell ist eine Anlehnung an das Modell der Maslowsche Bedürfnishierarchie. Während bei der Maslowschen Bedürfnishierarchie die menschlichen Bedürfnisse und Motive untersucht werden, um daraus menschliches Verhalten abzuleiten, werden mittels der Wertelemente, die Grundwerte, die das Verbraucherverhalten beeinflussen, identifiziert. Elements of Value ermöglicht Unternehmen eine datengesteuerte und bedürfnisorientierte Gestaltung ihrer Dienstleistung oder ihres Produkts, durch den Einsatz der richtigen Wertkombinationen [3] [4].

Ausblick

Erste Ideen für die Modellierung der approximativen Bestimmung von Nutzenempfinden für serviceorientierte Geschäftsmodelle befinden sich in der Konzeption. Die Annahme ist, den Service Value über den Value in Use des Kunden und die Wertelemente der Dienstleistung

annähernd zu quantifizieren und zu modellieren. Die Bestimmung des Value in Use soll über die Verwendung einer adäquaten, approximativen und quantitativen Bewertungsmethode der Bedürfnisse und Value-Elemente erfolgen.

Dazu könnte die Vorgehensweise wie folgt aussehen:

1. Identifikation der Bedürfnisse der Zielgruppe
2. Bestimmung der relevanten Wertelemente der Dienstleistung
3. Skalierung der Bedürfnisse der Zielgruppe und Wertelemente auf quantitativer Ebene
4. Modellierung des Value in Use

Für die Umsetzung der einzelnen Schritte und Realisierung des angestrebten Modellierungsansatzes bedarf es allerdings weiterer Evaluation und Analyse.

Literatur und Abbildungen

- [1] Team BeDien. Value in Use. <https://www.bedien.org/value-in-use#:~:text=Dieser%20besagt%2C%20dass%20Kunden%20als,sogenannte%20operande%20und%20operante%20Ressourcen.,> 07 2019.
- [2] Manfred Bruhn and Karsten Hadwich. Service Value - Eine Einführung in die theoretischen und praktischen Problemstellungen. In *Service Value als Werttreiber*, pages 3–31. Bruhn, Manfred; Hadwich, Karsten, 2014.
- [3] Almquist Eric, Senior John, and Bloch Nicolas. The Elements of Value. <https://hbr.org/2016/09/the-elements-of-value>, 09 2016.
- [4] Orit Gadiesh, Eric Almquist, Meacham Matthew, and Charlotte Apps. The Power of Delivering Elements of Value® in Consumer Products. <https://www.bain.com/insights/the-power-of-delivering-elements-of-value-in-consumer-products/>, 09 2019.
- [5] Christian F. Koof, Marion Büttgen, and Giancarlo Sapienza. Empirische Evidenz der Determinanten des Value in Use bei Professional Service Firms. In *Service Value als Werttreiber*, pages 427–456. Bruhn, Manfred; Harwich Karsten;, 201.
- [6] M Masterclass. How to Use the Elements of Value to Appeal to Customers. <https://www.masterclass.com/articles/elements-of-value>, 06 2022.
- [7] Susanne Robra-Bissantz, Christoph Lattemann, et al. *SERVICE FOR GOOD*. Lattemann, Christoph, 2022.

Konzeptentwicklung für die Einführung und Veränderung von Prozessen auf Basis des ITIL Change Managements und des Change of Mind-Ansatzes

Jan Lukas Autenrieth

Anke Bez

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Festo SE & Co.KG, Scharnhausen

Problemstellung

Die Implementierung von Änderungen innerhalb eines Unternehmens kann eine große Herausforderung darstellen, insbesondere in mittelständischen bis großen Unternehmen, bei denen hohe Abhängigkeiten zwischen Prozessen und Services bestehen. So können Veränderungen „[...] gravierend in das bestehende Gefüge einer Organisation eingreifen und dabei z.B. auf deren Aufbau/Strukturen und Abläufe/Prozesse einwirken“ [8]. In diesem Zusammenhang ist Change Management heutzutage ein wesentlicher Bestandteil des Unternehmensmanagements, indem es sich mit der Planung, Implementierung und Überwachung von Änderungen befasst und so zur Sicherstellung des Erfolgs in allen Geschäftsbereichen beiträgt. Dabei muss beim Change Management in eine fachliche und eine überfachliche Perspektive unterschieden werden. Diese beiden Perspektiven werden im Verlauf der vorliegenden Bachelorarbeit anhand des ITIL Change Managements (fachlich) und dem Change of Mind Ansatz (überfachlich) analysiert, um eine höhere Akzeptanz bei Mitarbeitern und Führungskräften zu schaffen und Widerstände möglichst gering zu halten.

ITIL

Die ITIL (Information Technology Infrastructure Library) ist eine „Sammlung von Best bzw. Good Practices im IT Service Management (ITSM), die über Regeln Aufbau- und Ablauforganisation sowie die benötigten Werkzeuge im IT-Dienstleistungsbereich beschreibt“ [7], also eine Sammlung von bewährten Vorgehensweisen, Methoden oder Verfahrensweisen, die in einem bestimmten Bereich oder in einer bestimmten Branche als besonders effektiv und effizient anerkannt werden. Dabei handelt es sich um Erfahrungswerte und Wissen, das über längere Zeit hinweg gesammelt wurde und sich in der Praxis bewährt hat. ITIL wurde in den 1980er Jahren in Großbritannien entwickelt und hat

sich seitdem als globaler Standard für ITSM etabliert. Die Bibliothek besteht aus einer Reihe von Verfahren, die Unternehmen dabei unterstützen sollen, ihre IT-Services effektiver zu verwalten und gleichzeitig die Kosten zu senken und die Qualität zu verbessern. ITIL V3 gliedert sich in fünf Kernbereiche: Service Strategy (Service-Strategie), Service Design, Service Transition (Service-Übertragung), Service Operation (Service-Betrieb) und Continual Service Improvement (Kontinuierliche Serviceverbesserung) (vgl. Abbildung 1). Dabei werden „die Phasen Service Design, Service Transition und Service Operation von dieser [festgelegten Service] Strategie geführt“ [1] und mithilfe des Continual Service Improvements ständig überwacht und basierend auf den strategischen Zielen optimiert. Jeder Kernbereich umfasst eine Reihe von Prozessen und Funktionen, die Unternehmen dabei helfen sollen, die IT-Services effektiv zu planen, zu entwerfen, zu implementieren und zu betreiben sowie kontinuierlich zu verbessern.

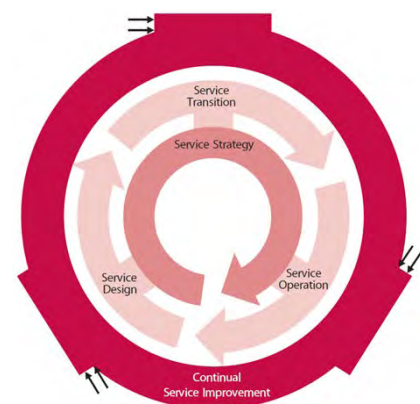


Abb. 1: ITIL V3 Lifecycle [5]

Das Change Management nach ITIL ist hierbei dafür verantwortlich, dass Changes kontrolliert implementiert werden. Das heißt, es stellt sicher, dass Changes nach

Eingang eines Request for Change (RFC) evaluiert, priorisiert, geplant, getestet, implementiert und dokumentiert werden (vgl. Abbildung 2). Es kann sowohl proaktiv, beispielsweise bei einer Kostenreduktion, als auch reaktiv, wie bei der Beseitigung von Serviceunterbrechungen, agieren. Für die Umsetzung genehmigter Changes ist schlussendlich jedoch der Service Owner verantwortlich.



Abb. 2: Change Management Lifecycle [2]

Change of Mind-Ansatz

Neben dem eher technischen Ansatz des Change Managements nach ITIL spielt in jedem Wandel auch der psychologische Aspekt eine sehr wichtige Rolle. Selten sind betroffene Personen mit einer Veränderung von Beginn an einverstanden, sei es im privaten Leben, bei neuen Arbeitsmethoden oder eben auch bei Änderungen in der IT-Landschaft. Eine Studie des österreichischen Herstein Institut fand bei einer Befragung von 1000 Unternehmen heraus, dass rund 38% der durchgeführten Change Prozesse scheiterten oder weniger erfolgreich als erwartet waren. Auf die Frage nach dem Grund des Scheiterns verwiesen 30% auf den Widerstand von Mitarbeitern gegenüber einem Wandel. 25% erachteten eine mangelhafte Prozesssteuerung als Ursache sowie 20% einer zu schnellen Durchführung [6]. Um diesem Widerstand in der Belegschaft entgegenzuwirken, entwickelte Kurt Lewin, der als einer „[...] der Urväter der modernen Psychologie [...]“ [4] gilt, ein Modell, welches auf den drei Phasen **Auftauen**, **Verändern** und **Verankern** beruht (vgl. Abbildung 3) und sich mit akzelerierenden (den Wandel antreibenden) und retardierenden (sich dem Wandel entgegensetzenden) Kräften befasst.



Abb. 3: 3-Phasen Modell nach Lewin [3]

Die große Herausforderung im Change Management ist es, diese Kräfte in ein Gleichgewicht zu bringen. Denn sollten retardierende Kräfte überwiegen, würde kein Wandel stattfinden. Andererseits würde bei einem Überwiegen der akzelerierenden Kräfte, also dem stetigen Drängen auf weiteren Wandel, eine Überforderung der Mitarbeiter und somit Leistungsverlust entstehen (siehe Abbildung 4).

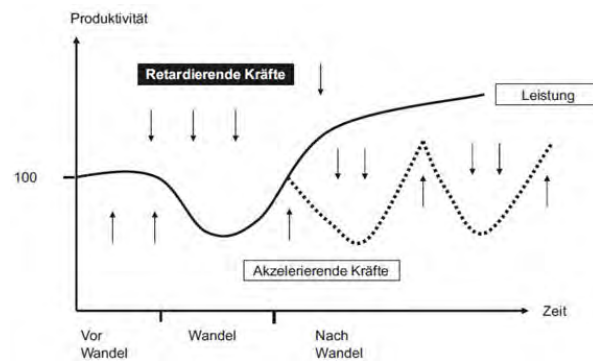


Abb. 4: Leistungskurve während eines Wandels [4]

Ausblick

Change Management wird in unserer schnelllebigen und sich stets wandelnden Geschäftswelt auch in Zukunft eine enorm große Rolle spielen. Zwar können Unternehmen auf Leitfäden wie die ITIL zurückgreifen, um Veränderungsprozesse ihren Wünschen und Visionen anzupassen, die Recherche für diese Bachelorarbeit hat jedoch gezeigt, dass auch wenn Change-Prozesse in der Theorie ordentlich implementiert sind, es ohne Einbeziehung des überfachlichen, psychologischen Aspekts kaum möglich ist, Mitarbeiter und Kollegen von der Umsetzung neuer oder sich verändernder Prozesse zu überzeugen.

Literatur und Abbildungen

- [1] Jan van Bon. *ITIL 2011 - Das Taschenbuch*. Van Haren Publishing, 2013.
- [2] Ayan Brahmachary. Change Management Lifecycle. <https://www.certguidance.com/change-management-itol/>, 2018.
- [3] Gerhild Deutinger. *Kommunikation im Change - Erfolgreich kommunizieren in Veränderungsprozessen*. Springer Gabler, 2017.
- [4] Thomas Lauer. *Change Management - Grundlagen und Erfolgsfaktoren*. Springer Gabler, 2019.
- [5] Axelos Limited. *The Official Introduction to the ITIL*. TSO - The Stationary Office, 2007.
- [6] Eric Schott and Christophe Campana. *Strategisches Projektmanagement*. Springer-Verlag, 2005.
- [7] Markus Siepermann. Definition ITIL. <https://wirtschaftslexikon.gabler.de/definition/itil-53471/version-276560>, 2018.
- [8] Kerstin Stolzenberg and Krischan Heberle. *Change Management - Veränderungsprozesse erfolgreich gestalten . Mitarbeiter mobilisieren*. Springer-Verlag, 2021.

The opportunity of creating a Data Hub for the Mercedes-Benz AG in order to optimize sales steering.

Rozin Babat

Dirk Hesse

Department of Computer Science and Engineering, Esslingen University

Work carried out at Mercedes-Benz AG, Stuttgart

Introduction

Data - a short word with huge importance. As time passes, data becomes more important. Nowadays the biggest companies in the world work on data. In 2006, the British mathematician and entrepreneur Clive Humby gained international attention with the quote "Data is the new oil" [4]. But how come it gained this much importance over the last years?

Data helps companies understand and solve problems and lastly make decisions based on data analytics. After the data is collected, it is being cleaned. Thus, the raw data is being sorted and filtered to prepare it for further analysis and lastly to display it to the user in a readable form like graphs, tables and charts, usually visualized in a dashboard. Due to the increasing importance of data analytics and data visualization over the last years, many big companies invested heavily into this section.

One of these companies is one of the biggest automobile manufacturers worldwide, the Mercedes-Benz AG. Mercedes-Benz introduced a variety of measures to optimize the data usage within the company. One of these measures was creating a data hub.

Data Hub

A data hub is a centralized system for data management. In the data hub, company data from various sources is being collected and afterwards spread throughout the company. The data hub is responsible for the entire data management process from collecting the data to cleaning it and finally displaying it in dashboards and spreading it to the different divisions in the company. While the data lake "is used for storing vast amounts of unprocessed data in its natural

or raw format" [2], the data hub is used for exchanging and sharing the data to make it readable for the user. Figure one shows the four different attributes a data hub consists of: the data models, the data governance, the data integration and the data persistence. The data models show how the data is structured and consumed while the data governance refers to the data privacy, access, control and security. The data integration describes the style and method of integrating the data into the data hub and the data persistence refers to the category of data store.

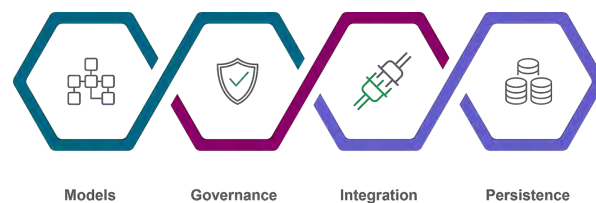


Fig. 1: Data Hub Attributes [1]

A data hub consists of different layers which can be seen in the second figure. The first layer is the source system layer, in which the data is being extracted from its original source and then integrated into the data hub. The second layer is the data integration layer, in which the data is being harmonized, mastered and also enriched with metadata. The storage layer is the third layer, and it is responsible for corresponding to the needs of database management and data modeling. The next layer is the data access layer which unites all access points connected to the data hub. Lastly, the orchestration layer is used to provide control over data integration, data flows, data transformations and data governance.

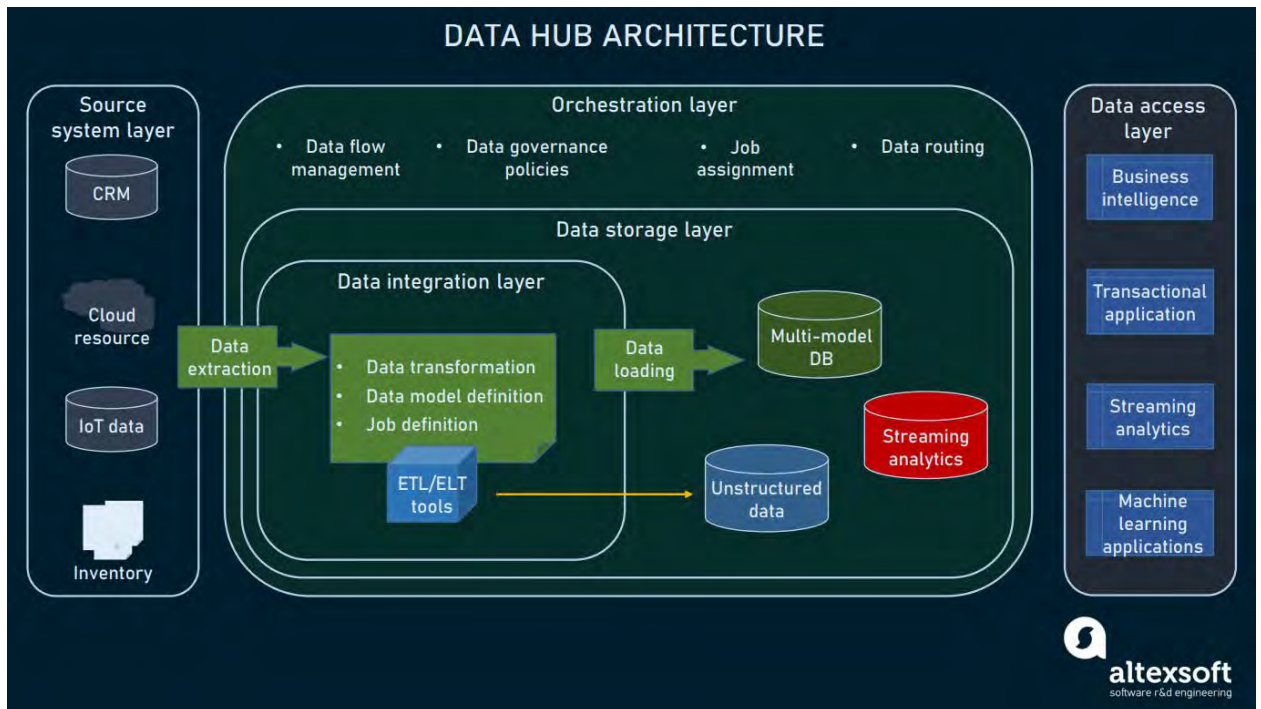


Fig. 2: Data Hub Architecture [3]

Data at Mercedes-Benz

The goal of this bachelor thesis is to find the opportunities the data hub holds for Mercedes-Benz to optimize their sales steering. It is important to understand how data from different departments can be used for the managers of the whole division to finally make decisions based on these data analytics. It is also essential to understand how the results of the data analysis can be displayed well for the decision makers to get the most information out of the analysis. Furthermore, it needs to be determined how the reports are being shared to the different departments. Mercedes-Benz uses Microsoft's Power BI to create dashboards with real-time transmission to let the user be aware of the latest updates concerning the data. After the reports are shared to the different departments, each department can decide themselves whether they take the reports as they are, whether they edit and adjust the reports the way they want them to be or whether they set-up a completely new dashboard which covers the information they need and they want to see on the dashboard. Either way of treating the received reports is being used at Mercedes-Benz.

Outlook

The data hub of the division this bachelor thesis is being completed in is located in Malaysia and managed by a team of eight employees, also being located in Malaysia. This team uses the centralized data hub to share the collected data to the different departments. In this way, they make it possible for their colleagues, especially the managers, to get as much relevant information as possible from the dashboards and to finally make decisions, which affect the whole business, based on the given information. For the team sitting in Malaysia, it is important to know which departments still need support to adjust the reports and which departments can do it on their own. Thus, the different departments are being interviewed to understand how they get to manage the reports they receive from the data hub team. Moreover, the team in Malaysia is being interviewed to find out what problems they have with managing the data and what information from the different departments is missing for them to create the optimal dashboards. In addition to that, it is crucial for the team in Malaysia to know whether the given reports display the current situation well and whether the information given on the reports is helpful for the colleagues and whether it can be used to support the decision-making process in order to optimize the sales steering of the vehicles in the company.

References and figures

- [1] Clive Bearman. Data Hub, Fabric or Mesh? Part 1 of 2. <https://www.qlik.com/blog/data-hub-fabric-or-mesh-part-1-of-2>, 2021.
- [2] Thorsten Heller. How is a Data Hub different from a Data Lake? <https://www.greenbird.com/solutions/data-hub-vs-data-lake>, 2021.
- [3] Oleksandr Medovoi. What is Data Hub: Purpose, Architecture Patterns, and Existing Solutions Overview. <https://www.altexsoft.com/blog/data-hub/>, 2021.
- [4] Matt Watts. Why data is the new oil. <https://futurescot.com/why-data-is-the-new-oil/>, 2021.

Re-Engineering einer Software für Smartphone-Reparatur Management: Usability, Performance und Security

Thibaut Becher

Astrid Beck

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Repair me GmbH, Stuttgart

Einleitung

In der heutigen digitalen Ära sind Smartphones allgegenwärtig und unverzichtbar geworden. Neben dem Verkauf neuer Geräte und von SIM-Karten ist die Reparatur von Smartphones zu einem wichtigen Geschäftsfeld der Mobilfunkbranche geworden. Die Sofortreparatur vor Ort ist eine beliebte Methode, bei der der Kunde sein defektes Handy in einen Repairshop bringt und die Reparatur sofort und vor seinen Augen durchgeführt wird. Dies ermöglicht es dem Kunden, den Reparaturprozess zu beobachten und insbesondere zu verfolgen, was mit seinem Gerät und den darauf gespeicherten, vertraulichen Daten geschieht. Diese Art der Reparatur ist effizient und nachhaltig, vermeidet den Versand des Geräts und sorgt für minimale Einschränkungen beim Kunden.

Ursprünglich war die Sofortreparatur nur auf Schäden beschränkt, die nicht von der Herstellergarantie abgedeckt waren. Jedoch haben Hersteller und Versicherungen die Vorteile dieser Reparaturmethode schnell erkannt. Hersteller haben dann ausgewählte Reparaturketten zertifiziert, um die Sofortreparatur für ihre Marke anzubieten. Versicherungen von Smartphones bevorzugen die Sofortreparatur aufgrund der geringeren Kosten und der hohen Kundenzufriedenheit. In den letzten Jahren haben die meisten Handyversicherungen das Modell des Geräteauswechsels (SWAP) durch Sofortreparaturen ersetzt. Die Verwaltung dieser Reparaturen ist jedoch zunehmend komplex geworden, weil Geräte technisch immer anspruchsvoller wurden, die Ersatzteile unterschiedliche Qualität haben etc.

Die Projektaufgabe ist folgende: Es soll ein neues Benutzer-Interface für die Auftragsverwaltung eines Sofortreparaturanbieters entwickelt werden, welches die wichtigsten Prozesse des Reparaturzyklus abbildet. Zusätzlich soll eine Statusseite erstellt werden, welche die Informationen aus dem Kernsystem und der Webseite verbindet. Besonderes Augenmerk muss dabei auf die Sicherheitsaspekte gelegt werden, wie z.B. Cross-Site-Scripting. Des Weiteren soll die API

(Schnittstelle) einer Versicherung integriert werden, um automatisch neue Aufträge zu übernehmen und den aktuellen Reparaturstand online mitteilen zu können. Auch soll die Korrespondenz, die bisher per E-Mail erfolgt, automatisch den Aufträgen zugeordnet werden.

Konzeption

Das Projekt wird in zwei Teile aufgeteilt: einerseits das Frontend (sichtbarer Bereich der Anwendung), das die Benutzungsoberfläche umfasst, und andererseits das Backend (Logik, Datenverwaltung und Interaktionen mit der Datenbank) mit einer API (siehe Bild 1). Es wird außerdem ein Sicherheitsstandard nach den Vorgaben des Bundesamts für Sicherheit in der Informationstechnik (BSI) implementiert. Die Kommunikation mit dem Endkunden erfolgt über einen Bot bei Telegram oder andere APIs so wie für E-Mail, SMS und Whatsapp. Das Frontend wird in JavaScript mit React.js entwickelt, während das Backend mit Python umgesetzt wird.

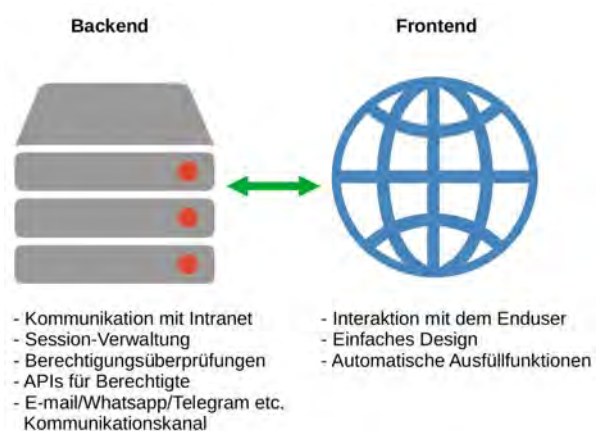


Abb. 1: Konzept Frontend/Backend [2]

Da die Authentifizierung der Benutzer am Backend ein kritischer Punkt der Anwendung ist, wird eine zertifizierte

katbasierte Authentifizierung (siehe Bild 2) anstelle von einer "einfachen" Benutzer-Passwort-Kombinationen verwendet. Dies bietet mehrere Vorteile, darunter eine vereinfachte Benutzenerfahrung und die Beseitigung des Problems schwacher Passwörter direkt und die

Verhinderung von Brute-Force-Angriffen. Phishing-Angriffe sind ineffektiv, da keine Passwörter gestohlen werden können. Externe Benutzer benötigen lediglich den Zugangsschlüssel, welcher in Missbrauchsfällen direkt gesperrt werden kann. (siehe: [5])

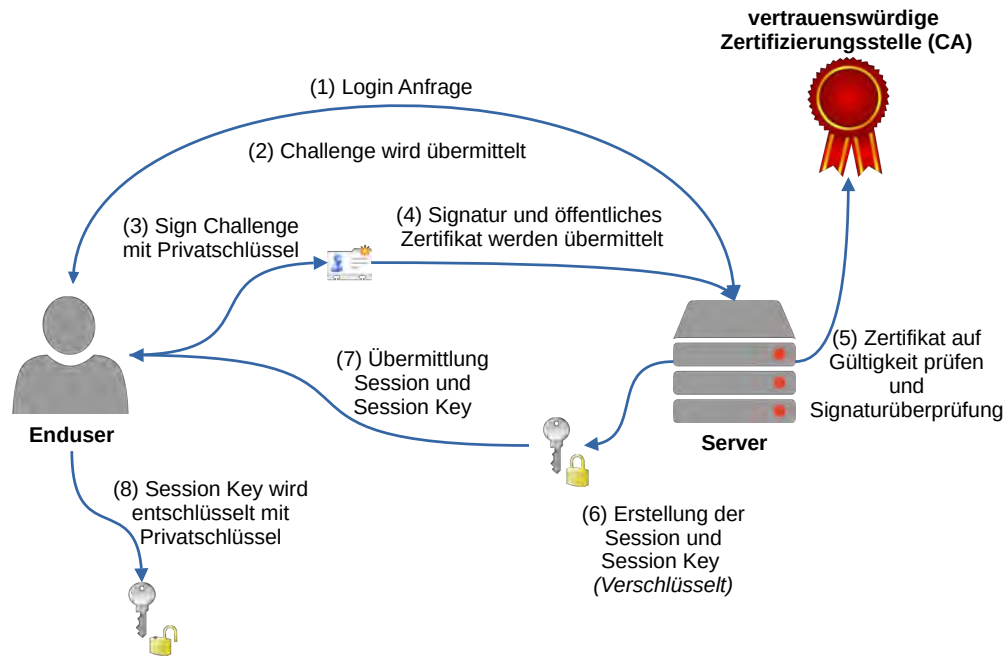


Abb. 2: Zertifikat basierte Authentifizierung [2]

Die Anbindung der Datenbank wird mittels "Connectors" realisiert. Dies ermöglicht dem Betreiber, eine beliebige Datenbank zu verwenden, außerdem können bestimmte Testfälle in beispielsweise lokalen SQLite-Dateien gespeichert werden. Mit einer gesonderten Testumgebung können neue Implementierungen getestet werden, bevor diese veröffentlicht werden, während das System selbst ungestört weiter voll in Betrieb ist. Diese Testumgebung kann die Applikation mithilfe des "Connectors" auf eine Test-Datenbank umleiten.

Sicherheitsstandard nach BSI

Eine sorgfältige Auswahl der Verschlüsselungsalgorithmen ist von großer Bedeutung für die Sicherheit von Anwendungen, die mit vertraulichen Daten arbeiten. Das BSI empfiehlt für die symmetrischen Verschlüsselungsverfahren die Verwendung von AES-128, AES-192 und AES-256. (siehe: [3])

AES (Advanced Encryption Standard) ist ein symmetrischer Verschlüsselungsalgorithmus, der zur Verschlüsselung von Daten verwendet wird. Die Zahlen 128, 192 und 256 beziehen sich auf die Schlüssellänge des Algorithmus, d.h. wie viele Bits der Schlüssel hat. Ein längerer Schlüssel kann eine höhere Sicherheit

bieten, jedoch kann dieser auch zu einer geringeren Leistung führen. (siehe: [1])

Für Signature-Algorithmen werden RSA, DSA und DSA-Varianten auf elliptischen Kurven empfohlen. Elliptische Kurven sind geometrische Objekte, die in der Mathematik und Kryptographie verwendet werden. Sie haben spezielle algebraische Eigenschaften und werden zur Verschlüsselung und digitalen Signatur verwendet. RSA ist ein asymmetrisches kryptografisches Verfahren, das sowohl für Verschlüsselung als auch für digitale Signaturen verwendet werden kann. DSA ist ein digitales Signaturverfahren, das auf diskreten Logarithmen basiert (mathematisches Konzept). DSA ist insbesondere für den Einsatz in Umgebungen mit begrenzten Ressourcen geeignet. Die DSA-Varianten auf elliptischen Kurven sind ähnlich wie DSA, verwenden jedoch elliptische Kurven anstelle von endlichen Körpern (mathematische Struktur). Diese Algorithmen haben sich als sicher und zuverlässig erwiesen und werden von vielen Organisationen und Regierungen weltweit eingesetzt. (siehe: [3], [4])

Umsetzung

Das Projekt wurde gemäß den vorher festgelegten Anforderungen erfolgreich umgesetzt. Ein Proof-of-Concept wurde entwickelt und in das bestehende System integriert, wobei die vorgegebenen Technologien (Python und JavaScript) verwendet wurden. Bei der Umsetzung wurden folgende Anforderungen erfüllt: präemptive Fehlerbehandlung, Behandlung von tatsächlich aufgetretenen Fehlern sowie die Gewährleistung von Datensicherheit, Authentizität, Integrität, Vertraulichkeit, Verfügbarkeit und Verbindlichkeit. Bei den ersten Tests des Systems wurden positive Ergebnisse erzielt. Es traten nur wenige unvorhersehbare Fehler auf, die größtenteils auf Bedienfehler

zurückzuführen waren. Diese Fehler wurden erfasst und werden in einer weiteren Korrekturrunde behoben.

Die Mitarbeiter, die die Software praktisch einsetzen, äußerten sich äußerst zufrieden. Sie bestätigten eine deutliche Verbesserung der Benutzerfreundlichkeit und der Bedienbarkeit.

Ausblick

Das System wird in seiner ersten Version im praktischen Einsatz weiter getestet. Die Erfahrungen der Anwender und auftretende Fehler werden fortlaufend erfasst. In einigen Monaten wird ein weiteres Evaluierungsgespräch durchgeführt, um zusätzliche Verbesserungsvorschläge zu identifizieren.

Literatur und Abbildungen

- [1] Corinne Bernstein and Michael Cobb. Advanced Encryption Standard (AES). <https://www.techtarget.com/searchsecurity/definition/Advanced-Encryption-Standard>, 2023.
- [2] Eigene Darstellung.
- [3] Bundesamt für Sicherheit in der Informationstechnik. BSI – Technical Guideline. https://www.bsi.bund.de/SharedDocs/Downloads/EN/BSI/Publications/TechGuidelines/TG02102/BSI-TR-02102-1.pdf?__blob=publicationFile, 01 2023.
- [4] Klaus SchmeH. *Kryptografie*. dpunkt.verlag GmbH, 6 edition, 2016.
- [5] Yubico Yubico. Certificate-Based Authentication Definition. <https://www.yubico.com/resources/glossary/what-is-certificate-based-authentication>, 2023.

Semantic Visual SLAM in Outdoor Environments

Maxim Becht

MarkusENZweiler

Department of Computer Science and Engineering, Esslingen University

Work carried out at Department of Computer Science and Engineering, Esslingen

Introduction

SLAM (Simultaneous Localization and Mapping) is an important component for many real-world applications in various fields, such as AR (Augmented Reality), robotics, and UAVs (Unmanned Autonomous Vehicles). The idea behind SLAM is described in its name: an autonomous system should be able to locate itself, while creating a map of the environment. This task can be achieved by observing the environment with different types of sensors e.g. LiDAR (Light Detection and Ranging), laser distance sensor, mono and stereo cameras. When mainly camera sensors are in use for this task, it is often referred to as V-SLAM (Visual SLAM). V-SLAM solves the problem of localization by analyzing positional image features (feature points), which represent unique markers. After evaluating the feature points, mapping can begin. Most SLAM architectures are re-evaluating previous saved feature points, when looping occurs. Loop detection is an important task within the SLAM process, because looping leads to re-observations of feature points already stored. By re-evaluating these feature points, in addition to newly detected feature points, a more robust and accurate description of the environment can be achieved. Similar detections are required when combining separately created maps.

Problem description

Dynamic environments are a threat to classical SLAM methods, because SLAM relies on static feature points and thus, on static objects. Many environments consist of dynamic objects, especially outdoor environments, which are less controllable to influences of natural events. The higher density of dynamic behaviour in these sorts of environments causes the need to adopt SLAM further towards detecting and avoiding dynamic objects.

Goal

The goal of this work is to stabilize SLAM in dynamic environments by masking out feature points of dynamic objects using Semantic Segmentation. For this task a framework similar to DS-SLAM (see figure 3) is given. The main objective can be divided into two sub-objectives, training the segmentation model on an outdoor environment dataset and integrating the trained model into the existing framework. After implementation, evaluation of the masking capabilities will be conducted (hit rate, real-time performance, etc.).

Data preparation

In order to train the segmentation model, data preparation on the given dataset has to be done. The given dataset comprises 2000 pairs of real (in contrast to synthetic) and labeled footage. Labeling in segmentation datasets is done by pixel color coding. The number of labeled classes equals the variety of colors used in annotation. There are three different variations of annotations over the same dataset from 23 different classes simplified to only two. Since all three different variations will be used in training, starting from training 23 different classes to only two in the process of transfer training (switching output layer), every set of annotations and the raw footage must be prepared before that. Modifications to dataset:

- divisibility by 32 -> resolution change from 1296 × 972 down to 640 × 480
- grayscale
- data split in 80/10/10 [%] (training/test/validation)

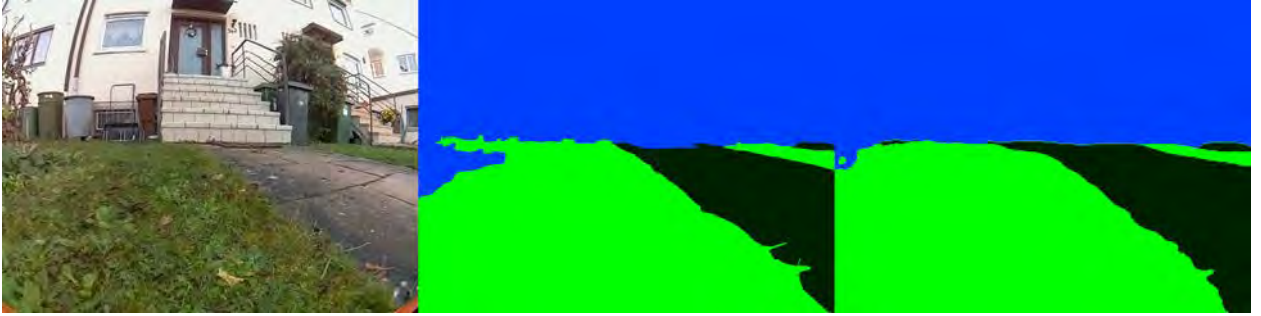


Fig. 1: left to right: raw, labeled, predicted [3]

Segmentation model training

The implementation of the applied image segmentation keras model uses the Tensorflow keras framework (is provided over the MIT license [2]). All models are trained using the dataset mentioned in previous paragraph.

Training results

Training results are given in different measurements: FwIoU (frequency weighted Intersect over Union), mIoU (mean Intersect over Union) and evaluation performance. The mIoU and FwIoU metrics are calculated using formulars proposed by [4]:

$$\text{mean IoU: } (1/n_{cl}) \sum_i n_{ii} / (t_i + \sum_j n_{ji} - n_{ii}) \quad (1)$$

$$\text{FwIoU: } (\sum_k t_k)^{-1} \sum_i t_i n_{ii} / (t_i + \sum_j n_{ji} - n_{ii}) \quad (2)$$

where n_{ij} is the number of pixels of class i predicted to belong to class j and n_{cl} the number of different classes.

$$t_i \text{ (total number of pixels of class } i) = \sum_j n_{ij}$$

model (classes)	FwIoU	mIoU	evaluation performance for 200 iterations [s]
SegNet (23)	0.6537	0.2479	≈18 (10.79it/s)
SegNet VGG16 encoder (23)	0.8123	0.4224	≈19 (10.17it/s)
SegNet VGG16 encoder (3)	0.9422	0.6756	≈14 (13.42it/s)

Fig. 2: evaluation results of trained SegNet models [3]

The 23 class SegNet model was trained for 60 epochs and has a slightly bigger Input layer (height=648

width=486) compared to dataset default 640 × 480. The VGG16 encoder of the other 23 class SegNet model was pretrained on ImageNet [1], and the complete model was trained for 180 epochs with the default input resolution of 640 × 480. The reduced 3 class SegNet model was trained with weights of the 23 class SegNet with pretrained VGG16 encoder for additional 5 epochs.

Additionally, class wise IoU can be evaluated: [Un-labeled: 0 (was ignored during training), Lawn: 0.95314635, Obstacle: 0.96455591, Flat: 0.78457187]. Comparing results, the 23 class SegNet model using a pretrained VGG16 encoder with three times the amount of epochs in training provided far better results. After transferring pretrained weights to a new model with simplified output layer, the results after only 5 epochs of additional training (mainly weights of connections to the output layer are being trained) are even better. As additional side effect, length of evaluation is reduced significantly. This leads to higher performance expectations regarding real time application, to be confirmed within the work of this thesis.

Integration of Semantic Segmentation in V-SLAM

After successful training, the semantic segmentation model is ready to be integrated into a framework, similar to the DS-SLAM framework displayed in figure 3. The tracking and loop closing threads are already implemented, as well as some backend threads (dense map creation without the use of Octo-tree maps and local mapping). Mainly the segmentation thread has to be implemented and integrated into the tracking task, with the addition of removing outliers and afterwards continue tracking and creating new keyframes. To integrate the new semantic thread like seen in figure 3, it is important to synchronize with the tracking thread, after the moving consistency checks are done, and the semantic segmentation is complete. The start of the tracking task (extract ORB, moving consistency check, wait for semantic segmentation results) can

be done in parallel to the semantic task. Features of the incoming image will be extracted with ORB (Oriented FAST and Rotated BRIEF) and afterwards checked for dynamic behaviour, mainly with the use of optic flow estimation. After synchronisation, detected

objects with dynamic behaviour can be removed prior to tracking and keyframe decision with the help of the segmented image (see figure 1) created by the segmentation model inside of the semantic thread.

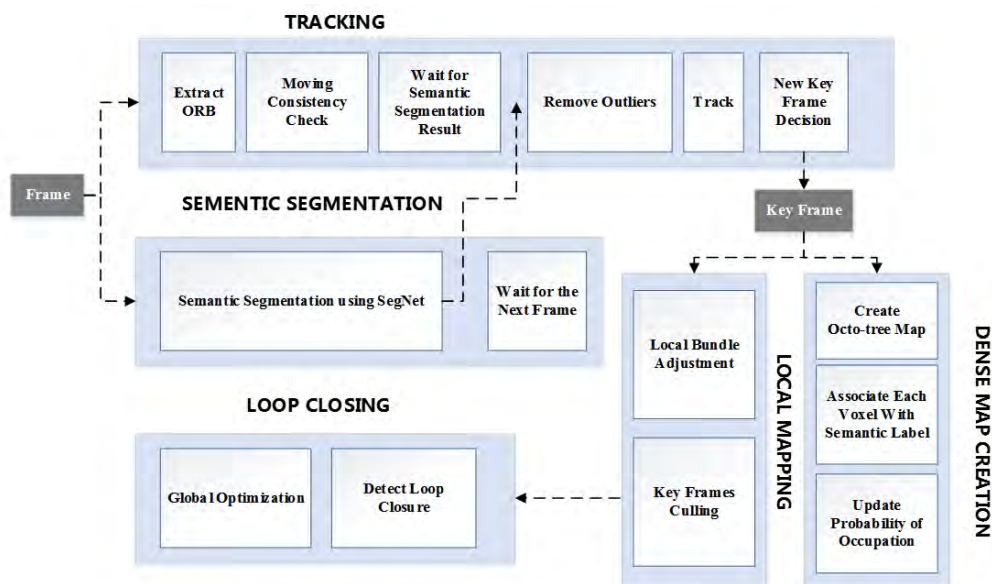


Fig. 3: Overview of the DS-SLAM framework [5]

References and figures

- [1] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. {ImageNet}: A large-scale hierarchical image database. In *2009 {IEEE} Conference on Computer Vision and Pattern Recognition*. IEEE, 2009.
- [2] Divam Gupta. Image Segmentation Keras : Implementation of Segnet, FCN, UNet, PSPNet and other models in Keras. <https://github.com/divamgupta/image-segmentation-keras>, 03 2017.
- [3] Own representation.
- [4] Evan Shelhamer, Jonathan Long, and Trevor Darrell. Fully Convolutional Networks for Semantic Segmentation. *{IEEE} Transactions on Pattern Analysis and Machine Intelligence*, pages 640–651, 2017.
- [5] Chao Yu, Zuxin Liu, Xinjun Liu, Fugui Xie, Yi Yang, Qi Wei, and Qiao Fei. DS-SLAM: A Semantic Visual SLAM towards Dynamic Environments. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, page 1169. IEEE, 2018.

Development, Simulation and Validation Environment for Autonomous Driving Algorithms based on a ROS Architecture

Constantin Blessing

Reiner Marchthaler

Department of Computer Science and Engineering, Esslingen University

Work carried out at Department of Computer Science and Engineering, Esslingen

Motivation

Testing and development for automotive applications can often be quite involved: In many scenarios hardware cannot be utilized by multiple independent teams working on different features, necessitating the requirement for additional copies of the same piece of hardware, which increases costs.

Furthermore, the ongoing trend for AI-based algorithms, which commonly rely on large amounts of training data, demands some kind of virtual environment which can be leveraged to train the employed AI-based algorithms.

Therefore, it makes sense to invest into the development of a virtual environment or simulator/digital twin that best mimics the actual hardware and the use case it is intended to be deployed in.

This investment aims not only to enable effective training of machine learning algorithms but also to alleviate the constraint on hardware availability, and to allow for faster iterative development, testing, and validation.

Task

In the scope of this work, exactly such a development, simulation and validation environment will be implemented for the it:movES group.

The it:movES group consists of a team of master and bachelor students developing and maintaining various model-sized vehicles, each of which features a Robot Operating System-based software stack enabling autonomous driving. The group is supervised by Prof. Dr. Reiner Marchthaler, Prof. Dr. Markus Enzweiler, Prof. Dr. Clemens Klöck and Prof. Dr. Thao Dang [2].

As far as the implementation is concerned, particular attention is paid to the following goals:

1. Standalone desktop application with minimal setup for users
2. Easily extensible and modular architecture
3. Testing and validation of a vehicle's software stack inside the virtual environment
4. Quick iterative prototyping of new vehicle algorithms right inside the development environment
5. Possibility for accessing the simulator from potentially multiple remote systems



Fig. 1: One of the vehicles maintained by the it:movES group. [1]

Approach and Current Status

In order to realize the aforementioned goals, the game engine Unity was chosen as the host of the virtual environment. Unity is responsible for simulating the environment, emulating the vehicle and its driving dynamics, offering configuration of various parameters, and communicating with the actual vehicle software stack. Figure 2 shows the basic architecture of the application.

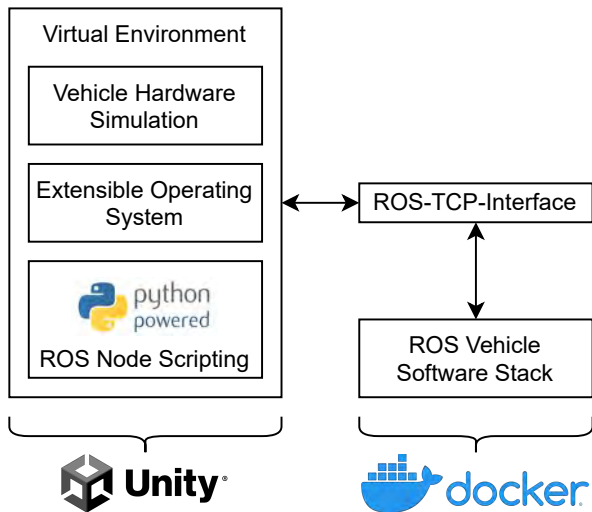


Fig. 2: Architecture of the simulator. [3]

As mentioned, the vehicle's software stack is virtualized inside a docker container. This has two benefits: It allows for automated setup of the required vehicle software stack and its dependencies - which would otherwise be a tedious endeavor - and it ensures that the simulator can run on operating systems that do not support ROS itself. Both benefits are aimed at achieving the first goal.

One noteworthy aspect of the vehicle software stack is that sensor/actuator-specific code needs to be simulator-aware because running inside the docker container means emulating the hardware, therefore entailing different mechanisms to address and control said hardware. This is achieved through conditional compilation of the vehicle software stack.

The second goal is primarily covered by the Unity-based virtual environment: An operating system-like framework, which lets developers integrate new features in an application like fashion (akin to applications in any standard operating system), allows for easy extensibility and modularity while avoiding unnecessary clutter of the user interface of the simulator. Various applications, such as a console, controls for the underlying docker container and the TCP-ROS-Interface, are already available.

The third and fourth goal are intended to be fulfilled

by a Python-based scripting environment embedded into the Unity simulator as an application. The idea is to enable users to quickly prototype new vehicle functionality inside the simulator by writing ROS nodes that do not run inside the docker container: The simulator provides a ROS facade inside the scripting environment through which ROS-like nodes can be written. The simulator will then take care of interfacing the prototype node with the real ROS software stack running inside the docker container upon execution. Through this workflow, changes in nodes do not need to be compiled first. Instead they are instantly reflected inside the simulation environment, allowing for rapid iterative development and testing/validation.



Fig. 3: Unity simulator showing the vehicle driving on a test track and the operating system. [3]

Outlook

At the time of writing this paper, work is still ongoing and not everything laid out has been accomplished yet. This mainly concerns the Python-based scripting environment, of which certain details and the overall feasibility are still uncertain. Also, remote access has not been yet set up.

Finally, the it:movES group is using an existing simulator that, albeit not as broad and powerful in scope, still has a slew of features that are worthwhile to port to the simulator developed over the course of this thesis.

References and figures

- [1] Robert Bosch GmbH. Bosch Future Mobility Challenge. <https://boschfuturemobility.com/>, 2023.
- [2] Anna Ioannidis. Hochschule Esslingen - Kleiner Flitzer ganz groß. <https://www.hs-esslingen.de/hochschule/aktuelles/news/artikel/news/kleine-flitzer-ganz-gross>, 2020.
- [3] Own representation.

Robuste Features für die visuelle Lokalisierung mobiler Robotik im Outdoor-Bereich

Marc Brugger

MarkusENZweiler

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Motivation

Wir Menschen besitzen die außergewöhnliche Fähigkeit, Objekte zu erkennen, Landschaften wahrzunehmen, Muster zu identifizieren und Personen zu unterscheiden. Von Geburt an nehmen wir Informationen mit unseren Augen auf, speichern sie tief in unserem Gehirn ab und können bei Bedarf darauf zugreifen [2]. Doch wie kann eine Maschine, die ein Foto lediglich als eine Matrix von Zahlen wahrnimmt, diese Art von Verständnis erreichen?

Zudem stellt die visuelle Lokalisierung mobiler Roboter im Outdoor-Bereich eine große Herausforderung dar, da die Umgebung oft unstrukturiert und dynamisch ist. Um jedoch präzise und zuverlässige Positionsinformationen zu erhalten, müssen mobile Roboter in der Lage sein, robuste Features zu erkennen und zu verarbeiten. Robuste Features spielen eine entscheidende Rolle bei der Bestimmung der Lage und Orientierung eines Roboters in einer unbekanntem Umgebung und ermöglichen eine präzise Navigation und Kartenbildung.

Klassische Algorithmen zur visuellen Feature Erkennung

Die klassischen Algorithmen folgen einem bestimmten Ablauf:

1. Erkennung (**Detection**): Bei der Erkennung versuchen verschiedene Algorithmen auf unterschiedliche Weise, interessante Punkte auf Bildern zu finden, die

markante Merkmale darstellen. Diese Punkte, auch "Keypoint Features" genannt, können beispielsweise Bergspitzen, Eckpunkte von Gebäuden oder Objekte mit ansprechenden Formen sein [3]. Die Bestimmung solch einzigartiger Merkmale ist entscheidend für einen guten Algorithmus und wird von verschiedenen Faktoren beeinflusst (siehe Abschnitt „Problematiken“).

2. Beschreibung (**Description**): Ein gutes Feature sollte idealerweise auf unterschiedlichen Bildern einer Szene, beispielsweise aus verschiedenen Perspektiven, wiedererkannt werden können. Daher versuchen verschiedene Algorithmen, ein Feature so zu beschreiben, dass es unabhängig von Lichtverhältnissen, Translation, Rotation, Skalierung und anderen Verzerrungen wiedergefunden werden kann. Die Beschreibung erfolgt anhand bestimmter Eigenschaften oder Deskriptoren, die die charakteristischen Merkmale des Features erfassen.

Darüber hinaus findet unabhängig von den spezifischen Algorithmen noch ein weiterer Schritt statt, das sogenannte "Matching". Dabei werden die gefundenen Features von zwei Bildern verglichen, um die entsprechenden Merkmale in beiden Bildern zu identifizieren. Hierbei wird eine Transformation, in Form einer Homographie, berechnet, um die Übereinstimmung zwischen den Features in den beiden Bildern herzustellen. Dies ermöglicht eine Zuordnung der entsprechenden Punkte und eine Verbindung zwischen den Bildern. Abb. 1 zeigt ein beispielhaftes Ergebnis nach der „Detection“, der „Description“ und dem „Matching“.

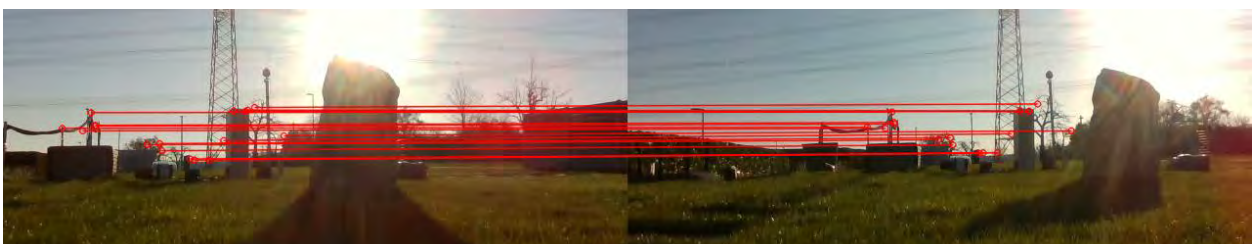


Abb. 1: Feature Keypoint Matching zweier Bilder [1]

Problematiken

Im Gegensatz zu strukturierten Innenräumen bieten Outdoor-Umgebungen eine breite Palette von visuellen Herausforderungen wie variierende Lichtverhältnisse, sich ändernde Wetterbedingungen, unterschiedliche Texturen und komplexe Szenen. Diese Faktoren führen dazu, dass herkömmliche visuelle Features wie Ecken und Kanten in solchen Umgebungen anfällig für Fehler und Verluste sind. Daher ist es von entscheidender Bedeutung, robuste Features zu identifizieren, die unter diesen schwierigen Bedingungen stabil bleiben und genaue Lokalisierungsinformationen liefern können.

Deep Learning Algorithmen

Deep Learning Algorithmen haben bereits in zahlreichen Anwendungen wie Gesichtsfilttern, Übersetzern, Chatbots, Assistenzsystemen, autonomen Autos und Drohnen ihre Bedeutung gefunden. Diese fortschrittlichen Verfahren zielen darauf ab, das Lernen und die Informationsverarbeitung nach dem Vorbild des menschlichen Gehirns zu simulieren. Dabei setzen sie auf mehrschichtige Strukturen, die als neuronale Netzwerke bezeichnet werden (siehe Abb. 2) [4]. Im Gegensatz zu klassischen Algorithmen haben Deep Learning Verfahren den Vorteil, dass sie aus den Fehlern und Limitierungen traditioneller Ansätze lernen und sich kontinuierlich verbessern können. Sie sind darauf ausgerichtet, Robustheit zu erlangen und sich an verschiedenste Szenarien anzupassen.

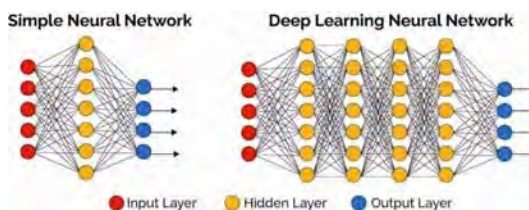


Abb. 2: Darstellungen neuronaler Netzwerke. [4]

Es gibt eine Fülle von Beweisen, die verdeutlichen, dass Deep Learning Algorithmen die klassischen Verfahren in vielerlei Hinsicht deutlich übertreffen. Durch die Nutzung großer Datenmengen und die Anwendung komplexer Modelle sind sie in der Lage, herausragende Leistungen in verschiedenen Kategorien zu erzielen. Sie können beispielsweise komplexe Muster und Zusammenhänge erkennen, sowie Prädiktionen und hochpräzise Entscheidungen treffen.

Ausblick

Im weiteren Verlauf der Arbeit werden eine festgelegte Anzahl an klassischen Algorithmen sowie Deep Learning Verfahren miteinander verglichen. Hierfür ist es essenziell, ein experimentelles Setup zu gestalten, sowie ein System zur präzisen Auswertung zu entwickeln. Besonderes Augenmerk liegt dabei auf Verfahren, die auf quantitativer Ebene miteinander vergleichbar sind, um die Leistung der Algorithmen zu bewerten. Hierbei werden sowohl die Vor- als auch die Nachteile der einzelnen Verfahren eingehend diskutiert. Dieser Vergleich auf quantitativer Ebene ermöglicht es, die Stärken und Schwächen der unterschiedlichen Ansätze besser zu verstehen und fundierte Schlussfolgerungen zu ziehen, um die Leistung der Algorithmen objektiv zu beurteilen und deren Eignung für die visuelle Lokalisierung mobiler Roboter im Outdoor-Bereich zu bewerten.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Peter Reid. The Difference Between Computer Vision and Human Vision. <https://visionaisuite.net/blog/computer-vision/the-difference-between-computer-vision-and-human-vision>, 2022.
- [3] Deepanshu Tyagi. Introduction To Feature Detection And Matching. <https://visionaisuite.net/blog/computer-vision/the-difference-between-computer-vision-and-human-vision>, 2019.
- [4] Laurenz Wuttke. Deep Learning: Definition, Beispiele & Frameworks. <https://datasolut.com/was-ist-deep-learning/>, 2022.

Konzeption und Implementierung einer webbasierten Peer-To-Peer-Infrastruktur für kooperative Virtual Reality Anwendungen

Jeremy Camier

Andreas Rößler

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Einleitung

Multiplayer Virtual Reality-Anwendungen sind normalerweise Anwendungen, die die produzierten Daten von den VR-Headsets eines Spielers über den Server des Anwendungsentwicklers leiten, von dem aus sie zu den VR-Headsets aller anderen teilnehmenden Nutzer verteilt werden. Diese Anwendungen werden üblicherweise vom Nutzer im Playstore des Herstellers heruntergeladen und auf dem internen Speicher des VR-Headsets abgelegt. Werden die Nutzer solcher Anwendungen stattdessen über Peer-To-Peer verbunden, dann werden Daten während der Nutzung direkt vom VR-Headset eines Senders zum VR-Headset eines Empfängers geschickt, ohne über einen zentralen Server geleitet zu werden. Das hat zum einen den Vorteil, dass durch den fehlenden zentralen Server, die Anwendung weniger anfällig für Komplettausfälle ist und zum anderen kann bei eventuellen Angriffen auf das Netzwerk, nicht die Kommunikation aller Teilnehmer auf einmal abgehört werden. Ist die Anwendung zusätzlich webbasiert, kann sie von den Nutzern über den Webbrowser aufgerufen und direkt verwendet werden, wodurch man einfachen und schnellen Zugriff auf direkte Kommunikation mit anderen Anwendern erhält, ohne eine Applikation herunterladen zu müssen.

Zielsetzung

Das Ziel der Arbeit ist es zunächst durch Verwendung von Peer.js, eine funktionierende und fehlerfreie Peer-To-Peer-Verbindung zwischen mehreren Teilnehmern mit VR-Headsets aufzubauen. Als Anwendungsbeispiel soll mit Hilfe von Three.js ein webbasiertes Multiplayer VR-Spiel entwickelt werden, über dessen Verbindungen sowohl Positions- und Rotationsdaten von Spielern, als auch Objekten versendet werden, mit denen interagiert werden kann. Mit Hilfe dieser Daten sollen Empfänger dann die Position anderer Spieler und Objekte berechnen und anzeigen können. Das entwickelte VR-Spiel

soll auf Meta Quest 2 VR-Headsets lauffähig sein und zwei gegen zwei gespielt werden.

Three.js

Three.js ist eine sehr mächtige JavaScript Bibliothek, die verwendet wird, um 3D-Grafiken und Animationen in webbasierten Applikationen zu erstellen. Es bietet viele Möglichkeiten 3D-Objekte zu rendern und zu manipulieren, wie zum Beispiel durch Belichtung, Shading oder Texturen. Außerdem hat es dank seiner einfachen Benutzbarkeit und starken Flexibilität in unterschiedlichen industriellen Bereichen, wie der Spieleentwicklung, architekturellen Visualisierung und auch wissenschaftlicher Simulation Anklang gefunden und erfreut sich steigender Beliebtheit. Es bietet verschiedene Werkzeuge, um VR-Erlebnisse zu kreieren, die durch VR-Headsets oder auf Webbrowsern mit WebXR Unterstützung begutachtet werden können. Entwickler können damit VR-Szenen mit 3D-Objekten erzeugen, in denen man sich durch die virtuelle Umgebung navigieren und dabei mit Objekten interagieren oder Ereignisse auslösen kann. Zusätzlich unterstützt Three.js auch stereoskopisches Rendering für realistische Tiefenwahrnehmung. [1]

Peer.js

Die JavaScript Bibliothek PeerJS macht es möglich verschiedene Peers, also in diesem Fall Nutzer mit VR-Headsets, miteinander zu verbinden, um so einen direkten Datenaustausch zwischen diesen zu ermöglichen. Peer.js bietet einen Server, der eine zufällige PeerID an anfragende Clients vergibt und diese verwaltet, damit keine mehrfach vorhandenen PeerIDs bestehen.

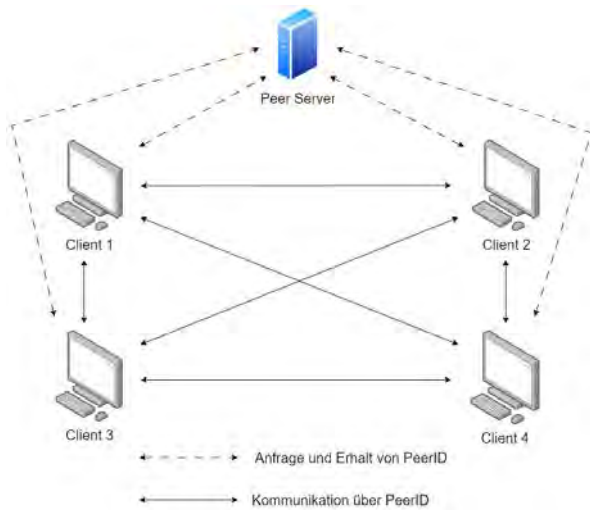


Abb. 1: PeerJS Kommunikation [2]

Dieser Server ist allein dafür zuständig den einzelnen Peers die Information zur Verfügung zu stellen, wie sie sich mit entfernt liegenden Peers verbinden können. Es kann jedoch auch eine selbst gewünschte PeerID vergeben werden, um einen spezifischen Peer immer

über dieselbe ID zu erreichen. Das ist beispielsweise dann nützlich, wenn man einen Client als Host deklarieren möchte, zu dem sich alle anderen Nutzer als erstes verbinden müssen, um dann eine Verbindung zu den restlichen Nutzern im Netzwerk aufbauen zu können. Sobald eine Verbindung zwischen zwei oder auch mehreren Peers besteht, läuft die gesamte Kommunikation nur noch über diese direkten Peer-To-Peer Verbindungen ab. Peer.js bietet mehrere Methoden und Listener, um Verbindungen aufzubauen und Daten zu versenden. Eine P2P Verbindung wird über die "connect" Methode hergestellt, die mit der PeerID des anderen Clients arbeitet. Alle Clients besitzen auch einen Listener der auf eingehende Verbindungsanfragen horcht und diese dann akzeptiert, sofern noch keine Verbindung mit diesem Peer besteht. Sobald die Peer-To-Peer Verbindung zwischen den Headsets steht, wird eine zweite Verbindung hergestellt, über die dann sowohl Daten versendet werden können, als auch auf eingehende Daten gelauscht wird. Über diese Datenverbindung sollen Positionsdaten und Informationen zu Spielern und Objekten mit denen interagiert wird, an alle anderen verbundenen Peers gesendet werden. [3]

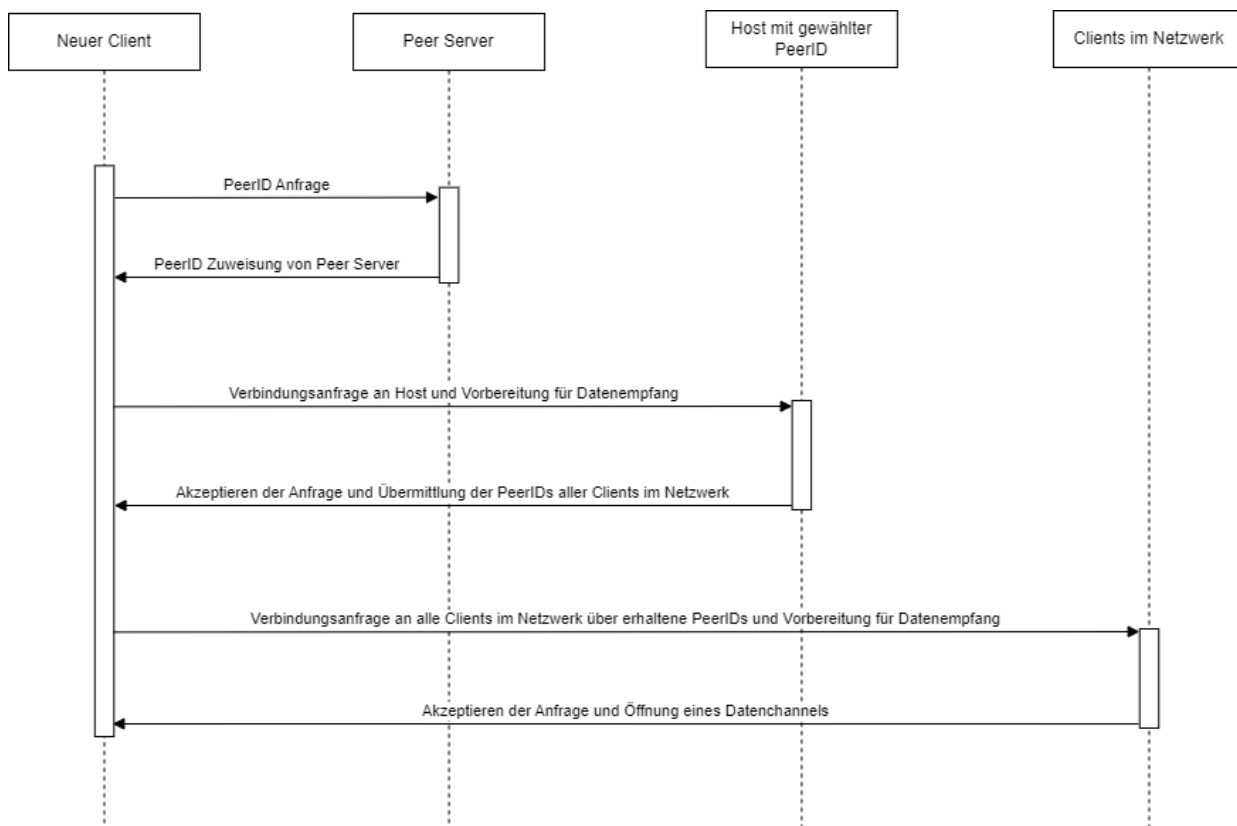


Abb. 2: Aufbau der Datenverbindung zwischen allen Clients des Netzwerks [2]

Problemstellung

Da die Kommunikation zwischen den Clients ohne größere Probleme ablaufen sollte, gibt es einige Herausforderungen, die es zu bewältigen gilt. Es soll in der virtuellen Welt mehrere Objekte geben die von Nutzern gegriffen und geworfen werden können. Dies kann zu Konflikten führen, da das Greifen eines Objekts mit sich bringt, dass kein anderer Spieler das Objekt zum gleichen Zeitpunkt greifen können darf. Weil Datenübertragung jedoch nicht unverzögert abläuft, kann es vorkommen, dass zwei Nutzer zur selben Zeit ein Objekt greifen, bevor beim jeweils anderen Nutzer die Nachricht ankommt, dass es sich bereits in der Kontrolle eines anderen Spielers befindet. Diesen Konflikt gilt es zu lösen. Außerdem steigen mit wachsender Spieleranzahl und somit mehr Datenverbindungen auch die Latenzen, weshalb ein Test durchgeführt werden soll, um auszumachen, was

für eine Last das Peer-To-Peer Netzwerk aushält. Das Ziel ist herauszufinden, ab wann die Nutzung nicht mehr zumutbar ist, da dies zu einer schlechten VR-Erfahrung durch Verzögerungen und eventuell auch damit verbundener Motion Sickness führen kann.

Ausblick

Da das sich für diese Arbeit in Entwicklung befindende Spiel die gleiche Art der Kommunikation nutzt, wie es auch andere Peer-To-Peer VR-Anwendungen könnten, kann der zur Datenübertragung genutzte Programmiercode anschließend mit entsprechenden Anpassungen auch für andere kooperative Anwendungen, wie VR-Socializing oder auch geschäftlich verwendet werden. Es könnten beispielweise ohne großen Aufwand, webbasiert, sichere und direkte Verbindungen in VR zwischen Geschäftspartnern realisiert werden, um virtuell neue Produkte zu präsentieren.

Literatur und Abbildungen

- [1] Ricardo Cabello. Three.js Docs. <https://threejs.org/docs/>, 2010.
- [2] Eigene Darstellung.
- [3] Eric Zhang and Michelle Bu. The PeerJS library. <https://peerjs.com/>, 2018.

Wohin führt uns die künstliche Intelligenz? Eine Analyse der Auswirkung von KI auf Wirtschaft, Gesellschaft und Ethik

Irem Nur Cankaya

Dirk Hesse

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Einleitung

In einer sich ständig verändernden Welt, die immer komplexer und digitaler wird, hat sich die künstliche Intelligenz (KI) zur zentralen Zukunftstechnologie entwickelt. Sie hat in zahlreichen Bereichen tiefgreifende Auswirkungen auf unser Leben und verändert bereits jetzt unser Verständnis von Arbeit, Gesellschaft und Ethik. Spätestens mit ChatGPT ist KI im Alltag angekommen. Doch wohin führt uns die KI wirklich? Welche Auswirkungen wird sie auf verschiedene Aspekte unseres Lebens haben?

Zielsetzung der Arbeit

Diese Bachelorarbeit befasst sich mit dem Thema künstlichen Intelligenz und die damit verbundenen wirtschaftlichen, gesellschaftlichen und ethischen Auswirkungen. Ziel dieser Arbeit ist es, herauszufinden, wie die künstliche Intelligenz diese drei Bereiche beeinflusst und wie sich diese in der Zukunft auswirken. Dabei wird insbesondere der wirtschaftliche Teil in den Fokus genommen.

KI und Datenanalyse

KI kann für eine intensive Datenanalyse genutzt werden, die zu völlig neuen Geschäftsmodellen führen kann, da sie neue Informationen generiert, die die Entwicklung eines Unternehmens schnell vorantreiben können.

Die Geschwindigkeit der Datenanalyse ist ebenfalls sehr wichtig, da sie zur schnellen Entwicklung neuer wettbewerbsfähiger Produkte genutzt werden kann. Es ist anzunehmen, dass die Vernetzung zwischen Menschen und Computer immer leistungsfähiger wird. Computer sind auch zunehmend in der Lage, menschliche Leistungen zu erfassen. Gleichzeitig helfen sie aber auch vielen Menschen, mehr Spaß und Erfolg bei ihrer Arbeit zu haben, weil ihnen die Arbeit erleichtert wird. Alles wird mit immer höherer Genauigkeit aufgezeichnet. Aber die Menschen sind sehr unterschiedlich und lassen sich nicht in willkürliche Muster pressen. Insofern müssen die Ergebnisse solcher Studien immer unter dem Aspekt betrachtet werden, wie aussagekräftig sie sind oder sein können [2].

Umfrage zur Zukunft mit KI

Eine Studie, Münchner Kreis Band VIII, befasste sich mit dem Thema der KI in der Zukunft. Die spannende Frage war, wie sich unsere Wirtschaft und Gesellschaft im Jahr 2035 durch eine KI-gelenkte Welt verändern wird. Die These wurde mit Hilfe von nationalen und internationalen Experten aus dem Bereich Digitalisierung und Technologie kritisch betrachtet [3].

These: Es wächst eine Generation auf, die völlig selbstverständlich und intuitiv KI-Systeme als integralen Bestandteil ihres täglichen Lebens verwendet (d.h. 24 Stunden am Tag, 7 Tage die Woche).



Abb. 1: Alltägliche Nutzung der KI-Systeme [1]

Mit dieser Befragung wird im Diagramm ersichtlich (siehe Abb. 1), dass 50 Prozent der Befragten glauben, dass es bis 2030 eine Generation gibt, die ohne Probleme täglich mit KI-Systemen auskommen können.

Darüber hinaus schätzen 47 Prozent die Realisierung nach 2030 ein und nur zwei Prozent weisen diese These zurück [3].



Abb. 2: Auswirkung durch alltägliche Nutzung der KI-Systeme [3]

Vor allem die Auswirkungen auf die Wirtschaft werden sehr positiv bewertet (siehe Abb. 2). Führungspersonen im Wirtschaftsbereich sollten danach streben, sich als positive Vorbilder in Bezug auf die Gestaltung und den Gebrauch von KI-Technologien zu bemühen. Sie sollten sowohl die positiven als auch die negativen Folgen dieser Technologie eindeutig aufzeigen und diskutieren.

Die Wirtschaft hat die Chance, durch transparent gestaltete Produkte und Services das Vertrauen der Bürger zu erlangen [3]. Ein bewusster Einsatz dieser Technologie ist essenziell, um sowohl soziale als auch ethische Konsequenzen besser abschätzen und proaktiv handhaben zu können. Des Weiteren sollte jedem Bürger die Möglichkeit gege-

ben werden, sich eingehend mit Künstlicher Intelligenz auseinandersetzen. Besonders für Schüler ist es von zentraler Bedeutung, den Umgang mit den neuesten Technologien zu erlernen und mit ihnen zu interagieren, jedoch immer auch kritisch zu hinterfragen [3].

Ausblick

Nach der Analyse der Auswirkungen von KI auf Wirtschaft und Gesellschaft, werden ethische Fragen beantwortet, um sicherzustellen, dass KI zum Wohl der Menschheit eingesetzt wird. Die Arbeit zielt darauf ab, die Chancen und Herausforderungen im Umgang mit KI besser zu verstehen.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Arnold Kitzmann. *Künstliche Intelligenz Wie verändert sich unsere Zukunft*. Springer, 2022.
- [3] Philipp Ramin, Felix Rothmund, Dennis Schmedt, Constanze Stedele, Verena Till, Nora Treiber-Lobenstein, Bernd Wiemann, Birgit Wintermann, Ole Wintermann, Malthe Wolf, and Petra Wolf. *Leben, Arbeit, Bildung 2035+ Durch Künstliche Intelligenz beeinflusste Veränderungen in zentralen Lebensbereichen*. <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/zukunftsstudie-leben-arbeit-bildung-2035#detail-content-6846-4>, 06 2020.

Konzeptionelle Entwicklung einer One Service Data Plattform im Bereich der Messanalyse mit Big und Small Data im industriellen Umfeld

Lars-Hendrik Crone

Mirko Sonntag

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Robert Bosch GmbH, Schwieberdingen

Einleitung und Problemstellung

Seit vielen Jahren befindet sich unsere Gesellschaft und die damit einhergehende Arbeitswelt in einem stetigen Wandel. Eine umfassende Digitalisierung sowie zukunftsorientierte Denkweisen sind ausschlaggebend für den Erfolg vieler Unternehmen. Die vierte industrielle Revolution, auch besser bekannt als Industrie 4.0, ist hochaktuell und bestimmt die direkte sowie intelligente Vernetzung zwischen Mensch, Maschine und Produkten mit Hilfe von Informations- und Kommunikationstechnik. [3] Einen essentiellen Bestandteil der Industrie 4.0 beschreibt der „Einsatz von Daten“ [3], um gegebenenfalls bestehende Prozesse optimieren oder offene Potentiale erschließen zu können. Hierzu werden die Daten in jeglichen Bereichen gesammelt, analysiert und evaluiert. Hierdurch erschließt sich die Möglichkeit bisher unbekannte Zusammenhänge in den Datenstrukturen zu entdecken. Dies ist gleichbedeutend mit einem Gewinn an potentiell wertvollem Wissen. Gerade deswegen sind Daten mittlerweile

ein wichtiges Hab und Gut vieler Firmen und bilden die Grundlage für Fortschritt und Verbesserung. Es werden mehr Daten denn je erfasst, gespeichert und ausgewertet. [5] Ein Begriff, welcher immer wieder im Zusammenhang mit dem stetigen Fortschritt in der Datenanalyse genannt wird und immer mehr an Bedeutung gewinnt, ist „Big Data“. Dieser umfasst Daten, welche in großer Vielfalt vorliegen, in ihrer Menge groß sind und in einer hohen Geschwindigkeit anfallen. [6] Die Entwicklung hin zu Big Data bringt dabei einige Herausforderungen mit sich. Es müssen Wege gefunden und Plattformen entwickelt werden, in denen mit den Mengen und Charakteristiken von Big Data umgegangen werden kann. Dies umfasst nicht minder die komplette Handelskette im Umgang mit Daten von der Erfassung, der Prozessierung, der Bereitstellung bis hin zur Konsumierung. Die einzelnen Schritte des Datenzyklus, welche auch bei der Robert Bosch GmbH umgesetzt werden müssen, sind in Abbildung 1 visualisiert.

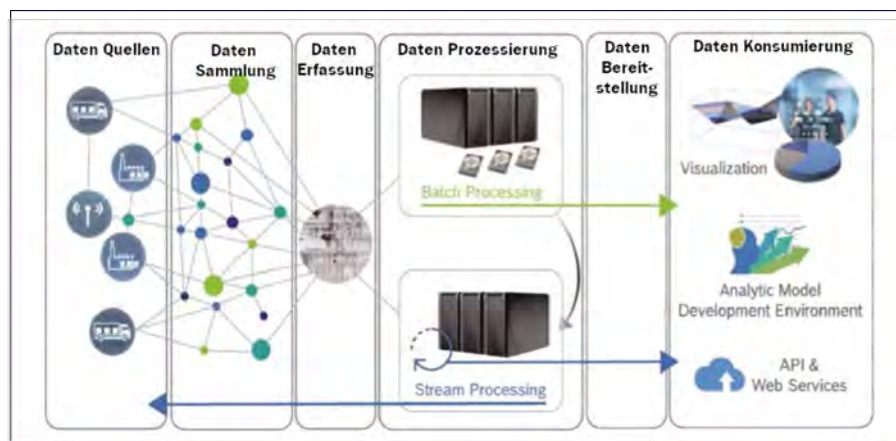


Abb. 1: Die Handelskette im Umgang mit Daten [4]

Als zukunftsorientiertes Unternehmen und weltweit größter Automobilzulieferer hat Bosch eine dementsprechend große Verantwortung, gerade weil die Frage nach den Fahrzeugantrieben der Zukunft noch nicht beantwortet ist. Dabei ist auch die Parametrierung von Steuergeräten und die damit einhergehende Messanalyse ein wichtiger Bestandteil. Hierbei steigt die Anzahl der nötigen Validierungsmessungen bei einer sinkenden Anzahl von verfügbaren Versuchsfahrzeugen. Des Weiteren steigt die Anzahl zusätzlich benötigter Simulationsmessungen. Das Prozessieren der immensen Datenmengen und den damit einhergehenden Algorithmen wird früher oder später nicht mehr durch Desktop-Anwendungen auf lokalen Rechnern bewältigt werden können. Dennoch kennzeichnen sich heutige Tätigkeiten im Bereich der Steuergeräte oftmals durch langwierige Prozessketten, bei welchen eine lange Abfolge kleinerer sich wiederholender Tätigkeiten erfolgen muss, um eine größere Aufgabe abschließen zu können. Die Tätigkeiten beinhalten insbesondere die Verwendung verschiedener Desktop-Anwendungen. Somit werden die Aufgaben nacheinander auf dem lokalen Rechner eines Mitarbeitenden erledigt. Dabei liegen auch keine systemübergreifenden Automatisierungen vor. Stattdessen müssen Daten aus einem System abgefragt, in ein anderes geladen, prozessiert, die Ergebnisse exportiert und in ein weiteres System geladen werden, damit sie weiter prozessiert werden können. Es muss somit ein Weg gefunden werden den Mitarbeitenden im Automobilbereich ihre benötigten Werkzeuge integriert, online und zentral zur Verfügung zu stellen. Es gilt somit Visualisierungen, Konfigurationen und das Steuern von Prozessierungen für Webanwendungen im Zusammenhang mit Small und Big Data zu ermöglichen. Dabei müssen Lösungen für User basierte Interaktionen und Steuerungen in der serverbasierten Arbeitsweise gefunden werden.

Ziel der Arbeit

Im Rahmen der Arbeit soll ein Konzept für eine One Service Plattform entstehen, die den Arbeitenden alle benötigten digitalen Arbeitswerkzeuge zentral zur Verfügung stellt. Über diese Plattform sollen alle benötigten Methoden, Tools, Messungen, Hilfsmittel oder Equipment für die Versuchsträger erreichbar sein. Dennoch muss die Plattform auf diverse in dieser Arbeit zu identifizierende Anwendungsfälle zugeschnitten und entsprechend personalisierbar sein. Das erstellte Konzept soll über die Erstellung eines Proof of Concept (PoC) für den Bereich der Messanalyse abgesichert werden. Ein essentieller Punkt für das Konzept ist dabei die Visualisierung und Bereitstellung der Messungen mit den einhergehenden Funktionen für die Nutzenden. Das PoC soll somit neben der Konzepterstellung auch die Implementierung eines vollumfänglichen Prototypen

in Form einer Webanwendung für die Verifizierung und die Ableitung weiterführender Tätigkeiten umfassen. Im PoC soll dabei ein Weg gefunden werden die bestehenden Systeme und Funktionen sinnvoll zu nutzen und zu verknüpfen, sodass verschiedenste Use Cases in der Messanalyse umgesetzt werden können.

Vorgehen

Das Vorgehen der Arbeit richtet sich nach dem Entwicklungsframework des hybriden Scrum. Dabei lassen sich die Schritte der Arbeit in die Anforderungsanalyse, das Grob- und Feindesign, den Übergang in die agile Entwicklungsphase, sowie die Test- und Integrationsphase unterteilen. Dementsprechend werden zunächst die relevanten Stakeholder definiert sowie die Use Cases mit höchster Priorität gesammelt. Mit den Stakeholdern werden Mockups für die entsprechenden Use Cases gezeichnet. Für die agile Entwicklung und das Erstellen eines Backlogs werden aus den Use Cases die zugehörigen User Stories abgeleitet. Anschließend wird die Konzeptfindung auf Basis der identifizierten Anforderungen in einem iterativen Prozess vorangetrieben. Über ein Grobkonzept geht der Weg hin zum Feinkonzept mit Mockups für die One Service Plattform und der Einordnung in die technische Infrastruktur eines Großkonzernes mit Auswahl der Technologien. Darauffolgend müssen die relevanten Punkte für das PoC festgelegt und umgesetzt werden. Daraus soll eine Offene Punkte Liste für eine vollumfängliche Umsetzung des Konzeptes in der Zukunft resultieren und eine Bewertung des Konzeptes erfolgen.

Entwicklungstechnologien

Im Zentrum der Implementierung steht der Prototyp der One Service Plattform. Hierbei steht eine Webanwendung der Form **Single Page Application (SPA)**. SPAs unterscheiden sich zu herkömmlichen Webanwendungen, bei denen der Browser bei einer Eingabe durch die Nutzenden jeweils eine neue HTML-Seite vom Server lädt, diese darstellt und den Inhalt bei neuer Interaktion wieder verwirft dahingegen, dass sie lediglich aus einem HTML-Element bestehen, welches dynamisch erweitert wird. Es werden somit zu Beginn auch Elemente und Funktionen geladen, die eventuell nicht direkt sichtbar sind. Bei einer Interaktion wird lediglich der geforderte Teil der Webseite angepasst und der Rest bleibt bestehen. Die statischen Elemente sowie bereits vorhandene Logik werden dabei als Hülle bezeichnet. Typische Elemente dafür sind der Header, Footer oder die Seitennavigation. Innerhalb der Hülle können dann verschiedene Ansichten dargestellt werden. [7] [8] Dieser Aufbau wird in Abbildung 2 visualisiert. Dabei ist die Hülle links hinterlegt, die einzelnen Ansichten finden sich rechts. Die Views werden dabei

in der Regel über das Konzept des Routings dargestellt, bei welchem Ansichten bestimmten URL-Mustern zugeordnet werden. In der Abbildung wird dies über den dunkelblauen Pfeil dargestellt. Eine Ansicht kann in sich noch weiter unterteilt werden, wie beispielsweise in Regionen, was ein weiteres Routing zur Folge hätte. In der Abbildung ist dies über den hellblauen Pfeil ersichtlich. Aus den genannten Besonderheiten der SPA ergeben sich auch die Vorteile der Nutzung. Diese liegen insbesondere in einem flüssigen Nutzungserlebnis sowie der Reduzierung von Anfragen an den Server. [7]

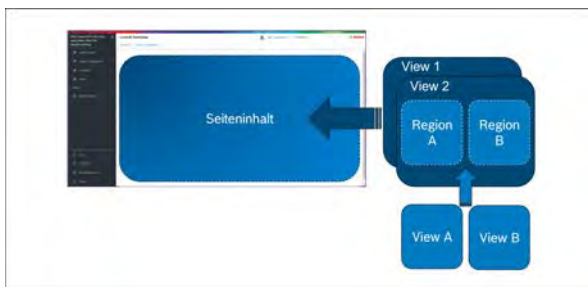


Abb. 2: Konzeptioneller Aufbau einer SPA [8]

Für das Frontend der SPA wird das Framework **Angular** verwendet, welches auf der Sprache TypeScript basiert und als Modul in der Laufzeitumgebung Node.js verfügbar ist. Da ein Webbrowser allerdings kein TypeScript interpretieren kann, wandelt der TypeScript-Compiler den Code automatisch in JavaScript-Code um. Angular zeichnet sich durch einen modularen Aufbau aus Komponenten, Services und weitere Klassentypen aus. Die Komponenten sind wiederverwendbar und können zusammengesetzt komplexe Anwendungen realisieren. Eine Komponente setzt sich aus einem HTML-Teil, einem TypeScript-Teil und einem oder mehreren CSS-Teilen oder SCSS-Teilen zusammen. Dabei erzwingt Angular die Entwicklung in einem MVC-Entwurfsmuster, wie in Abbildung 3 dargestellt. Für das dynamische Laden der Inhalte stellt Angular ein Routing-Modul bereit. [2]

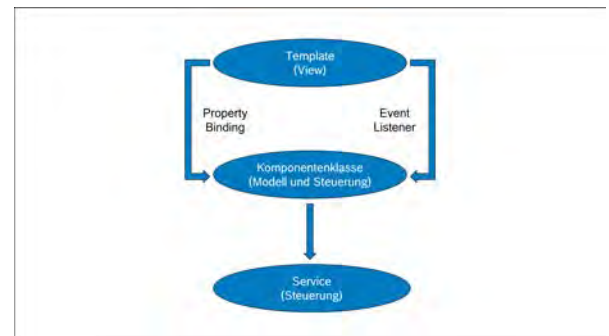


Abb. 3: MVC in Angular [4]

Das Backend wird ebenfalls in der Umgebung Node.js über **NestJS** und dem Object Relational Mapper (ORM) **TypeORM** über TypeScript realisiert. Somit können direkt über TypeScript relationale Datenbanken (DB) konfiguriert und mit diesen interagiert werden, ohne SQL-Abfragen schreiben zu müssen. ORMs im Allgemeinen bezeichnen eine Programmieretechnik, welche eine Interaktion zwischen DBs und objektorientierten Programmiersprachen ermöglicht. Als DB wird in dieser Arbeit auf eine Oracle DB zurückgegriffen. [1] NestJS wird hingegen als Framework für den Aufbau des Backends und somit auch für die Erstellung der Schnittstellen für das Frontend genutzt. Über NestJS kann TypeORM als Paket integriert werden.

Ausblick

Im Augenblick findet die Umsetzung des Prototypen, sowie weiterer Bestandteile für die Validierung des Konzeptes statt. Über das PoC werden alle Pfade des entwickelten Konzeptes abgedeckt werden. Es zeigt sich bereits eine hohe Akzeptanz und Notwendigkeit für das Ergebnis der Arbeit. Es entstehen immer weitere Use Cases für die die One Service Plattform eine große Relevanz hat. Das PoC stellt die Grundlage dar, um die nächsten Entwicklungsschritte zu planen und eine voll funktionsfähige Gesamtarchitektur zu entwerfen. Ebenfalls können die ersten Erkenntnisse über die Machbarkeit und die User Experience gewonnen werden.

Literatur und Abbildungen

- [1] Ihechikara Vincent Abba. What is an ORM – The Meaning of Object Relational Mapping Database Tools. <https://www.freecodecamp.org/news/what-is-an-orm-the-meaning-of-object-relational-mapping-database-tools/>, 2022.
- [2] . Angular. Angular Docs. <https://angular.io/docs>, 2023.
- [3] Wirtschaftu.Energie Bundesministerium. Was ist Industrie 4.0? Menschen, Maschinen und Produkte sind direkt miteinander vernetzt: die vierte industrielle Revolution hat begonnen. <https://www.plattform-i40.de/IP/Navigation/DE/Industrie40/WasIndustrie40/was-ist-industrie-40.html>, 2023.
- [4] Eigene Darstellung.
- [5] Stephen Kaisler, Frank Armour, J. Alberto Espinosa, and William Money. Big Data: Issues and Challenges Moving Forward. In *46th Hawaii International Conference on System Sciences*. IEEE, 2013.
- [6] . Oracle. Was versteht man unter Big Data? <https://www.oracle.com/de/big-data/what-is-big-data/>, 2023.
- [7] Jan Schulte. Single Page Applications (SPA): Erklärung, Vorteile und Beispiele. https://www.magnolia-cms.com/de_DE/blog/alles-wissenswertes-ueber-single-page-applications.html, 2021.
- [8] Emmit A. Scott. *SPA Design and Architecture Understanding Single-Page Web Applications*. Manning, 2015.

Untersuchung von Containern zur Entwicklung einer Build-Umgebung für Embedded Systems

Dimitar Dimov

Markus Enzweiler

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Robert Bosch GmbH, Schwieberdingen

Einleitung

Die kontinuierliche Weiterentwicklung der digitalen Technologie hat einen großen Einfluss auf die Softwareentwicklung, -bereitstellung und -verwaltung. In den letzten Jahren hat sich die Containertechnologie als bedeutsames Werkzeug für Unternehmen und Entwickler weltweit etabliert. Die Containerisierung ermöglicht der Entwickler, Software in isolierten und transportierbaren Einheiten zu verpacken und auszuführen.

Ziel der Arbeit

Diese Arbeit beschäftigt sich mit der Containerisierung einer Build-Umgebung für Embedded Systems. Ziel der Arbeit ist es dabei einen Überblick über die Grundlagen wie Virtuelle Maschinen und Containern zu schaffen, sowie die verwendeten Technologien wie Docker und Kubernetes genauer zu betrachten. Entsprechend wird die Erstellung von Dockerfile, Docker Image und Docker Container, die Verwendung von GitHub Actions hinsichtlich der Continuous Integration/ Continuous Deployment und die Anwendung und das Verhalten von Containern in einer lokalen und Cloud basierten Umgebung, wie zum Beispiel Openshift, erklärt.

Virtuelle Maschinen vs. Container

Virtuelle Maschinen sind Software, die eine virtuelle Umgebung innerhalb eines physischen Computers erstellen kann. Diese virtuelle Umgebung ist in der Lage, Hardware zu simulieren und eine komplette Hardwareumgebung nachzuahmen, einschließlich CPU, Arbeitsspeicher, Speicher und Betriebssysteme. Mit dieser Technologie können Anwendungen so ausgeführt werden, als würden sie auf einem separaten Computer laufen. Ein Container ist eine Virtualisierungstechnologie, die eine isolierte Umgebung für die Ausführung von Anwendungen und deren Abhängigkeiten erstellen. Im Gegensatz zu virtuellen Maschinen, die eine vollständige virtuelle Hardwareumgebung simulieren, nutzen Container die gleichen Host-Betriebssystemressourcen und den gleichen Kernel. Diese gemeinsame Nutzung von Ressourcen macht Container leichter und schneller als virtuelle Maschinen. Jeder Container verfügt über eigene Dateisysteme, Netzwerkschnittstellen und Prozessoren, die jedoch alle innerhalb des Host-Betriebssystems ausgeführt werden. [2]

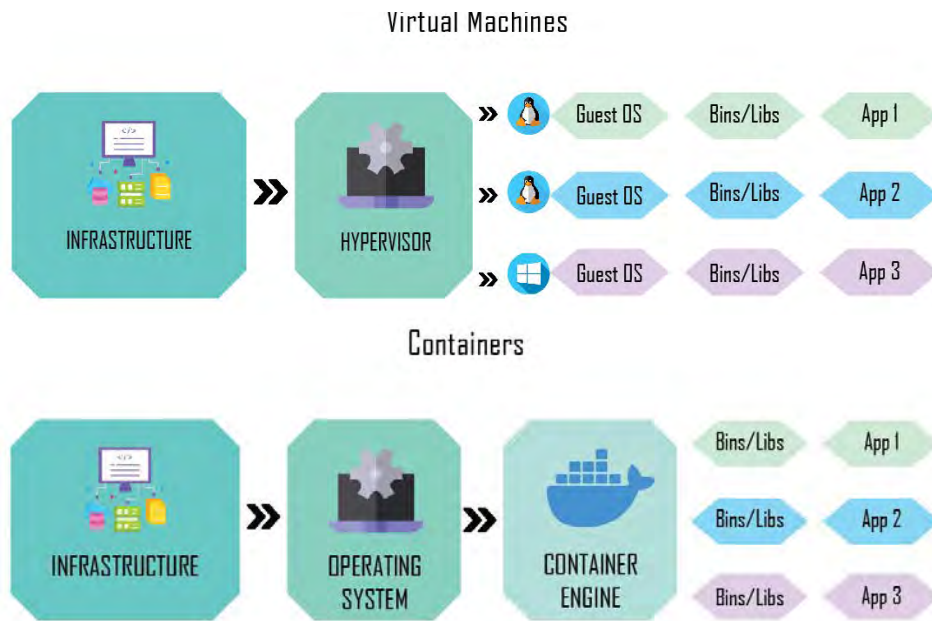


Abb. 1: Virtuelle Maschine vs. Container [1]

Docker

Im Mittelpunkt des Konzepts von Docker steht ein Open-Source-Tool, das die Virtualisierung von Betriebssystemen ermöglicht und die Ausführung von Anwendungen darauf erlaubt. Dieses Tool arbeitet auf dem Kernel des spezifischen Betriebssystems, auf dem es installiert ist, und nutzt die „runc“-Schnittstelle, um die Kommunikation mit diesem Kernel herzustellen. Der Hauptzweck von Docker besteht darin, die verschiedenen Phasen der Anwendungsentwicklung, -bereitstellung und -ausführung durch den Einsatz von Containertechnologie zu vereinfachen. Durch die Bündelung aller notwendigen Komponenten, die eine Anwendung benötigt, optimiert das Tool den gesamten Prozess, ohne sich dabei auf Bibliotheken, Datenbanken, Treiber und Konfigurationsdateien zu beschränken.

Problemstellung und Lösungsansatz

Aufgrund der aktuellen Entwicklung in der IT-Welt werden immer neuere und modernere Anwendun-

gen benötigt, um die großen Herausforderungen zu bewältigen. Dabei müssen die Entwickler viel Zeit investieren, da die Installation und Konfiguration von den neuen Tools Schwierigkeiten bereiten. Öfter werden konkrete Versionen vorausgesetzt, damit die Umgebung konfliktfrei ausgeführt werden kann. Mit Hilfe von Docker können die Entwickler Docker Containern erstellen, die die benötigten Werkzeuge für die konkreten Voraussetzungen enthalten. Diese werden dementsprechend auf einen Registry hochgeladen, um die Verwendung für andere Entwickler zu vereinfachen.

Ausblick

Sicherheit und Schnelligkeit sind in der heutigen Welt ein wichtiger Faktor. Mit der Implementierung von Docker in einem laufenden Projekt sollen die Entwicklungs-, Test- und Deploymentzeiten verkürzt werden. Im weiteren Verlauf dieser wissenschaftlichen Arbeit soll ein Vergleich zwischen der nativen Umgebung und der Docker Umgebung erfolgen.

Literatur und Abbildungen

[1] Eigene Darstellung.

[2] Hans-M. Hopp. *Docker Handbuch für Einsteiger: Der leichte Weg Zum Docker-Experten*. BMU Verlag, 2021.

Data Analytic Tools im Controlling am Beispiel der Prognose

Christian Duerr

Catharina Kriegbaum-Kling

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Einleitung

Die zunehmend steigende Verfügbarkeit großer Datenmengen und die Fortschritte im Bereich der Machine-Learning-Algorithmen haben das Potenzial, das Controlling maßgeblich zu verändern. In diesem Artikel wird ein Data-Analytic-Tool vorgestellt, das auf R basiert und verschiedene Algorithmen zur Prognose von zukünftigen Umsatzentwicklungen vergleicht. Dabei liegt das Hauptaugenmerk auf der Qualität der prognostizierten Umsatzentwicklungen sowie der Skalierbarkeit in Relation zur Anzahl der Datensätze. Effektives Controlling erfordert präzise Prognosen, um fundierte Entscheidungen treffen zu können. Traditionelle Ansätze stoßen jedoch oft an ihre Grenzen, insbesondere bei der Analyse großer und komplexer Datenmengen. Data Analytics und Machine-Learning-Algorithmen bieten hier neue Lösungsansätze, um große Datenmengen in Echtzeit zu analysieren und Muster zu erkennen, die dem menschlichen Auge sonst verborgen bleiben.

Zielsetzung der Arbeit

Das Ziel der Bachelorarbeit besteht analog darin, ein effektives Data-Analytic-Tool in R zu entwickeln, das Unternehmen bei der Generierung präziser Umsatzprognosen unterstützt und somit ihre Entscheidungsfindung verbessert. Gleichzeitig sollen dabei die Grenzen und Einsatzszenarien verschiedener Machine-Learning-Algorithmen eruiert werden.

Vorgehen

Im ersten Schritt werden verschiedene Machine-Learning-Algorithmen genauer betrachtet und auf ihre Anwendbarkeit im Kontext des Controllings untersucht. Die ausgewählten Algorithmen umfassen unter anderem Lineare Regression, Entscheidungsbäume, sowie den Random Forest und Gradient Boosting Algorithmus. Das Prinzip hinter Letzterem lässt sich in Abbildung 1 nachvollziehen. Diese Algorithmen haben sich in der Vergangenheit bereits bei der Vorhersage von Umsatzentwicklungen bewährt und werden in der Literatur als vielversprechend beschrieben [1].

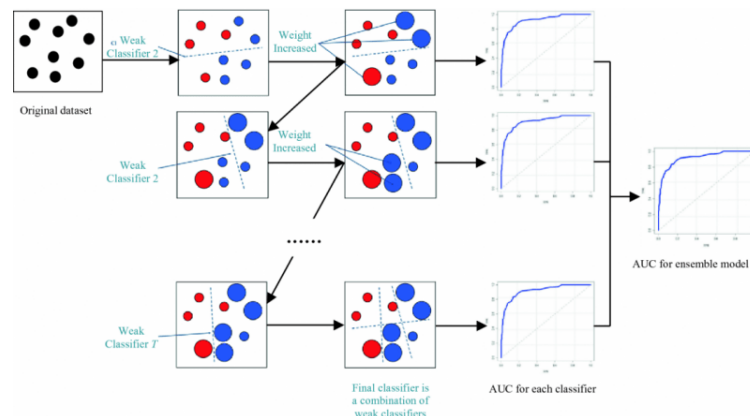


Abb. 1: Prinzip Gradientenverstärkung [2]

1

Ein weiterer Aspekt, der in der Untersuchung berücksichtigt wird, ist die Skalierbarkeit des Data-Analytic-Tools. Angesichts des stetigen Wachstums der Datenmengen ist es entscheidend, dass das Tool auch bei großen Datensätzen effizient arbeitet und aussagekräftige Ergebnisse liefert. Daher wird die Leistungsfähigkeit der einzelnen Algorithmen hinsichtlich der Skalierbarkeit geprüft, bewertet und gegenübergestellt.

Um die Leistung der Algorithmen zu vergleichen, werden sie auf umfangreiche Datensätze angewendet, die speziell für diese Untersuchung erstellt wurden. Die Daten basieren auf realistischen Szenarien und umfassen Informationen wie Alter, Geschlecht, Produktkategorie, Preis und Zeitstempel. Durch die Verwendung möglichst realistischer Daten können aussagekräftige Ergebnisse erzielt und der Mehrwert des Data-Analytic-Tools im Controlling umfassend untersucht werden.

Das entwickelte Data-Analytic-Tool nutzt darüber hinaus die vielfältigen Funktionen von R, um eine effektive Implementierung der ausgewählten Machine-Learning-Algorithmen zu ermöglichen. R bietet eine breite Palette von Bibliotheken und Paketen, die speziell für die Datenanalyse und Modellierung entwickelt wurden. Diese Bibliotheken erleichtern nicht nur die Implementierung der Algorithmen, sondern bieten auch umfangreiche Möglichkeiten zur Datenvisualisierung und -interpretation.

Die Architektur des Tools basiert auf einem modularen Ansatz, der es ermöglicht, verschiedene Algorithmen nahtlos zu integrieren und anzuwenden. Das Tool verfügt über eine intuitive und benutzerfreundliche Schnittstelle, die es den Benutzern ermöglicht, die Eingabeparameter anzupassen und so die Analyseprozesse zu steuern. Die Ergebnisse der Umsatzprognosen werden in übersichtlichen Grafiken und Diagrammen präsentiert (siehe Abbildung 2), um den Benutzern eine leicht verständliche Interpretation zu ermöglichen.

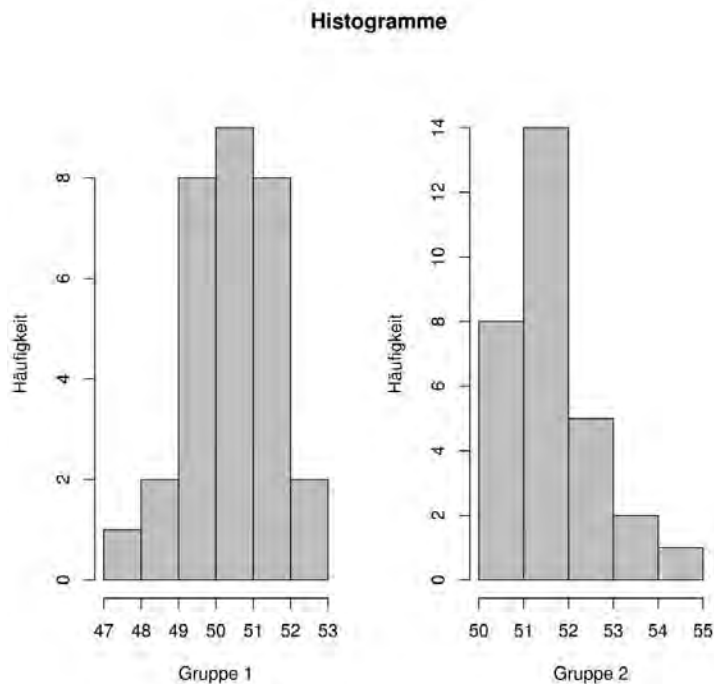


Abb. 2: Exemplarische grafische Ausgabe in R [3]

2

Ein wichtiger Schwerpunkt des entwickelten Data-Analytic-Tools liegt auf der kontinuierlichen Verbesserung der Prognosegenauigkeit. Durch den Einsatz von Machine-Learning-Algorithmen können Unternehmen ihre Umsatzentwicklungen präzise vorhersagen und so wertvolle Einblicke in zukünftige Geschäftstrends gewinnen. Das Tool bietet zudem die Möglichkeit, die

Leistung der Algorithmen regelmäßig zu überprüfen und gegebenenfalls anzupassen, um optimale Ergebnisse zu erzielen. Dieser iterative Ansatz ermöglicht es Unternehmen hierbei, ihre Prognosemodelle kontinuierlich zu optimieren und auf dem neuesten Stand zu halten, um fundierte Entscheidungen zu treffen und ihre Wettbewerbsfähigkeit zu steigern [4].

Auswertung

Um die Leistung des Tools auszuwerten, werden umfangreiche Versuche durchgeführt. Dabei werden verschiedene Metriken wie Genauigkeit, Präzision, Recall und F1-Score verwendet, um die Qualität der prognostizierten Umsatzentwicklungen zu bewerten. Darüber hinaus werden Laufzeitmessungen durchgeführt, um die Skalierbarkeit des Tools in Bezug auf die Größe der Datensätze zu überprüfen. Diese Evaluierung ermöglicht es, Stärken und Schwächen der einzelnen Algorithmen aufzudecken und eine fundierte Entscheidung bei der Auswahl des besten Ansatzes zu treffen.

Ein weiterer wichtiger Aspekt der Auswertung ist der Vergleich der prognostizierten Umsatzentwicklungen mit den tatsächlichen Werten. Hierfür werden historische Umsatzdaten verwendet, um die Vorhersagen des Tools mit den realen Umsatzwerten zu vergleichen. Dieser Vergleich ermöglicht es, die Genauigkeit und Zuverlässigkeit der prognostizierten Ergebnisse zu bewerten und Rückschlüsse auf die praktische Anwendbarkeit des Tools zu ziehen.

Ausblick

Die vorläufigen Ergebnisse deuten darauf hin, dass sowohl der Random Forest Algorithmus als auch der

Gradient Boosting Algorithmus vielversprechende Prognoseergebnisse liefern. Diese Ergebnisse sind jedoch vorläufig und müssen in weiteren Untersuchungen validiert werden, um ihre Zuverlässigkeit und Genauigkeit zu bestätigen. Es ist wichtig anzumerken, dass die Entwicklung und Auswertung des Tools zum Zeitpunkt des Verfassens dieses Artikels noch nicht abgeschlossen war. Die vorläufigen Ergebnisse zeigen jedoch vielversprechende Leistungen der implementierten Algorithmen und deuten darauf hin, dass das Data-Analytic-Tool einen wertvollen Beitrag zum Controlling leisten kann.

Zusammenfassend lässt sich sagen, dass das entwickelte Data-Analytic-Tool auf Basis von R und verschiedenen Machine-Learning-Algorithmen Unternehmen dabei unterstützen kann, präzisere Umsatzprognosen im Controlling zu generieren. Durch den Vergleich der Algorithmen können Unternehmen denjenigen Ansatz auswählen, der am besten zu ihren spezifischen Anforderungen und Datenstrukturen passt. Es bleibt jedoch zu betonen, dass weitere Untersuchungen und Validierungen in realen Unternehmensszenarien erforderlich sind, um die Leistungsfähigkeit und praktische Anwendbarkeit des Tools zu verifizieren.

Literatur und Abbildungen

- [1] Peter Buxmann and Schmidt Holder. *Künstliche Intelligenz*. Springer Gabler, 2019.
- [2] Scienceteam Data. Prinzip Gradientenverstärkung. <https://datascience.eu/de/maschinelles-lernen/gradientenverstarkung-was-sie-wissen-mussen/>, 01 2023.
- [3] Runge Statistik. Exemplarische grafische Ausgabe in R. <https://runge-statistik.de/statistische-auswertung-mit-r/>, 01 2020.
- [4] Phil Wenneker. *Künstliche Intelligenz in der Praxis*. Springer Gabler, 2020.

Konzeptentwicklung und prototypische Entwicklung eines First-Level-Support-Chatbots auf Basis von künstlicher Intelligenz

Nick Eschen

Anke Bez

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma ADS-TEC Administration GmbH, Nürtingen

Ausgangssituation

Die rasante Entwicklung der künstlichen Intelligenz hat in den letzten Jahren zu einer Vielzahl innovativer Anwendungen geführt. Eine solche Anwendung sind Chatbots, die auf der Grundlage von künstlicher Intelligenz entwickelt werden und Unternehmen neue Möglichkeiten und Chancen eröffnen können. Ein beliebtes Einsatzgebiet ist der Kundensupport, denn hier kommt es häufig zu wiederkehrenden Prozessen und Fragen, die von einem virtuellen Assistenten übernommen beziehungsweise beantwortet werden können. Dadurch wird den Unternehmen eine neuartige Möglichkeit geboten, die Effizienz und Produktivität zu erhöhen sowie Kosten einzusparen. Eine 24/7 Support-Verfügbarkeit kann zudem ein wichtiger Faktor in der Verbesserung der Kundenzufriedenheit darstellen [4].

Chatbots

Der Begriff 'Chatbot' setzt sich aus einer Kombination der englischen Wörter 'Chat' (Kommunikation) und 'Robot' (Roboter) zusammen. Beschrieben wird die Fähigkeit eines Computerprogramms, das

mithilfe von Audio- oder Messaging-Methoden menschenähnliche Konversation führt und Informationen bereitstellt. Chatbots werden als Dialogsysteme eingesetzt, um verschiedene Aufgaben, wie Kundenservice, Informationsbeschaffung und automatisierte Informationsabfrage, zu erfüllen. Es gibt verschiedene Arten von Chatbots, darunter regelbasierte Chatbots, KI-gestützte Chatbots und hybride Formen, die beide Ansätze kombinieren [3]. Die grundlegende Architektur eines Chatbots wird in Abbildung 1 veranschaulicht. Um eine Kommunikation mit dem Chatbot zu ermöglichen, ist ein Frontend erforderlich, mit dem der Benutzer interagieren kann. Das User Interface ist mit der Chatbot-Engine verbunden, in der sich der Dialog Manager befindet. Der Dialog Manager steuert die Eingaben des Benutzers und die Ausgaben des Chatbots. Natural Language Processing (NLP)-Methoden ermöglichen die Verarbeitung der Benutzereingaben und die Generierung passender Antworten. Die Applikation ist mit Backend-Systemen verbunden, die beispielsweise eine Wissensdatenbank oder ein analytisches Auswertungstool umfassen können.

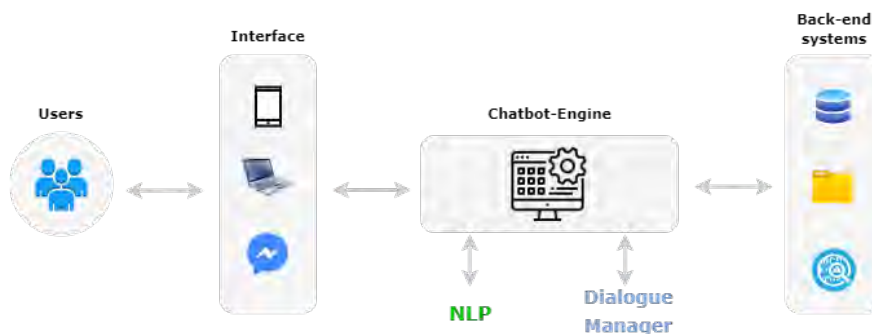


Abb. 1: Konversations Chatbot Architektur [1]

Zielsetzung

Das Ziel dieser Bachelorarbeit besteht darin, ein umfassendes Verständnis für den Einsatz von Kundensupport-Chatbots, die auf künstlicher Intelligenz basieren, in einem mittelständischen Unternehmen zu vermitteln. Im ersten Schritt werden zentrale Begriffe wie künstliche Intelligenz, Chatbots und die Aufgaben eines First-Level-Supports sowie die Vorteile und Möglichkeiten bei der Implementierung von Chatbots erläutert. Zudem werden aktuelle Trends und Herausforderungen analysiert, um das Projekt auf seine Durchführbarkeit hin zu untersuchen. Ein Marktvergleich verschiedener Anbieter soll zur Identifizierung der Kundenanforderungen durchgeführt werden. Basierend auf den gewonnenen Erkenntnissen wird ein Prototyp für den First-Level-Support entwickelt (siehe Abbildung 2). Der Chatbot soll als automatisiertes Support-Tool fungieren und in der Lage sein, häufig gestellte Fragen zu beantworten, kleinere Aufgaben zu übernehmen und Benutzer bei der Lösung von technischen Problemen zu unterstützen.

Ausblick

Der praktische Teil dieser Arbeit umfasst die Konzeptentwicklung eines Chatbots zur Unterstützung des First-Level-Supports in der IT-Abteilung von ADS-TEC. Die Durchführung von Experteninterviews mit Mitarbeitern der IT-Abteilung und eine Analyse des aktuellen Zustands des First-Level-Supports, um die Anforderungen an das Zielsystem zu definieren, sind hierbei relevante Bestandteile. Eine Risikoanalyse dient zur Einschätzung, Bewertung und Priorisierung der Risiken im Projekt und soll wichtige Erkenntnisse über mögliche Gefahren liefern. Im weiteren Verlauf werden die Anbieter anhand vordefinierter Kriterien bewertet, um eine geeignete Auswahl für eine Plattform oder eines Frameworks treffen zu können. Die Ergebnisse werden anschließend analysiert und beschrieben, um Rückschlüsse auf den Projekterfolg zu erzielen. Die finale Arbeit wird dem Unternehmen als zukünftige Machbarkeitsstudie dienen, um die Entscheidungsfindung für das weitere Vorgehen zu unterstützen.



Abb. 2: Entwurf eines Chatbot Widget [2]

Literatur und Abbildungen

- [1] Jerome Boyer. Build an IT support chatbot by using IBM Watson Assistant. https://www.ibm.com/cloud/architecture/tutorials/watson_conversation_support/, 2023.
- [2] Eigene Darstellung.
- [3] Sandeep A Thorat and V Jadhav. A review on implementation issues of rule-based chatbot systems. In: Proceedings of the international conference on innovative computing & communications. *ICICC*, 2020.
- [4] Raphael Meyer Von Wolff, Sebastian Hobert, and Matthias Schumann. *Einsatz von Chatbots am digitalen Büroarbeitsplatz. Eine praxisorientierte Betrachtung von Einsatzbereichen, Wirkungen und Handlungsempfehlungen.*, volume 3. Springer, 57 edition, 2020.

Vergleich und Evaluation von GraphQL-Implementierungen LowCode Framework gegenüber traditioneller Entwicklung

Julius Freudenberger

Jörg Nitzsche

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma doubleSlash Net-Business GmbH, Stuttgart

Einleitung

APIs umgeben uns in der modernen Applikationswelt mehr denn je. Immer mehr Softwareprodukte werden als einzelne Komponenten verpackt und öffentlich oder intern zur Verfügung gestellt [3].

Das 2015 von Facebook als Open-Source-Projekt veröffentlichte GraphQL ist eine Abfragesprache und eine Laufzeitumgebung für die Ausführung von Abfragen. Als Alternative zu REST bietet GraphQL mehr Flexibilität und Effizienz [4].

Jedoch müssen APIs weiterhin entwickelt und mit Datenquellen verknüpft werden. Diese Schritte werden größtenteils in traditioneller Softwareentwicklung durchlaufen. Oft unterstützt durch Frameworks oder Bibliotheken schreiben Entwickelnde Quellcode, der am Ende in einem Softwareprodukt ausgeliefert wird. LowCode-Frameworks sind eine Plattform, die – angetrieben durch den omnipräsenten Bedarf an Entwickelnden – versuchen, den Flaschenhals in der Softwareentwicklung zu weiten. Sie schaffen dies durch den klaren Fokus auf Entwicklungsproduktivität und die starke Reduzierung des Programmieranteils [2].

Zielsetzung der Arbeit

Zielsetzung dieser Arbeit ist der Vergleich von Implementierungen von GraphQL-APIs anhand eines traditionell entwickelten Services sowie mittels eines LowCode-Frameworks. Im Vergleich werden quantitative und qualitative Methoden genutzt. Ein weiteres Ziel ist herauszufinden, welcher Entwicklungsansatz in welchen Situationen geeignet ist.

GraphQL

Wie beim heutzutage häufig eingesetzten REST nutzt GraphQL das Request–Response–Schema. Allerdings werden keine verschiedenen Endpunkte für den Zugriff auf Ressourcen bereitgestellt. Vielmehr werden in einem GraphQL-Schema Objekttypen und ihre Beziehungen definiert. Abb. 1 zeigt beispielhaft die

Typdefinition eines Typs 'Customer' mit verschiedenen Attributen. Bei einer Abfrage an eine GraphQL-API können die benötigten Objekte mit ihren Attributen und Beziehungen anhand des Schemas angegeben werden. So wird das in REST-APIs auftretende Problem von Over- und Unterfetching vermieden. Bei Overfetching enthält die Antwort zu viele Daten, die der Client nicht benötigt. Unterfetching beschreibt die Notwendigkeit nach weiteren Anfragen, um alle relevanten Daten zu erhalten [4]. Die gesamte Abfragezeit für die gleichen Informationen ist bei GraphQL-APIs kürzer [3].

```
1  type Customer {
2      customerId: UUID
3      salutation: String
4      firstName: String
5      lastName: String
6      addresses: [Address]
7      contracts: [Contract]
8      cart: Cart
9      invoices: [Invoice]
10 }
```

Abb. 1: Definition des Objekttyps Customer in einer GraphQL-Schemadefinition [1]

Nutzung von GraphQL

In einem GraphQL-Schema werden neben den vorhandenen Objekttypen auch vorhandene Querys, Mutations und Subscriptions definiert. Diese Wurzeltypen bilden Einstiegspunkte in die GraphQL-API. Querys bieten die Möglichkeit, Daten abzufragen und dabei gegebenenfalls zu filtern. Sie kommen den GET-Requests in REST-APIs gleich. Manipulation von Daten ist mit Mutations möglich – ähnlich

wie POST-, PUT- oder DELETE-Requests in REST. Diese werden einzeln nacheinander abgearbeitet, um Race Conditions zu vermeiden. Durch Subscriptions können Clients benachrichtigt werden, wenn sich Daten geändert haben. So entfällt ein regelmäßiges Polling der aktuellen Daten seitens der Clients oder das Auslösen eines Webhooks [3].

Abb. 2 zeigt eine GraphQL, die für den Objekttyp 'Customer' die Attribute 'customerId', 'firstName' und 'lastName' abrufen. Die weiteren Attribute, die der Typ besitzt, werden nicht abgefragt und in der Antwort nicht zurückgegeben.

```

1 query getAllCustomers {
2   customers {
3     customerId
4     firstName
5     lastName
6   }
7 }
```

Abb. 2: Beispielhafte GraphQL-Query [1]

Konzept

Zur Entwicklung mit dem traditionellen Ansatz wird das Spring Framework genutzt. Das LowCode-Framework PostGraphile dient als Basis für die zweite Implementierung. Um verschiedene Möglichkeiten einer GraphQL-API darzustellen und Vergleichbarkeit zu gewährleisten, werden Use Cases und Anforderungen an das GraphQL-Schema definiert, die die Services umsetzen sollen. Außerdem werden ein Datenbankschema und Testdaten für das Datenbankmanagementsystem PostgreSQL definiert, die für beide Services identisch sind.

Die Use Cases bilden Anforderungen an Querys und Attribute ab. Darüber hinaus definieren sie Funktionen, die im Rahmen von Querys oder Mutations erfüllt werden sollen.

Der Vergleich der beiden implementierten Services nimmt qualitative und quantitative Merkmale in den Blick. Durch dieses Vorgehen können nicht nur die gemessenen Zeiten der Services verglichen werden, sondern auch weitere Aspekte geprüft werden, die sich direkt während der Entwicklung auswirken oder während der Laufzeit deutlich werden.

Es werden unter anderem die quantitativen Metriken Durchsatz und Anfrageausführungszeit ermittelt. Diese werden mithilfe der Werkzeuge zur Durchführung von Lasttests Gatling und k6 gemessen. Die qualitativen

Vergleichsaspekte umfassen unter anderem das Verhalten bei Pagination, das resultierende SQL-Statement und die Möglichkeit zur Implementierung eigener Logik.

Implementierung

Bei der Entwicklung mit Spring werden Resolver Functions implementiert. Diese rufen Daten aus der Datenbank ab und vervollständigen so die Attribute und Beziehungen des Antwortobjekts. Batch Loading kann verwendet werden, um dem Problem entgegenzuwirken, dass für jede Instanz eines Objekttyps die gleichen Daten aus der Datenbank geholt werden müssen. Dafür werden im Voraus alle Instanzen des gefragten Objekttyps aus der Datenbank abgerufen und in einer Map zwischengespeichert. Als Schlüssel dient dabei der Objekttyp, zu dem die Attribute abgerufen werden sollen. Bei Pagination von Ergebnislisten können im Schema sich wiederholende Muster teilweise generiert werden und auf der Datenbankseite auf Funktionen von Spring zurückgegriffen werden. Die Verknüpfung von Schema und Datenbankabfrage muss allerdings selbst entwickelt werden. Dabei wird jedem Eintrag in der Ergebnisliste ein Cursor hinzugefügt, mithilfe dessen in einer späteren Abfrage nachfolgende Einträge abgerufen werden können.

PostGraphile generiert nach der Konfiguration der Datenbankverbindung aus dem Tabellenschema ein GraphQL-Schema. Mithilfe von Plugins und Smart Comments kann die Benennung von Feldern und das Verhalten angepasst werden. Weitere Möglichkeiten bieten Funktionen, mithilfe derer weitere Objekttypen oder eigene Mutations hinzugefügt werden können. Dabei ist die in PostgreSQL integrierte Sprache *PL/pgSQL* hilfreich, die IF-Anweisungen und Schleifen beherrscht.

Erste Erkenntnisse und Ausblick

Während der Entwicklung sind bereits Unterschiede deutlich geworden. Die Flexibilität von Java in der Verwaltung von Listen sowie die strukturierte Ausnahmebehandlung vereinfachen die Implementierung.

Bei der Implementierung mit PostGraphile wird schnell ein GraphQL-Schema mit Standard-Querys und Resolver Functions generiert. Die in SQL definierten Objekttypen und berechneten Felder fügen sich nahtlos in das GraphQL-Schema ein. Vor allen Dingen bei unerfahreneren Entwicklenden benötigt die Implementierung von Mutations in *PL/pgSQL* mehr Einarbeitungszeit.

Die Messungen von Durchsatz und Anfrageausführungszeit werden weitere Aussagen im Vergleich quantitativer Aspekte treffen können.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Ulrich Frank, Pierre Maier, and Alexander Bock. Low Code Platforms: Promises, Concepts and Prospects: A Comparative Study of Ten Systems. In *ICB Research Reports*. DuEPublico: Duisburg-Essen Publications online, University of Duisburg-Essen, Germany, 2021.
- [3] Dominik Kress. *GraphQL: eine Einführung in APIs mit GraphQL*. dpunkt.verlag, 1 edition, 2021.
- [4] Eve Porcello and Alex Banks. *Learning GraphQL: Declarative data fetching for modern web apps*. O'Reilly Media, 2018.

Einsatz von Process Mining zur Optimierung von Geschäftsprozessen

Thimo Froehner

Dirk Hesse

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Einleitung

In wenigen Jahren hat sich die Gesellschaft von mehrheitlich analog zu digital gewandelt, was einen enormen Einfluss auf die Vorgehensweise hat, wie ein Mensch kommuniziert und Geschäfte macht. Der Mensch und Organisationen sind in ständiger Verbindung mit dem Internet und Daten werden jederzeit, überall und über jedes Thema gesammelt. Daraus ergibt sich auch, dass Mensch und Organisation immer mehr auf computergestützte Geräte angewiesen sind. Ebenso resultiert daraus ein enormes Datenwachstum in den letzten Jahren [4]. Eine Schätzung der International Data Corporation zeigt, dass die Menge der gespeicherten digitalen Informationen die im Jahr 2018 33 Zettabytes betrug, voraussichtlich bis ins Jahr 2025 auf unglaubliche 175 Zettabytes ansteigen wird [2]. Informationssysteme spielen nicht nur wegen der enormen Datenmenge eine wichtige Rolle, sondern auch wegen ihrem Einsatz in den heutigen Geschäftsprozessen. Ereignisse, die einen datengetriebenen Ansatz für Geschäftsprozesse liefern, können durch das Wachstum des digitalen Universums, das auch mit dem Begriff „Big Data“ zusammengefasst wird, heutzutage in Informationssystemen aufgezeichnet und danach abgeleitet und analysiert werden. Diese Entwicklung begründet den Einsatz von Process Mining, das genutzt werden kann, um Erkenntnisse zu gewinnen, Richtlinienverstöße aufzuzeichnen, Engpässe zu identifizieren, Probleme zu antizipieren, Gegenmaßnahmen zu empfehlen und Prozesse zu optimieren [4].

Ziel der Arbeit

Die Arbeit gibt einen Überblick über die theoretischen Grundlagen des Process Minings, zeigt das resultierende Potenzial bei Anwendung von Process Mining auf und befasst sich mit den zugrunde liegenden Techniken. Dabei wird der Fokus besonders auf die Transformation der Daten zu einem Ereignisprotokoll, sowie auf dessen Vorverarbeitung gelegt. Das Ziel der Arbeit ist somit,

ein Konzept vorzulegen wie Unternehmen an ihre Ereignisprotokolle herankommen.

Process Mining Grundlagen

Process Mining gilt als relativ junge Disziplin, die als ein Bindeglied zwischen den Bereichen Data Science und Process Science angesehen werden kann. Während Datenwissenschaften wie beispielsweise Data Mining, Statistik und Techniken des maschinellen Lernens keine End-to-End Prozesse berücksichtigen, sind prozesswissenschaftliche Ansätze prozessorientiert und konzentrieren sich auf die Modellierung und nicht auf das Lernen aus Ereignisdaten. Process Mining dagegen, als eine einzigartige Verbindung zwischen diesen zwei Bereichen, fungiert als Werkzeug, um End-to-End Prozesse zu verbessern, durch die Nutzung der immer weiter wachsenden verfügbaren Datenmengen. Die Grundidee hinter Process Mining ist somit, Prozesse die in der Realität ablaufen zu entdecken, diese zu überwachen und gegebenenfalls zu verbessern, mit Wissen das aus Ereignisprotokollen extrahiert wurde. Solche Ereignisprotokolle sind in heutigen Informationssystemen leicht verfügbar. Es gibt heutzutage viele unterschiedliche Systeme, die in Organisationen eingesetzt werden um die betrieblichen Prozesse zu unterstützen. Dabei wird eine enorme Menge an Ereignissen protokolliert, die detaillierte Informationen über die durchgeführten Aktivitäten geben. Informationssysteme können beispielsweise klassische WFM-Systeme, BPM-Systeme, ERP-Systeme, CRM-Systeme und Krankenhausinformationssysteme sein. Bei PAIS (Process Aware Information Systems) stehen Ereignisprotokolle direkt zur Verfügung, anders hingegen liegen die Informationen in den meisten Informationssystemen in unstrukturierter Form vor, d.h. die Ereignisdaten sind über mehrere Tabellen verteilt oder befinden sich in Subsystemen und müssen vor ihrer Verwendung aus diesen ausgelesen werden. Aus diesen Gründen gehört die Datenextraktion zu jedem

Process Mining Ansatz als wesentlicher Bestandteil dazu [4].

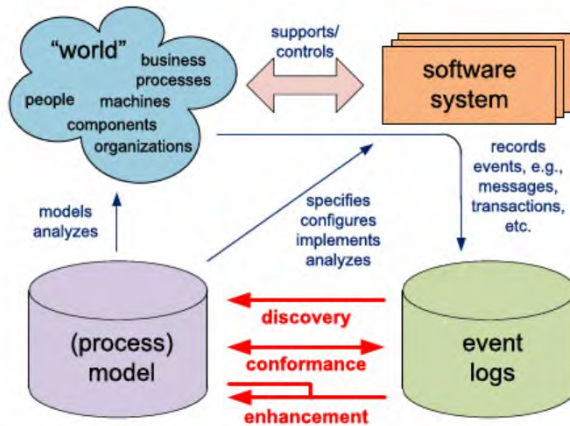


Abb. 1: Positionierung der drei Hauptarten des Process Mining [4]

Wie in Abbildung 1 zu sehen gibt es drei Process Mining Arten, bei deren Ausführung Ereignisprotokolle zum Einsatz kommen:

- Process Discovery:** Hierbei handelt es sich um die Generierung von Prozessmodellen basierend auf einem Ereignisprotokoll, ohne dabei vorherige Informationen zu verwenden. Zum Einsatz kommen verschiedene Algorithmen, beispielsweise der Alpha Algorithmus der mit dem zuvor in einem Ereignisprotokoll aufgezeichneten Verhalten ein Petrinetz erstellt. Liegen genügend Beispielausführungen des realen Prozesses vor, kann der Alpha Algorithmus ohne zusätzliches Wissen das dazugehörige Prozessmodell automatisch erstellen. Ebenso gibt es die Möglichkeit ressourcenbezogene Modelle zu entdecken, die die Zusammenarbeit innerhalb einer Organisation darstellen, wenn das Ereignisprotokoll ein Attribut mit Resource Informationen (Person oder Gerät, die eine Aktivität ausführt) enthält [4].
- Process Conformance:** Hierbei handelt es sich um die Feststellung, ob das Modell mit der aufgezeichneten Realität in einem Protokoll übereinstimmt und umgekehrt. Dabei wird ein bestehendes Prozessmodell und ein Ereignisprotokoll desselben Prozesses miteinander verglichen. Somit werden Abweichungen identifiziert, die behandelt werden können [4].

- Process Enhancement:** Bei der letzten Art des Process Mining handelt es sich um die Erweiterung, Anreicherung und Verbesserung eines bestehenden Prozessmodells. Dabei kommen ebenfalls Informationen aus dem Ereignisprotokoll zum Einsatz. Somit können in einem Prozessmodell beispielsweise Engpässe, Durchlaufzeiten, Häufigkeiten etc. angezeigt werden [4].

Event Log

Ein Ereignisprotokoll bildet die Grundlage jeder Prozessanalyse im Process Mining. In der Regel ist ein Ereignisprotokoll so aufgebaut, dass eine Spalte ein Attribut repräsentiert und eine Zeile einem Ereignis entspricht. Ein Ereignis steht dabei für die Durchführung eines Arbeitsschrittes innerhalb eines Prozesses. Jedes Ereignis wird einer Fall-ID zugeordnet, die einer Prozessinstanz entspricht. Eine zeitliche Abfolge von Ereignissen wird auch als Trace im Process Mining bezeichnet. Anhand der einzelnen Traces, ist ersichtlich auf welche unterschiedlichen Arten der Prozess ausgeführt wurde [3].

Case ID	Activity	Date	Time	Resource	Cost	...
CASE_0001	Order Request	2019-01-01	10:00	John Doe	100	...
CASE_0002	Order Request	2019-01-01	10:05	John Doe	100	...
CASE_0003	Order Request	2019-01-01	10:10	John Doe	100	...
CASE_0004	Order Request	2019-01-01	10:15	John Doe	100	...
CASE_0005	Order Request	2019-01-01	10:20	John Doe	100	...
CASE_0006	Order Request	2019-01-01	10:25	John Doe	100	...
CASE_0007	Order Request	2019-01-01	10:30	John Doe	100	...
CASE_0008	Order Request	2019-01-01	10:35	John Doe	100	...
CASE_0009	Order Request	2019-01-01	10:40	John Doe	100	...
CASE_0010	Order Request	2019-01-01	10:45	John Doe	100	...

Abb. 2: Aufbau eines Ereignisprotokolls [1]

Wie in Abbildung 2 durch die Zahlen 1 bis 3 gekennzeichnet, gibt es drei wichtige Datenanforderungen an ein Ereignisprotokoll. Erstens muss jedes Ereignis einer Fall-ID, die durch die erste Spalte repräsentiert wird, zugeordnet sein. Zweitens sollte jedes Ereignis einer Aktivität innerhalb des Prozesses entsprechen, hier repräsentiert von der zweiten Spalte. Als letzte Anforderung muss jedes Ereignis ein Zeitstempelattribut besitzen, repräsentiert durch die dritte Spalte, damit die Ereignisse eines Falles eine chronologische Abfolge bilden. Ebenso kann ein Ereignisprotokoll zusätzliche Attribute enthalten wie Resource, Kosten, usw., um die Analyse von Prozessen zu verbessern [1].

Literatur und Abbildungen

- [1] Jochen de Weerd and Thandar Wynn. Data Preprocessing - Foundations of Process Event Data. In *Process Mining Handbook*, chapter 6, pages 194–195. Springer Nature Switzerland, 2022.
- [2] Der Informationsdienst des Instituts der deutschen Wirtschaft. Datenmenge explodiert. <https://www.iwd.de/artikel/datenmenge-explodiert-431851/>, 06 2019.
- [3] Ralf Laue, Agnes Koschmider, and Dirk Fahland. *Prozessmanagement und Process Mining - Grundlagen*. Walter de Gruyter GmbH Berlin/Boston, 2021.
- [4] W.M.P. van der Aalst. *Process Mining: Data Science in Action*. Springer Berlin/Heidelberg, 2016.

Entwicklung von Echtzeit-Chatkomponenten mit den Web-Technologien SignalR und Blazor am Beispiel eines Azure-Cloud basierten 2-Personen Spieles

Frank Fuchs

Mirko Sonntag

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma SYSTECS Informationssysteme GmbH, Leinfelden-Echterdingen

Einleitung

Diese Bachelorarbeit beleuchtet die Funktionsweise und Möglichkeiten der Softwarebibliothek SignalR anhand des 2-Personen Spiels *Oware*. Aufgrund der zunehmenden Relevanz von KI wird es auch für Mittelständische Softwareunternehmen immer wichtiger sich in diesem Bereich Know-How anzueignen. Daher wird bei SYSTECS an einer KI für *Oware* gearbeitet, gegen die in Zukunft auch über diese SignalR-Webanwendung gespielt werden können soll. Die in dieser Arbeit erstellte Anwendung soll dann als eine Demo dienen um Interessierten die Fähigkeiten und Kompetenzen des Unternehmens im Bereich Künstliche Intelligenz und Echtzeit-Webanwendungen zu präsentieren.

Aufgabenstellung

Im Rahmen der Bachelorarbeit soll ein Konzept für eine Webanwendung für das 2-Personen Spiel *Oware* erarbeitet und umgesetzt werden. Dabei wird zuerst die notwendige Funktionalität anhand einer einfachen Chat Anwendung entwickelt in der alle Echtzeitkomponenten programmiert werden. Diese sollen dann auf das Spiel übertragen und gegebenenfalls angepasst werden. Am Ende ist das Ziel eine über Azure Cloud veröffentlichte Webanwendung, auf der Benutzer gegen andere Benutzer oder eine KI *Oware* spielen können. Dabei soll weder eine Registrierung noch eine andere Identifikation notwendig sein.

SignalR

SignalR ist Teil des ASP.NET Frameworks von Microsoft, welches zu den Hauptkompetenzen des Unternehmens zählt, bei dem diese Arbeit erstellt wird. Daher wird auch kein Vergleich zu anderen ähnlichen Frameworks vorgenommen. SignalR stellt eine API bereit, die es ermöglicht *Remote Procedure*

Calls (RPC) einfach zu erstellen. Dadurch können serverseitig Funktionen bei den gewünschten Klienten aufgerufen werden. [2] Für den Transport dieser RPCs verwendet SignalR automatisch die beste zur Verfügung stehende Technologie unter gegebenen Umständen (z.B. Browser-Version) in folgender Reihenfolge.

1. Web Sockets
2. Server-Sent Events
3. Long Polling

[3]

Konzept

Bei der Umsetzung wurde zuerst ein Konzept erstellt welches die Planung und Ausarbeitung der Anforderungen umfasst und es wurden alle Use-Cases festgehalten. Die wichtigsten dabei sind:

- Ein Benutzer kann ein neues Spiel erstellen oder einem bereits erstelltem Spiel beitreten
- Es können maximal 2 Benutzer an einem Spiel teilnehmen
- Ein Benutzer kann in Echtzeit gegen einen anderen Benutzer spielen und Chatnachrichten schicken
- Auf der Startseite werden jedem Benutzer alle verfügbaren Spiele angezeigt und die Ansicht wird automatisch aktualisiert

Dafür wurden verschiedene UML Diagramme entworfen die dann in einer Demo Version programmiert werden um die Umsetzbarkeit zu testen.

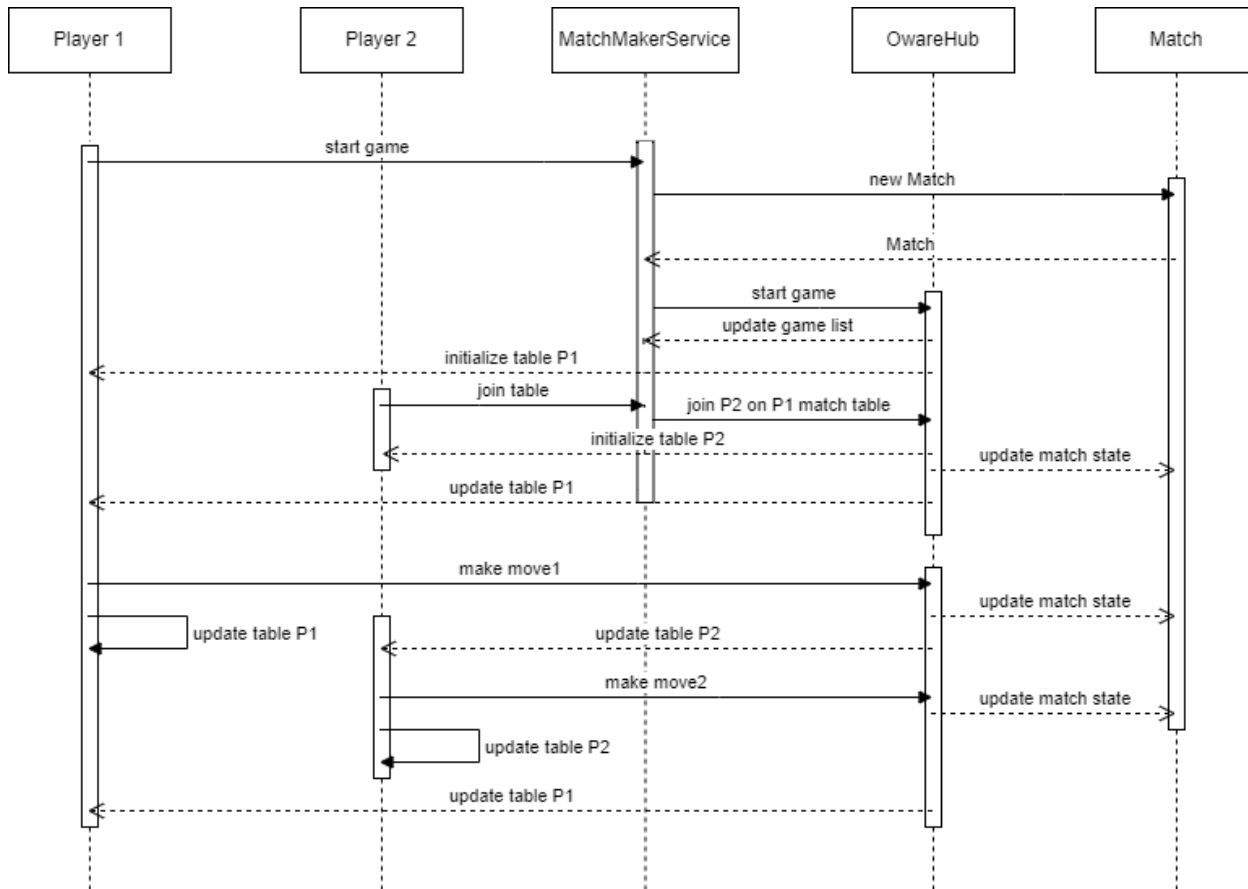


Abb. 1: Sequenzdiagramm Spiel Menschen gegen Mensch [1]

In Abbildung 1 ist zum Beispiel das Sequenzdiagramm für das Zusammenspiel der Codekomponenten im Fall eines Spiels zwischen zwei Benutzern zu sehen. Es wurde unter anderem auch ein Prototyp in Balsamiq

erstellt um die Funktionen visuell greifbarer zu machen. In der folgenden Abbildung 2 ist die Auflistung der verfügbaren Spiele zu sehen und die Möglichkeit einem Spiel beizutreten oder ein neues zu erstellen.

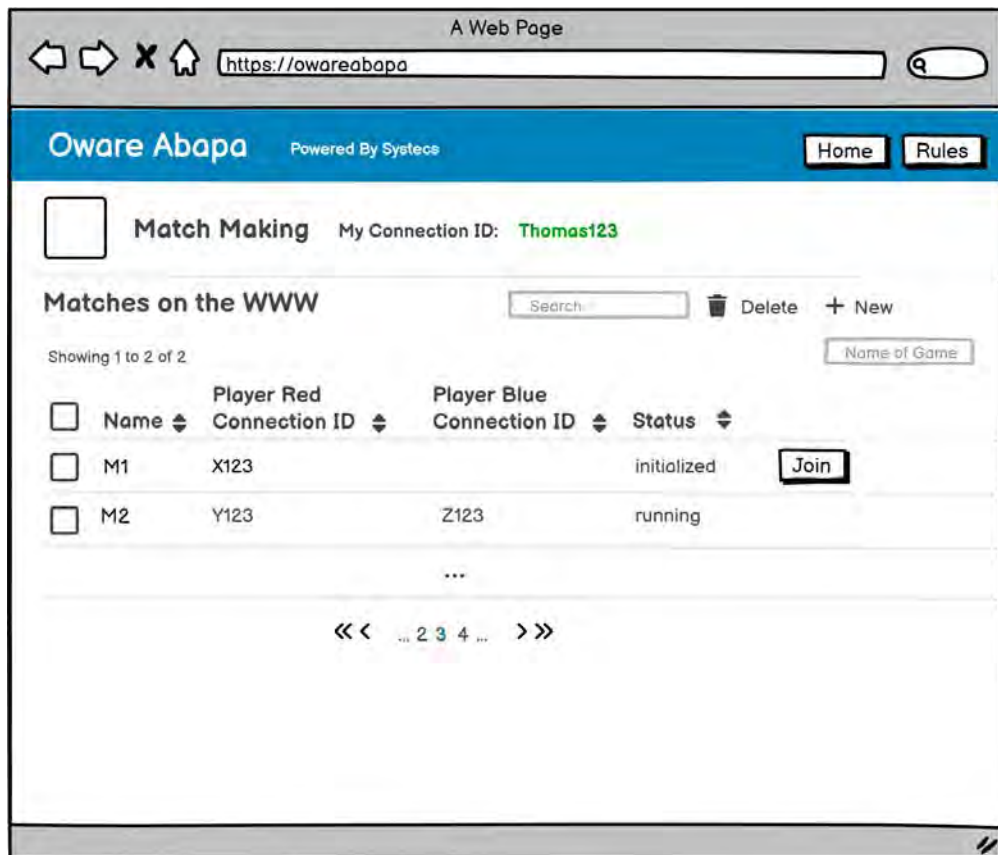


Abb. 2: Balsamiq Mockup Webanwendung [1]

Implementierung

Um sich mit der SignalR Technologie und den Funktionalitäten vertraut zu machen, wurde zuerst eine Demo Chat Webanwendung programmiert. Diese soll die wichtigsten Anforderungen abbilden. Zum Beispiel wird die Ansicht für alle, mit dem Server verbundenen Klienten aktualisiert, wenn ein neuer Chatraum erstellt, beziehungsweise ein bestehender Raum geschlossen wird. Außerdem wurden auch die einfacheren Remote Procedure Calls programmiert, wie zum Beispiel das Verschicken von Nachrichten an Teilnehmer eines Chatraums. Der Code soll später fast unverändert für die Spiel Anwendung übernommen werden. Dabei kristallisierten sich auch neue Probleme heraus, die im Laufe der Arbeit gelöst werden sollen, zum Beispiel *Was passiert, wenn ein Benutzer kurzzeitig*

die Verbindung verliert?

Ausblick

Zuerst soll das Spiel mit allen Funktionen fertiggestellt werden. Die Spiel Engine wird größtenteils zur Verfügung gestellt, muss aber noch optimiert werden. Dann sollen die aus der bereits erwähnten Chat Demo gewonnen Erkenntnisse auf die Spieleanwendung übertragen werden. Im weiteren Schritt soll versucht werden, das in **Implementierung** genannte Problem zu lösen. Die aktuelle Idee ist an JSON Web Tokens agelehnt. Sie sieht vor, dass bei der ersten Verbindung eines Benutzers ein Token vergeben und im Browser gespeichert wird. Dieser ermöglicht bei einer Wiederverbindung eine Identifikation des Benutzers.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Patrick Fletcher et al. Introduction to SignalR. <https://learn.microsoft.com/de-de/aspnet/signalr/overview/getting-started/introduction-to-signalr>, 09 2020.
- [3] Brady Gaster et al. Overview of ASP.NET Core SignalR. <https://learn.microsoft.com/en-us/aspnet/core/signalr/introduction?view=aspnetcore-7.0>, 02 2023.

Entwicklung einer Softwarelösung zur Verifizierung und Lizenzierung eines Ladesäulen-Prüfgeräts

Ismet Gezer

Mirko Sonntag

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Carico GmbH, Esslingen

Einführung

Der Ausbau der Ladeinfrastruktur in Deutschland ist ein wichtiger und stetiger Prozess. Zum 1. März 2023 waren, laut der Bundesnetzagentur, **71.395 Normalladepunkte** (AC-Ladesäulen) und **13.714 Schnellladepunkte** (AC- & DC-Ladesäulen) im Betrieb. [5] Diese Zahlen beziehen sich nur auf die Anzahl der öffentlichen Ladepunkte, da die privaten Ladepunkte nicht erfasst werden. Die Installation, Wartung und Reparatur der Ladepunkte spielen dabei eine wichtige Rolle. Für diese Aufgaben existieren Ladesäulen-Prüfgeräte. Mit diesen Prüfgeräten werden Prüfnormen durchgeführt, die die Funktionalität der Ladesäule sicherstellen. Hierbei wird zwischen Erst-, Regel- und Sicherheitsprüfungen wie der DIN VDE 0100-600 und der DIN VDE 0105-100 unterschieden. [6] Die Carico GmbH entwickelt ein Prüfgerät für AC und DC-Ladesäulen, welches die ServicetechnikerInnen dabei unterstützt, diese schneller zu warten, zu reparieren und zu installieren. Zum aktuellen Zeitpunkt liegt der Aufbau eines AC-Ladesäulen-Prüfgeräts im Fokus (siehe Abbildung 1).



Abb. 1: AC-Prüfgerät Prototyp [6]

Produktfälschungen

Des Weiteren haben viele Produkte, unabhängig der Kategorie (Kleidung, Schmuck, Uhren, Medizin, ...), mit Plagiaten und Produktpiraterie zu kämpfen. Um die Gesellschaft auf dieses Problem aufmerksam zu machen, existiert unter anderem ein Negativ-Preis für Produktfälschungen, der sogenannte *Plagiarus-Preis* (siehe Abbildung 2). Dieser wird, seit 1977, jährlich an Hersteller und Händler von ausdrücklich dreisten Produktfälschungen vergeben. [3]



Abb. 2: Plagiarus-Preis [3]

Damit sowohl eine potentielle Produktfälschung als auch die unbefugte Nutzung des AC-Ladesäulen-Prüfgeräts ausgeschlossen werden kann, soll ein Verifizierungs- und Lizenzierungsmodell ausgearbeitet werden.

Aufgabenstellung

Folglich sollen neue Software- und Hardware-Komponenten erstellt und bestehende erweitert werden. Darunter fällt die Ausarbeitung und Dokumentation der Nutzer-Produkt Verifikation, das Anpassen der Smartphone-Applikation sowie die Erweiterung dieser mit der Verify Origify Technologie, das Anpassen der Backend-Datenbank und das Ausarbeiten von möglichen Extremfällen (beispielweise Sabotagen, Sicherheitslücken, etc.).

Anforderungen der Bachelorarbeit

Die Kommunikation zwischen dem Kunden und dem AC-Ladesäulen-Prüfgerät erfolgt mittels einer Smartphone-Applikation. Durch diese App erhält der AC-Koffer seine gesamte Funktionalität. Demnach ist die Ausarbeitung eines Verifizierungs- und Lizenzierungsmodells in direkter Abhängigkeit mit der App. Hierbei soll die bereits existierende Smartphone-Applikation so adaptiert werden, dass die Nutzbarkeit und die Funktionalitäten mit dem AC-Koffer vorerst eingeschränkt werden. Dafür müssen neue Funktionen in die App implementiert werden. Die Nutzung soll erst dann freigeschaltet werden, wenn der Nutzer sein registriertes Produkt in der App verifiziert und zugleich eine Lizenz für die Funktionalitäten in der App erworben hat.

Verifizierung

Dabei definiert der Begriff *Verifizierung* die „Bestätigung durch Bereitstellung eines objektiven Nachweises, dass festgelegte Anforderungen erfüllt worden sind“ [4]. Dies kann durch unterschiedliche Methoden erfolgen. Die Bachelorarbeit beschränkt sich auf die folgenden drei Verfahren:

- Verifizierung mithilfe der Verify Origify App
- Verifizierung mithilfe von Software/Firmware
- Verifizierung mithilfe von Hardware

Die Verify Origify App ist hierbei eine Smartphone-Applikation von der Robert Bosch GmbH, die die Produktauthentifizierung ermöglichen und vereinfachen soll. [7] Hierbei erfolgt die Produktregistrierung durch eine Mustersignatur. [8]

Lizenzierung

Die *Lizenzierung* dient als „Sammelbegriff für Maßnahmen, um [...] Software gegen illegale Nutzung zu schützen und um flexible Vermarktungsmodelle zu realisieren“ [1]. Hierbei unterscheiden sich die Lizenzierungsmodelle anhand der gewählten Lizenz. Es gilt zu unterscheiden zwischen folgenden Lizenzen:

- Unbefristete Lizenz
- Concurrent-User Lizenzen
- Abonnement-basierte Lizenzen
- Proprietäre Lizenzen
- Floating-Feature Lizenzen
- Feature-basierte Lizenzen
- Netzwerk Lizenzen
- Cloud-basierte Lizenzen

Dabei sollen diese Lizenzierungsmethoden sowie das Sicherstellen der Einhaltung in der Bachelorarbeit thematisiert werden. Des Weiteren muss geklärt werden, wie es zum Lizenzerwerb kommt, was für eine Lizenzlaufzeit angestrebt wird und ob es unterschiedliche Lizenzstufen zum Erwerb geben wird.

Ausblick

Zum aktuellen Stand der Bachelorarbeit wurde bereits ein Layout-Screen für die Implementierung der Verify Origify SDK vorbereitet und ausprogrammiert. Dieses soll die Verifizierung für das Ladesäulen-Prüfgerät bereitstellen. Infolgedessen muss aktuell geklärt werden, welche Position auf der Frontplatte (siehe Abbildung 3) sich für die Produktverifizierung eignen würde. Im Anschluss muss das Produkt durch die Robert Bosch GmbH ins System eingepflegt werden, sodass eine Statusabfrage mit der SDK ermöglicht werden kann. Des Weiteren wurden die Grundfunktionen zwischen App und Prüfgerät durch einen Boolean-Eintrag im Backend eingeschränkt. Dieses soll nach erfolgreicher Verifizierung des Prüfgeräts auf *true* gesetzt werden. Die Freischaltung der Funktionen wird dementsprechend zum Teil gewährleistet. Damit die volle Nutzbarkeit der App erreicht werden kann, soll infolgedessen noch ein Lizenz-Parameter abgefragt werden. Dieses Lizenzmodell befindet sich aktuell in der Implementierung.

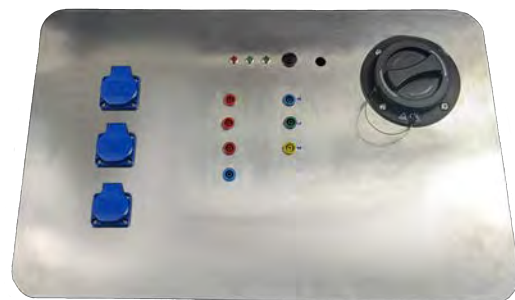


Abb. 3: Frontplatte [2]

Literatur und Abbildungen

- [1] Reimer M. Bürkner. *Erfolgreiche Software-Lizenzierung*. Springer Berlin Heidelberg, 2003.
- [2] Eigene Darstellung.
- [3] Aktion Plagiarius eV. Aktion Plagiarius - Innovation statt Imitation. <https://www.plagiarius.com/index.php?ID=39>, 2023.
- [4] DIN Deutsches Institut für Normung eV. *DIN EN ISO 9000 – Qualitätsmanagement – Grundlagen und Begriffe*. Beuth Verlag GmbH, 2005.
- [5] Bundesnetzagentur für Elektrizität und Gas. Elektromobilität: Öffentliche Ladeinfrastruktur. <https://www.bundesnetzagentur.de/DE/Fachthemen/ElektrizitaetundGas/E-Mobilitaet/start.html>, 2023.
- [6] Carico GmbH. SMARTE TESTSYSTEME FÜR AC UND DC-CCS LADESÄULEN. <https://www.carico.de/de>, 2023.
- [7] Robert Bosch GmbH. Bosch Origify. <https://www.bosch-origify.com/de/>, 2022.
- [8] Robert Bosch GmbH. Wie es funktioniert. <https://www.bosch-origify.com/de/how-it-works/>, 2022.

Analyse der Anforderungen für den Einsatz eines Systems zur Angriffserkennung für kritische Infrastrukturen

Maximilian Gittel

Dominik Schoop

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Fichtner IT-Consulting, Stuttgart

Motivation und Problemstellung

Ein moderner Geschäftsbetrieb ist heute aufgrund der Digitalisierung und Menge an Daten in Wirtschaft und Verwaltung ohne IT-Unterstützung nicht möglich, da alle Geschäftsbereiche und Prozesse auf IT basieren. Die damit verarbeiteten Informationen sind ein wesentlicher Wert für Unternehmen und Behörden und müssen daher angemessen geschützt sein [4]. Durch Integrationsanforderungen und der Entwicklung zur Optimierung der Fertigung, speziell im Rahmen der Industrie 4.0 und Digitalisierung setzen Unternehmen zudem immer mehr auf Prozessleit- und Automatisierungstechnik, der sogenannte Operational Technology (OT), die mit den Netzwerken des Unternehmens verbunden und sogar über Organisationsgrenzen hinweg verfügbar ist [2].

Mit der zunehmenden Digitalisierung und Vernetzung dieser Infrastrukturen steigt jedoch auch das Risiko von Cyberangriffen und Cyberkriminalität, die gravierende Auswirkungen auf die Funktionsfähigkeit und Stabilität dieser Systeme haben können. Die IT-Sicherheit wird zu einer strategischen Schlüsselfunktion, deren Umsetzung essenziell ist. Durch ein nicht ausreichendes Sicherheitsniveau kann es bei einem erfolgreichen Angriff zu schwerwiegenden Folgen für das Unternehmen kommen. Je nach Schwere des Angriffs kann dieser existenzbedrohend für das betroffene Unternehmen sein.

Im Bereich der Kritischen Infrastrukturen (KRITIS, Abb. 3) ist dies besonders gefährlich, da ein Ausfall oder eine Störung des Betriebs dramatische Folgen für Wirtschaft, Staat und Gesellschaft zur Folge haben kann.



Abb. 1: Kritische Infrastrukturen in Deutschland [6]

Daher wurde im Zuge der Cyber-Sicherheitsstrategie 2011 der Bundesregierung durch das IT-SiG 1.0 2015 die Befugnisse des Bundesministerium für Sicherheit in der Informationstechnik (BSI) entsprechend erweitert und Betreiber von KRITIS zur Umsetzung von IT-Sicherheit nach dem Stand der Technik verpflichtet [3].

Als eine der Neuerungen des IT-Sicherheitsgesetz 2.0, dass im Mai 2021 in Kraft getreten ist, wird der Einsatz und Nachweis eines Systems zur Angriffserkennung (SzA) als Teil des „Stand der Technik“ zur Gefahrenabwehr gefordert. Diese sollen im laufenden Betrieb kontinuierlich Angriffe durch den Abgleich von Mustern erkennen und vermeiden. Dazu sind aber explizit neben den technischen Werkzeugen auch unterstützende Prozesse gefordert [7].

Insbesondere den kleineren und mittleren betroffenen Unternehmen fehlt es jedoch an technischem Know-how, Verständnis der Thematik und Ressourcen, um die gesetzlichen Anforderungen zu verstehen und

erfolgreich umzusetzen.

ANFORDERUNG AN SZA

Als Unterstützung für die Umsetzung und Nachweis der definierten Verpflichtungen wurde durch das BSI die „Orientierungshilfe zum Einsatz von Systemen zur Angriffserkennung“ (OH SzA) veröffentlicht, in der die Grundlagen, Anforderungen und der Nachweis von SzA aus Sicht des BSI erläutert wird. Die technischen Funktionalitäten eines SzA lassen sich im Wesentlichen

in den Bereichen Protokollierung, Detektion und Reaktion zuordnen, die auch zur Einteilung der definierten Anforderungen genutzt werden. Für die Protokollierung und Detektion werden die Anforderungen zusätzlich in die Unterbereiche Planung und Umsetzung aufgeteilt. Zusätzlich zu den direkt in der OH SzA beschriebenen Anforderungen werden auch Anforderungen aus Bausteinen des IT-Grundschutz definiert, die umgesetzt werden müssen [5]. Eine Übersicht über den Aufbau der Anforderungen liefert Abbildung 2.



Abb. 2: Aufbau der Anforderungen nach OH SzA [1]

Zielsetzung und Ergebnis

Die Zielsetzung dieser Arbeit besteht darin, die gesetzlichen Anforderungen für Angriffserkennungssysteme in KRITIS zu analysieren und damit verbundene Pflichten möglichst verständlich zu machen. Den durch die KRITIS-Regulierung betroffenen Unternehmen soll so die Umsetzung der Anforderungen und Prüfung ihrer Systeme erleichtert werden, damit diese den erforderlichen Umsetzungsgrad erreichen. Zudem sollen die Anforderungen als Teil der Analyse mit gängigen Standards für Informationssicherheitsmanagementsysteme (ISMS) verknüpft werden, um betroffenen Unternehmen mit konkreten Maßnahmen zu den Anforderungen bei der Umsetzung zu unterstützen.

Somit stellt die durchgeführte Analyse der Anforderungen auch das Ergebnis der Arbeit dar, welches als Leitfaden zur Umsetzung der Anforderungen und dem Einsatz von SzA genutzt werden kann. Ergänzend dafür werden die strukturierten Anforderungen zusätzlich in eine Checkliste eingebaut die als Grundlage zur Prüfung des derzeitigen Umsetzungsgrades und für den Nachweis der Umsetzung in Unternehmen verwendet werden kann.

DESIGN UND VORGEHENSWEISE

Vor der Analyse der Anforderungen wird zunächst eine grundlegende Übersicht über die Kernthemen der

Arbeit erarbeitet, die die Basis für die Analyse der Anforderungen bilden. Bei diesen handelt es sich um:

- IT- und Informationssicherheit in der digitalisierten Wirtschaft und Gesellschaft.
- KRITIS und deren gesetzliche Regulierung
- die gesetzlichen Anforderungen und Pflichten zum Einsatz von Systemen zur Angriffserkennung

Als Teil der Analyse werden die Anforderungen umstrukturiert und kategorisiert, da die Strukturierung (Vgl. Abb. 2) speziell die der Anforderungen, die sich direkt aus der OH SzA ergeben unübersichtlich dargestellt ist und übergreifende Zusammenhänge nicht klar sind. Als Basis für die Umstrukturierung dient eine Ausarbeitung durch das Online-Portal OpenKritis und Strukturen des IT-Grundschutz. Damit die Planung und Umsetzung der Anforderungen sowie zukünftige Änderungen möglichst effizient ablaufen und die SzA somit so gut wie möglich arbeiten können, wird

zusätzlich empfohlen diese in ein ISMS zu integrieren. Durch die Umstrukturierung wird zusätzlich ermöglicht, Anforderungen präzise mit den gängigen ISMS Standards IT-Grundschutz und die ISO/IEC 2700X Reihe zu verknüpfen. Dadurch können Schnittstellen sowie direkte Handlungs- und Umsetzungsmaßnahmen aufgezeigt werden.

Ausblick

Um zusätzlich das Verständnis zu verbessern, sollen die umstrukturierten Anforderungen in Grafiken und Diagrammen aufbereitet und dargestellt werden. Durch die Visualisierung lassen sich die Schnittstellen Interaktionen und Abhängigkeiten zwischen einzelnen Strukturbausteinen besonders gut aufzeigen und erklären. Dazu sollen diese zunächst in die übergeordneten Kategorien der Organisatorischen-, Prozess- und technischen Anforderungen zugeordnet werden und jeweils separat visualisiert werden.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Bundesamt für Sicherheit in der Informationstechnik. *BSI ICS-Security-Kompendium*. BSI, 2013.
- [3] Bundesamt für Sicherheit in der Informationstechnik. *Das IT-Sicherheitsgesetz - Kritische Infrastrukturen schützen*. BSI, 2016.
- [4] Bundesamt für Sicherheit in der Informationstechnik. *BSI-Standard 200-1 IT-Grundschutz-Methodik*. BSI, 2017.
- [5] Bundesamt für Sicherheit in der Informationstechnik. *Orientierungshilfe zum Einsatz von Systemen zur Angriffserkennung*. BSI, 2022.
- [6] Bundesamt für Sicherheit in der Informationstechnik. BSI - Was sind Kritische Infrastrukturen? https://www.bsi.bund.de/DE/Themen/KRITIS-und-regulierte-Unternehmen/Kritische-Infrastrukturen/Allgemeine-Infos-zu-KRITIS/allgemeine-infos-zu-kritis_node.html, 2023.
- [7] Open Kritis. Das IT-Sicherheitsgesetz 2.0.: <https://www.openkritis.de/it-sicherheitsgesetz/ausblick-it-sicherheitsgesetz-2-0.html>, 2023.

Migration und Optimierung einer Schwingungsanalyse-Software: Implementierung in Python und Anwendung zur Untersuchung von Abgasanlagen

Luisa Glass

Clemens Klöck

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Purem GmbH, Esslingen

Einleitung

Die Automobilindustrie steht immer wieder vor neuen Herausforderungen, um die Anforderungen von Kunden und Gesetzgebern zu erfüllen. Eine der wichtigsten Komponenten eines Fahrzeugs ist die Abgasanlage. Sie ist dafür verantwortlich, schädliche Emissionen und Geräusche zu reduzieren und gleichzeitig eine optimale Motorleistung zu gewährleisten [1].

Durch Motor- und Fahrbahnanregungen geraten Abgasanlagen unerwünscht ins Schwingen. An manchen Stellen können durch diese Schwingungen Beschädigungen wie z.B. Risse im Material entstehen. Deshalb sollen kritische Stellen und mögliche Resonanzfrequenzen frühzeitig erkannt werden, damit Verbesserungen vorgenommen werden können. Um sicherzustellen, dass eine Abgasanlage den Belastungen während des Lebenszyklus standhält und Emissionen sowie vorgeschriebenen Grenzwerten entspricht, ist eine detaillierte Analyse der Abgasanlage unerlässlich. Das Schwingverhalten und die Festigkeit dieser Systeme können mit dem Einsatz von experimentellen Messungen, sowie mit numerischen Simulationen bewertet und optimiert werden.

Purem by Eberspächer ist ein international tätiges Unternehmen, das sich auf die Entwicklung von Schalldämpfern, Katalysatoren, Rohren und Krümmern sowie kompletten Abgasanlagen für Pkw und Nutzfahrzeuge spezialisiert hat. Um die Qualität und Leistungsfähigkeit ihrer Produkte sicherzustellen, werden verschiedene Simulationen und Versuche durchgeführt. Dazu gehört auch die Identifikation und Analyse von Schwingungsverhalten und Resonanzfrequenzen.

Für diese Zwecke hat das Unternehmen ein eigenes Tool zur Schwingungsanalyse von Abgasanlagen entwickelt. Das Programm vergleicht Simulations- und Messdaten basierend auf dem Modal Assurance Criterion (MAC).

Modal Assurance Criterion (MAC)

Annahmen und Vereinfachungen im Simulationsmodell können dazu führen, dass die simulierten Schwingungsdaten von den tatsächlichen Werten abweichen. Durch den Vergleich von Simulation und Messung kann die Genauigkeit der Simulation überprüft werden.

Das Modal Assurance Criterion (MAC) ist ein wichtiges Verfahren zur Schwingungsanalyse, das die Übereinstimmung von Moden aus Simulationen und experimentellen Messungen vergleicht. Die MAC-Formel basiert auf dem normierten Skalarprodukt von zwei Vektoren. Der MAC-Wert liegt im Bereich von 0 bis 1. Ein MAC-Wert von 1 zeigt eine vollständige Übereinstimmung der Vektoren, während ein Wert von 0 keine Übereinstimmung anzeigt [2].

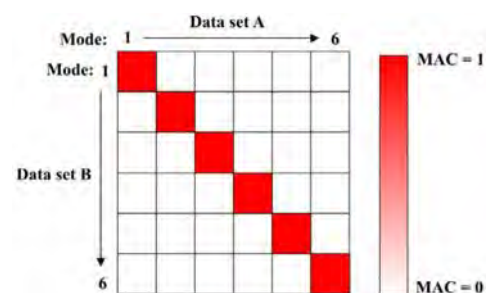


Abb. 1: Beispielhafte Darstellung einer MAC-Matrix [2]

Die Formel zur Berechnung des MAC-Werts liefert einen einzelnen Skalar. Wenn eine Menge an Vektoren vorliegt und diese untersucht werden soll, kann eine MAC-Matrix (vgl. Abbildung 1) erstellt werden. Die Skalare werden dann in einer MAC-Matrix angeordnet. Jedes Element der Matrix repräsentiert einen Vergleich von einem Vektorpaar. Die MAC-Matrix zeigt an, welche Vektoren aus den beiden Mengen zueinander in Beziehung stehen. Die MAC-Werte können dabei zur besseren Interpretation mit verschiedenen Farben

visualisiert werden. In Abbildung 1 ist ein Spezialfall zu sehen. Die Werte der Hauptdiagonalen besitzen den MAC-Wert 1, wohingegen die Nebenelemente den MAC-Wert 0 besitzen. In der Regel sind die Werte der Matrix nicht exakt 0 oder 1, sondern befinden sich dazwischen.

Das MAC-Verfahren bietet somit eine solide Grundlage für die Validierung von Simulationsmodellen im Schwingungsbereich.

Problemstellung

Das derzeitige MAC-Programm von Purem ist in der Programmiersprache Fortran implementiert und nur für das Betriebssystem Linux verfügbar. Obwohl das MAC-Programm effektiv ist und den Anforderungen des Unternehmens entspricht, ist seine derzeitige Inkompatibilität mit dem Betriebssystem Windows ein Hindernis für die globale Anwendung. Insbesondere an den ausländischen Standorten von Purem, wo Windows-Betriebssysteme verwendet werden und wenig oder kein Linux Know-how vorhanden ist, ist eine reibungslose Nutzung des Programms problematisch. Darüber hinaus kann das MAC-Programm aufgrund seiner Linux-Kompatibilität nicht in das firmeninterne Simulationsdatenmanagement-System integriert werden, das ausschließlich Windows-Anwendungen unterstützt.

Ziele der Arbeit

Ziel ist es, das MAC-Programm in die betriebssystemunabhängige Programmiersprache Python zu migrieren. Dies ermöglicht die Kompatibilität mit dem Betriebssystem Windows und erlaubt die globale Verwendung des Programms an allen Purem-Standorten. Durch die Migration des MAC-Programms wird die Einbindung in das firmeninterne Simulationsdatenmanagement-System ermöglicht. Dies fördert die systematische Verwaltung von Simulationsdaten und stellt sicher, dass Know-how zentral konserviert und allen Benutzern bereitgestellt wird.

Die Leistung und Effizienz des implementierten MAC-Programms sollen untersucht und optimiert werden, um eine verbesserte Identifikation von Resonanzfrequenzen und MAC-Werten zu ermöglichen. Es soll evaluiert werden, ob es geeignete alternative Methoden und Kennzahlen für den Vergleich der Schwingungsformen gibt. Diese sollen auf ihre Eignung

hin überprüft und als Entscheidungsbasis für die zukünftige Verwendung vorgeschlagen werden.

Insgesamt soll die Arbeit dazu beitragen, das MAC-Programm von Purem effizienter und anpassungsfähiger zu gestalten, um die Betriebsfestigkeit und Qualität der entwickelten Abgasanlagen weiterhin sicherzustellen.

Aktueller Stand

In der aktuellen Phase der Arbeit wurde eine Schwachstellenanalyse am bestehenden MAC-Programm durchgeführt und somit Defizite identifiziert. Das MAC-Programm wurde, unter Berücksichtigung der gewonnenen Erkenntnisse aus der Schwachstellenanalyse, in die Programmiersprache Python migriert. Die Hauptfunktionen sind bereits nutzbar und auch eine grafische Benutzeroberfläche zur einfacheren Bedienung ist implementiert.

Erste Untersuchungen zu alternativen Kennzahlen wurden durchgeführt und die gefundenen Ansätze bewertet. Zum Beispiel gibt es einige Erweiterungen des MAC, welche zusätzliche Informationen zur Analyse des Schwingverhaltens ermöglichen.

Auch die eigentliche MAC-Bestimmung wurde auf mögliche Optimierungen hin untersucht. Beispielsweise kann ein Phasenwinkel verwendet werden, um die mögliche Phasenverschiebung zwischen Messung und Simulation zu berücksichtigen.

Während der MAC-Berechnung ist die Bestimmung von lokalen Maxima erforderlich. Daher wurden verschiedene Ansätze zur Identifizierung von lokalen Maxima untersucht. Sowohl die manuelle Implementierung als auch die Nutzung von vorhandenen Python-Bibliotheken wurden in Betracht gezogen und bewertet.

Ausblick

In zukünftigen Entwicklungen der Bachelorarbeit werden weitere alternative Kennzahlen und Ansätze zum Vergleich von Schwingformen untersucht. Diese sollen bewertet und auf ihre Eignung für das Projekt hin evaluiert werden.

Weiterhin kann untersucht werden, ob eine Verwendung von komplexen Zahlen zur Beschreibung der Schwingungen Vorteile bietet. Derzeit werden im MAC-Programm reelle Zahlen verwendet.

Darüber hinaus können die gefundenen, alternative Ansätze implementiert und getestet werden.

Literatur und Abbildungen

- [1] Harald Kutka and Herbert Schnattinger. Abgasanlagen auf dem Prüfstand. *Materials Testing*, 41, 1999.
- [2] Miroslav Pastor, Binda Michal, and Tomáš Harčarik. Modal Assurance Criterion. *Procedia Engineering*, 48, 2012.

Verwendung von Software-Bill-of-Materials zur automatischen netzwerkweiten Erfassung von Softwarekomponenten

Julian Groshaupt

Tobias Heer

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Motivation und Problemstellung

Neben zunehmender Häufigkeit von Cyber-Angriffen auf Unternehmen [4] und steigender Komplexität von Netzwerken [3] tragen auch unterschiedliche Software verschiedener Hersteller zu unübersichtlicheren IT-Infrastrukturen bei. Um Schwachstellen zu finden werden häufig Netzwerk-Scanner eingesetzt, welche Hosts über von außen erreichbare Schnittstellen auf bekannte Schwachstellen untersuchen. Das Ergebnis sind Reports die keinen Blick "in eine Anwendung" (Abhängigkeiten).

Dies wird besonders dann zur Herausforderung, wenn Software von Drittanbietern eingesetzt wird, deren Software-Abhängigkeiten nicht bekannt sind, wie bspw. der Fall Log4J zeigte [6]. Genau dieses Problem kann mithilfe von Software-Bill-of-Materials (SBOMs) gelöst werden. Eine SBOM ist eine Auflistung mit Versionsinformationen aller Abhängigkeiten und verwendeten Bibliotheken einer Software. Format und Detailgrad einer SBOM können variieren [2] [1] [7].

Design

Das Ziel der Arbeit liegt darin ein Konzept zu entwerfen um Informationen über Software-Abhängigkeiten sowie verwendete Libraries und Frameworks als SBOMs in einem einheitlichen Format bereitzustellen. Diese SBOM kann dann von einer Software entgegengenommen/abgefragt und gespeichert werden um mit ihr weiterzuarbeiten und einen Blick "in die Software" zu erlangen.

Vollständigkeit und Erzeugung der SBOM

Es muss sichergestellt werden, dass die SBOM von den einzelnen Komponenten vollständig und korrekt ist; z.B. können SBOMs beim Kompilieren erzeugt, oder auch beim Installieren durch einen Paket-Manager zusammengetragen, werden. Da Software über unterschiedliche Wege (z.B. über Paket-Manager, manueller Download, portable Installationen) installiert und

ausgeführt werden kann, muss sichergestellt werden, dass eine SBOM für alle Arten von Software erzeugt wird.

Übertragung der SBOM

Ein Angreifer hat ein Interesse daran eine SBOM zu manipulieren um eine ausgenutzte Schwachstelle zu verstecken – oder auch um mögliche Schwachstellen auszumachen. Die SBOMs müssen deshalb zwischen allen beteiligten Komponenten signiert, verschlüsselt und authentifiziert übertragen werden. Hierfür können existierende Protokolle (z.B. HTTPS oder NETCONF) aber auch neue Protokolle oder -erweiterungen betrachtet werden. Zusätzlich muss das Konzept vorsehen, welche Stelle überhaupt Informationen über SBOMs austauscht. Beispielsweise kann eine zentrale Stelle (SBOM-Manager) SBOMs aktiv abfragen, jede Software selbständig ihre SBOM übertragen oder auf einem Host sammelt ein Dienst alle SBOMs von laufenden Prozessen und leitet diese gesammelt weiter.

Unterstützung von "Legacy-Anwendungen"

Da in einem Netzwerk nicht sichergestellt werden kann, dass alle Komponenten bereits über eine Funktion verfügen die eigene SBOM zu verbreiten, muss in der Lösung eine Möglichkeit berücksichtigt werden die SBOM auch für sogenannte "Legacy-Anwendungen" auswerten zu können. Die Pflege von SBOMs kann in diesen Fällen ggf. dennoch automatisiert erfolgen, indem aktuelle SBOMs bspw. von Hersteller-Webseiten heruntergeladen, bei Updates über eine extra SBOM-Datei bereitgestellt, Binaries mithilfe einer Agenten-Software analysiert werden oder herkömmliche Netzwerk-Scanner die SBOM mit so vielen Informationen wie möglich erstellen.

Verwandte Arbeiten

Die NTIA (National Telecommunications and Information Administration) der USA hat 2019 einen Bericht herausgebracht, der u.a. ein einheitliches Format für

SBOMs beschreibt. [2] Andere Arbeiten, wie [1] und [7], nutzen und schlagen andere Formate vor. In dieser Arbeit werden die verschiedenen Motivationen um Details mit der Kompatibilität zum automatischen Generieren gegenübergestellt.

Der NTIA-Bericht beschreibt außerdem, dass eine SBOM bei jedem Release einer Software erzeugt werden muss und dies im Idealfall in den Entwicklungsprozess integriert wird. [2] Während der Bericht ein umfangreiches Konzept zu Inhalt und Struktur von SBOMs liefert wird wenig zu den Möglichkeiten der Erzeugung von SBOMs erläutert. Dieser Aspekt soll in dieser Arbeit betrachtet werden.

In der Studie [8] wurden Open-Source Projekte von 14 Paket-Managern hinsichtlich der Aktualität von Software-Abhängigkeiten untersucht. Ein referenzierter Artikel beschreibt, dass in Projekten mit veralteten Abhängigkeiten Sicherheitslücken viermal so häufig auftreten als in Projekten mit aktuellen Abhängigkeiten [5]. Als Möglichkeit zur Lösung dieses Problems werden Tools wie Dependabot von GitHub vorgeschlagen, welche Projekte automatisch auf veraltete Versionen hin überprüfen. Dieses Tool kann in dieser Arbeit als Grundlage für die automatisierte Erzeugung von SBOMs sowie für Vorschläge zur Prävention dienen.

Ergebnis

Das Ziel der Arbeit ist ein Konzept, wie SBOMs erzeugt und zwischen Komponenten in einem Netzwerk ausgetauscht werden können. Dabei werden insbesondere die Gegebenheiten der industriellen Anwendungsdomänen berücksichtigt. Aspekte anderer Domänen werden in der Arbeit diskutiert.

Vollständigkeit und Erzeugung der SBOM

Wird bei der Software-Entwicklung ein Paket-Manager zur Verwaltung von Abhängigkeiten eingesetzt, kann dieser auch für das Erzeugen einer SBOM verwendet werden. Mithilfe von Tools wie bspw. syft können Informationen zu rekursiven Beziehungen zwischen Abhängigkeiten aus dem Paket-Manager extrahiert werden. Die SBOM kann im Rahmen des Entwicklungsprozesses (bspw. bei jedem Build) automatisiert erzeugt werden.

Eine andere Möglichkeit zum Erstellen von SBOMs bietet die Analyse von Quellcode indem z.B. include-Angaben (bei C: Zeilen beginnend mit `#include`) und Programmiersprachen-spezifische Dateien (bei C: bspw. Makefiles) untersucht werden. Aus diesen Informationen lässt sich dann ebenfalls SBOM zusammenbauen. Der Nachteil dieser Methode liegt darin, dass die erzeugte SBOM ggf. nicht vollständig ist, da die rekursiven Abhängigkeiten einer Abhängigkeit nicht in jedem Fall ermittelt werden können. Hinzu kommt, dass auch die Versionen der gefundenen Abhängigkeiten schwierig und ggf. nur manuell herauszufinden sind.

Ebenso können ausführbare Dateien und Container auf Abhängigkeiten skriptbasiert untersucht werden. Auch hier liegen die Herausforderungen im Finden der rekursiven Abhängigkeiten und im Bestimmen der eingesetzten Versionen.

Als letzte Option bleibt noch das manuelle Erstellen einer SBOM.

Übertragung der SBOM

Je nach Ausprägung eines Netzwerks kann es einfacher und sicherer sein, eine SBOM von den verschiedenen Hosts zu erfragen (z.B. über NETCONF oder HTTP well-known Endpoints). Dabei ist jedoch die Aktualität nicht zuverlässig gegeben, da Änderungen immer erst zu bestimmten Zeiten angefragt werden. Diese Anfragen können jedoch in Zeiten gelegt werden, in denen das Netzwerk bekannterweise unausgelastet ist. Alternativ kann eine Software ihre SBOM auch selbständig beim Installieren oder Aktualisieren übertragen. Dadurch wird die Aktualität der SBOM-Informationen sichergestellt. Als letzte Möglichkeit kann eine SBOM auch unabhängig von Geräten und Software übertragen werden, indem bspw. manuelle Downloads in Kundenportalen oder der Austausch per E-Mail angeboten werden. Auch hier ist wieder die Aktualität eventuell nicht gegeben, da sie von weiteren Faktoren der out-of-band Kommunikation abhängig ist.

Für alle obengenannten Übertragungsmöglichkeiten muss auch der Schutz der SBOM betrachtet werden. Da eine SBOM ggf. Informationen über Schwachstellen einer Software enthält, sollte diese – oder zumindest die Information über den Einsatz der Software – nicht unbedingt öffentlich gemacht werden. Deshalb soll nachfolgend für die drei Übertragungsmöglichkeiten 1.) (teil-)öffentliche Bereitstellung im Internet, 2.) Abfragen von einem Gerät oder einer Anwendung und 3.) out-of-band Kommunikation die Punkte a) Notwendigkeit für Authentifizierung (Wer kann auf die SBOM zugreifen?) und b) Integritätsschutz (Wie wird sichergestellt, dass die SBOM beim Übertragen nicht verändert wird?) betrachtet werden.

Ist eine SBOM öffentlich im Internet verfügbar ist keine Authentifizierung notwendig (ausgenommen z.B. Kundenportale mit Login). Für den Integritätsschutz auf dem Übertragungsweg kann eine TLS-Verbindung verwendet werden. Wird die SBOM direkt von einem Gerät oder einer Anwendung abgefragt, kann bspw. mithilfe von Firewall-Regeln, oder eines Private-Public-Key-Verfahrens (z.B. bei NETCONF über SSH) der Zugriff limitiert werden. Integritätsschutz kann auch hier durch die Verwendung von bspw. TLS oder SSH-signing erzielt werden. Im Fall der out-of-band Kommunikation kann davon ausgegangen werden, dass die SBOM nur an berechtigte Empfänger zur Verfügung gestellt wird. Der Schutz der Integrität der SBOM ist dann abhängig von dem gewählten

Kommunikationsmedium – im Falle von E-Mail bspw. PGP-Signaturen.

Unterstützung von “Legacy-Anwendungen”

Es gibt Situationen in denen eine Bereitstellung der SBOM durch den Hersteller nicht möglich ist. Für diese Fälle gibt es unterschiedliche Möglichkeiten dennoch eine SBOM mit möglichst vielen und korrekten Informationen zusammenzustellen.

Als eine erste Möglichkeit können Netzwerk- und Schwachstellen-Scanner eingesetzt werden, welche eine Software anhand der nach außen verfügbaren Schnittstellen analysieren (“externe Sicht”). Ein Netzwerk-Scanner kann z.B. über (TCP)-Banner-Informationen beim Verbindungsaufbau Informationen über einen Dienst sammeln. Simuliert man bspw. einen SSH-Verbindungsaufbau mit einem OpenSSH Server, wird man mit der Version des OpenSSH Servers begrüßt. Ebenso können Schwachstellen-Scanner aufgrund eines bestimmten Verhaltens und der daraus resultierenden Schwachstelle auf eine mögliche Software schließen.

Eine andere Möglichkeit zum Erzeugen von SBOMs mithilfe von Netzwerk-Scannern sind authentifizierte Scans, bei denen die Netzwerk-Scannern Zugangsdaten für ein System besitzen (um sich anzumelden und Befehle auszuführen). Sind auf einem System alle Voraussetzungen erfüllt um mithilfe eines Skripts eine SBOM des gesamten Systems zu erstellen, kann der

Netzwerk-Scanner im Rahmen eines authentifizierten Scans dieses Skript triggern und dadurch eine SBOM erzeugen.

Schlussfolgerung

Die zuvor beschriebenen Ergebnisse zeigen, dass eine Arbeit mit SBOMs grundsätzlich möglich ist. Mit den verschiedenen Möglichkeiten SBOMs zu erzeugen werden alle gängigen Entwicklungsmöglichkeiten abgedeckt um für jede Software eine SBOM zu erstellen. Auch die Übertragung der SBOM kann wie aufgezeigt über unterschiedliche Wege stattfinden um sich an die Gegebenheiten eines Netzwerks anzupassen. Die größte Herausforderung liegt in der Unterstützung von Legacy Anwendungen. Können Organisationen hier nicht auf die Hersteller zählen, müssen die SBOMs über alternative Wege erzeugt werden, gegebenenfalls unvollständige Informationen liefern. Daher ist der Einsatz von SBOMs als unterstützendes Mittel im Schwachstellenmanagement insbesondere in Umgebungen von Vorteil, die über aktuelle Software verfügen, für welche Updates bereitgestellt werden – oder in denen bereits bei der Auswahl von Software auf das Vorhandensein von SBOMs geachtet wird. Am Ende steht und fällt der Nutzen von SBOMs jedoch auch mit der tatsächlichen Verwendung der Informationen.

Literatur und Abbildungen

- [1] syft anchore. syft. <https://github.com/anchore/syft>, 2022.
- [2] NTIA National Telecommunications and Information Administration. Framing Software Component Transparency: Establishing a Common Software Bill of Material (SBOM). https://ntia.gov/files/ntia/publications/framingsbom_20191112.pdf, 2019.
- [3] Dominik Bredel. Drei Maßnahmen um IT-Sicherheit und Resilienz zu stärken. <https://www.security-insider.de/drei-massnahmen-um-it-sicherheit-und-resilienz-zu-staerken-a-733f535059d46a4f2f3d1ee15352324f/07> 2022.
- [4] BKA Bundeskriminalamt. Bundeslagebild Cybercrime 2021. https://www.bka.de/DE/Presse/Listenseite_Pressemittelungen/2022/Presse2022/220509_PM_CybercrimeBLB.html, 2022.
- [5] Joël Cox, Eric Bouwers, Mark van Eekelen, and Joost Visser. Measuring Dependency Freshness in Software Systems. *Proceedings of the 37th International Conference on Software Engineering - Volume 2*, pages 109–118, 2015.
- [6] Raphael Hiesgen, Marcin Nawrocki, Thomas Schmidt, and Matthias Wählisch. The Race to the Vulnerable: Measuring the Log4j Shell Incident. *Proc. of Network Traffic Measurement and Analysis Conference (TMA '22)*, 2022.
- [7] MS Microsoft. Component Detection. <https://github.com/microsoft/component-detection>, 2022.
- [8] Jacob Stringer, Amjed Tahir, Kelly Blincoe, and Jens Dietrich. Technical Lag of Dependencies in Major Package Managers. *2020 27th Asia-Pacific Software Engineering Conference (APSEC)*, pages 228–237, 2020.

Data Insights im Connected Car Umfeld: Verbesserung der Entscheidungsfindung durch den Business Intelligence Prozess mittels manuellem Testing

Raphael Gschwind

Dirk Hesse

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Mercedes-Benz Group AG, Sindelfingen

Einleitung

In einer zunehmend datengetriebenen Welt gewinnt die Fähigkeit, fundierte Entscheidungen auf Grundlage von Daten zu treffen, eine immer größere Bedeutung für Unternehmen und Organisationen. Unternehmen müssen flexibel auf Veränderungen reagieren und ihre Kosten kontrollieren, um langfristig erfolgreich zu sein. Eine präzise Kenntnis der aktuellen Unternehmens- und Marktlage ist dabei von entscheidender Bedeutung. Im Zeitalter der Digitalisierung und des Internet of Things (IoT) eröffnen sich Unternehmen neue Möglichkeiten, wertvolle Informationen aus verschiedenen Datenquellen, wie Sensordaten oder Protokolldaten, zu gewinnen [7]. Eine wichtige Herausforderung besteht jedoch darin, diese heterogenen Daten effektiv zu erfassen und auszuwerten, um daraus fundierte Entscheidungen ableiten zu können. Die Auswertung dieser Daten ist oft zeitaufwendig und komplex. Daher bedarf es leistungsfähiger Analysetechniken und -werkzeuge, um relevante Informationen aus Daten extrahieren zu können. Zudem müssen die Ergebnisse der Datenanalyse in aussagekräftigen Visualisierungen präsentiert werden, um sie für Entscheidungsträger verständlich und handlungsrelevant zu machen. Um das Management effektiv bei der Planung und Entscheidungsfindung zu unterstützen, greifen Unternehmen heute vorwiegend auf Business Intelligence (BI) zurück. Business Intelligence umfasst verschiedene Methoden, um Unternehmensdaten zu sammeln, zu speichern, aufzubereiten, zu analysieren und zu präsentieren. Dadurch werden wertvolle Erkenntnisse gewonnen, die dazu beitragen, die Geschäftsprozesse zu optimieren und fundierte Entscheidungen zu treffen.

Zielsetzung

Das Ziel der vorliegenden Bachelorarbeit besteht darin, Transparenz über den Datenverbrauch einzelner Services im Connected Car Umfeld von Mercedes-Benz zu

schaffen. Dies soll durch die Auswertung, Verarbeitung und Visualisierung von Diagnoseprotokollen und Trace (DLT) Messages im Kontext des Business Intelligence Prozesses erfolgen, indem mehrere Testings an Connected Cars durchgeführt werden. Auf dieser Basis resultiert ein Ansatz, der zukünftig eine Orientierung bieten soll, Connect Services zu tracken. Dadurch können anfallende Kosten für Datenverbräuche besser kalkuliert und Angebotskonzepte optimiert werden.

Business Intelligence

Business Intelligence befasst sich mit den Prozessen und Werkzeugen zur Analyse von Unternehmensdaten, um daraus verwertbare Erkenntnisse zu generieren. Dies unterstützt Mitarbeiter dabei, qualifizierte Entscheidungen zu treffen. Ein BI-System analysiert sowohl gegenwärtige als auch historische Daten und präsentiert die Ergebnisse in verständlichen Berichten, Dashboards, Grafiken und Diagrammen, die von allen Abteilungen im Unternehmen genutzt werden können. Häufig wird BI auch als „deskriptive Analyse“ bezeichnet, da sie die Leistung eines Unternehmens in der Vergangenheit und Gegenwart beschreibt [5]. Sie beantwortet Fragen darüber, was geschehen ist und welche Veränderungen erforderlich sind. Allerdings fokussiert sie sich nicht auf die Ursachen bestimmter Ereignisse oder zukünftige Entwicklungen [2]. Wie in Abbildung 1 zu erkennen ist, kann der BI-Prozess in folgende Schritte unterteilt werden:

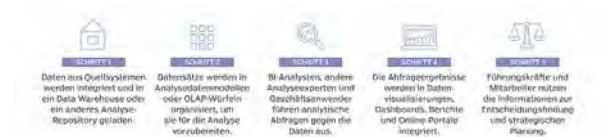


Abb. 1: Der Business Intelligence Prozess [3]

Vorgehensweise

Die methodische Herangehensweise zur Erforschung von Connected Car Systemen umfasst mehrere aufeinanderfolgende Schritte. Zuerst erfolgt eine umfassende Auseinandersetzung mit der Struktur und den Funktionalitäten der Connected Car Systeme. Hierbei liegt ein besonderes Augenmerk auf der eingehenden Analyse des Datenverkehrs, welcher mithilfe eines spezialisierten Datenloggers, dem Bluepirat, erfasst wird. Dieser hat verschiedene Schnittstellen, was in Abbildung 2 zu erkennen ist. Der Bluepirat wird mit dem Bus-System des Fahrzeugs verbunden und ermöglicht die umfassende Aufzeichnung sämtlicher Busprotokolle sowie den Datenverkehr und die Kommunikation zwischen den verschiedenen Systemkomponenten wie etwa elektronischen Steuergeräten. Der Datenverkehr wird daraufhin per Ethernet auf einen Archiv-Server namens CarLA übertragen.



Abb. 2: Schnittstellen eines Bluepirat Datenloggers [6]

CarLA, kurz für Car Log Analyser, stellt ein Tool dar, welches die Analyse und Verwaltung umfangreicher Logdateien von Fahrzeug-Datenbussen verschiedener Steuergeräte ermöglicht. Zusätzlich bietet CarLA präzise Analysefunktionen, die die Extraktion relevanter Informationen für eine detaillierte Untersuchung der Fahrzeugdaten ermöglichen. Um die riesigen Datenmengen auf die relevanten Informationen zu beschränken, wird ein Java-Filter entwickelt, der innerhalb von CarLA ausgeführt wird. Dieser Filter ermöglicht es, gezielt die relevanten Informationen aus dem Standard Header, dem Extended Header und die Payload Informationen aus dem Daten-Frame zu extrahieren, um eine weiterführende Analyse durchzuführen. Der Frame wird in Abbildung 3 dargestellt.

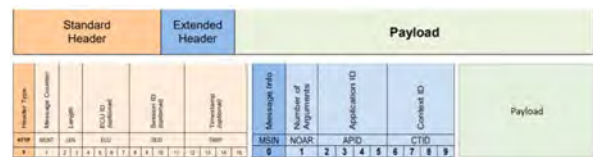


Abb. 3: DLT Message Format [4]

Im Rahmen dieses Prozesses werden die Logs, die von der Steuereinheit der Headunit gesendet und empfangen werden, nach den übertragenen Bytes durchsucht. Anschließend werden die Informationen in ein lesbare Format konvertiert und in einer übersichtlichen CSV-Datei gespeichert. Die weiterführende Analyse und Auswertung der Daten erfolgen unter Verwendung eines geeigneten Reporting-Tools. Hierbei kommt das Tool Splunk zum Einsatz, welches es ermöglicht, die Daten in aussagekräftigen Grafiken und Diagrammen darzustellen. Durch diese visuelle Aufbereitung können komplexe Zusammenhänge veranschaulicht und bedeutsame Erkenntnisse gewonnen werden.



Abb. 4: Datentransfer Carla - Splunk [1]

Ausblick

Die vorliegende Vorgehensweise ermöglicht es, auf Serviceebene Einblicke in Daten zu gewinnen. Man hat die Möglichkeit, den Java-Filter zu übernehmen und auf die eigenen Services anzuwenden, um zukünftig einen verbesserten Überblick über den Datenverbrauch der Connect-Services zu erhalten. Dadurch kann man die Logdaten analysieren und auswerten, was fundierten Entscheidungen und der Kalkulation von Angebotskonzepten durch das Management dient. Diese Herangehensweise eröffnet somit die Möglichkeit, den Datenverbrauch effizienter zu kontrollieren und optimierte Strategien für die Zukunft zu entwickeln.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Stefan Döring. Business Intelligence erklärt. <https://muenchen.digital/blog/business-intelligence-erklaert/>, 01 2021.
- [3] Kevin Ferguson. Business Intelligence Trends für 2021 und darüber hinaus. <https://www.computerweekly.com/de/feature/Business-Intelligence-Trends-fuer-2021-und-darueber-hinaus>, 01 2021.
- [4] Autosar GbR. Log and Trace Protocol Specification. https://www.autosar.org/fileadmin/standards/R22-11/FO/AUTOSAR_PRS_LogAndTraceProtocol.pdf, 11 2022.
- [5] Patrick Kinter. Business Intelligence: Kompakt erklärt. <https://www.alexanderthamm.com/de/blog/business-intelligence-kompakt-erklaert/>, 03 2023.
- [6] Magna Telemotive. Benutzerhandbuch Bluepirat. https://sc.telemotive.de/4/uploads/media/blue_PiraT_Rapid_Benutzerhandbuch.pdf, 04 2020.
- [7] Felix Weber. *Künstliche Intelligenz für Business Analytics - Algorithmen, Plattformen und Anwendungsszenarien*. Springer Vieweg Wiesbaden, 2020.

Untersuchungen zum Bidirektionalen Laden von batterieelektrischen Fahrzeugen mit ISO15118-20 und Aufbau eines Hardware- und Software-in-the-loop Testers

Rahim Haidari

Clemens Klöck

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma eSystems MTG GmbH, Wendlingen am Neckar

Motivation

Die Nutzung der Elektromobilität wird immer wichtiger in der heutigen Gesellschaft. Immer mehr Menschen entscheiden sich für Elektrofahrzeuge, um den CO₂-Ausstoß zu senken und die Umwelt zu schonen. Abgesehen davon, dass immer mehr Elektrofahrzeuge auf den Straßen unterwegs sind, schreitet ebenso die Technologie der Elektromobilität immer schneller voran. Doch Elektroautos sind nicht nur leise und emissionsfrei, sondern haben ebenso das Potenzial, als mobile Energiespeicher eine wichtige Rolle im Energiesystem zu spielen. Hierbei spielt die bidirektionale Ladetechnologie eine wichtige Rolle, die im ISO 15118 Standard definiert ist. Dieser regelt die Kommunikation zwischen Elektrofahrzeug und Ladestation. Der Standard besteht aus einer Reihe von Dokumenten, die sich mit den Anforderungen der OSI-Schichten 1-7

befassen. Die ISO 15118-2 stellt die erste Generation dar, die sich mit den Anforderungen der OSI-Schichten 3-7 befasst, jedoch das bidirektionale Laden nicht unterstützt. Im Gegensatz dazu ermöglicht die neuere Version, die ISO 15118-20, das bidirektionale Laden zwischen Elektrofahrzeug und Ladestation. Ebenso wurden in der ISO 15118-20 die bereits bekannten Ladearten, nämlich das Gleichstromladen (DC) und Wechselstromladen (AC) aus der ersten Generation, um das Laden mit ACDP (Automated Connection Device) und WPT (Wireless Power Transfer) erweitert. Der Fokus in der vorliegenden Arbeit liegt hauptsächlich auf dem AC-Laden, da das Unternehmen sich primär auf die Entwicklung und Implementierung des AC-Ladens konzentriert. Die Abbildung 1 veranschaulicht die Zuordnung der ISO 15118-Dokumentenreihe in der OSI-Schicht.

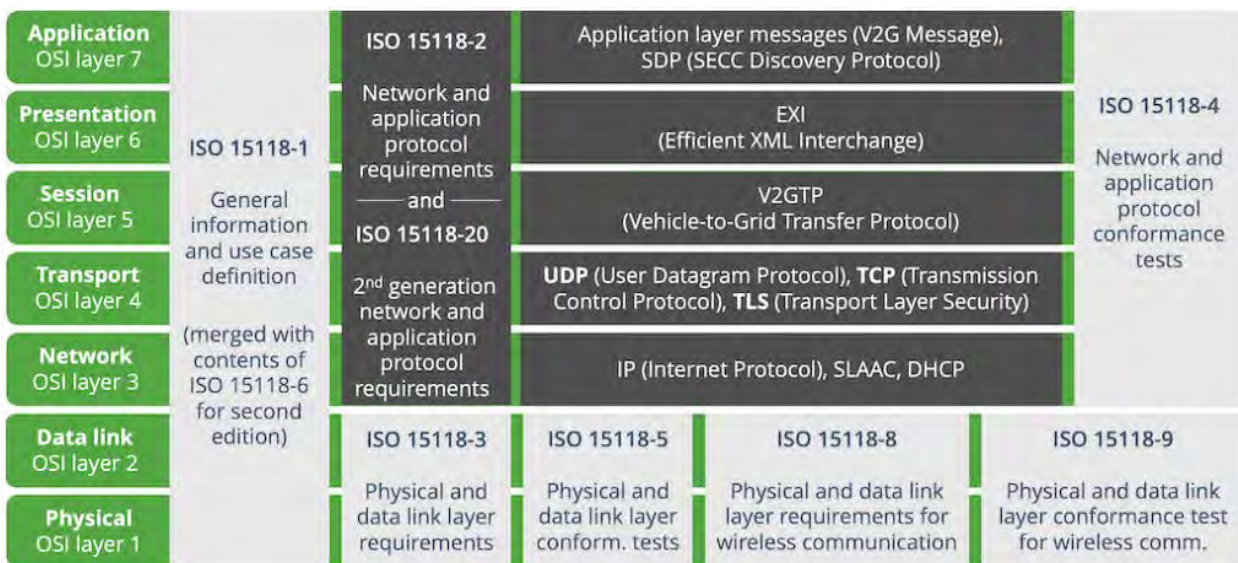


Abb. 1: ISO 15118 Dokumentenübersicht in der ISO/OSI-Schicht [5]

Bidirektionales Laden und Vehicle-to-X

Beim bidirektionalen Laden kann der Strom in beide Richtungen fließen, was bedeutet, dass er entweder vom Ladegerät in das Elektrofahrzeug fließt oder umgekehrt. Beim Ladevorgang wird der Strom von Wechselstrom in Gleichstrom mithilfe eines Gleichrichters umgewandelt. Beim Entladevorgang hingegen muss der Strom im Elektrofahrzeug von Gleichstrom wieder in Wechselstrom umgewandelt werden. Diese Umwandlung wird durch einen Wechselrichter in der Ladestation umgewandelt. Bidirektionales Laden ermöglicht daher, den aus erneuerbaren Energien gewonnenen Strom zu speichern, wenn er im Netz aktuell nicht gebraucht wird. [2] Der Strom aus der Fahrzeugbatterie wird dann bei Bedarf wieder ins Netz eingespeist, sodass eine effizientere Nutzung erneuerbarer Energien ermöglicht wird. Für die Nutzung der bidirektionale Ladetechnologie gibt es verschiedene Anwendungsmöglichkeiten.

- **V2G** steht für „Vehicle-to-Grid“ und bezeichnet die Form, bei der das Elektrofahrzeug mit dem öffentlichen Stromnetz verbunden ist. Elektroautos können während der Standzeit genutzt werden um überschüssige Energie, die aus Sonnen- oder Windkraft gewonnen wird zu speichern. Diese Energie kann dann zu Spitzenverbrauchszeiten oder nachts wieder ins öffentliche Netz zurückgespeist werden, um es zu stabilisieren. Elektrofahrzeuge können erneuerbare Energiequellen effizienter nutzen, so dass Wind- oder Solaranlagen nicht mehr abgeschaltet werden müssen, weil das Stromnetz nicht den Strom aufnehmen kann. [3]
- Mit **V2H**, also „Vehicle-to-Home“, bezeichnet die Form, bei der das Elektrofahrzeug mit dem Haus oder einem Gebäude verbunden ist. Elektroautos können tagsüber an das eigene Hausstromnetz angeschlossen werden, um überschüssigen Strom aus den Solarzellen auf dem Hausdach zu speichern. In der Nacht oder wenn der Strom zu den bestimmten Tageszeiten besonders teuer ist, kann der gespeicherte Strom zurückgespeist werden. [3]
- **V2L** steht für „Vehicle-to-Load“ und bezieht sich darauf, dass der Strom aus einem Elektrofahrzeug für andere Geräte nutzbar ist. [3]

Gridforming und Gridfollowing

Der „GeneratorMode“ ist eine Neuerung in der ISO 15118-20 Norm und war im vorherigen ISO 15118-2 Standard noch nicht enthalten. Der „GeneratorMode“ gibt an, ob das System, welches aus Elektrofahrzeug und Ladestation besteht, als netzfolgender (Gridfollowing) Generator oder als netzbildender (Gridforming)

Generator arbeitet. Beim „Gridfollowing“ gibt das Netz die Spannung und Frequenz vor, denen das System folgen muss. Das System speist lediglich Wirk- und Blindleistung entsprechend den Anforderungen des Netzes ein. Beim „Gridforming“ hingegen ist das System in der Lage, die Spannung und Frequenz des Netzwerks zu steuern und Leitungen mit Strom zu versorgen, die andernfalls nicht versorgt würden. Gewählt werden sollte der Gridfollowing-Generator-Modus, wenn das System an das Verteilernetz angeschlossen ist und nicht als einer der Hauptnetzgeneratoren fungiert. Hingegen sollte das Gridforming-Generator Modus gewählt werden, wenn das System eine entfernte Last oder das Netz eines Hauses mit Strom versorgt. [6]

Control modes

Die ISO 15118-20 stellt zwei Kontrollmodi zur Verfügung, den „Dynamic Mode“ und den bereits in der ISO 15118-2 vorhandenen „Scheduled Mode“. Im „Scheduled Mode“ wird ein Energieprofil zwischen Ladestation und Elektrofahrzeug ausgehandelt. In diesem Modus liegt die Verantwortung beim Elektrofahrzeug, die Mobilitätsbedürfnisse sicherzustellen. Dazu gehört die Berechnung der benötigten Energiemenge, gemäß den Anforderungen des Fahrers aufzuladen, sowie die Festlegung des Zeitpunkts, zu dem der Fahrer die Ladestation verlassen muss. Dabei werden auch die Leistungsgrenzen der Ladestation berücksichtigt. Im „Dynamic Mode“ hingegen unterliegt es der Ladestation, die Mobilitätsbedürfnisse zu erfüllen, ohne dass eine Verhandlung mit dem Elektrofahrzeug erforderlich ist. Das Elektrofahrzeug übermittelt an die Ladestation die Parameter, welche die Leistungsdaten der Batterie bestimmen. Die Ladestation sendet dann einzelne Leistungswerte an das Elektrofahrzeug, an die das Elektrofahrzeug halten soll. Der „Dynamic Mode“ ist besonders nützlich, wenn das Stromnetz instabil ist und eine schnelle Reaktion zur Regulierung der Frequenz erforderlich ist. [4]

V2G-Kommunikation in der ISO 15118-20

In Abbildung 2 wird der Kommunikationsfluss mit den entsprechenden Nachrichten veranschaulicht, welche im Folgenden näher erläutert werden.

- **SECC discovery protocol Req/-Res (SDP):** Die Allererste Nachricht, die vom Elektrofahrzeug gesendet wird, ist die SDP-Anfrage, mit der das Elektrofahrzeug die IP-Adresse und Portnummer der Ladestation erhält. Diese Anfrage wird über UDP (User Datagram Protocol) als Multicast-Nachricht auf Port 15118 gesendet. [6]

- **SupportedAppProtocolReq/Res:** Das erste XML-basierte Nachrichtenpaar, das über TLS erfolgt, ist das „SupportedAppProtocol“. Mit dieser Nachricht kann das Elektrofahrzeug alle unterstützten Protokolle wie ISO 15118-2 oder ISO 15118-20 mitteilen. Die Protokolle werden in einer bestimmten Priorität übermittelt, basierend darauf, welche das Elektrofahrzeug bevorzugt, nutzen möchte. Je nach was die Ladestation unterstützt wird das entsprechende Protokoll nach Priorität ausgewählt. [6]
- **SessionSetupReq/Res:** Durch die Nachricht „SessionSetup“ baut das Elektrofahrzeug eine V2G-Kommunikationssitzung auf. Das Elektrofahrzeug erhält durch diese Anfrage eine Session-ID von der Ladestation zugewiesen, mit der eine Sitzung aufgrund einer Unterbrechung, wie beispielsweise einer Pause, wieder aufgenommen werden kann. [6]
- **AuthorizationSetupReq/Res:** Durch das Senden der „AuthorizationSetupReq“ fordert das Elektrofahrzeug die verfügbaren Autorisierungsmethoden von der Ladestation an. Die Ladestation antwortet mit der „AuthorizationSetupRes“ Nachricht und stellt Informationen über die verfügbaren Autorisierungsmodi (Plug & Charge und EIM) sowie die Verfügbarkeit des Zertifikat-Installationsdienstes bereit. [6]
- **CertificateInstallationReq/Res:** Wenn das Elektrofahrzeug die Autorisierungsmethode PnC (Plug and Charge) auswählt, erfordert dies ein gültiges digitales Zertifikat. Das Elektrofahrzeug kann ein solches Zertifikat mithilfe der „CertificateInstallation“ Nachricht installieren. Diese Nachricht wird ebenso verwendet, wenn das Elektrofahrzeug kein gültiges Vertragszertifikat besitzt, das aktuelle Zertifikat bald abläuft, ein neues Zertifikat benötigt wird oder das Elektrofahrzeug einfach prüfen möchte, ob ein neues Vertragszertifikat verfügbar ist. [6]
- **AuthorizationReq/Res:** Das Elektrofahrzeug wählt über die Nachricht „Authorization“ eine Autorisierungsmethode aus, entweder PnC (Plug and Charge) oder einer RFID-Karte (EIM). [6]
- **ServiceDiscoveryReq/Res:** Über die „ServiceDiscovery“ Nachricht fragt das Elektrofahrzeug alle verfügbaren Dienste von der Ladestation ab. In der Antwort kann die Ladestation angeben, ob sie den bidirektionalen oder nur den unidirektionalen Ladedienst anbietet. [6]
- **ServiceDetailReq/Res:** Mit der Nachricht „ServiceDetail“ kann das Elektrofahrzeug detaillierte Informationen über einen Ladedienst erhalten, den es nutzen möchte. [6]
- **ServiceSelectionReq/Res:** Mithilfe der „ServiceSelection“ Nachricht kann das Elektrofahrzeug einen Ladedienst auswählen, wie beispielsweise AC, AC-BPT, DC, DC-BPT, DC-ACDP, DC_ACDP_BPT oder WPT. [6]
- **AC_ChargeParameterDiscoveryReq/Res:** Die „AC_ChargeparameterDiscovery“ Nachricht ermöglicht die Verhandlung von Parametern wie der maximalen Ladeleistung (MaximumChargepower) und der maximalen Entladeleistung (MaximumDischargePower) zwischen dem Elektrofahrzeug und der Ladestation. Im Dynamic Mode können sich die Parameter während des „AC_Charge_Loop“ dynamisch verändern. [6]
- **ScheduleExchangeReq/Res:** Durch das Senden der „ScheduleExchange“ Nachricht teilt das Elektrofahrzeug der Ladestation seine Energieübertragungsparameter mit. Diese Parameter umfassen die geschätzte Energiemenge, die für das Laden des Elektrofahrzeugs benötigt wird, sowie den Zeitpunkt, zu dem der Fahrer beabsichtigt, die Ladestation zu verlassen. [6]
- **PowerDeliveryReq/Res:** Die „PowerDelivery“ Nachricht kennzeichnet den Zeitpunkt, an dem die Ladestation mit der Energieversorgung beginnen kann oder das Elektrofahrzeug mit dem Entladen beginnen kann. Das Elektrofahrzeug tauscht über die Nachricht auch ein Ladeprofil aus und hat die Möglichkeit, sowohl das Laden zu stoppen als auch zu pausieren. [6]
- **AC_ChargeLoopReq/Res:** Wenn die „AC_ChargeLoop“ Nachricht zwischen dem Elektrofahrzeug und der Ladestation ausgetauscht wird, findet ein kontinuierlicher Datenaustausch sowohl für das Laden als auch für das Entladen statt. Dabei werden fortlaufend Daten zwischen dem Elektrofahrzeug und der Ladestation ausgetauscht, um den Lade- und Entladevorgang zu steuern und zu überwachen. Im Dynamic Mode besteht die Möglichkeit, die Energieübertragungsparameter, die in der „ChargeparameterDiscovery“ Nachricht festgelegt sind, dynamisch zu verändern, je nach den Anforderungen der Ladestation. [6]
- **SessionStopReq/Res:** Die „SessionStop“ Nachricht ermöglicht die Beendigung der V2G-Kommunikationssitzung. [6]

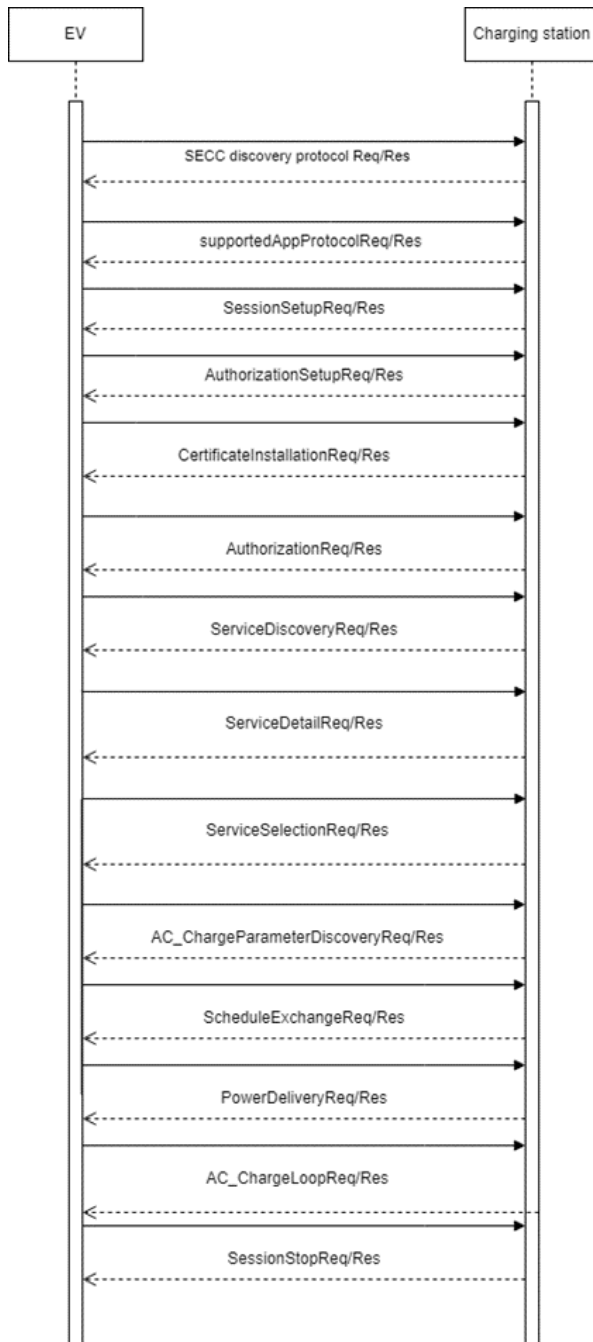


Abb. 2: AC-Nachrichtensequenz [1]

Ziel der Arbeit und Ausblick

Das Ziel dieser Abschlussarbeit besteht darin, eine eingehende Untersuchung der ISO 15118-20 Norm durchzuführen und einen Elektrofahrzeugsimulator zu entwickeln, der die Möglichkeit bietet, das zukünftig im Unternehmen entwickelte Ladegerät gemäß den Spezifikationen des ISO 15118-20-Standards zu testen. Für die Implementierung des Elektrofahrzeugsimulators wird das Framework Josev von der Firma Switch verwendet, welches auf GitHub verfügbar ist. Das Josev Framework ist eine Referenzimplementierung, die die Entwicklung einer Vehicle-2-Grid-Kommunikationsschnittstelle nach ISO 15118 Standard unterstützt. Im Rahmen dieser Arbeit soll das Framework untersucht und erweitert werden, da diese nicht alle Spezifikationen des Standards umsetzt. Die Steuerung des Simulators soll über ein Webinterface erfolgen, das es dem Nutzer ermöglicht, das simulierte Elektrofahrzeug an der Ladestation anzuschließen oder zu trennen sowie andere Aktionen durchzuführen. Für die Entwicklung des Webinterface wird das Vue.js-Framework verwendet.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Team. energielösung. Bidirektionales Laden: Wenn der Ladevorgang beim Elektroauto in beide Richtungen funktioniert. <https://www.energieloesung.de/magazin/bidirektionales-laden-wenn-der-ladevorgang-beim-elektroauto-in-beide-richtungen-funktioniert/>, 02 2022.
- [3] Maltzan Jörg. So wird das Elektroauto zum Energiespeicher. <https://www.autobild.de/artikel/wallbox-e-auto-bidirektionales-laden-v2g-v2h-v2l-22300157.html>, 03 2023.
- [4] Marc Mültin. The new features and timeline for ISO 15118-20. <https://www.switch-ev.com/blog/new-features-and-timeline-for-iso15118-20>, 2020.
- [5] Marc Mültin. What is ISO 15118? <https://www.switch-ev.com/blog/what-is-iso-15118>, 10 2021.
- [6] International Organization for Standardization. *ISO 15118-20:2022-04 Road vehicles - Vehicle to grid communication interface - Part 20: 2nd generation network layer and application layer requirements. International Organization for Standardization, 2022.*, volume 1. International Organization for Standardization, 2022.

Entwicklung einer Anwendung zur Erkennung und Darstellung von Sensoren und deren Wirkbereichen in Augmented Reality

Sami Hazzouri

Jürgen Koch

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Pilz GmbH & Co. KG, Ostfildern

Motivation und Zielsetzung

In unserer zunehmend digitalisierten Welt befinden sich Mensch und Maschine vermehrt in gemeinsamen Arbeitsumgebungen und interagieren direkt miteinander. Diese Entwicklung erfordert erhöhte Sicherheitsmaßnahmen, um Arbeitsunfälle zu vermeiden und die Sicherheit der Menschen zu gewährleisten [2] [5]. Eine mögliche Lösung zur Erfüllung dieser Sicherheitsanforderungen besteht in der Verwendung von Sicherheitssensoren, die in der Lage sind, die Arbeitsumgebung kontinuierlich zu überwachen und die Maschinen bei potenzieller Gefahr in den Sicherheitszustand zu versetzen. Um diese Lösung erfolgreich umzusetzen, ist eine präzise Platzierung und Konfiguration der Sicherheitssensoren unerlässlich, was hohe Ansprüche an die Sicherheitsplanung stellt. Angesichts dieser herausfordernden Situation erweist sich Augmented Reality (AR) als äußerst vielversprechendes Werkzeug, um sowohl die Sicherheitsplanung als auch den Vertrieb digital zu unterstützen.

Das Ziel dieser Arbeit ist die Entwicklung einer mobilen Anwendung, die Augmented Reality und LiDAR-Scanner verwendet, um den Vertrieb und die Sicherheitsplanung zu unterstützen.

Augmented Reality

Unter Augmented Reality (AR) wird die Integration virtueller computergenerierter Objekte in die Wahrnehmung der realen Welt verstanden. Dabei werden keine komplett neuen virtuellen Welten generiert, sondern die Wahrnehmung der realen Welt durch virtuelle Objekte erweitert, sodass diese virtuellen Objekte als Teil der realen Welt erscheinen [6] [12].

Ein Augmented-Reality-System muss die folgenden drei Kriterien erfüllen [1] [8]:

- Kombiniert real und virtuell
- Interaktiv in Echtzeit

- Dreidimensionaler Bezug virtueller und realer Objekte

Um die Wahrnehmung der Realität durch virtuelle Elemente zu erweitern, muss ein AR-System als Erstes die reale Umgebung erfassen. Die erfassten Daten der realen Umgebung werden anschließend analysiert. Basierend auf der Erfassung und Analyse der realen Umgebung erfolgt zunächst die Mischung der realen und virtuellen Elemente. Zuletzt wird das Ergebnis dem Benutzer präsentiert. Der Benutzer kann sowohl mit der realen Umgebung als auch mit den virtuellen Objekten interagieren [9] (siehe Abbildung 1).

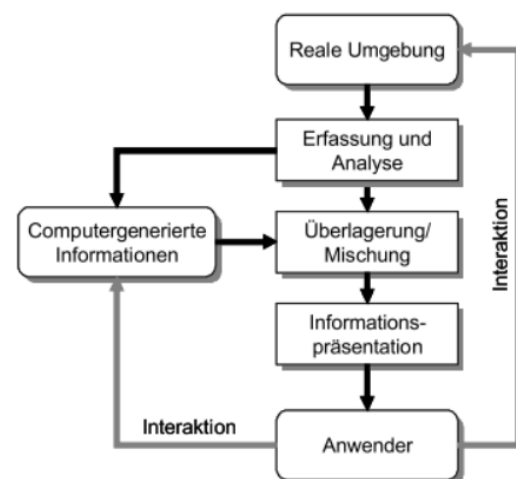


Abb. 1: Funktionsprinzip eines Augmented-Reality-Systems [9]

LiDAR

Unter LiDAR (Light detection and ranging oder Laser imaging, detection and ranging) versteht man ein aktives, optisches Messverfahren, das verwendet wird,

um Eigenschaften von Objekten durch das Senden und Empfangen elektromagnetischer Strahlung zu erfassen. Ein LiDAR-Sensor besteht schematisch aus einem Laser, einem Sendeobjektiv, einem Detektor und einem Empfangsobjektiv [13] [4].

Man kann bei LiDAR-Sensoren zwischen vier verschiedenen Messprinzipien unterscheiden [13]. Für die vorliegende Arbeit ist die Laufzeitmessung (Time-of-Flight (ToF)) relevant. Ein ToF-LiDAR-Sensor sendet

Lichtstrahlung in Form von kurzen Pulsen im Nanosekundenbereich auf ein Objekt. Im Anschluss wird die Leistung der Pulse, die vom Objekt zurück in Richtung des Detektors gestreut werden, zeitlich versetzt zum Sendepuls gemessen. Die Distanz zwischen dem Sensor und dem Objekt wird aus der Lichtgeschwindigkeit im Vakuum und der Zeitdifferenz zwischen dem Empfangs- und Sendepuls berechnet [13] (siehe Abbildung 2).

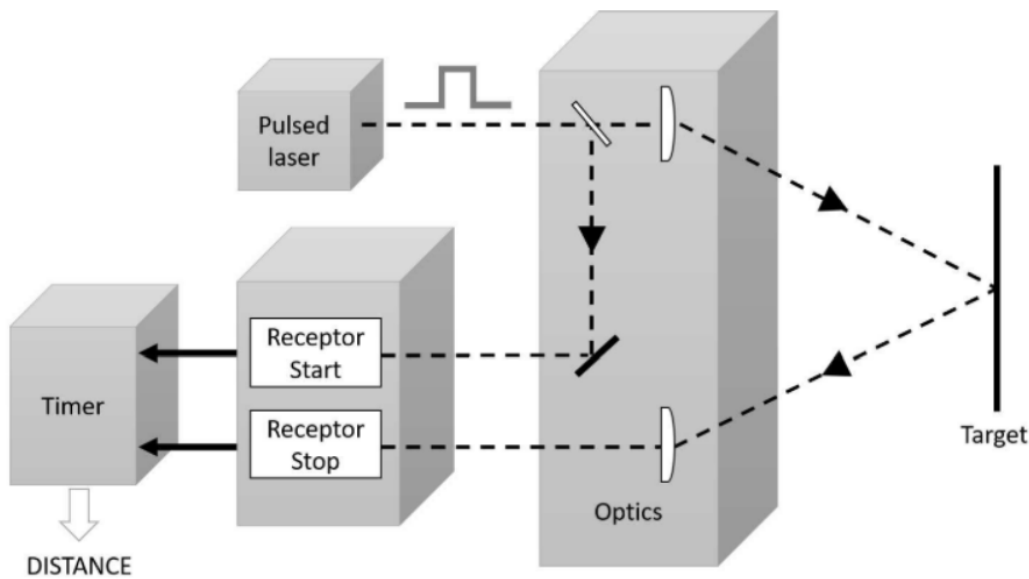


Abb. 2: ToF-Messprinzip [11]

Durch das ToF-Messprinzip wird die Distanz zwischen dem Sensor und einem festen Punkt des Zielobjekts berechnet, dadurch wird aber nicht das gesamte Zielobjekt erfasst. Für die 3D-Objekterfassung müssen diese punktuellen Distanzmessungen auf das gesamte Zielobjekt bzw. auf das gewünschte Sichtfeld ausgedehnt werden. Diese Distanzmessungen werden im Anschluss in einem kartesischen Koordinatensystem abgebildet und mit Tiefeninformationen in eine 3D-Punktwolke oder ein 3D-Mesh transformiert [13] [11] [10] [7].

LiDAR-Sensoren gelten als große Unterstützung für AR-Systeme, da sie eine hohe Genauigkeit und Auflösung bei der Erfassung der realen Umgebung bieten und auch bei schlechten Lichtverhältnissen funktionieren.

Ergebnisse

Der erste Prototyp der zu entwickelnden Anwendung unterstützt zwei Anwendungsfälle. Der erste Anwendungsfall tritt auf, wenn beispielsweise ein Vertriebsmitarbeiter die Funktionalität der Sensoren erklären oder ein Kunde die Sicherheitsplanung für eine Anlage durchführen will, in der realen Umgebung

aber keine physischen Sensoren vorhanden sind. Die Anwendung ermöglicht dem Benutzer, beliebig viele und gegebenenfalls unterschiedliche virtuelle Sensoren in AR darzustellen, zu positionieren und zu rotieren. Darüber hinaus bietet die Anwendung die Möglichkeit, die Wirkbereiche der dargestellten Sensoren zu visualisieren und ihre Darstellung in Echtzeit zu konfigurieren (siehe Abbildung 3). Diese Funktion unterstützt den Bestells- und Planungsprozess, indem sie dem Benutzer eine präzise Vorstellung von den Sensorplatzierungen vermittelt und ihm ermöglicht, die optimalen Konfigurationen für die gegebene Situation zu planen.

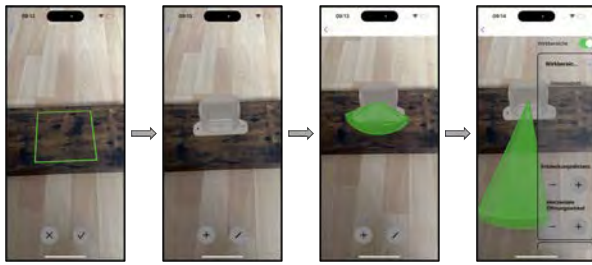


Abb. 3: Anwendungsfall 1 (Von links nach rechts: (1) Kein Sensor existiert (2) Ein virtueller Sensor wurde in AR dargestellt (3) Die Wirkbereiche des Sensors wurden in AR dargestellt (4) Die Wirkbereiche sind in Echtzeit konfigurierbar) [3]

Der zweite Anwendungsfall tritt beispielsweise auf, wenn ein Kunde bereits einen Sensor besitzt und er ihn neu konfigurieren oder umplatzieren will. Die Anwendung ermöglicht dem Benutzer, die physisch existierenden Sensoren zu erkennen und die Darstellung ihrer Wirkbereiche in AR zu visualisieren und in Echtzeit zu konfigurieren (siehe Abbildung 4). Hierbei handelt es sich aber nur um die visuelle Darstellung und Konfiguration der Wirkbereiche der Sensoren,

auf die Sensoren selbst und ihre Konfigurationen greift die Anwendung jedoch nicht zu. Durch diese Funktion kann der Anwender die tatsächliche Position der Sensoren erfassen und die Darstellung ihrer Wirkbereiche visualisieren und konfigurieren, um auf dieser Grundlage präzise Anpassungen und Konfigurationen zu planen. Dies erleichtert dem Kunden die Umsetzung von Veränderungen in der Sensorplatzierung und trägt dazu bei, die Sicherheit und Effizienz in der Arbeitsumgebung zu optimieren.



Abb. 4: Anwendungsfall 2 (Von links nach rechts: (1) Ein Sensor existiert (2) Beim Drücken auf den Suche-Button beginnt die Suche nach einem Sensor (3) Der Sensor wird erkannt (4) Die Wirkbereiche werden in AR dargestellt) [3]

Literatur und Abbildungen

- [1] Ronald T. Azuma. A Survey of Augmented Reality. *Presence: Teleoperators & Virtual Environments*, pages 355–385, 1997.
- [2] Eva A. C. Bittner, Sarah Oeste-Reiß, Philipp Alexander Ebel, and Matthias Söllner. Mensch-Maschine-Kollaboration: Grundlagen, Gestaltungsherausforderungen und Potenziale für verschiedene Anwendungsdomänen. *HMD Praxis der Wirtschaftsinformatik*, 56:34–49, 2019.
- [3] Eigene Darstellung.
- [4] Patrick Dietrich. Methoden zur Reduktion der Messlatenz von GOBO-Projektor-basierten 3D-Sensoren, 2022.
- [5] Fraunhofer-Institut für Kognitive Systeme IKS. Mensch-Roboter-Kollaboration: Wie wir in Zukunft mit dem Roboter zusammenarbeiten werden. <https://www.iks.fraunhofer.de/de/themen/industrie-40/mensch-roboter-kollaboration.html>, 2020.
- [6] Anke Hafner. Mobile Assistenzsysteme in der Intralogistikplanung der Automobilindustrie – Gestaltung, Nutzen und Akzeptanz Augmented Reality-basierter Mensch-Maschine-Schnittstellen, 2022.
- [7] Olaf Keitsch. Entwicklung einer Sensorfusion basierten, UAV gestützten Navigationslösung in geschlossenen baulichen Strukturen unter Verwendung von 3D-Raummodellen, 2020.
- [8] Anett Mehler-Bicher and Lothar Steiger. *Augmented Reality: Theorie und Praxis*. De Gruyter Oldenbourg, 2 edition, 2014.
- [9] Christian Patron. Konzept für den Einsatz von Augmented Reality in der Montageplanung, 2005.
- [10] Bastian Plaß and Thomas Klauer. Digital Flash LiDAR Punktwolken – Consumer-Produkt oder geodätische Zukunftstechnologie? *DVW Schriftenreihe*, 2022.
- [11] Santiago Royo and Maria Ballesta-Garcia. An Overview of Lidar Imaging Systems for Autonomous Vehicles. *Applied Sciences*, 2019.
- [12] Dieter Schmalstieg and Tobias Höllerer. *Augmented Reality: Principles and Practice*. Addison-Wesley, 2016.
- [13] Manuel Trierweiler. Detektion und Klassifizierung unterschiedlicher Bedeckungen auf Abschlusscheiben von LiDAR-Sensoren, 2022.

Evaluation und Vergleich verschiedener Methoden zur prädiktiven Instandhaltung von CNC-Maschinen anhand von zeitbasierten Maschinendaten

Johannes Hofmayer

Clemens Klöck

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma WAFIOS AG, Reutlingen

Einleitung

Die WAFIOS AG entwickelt und produziert mit ihren rund 1000 Mitarbeitern über 120 verschiedene Maschinentypen. Durch hohe Innovationskraft zählt die WAFIOS AG zu den weltweit führenden Herstellern von Draht, Rohr, Ketten und Federbiegemaschinen. Auch die WAFIOS AG ist dem Trend von Industrie 4.0 und IoT gefolgt und hat deshalb seine neusten Maschinengenerationen mit einer OPC-UA Schnittstelle ausgestattet. Neben der reinen Zustandsüberwachung sollen über diese Schnittstelle auch Diagnosedaten zugänglich gemacht werden, um Defekten an den Maschinen vorzubeugen und so Ausfälle zu vermeiden.

Das Fingerprintkonzept

Mit dem Fingerprint-Konzept versucht die WAFIOS AG, erste Auffälligkeiten an den Achsen durch Datenanalysen festzustellen. Durch die Vorhersage eines sich anbahnenden Defektes kann dieser frühzeitig behoben und ein Ausfall der Maschine verhindert werden. Die Datengrundlage zur Bewertung der Achsen, wird anhand eines standardisierten CNC-Messprogramms geschaffen. Position, Geschwindigkeit und Beschleunigung werden durch das Fingerprint Messprogramm vorgegeben und durch die Regler der Maschine ausgegelt. Das Drehmoment ist hierbei die Stellgröße, weshalb sich defekte in den Drehmomentdaten am leichtesten zu identifizieren lassen. [2]

Ziel der Arbeit

Ziel der Arbeit ist es, die Achsen auf Grundlage des Fingerprint-Konzepts automatisiert zu bewerten. Die zeitbasierten Maschinendaten sollen so aufbereitet werden, dass diese eine einfache Klassifikation in 'Gut' / 'Schlecht' zulassen. Es sollen hierfür Kennzahlen berechnet werden, die sehr sensitiv auf Abweichungen einer Achse zur Norm reagieren und diese messbar

machen. Für die Klassifikation sollen mehrere geeignete Ansätze miteinander verglichen werden.

Data Preprocessing

Aus Datenreihen defekter Achsen wurden Merkmale extrahiert, die die bekannten Fehler möglichst gut repräsentieren. Um die Merkmale messbar zu machen, wurde aus diesen jeweils eine Kennzahl abgeleitet, die als Grundlage der Klassifikation dient. Um Schwingungen zu erkennen, wurde eine Frequenzanalyse durchgeführt. Verkippungen im Drehmomentverlauf wurden durch Regressionsgeraden messbar gemacht. Erhöhtes Drehmoment wurde durch die Analyse des Momentes im Vor- und Rückwärtslauf ermittelt. Elf gelabelte Datensätze, mit einem Umfang von durchschnittlich 25,18 Datenpunkten, wurden auf Grundlage der berechneten Kennzahlen durch die vier Klassifikationsmethoden klassifiziert.

Verwendete Methoden zur Klassifizierung

Zur Klassifikation wurden die vier Ansätze DBSCAN, OPTICS, Support Vector Machine (SVM) und ein Naiver Ansatz miteinander verglichen. Die Dichte basierten Clustering Verfahren DBSCAN und OPTICS teilen die Daten anhand ihrer Merkmale in Gruppen ein und sind alleine nicht in der Lage die Daten zu klassifizieren [3]. Die beiden Ansätze wurden deswegen durch einen Nearest Neighbour Ansatz erweitert. Wird ein Cluster durch DBSCAN oder OPTICS gebildet, wird dem gesamten Cluster das Label zugewiesen, welches am häufigsten innerhalb des Clusters auftritt. Punkte die durch das Clustering als Outlier erkannt wurden, werden immer als 'Schlecht' klassifiziert. Zur Bewertung der Daten durch den Naiven Ansatz wurden die einzelnen Messkurven mit einer Durchschnittskurve verglichen und die Abweichungen voneinander als

Fehlerflächen berechnet. Annahme ist, dass defekte Achsen größer von der Durchschnittskurve abweichen, als unauffällige Achsen und diese so voneinander getrennt werden können. Alle Achsen mit Flächenwerten größer dem optimierten Trennwert werden später als 'Schlecht' eingestuft. Die SVM wird anhand von gelabelten Datensätzen auf die Achsdaten trainiert und ist im Anschluss in der Lage diese zu Trennen. Als Kernelfunktion wird eine Polynom dritten Grades verwendet.

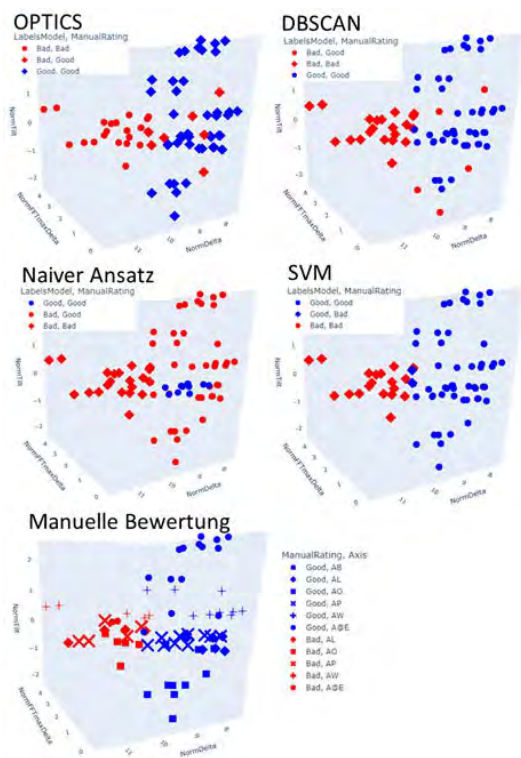


Abb. 1: Klassifikationsergebnisse durch OPTICS, DBSCAN, Naiver Ansatz und SVM auf einem Achsdatensatz. Zusätzlich sind die manuell erstellten Labels zusehen. Rot dargestellte Punkte sind als 'Schlecht', blau dargestellte Punkte als 'Gut' klassifiziert [1]

Ergebnisse der Arbeit

Problematisch bleibt auch zum Ende der Arbeit hin die Datenlage. Insgesamt wurden 277 Fingerprints auf elf Datensätzen verteilt validiert. Durchschnittlich besitzt ein Datensatz dabei 25,2 Fingerprints. Der größte

Datensatz besitzt einen Umfang von 72 Fingerprints, der kleinste acht. Die errechneten Kennzahlen zur Verkipfung (tilt), dem Drehmomentsprung (delta) und der Frequenzanalyse (FFT), können die Achszustände gut abbilden. Werden die Kennzahlen als Datenpunkte im 3D-Raum dargestellt, können visuell Cluster erkannt werden. Aufgrund der Lage der Datenpunkte können auch Rückschlüsse auf die mechanischen Eigenschaften der Achsen getroffen werden.

Im Allgemeinen liefert der naive Ansatz die schlechtesten Kennzahlen über alle Datensätze mit einem durchschnittlichen F-Maß von 0,692 und einer Treffgenauigkeit $Acc = 0,7$. Eine Klassifizierung rein aufgrund der abweichenden Flächen des Drehmomentverlaufs zum durchschnittlichen Drehmomentverlauf scheint der Problemstellung nicht gerecht zu werden.

OPTICS unterliegt DBSCAN mit einem F-Maß = 0,730 und einem $Acc = 0,838$. Abgesehen zweier Datensätze zeigt OPTICS eine passable Klassifizierungsleistung. Die SVM kann häufig die Daten sehr gut klassifizieren und erreicht bei acht von elf Datensätzen das beste F-Maß. Aufgrund der großen Schwierigkeiten der SVM bei zwei Datensätzen, erreicht das F-Maß dennoch nur einen Durchschnittswert von 0,82. Die große Standardabweichung von 0,296 spiegelt diese Schwankungen in den Fähigkeiten des Modells sich an die Daten anzupassen wider.

Den anderen Methoden überlegen zeigt sich DBSCAN mit einem F-Maß von 0,897 und einer Treffgenauigkeit während der Klassifizierung von $Acc = 0,923$. Außerdem sind die Standardabweichungen der beiden Kennzahlen die geringsten und deuten auch auf die höchste Konsistenz der Klassifizierungsleistung hin.

Zur Klassifizierung der Achsen eignet sich DBSCAN im Vergleich zu den anderen Ansätzen am besten, weshalb eine weitere Verwendung von DBSCAN auch im späteren System am sinnvollsten zu sein scheint. Die Treffgenauigkeit von 92,3 % scheint ausreichend groß, um robuste Ergebnisse zu erzielen und so den manuellen Klassifizierungsaufwand zu reduzieren. Die Kombination aus DBSCAN und dem Nearest Neighbor Ansatz tendiert eher dazu Achsen falsch negativ zu klassifizieren, da diese als Outlier erkannt wurden. Defekt markierte Achsen werden nachkontrolliert, weshalb die falschen Klassifikationen in der Nachkontrolle meist identifiziert werden.

Die Ergebnisse der bisherigen Analyse werden als vielversprechend angesehen. Das Fingerprint-Projekt soll fortgeführt werden.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Claus Denkinger, Manuel Euper, and Joschija Merkle. *Dokumentation Fingerprint*. WAFIOS AG, Digital Services, 2022.
- [3] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Knowledge Discovery and Data Mining*, 1996.

Evaluierung von Einsatzmöglichkeiten für Post-Quanten-Kryptografie-Algorithmen im Automobilumfeld

Tolgahan Kandemir

MarkusENZweiler

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Vector Informatik GmbH, Stuttgart

Motivation und Problemstellung

Die Entwicklung von Quantencomputern stellt mithilfe effizienterer Algorithmen ein Risiko für die heutzutage eingesetzte Kryptografie dar. Die Quantentechnologie ermöglicht es, große Primzahlen in Polynomialzeit mit Algorithmen wie dem Shor-Algorithmus [5] oder dem Grover-Algorithmus [2] zu faktorisieren und das Problem des diskreten Logarithmus in linearer Zeitkomplexität zu lösen. Für die Ausführung dieser Algorithmen sind Quantencomputer mit ausreichend großen und fehlerfreien Qubits erforderlich, die bei der Konzeption dieser Algorithmen noch nicht verfügbar waren. Mit den aktuellen Fortschritten bei Quantencomputern wächst das Potenzial, dass die heutige Kryptografie in naher Zukunft von leistungsstarken Quantencomputern gebrochen wird.

Wie viele andere Industriezweige ist auch die Automobilbranche durch die aufkommende Bedrohung durch Quantencomputer betroffen. In vielen Bereichen des Fahrzeuges wird die asymmetrische Kryptografie für den sicheren Schlüsselaustausch, der Verschlüsselung von Kommunikationskanälen oder der Authentizitätsprüfung von Nachrichten und Schlüsseln verwendet. Aufgrund der zunehmenden Bedrohung besteht die Möglichkeit, dass digitale Signaturen von Softwarekomponenten gefälscht werden. Dadurch könnte potenziell manipulierte Software in das Fahrzeug eingeschleust und ausgeführt werden, was ein Risiko darstellt. Dies hat zur Folge, dass beliebiger Code ausgeführt und katastrophaler Schaden angerichtet werden kann.

Zielsetzung

Ziel dieser Arbeit ist es, Aussagen über die Tauglichkeit von PQC-Verfahren für Anwendungsgebiete im Automobilbereich zu treffen. Zunächst müssen die Anwendungsgebiete, welche von der aufkommenden Bedrohung betroffen sind, ermittelt werden. Anschließend sollen die Anforderungen des jeweiligen Use-Cases

an das Signaturverfahren abgebildet werden. Die Länge des öffentlichen Schlüssels, Länge der Signatur oder die Verifikationsgeschwindigkeit sind Beispiele für Kenngrößen, die bei der Anforderungsanalyse eine wichtige Rolle spielen. Anhand der ausgearbeiteten Anforderungen sollen geeignete PQC-Verfahren ausgewählt werden, die im nächsten Schritt mithilfe von Benchmarks miteinander verglichen werden. Für die Durchführung der Benchmarks wird ein Messkonzept entwickelt. Zunächst werden allgemeine und anschließend Use-Case spezifische Benchmarks durchgeführt. Es werden die Use-Cases Secure-Boot und Firmware-Updates betrachtet.

Standardisierungsverfahren des NIST

Das *National Institute of Standards and Technology* (NIST) hat die aufkommende Bedrohung durch Quantencomputer erkannt und 2016 ein Auswahlverfahren zur Standardisierung von quantensicheren Verschlüsselungs-/Schlüsselaustauschalgorithmien und Signaturverfahren gestartet [3]. Unter anderem wurden *gitterbasierte*, *hashbasierte*, *codebasierte* und *multivariatebasierte* Verfahren eingereicht. Mit der Beendigung der dritten Auswahlrunde hat das NIST 2022 bekannt gegeben, dass insgesamt drei Signaturverfahren zur Standardisierung ausgewählt werden:

- **CRYSTALS-Dilithium** (gitterbasiert)
- **Falcon** (gitterbasiert)
- **SPHINCS+** (hashbasiert)

Konzeption

Für Evaluierungszwecke wurde ein Messinstrument entwickelt, das die Nachbildung und Messung von Automotive Use-Cases ermöglicht. Im Rahmen dieser Arbeit wurden die Use-Cases Secure-Boot und

Firmware-Updates nachkonstruiert und das Signieren von Nachrichten sowie das Verifizieren von Signaturen gemessen. Durch häufige Iterationen entstehen Cache-Effekte, die je nach Ausgangssituation gefördert oder unterbunden werden können.

Zur Vergleichbarkeit mit konventionellen Verfahren werden die Signaturalgorithmen RSA, EdDSA und ECDSA mit Sicherheitsstufen von 128-, 192- und 256-Bit unterstützt. Für PQ-Signaturverfahren wurde die Open-Quantum-Safe-Bibliothek verwendet, während für konventionelle Signaturverfahren die wolfCrypt-Bibliothek eingesetzt wurde.

Ergebnisse

Im Rahmen des allgemeinen Benchmarks konnte gezeigt werden, dass die PQ-Signaturverfahren Dilithium und Falcon für die Verifizierung der Signatur performanter sind, als konventionelle Signaturverfahren wie ECDSA, EdDSA und insbesondere RSA. Bei der Verifizierung von kurzen Nachrichten hat sich Falcon als etwa doppelt so schnell herausgestellt wie Dilithium. Für das Signieren einer Nachricht haben die PQ-Signaturverfahren zwar nicht die Geschwindigkeit konventioneller Verfahren übertroffen, aber dennoch mithalten können. Hierbei hat sich Dilithium als das zweitschnellste Verfahren erwiesen, langsamer als EdDSA, aber dennoch schneller als ECDSA.

Die Messung der Stack- und Heap-Speichernutzung in Abbildung 1 hat ergeben, dass Falcon im direkten Vergleich zu Dilithium über alle Operationen hinweg eine geringere Speichernutzung aufweist. Falcon benötigt für die Generierung von Schlüsselpaaren, zum Signieren einer Nachricht und Verifizieren der Signatur ein vielfaches weniger Speicher, als Dilithium.

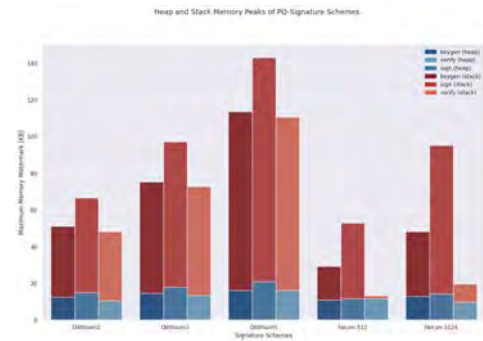


Abb. 1: Heap- und Stack-Speicherpeaks von PQ-Signaturverfahren [1]

Für die Use-Cases hat sich ergeben, dass die Laufzeit für das Signieren und Verifizieren von Nachrichten weniger vom verwendeten Signaturverfahren abhängt, sondern vielmehr von der verwendeten Hashfunktion. Insbesondere beim Test-Case Secure-Boot beträgt die Laufzeit zum Hashen der Nachricht bis zu 97 % der Gesamtlaufzeit.

Ausblick

Um die Gesamtlaufzeit zu verringern, muss das Hashen der Nachricht in Hardware beschleunigt werden. Die Messungen haben gezeigt, dass Falcon aufgrund der besseren Performance und insbesondere der geringeren Schlüssel- und Signaturlänge besser für die Use-Cases geeignet sind, als Dilithium. Das NIST hat bekannt gegeben, dass ein Auswahlverfahren für die Standardisierung von Signaturverfahren, welche nicht auf Gittern basieren, durchgeführt wird [4]. Falls Signaturverfahren mit kürzeren Schlüsseln und Signaturen zur Standardisierung ausgewählt werden, so kann sich die aktuelle Auswahl drastisch ändern.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Lov Grover. A Fast Quantum Mechanical Algorithm for Database Search. In *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, pages 212–219. Association for Computing Machinery, 1996.
- [3] Computer Security Resource Center. Announcing Request for Nominations for Public-Key Post-Quantum Cryptographic Algorithms. <https://csrc.nist.gov/News/2016/Public-Key-Post-Quantum-Cryptographic-Algorithms>, 2016.
- [4] Computer Security Resource Center. Call for Additional Digital Signature Schemes for the Post-Quantum Cryptography Standardization Process. <https://csrc.nist.gov/csrc/media/Projects/pqc-dig-sig/documents/call-for-proposals-dig-sig-sept-2022.pdf>, 2022.
- [5] Peter Shor. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. In *SIAM Journal on Computing*, pages 1484–1509. Society for Industrial & Applied Mathematics SIAM, 1997.

Disparity Estimation with Stereo Event-Based Cameras and Sparse Coding

Leon David Kiefer

Steffen Schober

Department of Computer Science and Engineering, Esslingen University

Work carried out at Mercedes-Benz AG, Böblingen

Introduction

In the field of Advanced Driver Assistance Systems (ADAS), the accurate perception of the vehicle's environment plays an important role. A component for achieving this goal is to be able to determine the distance between the vehicle and objects in the environment. Depth sensors such as LiDAR or RGB-D cameras can be used for this purpose, but their range of application is limited due to their cost, size and power consumption [1]. For this reason, geometric methods are often used, which make statements about the distance to an object in three-dimensional space using stereo images with two different viewpoints. This task can be considered as stereo matching problem, aiming to determine which part of the left image corresponds to which part of the right image. Through the identification of corresponding points within stereo images, it becomes possible to make assertions regarding the disparity exhibited by these points. Disparity is a quantity of the epipolar geometry that is inversely proportional to depth, which is why depth can ultimately be calculated on the basis of disparity. Disparity is defined as the relative displacement of the projections of a 3D point in both stereo image planes [14]. In axis-parallel systems, in which the cameras are only shifted horizontally and not rotated against each other, the disparity corresponds to a purely horizontal offset. Such a system is shown in Fig. 1. Here, M represents the 3D point projected onto points m_1 and m_2 in projection planes I_1 and I_2 , respectively. C_1 and C_2 represent the projection centers of the two cameras, which are only shifted horizontally, and f corresponds to the focal length. The connecting line B indicates the distance of the projection centers. Together with C_1 and C_2 , the 3D point M spans the epipolar plane, which intersects the projection planes in the two epipolar lines E_1 and E_2 . The parallelism of the projection planes, as well as the sole horizontal offset, result in the epipolar lines all being horizontal and on the same

image line for corresponding projections. Due to the different perspective with respect to M , this leads to a purely horizontal offset of the projections m_1 and m_2 along their epipolar lines E_1 and E_2 . The disparity δ represents this relative displacement of the projections of point M . The unit of disparity is generally given in pixels and is calculated as $\delta = u_1 - u_2$. Depth p can ultimately be determined through triangulation using the calculated disparity.

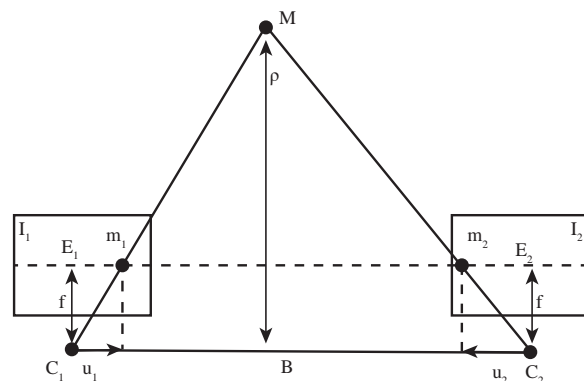


Fig. 1: Characteristics in an axis-parallel stereo camera system. Disparity corresponds to the relative displacement of the projections of 3D point M . Adapted from [14]

As Scharstein et al. [13] point out in their taxonomy, in the early phase of disparity estimation, hand-crafted cost measures were used for finding correspondences, such as squared intensity differences of pixels, often aggregated in the form of sums or averages over a local image region. For computation of disparity it is, based on the reference image, ultimately examined which pixel in the secondary image minimizes the cost. Modern frame-based stereo matching approaches no longer use hand-crafted cost functions, but estimate disparities in an end-to-end manner using Deep Neural Networks (DNNs), which proved to be superior compared to the early methods. These modern

stereo matching methods use deep learning based approaches with convolution operations for feature extraction, which either estimate a distribution of disparities [16] or predict it directly [9]. Further, there are stereo matching methods that consider the task of finding correspondences in stereo images as a sparse coding problem. Sparse coding has a close connection to neurobiology, which suggests that sensory information in the brain is encoded by a small, sparse number of active neurons, which serves as an effective representation of information [11] [10]. Sparse Convolutional Artificial Neural Networks (SCANN) attempt to find unique, sparse representations of image segments that can be used to establish a correspondence between the stereo images [8]. For this, the Locally Competitive Algorithm (LCA) can be used, which utilizes local inhibitory competition to generate sparse representations [2]. An extension of the LCA is the spiking version S-LCA [17], which mainly differs from the classical variant in that neurons communicate with spikes and thus falls into the category of Spiking Neural Networks (SNN), which mimic the functioning of biological neurons more realistically than ANNs.

In addition to approaches that address the stereo matching problem with conventional frame-based cameras, there are also approaches in literature that make use of event cameras. Event cameras use sensors that detect changes in brightness for individual pixels with very high temporal resolution and thus generate asynchronous events or spikes if a change in brightness crosses an intensity change threshold [3]. The polarity indicates whether a brightness increase or decrease has occurred. Besides the high temporal resolution, event cameras usually also have a higher dynamic range ($>120\text{dB}$) than conventional frame-based cameras, coming from the use of a logarithmic scaling in the photoreceptor and the fact that every pixel is operating independently and not waiting for a global shutter [3]. Fig. 2 depicts the inherent advantage of event cameras with respect to their higher dynamic range. The right image illustrates an intensity frame captured using a conventional frame camera. It is evident that the intensity frame suffers from overexposure at the tunnel exit, hindering accurate perception of the surroundings. Conversely, the left image shows the accumulated events of the identical scene recorded by an event camera. By leveraging its superior dynamic range, the event camera enables precise perception despite challenging lighting conditions. Notably, the red and blue pixels in the event representation correspond to individual events characterized by an increase and decrease in brightness, respectively.

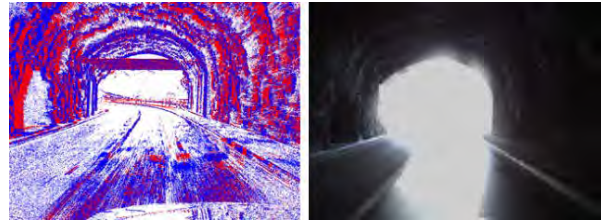


Fig. 2: Advantages of event cameras over conventional frame cameras. While the frame camera overexposes, the scene is captured accurately by the event camera. Events are marked as red and blue pixels corresponding to increases and decreases in brightness. From [4]

In order to process asynchronous event data with known stereo matching methods, which are designed for frame-based images, these must be converted into frame-based representations. Simple representations are 2D histograms [6] [3], counting or accumulating the events at a certain location, for example. Although precise temporal information is lost by quantizing timestamps, 2D histograms contain spatial information about edges in the scene, which are the most informative regions in natural images and also provide information about the presence and absence of events [3]. An alternative representation are time surfaces [5] [7], which integrate both temporal and spatial information. Time surfaces correspond to a 2D map for each polarity in which each pixel contains the timestamp of the last event at that particular position [3]. To highlight more recent events, an exponential decaying kernel is used in [5] [7]. A more complex approach, which tries to preserve the continuity of the data, is the representation as event sequence embedding [15]. The approach differs from conventional aggregation methods since the representation is not hand-crafted but learned. It considers that event sequences are sampled continuously and irregularly. Prior to the embedding, events are continuously stored in an event queue. For each new event, the oldest event is pushed out of the queue and the sequence is subsequently embedded.

There are a couple of event-based datasets in the literature that can be used for training event-based stereo matching models for disparity estimation. One dataset that can be used for this purpose is known as Stereo Event Camera Dataset for Driving Scenarios (DSEC) [4], which contains various driving sequences and disparity ground truths. It consists of driving scenarios in Zurich city and the surrounding area, divided into a total of 53 sequences with a total duration of 3193 seconds. The stereo events were recorded with two Prophesee Gen3.1 event cameras with a resolution of 640×480 pixels. The dataset additionally contains stereo RGB frames recorded at a

frequency of 20Hz. Fig. 3 displays two event frames, as well as two RGB frames with corresponding disparity ground truth.



Fig. 3: Demo images of the DSEC dataset with events represented as frame on the left and its disparity ground truth right beside. To the right, RGB frames are shown with their corresponding disparity ground truth. From [4]

Research Concept and Aim of This Work

The aim of this work is to address the technical challenge of disparity estimation by mimicking biological visual perception starting from the retina to the processing in the neural system. For this purpose stereo event cameras and sparse coding are used. Training of the different models in this work is performed with the stereo event-based dataset DSEC [4]. To achieve the goal of solving the stereo matching task, this work is divided into three individual steps.

The first step is to establish a baseline against which performance of the sparse coding model is measured and to evaluate the potential of different event representations. Here, the focus is on deep learning models, which proved to achieve state of the art performance. For this purpose the stereo matching model DispNet by Mayer et al. [9] is used, which consists of a contracting (encoder) and expanding (decoder) part with long ranged connections in between.

The encoder consists of convolutional layers, resulting in a downsampling of the input. Through a sequence of up-convolutional (also called transposed convolution) and convolutional operations, the decoder gradually and nonlinearly upsamples the feature maps, also taking into account the features of the encoder through long ranged connections. To use existing frame-based stereo matching models, event data must be converted into a representation that can be processed. Both 2D histograms [6] and time surfaces [7] are used as representations, because it has already been shown in literature that the stereo matching problem can be solved with these representations [6] [12] [5]. As part of the first step, the extent to which the predictive quality of event representations and conventional intensity frames differs is also examined.

The next step foresees the use of a more complex representation. For this purpose, we follow the work of Tulyakov et al. [15] in which an event sequence embedding is learned to predict disparity. This approach preserves the continuity and irregular sampling of the event stream while integrating information about the K most recent events and is therefore a promising approach.

Although the first and second steps are based on the perception of the receptive field of the retina, using event-based vision sensors, the processing by DispNet does not correspond to the biological neural system. Thus, in the third step the processing of the perceived events with DispNet is replaced by a sparse convolutional artificial neural network, that mimics the natural processing of visual information in mammalian brains. For this purpose, an approach is pursued that is strongly oriented towards the work of Ecker et al. [2] and Lundquist et al. [8]. Their work describes how sparse code can be generated of stereo grayscale images, which, in this work, is replaced by the use of stereo event data representations. In the work of Ecker et al. [2] LCA is used to generate the sparse code.

References and figures

- [1] Zhao Chaoqiang et al. Monocular depth estimation based on deep learning: An overview. In *Science China Technological Sciences*, volume 63, pages 1612–1627. Science China Press, 2020.
- [2] Gerrit Ecke, Harald Papp, and Hanspeter Mallot. Exploitation of image statistics with sparse coding in the case of stereo vision. In *Neural Networks*, volume 135, pages 158–176. Elsevier, 2021.
- [3] Guillermo Gallego et al. Event-based vision: A survey. In *IEEE transactions on pattern analysis and machine intelligence*, volume 44, pages 154–180. IEEE, 2020.
- [4] Mathias Gehring et al. Dsec: A stereo event camera dataset for driving scenarios'. In: In *IEEE Robotics and Automation Letters*, volume 6, pages 4947–4954. IEEE, 2021.
- [5] Sio-Hoi Ieng et al. Neuromorphic event-based generalized time-based stereovision. In *Frontiers in neuroscience*, volume 12, page 442. Frontiers Media SA, 2018.
- [6] Jürgen Kogler et al. *Neuromorphic event-based generalized time-based stereovision*. IntechOpen London, 2011.
- [7] Xavier Lagorce et al. Hots: a hierarchy of event-based time-surfaces for pattern recognition. In *IEEE transactions on pattern analysis and machine intelligence*, volume 39, pages 1346–1359. IEEE, 2016.
- [8] Sheng Lundquist et al. Sparse encoding of binocular images for depth inference. In *IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, pages 121–124. IEEE, 2016.
- [9] Nikolaus Mayer et al. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4040–4048. IEEE, 2016.
- [10] Bruno Olshausen and David Field. Sparse coding with an overcomplete basis set: A strategy employed by V1? In *Vision Research*, volume 37, pages 3311–3325. Elsevier, 1997.
- [11] Bruno Olshausen and David Field. Sparse coding of sensory inputs. In *Current opinion in neurobiology*, volume 14, pages 481–487. Elsevier, 2004.
- [12] Paul Rogister et al. Asynchronous event-based binocular stereo matching. In *IEEE Transactions on Neural Networks and Learning Systems*, volume 23, pages 347–353. IEEE, 2011.
- [13] Daniel Scharstein and Richard Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. In *International Journal of Computer Vision*, volume 47, pages 7–42. Springer Berlin, 2002.
- [14] Oliver Schreer. *Stereoanalyse und Bildsynthese*. Springer Berlin, 1 edition, 2005.
- [15] Stepan Tulyakov et al. Learning an event sequence embedding for dense event-based deep stereo. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1527–1537. IEEE, 2019.
- [16] Stepan Tulyakov, Anton Ivanov, and Francois Fleuret. Practical deep stereo (pds): Toward applications-friendly deep stereo matching. In *Advances in Neural Information Processing Systems*, volume 31. NIPS Foundation, 2018.
- [17] Yijing Watkins et al. Unsupervised dictionary learning via a spiking locally competitive algorithm. In *Proceedings of the International Conference on Neuromorphic Systems*, pages 1–5. IEEE, 2019.

Design und Implementierung einer Synchronisationselektronik für digitale Messgeräte im Pikosekunden-Bereich

Florian Korotschenko

Markus Enzweiler

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Swabian Instruments GmbH, Stuttgart

Motivation

Zahlreiche Anwendungen wie die zeitkorrelierte Einzelphotonenzählung haben Bedarf an präzisen Zeitmessungen und dem Generieren von Pulsen mit Genauigkeiten im Pikosekunden-Bereich. Besonders hohe Ansprüche entstehen hier im wichtigen Feld der Linear Optics Quantum Computation (LOQC). [5] Mehr verschränkte Photonen, respektive Qubits, also Träger quantenmechanischer Informationen, erlauben es komplexere Probleme in weniger Zeit zu lösen. Ausgelesen werde die Informationen mittels Photonen-detektoren und zur Skalierung dieser Computer ist eine erhebliche Menge an korrelierbaren Eingangskanälen notwendig. [2] Hierfür verwendete Messgeräte wie die Time Tagger der Firma Swabian Instruments GmbH bieten jedoch nur maximal 18 Kanäle pro Gerät, somit besteht der Bedarf eine Synchronisationselektronik zu entwickeln, welche mehrere Messgeräte auf kohärente Weise verbindet.

Einleitung

Da die, zu synchronisierenden, Messgeräte hochpräzise arbeiten, ergeben sich ebenso hohe Ansprüche an die Synchronisationselektronik, beziehungsweise den Synchronizer. Hauptziel einer jeden Synchronisationselektronik sollte sein, zeitliche Bezüge zwischen verschiedenen Geräten herzustellen. Hierzu verteilt der Synchronizer einen einheitlichen Takt an alle Geräte um eine einheitliche Zeitbasis zu schaffen. Gleichzeitig muss auch eine Kommunikation mit den Geräten stattfinden. Hierzu werden taktsynchron unidirektional über eine dedizierte SYNC-Leitung Datenpakete an alle Geräte versendet, welche Nummerierung der Einzelgeräte und Gruppen und die Synchronisation von Pufferüberläufen ermöglichen. Über eine Feedback-Leitung können Messgeräte auch Nachrichten an den Synchronizer zurückschicken, beispielweise eine Bereitschaftsbestätigung oder Fehlercodes. Überdies

besitzt der Synchronizer Trigger Eingänge, deren Zustand an angeschlossene Pulse Streamer übertragen werden. Darüber hinaus soll die Elektronik flexibel einsetzbar sein und auch mit zukünftigen Produkten kompatibel sein.

In nachfolgender Abbildung 1 ist ein Beispiel für den Anschluss von 3 Time Tagger Messgeräten der Firma Swabian Instruments an einen Synchronizer zu sehen. Hier werden über gleichlange SMA-Koaxialkabel Takt (CLK) und Daten (SYNC) an die Geräte verteilt und über eine Ringtopologie wird ein Feedback (FDBK) übertragen, worüber die physische Reihenfolge bestimmt wird. [6]

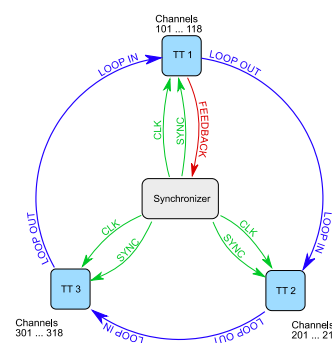


Abb. 1: Beispielhafter Aufbau für Synchronizer mit 3 Time Taggern [6]

Zielsetzung

In dieser Arbeit ist die Erstellung eines vollständigen Schaltplanes gefordert. Dieser beinhaltet den vollen (HF)-Signalpfad sowie eine rauscharme Stromversorgung über multiple Schalte- und Linearwandler. Im Hinblick auf die Performance des HF-Pfades wird eines der kritischsten Komponenten, der Fanout-Buffer

für Takt- und Synchronisationssignale, vermessen und verifiziert. Ein Teil des Layout wird erarbeitet, mit sinnvollem Stackup und Aufarbeitung des notwendigen Wissens aus dem Bereich der High-Speed Leitungstheorie. Als Programm für die Erstellung von Schaltplan und Layout wird Altium Designer verwendet.

Umsetzung

Um Konvertierungen, und daher rührende zeitliche Ungenauigkeiten des Taktsignals, zu vermeiden, wird direkt ein 500 MHz Takt an die Messgeräte verteilt, welcher von Time Taggern verwendet wird. Erzeugt wird dieser Takt wahlweise über einen Quarzofen (OXCO) und eine PLL, ersterer garantiert eine hohe Stabilität des Taktes, oder mittels der Multi-Gigabit Transceiver des verwendeten FPGA. Dieser stellt gleichzeitig auch das zentrale Element dieser Elektronik

dar. Ein FPGA erlaubt hohe Flexibilität und besitzt schnelle I/O-Ports, welche auch die Synchronisationspakete versenden, die Trigger Eingänge abtasten oder diverse Multiplexer, Buffer und Komparatoren ansteuern. Über all jene Bausteine werden die Signale von und zum FPGA geführt. Gleichzeitig sorgt ein Mikrocontroller (BMC) dafür, dass Initialisierungs- und Power-down-Sequenzen kontrolliert ablaufen, Versorgungsspannungsniveau und Temperaturen überwacht und Lüfter und Knöpfe korrekt angesprochen werden. Einen Überblick verschafft Abbildung 2 als logisches Blockschaltbild. Es werden in dieser vereinfachten Darstellung nur Komponentengruppen aufzeigt. Neben diesen, werden auch die verwendeten differenziellen elektrischen Pegel der jeweiligen Ausgangstreiber farblich unterteilt. Dies soll eine der größeren Herausforderungen dieser Arbeit unterzeichnen.

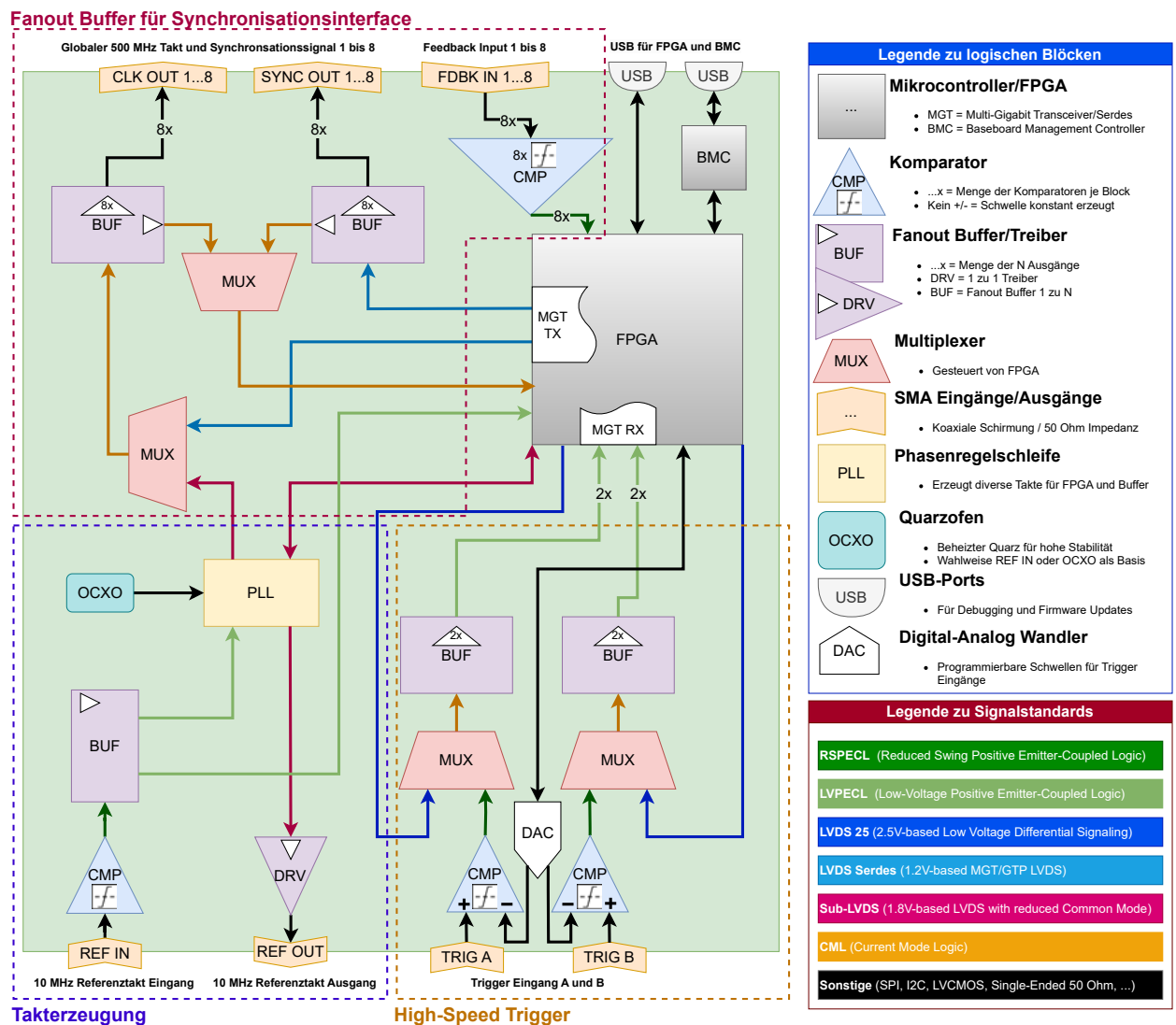


Abb. 2: Logisches Blockschaltbild der Synchronisationselektronik [3]

Herausforderungen

Eine solch große Anzahl an verwendeten elektrischen Signalstandards, bedingt durch die Auswahl der verwendeten aktiven Komponenten, bedingt Pegelanpassungen. Stets müssen die Hoch- und Tiefpunkte (Peaks), respektive der Spannungshub (Swing) und der neutrale Bezugspunkt (Common Mode), um den oszilliert wird, für den Empfänger kompatibel gemacht werden. Hierzu sind Spannungsanpassungen mittels Spannungsteilern oder eine kapazitive Kopplung in Kombination mit Rebiasing-Netzwerken zur Erzeugung einer neuen Common Mode notwendig. [4] Die Entscheidung und exakte Implementation hängt hier wiederum vom internen Aufbau der Sende- und Empfangselektronik der Bausteine und den Limits der jeweiligen Datenblätter ab. Einen Überblick von üblichen, Single-Ended betrachteten, Pegeln gibt Abbildung 3.

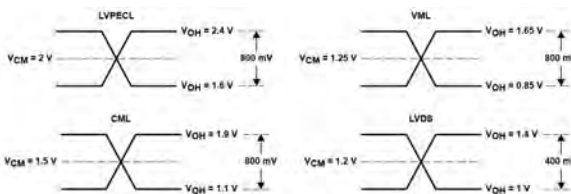


Abb. 3: Verschiedene differentielle Pegel im Überblick [4]

Neben diesen Maßnahmen zur Kompatibilität sind auch allgemeine Faktoren der Signalstandards zu beachten. PECL-basierte Treiber benötigen in der Regel externe Widerstände zur Einstellung des Stroms durch die Ausgangstransistoren. Hier muss zwischen Stromverbrauch und Geschwindigkeit der Flanken entschieden werden. [9] Auch muss der gesamte Strom, durch Sender und Empfänger, unterhalb des zulässigen Maximums liegen, sofern der Sender keine geregelte Stromquelle verwendet. Bei CML-basierten Treibern hingegen ist ein externer Pull-up-Widerstand notwendig, um die Übertragungsleitung korrekt zu betreiben. [4] Eine Beispielübersetzung von LVPECL auf LVDS, welche auf Kompatibilität angepasst wurde, ist in Abbildung 4 dargestellt. Dabei wird der 100 Ohm Terminierungswiderstand, über den der Wechselspannungshub abfällt, als intern angenommen. Der Arbeitspunkt des LVPECL-Treibers wird mittels Widerstandswert R_E eingestellt, während die seriellen Widerstände R_S zusammen mit der Terminierung im Empfänger einen Spannungsteiler bilden und den Swing des Signals dämpfen. R_{E3} und R_{E4} bilden überdies einen weiteren Spannungsteiler, welcher den Gleichspannungsanteil (Common Mode) auf der

Empfängerseite der kapazitiven Kopplung auf einen LVDS-kompatiblen Wert legt. [8]

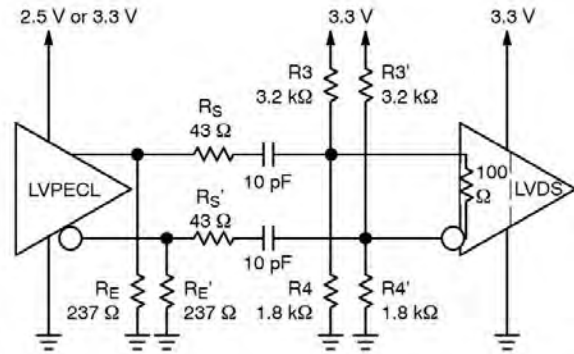


Abb. 4: Beispielschaltung kapazitiver Kopplung mit Biasing und interner Terminierung [8]

Der vorherige Abschnitt ist ein Beispiel dessen, was beim Erstellen des Schaltplan beachtet werden muss. Weitere Faktoren der Transmission Line Theory müssen, insbesondere bei der Übersetzung in ein Layout, beachtet werden. Neben der korrekten Terminierung, muss auch die Leitungsimpedanz im physischen Layout übereinstimmen, um Reflektionen und Leistungsverluste zu vermeiden. Die Bandbreite muss zur Minimierung von Jitter und Maximierung von Flankensteilheit über 5 GHz betragen, was Implikationen auf die Selektion der Komponenten und deren Platzierung hat. [7] Auch die Leitungen, beziehungsweise Microstrips auf der Platine, müssen spezifische Dimensionen aufweisen, welche in Kombination mit dem Stackup die Zielimpedanz ergeben. Ebenfalls abhängig vom Stackup ist die Propagationsgeschwindigkeit der Signale über genannte Leitungen. Um zeitlichen Versatz der Ausgangskanäle des Synchronizers zu minimieren, muss an entsprechender Stelle ein Längenausgleich auf den Übertragungsleitungen implementiert werden. [10] Ebenso müssen Wechsel von Ebenen durchdacht und Anschlüsse an Ports optimiert werden, um Abweichungen der Leitungsimpedanz entlang der Signalstrecke zu minimieren. [1]

Ausblick

Nachdem das Schematic fertiggestellt ist, muss auch das Layout komplett erarbeitet werden. Dies ist mindestens genauso relevant für die Performance des physischen Designs. Die gesamte Platine soll nahezu die Maße eines 19 Zoll Racks einnehmen. Nach dieser Arbeit sind weitere Schritte wie die Erstellung der Firmware für FPGA und BMC, sowie Verifikation des physischen Aufbaus notwendig, bevor ein finales Produkt entstehen kann.

Literatur und Abbildungen

- [1] Tsun-Kit Chin and Texas Instruments. Lab Report SNLA186: Getting Signal Launch Right for High Speed Sub-systems. https://www.ti.com/lit/an/snla186/snla186.pdf?ts=1685908974677&ref_url=https%253A%252F%252Fwww.google.com%252F, 2009.
- [2] Charles Q. Choi. How Many Qubits Are Needed for Quantum Supremacy? *IEEE Spectrum*, pages 1–2, 2020.
- [3] Eigene Darstellung.
- [4] Nick Holland and Texas Instruments. Application Report SLLA120: Interfacing Between LVPECL, VML, CML, and LVDS Levels. https://www.ti.com/lit/an/slla120/slla120.pdf?ts=1668583179460&ref_url=https%253A%252F%252Fwww.google.ru%252F, 2002.
- [5] Swabian Instruments. Linear Optics Quantum Information. <https://www.swabianinstruments.com/applications/linear-optics-quantum-information/>, 2023.
- [6] Swabian Instruments. Synchronizer — Time Tagger User Manual 2.15.0.0 documentation. <https://www.swabianinstruments.com/static/documentation/TimeTagger/synchronizer/index.html#>, 05 2023.
- [7] Maxim Integrated. APPLICATION NOTE 1916: An Introduction to Jitter in Communications Systems. <https://www.analog.com/media/en/technical-documentation/tech-articles/an-introduction-to-jitter-in-communications-systems.pdf>, 2003.
- [8] Paul Lee and ON Semiconductor. AN1568/D: Interfacing Between LVDS and ECL. <https://www.onsemi.com/pub/Collateral/AN1568-D.PDF>, 2013.
- [9] Paul Shockman and ON Semiconductor. AND8020/D: Termination of ECL Devices with EF (Emitter Follower) OUTPUT Structure. <https://www.onsemi.com/pub/Collateral/AND8020-D.PDF>, 2007.
- [10] Cadence System-Analysis. Transmission Line Propagation Delay: Characteristic Impedance and Dielectric Material. <https://resources.system-analysis.cadence.com/blog/msa2021-transmission-line-propagation-delay-characteristic-impedance-and-dielectric-material>, 2023.

Erweiterung eines HF Performance Messplatzes mit automatischer Erzeugung normkonformer Reports

Markus Lang

Clemens Klöck

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Hirschmann Car Communication, Neckartenzlingen

Abstract

Im vorliegenden Dokument wird die Erweiterung eines HF Performance Messplatzes mit automatischer Erzeugung normkonformer Reports vorgestellt. Die angestrebte Norm ist die Radio Equipment Directive (RED). Im RED Standard wird beschrieben, welche Anforderungen für die Bereitstellung von funkfähigen Geräten auf dem europäischen Markt eingehalten werden müssen. Dieser Standard wurde vom European Telecommunications Standards Institute (ETSI) entworfen und fundiert auf europäischen Richtlinien. [6] Im Folgenden wird beschrieben, mit welchen Methoden ein vorhandener HF Performance Messplatz erweitert wird, um diesen Standard einzuhalten.

Einleitung

In der heutigen Zeit spielt die reibungslose drahtlose Kommunikation vieler Geräte untereinander eine wichtige Rolle. Eine der ältesten Kommunikationsmethoden ist dabei der Mittelwellenfunk (AM). Dieser wurde bereits 1924 als Massenmedium verwendet. In den 1950er Jahren ist das Ultrakurzwellen-Radio (UKW-Radio), besser bekannt als FM-Radio eingeführt worden. Dieses bot im Gegensatz zum AM Standard einen Stereo-Ton. Beide Standards sind bis heute weltweit vertreten. [1] 1999 wurden die ersten Sender über den Digital Audio Broadcasting (DAB) Service verfügbar. Diese Technologie ermöglicht das Übertragen von Radiosendern in einem digitalen Format. [8] Trotz des digitalen Fortschritts in unserer Welt wird nach wie vor AM und FM beim Rundfunkempfang verwendet. Besonders in Krisenzeiten sind diese Übertragungsmöglichkeiten von großer Bedeutung. Die Empfänger für AM und FM können mithilfe weniger elektronischer Bauteile problemlos selbst hergestellt werden. Dies stellt sicher, dass sich jedermann jederzeit über dieses Medium informieren und kommunizieren kann. Diese Übertragungsmethoden fallen in den Hochfrequenzbereich (HF). Das European Telecommunications Standards Institute (ETSI) gibt dabei ein Regelwerk

vor, wie sich beispielsweise Empfangsgeräte des AM / FM / DAB Standards verhalten müssen. Im europäischen Raum ist die RED seit 2015 in Kraft. Das Ziel der RED ist es, einheitliche Regeln für den freien Verkehr von Funkgeräten innerhalb des Europäischen Wirtschaftsraums (EWR) zu schaffen und die Sicherheit sowie den Schutz der Gesundheit von Benutzern und Verbrauchern zu gewährleisten. [6] Um der RED-Norm gerecht zu werden, müssen die definierten Ziele aus der ETSI EN 303 345-1 V1.1.1 (2019-06) eingehalten werden. Die Normen der ETSI sind in spezifische Anforderungen für den jeweiligen Broadcast Service unterteilt. Dabei relevant für AM ist die ETSI EN 303 345-2 V1.2.1 (2021-12), für FM die ETSI EN 303 345-3 V1.1.1 (2021-03) und für DAB die ETSI EN 303 345-4 V1.1.1 (2021-06). [4] Nach diesen drei Anforderungsdokumenten werden Messungen implementiert, um sicherzustellen, dass ein Device Under Test (DUT), welches einem empfangsfähigen Gerät entspricht, den Anforderungen der oben genannten Normen einhält. Ziel dieser Arbeit ist es, die in der RED geforderten Messungen entsprechend derer Vorgaben zu implementieren. Die Messergebnisse werden automatisiert mit den Grenzen der Norm verglichen und visuell dargestellt.

Methoden

Der Messablauf wird in der Keysight TestExec [7] implementiert. Mithilfe dieses Tools können automatisierte Testablaufpläne erstellt und ausgeführt werden. Dabei greift diese auf eine funktionale Anwendung im Hintergrund zu, welche in Microsoft .NET geschrieben wurde. Diese Bibliothek bietet unter anderem eine Signalauswertung an, welche den Signal to Noise (SNR) als Verhältnis in dB des Nutzsignals zum Rauschen ausgibt. Bei der SNR-Messung wird ein modulierte HF-Nutzsignal mit einem Sinus-Ton eingespeist. Dabei wird der Leistungsunterschied des Sinus-Tons zum Rauschen gemessen. Gemäß Spezifikation werden zwei verschiedene Messungen gefordert, welche sich in drei

Messblöcke aufteilen. Dabei ist der Ablauf für alle Standards ähnlich. Es wird zunächst die Sensitivity gemessen. Diese prüft die Eingangsempfindlichkeit mittels SNR. Weiterführend wird die Adjacent Channel Selectivity Messung vorgenommen. Diese prüft den SNR bei Nachbarkanalstörungen in definierten Abständen zur Trägerfrequenz. Des Weiteren folgt die Blocking Messung. Diese prüft auf ähnliche Weise bei größerem Frequenzabstand zur Trägerfrequenz den SNR. Zu allen Messverfahren werden Streams als IQ-File von der ETSI zur Verfügung gestellt. Diese werden mithilfe eines Signalgenerators in das System eingespeist. In den Spezifikationen EN 303 345 gemäß den Standards selbst ist exakt beschrieben, mit welchem Pegel bei welcher Frequenz die Signale eingespeist werden müssen. Über alle Messungen hinweg der Signal to Noise (SNR) Pegel ausgewertet. Dieser muss stets größer gleich 40 dBQ sein. Die Einheit dBQ wird in der ITU-R BS.468-4 angedeutet. Die Interpretation dieses Wertes ist mit internen HF-Experten abgestimmt worden. Aus dem geforderten Messverfahren mit Vorschaltung eines speziellen Filters (ITU-R BS.468-4) wird das Signal wie in Abbildung 1 beschrieben gefiltert.

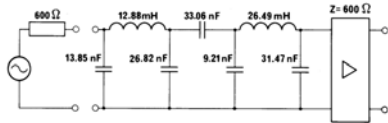


Abb. 1: ITU Filter Netzwerk nach BS.468-4 [3]

Um diesen Filter zu implementieren, wurde dieser simuliert. Anhand dieser Simulation ist eine Werte-

tabelle von 1 Hz bis 48000 Hz mit der passenden Verstärkung (Gain) in dB entstanden. Um den Filter in der Signalauswertung zu verwenden, wird aus der Verstärkung in dB ein Leistungsfaktor bestimmt. Dieser errechnet sich aus $P_{Faktor} = 10^{\frac{Gain[dB]}{20}}$. Um die Ergebnisse zur Laufzeit des Programmes zur Verfügung zu stellen, werden diese in einem Caching-Verfahren zwischen 1 Hz und 48000 Hz vorgerechnet. Das angewendete Verfahren ermöglicht es, die Ergebnisse simultan zur Laufzeit ohne Verzögerung während der Fast Fourier Transformation (FFT) zu nutzen. Die Verrechnung des Filters erfolgt als Leistungsfaktor direkt mit dem Ergebnis der FFT. Durch das Verwenden des BS-468-4 Filters ergibt sich ein Qualitätsmerkmal des Pegels. Somit ist die Einheit des SNR in dBQ. In Abbildung 2 ist der Messaufbau vom Signalgenerator bis zur Messauswertung ersichtlich. Dieser Verlauf wird von links nach rechts gelesen. Ein Signalgenerator speist ein Signal in das DUT. Dieses gibt ein Audiosignal aus. Im Anschluss wird dieses durch eine FFT in seine Frequenzkomponente zerlegt. Diesem Frequenzspektrum gibt der BS-468-4 Filter passend zu jeder ganzzahlig gerundeten Frequenz den passenden Verstärkungsfaktor bei. Anschließend wird das verrechnete Frequenzspektrum auf den aus der EN 303 345 passenden Bandbereich begrenzt, für FM beispielsweise gemäß EN 303 345-3 15 kHz [5]. Daraufhin folgen zwei weitere Filter, welche das Nutzsignal vom Rauschen trennen. Das Nutzsignal wird durch einen Schmalbandfilter welcher auf die Nutzfrequenz eingestellt wird gefiltert. Das Rauschen erhält man, indem der gesamte Rest des Spektrums in ein Verhältnis gesetzt wird $SNR = \frac{Signalleistung[dB]}{Rauschleistung[dB]}$

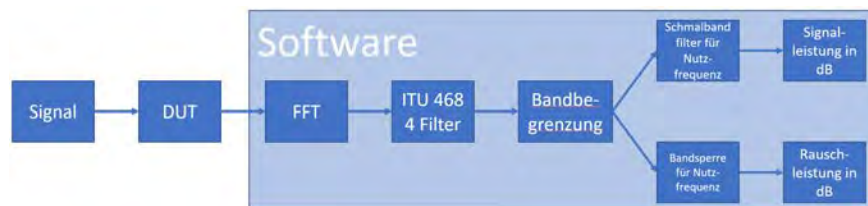


Abb. 2: Signalverlauf des Messaufbaus [2]

Ergebnis

Der Filter wurde nach ITU BS.468-4 implementiert und getestet. Nach einer Laufzeitanalyse mit anschließender

Optimierung des Filters wurden passende Caching-Verfahren implementiert. In der TestExec sind die Testfälle für den AM und FM Standard ebenfalls implementiert und deren Ergebnisse werden verifiziert.

Literatur und Abbildungen

- [1] Gorse Christiane and Schneider Daniel. Geschichte des Radios. https://www.planet-wissen.de/kultur/medien/geschichte_des_radios/index.html, 2019.
- [2] Eigene Darstellung.
- [3] . ETSI. ITU-R BS.468-4. https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.468-4-198607-III-PDF-E.pdf, 1986.
- [4] . ETSI. ETSI EN 303 345-1 V1.1.1. https://www.etsi.org/deliver/etsi_en/303300_303399/30334501/01.01.01_60/en_30334501v010101p.pdf, 2019.
- [5] . ETSI. ITU 303 345-3. https://www.etsi.org/deliver/etsi_en/303300_303399/30334503/01.01.01_60/en_30334503v0102021.
- [6] . ETSI. Radio. <https://www.etsi.org/technologies/radio>, 2022.
- [7] . KeySight. KeySight. <https://www.keysight.com/de/de/product/E2011GC/testexec-sl-software.html>, 2000.
- [8] Mäusli Theo and Steigmeier Andreas. Geschichte. <https://www.srgssr.ch/de/news-medien/dossiers/abloesung-ukw/geschichte>, 2023.

Eine vergleichende Evaluierung von Flutter und React als Alternative zu Qt für die Programmierung der Bedienoberfläche von professionellen Kaffeemaschinen

Steffen Lang

Jörg Nitzsche

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma WMF GmbH, Geislingen an der Steige

Einleitung

Die WMF GmbH setzt seit einiger Zeit Touchscreen-Displays (siehe Abb. 1) für die Bedienung ihrer professionellen Kaffeemaschinen ein. Bisher wurde die GUI-Entwicklung für die Bedienoberfläche mithilfe des Qt/QML-Frameworks durchgeführt.



Abb. 1: Bedienoberfläche eines WMF 5000S+ Kaffeevollautomaten [2]

Allerdings hat die WMF GmbH mit steigenden Lizenzierungskosten für das Qt Framework zu kämpfen. Die Long-term-supported (LTS) Releases des Qt Frameworks sind seit 2020 nur noch im Rahmen der kommerziellen Lizenzierung verfügbar [1]. Vor diesem Hintergrund besteht die Motivation, alternative Frameworks für die GUI-Entwicklung zu evaluieren. Es gibt kostenlose Alternativen wie z.B. das Cross-Plattform Framework Flutter und die JavaScript Bibliothek React mit denen benutzerfreundliche Oberflächen

entwickelt werden können. Darüber hinaus gibt es Frameworks wie Electron und eingebettete Webbrowser wie Chromium die den Einsatz von Web-Technologien ermöglichen. Die gestiegene Rechenleistung der verwendeten Hardware, insbesondere durch den Einsatz eines leistungsfähigen Prozessors und einer dedizierten Grafikeinheit, eröffnet die Möglichkeit, nicht nativ kompilierte Software auf dem Board einzusetzen, ähnlich wie es bei Web-Technologien der Fall wäre [4]. Vor diesem Hintergrund soll in der vorliegenden Arbeit eine vergleichende Evaluierung von Flutter und React als Alternative zu Qt für die Programmierung der Bedienoberfläche von professionellen Kaffeemaschinen durchgeführt werden.

Technische Anforderungen

Die technischen Anforderungen an die Bedienoberfläche beziehen sich auf verschiedene Aspekte. Erstens müssen die Anforderungen bezüglich der Hardware erfüllt werden, einschließlich der Unterstützung verschiedener Prozessorarchitekturen und Betriebssysteme. Dies umfasst die Kompatibilität mit einem ARM-basierten 4-Kern-Prozessor mit Hardwarebeschleunigung, sowie die Unterstützung von Windows-Rechnern mit x86-Architektur für die Simulation.

Die Performance der Software ist ebenfalls ein wichtiger Aspekt. Hierbei geht es darum bestimmte Kriterien zu erfüllen. Neben subjektiven Kriterien, wie der flüssigen Darstellung von Animationen sind auch messbare Kriterien zu erfüllen, wie die Nutzung von Systemressourcen, wie CPU und Speicher und die Erfüllung spezifischer Test-Cases. Diese sind z.B. die Reaktionszeit auf Bestellungen, die Aktualisierungsrate der Getränkeverfügbarkeit und die maximale Aktualisierungsrate der Echtzeit-Brühparameter. Ziel ist es, eine gute Benutzererfahrung sicherzustellen und bestimmte Performance-Ziele, wie 60 Frames

pro Sekunde und die Vermeidung von Rucklern, zu erreichen.

Die Internationalisierung und das Testing sind weitere wichtige Anforderungen. Die Software muss in der Lage sein, verschiedene Sprachen zu unterstützen, einschließlich einer virtuellen Tastatur mit mehreren Sprachen.

Weitere Punkte, die zu prüfen sind, umfassen die Verfügbarkeit aktueller GUI-Komponenten in der entsprechenden Sprache/in dem entsprechenden Framework, die Unterstützung von Video- und Fotoformaten, die technische Integration in aktuelle Apps sowie die Erfüllung der von der Entwicklung definierten vier wichtigsten Qualitätsmerkmale der Softwarearchitektur. Diese sind Modularität, Modifizierbarkeit, Testbarkeit und Verfügbarkeit.

Auswahl der Frameworks

Die Wahl von Flutter und React als zu evaluierende Frameworks bzw. Bibliothek basiert auf mehreren Kriterien. Einerseits ist es wichtig, dass das Framework, eine geeignete Entwicklungsumgebung und ggf. das Testframework kostenlos zur Verfügung stehen, um Kosten zu minimieren. Darüber hinaus ist die Beliebtheit ein entscheidender Faktor, da populäre Frameworks in der Regel eine aktive und wachsende Community und eine breite Unterstützung haben. Ein weiterer Aspekt ist die Verbreitung des Frameworks, da es einfacher sein kann, mit dem Framework vertraute Entwickler zu finden. Eine weite Verbreitung erleichtert den Zugang zu Ressourcen wie Tutorials, Beispielcode und Hilfestellung von anderen Entwicklern. Es ist auch wahrscheinlicher, dass häufig benötigte GUI-Komponenten vorhanden sind. Die Reife des Frameworks und die Open-Source Lizenzierung sind ebenfalls wichtig. Eine Überprüfung der Stellenangeboten auf Plattformen wie Xing und LinkedIn in Abb. 2 zeigt, dass es für beide Frameworks eine hohe Anzahl von Stellenangeboten gibt, woraus eine hohe Verbreitung geschlossen wird. Auch GitHub Stars, die Relevanz auf Google Trends und die Verfügbarkeit von optimalerweise deutschsprachigen Udemy Online-Kursen wurden berücksichtigt.

Stellenanzeigen Frameworks/Bibliotheken

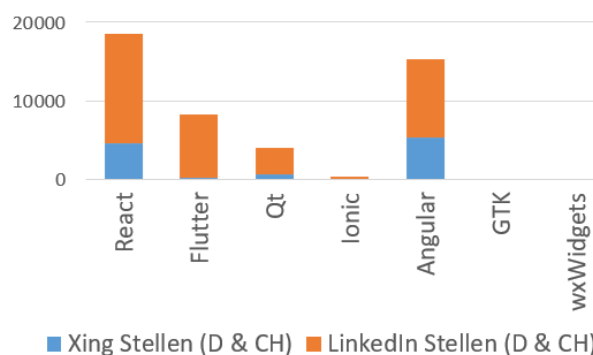


Abb. 2: Stellenanzeigen für Frameworks & Bibliotheken (D & CH) [5]

Diese positiven Indikatoren unterstreichen die Entscheidung, das Cross-Plattform Framework Flutter und die JavaScript-Bibliothek React genauer zu evaluieren. Das Web-Framework Angular wird als Konkurrenz zu React betrachtet und wird wegen des zusätzlichen Aufwands nicht berücksichtigt.

Methodik

Flutter und React werden anhand der o.g. technischen Anforderungen gemessen. Um allgemeine und spezifische Performance-Anforderungen zu verifizieren, wird ein Proof of Concept (PoC) der Bedienoberfläche, wie in Abb. 3 dargestellt, mit Flutter und React implementiert. Der PoC soll verschiedene Features nutzen, die auch in der aktuellen PCM-GUI verwendet werden, wie z.B. das Darstellen von Videos und die Internationalisierung der App. Einige unwichtigere Komponenten werden nicht implementiert, um den Aufwand zu begrenzen. Die React Anwendung kann entweder auf einem Webserver ausgeliefert werden, um sie in einem Webbrowser anzuzeigen oder als Electron Anwendung. Electron ermöglicht es, die Anwendung in eine eigenständige Desktop-Anwendung umzuwandeln, die plattformübergreifend auf verschiedenen Betriebssystemen funktioniert. Bekannte Beispiele für Electron Anwendungen sind Visual Studio Code oder Slack [3]. Zusätzlich zur App selbst wird eine weitere App als Backend implementiert, um den Bestellprozess, die Getränkeverfügbarkeit und das Liefern von Prozesswerten zu simulieren. Die Erfahrungen während des Entwicklungsprozesses für den PoC, einschließlich Testing, Debugging, Deployment und Tooling, sind von Relevanz.

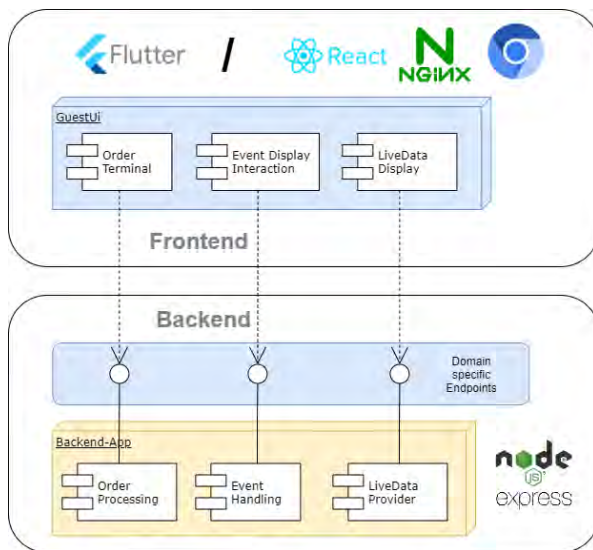


Abb. 3: Proof of concept - Fachliche Komponenten & verwendete Technologien [2]

Um die Performance im produktiven Einsatz zu testen, wird die App und das Backend auf einem Raspberry Pi 3 B+ getestet. Dieser Einplatinencomputer eignet sich gut für das Rapid Prototyping und verfügt über die erforderlichen Ressourcen. Der Raspberry Pi 3 B+ besitzt vergleichbaren Hardware-Ressourcen wie das von WMF zuletzt verwendete Board für die professionellen Kaffeemaschinen.

Ergebnisse & Ausblick

Einige bereits vorliegende Ergebnisse deuten darauf hin, dass Flutter die bessere Wahl ist. Sofern ein plattformspezifischer Flutter-Embedder, der eine Hardwarebeschleunigung ermöglicht, vorhanden ist, bietet Flutter eine sehr gute Performance wie Abb. 4 zeigt.

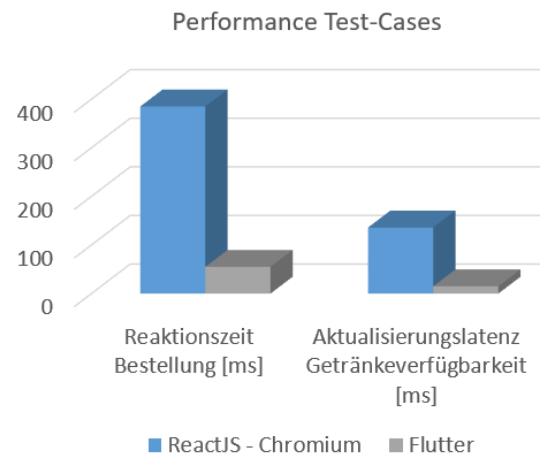


Abb. 4: Spezifische Performance Test-Cases der PoC Apps auf Raspberry Pi 3 B+ [2]

Beim UI Testing sind beide Frameworks im Vergleich zu Squish (aktuelles Test-Framework für Qt) etwas im Nachteil, da sie kein programmierfreies Record & Replay Testing anbieten. Sowohl React, als auch Flutter lassen sich hervorragend internationalisieren, sind beliebt, verbreitet und vor allem kostenlos. SPAs, die mit React programmiert sind, erfordern durch den Browser sehr viel Speicher. Der RAM eines Raspberry Pi 3 B+ (1 GB), sowie dessen Swap-Speicher sind bei einer mittelgroßen SPA schon voll ausgelastet. Auch die CPU ist durch die Anzeige von Videos, und dem animierten Wechseln einer Ansicht sehr stark ausgelastet. Es wird im Folgenden noch Untersuchungen mit einem Raspberry Pi 4 mit vier GB Speicher und dem Deployment als Electron App geprüft. Auch Optimierungen im Code sind noch denkbar. Sollten diese keine Besserung der größtenteils ruckeligen Darstellung trotz hardwarebeschleunigtem Chromium Browser ergeben, wird eine klare Empfehlung für Flutter ausgesprochen. Die Nachteile von Flutter beschränken sich auf dem jungen Alter des Frameworks und der damit einhergehenden Unerfahrenheit der Lieferanten mit dem Framework und der relativ wenig verbreiteten Programmiersprache Dart.

Literatur und Abbildungen

- [1] Qt Company. Qt offering changes 2020. <https://www.qt.io/blog/qt-offering-changes-2020>, 2020.
- [2] Eigene Darstellung.
- [3] Website Electron. Electron JS. <https://www.electronjs.org/de/>, 2023.
- [4] Device Hive. User Interface for Embedded Devices with Modern Web Technologies. <https://medium.com/@DeviceHive/user-interface-for-embedded-devices-with-modern-web-technologies-921dfed25e89>, 2021.
- [5] Website LinkedIn. LinkedIn. <https://linkedin.com>, 2023.

Finding States in Black-Box Fuzzing

Andreas Lautner

Dominik Schoop

Department of Computer Science and Engineering, Esslingen University

Work carried out at Robert Bosch GmbH, Renningen

Motivation

Automated security testing is indispensable for modern software development. The sheer scope of digitalization and the complexity of code bases increase potential attack vectors and make manual security testing or auditing impossible. Fuzzing, a dynamic software testing technique coined by Miller *et al.* [2], proved tremendously beneficial for finding bugs. A fuzz tester (or fuzzer) is a tool that iteratively generates unexpected, malformed, or random data and uses this data as input for software under test [1]. The executed software under test is then monitored for e.g., crashes or hangs. The most popular form of fuzzing is grey box fuzzing, i.e., when the target is instrumented providing lightweight coverage feedback to the fuzzer. This sheds some insights into the internal behaviour or structure of the software.

While grey box fuzzing works well when software can be recompiled or run in a special environment, the application of fuzzing for the embedded domain – such as a car – remains challenging. Embedded software can often only be tested on the target hardware itself. Additionally, embedded software has all sorts of real-time constraints, hardware dependencies, and specialized build chains, so a port to a fuzzable target alters the results to not be representative with the original software. Because of all these constraints where instrumentation is hard, this thesis covers black box fuzzing, where no information of the internal behaviour is known. In a black box setting, where the behaviour of the software is not known, the target can be seen as an oracle, where the input can be generated and output messages, as well as its timing, can be observed.

Automata Learning

The internal behaviour of a black box can be modelled by its state machine. The state machine contains the states a black box can have and the transitions between states. Since no instrumentation through coverage is possible in black box fuzzing, some approaches use the state machine for instrumenting the fuzzer. A state

machine can be estimated through observations that are processed by so-called learning algorithms. The easiest form of observation is inputs and outputs sent to and by the black box.

By capturing and analyzing the observations a state machine can be built. A Mealy Machine is a common representation of a state machine and is used in this thesis. It has been shown that Mealy Machines are suitable to model reactive systems [4]. A Mealy Machine is deterministic and only has one transition for each state and input symbol. It can be described by the 6-tuple $M = (Q, \Sigma, \Gamma, \delta, \lambda, q_0)$, where

- Q is a finite set of states,
- Σ is a finite set of input symbols (the alphabet),
- Γ is a finite set of output symbols,
- $\delta: Q \times \Sigma \rightarrow Q$ is the transition function,
- $\lambda: Q \times \Sigma \rightarrow \Gamma$ is the output function,
- $q_0 \in Q$ is the initial state.

Most automata learning algorithms follow the Minimally Adequate Teacher (MAT) framework. The MAT framework consists of two parties: The teacher knows the language of the system under learning (SUL), in learning algorithms he is mostly represented by the SUL itself. The learner tries to learn the behavioural model of the SUL by processing queries to the teacher. Membership queries are used to check if a word is part of the language of the SUL. Here, the learner gains knowledge of the SUL and with this knowledge, he creates a hypothesis of the state machine of the SUL. In equivalence queries, the learner gives the teacher the hypothesis and the teacher checks if the query output of the hypothesis is equivalent to the output of the SUL. When the hypothesis output is not equivalent to the SUL's, a counter-example is provided by the teacher. With this counter-example, the hypothesis gets refined and the steps repeat. Iteratively, a state machine is generated until no counter-example can be found.

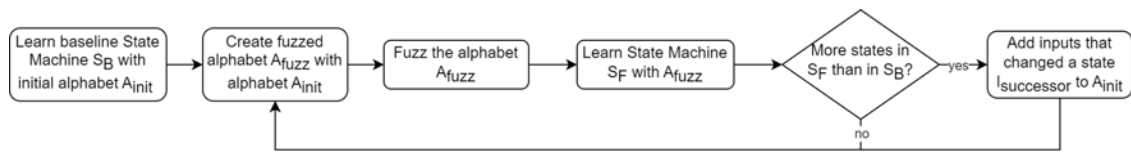


Fig. 1: Activity diagram of the automata learning with integrated fuzzing [3]

Solution Design

The goal of this work is to create a one-fits-all state machine estimation tool. Prerequisites to tackle this goal are that the test object is some kind of executable with some specifications e.g., by the vendor and no available source code. For the state machine estimation, inputs and outputs will be passed to the learning algorithm. The algorithm needs an initial alphabet of accepted inputs for the SUL. The initial alphabet will be extracted from the specification. For the tool, it is also required to know how to communicate with the test object to send inputs to it and receive its outputs. The communication of the learning algorithm with the SUL happens over an interface which can e.g., be TCP.

This thesis also tackles the use of fuzzing in automata learning. In combination with automata learning, there are two ways to use fuzzing. In the first approach, the state machine of the test object is initially learned. The fuzzer uses the information of the state machine to get itself into a specific state and only then start to generate random inputs. A simple server can be used to illustrate this concept, where one first needs to be logged in to unlock a majority of functions. This should make the fuzzer more efficient and function-oriented. The second proposed approach is the use of fuzzing in the learning process. The process is shown in Figure 1. With this approach, new valid inputs can be found which results in a deeper exploration of the test object and a more complete state machine.

Implementation

The whole implementation of the approach taken in this thesis is based on the open-source Java learning

framework LearnLib. LearnLib was released in 2013 and is developed and maintained by the TU Dortmund. LearnLib provides various algorithms for automata learning for different target models as well as a variety of equivalence approximation strategies. The features also include learning statistic reports, interfaces for SULs and easy extensibility. LearnLib has been used widely in research.

The implementation of this thesis consists of three major parts. The first part is the SUL interface which currently holds a SocketSUL for test objects that communicate over TCP and UDP. The second part is the learning factory which is responsible for automata learning. In the learning factory, different combinations of learning algorithms and equivalence strategies can be run. It is also used for benchmarking the learning process which is relevant for the evaluation. The third part is where fuzzing is used.

Validation and Outlook

The first tests of the automata learning were done with an FTP server where a client gets learned with a fraction of the functions. All implemented parts for the learning factory worked well and the resulting state machine matched with the protocol specification. Some runs were conducted where fuzzing was incorporated into the learning of an FTP client. Since this process heavily relies on randomness, a large set of iterations need to be conducted and until now, no new states or valid inputs were found. More tests need to be done to get a chance for interesting findings. The aim is to also incorporate more interfaces like CAN into the framework that allow learning various test-objects.

References and figures

- [1] George Klees et al. Evaluating fuzz testing. *CCS '18: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*, pages 2123–2138, 2018.
- [2] Barton Miller, Lars Fredriksen, and Bryan So. An empirical study of the reliability of UNIX utilities. *Communications of the ACM*, 33:32–34, 1990.
- [3] Own representation.
- [4] Bernhard Steffen, Falk Howar, and Maik Merten. Introduction to Active Automata Learning from a Practical Perspective. *SFM*, pages 256–296, 2011.

Design und Implementierung einer Datenpipeline-Architektur als Grundlage für IIoT-Anwendungen zur Nutzung industrieller Prozessdaten

Dennis Lehmann

Rainer Keller

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Einleitung

Immer mehr produzierende Unternehmen haben das Ziel sich verstärkt Richtung Industrie 4.0 auszurichten, um Prozesse, Routinen und Systeme effizienter zu gestalten. Dies betrifft neben Geschäftsbereichen wie Entwicklung, Service, Marketing und Vertrieb auch die Produktion. Daten werden hier zu einem wesentlichen Schlüsselfaktor, um die Wettbewerbsfähigkeit der Fertigung zu steigern. Der von der Deutschen Akademie der Technikwissenschaften (acatech) entwickelte Maturity Index und dessen zugehöriger Industrie 4.0-Entwicklungspfad (siehe Abbildung 1) zeigt die notwendigen Schritte für ein Unternehmen auf, um die digitale Transformation erfolgreich zu implementieren [5].

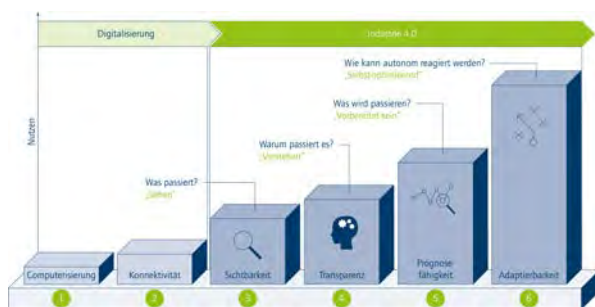


Abb. 1: Stufen des Industrie 4.0-Entwicklungspfads [5]

Dieser Transformationsprozess wird dabei in sechs Stufen unterteilt, wobei das Erreichen jeder einzelnen Stufe bereits einen Nutzenzuwachs verspricht. Die ersten beiden Stufen *Computerisierung* und *Konnektivität* sind der Digitalisierung eines Unternehmens zugeordnet. Hier liegt der Hauptfokus auf der Integration operativer Technologien (OT) und Informationstechnologien (IT). Während die OT hauptsächlich die Steuerungs- und Überwachungssysteme von Anlagen und Maschinen umfasst, gehören Rechen-, Netzwerk- und Speicherres-

ourcen zur IT [2]. Ist die Anbindung der verschiedenen Systeme erreicht, sind die Voraussetzungen für die dritte Stufe *Sichtbarkeit* geschaffen. Ab dieser Stufe wird von Industrie 4.0 gesprochen. Eines der zentralen Ziele dieser Stufe ist, durch Bereitstellung von Daten aus der Produktion datenbasierte Entscheidungen unterstützen zu können. Typischerweise haben solche Daten Charakteristika wie großes Volumen, hohe Geschwindigkeiten, hohe Variabilität und Zeitabhängigkeit, weshalb man diese oft als Industrial Internet of Things (IIoT) Daten klassifiziert [3]. Dementsprechend werden Datenpipeline und Anwendungen, die auf solchen Daten agieren als IIoT-Pipeline oder IIoT-Anwendungen bezeichnet.

Die hier vorgestellte Masterarbeit untersucht, wie in einem mittelständischen Unternehmen Prozessdaten mithilfe einer passenden IIoT-Architektur verfügbar gemacht werden können, um erhöhte Sicherheit in der Produktion zu schaffen. Zusätzlich wird durch die Implementierung eines Proof-of-Concepts (PoC) gezeigt, wie als eine erste Anwendung verschiedene Stakeholder aus dem Unternehmen ohne tiefere Programmierkenntnisse Produktionsdaten zu datenbasierten Entscheidungen nutzen können. Die zu untersuchenden Forschungsfragen sind folgendermaßen formuliert:

FF1: Wie können globale technische Prozessdaten mit unterschiedlichsten Eigenschaften standardisiert nutzbar gemacht werden, um datengetriebene Fertigung zu ermöglichen? **UF1:** Wie kann eine solche IIoT Architektur aussehen? **UF2:** Welche Rahmenbedingungen haben einen wesentlichen Einfluss auf das Design der Architektur? **UF3:** Welche neuen Möglichkeiten entstehen durch die Implementierung einer IIoT-Infrastruktur für den Ausbau der datengetriebenen Fertigung?

Grundlagen IIoT Architektur

Um Daten aus der Produktion zur einfachen Visualisierung bereitstellen zu können, ist eine entsprechende Infrastruktur in Form einer IIoT-Pipeline notwendig. Diese hat als Kernaufgaben das Sammeln, Aufbereiten und Speichern dieser Daten [5]. Um die Transparenz der in der Produktion ablaufenden Prozesse zu maximieren, müssen sowohl Echtzeit-Visualisierungen (Latenzen im Sekundenbereich) als auch historische Analysen unterstützt werden. Eine weitere Dimension der Daten ist die hohe Variabilität an Datenquellen. Im Produktionskontext werden häufig verschiedenste Maschinen und Anlagen, aber auch externe Sensoren angebunden. Dabei unterscheiden sich die Anlagen nicht nur in Aufbau und Konfiguration, sondern (vor allem aus informationstechnischer Sicht relevant) auch in Kommunikationsschnittstellen und Rechenpower. Werden Digitalisierungsprojekte in einer bereits existierenden Produktionslandschaft ausgerollt, spricht man auch von einem Brownfield. Im Vergleich zu einer neuen Produktionsumgebung, die als Greenfield bezeichnet wird, treten hier deutlich mehr Varianzen auf, da bereits bestehende Anlagen mehrere Jahrzehnte alt sein können und dadurch nicht für moderne Interoperabilität ausgelegt sind [6]. Diese Varianzen stellen besondere Herausforderungen sowohl für die IIoT-Architektur als auch für Datenmodelle dar.

Da Produktionsdaten teilweise nur in Echtzeit einen Mehrwert liefern und in diesem Zusammenhang schnell an Wert verlieren können, ist es notwendig, Technologien zu implementieren, die diese Daten mit möglichst geringer Latenz zur Verfügung zu stellen. Verschiedene Referenzarchitekturen, wie die Lambda- oder Kappa-Architektur versuchen diese Bedingung umzusetzen, indem einzelne Datenpunkte als Event klassifiziert werden und Streaming-Lösungen eingesetzt werden. Allgemein wird zwischen der Stream- und Batch-Verarbeitung unterschieden. Während die Stream-Verarbeitung potenziell echtzeitfähig ist, werden bei der Batch-Verarbeitung Datenpakete vorab gesammelt, transformiert und anschließend transportiert. Allgemein ist die Stream-Verarbeitung vor allem bei hohem Datenvolumen deutlich kostenintensiver,

weshalb hier eine ständige Abwägung zwischen verschiedenen Architekturkonzepten stattfindet [1]. Das hohe Volumen und die Geschwindigkeit aufkommender Produktionsdaten betrifft sowohl die Übertragungs-, die Bearbeitungs- als auch die Abspeicherungsprozesse. Um Übertragungsprozesse zu entlasten, werden vor allem hochfrequente Daten häufig durch Services auf lokalen Servern in der Produktion (on-premise) vorangegriert und gefiltert. Aktuelle Forschung beschäftigt sich zudem damit, auch nachfolgende Services lokal anzusiedeln, um möglichst echtzeitnah agieren zu können. Bei der Verteilung verschiedener Services auf lokale Server oder Cloud-Instanzen wird allgemein von Edge- und Cloud-Computing gesprochen. Je nachdem, wie die Anforderung an Latenz, Rechenpower und Skalierbarkeit gestaltet sind, gibt es auch hier unterschiedliche Architekturkonzepte und Best Practices [4].

Methodisches Vorgehen und Ausblick

Das angewandte Research-Design, um die gestellten Forschungsfragen beantworten zu können, basiert auf einer Literaturanalyse, Interviews mit Industriekontakten und der Analyse der Anforderungen und Rahmenbedingungen des Unternehmens, bei welchem die IIoT-Architektur umgesetzt werden soll. Die Ergebnisse aus der Literaturrecherche und den Industrieinterviews werden mit den Einschränkungen aus der Unternehmensanalyse kombiniert und herausgearbeitet, welche Konzepte im Unternehmensumfeld zur Umsetzung geeignet sind. Dieser Abgleich geht dabei auf zentrale Bewertungskriterien wie der Kompatibilität zu bereits bestehender IT-Umgebung, der Struktur und Aufbau der verantwortlichen IT-Abteilungen und den zukünftigen Plänen für Erweiterungen im Hinblick auf zusätzliche IIoT-Anwendungen. Aufbauend auf den daraus gewonnenen Erkenntnissen werden anschließend verschiedene IIoT-Architekturdesigns diskutiert. Basierend auf diesen drei Analysen werden mögliche IIoT-Architekturen diskutiert und bewertet, um einen Design-Vorschlag zu entwickeln, welcher dann als PoC umgesetzt und implementiert wird.

Literatur und Abbildungen

- [1] Moneer Helu, Timothy Sprock, Daniel Hartenstine, Rishabh Venketesh, and William Sobel. Scalable data pipeline architecture to support the industrial internet of things. *CIRP Annals - Manufacturing Technology*, 2020.
- [2] Yan Lu, Frank Riddick, and Nenad Ivezic. The Paradigm Shift in Smart Manufacturing System Architecture. *Advances in Production Management Systems*, 2016.
- [3] H.D. Nguyen, K. P. Tran, X. Zeng, L. Koehl, P. Castagliola, and P. Bruniaux. Industrial Internet of Things, Big Data, and Artificial Intelligence in the Smart Factory: a survey and perspective. *ISSAT International Conference on Data Science in Business, Finance and Industry*, pages 72–76, 2019.
- [4] Tie Qiu, Jiancheng Chi, Xiaobo Zhou, Zhaolong Ning, Mohammed Atiquzzaman, and Dapeng Oliver Wu. Edge Computing in Industrial Internet of Things: Architecture, Advances and Challenges. *IEEE Communications Surveys & Tutorials*, 2020.
- [5] Günther Schuh, Reiner Anderl, Roman Dumitrescu, Antonio Krüger, and Michael ten Hompel. Industrie 4.0 Maturity Index. Die digitale Transformation von Unternehmen gestalten. *acatech STUDIE*, 2020.
- [6] Tuan-Anh Tran, Tamas Ruppert, Gyorgy Eigner, and Janos Abonyi. *Retrofitting-Based Development of Brownfield Industry 4.0 and Industry 5.0 Solutions*. IEEE Access, 10 edition, 2022.

Analyse zeitbasierter Seitenkanäle und deren Einfluss auf die Sicherheit von Automotive-Software

Simon Leibfritz

Tobias Heer

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Vector Informatik GmbH, Stuttgart

Motivation und Problemstellung

Fahrzeuge bzw. deren Steuergeräte werden zum Schutz vor unbefugten Zugriffen längst mit kryptografischen Grundfunktionalitäten ausgerüstet. Ein Fokus der Automobilindustrie liegt darauf, die diversen Algorithmen auf ihre Sicherheit hin zu analysieren und an den Automobilkontext anzupassen. Oft wird dabei auf deren abstrakte mathematische oder algorithmische Beschreibung zurückgegriffen. Diese Betrachtung vernachlässigt jedoch die Implementierung sowie den Einsatz der Systeme, was je nach Szenario aber nicht minderwichtig für deren Sicherheit ist. So ist es möglich, dass eine kryptografische Bibliothek durch ihre Verhaltensweise unwillentlich brisante Informationen preisgibt. Derartige Schwachstellen werden als Seitenkanäle bezeichnet. Diese schwächen die theoretisch sicheren kryptografischen Algorithmen. Die Existenz dieser stellt somit eine Gefahr für das gesamte System dar. Eine Kategorie solcher Seitenkanäle sind die zeitbasierten Seitenkanäle. Diese beruhen auf der Zeitdifferenz, welche entsteht, wenn die kryptografische Bibliothek verschiedene, geheime Informationen verarbeitet.

In der hier vorliegenden Arbeit wird die Existenz, sowie die Auswirkung solcher zeitbasierter Seitenkanäle auf die Sicherheit von Steuergeräten im Automobilbereich untersucht. Hierzu wird exemplarisch, eine von der Vector Informatik GmbH entwickelte, kryptografische Bibliothek untersucht. Des Weiteren wird betrachtet, wie sich typische Einsatzszenarien im Automobilumfeld auf die Messbarkeit der erzeugten Seitenkanäle auswirken. Das Ziel ist, einen realistischen Eindruck von der Verbreitung zeitbasierter Seitenkanal-Schwachstellen in automobiler Software zu erhalten. Letztendlich kann so eine Bewertung der Plausibilität von Angriffen auf eine solche Art von Schwachstellen gewonnen werden.

Entstehung und Auswirkung

Zeitbasierte Seitenkanäle entstehen, wenn die Ausführungsdauer einer kryptografischen Funktion messbar

von geheimem Material abhängt. Ein Beispiel hierfür stellt eine Passwortkontrolle dar, welche beim ersten falschen Buchstaben im Passwort abbricht. Über das einfache Variieren des ersten Buchstabens und das Messen der benötigten Zeit kann ein Angreifer den korrekten Buchstaben herausfinden. Als Konsequenz dieser Schwachstelle wird die Komplexität eines Brute-Force-Angriffes drastisch reduziert (von $|A|^n$ auf $|A|*n$ für die möglichen Buchstaben im Passwort $|A|$ und die Länge des Passworts n).

Vorgehen

Grundlegend besteht das betrachtete Gesamtsystem aus

1. der Software
2. der Hardware
3. dem Umfeld, in welchem die Bibliothek eingesetzt wird

Jeder Teil des Systems trägt auf seine Weise dazu bei, die Ausprägung einzelner Seitenkanäle zu verstärken oder abzuschwächen. Will ein Angreifer präzise Informationen über die Laufzeit-Dauer einer Operation erlangen, so ist dies bedeuten schwerer, wenn das zu testende System nur über ein Netzwerk mit stark variierender Latenz zu erreichen ist. Auf der anderen Seite können jedoch Features, wie Cache-Zugriffsoptimierung oder Compiler-Optionen, eine solche Attacke erleichtern. Um das System also adäquat analysieren zu können, ist es essenziell, alle Teilaspekte und deren Verknüpfung zu betrachten.

Da die meisten Programme jedoch von einer großen Menge an Parametern abhängen, ist die Ursache zeitbasierter Seitenkanäle leider sehr schwierig zu erkennen und noch schwerer formal nachzuweisen. Ein möglicher Ansatz besteht darin, die benötigte Zeit des Algorithmus bei verschiedenen Inputs zu messen. Jedoch bedeutet die Absenz messbarer Zeitunterschiede nicht zwangsweise das Nichtvorhandensein

eines zugehörigen Seitenkanals. Durch Faktoren, wie beispielsweise die Änderung der System-Auslastung oder dem Austausch der Hardware, können bisher nicht nachweisbare Unterschiede plötzlich messbar, und dementsprechend angreifbar, werden. Auch eine simple Code-Analyse ist nicht vollkommen ausreichend. Mechanismen wie Compiler-Optimierung können den zeitlichen Verlauf einzelner Operation beeinflussen und in diesem Zuge Zeitunterschiede erneut messbar machen.

Um möglichst fundiert an das Problem heranzugehen, stützt sich die vorliegende Arbeit auf eine Kombination beide dieser Ansätze. Dieser Fakt geht auch aus Abbildung 1 hervor, welche eine grafische Darstellung des erarbeiteten Analyse-Konzepts darstellt. Da eine existierende Schwachstelle in beiden der angewendeten Methoden nachweisbar sein muss, validieren sich die beiden Ansätze selbst. Somit ist der angewendete Prozess stabil gegenüber falsch-positiven Ergebnissen.

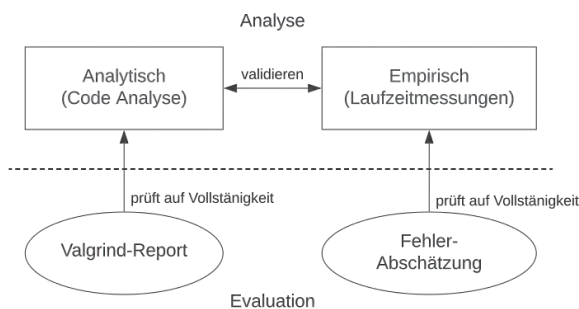


Abb. 1: Grafische Darstellung des zum Einsatz kommenden Analyse-Konzepts [1]

Diesem Prozess kommt zugute, dass er in gewissem Maße bereits selbst validierend ist. Wird ein potenzieller Seitenkanal im Code festgestellt, so muss dessen Existenz auch messbar sein. Analog gilt, dass für gemessene Zeitunterschiede auch eine Ursache im Code zu finden sein muss. Zusätzlich ist der Prozess stabil gegenüber falsch-positiven Ergebnissen, welche sich nur bei einer der beiden Analyse-Methoden nachweisen ließen.

Ein größeres Problem stellt der Nachweis der Vollständigkeit dar. Um diesbezüglich die Genauigkeit der angewendeten Methoden beurteilen zu können, werden diese mit einer Reihe von Test-Cases auf die Probe gestellt. Hierfür wurde eine Test-Applikation entwickelt, welche präzise bekannte Zeitvarianzen beinhaltet. Durch das Variieren dieser Zeitunterschiede im Code kann eine Einschätzung der maximal messbaren Genauigkeit und somit auch eine Einschätzung des Fehlers gemacht werden.

Ergebnisse

Die hier thematisierte Arbeit beschäftigt sich mit drei Algorithmen. Konkret handelt es sich hierbei um die Algorithmen Advanced Encryption Standard (AES), Rivest-Shamir-Adleman (RSA) und Elliptic Curve Digital Signature Algorithm (ECDSA). Die Implementierung aller analysierten Algorithmen zeigt Anzeichen für zeitbasierte Seitenkanal-Schwachstellen. Dabei scheint es keinen direkten Zusammenhang zwischen der Komplexität des Algorithmus und der Ausprägung des Seitenkanals zu geben. So sind selbst simple Algorithmen und Funktionen oft von dieser Art Schwachstellen betroffen.

Im Falle der analysierten Bibliothek stellte sich heraus, dass vor allem Basisfunktionalitäten Ursachen für Seitenkanäle sind. Ein prominentes Beispiel hierfür ist die Modulare-Multiplikation ($\tilde{x} * \tilde{y} \bmod n$) zweier beliebig großen Zahlen. Solche Multiplikationen finden vor allem im Rahmen asymmetrischer Algorithmen Verwendung. Um die Berechnungen möglichst effizient zu gestalten, verwendet die analysierte Bibliothek hier den sogenannten Montgomery-Reduktion-Algorithmus [3]. Dieser Algorithmus benötigt in seltenen Fällen eine zusätzliche Subtraktion, welche sich durch Zeitmessungen erkennen lässt. Dies ist vor allem ein Problem, wenn eine gleichbleibende Zahl \tilde{x} (z. B. der private Schlüssel) mit Nutzer-Input multipliziert wird. Durch das zufällige Variieren des Inputs kann ein Angreifer verschieden große Zahlen erzeugen und jeweils aufzeichnen, ob diese eine zusätzliche Subtraktion erzeugen. Wird dies genügend oft wiederholt, approximiert das gemessene Verhältnis die eigentliche Wahrscheinlichkeit. Dieses Konzept ist vergleichbar mit einer Monte-Carlo-Integration. Sofern nun also ein direkter Zusammenhang zwischen dem privaten Schlüssel und dieser Wahrscheinlichkeit besteht, kann ein Angreifer Informationen über den Schlüssel gewinnen. Und tatsächlich bewies W. Schindler diesen Zusammenhang bereits im Jahre 2000 [4]. Schindler zeigte, den nahezu linearen Anstieg der Wahrscheinlichkeit in Abhängigkeit der Größe des privaten Schlüssels \tilde{x} , was auch in Abbildung 2 dargestellt ist.

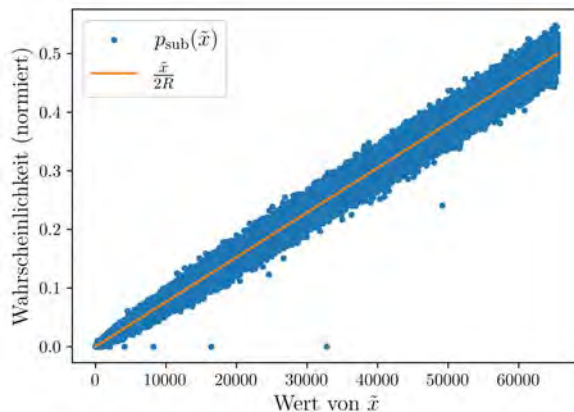


Abb. 2: Wahrscheinlichkeit einer zusätzlichen Reduktion in Abhängigkeit des privaten Schlüssels. [1]

Diese Schwachstelle in der Multiplikation zweier Zahlen ist verantwortlich für eine Reihe anderer Schwachstellen in den Algorithmen RSA und ECDSA. Mit entsprechend genauen und vielen Messungen erlauben manche dieser Schwachstellen das Auslesen des gesamten privaten Schlüssels (es existieren hier einige Angriffe aus verwandter Literatur wie beispielsweise [4] oder [2]).

Auswirkungen des Automobilumfelds

Aus dem oben erwähnten Beispiel der Multiplikation zweier Zahlen wird klar, dass zwei Faktoren besonders wichtig für die Angreifbarkeit einer solchen Seitenkanal-Schwachstelle sind.

1. Wie genau ein Angreifer Zeitmessungen am System durchführen kann. Messungen mit viel Ungenauigkeit verdecken die erzeugten Zeitdifferenzen, und schwächen den Seitenkanal ab.
2. Wie viele Messungen ein Angreifer durchführen kann. Um beispielsweise die Wahrscheinlichkeit

einer zusätzlichen Subtraktion im obigen Beispiel zu approximieren benötigt ein Angreifer eine große Menge an Messungen. Bei realen Angriffen auf einen RSA werden beispielsweise mehrere Millionen Messungen benötigt.

Die Betrachtung des Automobilumfelds wird entsprechend dieser beiden Kriterien durchgeführt.

Um die Genauigkeit der zur Verfügung stehenden Messungen beurteilen zu können, werden geläufige Szenarien der Interprozesskommunikation (IPC) betrachtet. Neben Shared-Memory wird hier auch die Kommunikation zweier Geräte über eine UDP/IP Verbindung betrachtet. Bei den meisten getesteten Methoden bleiben, trotz großer Latenzunterschiede, selbst kleinste Zeitunterschiede gut messbar. Eine nennenswerte Ausnahme stellen hier IPC-Methoden dar, welche den Datenaustausch über zyklisches Prüfen (Polling) realisieren. Solche Methoden sind demnach automatisch resistenter gegen Seitenkanalattacken. Um die Frage nach der Menge erlangbarer Messungen zu beantworten, werden typische Automobilprotokolle betrachtet. Hier existieren mit Diagnostics over IP (DoIP) und Unified Diagnostic Services (UDS) bereits zwei Protokolle, welche dem Angreifer erlauben, eine beliebige Menge an kryptografischer Operationen anzustoßen.

Fazit

Die zugrundeliegende Arbeit zeigt eindeutig die Existenz von zeitbasierten Seitenkanäle in Automotive Software. Oftmals lassen sich diese Seitenkanäle selbst über komplexere Kommunikations-Szenarien messen. Dabei begünstigen manche typischen Automobilprotokolle sogar die Angreifbarkeit der jeweiligen Schwachstellen. Einige dieser Schwachstellen könne, unter den richtigen Umständen, zur Freilegung des gesamten Schlüsselmaterials führen und stellen damit durchaus ernstzunehmende Schwachstellen dar.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Paul C Kocher. Timing attacks on implementations of Diffie-Hellman, RSA, DSS, and other systems. *Annual International Cryptology Conference*, 1996.
- [3] Peter L. Montgomery. Modular Multiplication Without Trial Division. *Springer*, 1985.
- [4] Werner Schindler. A timing attack against RSA with the chinese remainder theorem. *Cryptographic Hardware and Embedded Systems—CHES 2000: Second International Workshop Worcester, MA, USA, August 17–18, 2000 Proceedings 2*, 2000.

rungsverlinkung erarbeitet werden, indem anfallende Daten aufbereitet und zur Verfügung gestellt werden, wodurch die Produktprojekte in der Vorbereitung von Meilensteinen ihren inhaltlichen Fortschritt über Kennzahlen aus dem System aufzeigen können.

Anforderungsmanagement



Abb. 2: Einflussfaktoren auf den RE-Prozess [3]

Requirements Engineering, englisch für Anforderungsmanagement, trägt bei der Systementwicklung entscheidend zum Projekterfolg bei. Es beschreibt den Umfang eines Projekts und bildet die Kommunikationsbasis für alle Bereiche, die an dem Projekt beteiligt sind. Requirements Engineering (RE) kann sich positiv auf die Projektkosten auswirken: je besser das RE, desto weniger Fehler treten in der Entwicklung auf, welche später ressourcenintensiv behoben werden müssen. Unter dem Strich stehen somit geringere Gesamtkosten, obwohl Anforderungsmanagement an sich erst einmal Mehrkosten verursacht, vgl. [2]. Das Requirements Engineering ist ein wichtiger Bestandteil des Systems Engineering und unterstützt Unternehmen dabei, den Überblick über immer komplexer werdende Systeme zu behalten, während sich das Systems Engineering mit der gesamtheitlichen Entwicklung eines Systems befasst. In 1 sind die relevanten Nutzenaspekte des Anforderungsmanagements zu sehen. Während sich jeder Anforderungsmanagementprozess normalerweise von Projekt zu Projekt unterscheidet, da sich die äußeren und inneren Einflussfaktoren, wie zum Beispiel Stakeholder, Zeitpläne, Vertrags- beziehungsweise Auftraggeberverhältnisse, Vorkenntnisse aus Vorentwicklungen oder aber auch Kundenbedürfnisse, ändern, soll ebendieser Prozess bei Stahl möglichst für alle Entwicklungsprojekte einheitlich sein, vgl. [3]. Der Requirements Engineer ist mit der Aufgabe betraut, die Methoden und Hilfsmittel so zu wählen, dass sie die eigenen Kompetenzen dabei nicht übersteigen, die Ziele und Anforderungen der Stakeholder vollständig

und in der richtigen Art und Weise erarbeitet und dokumentiert werden und im Optimalfall die Artefakte seiner/ihrer Arbeit in folgenden Projekten wiederverwendet werden können. Hierbei gilt zu beachten, dass Unternehmen oft vor allem in Phasen des Aufbaus des Anforderungsmanagements keine eigenen Requirements Engineer Stellen schaffen, sondern bestimmte Mitarbeiter zu sogenannten Experten fortbilden und diese sich die Rolle des Requirements Engineer teilen. Je nach Entwicklung des Requirements Engineering in einem Unternehmen kann dies dann im Laufe der Zeit nachgeholt werden, sodass ein Team aus Experten geschaffen werden kann, welches sich nur mit Themen rund um das Anforderungsmanagement auseinandersetzt. Die Wiederverwendbarkeit ist nach ersten Auswertungen einer der wichtigsten Nutzenaspekte für Fachabteilungen in der Entwicklung. Mit Wiederverwendbarkeit wird im Kontext des Requirements Engineering der Prozess der Auswertung auf Aufbereitung genutzter Anforderungen und Inhalte beschrieben, sodass diese für folgende Projekte zur erneuten Nutzung verfügbar sind. Unternehmen können so von verkürzten Entwicklungszeiten profitieren, da bei Folgeprojekten nicht immer wieder von Null an gestartet werden muss und so erheblich weniger Ressourcen in gleichbleibende Inhalte gesteckt werden müssen. Hierfür werden am Ende des Projekts die aufbereiteten Anforderungen in ein sogenanntes Master-Dokument zusammengeführt/hinzugefügt, aus welchen sich Folgeprojekte bedienen und qualitativ hochwertig formulierte Anforderungen beziehen können.

Ausblick

Noch während der Durchführung der Interviews zur Analyse der aktuellen Vorgehensweisen und Bedarfen der Fachabteilungen kristallisiert sich der Bedarf an niederschwelliger Bereitstellung von projektbezogenen Informationen ab. Aufgrund der aktuell stattfindenden schrittweisen Veränderung und Entwicklung der Requirements Engineering Prozesse, welche die Einführung eines neuen, schwergewichtigen Softwaretools beinhaltet, soll ein Konzept erarbeitet werden, welches diese Bedarfe abdeckt. Durch leichtgewichtige Lösungen wird im Optimalfall der Bedarf, die im Rahmen des Anforderungsmanagements erstellten Daten automatisiert auszuwerten, vollständig abgedeckt, indem unter anderem auf die jeweiligen Meilensteine maßgeschneiderte Reports und Auswertungen in die Projektarbeit und Prozesse integriert werden. Tiefergehendere Auswertungen der durchgeführten Interviews sollen im Anschluss an die Abschlussarbeit durch den Autor sowie die Verantwortlichen stattfinden, um die erhaltenen Rückmeldungen nachhaltig in die Prozessgestaltung einfließen lassen zu können.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] IREB International Requirements Engineering Board. IREB ausführlich - International Requirements Engineering Board. <https://www.ireb.org/de/about/ireb>, 2023.
- [3] Christine Rupp and die SOPHISTEN. *Requirements- Engineering und -Management*. Carl Hanser Verlag; SOPHIST-Gesellschaft für Innovatives Software-Engineering, 7 edition, 2021.

Vergleich der Webframeworks Spring und Phoenix anhand der Implementierung einer Cloud-nativen Messenger-Anwendung

Nico Linder

Jörg Nitzsche

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma doubleSlash Net-Business GmbH, Stuttgart

Einleitung

Um von der Flexibilität, Skalierbarkeit und Redundanz einer verteilten Infrastruktur zu profitieren, werden viele Webanwendungen heute von Anfang an für den Betrieb in der Cloud konzipiert. Für die Entwicklung solcher Cloud-nativen Anwendungen kommen bei der doubleSlash Net-Business GmbH oft Frameworks zum Einsatz, die es ermöglichen, verteilte Anwendungen mit etablierten Programmiersprachen wie Java zu entwickeln, während weniger verbreitete Sprachen oder Frameworks oft nicht berücksichtigt werden, obwohl sie Vorteile bei der Entwicklung Cloud-nativer Software bringen könnten.

Elixir und Phoenix

Elixir ist eine funktionale Programmiersprache auf Basis der Erlang VM, die speziell für verteilte, fehlertolerante Systeme entwickelt wurde. Die Elixir-Plattform bietet einige Features, die sie für die Entwicklung verteilter Anwendungen attraktiv machen, wie Skalierbarkeit durch nebenläufige Prozesse und automatischem Neustart abgestürzter Komponenten. [2] Zusammen mit dem Webframework Phoenix, das auf Elixir basiert, soll sie die Entwicklung von Microservices oder vollständiger Webanwendungen ermöglichen, bei denen Skalierbarkeit und Ausfallsicherheit zu den Prinzipien der Programmiersprache gehören und nicht durch zusätzliche Abhängigkeiten bereitgestellt werden müssen. [3]

Zielsetzung

Ziel der Bachelorarbeit ist es, die Frameworks Spring für Java und Phoenix für Elixir im Hinblick auf die Entwicklung einer verteilten Anwendung zu evaluieren und zu vergleichen. Besonderes Augenmerk beim Vergleich liegt darauf, inwiefern die beiden Frameworks

Konzepte der Cloud-nativen Entwicklung unterstützen. Zu den Faktoren, die betrachtet werden sollen, gehören:

- Startzeit der Anwendung
- Antwortzeiten
- Horizontale und vertikale Skalierbarkeit
- Ressourcenverbrauch
- Docker-Integration und Image-Größe

Weitere Vergleichskriterien sind die Developer Experience und allgemeine Features der Frameworks bzw. der zugrundeliegenden Programmiersprachen wie Datenbankintegration, Logging und Error Handling. Schließlich soll auf die Frage eingegangen werden, ob die bei Elixir bereits in die Sprache integrierten Mechanismen zur Fehlertoleranz und Skalierung durch Nebenläufigkeit, die oft als besondere Stärken der Sprache hervorgehoben werden [5], einen Mehrwert bei der Entwicklung verteilter Anwendungen bringen.

Lösungsansatz

Zunächst wird ein Prototyp einer Messenger-Anwendung konzipiert und jeweils mit Spring und Phoenix implementiert. Die Anwendung ermöglicht es Clients, über eine Websocket-Verbindung bidirektional Nachrichten nach dem Pub/Sub-Pattern auszutauschen (siehe Abbildung 1). In der Spring-Anwendung wird hierfür das STOMP-Protokoll verwendet, die Phoenix-Anwendung nutzt die Phoenix Channels-Implementierung. Beide Implementierungen erlauben es Clients, Nachrichten an bestimmte Zielgruppen („Topics“) zu senden, sich für das Empfangen von Nachrichten bei Topics zu registrieren, sowie die serverseitige Verarbeitung von Nachrichten, bevor diese weitergeleitet werden. [3] [4]

Beide Versionen der Anwendungen sollen dann als Docker-Images gebaut werden. Hier soll die Docker-Integration der Frameworks verglichen werden, also wie hoch der Aufwand ist, ein lauffähiges Image zu erzeugen und welche Unterstützung die Frameworks bieten. Mit den Images sollen dann alle weiteren Vergleiche durchgeführt werden.

Zum Vergleich der Ressourceneffizienz soll mithilfe von Lasttests ermittelt werden, wie viele Anfragen die jeweilige Anwendung mit begrenzten Ressourcen gleichzeitig bearbeiten kann. Auch die Zeit, die zur Bearbeitung einer Anfrage bei niedriger bzw. hoher Last benötigt wird, soll so verglichen werden, insbesondere sind hier Extremwerte von Interesse.

Um die vertikale Skalierbarkeit zu bewerten, werden den Systemen mehr Ressourcen zur Verfügung gestellt und die Lasttests wiederholt. Durch Vergleich mit den vorherigen Tests kann bewertet werden, inwiefern die Anwendungen von mehr Ressourcen profitieren.

Zur Evaluierung der horizontalen Skalierbarkeit wird untersucht, welche Möglichkeiten die Frameworks bieten, die Last auf mehrere Instanzen der Anwendung zu verteilen. Insbesondere soll betrachtet werden, ob der designbedingte Fokus von Elixir auf skalierbare

Prozesse hier einen Vorteil gegenüber einer Implementierung in Java darstellt.

Für eine bessere Vergleichbarkeit sollten möglichst alle betrachteten Features von den Frameworks selbst bereitgestellt werden und nicht von zusätzlichen externen Komponenten abhängig sein.

Für die Bewertung der Developer Experience können Features oder Vorgehensweisen, die den Entwicklern bereits aus Firmenprojekten mit Spring bekannt sind, als Kriterien herangenommen werden. Indem geprüft wird, ob Phoenix bzw. Elixir vergleichbare Features bietet, kann darauf geschlossen werden, ob Phoenix Entwicklern, die es gewohnt sind mit Spring zu arbeiten, potenziell eine positive Developer Experience bieten kann. Kriterien der Developer Experience können die Sprache direkt betreffen, wie Typsicherheit, Komplexität und Umfang des Ökosystems, oder beschreiben, welches Feedback Entwickler beim Entwickeln, Testen und Debuggen bekommen. Je nach Definition umfasst der Begriff Developer Experience noch viele weitere Aspekte, im Rahmen dieser Arbeit werden jedoch nur solche betrachtet, die auf der verwendeten Sprache bzw. Framework basieren.

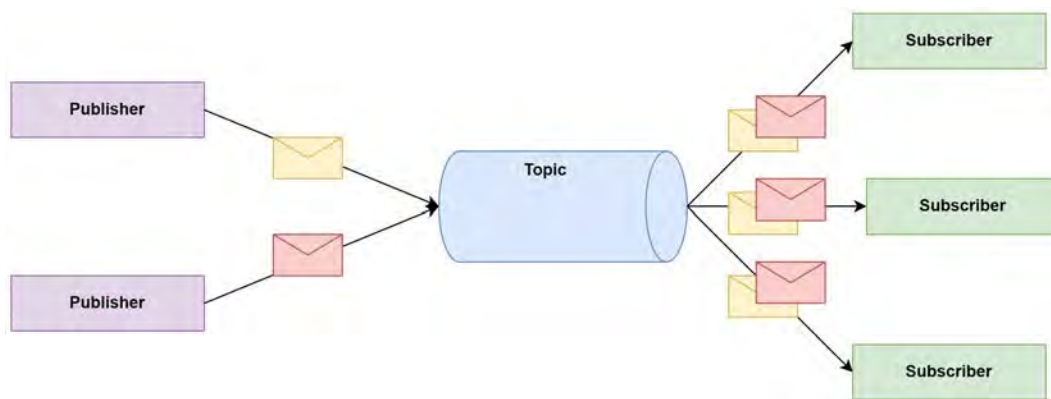


Abb. 1: Nachrichtenaustausch nach dem Publish-Subscribe-Pattern [1]

Ausblick

Im weiteren Verlauf sollen die konkreten Vergleichskriterien für Developer Experience und allgemeine Sprachfeatures definiert werden. Die Evaluierung der Ergebnisse aus dem Vergleich der Cloud-native-Kriterien soll zeigen, ob Phoenix im betrachteten Anwendungsfall eine technisch sinnvolle Alternative zu etablierteren Frameworks darstellt. Die Evaluierung

der Developer-Experience kann einen Hinweis auf die erwartbare Akzeptanz und Produktivität in einem Entwicklungsumfeld geben, in dem bisher mit Spring entwickelt wurde. Ein möglicher nächster Schritt wäre die Integration eines auf Phoenix basierenden Service in ein reales Projekt, entweder als Erweiterung oder Ersatz für einen bestehenden Service, um die Ergebnisse dieser Arbeit in einer produktiven Umgebung zu überprüfen.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Gints Dreimanis. Introduction to Elixir. <https://serokell.io/blog/introduction-to-elixir>, 2020.
- [3] Geoffrey Lessel. *Phoenix in Action*. Manning Publications, 2019.
- [4] Aliaksandr Liakh. WebSockets With Spring, Part 3: STOMP Over WebSocket. <https://medium.com/swlh/websockets-with-spring-part-3-stomp-over-websocket-3dab4a21f397>, 2020.
- [5] Elvio Vicoso. 5 reasons Elixir can be your competitive advantage. <https://medium.com/@elviovicosa/5-reasons-elixir-can-be-your-competitive-advantage-7ad884e9cb98>, 2019.

Hardware-Tracing von Standard Utility Modules (GM). Eine Laufzeitanalyse auf einem Infineon Aurix System.

Svenja Luik

Rainer Keller

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Vector Informatik GmbH, Weilimdorf, Stuttgart

Motivation

Im Laufe der letzten Jahre sind die Anforderungen an Computerchips und Mikrocontroller stetig gewachsen. Deshalb werden unter anderem auch Multi-Core Mikrocontroller eingesetzt. Solche Systeme besitzen mehrere Kerne und können so Prozesse im besten Fall parallel und effizient ausführen. Auch in der Automobilindustrie werden zunehmend solche leistungsstarken Mikrocontroller eingesetzt. Je nach Einsatzbereich muss für das Steuergerät jedoch genau überprüft werden, ob die Verwendung eines Multi-Core-Systems sinnvoll ist, denn die Systeme sind deutlich komplexer und kostenintensiver. Um die Laufzeit eines solchen Multi-Core-Systems analysieren und vergleichen zu können, soll ein realistischer Anwendungsfall eines Steuergeräts für den Single- und Multi-Core-Fall verglichen werden. Dazu wird zum Vergleich die fahrzeugherstellerspezifische Applikationssoftware von General Motors, sogenannte Standard Utility Modules (SUMs), für beide Fälle konfiguriert. Durch Laufzeitmessungen soll die Systemauslastung beider Konfigurationen bestimmt und miteinander verglichen werden. Dazu wird im Folgenden die verwendete Tracing-Methode sowie das Hardware-Setup beschrieben, um so eine erste Einschätzung für die zu erwartenden Ergebnisse der Laufzeitmessung geben zu können.

Aufbau

Auf den Steuergeräten ist MICROSAR Classic von Vector integriert, welches als AUTOSAR-Classic-Basissoftware mit Betriebssystem die Hardware abstrahiert und zudem für die darüberliegende Applikationssoftware Schnittstellen bereitstellt. Die einzelnen Prozesse von Applikation und Betriebssystem werden über Runnables (nebenläufige Prozesse) und Tasks ausgeführt und verwaltet. Um nun die Laufzeit der Tasks und Runnables überwachen und Tracen zu können, wird der Programmcode zunächst instrumentiert. Das heißt, der Programmcode muss so angepasst werden, dass vor und nach dem Aufruf der zu messenden

Funktion ein Zeitstempel (Timestamp) gespeichert wird. Der Zeitstempel enthält dann Informationen über den Start (oder Stopp) und die Identifikationsnummer des jeweiligen Tasks oder Runnables. Dazu werden sogenannte „Hooks“ implementiert. Die ursprünglichen Funktionsaufrufe werden umdefiniert, sodass anstelle der Funktion, zunächst eine Start-Hook aufgerufen wird. Diese speichert den Zeitstempel und es wird anschließend die eigentliche Funktion ausgeführt. Nach Rückkehr (Return) der zu messenden Funktion, wird durch die zugehörige Stopp-Hook ein weiterer Zeitstempel gespeichert. Damit wird die Laufzeit des Tasks oder Runnables durch die beiden Zeitstempel abgebildet. Wird der Task oder das Runnable unterbrochen, werden für die jeweilige unterbrechende Funktion ebenfalls Zeitstempel generiert, welche dann in der späteren Auswertung die entsprechende Unterbrechung des ursprünglichen Runnables darstellen. Durch die Hooks wird zusätzlicher C-Code zum Programm hinzugefügt und ausgeführt, was zu einem geringen Overhead im Programmablauf führt. Die Hooks werden dabei so implementiert, dass ihr Einfluss auf die Laufzeit und Auslastung der Soft- und Hardware möglichst gering ist. Das Umdefinieren der zu messenden Funktionen und die Generierung der Hooks wird durch Verwenden des „Vector Timing-Bundles“ automatisch ausgeführt [5]. Angepasst an die Auswahl des Debuggers und der Tracing-Methode, stellt Vector die Timing-Bundles der gängigsten Systeme frei zur Verfügung. Das hier verwendete Lauterbach Timing-Bundle besteht aus wenigen C-Dateien, welche dem MICROSAR-Classic Projekt hinzugefügt und mitkompiliert werden. Wird im Projekt das Virtual-Funktion-Bus-Tracing (VFB-Tracing) aktiviert, können die zu messenden Runnables ausgewählt werden. Durch Automationskripte wird dann das Tracing initialisiert und die Hooks anschließend generiert. Damit die Laufzeit der gesamten Software auf dem Steuergerät aufgezeichnet und analysiert werden kann, wird eine Tracing-fähige Hardware benötigt. In diesem Fall besteht der Aufbau aus einem Infineon Aurix TriCore TC397XE mit

sechs Kernen und einem Emulationsgerät (Emulation Device), welches den Mikrocontroller mit zusätzlichen Debugging- und Tracing-Eigenschaften erweitert [1]. Neben dem Mikrocontroller gehört ein Tracing-fähiger Debugger zum Aufbau. Der hier verwendete Debugger erlaubt es Laufzeitinformationen der sechs Kerne über mehrere Minuten hinweg aufzuzeichnen. Der Debugger ist über eine Infineon DAP-Schnittstelle (Device Access Port) mit dem TC397 verbunden. Über die DAP-Schnittstelle können die Laufzeitdaten des Systems in Echtzeit aufgezeichnet und mit hohen Frequenzen übertragen werden. Als Messmethode wurde das OLDA-Tracing verwendet. Beim „OnLine Data Acquisition“-Tracing werden die Messdaten aus den Tracing-Hooks zur Laufzeit in einem Multibuffer zwischengespeichert und später über die verfügbare Debug bzw. Trace Schnittstelle, hier über die DAP-Schnittstelle, ausgelesen. Mit dieser Messmethode lassen sich Ereignisse im Bereich von Sekunden bis hin zu Nanosekunden aufgezeichnet werden. Die Auswertung und Visualisierung der Laufzeitanalyse findet nach der Aufzeichnung in der Timing Architects Tool Suite von Vector statt. Diese Software kann unter anderem die Messdaten auswerten, sodass die CPU-Auslastung und der Einfluss der jeweiligen Runnables, Tasks und Interrupts auf das System berechnet werden können. Im „Inspection“ Modus kann dargestellt werden, zu welcher Zeit, welche Tasks ausgeführt wurden. Die Abbildung 1 zeigt eine solche Visualisierung [4].

Buskommunikation und Messung

Mit dem zuvor beschriebenen Aufbau wird nun ein MICROSAR Projekt mit integrierten Standard Utility Modules (SUM) von General Motors ausgemessen und analysiert. Die verschiedenen SUMs bilden zusammen eine fahrzeugherstellerspezifische Applikationssoftware, welche unter anderem die Interaktion zwischen Basissoftware und Applikationssoftware abstrahiert und spezifiziert. Um die Funktionalität der SUMs zeigen zu können, enthält das Projekt eine übergeordnete Softwarekomponente, welche die Standard Utility Modules als Komposition aufruft und ihren Funktionsumfang sichtbar macht. Dazu werden Nachrichten in Form von CAN-Botschaften zyklisch geprüft, ver- oder entschlüsselt und versendet bzw. empfangen. Um die CAN-Kommunikation auf dem CAN-Bus sichtbar zu machen, wird mit Hilfe der Simulationssoftware CANoe von Vector, ein zweiter Knoten im Netzwerk

simuliert, sodass die vom Steuergerät gesendeten CAN-Botschaften quittiert (Acknowledged) werden. Ohne die Quittierungen würde das Steuergerät in einen Ruhezustand versetzt werden. Nachrichten, welche das Steuergerät erwartet zu empfangen, werden ebenfalls über das Simulationstool auf den CAN-Bus gesendet und so dem Steuergerät übermittelt [3]. Die zu sendenden oder zu empfangenen Nachrichten sind über Systemsignale in der Datenbank des Projekts definiert und über das Werkzeug DaVinci Configurator Classic von Vector konfiguriert. Hierbei wird zwischen Tx (Transmit) und Rx (Receive) Nachrichten und deren jeweilige Eigenschaften unterschieden. Neben einzelnen Nachrichten werden auch Nachrichtenbündel versendet. Die Bündel bestehen aus zusammengehörigen Botschaften, welche in kurzen Abständen als CAN-Frames gesendet werden. Wie die normalen Botschaften, müssen auch die gebündelten Signale von der Applikationssoftware ausgelöst werden um in Form von CAN-Botschaften gesendet zu werden. Damit muss der Aufruf zum Senden einer Botschaft von der Applikationsebene, über die standardisierten Schnittstellen und der SUMs, bis zur Basissoftware durch den gesamten Software-Stack geleitet werden, um dann schließlich über die hardwarenahe Software der Interfaces gesendet zu werden. Dieser Prozess benötigt nicht nur Zeit, sondern auch CPU-Ressourcen. Durch das Hinzufügen neuer Nachrichten, sollte die Laufzeit und die Beanspruchung der CPU mit jeder zusätzlichen Nachricht erhöht deutlich werden. Um die Belastung des Single-Core-Systems mit einem Multi-Core System mit gleicher Konfiguration vergleichen zu können, wird zunächst das Baseline-Projekt, welches SUMs und Basissoftware enthält, so umkonfiguriert, dass mehrere Kerne des Mikrocontrollers verwendet werden. Aufgrund des relativ großen Integrationsaufwands, beschränkt sich das Multi-Core-Projekt auf zwei Kerne des Aurix TC397. Dafür werden Basissoftware und Applikationssoftware voneinander getrennt und auf verschiedenen Kernen verteilt. Um den CAN-Bus des Mikrocontrollers verwenden zu können, muss die Applikationssoftware nun also mit der Basissoftware auf dem Hauptkern kommunizieren.

Ergebnisse

Die beiden Konfigurationen können dann, mit Hilfe des Tracings und der anschließenden Auswertung in der TA Tool Suite, verglichen und analysiert werden.

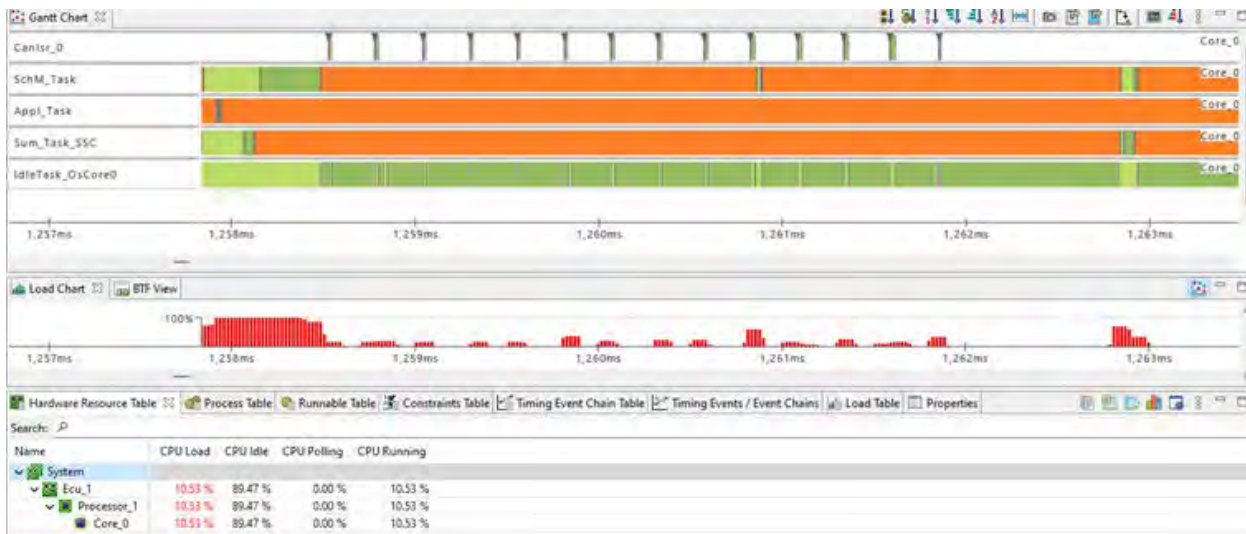


Abb. 1: Darstellung einer Laufzeitmessung und CPU-Auslastung des Single-Core-Systems in der TA Toolsuite. [2]

Das Baseline-Projekt enthielt bereits wenige vorkonfigurierte Tx- und Rx- Nachrichten und Signalbündel. Bei zyklischen Aufrufen der Nachrichten wies das Single-Core Projekt eine CPU-Auslastung von etwa 10-11% auf, das Multi-Core Projekt im Durchschnitt ebenso. Bei genauerer Betrachtung der einzelnen Kerne, zeigten die Messungen (siehe Abbildung 2) eine Auslastung von ca. 10% pro Kern. Damit wurde keine Verbesserung der Auslastung oder Laufzeit durch eine Migration auf Multi-Core erzielt und stattdessen ein doppelter Overhead erzeugt. Grund hierfür ist die aufwändige Kommunikation von Basis- und Applikationssoftware zwischen den Kernen. Bei der Cross-Core-Kommunikation müssen Speicherzugriffe durch Spinlocks abgesichert werden, um Inkonsistenzen der Daten zu verhindern [6]. Damit die beiden Kerne kommunizieren können, müssen außerdem X-Signals verwendet werden, welche Aufrufe auf dem anderen Kern durch Interrupts verarbeiten. Vergleicht man Laufzeit des Single-Core-Systems, mit der des Multi-Core-Systems für eine Konfiguration mit wenigen

Nachrichten, also einer geringen Auslastung, scheinen beide Systeme ähnlich zu performen. Mit nur einem Kern ergibt sich für das System eine Auslastung von ca. 10-11%. Für das System mit zwei Kernen ergibt sich im Durchschnitt ein ähnlicher Wert von ca. 11% CPU-Auslastung. Betrachtet man die einzelnen Kerne genauer, ist zu erkennen, dass beide Kerne mit einer Auslastung von ca. 9-11% arbeiten. Hier führt der zusätzliche Kern zu einem größeren Overhead durch die zusätzlich benötigte Delegation und Cross-Core Kommunikation. Das führt zu einer insgesamt schlechteren Leistung, verglichen mit dem Single-Core-System. Die Anzahl der Nachrichten soll nun so erhöht werden, dass das Single-Core-System mit ca. 30% Auslastung arbeitet. Für das Multi-Core-System sollte sich für die gleiche Konfiguration nun ein besserer durchschnittlicher Wert der CPU-Auslastung ergeben, da Basis- und Applikationssoftware getrennt voneinander ausgeführt werden und so die CPU-Ressourcen effizienter genutzt werden können.

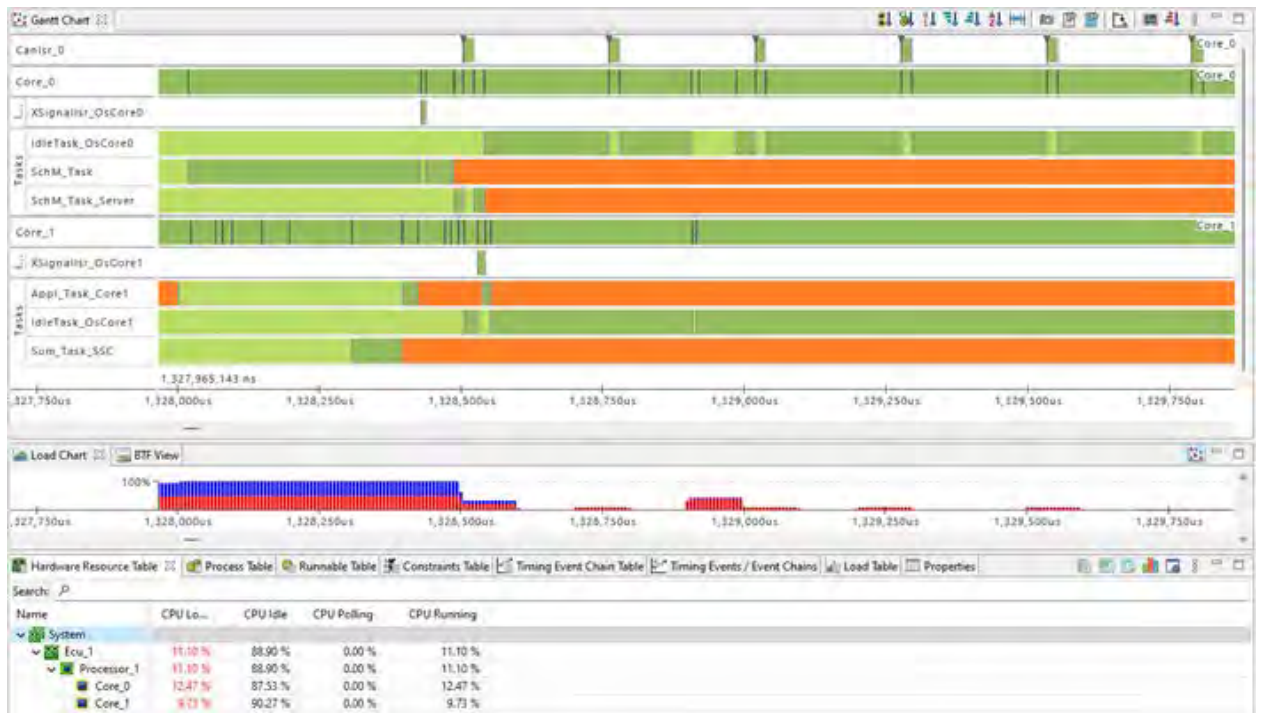


Abb. 2: Darstellung einer Laufzeitmessung und CPU-Auslastung des Multi-Core-Systems in der TA Toolsuite. [2]

Für den Allgemeinen Fall eines Steuergeräts im Automotive-Bereich sollte jedoch nicht pauschal auf ein Multi-Core-System gewechselt werden. Die potenzielle Verbesserung durch den Einsatz mehrerer Kerne muss für den jeweiligen Anwendungsfall abgeschätzt oder durch fundiertes Tracing überprüft werden. Der zusätzliche Overhead, welcher durch die aufwändige

Kommunikation zwischen den Kernen entsteht, kann unter Umständen die zusätzliche Leistung aufheben. Zu beachten ist außerdem, dass Integrationsaufwand und Kosten für Multi-Core-Systeme oft sehr hoch sind, sodass eine Redundanz von Single-Core-Systemen eine preiswertere Alternative sein können.

Literatur und Abbildungen

- [1] Infineon Technologies AG. Aurix TC3XX User Manual 1. https://www.infineon.com/dgdl/Infineon-AURIX_TC3xx_Part1-UserManual-v02_00-EN.pdf?fileId=5546d462712ef9b701717d3605221d96, 02 2021.
- [2] Eigene Darstellung.
- [3] Vector Informatik GmbH. Product Information CANoe V6.5. [https://cdn.vector.com/cms/content/products/canoe/canoe/docs/Product Informations/CANoe_ProductInformation_EN.pdf](https://cdn.vector.com/cms/content/products/canoe/canoe/docs/Product%20Informations/CANoe_ProductInformation_EN.pdf), 10 2022.
- [4] Vector Informatik GmbH. TA Tool Suite Fact Sheet, DE. https://cdn.vector.com/cms/content/products/TA_Tool_Suite/Docs/TA_Tool_Suite_FactSheet_DE.pdf, 02 2022.
- [5] Andrea Martin. Case Study Profiling. https://cdn.vector.com/cms/content/products/TA_Tool_Suite/Docs/Lauterbach_Case_Study_Profiling.pdf, 2023.
- [6] R. Martins Gomes, F. Mauroner, and M. Baunach. Collaborative Resource Management for Multi-Core AUTOSAR OS. In *Betriebssysteme und Echtzeit*, pages 99–108. Springer Vieweg, 2015.

Methodik zur Bewertung und Verwaltung der Datenqualität in einer Powertrain Solution Pipeline

Christof Meyer

MarkusENZweiler

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Robert Bosch GmbH, Schwieberdingen

Einleitung

Der digitale Wandel spielt heutzutage eine existenzielle Rolle in den Arbeitsabläufen eines Unternehmens. Es werden immer größere Mengen an Daten gesammelt, verarbeitet und ausgewertet. Hierbei hat die Qualität der Daten eine immer größer werdende Wichtigkeit. Um die Qualität der Daten sicherzustellen ist es wichtig, ein entsprechendes Datenqualität Konzept zu entwickeln, welches bei der Erkennung und Behebung von fehlerhaften Daten zur Hilfe kommt. Dabei wird ein Datenqualitätsmanagement Framework ausgewählt und im Prozess der Datenverwaltung verankert. Zur Messung der Datenqualität, werden Datenqualitätskriterien definiert und den Daten zugeordnet. Die meistverwendeten Datenqualitätskriterien sind Genauigkeit, Vollständigkeit, Konsistenz, Pünktlichkeit, Gültigkeit und Einzigartigkeit. Es existieren allerdings auch noch weitere Kriterien anhand denen die Datenqualität gemessen werden kann.

Ziel der Arbeit

Ziel dieser Masterarbeit ist es, die Datenqualität in einer, von Robert Bosch GmbH verwalteten, Powertrain

Solution Pipeline am Ende zu bewerten und aktive Maßnahmen anzuwenden, um die Datenqualität zu verbessern. Die Herausforderung hierbei ist es, fehlerhafte Daten zu erkennen und entsprechende Vorkehrungen zu treffen, wie zum Beispiel die Filterung der Daten oder die Alarmierung von potenziellen Fehlern. Um dies zu ermöglichen, soll ein entsprechendes Datenqualitätsmanagement Framework entwickelt und angewendet werden.

Powertrain Solution Pipeline

Die zu analysierenden Daten werden durch eine Powertrain Solution Pipeline gesammelt. Der sogenannte *Powertrain* ist das Herzstück eines Fahrzeugs, bestehend aus Motor, Getriebe und weiteren Komponenten, die den Antrieb des Fahrzeugs ermöglichen. Während der Entwicklung eines Powertrain werden eine Vielzahl an Sensoren verbaut, welche Daten bezüglich des Zustandes des Motors liefern. Diese Daten werden anhand einer Pipeline zur Analyse von unterschiedlichen Anwendungsfällen gesammelt und aufbereitet. In Abbildung 1 ist eine vereinfachte Darstellung des Datenflusses dargestellt.

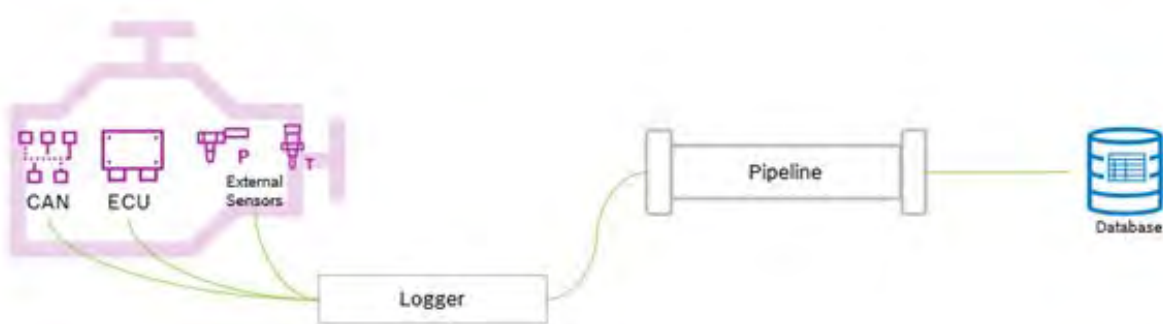


Abb. 1: Vereinfachte schematische Übersicht der Powertrain Solution Pipeline [1]

Total Data Quality Management (TDQM)

TDQM [3] ist ein umfassender Ansatz zum Datenqualitätsmanagement, der die Festlegung von Datenqualitätszielen, die Implementierung von Datenqualitätskontrollen, die Überwachung der Datenqualität und die kontinuierliche Verbesserung der Datenqualität im Laufe der Zeit umfasst. Die Methodik kann in vier Phasen eingeteilt werden, welche in Abbildung 2 abgebildet sind.

- *Define*: Umfasst die Datenanalyse, Definition der Datenqualitätskriterien und Modellierung des Informationsverarbeitungsprozesses
- *Measure*: Messung der Datenqualität
- *Analyze*: Identifizierung der Ursache für die Diskrepanz
- *Improve*: Definition und Umsetzung eines vorgehen zur Verbesserung

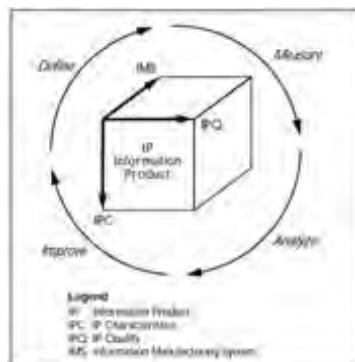


Abb. 2: Schema der TDQM-Methodik [3]

Umsetzung des TDQM

Zur Umsetzung des TDQM, wird in einem ersten Schritt eine Übersicht der tool-Kette erstellt, um zu verstehen, welche Daten gemessen werden und wie diese zusammenhängen. Anhand der definierten

Qualitätskriterien und Metriken wird anschließend die Datenqualität gemessen. Um einen Eindruck der aktuellen Datenqualität zu erhalten, soll ein Dashboard entwickelt werden, das eine Übersicht der gemessenen Kriterien darstellt. Zur Verbesserung der Datenqualität wird ein *Plan-Do-Check-Act (PCDA)* oder auch Deming-Zyklus verwendet [2]. PDCA ist ein vierstufiges iteratives Verfahren zur Problemlösung und kontinuierlichen Verbesserung. In der „Plan“ Phase werden die Ziele und Probleme der Datenqualität festgelegt und eine Verbesserungsmethode entwickelt, um das gewünschte Ziel zu erreichen. Während der „Do“ Phase wird dieser Plan angewendet. Nach der „Do“ Phase wird in der „Check“ Phase überprüft, ob das Qualitätsproblem bereits behoben wurde. Wenn das Problem weiterhin besteht, wird ein neuer Plan entwickelt. Sobald die „Check“ Phase erfolgreich war, werden in der „Act“ Phase Maßnahmen ergriffen, um dauerhaft die Datenqualität zu sichern. Dies kann durch eine Filterung der Daten oder bei komplexen Problemen durch eine Benachrichtigung der zuständigen Person erfolgen.

Ausblick

Nach der Anwendung des TDQM-Methodik, in welchem mehreren Qualitätskriterien angewendet wurden, soll die Qualität der Daten einfacher zu bewerten sein und mithilfe des Dashboards eine gute Übersicht liefern. Darüber hinaus soll es durch die zuvor identifizierten Filter möglich sein, die Daten auf bestimmte Qualitätsmerkmale zu überprüfen. Bei Problemen, welche nicht durch einen Filter behoben werden können, soll eine Benachrichtigung an die betroffenen Nutzer ausgegeben werden. Dabei ist es wichtig, Informationen und eine mögliche Methode zur Problembhebung zu liefern. Wie die Filter und Benachrichtigungen zukünftig umgesetzt werden, steht aktuell in der Planung mit darauffolgender Implementierung. Darüber hinaus wird das Dashboard erweitert, um eine genauere Analyse der Daten und weiterer Qualitätskriterien zu ermöglichen. Es existieren allerdings auch noch weitere Kriterien anhand denen die Datenqualität gemessen werden kann.

Literatur und Abbildungen

[1] Eigene Darstellung.

[2] Darr Kurt. Quality Improvement: The Pioneers. *Informa {UK} Limited*, 85:35–38, 2007.

[3] Y. Wang Richard. A product perspective on total data quality management. *Association for Computing Machinery ({ACM})*, 41:58–65, 1998.

Integration von PaaS und SaaS zur Verwaltung sowie Verknüpfung verschiedener IoT-Edge-Geräte und Services für intelligente Analyseprozesse

Nikolai Miguel Mora Lozano

Rainer Keller

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Tecomon GmbH, Denkendorf

Motivation

Heutzutage gibt es immer mehr smarte Geräte die eigenständig Daten sammeln und diese verarbeiten können, auch IoT (Internet of Things) -Edge-Geräte genannt, da sie am Rand des Netzwerks arbeiten. Sie können meistens selbst oder über das Gateway mit dem sie verbunden sind, mit dem Internet verbunden werden und viele Daten aus ihren eingebauten Sensoren sammeln. Dies löst bei vielen Anwendern Privatsphäre bedenken aus, doch sicher integrierte und anonyme Datenweitergabe, können IoT-Edge-Geräte anhand der Kombination vieler Daten, aus verschiedenen Geräten sehr hilfreich sein. Da aber jeder Hersteller meistens seine eigenen Lösungen für die eigenen Gateways (welche die eigenen Geräte verwalten) oder für eigenständige Geräte ohne Gateway entwickelt, ist es schwer Daten verschiedenster Geräte zu kombinieren und daraus intelligente Analysen herzuleiten. Dieses Problem wird auch als Schnittstellen-Problem bezeichnet. Um dies zu lösen, benötigt man Software, die diese Arbeit im Hintergrund übernimmt und den Fokus auf das Entwickeln von Services zur Kombination und Analyse der gewonnenen Daten setzt.

Ziel der Arbeit

Das Ziel der Arbeit ist es eine funktionstüchtige Cloud-Architektur zu schaffen, zur Verwaltung von IoT-Edge-Geräten und deren Verknüpfung mit Services zur Verarbeitung der bereitgestellten Daten. So soll es später möglich sein, neue Services anhand gegebener Vorlage zu entwickeln und der bestehenden Architektur beizufügen. Die Lösung des Schnittstellen-Problems übernehmen hier schlanke Microservices. Sie wissen wie man mit dem verbundenen Gateway oder Gerät kommuniziert und geben die gewonnenen Daten in einem so genannten „kanonischen Format“ aus, eine Form der Daten, die intern nun standardisiert ist und von jedem Service verstanden wird. Hierbei sollen die Cloud-

Service Modelle PaaS (Platform as a Service) und SaaS (Software as a Service) integriert werden, da das System so aufgebaut werden soll, dass es später auch als Plattform genutzt werden kann, um neue Services zur Verarbeitung der Daten auf diesem zu entwickeln.

Portal

Das Portal übernimmt die Funktion des Frontend, für die dahinterliegende Microservice-Architektur in der Cloud. Die Portal Web-App wurde mit React entwickelt und wird auf dem Web-Server der Firma gehostet. Über das Portal kann sich der Nutzer mit seinem Account einloggen und seine registrierten Gateways, gemanagten Geräten und verknüpften Services einsehen (siehe Abbildung 1).



Abb. 1: Portal Übersicht [2]

Die Gateways stellen dabei physische Geräte oder Clouds dar, die mehrere IoT-Edge-Geräte registriert haben und diese mit nur einer Registrierung im Portal mitbringen. Dabei können Clouds auch verknüpfte Benutzer-Accounts sein, die neben physikalischen Rohdaten, auch bereits vorverarbeitete physiologische Daten zum Gerät anbieten. Hat man nun Geräte im Portal, kann man diese managen, dies beinhaltet die Vergabe von Namen, Icons, Zonen (Räume) oder Tags.

Es ist auch möglich durch eine erweiterte Ansicht die UID (Unique Identifier), die aktuell empfangenen Daten und die unterstützten Einheiten für die Gerätedaten einzusehen z.B. „W“ für Watt. Die unterstützten Einheiten der Geräte sind essenziell, um herauszufinden mit welchen Services die Geräte verknüpft werden können. Das Portal zeigt dem Nutzer verfügbare Services an und er kann anhand der kompatiblen Geräte, diese einem Service zuweisen. So ist es beispielsweise möglich mehrere Geräte, die eine Unterstützung für die Einheit Watt bieten, mit einem Stromkosten-Service zu verbinden. In diesem Service ist es dann möglich unabhängig des Ursprungs der Geräte, ihre Daten zu bündeln und so wertvolle Informationen zu sammeln.

Microservices

Die Services im Backend werden in einer Microservice-Architektur betrieben. Für diese Microservices wird momentan Node.js mit dem Webframework Express.js genutzt. Sie laufen alle in einem Container, dieser kann von Docker und in Kubernetes ausgeführt werden. Die Nutzung dieser Architektur bietet einige Vorteile, einer ist beispielsweise eine bessere Stabilität, da das System nicht komplett abstürzt, wenn ein Microservice für die Verbindung zu einem Gateway abstürzt. Das Gegenteil der Microservice-Architektur ist die Monolith-Architektur, hier wird eine große Anwendung nicht in einzelne schlanke Anwendungsteile aufgeteilt. In Abbildung 2 sind diese beiden Architekturen am Beispiel einer Datenbanknutzung gegenübergestellt. „Microservices lassen lieber jeden Service seine eigene Datenbank verwalten. Dies können entweder verschiedene Datenbanken der gleichen Technologie oder ganz unterschiedliche Datenbank-Systeme sein – ein Ansatz, der Polyglot Persistence genannt wird“ [3].

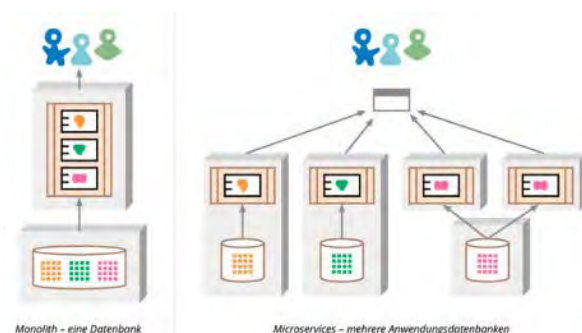


Abb. 2: Gegenüberstellung: Monolith vs. Microservices [3]

Kubernetes

Um die Vielzahl der Microservice-Container zu orchestrieren, wird in dieser Arbeit Kubernetes genutzt. Kubernetes bietet eine sehr große Vielfalt an Funktionen. Zu den Funktionen gehören beispielsweise, Automatisierte Rollouts und Rollbacks (siehe Abbildung 3). Beim Ausrollen eines Deployment Updates nutzt Kubernetes die Methode „Slowing Rollouts to Ensure Service Health“, dabei beinhaltet ein Deployment in Kubernetes mehrere Einstellungen, wie das genutzte Container-Image oder die Anzahl der Pods (ein Pod beinhaltet immer mindestens einen Container). Der Vorteil eines stufenweisen Rollouts ist es, sicherzustellen, dass ein Rollout in einen fehlerfreien, stabilen Service mit neuer Softwareversion resultiert. Um dies zu garantieren, wartet der Deployment Controller bis ein Pod seine Bereitschaft meldet, bevor mit der Aktualisierung des nächsten Pods fortgefahren wird. [1]

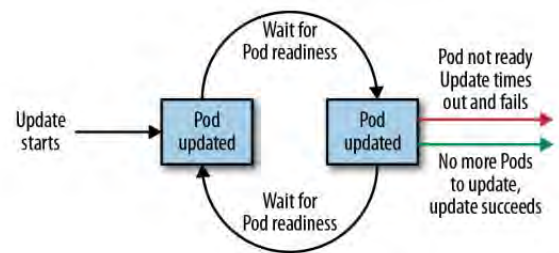


Abb. 3: Kubernetes Rollout [1]

Ausblick

Da diese Infrastruktur auch als PaaS zur Entwicklung von Services zur Verarbeitung von IoT-Edge-Geräte Daten genutzt werden soll, kann man diese noch durch Frontend- und Backend Erweiterungen erweitern, um das System einfacher mit neuen Services zu ergänzen. Die Nutzung der Infrastruktur als PaaS kann sogar so weit gehen, dass die Plattform als Marktplatz für Drittanbieter-Services dient. Außerdem können an diese Arbeit weitere Projekte angeknüpft werden, um mithilfe dieser Plattform intelligente Analyseprozesse zu realisieren, wie z.B. ein intelligenter Service zur Anwesenheitskontrolle in Räumen.

Literatur und Abbildungen

- [1] Brendan Burns, Joe Beda, and Kelsey Hightower. *Kubernetes Up & Running: Dive into the Future of Infrastructure*. O'Reilly, 2 edition, 2019.
- [2] Eigene Darstellung.
- [3] Martin Fowler and James Lewis. *Microservices: Nur ein weiteres Konzept in der Softwarearchitektur oder mehr?* OBJEKTSpektrum, 2015.

Konzeption und Implementierung eines chinesischen V2X-Stacks für ein POSIX-Betriebssystem

Aaron Mueller

Michael Scharf

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Vector Informatik GmbH, Stuttgart

Motivation

Der Straßenverkehr ist in den letzten Jahren immer sicherer, umweltfreundlicher und effizienter geworden, u. a. mit der Einführung von Fahrerassistenzsystemen und der steigenden Popularität von Elektrofahrzeugen. Um die Fortschritte nochmals zu erhöhen, ist die Vernetzung von Fahrzeugen mit ihrer Umwelt sinnvoll [6]. Man spricht hierbei von Vehicle to Everything (V2X) Kommunikation. Dieser Begriff umfasst direkte Kommunikation von Fahrzeugen untereinander, sowie von Fahrzeugen zu Netzwerkbasistationen (z.B. für Internetzugriff), Fußgängern oder Infrastruktur, wie z.B. Ampelanlagen. Der Fokus der Bachelorarbeit liegt auf Kommunikation zwischen mehreren Fahrzeugen, auch Vehicle to Vehicle (V2V) genannt.

Einer der größten Vorteile von V2X ist das Sammeln von Umgebungsdaten. Das ist besonders wichtig für Fahrzeuge mit höheren Automatisierungsgraden [8]. Bisher stammen diese allein von Sensoren, deren Sichtfeld begrenzt ist. Daten aus V2X Kommunikation können einen größeren Bereich abdecken, etwa hinter blinden Kurven oder bei schlechten Witterungsverhältnissen. Außerdem können sie zusätzliche Statusinformationen, wie z.B. die exakte Position und Bewegungsrichtung, enthalten.

Vergleich der Standards

Für V2X gibt es verschiedene regionale Standards; u. a. der europäische Standard (ETSI ITS [7]) und der chinesische Standard, veröffentlicht von der China Communications Standards Association ([1], [2]). Die grundlegende Architektur der Protokollstapel (Stacks) dieser Standards ist sehr ähnlich, es werden jedoch verschiedene Protokolle verwendet. Beide Stacks sind nebeneinander in Abb. 1 dargestellt.

Ein wesentlicher Unterschied zwischen beiden Stacks ist die verwendete Übertragungstechnologie. Während der europäische Standard auf dem WLAN Standard 802.11p basiert, sieht der chinesische Standard die Verwendung von Mobilfunkstandards, wie z.B. LTE, vor. Man spricht dabei auch von C-V2X (Cellular V2X). Die darüber liegenden Layer abstrahieren die Übertragungstechnologie zu höheren Schichten. Eine der Hauptaufgaben ist die Adressierung der Verkehrsteilnehmer, mithilfe sogenannter Layer-2 IDs, oder anhand der geografischen Position. Der europäische Standard definiert hierfür die Protokolle GeoNetworking und Basic Transport Protocol (BTP), der chinesische Standard verwendet einen Abstraction Layer sowie das Dedicated Short Message Protocol (DSMP).

Die V2X Nachrichten sind im Facilities bzw. Message Layer definiert. Die wichtigsten Nachrichtentypen des europäischen Stacks sind die Cooperative Awareness Message (CAM) und die Decentralized Environmental Notification Message (DENM).

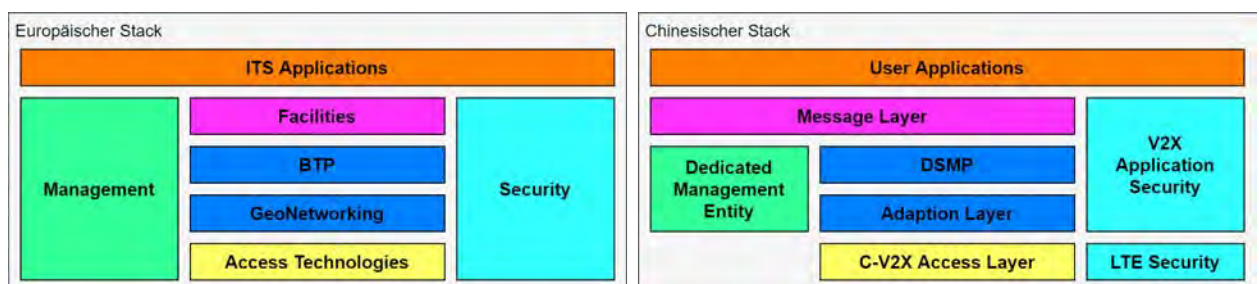


Abb. 1: Architektur der europäischen und chinesischen Stacks im Vergleich [3]

CAMs werden periodisch gesendet und enthalten Statusinformationen. Ergänzend dazu informieren DENMs andere Verkehrsteilnehmer über auftretende Ereignisse, wie z.B. Vollbremsungen. Der chinesische Standard vereint beide in einem Nachrichtentypen, der Basic Safety Message (BSM).

Die Daten aus den empfangenen Nachrichten werden sogenannten User Applications zur Verfügung gestellt, welche die Verkehrssicherheit und Effizienz erhöhen, sowie zusätzliche Services ermöglichen. Beispiele für Anwendungsfälle sind z.B. Kollisionswarnung, automatische Mautzahlung, Teilen von Sensordaten oder kooperative Manöver von autonomen Fahrzeugen. User Applications sind eigenständige Anwendungen, und kommunizieren über Interprocess Kommunikation (IPC) mit dem Facilities bzw. Message Layer.

Beide Standards sehen außerdem zwei Layerübergreifende Module vor: Ein Management Modul, was z.B. für Congestion Control Mechanismen verwendet wird, und ein Security Modul, welches V2X Nachrichten signiert und verschlüsselt.

Ziel der Arbeit

Das Ziel dieser Bachelorarbeit ist die Konzeptionierung und Implementierung eines V2X-Stacks, basierend auf dem chinesischen Standard. Dieser soll in ein bestehendes Projekt integriert werden, welches aktuell eine Implementierung des europäischen Stacks verwendet. Das bestehende Projekt wurde in C++ entwickelt, und baut auf der POSIX Programmierschnittstelle PSE 51 auf.

Es sollen Gleichanteile zwischen dem chinesischen und europäischen Stack identifiziert, und wenn möglich wiederverwendet werden. Am Ende soll eine Konfigurationsmöglichkeit implementiert werden, um zur Kompilierzeit zu entscheiden, ob der chinesische oder der europäische Standard verwendet wird. Der V2X Stack wird in eine Virtuelle Maschine mit Linux Betriebssystem integriert. Für den Test des Stacks kann das von Vector entwickelte Simulations- und Analysetool CANoe zur Restbussimulation verwendet werden. Das Programm unterstützt mit der Erweiterung CANoe.Car2x [5] beide relevante Standards, und ermöglicht die Simulation komplexer Verkehrsszenarien.

Anbindung an CANoe

Zur Anbindung an CANoe wird ein Remote Access Layer (RAL) nach dem Vorbild der Autosar Spezifikation [4] implementiert. Der RAL trennt den Access Layer vom Rest des V2X-Stacks, sodass die Nachrichten von einer externen Antenne ausgesendet werden können. Benötigte Sendeparameter werden dazu in einen dedizierten Header gepackt, welcher mitsamt der V2X Nachricht an die Antenne gesendet wird. Auf der Gegenseite, dem RAL Gateway, wird dann der Header entpackt und die Nachricht über die Antenne ausgesendet.

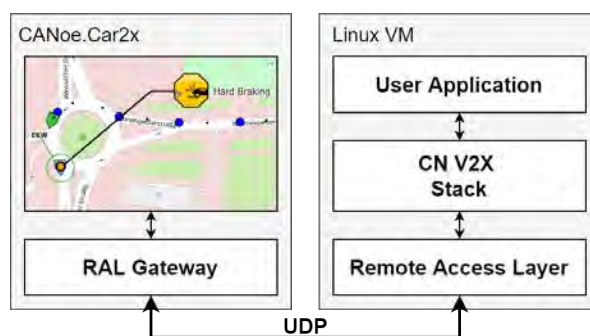


Abb. 2: Simulationsaufbau [3]

Das RAL Gateway wird als ein Knoten im Simulationsnetzwerk implementiert, und kann somit Nachrichten senden und empfangen. Für die Kommunikation zwischen dem V2X Stack und CANoe wird UDP verwendet. Der komplette Simulationsaufbau ist in Abb. 2 zu sehen.

Ausblick

Für Softwarehersteller ist es wichtig, alle regionalen Standards vor der Entwicklung neuer Produkte zu berücksichtigen. Deshalb ist eine modulare und konfigurierbare Lösung für die V2X Stacks sinnvoll. Der Vergleich der europäischen und chinesischen Standards hat gezeigt, dass architektonische Ähnlichkeiten bestehen, obwohl unterschiedliche Protokolle und Technologien verwendet werden. Das Projekt wurde daraufhin umstrukturiert, um Gleichanteile von Standard-spezifischen Teilen zu trennen und den Einsatz der Stacks konfigurierbar zu gestalten. Als nächster Schritt wird eine CANoe Konfiguration erstellt, in der BSM Nachrichten generiert und über den RAL an den V2X Stack gesendet werden.

Literatur und Abbildungen

- [1] China Communications Standards Association. *YD/T 3707-2020*. CCSA, 2020.
- [2] China Communications Standards Association. *YD/T 3709-2020*. CCSA, 2020.
- [3] Eigene Darstellung.
- [4] AUTOSAR Gbr. Vehicle-2-X Remote Access Layer Protocol Specification. https://www.autosar.org/fileadmin/standards/R22-11/FO/AUTOSAR_PRS_V2XRemoteAccessLayer.pdf, 11 2022.
- [5] Vector Informatik GmbH. CANoe.Car2x - Produktinformation. https://cdn.vector.com/cms/content/products/canoe/_car2x/Docs/CANoe_Car2x_Product_Information_DE.pdf, 11 2022.
- [6] Qualcomm Technologies Inc. Introduction to Cellular V2X. https://www.qualcomm.com/content/dam/qcomm-martech/dm-assets/documents/c-v2x_intro.pdf, 10 2019.
- [7] European Telecommunications Standards Institute. ETSI EN 302 665. https://www.etsi.org/deliver/etsi_en/302600_302699/302665/01.01.01_60/en_302665v010101p.pdf, 09 2010.
- [8] Chanyoung Jung et al. V2X-communication-aided autonomous driving: System design and experimental validation. *Sensors*, 20, 2020.

Workflow zur Frühwarnerkennung von gesundheitlichen Veränderungen bei Patienten auf Basis von Web-Technologien

Robin Mueller

Astrid Beck

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Philips Medizin Systeme Böblingen GmbH, Böblingen

Einleitung

Die Möglichkeit, Menschen mit diversen Erkrankungen in einem klinischen Kontext zu helfen, hat sich in den letzten Jahrzehnten rapide verbessert. Die Grundlage für diese Verbesserung basiert auf der kontinuierlichen Evaluation von derzeitigen Prozessen und der Identifizierung von Problemen in diesen. Neue Behandlungsmethoden, die bessere Schulung von Personal und die Einführung neuer technischer Hilfsmittel konnten dazu beitragen, klinische Ergebnisse nachhaltig zu verbessern. Trotz dieser Verbesserungen verbleiben einige Probleme, die im schlimmsten Falle zum Tod eines Patienten führen können. Eines der Hauptprobleme ist die nicht erkannte Verschlechterung im Gesundheitszustand von Patienten. So beschreibt beispielsweise ein Bericht des *Patient Safety Observatory* des NHS im Jahre 2007, dass in 66 % der Zwischenfälle, die als Konsequenz den Tod eines Patienten nach sich zog, die klinische Verschlechterung des Patienten nicht erkannt wurde, bzw. keine Reaktion des Personals erfolgte [4].

Um eine gute Versorgung der Patienten zu erreichen, werden eine wachsende Anzahl an softwarebasierten Lösungen genutzt, um verschiedenste Prozesse im Klinikalltag effizienter und auch sicherer zu gestalten. So werden beispielsweise Patientenmonitore weitverbreitet zur Überwachung von Patienten eingesetzt. Neben der reinen Überwachung der Vitalwerte eines Patienten verfügen diese mittlerweile über viele weitergehende Funktionen, wie z.B. Frühwarnsysteme zum Erkennen von Verschlechterungen im Gesundheitszustand. Hierzu werden oftmals so genannte *Early Warning Score* (EWS) Systeme genutzt, diese sind klinische Prognoseinstrumente, die Veränderungen in den Vitalzeichen von Patienten erfassen und eine frühzeitige Erkennung und Eskalation der Versorgung bei klinischer Verschlechterung ermöglichen.

Ziel der Arbeit

Das Ziel der Arbeit ist es einen Prototyp eines EWS Protokoll Editors zu konzipieren, zu implementieren und zu evaluieren. Ein solches Werkzeug ermöglicht es medizinischem Personal, individuelle EWS Systeme zu erstellen. Dies ist nötig, da kein einheitliches internationales System zur Bewertung des Gesundheitszustandes von Patienten existiert. Zudem werden für verschiedene Patientengruppen oft verschiedene Bewertungssysteme eingesetzt. Ein sehr wichtiger Aspekt im Erstellungsprozess ist die Usability der zu erstellenden Software, da die Akzeptanz beim Personal in großem Maße davon abhängt. Daher wird ein besonderes Augenmerk auf diesen Usability-Aspekt gelegt.

Vorgehensweise

Um die Anforderungen an die Software aufzustellen, werden zu Beginn der Arbeit bestehende EWS Systeme analysiert. Zudem wird der derzeitige Markt für Produkte dieser Art erörtert und verglichen. Hierdurch wird mögliches Verbesserungspotential gegenüber diesen herausgestellt. Aufbauend auf diesen Grundlagen wird ein grundlegendes Konzept mittels des Prototyping Werkzeugs *Figma* erstellt. Dabei werden verschiedene Designansätze erprobt und verglichen. Besonders wichtig für die spätere Usability ist die Befolgung der Empfehlungen der ISO Norm 9241 für Ergonomie in der Mensch-System-Interaktion [3] und das Einbeziehen des *Philips digital Design Language System*. Ausgehend vom erstellten Konzept wird eine prototypische Implementierung mithilfe von Web-Technologien vorgenommen. Dabei werden Web Components genutzt, um die verschiedenen Oberflächen des Editors zu realisieren. Web Components bieten den Vorteil, dass sie eine auf Browserebene standardisierte Technologie darstellen. Hierdurch können die erstellten UI-Komponenten unabhängig vom gewählten Framework (bspw. React,

Angular, Vue) verwendet werden. So können die erstellten Komponenten in anderen Projekten mit ggf. anderen Technologien frei wiederverwendet werden. In Abbildung 1 wird ein beispielhafter Ausschnitt aus dem erstellten EWS-Editor gezeigt. Diese Oberfläche baut sich aus mehreren Web Components auf.

Abschließend wird die erstellte Applikation mithilfe eines Usability-Tests mit 20 Testpersonen evaluiert, um die Benutzungsfreundlichkeit sicherzustellen. Zur Bewertung wird die *System Usability Scale* (SUS) verwendet [2].

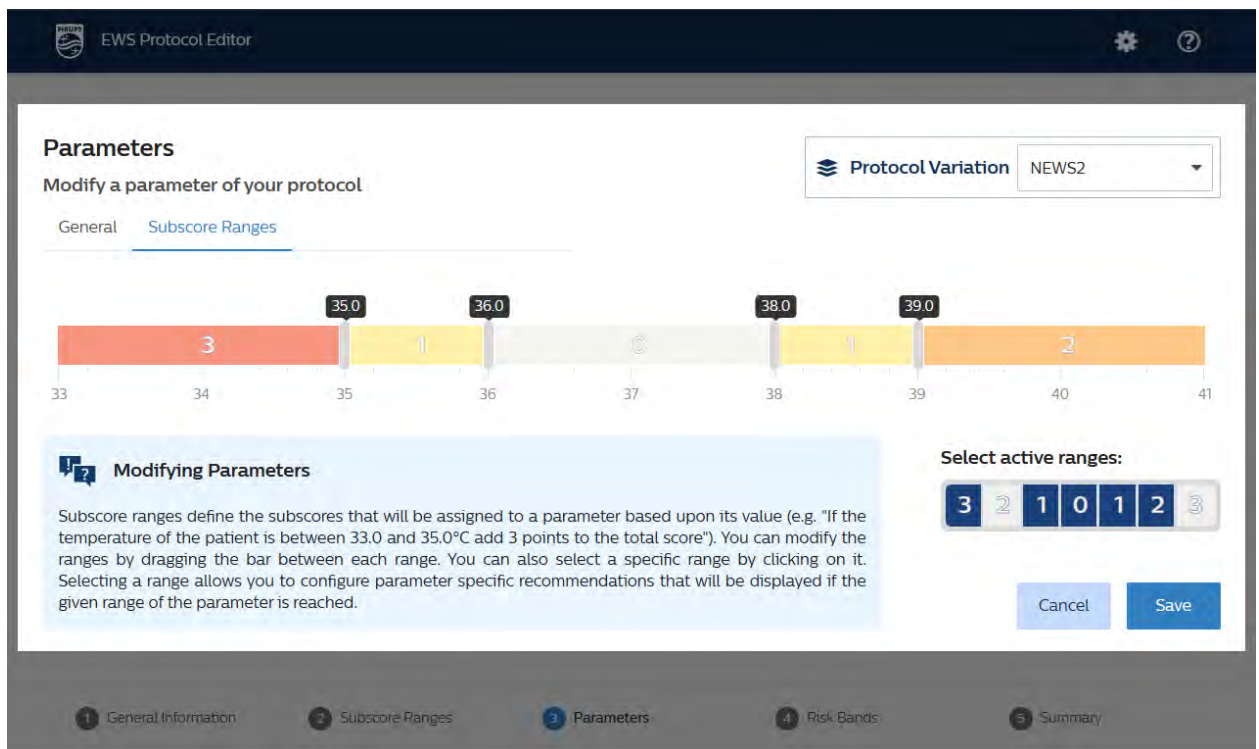


Abb. 1: Auszug aus der erstellten EWS-Editor Applikation [1]

Ergebnisse der Arbeit

Im Rahmen der Arbeit wurde ein vollständiger EWS-Editor konzeptioniert, implementiert und evaluiert. Zur Erreichung einer möglichst hohen Benutzungsfreundlichkeit wurden bei der Konzeptionierung relevante Designprinzipien und Best Practices angewandt. Es wurden neue Bedienkonzepte erarbeitet, die von Anwendern gut angenommen wurden. Anhand der durchgeführten Evaluation konnte der Prototyp pra-

xisnah untersucht und auf Mängel überprüft werden. Der Prototyp hat sich hierbei bereits, gemessen an den festgelegten Bewertungsmetriken, als gebrauchstauglich erwiesen. Die verschiedenen Schritte des Editors wurden durchweg als intuitiv verständlich und einfach zu bedienen wahrgenommen. Der erzielte *System Usability Score* von 86.75 Punkten liegt nach dem Bewertungsschema im Bereich "*Best Imaginable*" und spricht für eine benutzungsfreundliche Applikation.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Patrick Jordan et al. *Usability Evaluation In Industry*. CRC Press, 1996.
- [3] International Organization for Standardization. *Ergonomie der Mensch-System-Interaktion*. International Organization for Standardization, 2017.
- [4] Richard Thomson et al. *Safer care for the acutely ill patient: Learning from serious incidents*. National Patient Safety Agency, 5 edition, 2007.

Virtualisierung von GPUs: Einbindung in SLURM, Accounting und Energieeffizienz

Silas Mueller

Rainer Keller

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Einleitung

GPUs (Graphics Processing Unit) sind längst nicht mehr nur für Computerspiele und aufwendige 3D-Anwendungen wie CAD- oder Videoschnittprogramme interessant. Überall dort, wo viele parallele Berechnungen in kurzer Zeit durchgeführt werden müssen, kommen Grafikprozessoren immer häufiger zum Einsatz, da sie der CPU bei der parallelen Datenverarbeitung überlegen sind. Diese Vorteile machen sich besonders bei wissenschaftlichen Simulationen wie z.B. in der Strömungsmechanik bemerkbar, wo häufig numerische mathematische Methoden zum Einsatz kommen [4]. Weitere Einsatzgebiete von GPU Computing sind Big Data und Machine Learning, auch hier werden große Datenmengen analysiert, was durch die parallele Architektur einer GPU deutlich schneller und effizienter als mit der klassischen sequentiellen Arbeitsweise einer CPU umgesetzt werden kann. Mit dem Aufschwung der genannten Geschäftsfelder steigt auch die Nachfrage nach GPUs und deren Rechenleistung. Damit wird auch die Virtualisierung von Grafikkarten immer interessanter, um die von der Karte bereitgestellten Ressourcen möglichst effizient nutzen zu können, beispielsweise in HPC-Clustern (High Performance Computing) oder online bereitgestellten GPU-Servern.

Ziel der Bachelorarbeit

Die Bachelorarbeit ist in zwei Teile unterteilt, ein Teil der Arbeit ist zu verstehen, wie die Virtualisierung von NVIDIA GPUs funktioniert und wo virtuelle GPUs an ihre Grenzen kommen. Der Zweite Teil der Bachelorarbeit beschäftigt sich mit der Einbindung von (virtuellen) Grafikkarten in SLURM und der Messung des Stromverbrauchs eines Prozesses oder SLURM-Job der auf der Karte ausgeführt wird, um dem Benutzer Feedback über den Stromverbrauch seiner Berechnungen zu geben.

Anwendungsbeispiele für virtuelle GPUs

Virtualisierte GPUs, auch vGPUs genannt, haben viele verschiedene Anwendungsfälle. Der offensichtlichste Grund für die Virtualisierung ist die effizientere Nutzung der Ressourcen, da mehrere virtuelle Maschinen auf einer GPU ausgeführt werden können, um die Rechenleistung der Karte optimal zu nutzen. Durch den Einsatz von virtuellen Maschinen kann die zugewiesene Rechenleistung wie bei virtualisierten CPUs an das aktuell benötigte Pensum angepasst werden, um eine optimale Auslastung der Hardware zu gewährleisten. Durch die effizientere Nutzung der Ressourcen und die erhöhte Flexibilität können Kosten eingespart werden. Ein weiteres Anwendungsbeispiel für vGPUs sind virtuelle Desktops [2]. Diese werden in Zeiten von Home-Office und mobilem Arbeiten immer wichtiger, da über Technologien wie Remote Desktop oder ähnliche Tools von jedem Laptop aus auf einen virtuellen Desktop mit hoher Rechenleistung zugegriffen werden kann. Dies ist vor allem für CAD- und Videoschnittprogramme interessant, da nicht mehr jeder Arbeitsplatz eine eigene Grafikkarte benötigt, sondern Ressourcen geteilt werden können.

Virtualisierung von NVIDIA GPUs

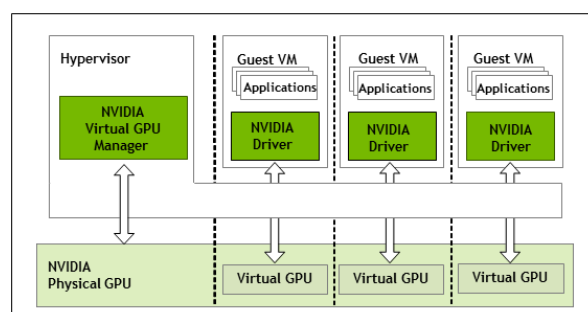


Abb. 1: Mehrere virtuelle Maschinen greifen auf eine physische Grafikkarte zu [3]

Für die Virtualisierung von NVIDIA GPUs werden neben der entsprechenden Hardware auch spezielle Treiber benötigt, die die Virtualisierung ermöglichen. Diese werden von NVIDIA kostenpflichtig zur Verfügung gestellt. Bei richtiger Konfiguration kann ein Hypervisor mit NVIDIA Virtual GPU Manager, wie z.B. VMWare vSphere, die Ressourcen einer physischen Karte auf beliebig viele virtuelle Maschinen aufteilen. Dies geschieht mit nur geringen Leistungseinbußen im Vergleich zur Bare-Metal-Leistung der Grafikkarte.

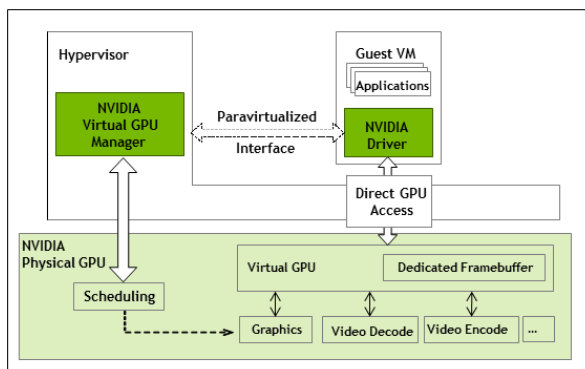


Abb. 2: NVIDIA vGPU Architektur [3]

Bei der Aufteilung der Ressourcen der Karte werden sowohl die Recheneinheiten als auch der Grafikspeicher aufgeteilt. Die spezialisierte Hardware der Karte, wie

z.B. Video Decoder und Encoder, wird zwischen den virtuellen Maschinen geteilt. So verfügt jede virtuelle GPU über einen Framebuffer fester Größe, der bei der Erstellung der virtuellen GPU auf der physischen Karte reserviert wird und erst beim Löschen der virtuellen GPU freigegeben wird, sowie über einen oder mehrere virtuelle Display-Ausgänge, auch Heads genannt. Auf jeder virtuellen Maschine, die über eine vGPU verfügt, können wie gewohnt die NVIDIA-Treiber installiert werden, die dann direkten Zugriff auf die GPU haben.

Messung des Stromverbrauchs mit DCGM

Mit Hilfe der Library DCGM (Data Center GPU Manager) ist es möglich, den Stromverbrauch eines Prozesses oder eines Jobs auf dem Grafikprozessor zu messen [1]. Um den Stromverbrauch eines Prozesses messen zu können, muss zunächst die GPU oder die Gruppe von GPUs, auf der der Prozess läuft, in den sogenannten Watch-Modus versetzt werden, wodurch DCGM verschiedene Daten, wie z.B. den Stromverbrauch des von der Karte ausgeführten Prozesses, messen kann. Ein nächster Schritt wäre die automatische Messung der Daten in SLURM bei jedem GPU-Job zu triggern, um dem Benutzer ein zuverlässiges Feedback zum Stromverbrauch seiner Berechnungen geben zu können.

Literatur und Abbildungen

- [1] NVIDIA Corporation. NVIDIA DCGM Documentation. <https://docs.nvidia.com/datacenter/dcgm/latest/>, 2023.
- [2] NVIDIA Corporation. NVIDIA Virtual GPUs multiplied. <https://www.nvidia.com/en-us/design-visualization/solutions/multi-virtual-gpus/>, 2023.
- [3] NVIDIA Corporation. Virtual GPU Software Documentation. <https://docs.nvidia.com/grid/10.0/grid-vgpu-user-guide/index.html>, 2023.
- [4] Michael Griebel and Peter Zaspel. A multi-GPU accelerated solver for the three-dimensional two-phase incompressible Navier-Stokes equations. https://ins.uni-bonn.de/media/public/publication-media/zaspel_gpu.pdf, 2010.

Entwicklung und prototypische Implementierung einer Cross-Media-Marketingstrategie für ein IT-Software-Systemhaus

Brikena Mushkolaj

Anke Bez

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma evia consulting GmbH, Stuttgart

Einleitung

Die Nutzung von Online- und Offline-Plattformen sowie die allgemeine Digitalisierung von Geschäftsprozessen zwingen Unternehmen dazu, ihre Marketingstrategien an aktuelle Trends und Entwicklungen anzupassen, um die eigene Wettbewerbsfähigkeit zu steigern. Dabei kann eine Cross-Media-Marketingstrategie helfen. Diese verknüpft verschiedene Kommunikationsträger, um die eigenen Produkte oder Dienstleistungen effektiv und effizient an die ausgewählte Zielgruppe zu vermarkten. Es muss ein roter Faden in der Kommunikation mit dem Kunden erkennbar sein und der Kunde sollte auf ein bestimmtes Zielmedium (z.B. Landingpage) geleitet werden. [3] Die strategischen Ziele können von Unternehmen zu Unternehmen unterschiedlich sein, umfassen aber in der Regel die Stärkung der eigenen Marke, die Gewinnung neuer Kunden und die Steigerung des Umsatzes.

Zielsetzung und Forschungsfrage

Das Ziel dieser Bachelorarbeit liegt darin, eine Cross-Media-Marketingstrategie für ein IT-Software-Systemhaus zu entwickeln und prototypisch umzusetzen. Dabei werden verschiedene Online-Medien aufeinander abgestimmt eingesetzt, um die relevante Zielgruppe an unterschiedlichen Touchpoints mit den wichtigsten Informationen zu versorgen. Es soll folgende Forschungsfrage beantwortet werden: Wie kann ein IT-Software-Systemhaus eine erfolgreiche Cross-Media-Marketingstrategie entwickeln und implementieren, um die Effektivität seiner Marketingaktivitäten zu steigern und seine B2B-Kunden auf verschiedenen Kanälen anzusprechen?

Begriffsabgrenzung

Häufig werden Multi-Channel-Marketing, Omni-Channel-Marketing und Marketing-Orchestrierung mit

Cross-Media-Marketing in Verbindung gebracht. Die Unterschiede liegen im Detail. Alle zielen darauf ab, viele verschiedene Marketingkanäle aufeinander abzustimmen, um eine konsistente Kundenerfahrung zu schaffen. Beim Multi-Channel-Marketing werden die Kanäle jedoch nicht miteinander verknüpft und somit unterschiedliche Zielgruppen über verschiedene Kanäle angesprochen. [6] Im Gegensatz dazu müssen beim Omni-Channel-Marketing alle Marketingkanäle nahtlos miteinander verbunden sein. [5] Bei der Marketing-Orchestrierung geht es um die Planung und Koordination aller Marketingaktivitäten, nicht nur der Marketingkanäle. [4]

Vorgehensweise

Zu Beginn wird eine Bestandsaufnahme der Kommunikationsträger und Tools des Unternehmens durchgeführt. Dabei werden bereits existierende und benötigte Marketingkanäle und -tools identifiziert. Anschließend wird eine SWOT-Analyse (siehe Abbildung 1) der Chancen, Risiken, Stärken und Schwächen durchgeführt, um das Geschäftsumfeld richtig einschätzen und gezielt auf Chancen und Risiken reagieren zu können. Die Markt- und Zielgruppenanalyse rundet die Strategie ab. Dabei entstehen Personas (siehe Abbildung 2), die Identifikation des relevanten Marktes und die verschiedenen Customer Journeys. Diese Ergebnisse werden genutzt, um die Strategieziele, die Rahmenbedingungen, den Cross-Media-Plan sowie die Auswahl der geeigneten Kommunikationsträger festzulegen. Als Nächstes wird die prototypische Umsetzung dokumentiert, indem gezeigt wird, wie die Marketing-Maßnahmen umgesetzt werden. Des Weiteren wird Google-Analytics eingesetzt, um die Erfolgsergebnisse der Strategie mithilfe verschiedener Kennzahlen darzustellen.

		POSITIVE	NEGATIVE
		STRENGTHS	WEAKNESSES
INTERNAL		<ul style="list-style-type: none"> - Qualifizierte Mitarbeiter - Hohes technologisches Know-How - Kurze Entscheidungswege - Keine Stakeholder/Investoren - Breites Angebotsspektrum - Hohe Kundenzufriedenheit 	<ul style="list-style-type: none"> - Keine starke Marke/Bekanntheit - Geringe Online-Präsenz - Geringes Budget für Marketing - Qualifiziertes Personal zu rekrutieren und zu halten
EXTERNAL		<ul style="list-style-type: none"> - KI Trends - Veränderung im Kundenverhalten auf Grund von COVID - Steigende Nachfrage nach IT-Dienstleistungen im Zuge der Digitalisierung 	<ul style="list-style-type: none"> - Hohe Konkurrenz - Schnellebigkeit des technologischen Fortschritts - Wirtschaftliche Instabilität durch Marktlage (Rezession) - Abwerbung von qualifizierten Mitarbeitern durch Headhunter

Abb. 1: SWOT-Analyse evia [1]

THE PERSONA PROFILE

Name Sarah Petersen



Target group/customers: Decider

A short version of the life-story: Sarah Petersen hat einen Bachelor-Abschluss in Informatik und hat in den letzten 10 Jahren in der IT-Branche gearbeitet. Sie hat in verschiedenen Rollen in verschiedenen Unternehmen gearbeitet und hat dabei umfangreiche Erfahrungen in der Entwicklung, Implementierung und Wartung von Anwendungen gesammelt. Nun arbeitet Sie als IT-Teamleiterin in einem Konzern. Sarah ist eine motivierte und engagierte Person, die gerne neue Herausforderungen annimmt und ihre Fähigkeiten ständig erweitert.

My skills

- Gute Kenntnisse in verschiedenen Programmiersprachen wie Java, Python und C#.
- Erfahrung in der Zusammenarbeit mit Teams und der Führung von Projekten.
- Ausgezeichnete Kommunikationsfähigkeiten
- Fähigkeit, komplexe Probleme zu identifizieren und effektive Lösungen zu entwickeln.
- Erfahrung in der Verwaltung von Budgets und Zeitplänen für IT-Projekte.

My dreams

- Sie träumt davon, eine erfolgreiche Führungskraft zu sein, die ihr Team inspiriert und motiviert.
- Sarah strebt danach, sich in der IT-Branche weiterzuentwickeln und ihr Wissen und ihre Fähigkeiten auszubauen.
- Sie möchte dazu beitragen, dass das Unternehmen innovative Technologien nutzt und auf dem neuesten Stand der Technik bleibt.

My rituals

- Sarah legt großen Wert auf regelmäßige Teammeetings und individuelle Gespräche mit ihren Mitarbeitern.
- Sie plant regelmäßig Zeit für die Überprüfung des Fortschritts und der Ergebnisse ihrer Projekte ein und sorgt dafür, dass sie innerhalb des Budgets und des Zeitplans bleiben.
- Sarah nutzt ihre Freizeit, um ihre Fähigkeiten zu verbessern, indem sie Online-Kurse und technische Bücher liest.
- Sie geht regelmäßig joggen, um Stress abzubauen und den Geist zu klären.

My relationship with technology

Sie nutzt Technologie, um Geschäftsprozesse zu optimieren und das Unternehmen wettbewerbsfähig zu halten. Gleichzeitig ist sie sich bewusst, dass Technologie nur ein Mittel zum Zweck ist und dass es darauf ankommt, sie effektiv zu nutzen, um die Ziele des Unternehmens zu erreichen.

Abb. 2: Persona "Decider-[2]

Ausblick auf zukünftige Entwicklung

Basierend auf den Erkenntnissen und Erfahrungen der Bachelorarbeit können mögliche Erweiterungen und zukünftige Forschungsfelder identifiziert werden. Beispielsweise können weitere, für das Unternehmen evia relevante, Strategien untersucht und verglichen werden. Auch könnte es interessant sein, spezielle Technologien wie Chatbots, AI oder personalisierte Landingpages in die Marketingstrategie zu integrieren,

um die Effizienz weiter zu steigern. Bestehende Kontakte können durch eine Lead-Nurturing-Kampagne in Form eines Newsletters weiter mit Informationen über andere Dienstleistungen versorgt werden. Es kann auch sinnvoll sein, Umfragen oder Interviews durchzuführen, um die Meinungen und Bedürfnisse der Kunden besser zu verstehen. Auf der Grundlage dieses Feedbacks können gezielte Anpassungen vorgenommen werden, um die Strategie zu optimieren.

Literatur und Abbildungen

- [1] Havard Business Review. HBR Tools: SWOT Analysis. *HBR Tools: SWOT Analysis*, page 4, 2015.
- [2] Eigene Darstellung.
- [3] Philipp Kolbeck. Werbeagentur in Nürnberg und Fürth - Lingner Marketing GmbH. https://www.lingner.de/dnl/crossmediale_kommunikation.pdf, 2023.
- [4] Uwe Seebacher. *Praxishandbuch B2B-Marketing: Neuste Konzepte, Strategien und Technologien sowie praxiserprobte Vorgehensmodelle - mit 11 Fallstudien*. Springer Gabler, 2021.
- [5] Venkatesh Shankar and Kushwaha Turan. Omnichannel Marketing: Are cross-channel effects symmetric? <https://www.sciencedirect.com/science/article/pii/S016781162030077X>, 06 2021.
- [6] Bernd Wirtz. *Multi-Channel-Marketing: Grundlagen - Instrumente - Prozesse*. Gabler Verlag, 2013.

Darstellung und Bewertung der Trends der Arbeitswelt

Maximilian Nikitsch

Catharina Kriegbaum-Kling

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

I. EINLEITUNG UND FRAGESTELLUNG

Die Arbeitswelt ist im stetigen Wandel, die Arbeitnehmer und Arbeitgeber müssen sich den Anforderungen der Zeit stellen und sich den Trends der Zukunft öffnen. Und sich mit diesen Trends auseinandersetzen. Sonst könnten die Unternehmen einen entscheidenden Wettbewerbsnachteil erhalten, da die Mitarbeiter zu den Unternehmen wechseln könnten, die diese Trends mitmachen und einführen. Um das zu verhindern muss ausgearbeitet und analysiert werden welche Trends sich immer weiter festsetzen. Welche Vorteile und

Nachteile sie bringen. Ob diese Trends eine Zukunft haben oder nur eine kurzfristige Erscheinung sind. Diese Auswertung darf keine Ausnahmen bilden, egal wie absurd sie klingen mögen. Auf diesen Weg, lässt sich erarbeiten welcher Trend in welchem Maße sinnvoll für das Unternehmen ist und welcher nicht.

II. ERARBEITUNG

In der Arbeitswelt gibt es unzählige Trends die sich festsetzen. Die bekanntesten wären die 5-Stunden-Regel oder die 4-Tage-Woche. Die analysierten Trends sind in der Abbildung [1] dargestellt.

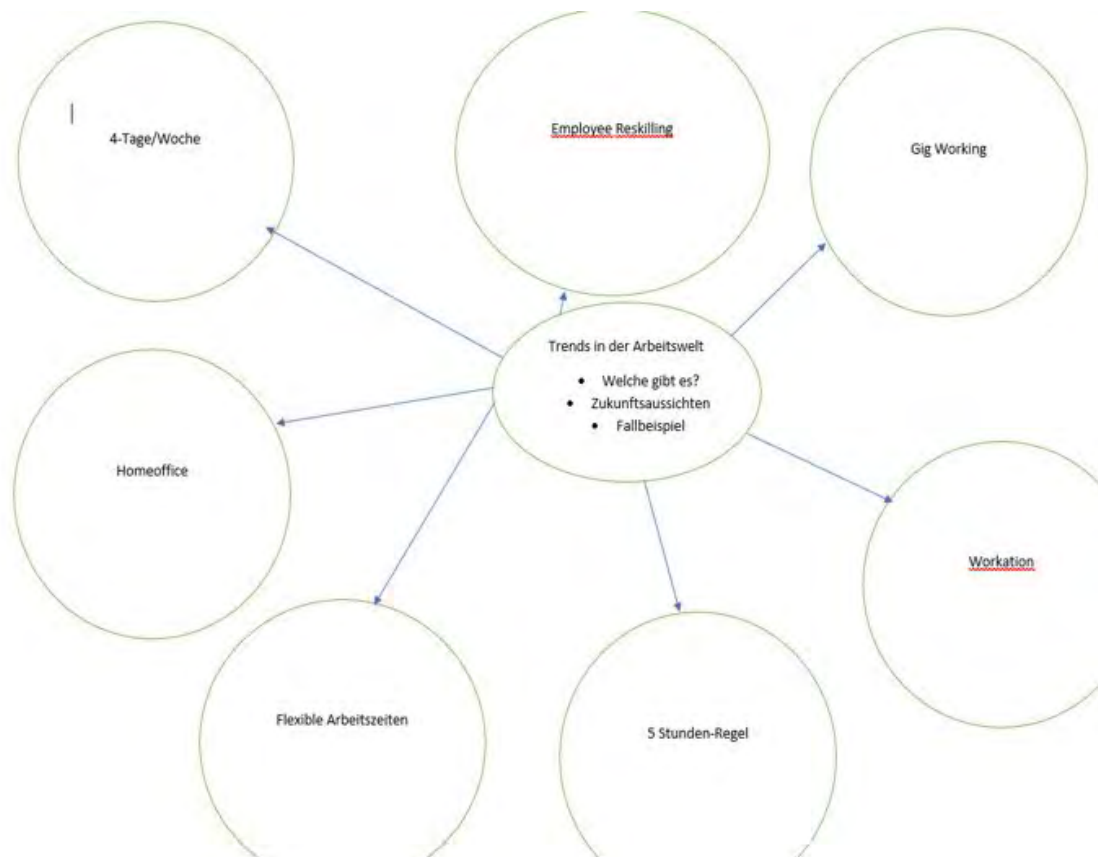


Abb. 1: Trends der Arbeitswelt [1]

Diese dargestellten Trends gehören zu den bekanntesten und wichtigsten. Als erstes muss der Trend beschrieben werden, damit gewusst wird worum es bei dem Trend überhaupt geht. Um den Trend vollständig zu verstehen darf man nicht nur den Trend beschreiben, es müssen auch Vorteile und Nachteile analysiert werden. Jeder Trend hat Vorteile und Nachteile. Als letzter Punkt wird die Zukunft des Trends betrachtet. Hat der Trend überhaupt eine Zukunft oder ist es eine modische Erscheinung? Wenn er eine Zukunft hat, wie sieht diese aus?

III. FALLBEISPIEL

Um ein Beispiel für die heutige Arbeitswelt zu bekommen, wird eine Umfrage an mehrere Unternehmen geschickt, das Ergebnis wird anschließend ausgewertet. Durch diese Umfrage lässt sich erkennen, wie weit sich deutsche Unternehmen auf neue Trends einlassen und sie umsetzen.

Durch diese Befragung, wird ebenfalls in Erscheinung gebracht, ob die Unternehmen die Trends kennen und was diese Unternehmen unter dem Trend verstehen. Bei diesem Punkt gibt es Ausnahmen, es wird erwartet, dass jedes Unternehmen mit den Trends des Homeoffice und der flexiblen Arbeitszeiten etwas anfangen kann. Deshalb wird bei diesen beiden Trends auf diese Frage verzichtet. In Abbildung [2] ist ein Beispiel zu sehen.

1. Was verstehen Sie unter der 4-Tage/Woche

Ihre Antwort eingeben

Abb. 2: Frage zum Verständnis [1]

Ebenso lässt sich erkennen wie viele Mitarbeiter diesen Trend nutzen. Dies wird im prozentualen Vergleich des Unternehmens abgefragt, in dem untersucht wird ob in dem Unternehmen mehr als 50% oder weniger als 50% mit dem Trend arbeiten. Oder alle beziehungsweise keiner.

Da diese Abfragung bei den Trends Gig Working und Employee Reskilling nicht geht. Aufgrund der Art des

Trends. Wird nach der Häufigkeit abgefragt. Also wie oft der Trend eingesetzt wird. So oft es geht, manchmal oder als aller letzter Ausweg.

Die letzten Fragen beziehen sich auf die Zukunft und das Einführen der Trends. Es wird gefragt, auf welche Trends die Corona- Pandemie einfluss hatte und auf welche Weise [3].

13. Hatte die Corona-Pandemie einen Einfluss auf einen der Trends, wenn ja bitte den betreffenden Trend ankreuzen (Mehrfach Wahl möglich)

- 4-Tage/Woche
- Employee Reskilling
- Gig Working
- Workation
- 5-Stunden-Regel
- Flexible Arbeitszeiten
- Homeoffice

14. Auf welche Weise wurden die Trends beeinflusst?

Ihre Antwort eingeben

Abb. 3: Corona und die Trends [1]

Zum Abschluss wird noch nach den Gründen für das erfolgreiche Einführen eines Trends oder nach den Gründen für das Scheitern gefragt [4].

15. Nennen Sie die Hauptgründe für das einführen der Trends

Ihre Antwort eingeben

16. Nennen Sie die Hauptgründe für das scheitern/nicht einführen der Trends

Ihre Antwort eingeben

Abb. 4: Gründe [1]

Am Schluss kann noch ein Vergleich zwischen den Unternehmen erstellt werden.

Literatur und Abbildungen

[1] Eigene Darstellung.

Integration einer datenbasierten Schnittstelle zwischen SAP SuccessFactors und Payroll-Systemen in die bestehende IT-Landschaft der Alfred Kärcher SE & Co. KG - Eine Potenzialanalyse und Handlungsempfehlung

Fawad Rasuly

Thomas Rodach

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Alfred Kärcher SE & Co. KG, Winnenden

Einleitung

Die Digitalisierung hat auch das Personalmanagement stark verändert. Die klassische Personalabteilung hat sich zu einem modernen Management weiterentwickelt, das digitale Technologien nutzt. Eine entscheidende Rolle spielen dabei die in der Vergangenheit lokal im Unternehmen installierten Personal-Informationssysteme. Mit dem Aufkommen von Cloud-Technologien eröffnen sich jedoch neue Möglichkeiten. Auch wenn sich HR lange gegen den Einsatz von Cloud Computing gewehrt hat, ist dieser Weg nun unumkehrbar. [3] In Unternehmen mit vielen Systemen und Schnittstellen, die alle miteinander kommunizieren müssen, kann die Anbindung der verschiedenen HR-IT-Schnittstellen problematisch werden.

Ziel

Das Ziel ist es zu definieren, welche Schnittstelle in welchem Fall am besten geeignet ist, damit SAP SuccessFactors als Personaldatensystem mit beispielsweise der Payroll oder dem Zeiterfassungssystem der verschiedenen Tochtergesellschaften weltweit integriert werden kann.

SuccessFactors

SAP SuccessFactors-Lösungen sind HCM-Softwareanwendungen in der Cloud, die Kern-HR und Payroll, Talentmanagement, Personalplanung und -analyse sowie Employee Experience Management unterstützen. Die Suite, jetzt SAP SuccessFactors HXM Suite genannt, bietet personalisierte Erlebnisse, die Mitarbeiter glücklich, produktiv und zufrieden stellen. SuccessFactors-Lösungen wurden 2001 eingeführt und sind seit 2012 Teil der SAP-Familie. Sie werden von mehr als 242 Millionen Benutzern in

mehr als 200 Ländern und Regionen auf der ganzen Welt verwendet. [4]

Employee Central

Kernmodul von SuccessFactors ist Employee Central. Es ist eine Softwarelösung zur Verwaltung des gesamten Lebenszyklus von Mitarbeitern, von der Einstellung bis hin zur Kündigung. Es wurde entwickelt, um Unternehmen dabei zu helfen, HR-Prozesse zu optimieren, indem es eine zentralisierte Mitarbeiterdatenbank bereitstellt und verschiedene Funktionen zur Verbesserung des HR-Managements bietet. In der nachfolgenden Abbildung sind die Module von SuccessFactors zu sehen, wobei in der Firma Kärcher noch nicht alle Module global ausgerollt wurden. Damit ist gemeint, dass nur die Module, welche farblich hervorgehoben sind, bereits in allen Gesellschaften der Firma Kärcher im Einsatz sind.



Abb. 1: Welche Module von SF live sind [1]

Schnittstellenarten

Um Employee Central als führendes HR-System zu nutzen, müssen die lokalen Abrechnungssysteme der Tochtergesellschaften der Firma Kärcher mit Daten aus Employee Central versorgt werden. Damit dies möglich gemacht wird, können die lokalen Abrechnungssysteme mit Schnittstellen angebunden werden. Vorteil dieser Vorgehensweise ist unter anderem eine erhöhte Datenkonsistenz, welche zuvor durch aufwändiger Doppelpflege nicht in gleicher Höhe gewährleistet werden konnte. Abgesehen von der mühsamen Doppelpflege, auch Export/Import genannt, gibt es die Möglichkeiten der individuellen Programmierung einer Schnittstelle, der Nutzung des Integration Centers, welches von SAP SuccessFactors selbst angeboten wird und der Cloud Platform Integration, welche ebenfalls von SAP, aber auch von 3. Anbietern angeboten wird. In der folgenden Abbildung wird die einfachste Form, der File Transfer visualisiert, in welcher die Application A in unserem Fall Employee Central die aktuellsten Daten exportiert und diese dann in die lokalen Abrechnungssysteme importiert werden.



Abb. 2: Export/Import [2]

Es ist zudem möglich, mehrere Gesellschaften regional oder sogar global abzurechnen, womit viele Schnittstellen gespart werden können und Employee Central mit wenigen oder nur einem globalen Abrechnungssystem integriert wird. Das Prinzip der shared database kann in diesem Fall zur Veranschaulichung genutzt werden, wobei die shared database Employee Central darstellt und wenige Schnittstellen die Regionen A, B und C bedienen. Die Verantwortung für die Schnittstellen zwischen den Regionen und dem globalen Abrechner kann zudem der globale Abrechner übernehmen.

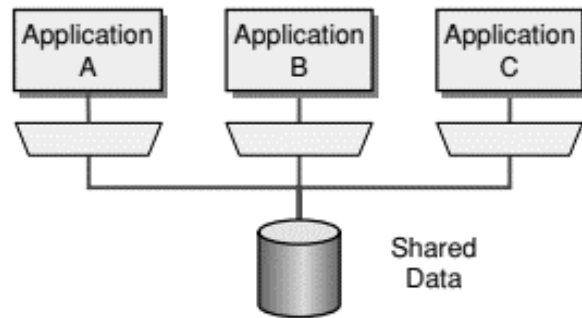


Abb. 3: shared services [2]

Kriterien zur Wahl einer Schnittstelle

Paarweise Vergleich der definierten Kriterien, um den Kriterien jeweils eine Gewichtung zu geben: Und die daraus resultierende Nutzwertanalyse, in der die Kriterien mit den jeweiligen Schnittstellenarten zusammengerechnet werden, wobei die Relevanz der Kriterien anhand vordefinierter Fragen festgelegt werden, um verschiedenste Szenarien darstellen zu können. Die mit der IT erstellten Nutzwertanalyse lässt sich in der Bachelorarbeit unter Kapitel 3 finden.

Praxisteil - Qualitative Forschungsmethode

Die Bachelorarbeit umfasst über 15 Experteninterviews/Gruppendiskussionen, welche anonymisiert ausgewertet und in Kapitel 4 in geclusterten Kategorien zusammengefasst und interpretiert werden. Daher sind diese nur in der Bachelorarbeit enthalten und lassen sich im Detail als Transkripte im Anhang durchlesen. Es wurde mit über 20 Experten von 12 unterschiedlichsten Unternehmen über das Thema der Arbeit diskutiert, wobei Eindrücke gegeben wurden, wie die Unternehmen in den verschiedensten Situationen mit der Problematik umgegangen sind und welche Erkenntnisse diese daraus gezogen haben.

Ausblick

Die Strategien bauen auf die erarbeiteten Kenntnisse über die verschiedenen Schnittstellentypen und die relevanten Kriterien zusammen mit den Erkenntnissen der zahlreichen Experten. Jede Strategie hat ihre Vor- und Nachteile und diese gilt es in jedem Fall abzuwägen. In meinem Ausblick schreibe ich der Firma Kärcher meine persönliche Handlungsempfehlung aus und werde diese mit weiteren möglichen Strategien vergleichen, wobei ich meine Wahl für meine Handlungsempfehlung begründen werde.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Gregor Hohpe and Bobby Woolf. *Enterprise Integration Patterns*. Addison-Wesley, 2011.
- [3] Ulli Pesch. Digitalisierung – die Transformation von HR. <https://www.personalwirtschaft.de/news/hr-organisation/digitalisierung-die-transformation-von-hr-96136/>, 10 2021.
- [4] Seite SAP. HR-Software und Personalmanagement. <https://www.sap.com/germany/products/hcm.html>, 01 2023.

Untersuchung einer Schnittstelle zwischen PyScript und Angular

Cornelius Rost

Jürgen Koch

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma adesso SE, Stuttgart

Einleitung

Das Wachstum des Internets und die steigende Nachfrage nach Webanwendungen bringen ständig neue Herausforderungen mit sich. Ursprünglich gab es statische Webseiten, die auf den Grundlagen von HTML und CSS erstellt wurden. Mit der Einführung von JavaScript wurde es möglich, dynamische und interaktive Webseiten zu entwickeln. Im Laufe der Zeit entstanden Web-Entwicklung-Frameworks. Diese Frameworks bieten Entwicklern einen strukturierten Ansatz für die Entwicklung komplexer Anwendungen. Eines der bekanntesten und beliebtesten Web-Entwicklung-Frameworks ist Angular. Trotz aller Fortschritte gibt es aber immer noch Anwendungsfälle, für die JavaScript nicht ideal ist. In solchen Szenarien wäre es von Vorteil, eine Programmiersprache zu verwenden, die besser auf den spezifischen Anwendungsfall zugeschnitten ist, um sowohl die Entwicklungseffizienz als auch die Benutzererfahrung zu optimieren. Ein Ansatz zur Lösung bestimmter Anwendungsfälle ist die Verwendung von WebAssembly (WASM). WebAssembly erweitert die Fähigkeiten des Browsers, indem es die Ausführung von Code ermöglicht, der in einer Vielzahl von Sprachen geschrieben werden kann, nicht nur in JavaScript [1]. Mit dem Aufkommen von WASM und anderen modernen Webtechnologien wurde das Framework PyScript entwickelt. Diese Entwicklung ermöglicht die direkte Ausführung von Python im Browser und eröffnet damit neue Möglichkeiten innerhalb von Webanwendungen [5].

PyScript

Python ist einer der beliebtesten und eine der am weitesten verbreiteten Programmiersprachen [4]. Durch die intuitive Syntax, die Plattformunabhängigkeit und die große Auswahl an Bibliotheken ist Python vor allem in den Bereichen Data Science und Machine Learning sehr beliebt. Das Open-Source Framework PyScript wurde 2022 vorgestellt und von Anaconda entwickelt. PyScript ist eine TypeScript-Bibliothek

und ermöglicht die Ausführung von Python-Skripten im Webbrowser, die normalerweise auf einem Server ausgeführt werden. Damit ist es nun möglich, Python in der Frontend-Entwicklung einzusetzen und auf eine Vielzahl bekannter Python-Bibliotheken zuzugreifen. Dabei soll PyScript JavaScript nicht ersetzen, sondern mit JavaScript koexistieren. Daher verspricht PyScript eine bidirektionale Kommunikation zwischen PyScript und JavaScript [5]. Abbildung 1 verdeutlicht den Aufbau von PyScript sowie die zugrunde liegenden Technologien. PyScript basiert auf Pyodide, einer Python-Distribution für die CPython-Implementierung (CPython ist die Standardimplementierung) von Python, die speziell für den Einsatz im Browser entwickelt wurde. Um CPython im Browser ausführen zu können, kommt der Compiler Emscripten zum Einsatz, der C- und C++-Code in WebAssembly umwandelt [2]. Da CPython in C geschrieben ist, kann es mithilfe von Emscripten in WebAssembly kompiliert werden. In WebAssembly kann nun der kompilierte Code, der im Binärcode Format vorliegt, ausgeführt werden. PyScript und die zugrunde liegenden Technologien ermöglichen somit die Ausführung von Python im Browser.

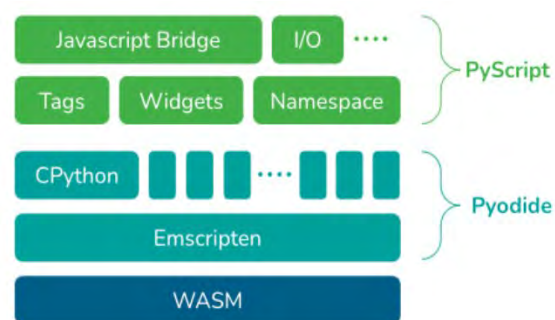


Abb. 1: PyScript Techstack [8]

Angular

Angular wurde von Google entwickelt und ist ein client-seitiges Open-Source Framework für die Entwicklung von Single-Page-Anwendungen. Es ist in TypeScript geschrieben, das in JavaScript kompiliert wird. Angular bietet eine Reihe von Eigenschaften, die es Entwicklern ermöglichen, komplexe Anwendungen zu erstellen. Ein Vorteil von Angular ist, dass der Quellcode sehr modular aufgebaut ist. Eine Angular-Anwendung besteht aus mehreren Komponenten, die ineinander verschachtelt werden können und dabei miteinander kommunizieren können. Neben Komponenten gibt es auch Services. Services sind für die Kommunikation mit dem Server (Backend) zuständig und stellen somit die Daten der Anwendung bereit. Zusätzlich bietet Angular die Möglichkeit, Services und Komponenten zu Modulen zusammenzufassen, wodurch einzelne Features noch modularer aufgebaut werden können [6]. Abbildung 2 stellt den Aufbau eines Angular Moduls dar.

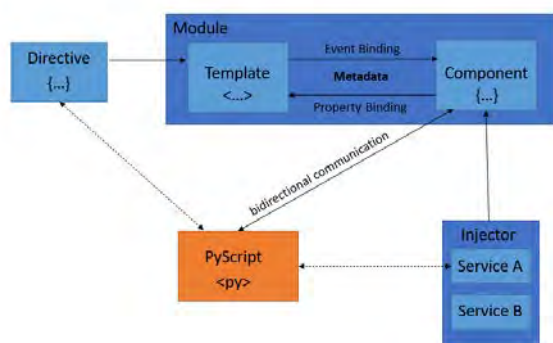


Abb. 2: Angular und PyScript [3]

Ziele der Arbeit

Ziel dieser Bachelorarbeit ist es, zu untersuchen, wie PyScript zusammen mit Angular eingesetzt werden kann. Dazu werden zunächst die beiden Technologien untersucht. Anschließend wird anhand der Eigenschaften von Angular praktisch gezeigt, wie PyScript sinnvoll in das Angular Ökosystem integriert werden kann. Durch die gewonnenen Erkenntnisse soll ein umfangreiches Beispiel entwickelt werden, das die Möglichkeiten von PyScript und Angular aufzeigt. Anhand der Untersuchungen soll auch abgeschätzt werden, was PyScript für die Zukunft der Webentwicklung bedeuten könnte.

Ergebnisse und Ausblick

Zuerst wurden PyScript und Angular getrennt untersucht. Anschließend wurde untersucht, ob PyScript in Angular integriert werden kann. Die Ergebnisse zeigen, dass eine Integration von PyScript in Angular möglich ist. Allerdings erfordert die Einbindung ein dynamisches Laden von PyScript innerhalb einer Komponente oder eines Moduls. Moderne Frameworks wie Angular erlauben aus Sicherheitsgründen keine Skript-Tags innerhalb der HTML-Datei. Zusätzlich ist es notwendig, die Verwendung von benutzerdefinierten HTML-Tags innerhalb eines Moduls zuzulassen. Dies stellt sicher, dass PyScript-Tags und damit Python-Skripte ausgeführt werden können.

Der Schwerpunkt der Untersuchung lag bisher auf der Kommunikation. Während PyScript eine bidirektionale Kommunikation zwischen JavaScript und PyScript verspricht, stellt sich die Frage, wie die Kommunikation und der Zugriff zwischen PyScript und Angular-Komponenten aussieht. Innerhalb von Komponenten kann auf PyScript-Objekte und -Methoden über das Window-Objekt zugegriffen werden. PyScript bietet die Möglichkeit, JavaScript-Objekte direkt zu importieren. Dieser Importvorgang erfolgt ebenfalls über das Window-Objekt. Angular arbeitet jedoch mit Modulen, die mit ES6 eingeführt wurden und einen eigenen isolierten Namensraum pro Modul versprechen, dadurch wird nicht mehr der globale Scope überflüssig überladen. Daher müssen Angular-Komponenten zunächst in den globalen Scope gebracht werden, um in PyScript importiert werden zu können. Die Darstellung in Abbildung 2 veranschaulicht die Kommunikation zwischen einer Komponente und PyScript. Ebenfalls denkbar ist eine Kommunikation zwischen Services/Direktiven und PyScript.

Im Rahmen der Kommunikation muss zusätzlich beachtet werden, wie verschiedene Typen übersetzt werden. Man unterscheidet zwischen "implicit converting" und "proxying". Implicit converting bezieht sich auf unveränderliche Typen und erstellt ein neues Objekt in der Zielsprache, das äquivalent zu dem Objekt der Ursprungssprache ist. Proxying bezieht sich auf veränderliche Typen und erstellt ein spezielles Objekt in der Zielsprache, das Anfragen an die Ursprungssprache weiterleitet [7]. Durch einen solchen Proxy kann beispielsweise innerhalb von PyScript auf eine Angular-Komponente und deren Membervariablen und Membermethoden zugegriffen werden.

Die bisherigen und weiteren Erkenntnisse werden in Form einer Webanwendung dargestellt, die einem Wiki ähnelt und mit Codebeispielen präsentiert werden. Darüber hinaus ist geplant, nach Abschluss der Untersuchungen ein umfassendes Beispiel zu entwickeln, das die Möglichkeiten von Angular und PyScript verdeutlicht.

Literatur und Abbildungen

- [1] Bytecode Alliance. WebAssembly Dokumentation. <https://webassembly.org>, 2023.
- [2] Emscripten Contributors. emscripten. <https://emscripten.org/index.html>, 2023.
- [3] Eigene Darstellung.
- [4] Statista Research Department. Die beliebtesten Programmiersprachen weltweit laut PYPL-Index im März 2023. <https://de.statista.com/statistik/daten/studie/678732/umfrage/beliebteste-programmiersprachen-weltweit-laut-pypl-index/>, 2023.
- [5] Anaconda Inc. PyScript Dokumentation. <https://docs.pyscript.net/latest/>, 2023.
- [6] Google LLC. Angular Dokumentation. <https://angular.io/docs>, 2023.
- [7] Pyodide contributors and Mozilla. What is Pyodide? <https://pyodide.org/en/stable/index.html>, 2023.
- [8] Sophia Yang. PyScript: Python in the Browser. <https://anaconda.cloud/pyscript-python-in-the-browser>, 2023.

Unified UX - Analyse und Optimierung der produktübergreifenden User Experience am Beispiel des Produktportfolios der USU Software AG

Moritz Ruider

Astrid Beck

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma USU Software AG, Möglingen

User Experience als unausgeschöpftes Potenzial in der Softwareindustrie

Obwohl heutzutage die Bedeutung der User Experience (UX) allgemein anerkannt wird und viele Unternehmen diese als wertvoll erachten, besteht dennoch das Problem, dass das Nutzererlebnis nicht ausreichend fundiert in Softwareentwicklungs- und Gestaltungsprozessen der Industrie verfolgt wird [9]. Unternehmen erkennen die Vorteile einer positiven UX wie zufriedene Kunden, höhere Benutzerakzeptanz und eine bessere Marktpositionierung. Jedoch kann beobachtet werden, dass UX meistens nur oberflächlich in Produktentwicklungsprozesse integriert wird.

Die Entwicklung von kohärenten und qualitativ hochwertigen Systemen erfordert neben der Betrachtung der reinen Funktionsweise aber auch die Berücksichtigung einer Vielzahl weiterer Merkmale [3]. Dazu zählen beispielsweise die Nutzererfahrungen, Fähigkeiten, Bedürfnisse, das Verhalten und kognitive Eigenschaften [2]. Auch ist allgemein anerkannt, dass eine fundierte UX-Strategie und -Umsetzung langfristig zu einem Wettbewerbsvorteil führen kann [1]. So kann die Beachtung und Erfüllung grundlegender Nutzerbedürfnisse für den langfristigen Erfolg von technologischen Produkten entscheidend sein. Produkte, die viele positive Erfahrungen gemacht haben, werden häufiger genutzt und daher bevorzugt [13].

Viele Unternehmen investieren zwar in UX-Design oder verbessern einzelne Benutzeroberflächen, vernachlässigen jedoch die ganzheitliche Betrachtung der Nutzererfahrung über alle Touchpoints hinweg. Dies kann dazu führen, dass potenzielle Probleme oder Verbesserungspotenziale übersehen werden, die letztendlich die Effektivität und Effizienz der Produkte oder Dienstleistungen beeinträchtigen.

Problemstellung & Zielsetzung

Einleitend wurde aufgezeigt, inwiefern die Wissenschaft zunehmend die Relevanz von User Experience für die Wettbewerbsfähigkeit und Produktakzeptanz betont – die theoretischen Erkenntnisse, User Experience verstärkt bei Produktentwicklungsprozessen in den Fokus zu rücken, müssen auch in praxisnahen Kontexten Anwendung finden. Im Zuge der Bachelorarbeit soll versucht werden, diesen Ansatz anhand des Produktportfolios der USU Software AG umzusetzen.

Das Angebot der USU Software AG umfasst im Allgemeinen Software und Dienstleistungen für das IT- und Customer-Service-Management. Der Fokus des Produktportfolios liegt auf der Digitalisierung und Automatisierung von Prozessen, die von der Service-Planung über die Steuerung und Nutzung bis hin zur Verrechnung des Service reichen [12]. Aufgrund der Vielseitigkeit des weiterhin wachsenden USU Produktportfolios und der für jedes Produkt separat ablaufenden Softwareentwicklungsprozesse besteht für existierende und zukünftige Produktservices des Unternehmens zunehmend Bedarf der Vereinheitlichung. Es besteht unter anderem der Anspruch, die Zugehörigkeit zwischen dem Unternehmen und den Produkten ganzheitlich zu verstärken. Dabei gilt es als besonders wertvoll, neben der visuellen Anmutung (UI) und Funktionalität (Usability) auch das Nutzererlebnis (UX) in den Fokus der Optimierungsprozesse zu rücken. Der holistische Ansatz für die Vereinheitlichung der Produktzusammengehörigkeit umfasst zu einem großen Teil die produktübergreifende User Experience. Hierbei müssen die Grenzen und Möglichkeiten der "Unified UX" analysiert werden. Die ausführliche Analyse des Status quo der User Experience und den resultierenden Optimierungsmöglichkeiten beschreiben das Ziel der Bachelorthesis. Dies gilt es mithilfe von Nutzer:innen im bestehenden Nutzungskontext und unter Berücksichtigung von UI und Usability zu untersuchen.

Überblick über die Zusammenhänge der menschenzentrierten Gestaltung

Der Ansatz der menschenzentrierten Gestaltung versucht, Verbraucher innen auf einer individuellen Ebene anzusprechen. Neben dem Berücksichtigen menschlicher Faktoren (wie Ergonomie, Kognition, Psychologie) bei der Gestaltung von Produkten werden Gestal-

ter innen dazu ermutigt, sich auch auf menschliches Wohlbefinden, Benutzerzufriedenheit, Zugänglichkeit und Nachhaltigkeit zu konzentrieren [6], [10]. Die menschenzentrierte Gestaltung legt den Fokus auf die Bedürfnisse und Anforderungen der Nutzer:innen, um Produkte und Systeme optimal auf ihre Bedürfnisse abzustimmen (Abbildung 1).

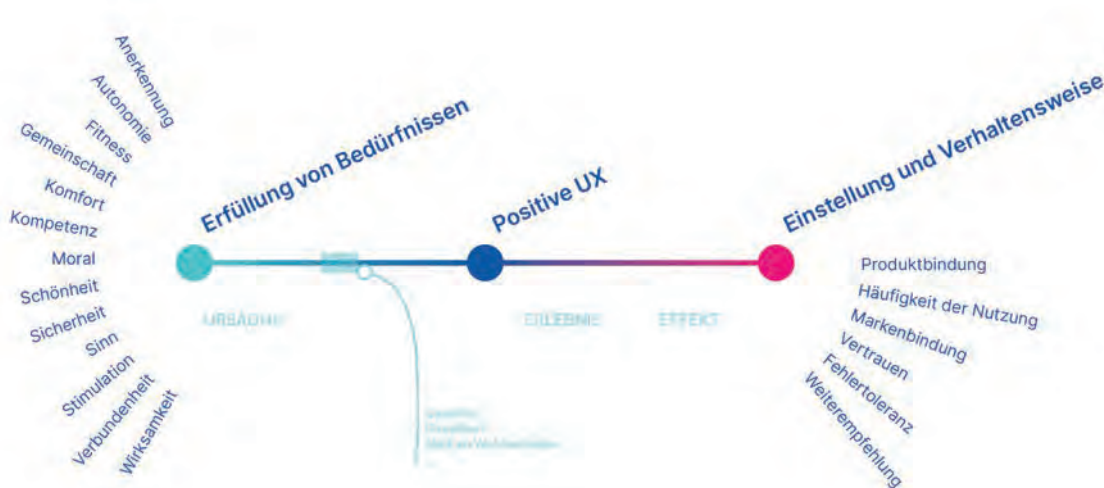


Abb. 1: Wirkungsverhältnisse bei positiver UX [4]

Ein wichtiger Aspekt dieser Gestaltungsphilosophie ist die Berücksichtigung der User Experience.

Durch einen menschenzentrierten Gestaltungsansatz können Lösungen entwickelt werden, die eine emotionale Bindung zu Nutzer:innen herstellen. Dabei werden wissenschaftliche Erkenntnisse über die physischen, wahrnehmungsbezogenen, kognitiven und emotionalen Eigenschaften des Menschen beachtet, um dieses Ziel zu erreichen [8].

Das Deutsche Institut für Normung (DIN) definiert in verschiedenen Normen einen internationalen Standard, der sich mit der Ergonomie der Mensch-Computer-Interaktion beschäftigt [5] [6] [7]. Zudem wird hier die User Experience als zentraler Bestandteil adressiert. Die Norm gibt vor, welche Anforderungen und Empfehlungen für die Gestaltung von interaktiven Systemen gegeben sein müssen, um die Benutzerfreundlichkeit, Effektivität und Zufriedenheit der nutzenden Personen zu verbessern. Des Weiteren werden grundlegende Kriterien in Bezug auf Individualisierbarkeit, Aufgabenmanagement und Dialoggestaltung festgelegt [6]. User Experience wird in der Norm als Wahrnehmung und Reaktion einer Person definiert, die sich aus

der tatsächlichen und/oder erwarteten Nutzung eines Produkts ergibt. Während User Experience das Erlebnis vor, während und nach der Nutzung beschreibt, ist eine angemessene Usability nur in der Phase der aktiven Nutzung erfolgsentscheidend, um eine effektive, effiziente und zufriedenstellende Nutzung zu gewährleisten [5]. Damit wird Usability zu einem integralen Bestandteil der User Experience. Allerdings zeigt diese Standardisierung auch, dass es für eine Analyse der User Experience nicht ausreicht, ausschließlich Aspekte der Usability eines Produktes zu betrachten [6].

Bedeutung des menschenzentrierten Gestaltungsprozesses für unified UX

Der Prozess in der menschenzentrierten Gestaltung wird abermals durch das Deutsche Institut für Normung in einen standardisierten Rahmen gesetzt. Der Prozess basiert auf sechs Schritten, die iterativ wiederholt werden können (Abbildung 2). Es soll sichergestellt werden, dass bereits bei den ersten Betrachtungen die Bedürfnisse, Anforderungen und Erfahrungen der Nutzer:innen von Anfang an berücksichtigt werden.



Abb. 2: Ablauf des menschlichen Gestaltungsprozesses [6]

In der Abschlussarbeit wird besonders der Fokus auf die ersten Phasen des Gestaltungsprozesses gelegt. Untersuchungsgegenstand sind fünf Produkte aus dem USU Produktportfolio, die zukünftige Fokusprodukte der USU im IT-Servicebereich darstellen. Diese gilt es im Hinblick auf den Nutzungskontext zu analysieren. Außerdem kann mithilfe qualitativer Methoden die Nutzungsanforderungen von realen Nutzer:innen und Stakeholdern definiert werden, die Gestaltungslösungen als Resultat erbringen können und somit zur Optimierung der Einheitlichkeit der UX beitragen. Aufgrund der qualitativen Vorgehensweise wird eine hohe Inhaltsvalidität und ein tiefer Informationsgehalt erwartet. Qualitative Methoden eignen sich insbesondere in den frühen Phasen des menschenzentrierten Gestaltungsprozesses, weil Hintergründe erfragt und Unklarheiten geklärt werden können, womit sich die Möglichkeit eröffnet, neue und bisher unbekannte Sachverhalte zu entdecken [11].

Ausblick

Die Herausforderung, ein konsistentes und ganzheitliches Benutzererlebnis zu schaffen, liegt darin, Design und Funktionalität über verschiedene Plattformen

und Produkte hinweg effektiv zu koordinieren. Im Zuge der Bachelorthesis soll ein einheitliches Benutzererlebnis erarbeitet werden, das auf einer sorgfältigen und tiefgründigen Analyse der Nutzer:innen und des Nutzungskontextes basiert. Durch die qualitative Forschungsmethodik soll außerdem ein Verständnis geschaffen werden, wie Nutzende mit den verschiedenen Produkten interagieren und welches produktübergreifende Nutzungsverhalten existiert. Basierend auf diesen Kenntnissen soll eine Optimierung der User Experience gewährleistet werden.

Diese Verbesserungen können in Form von Oberflächenanpassungen und einer generellen Überarbeitung von Arbeitsabläufen für holistische Anwendungsfälle auftreten. Bestandteil der Arbeit ist zudem, Marktforschung durchzuführen, um einen Überblick über die Wettbewerbssituation zu erhalten. Durch diese Analyse könnte die eigene Positionierung und Ausrichtung des Produktportfolios verstanden und gegebenenfalls verbessert werden. Zukünftig gilt es, die in der Bachelorthesis erarbeiteten Hypothesen und Konzepte wiederholt durch Nutzerstudien zu evaluieren und in Iterationen durch prototypische Ausarbeitungen zu verfeinern und zu verbessern.

Literatur und Abbildungen

- [1] A. Ant Ozok and Gavriel Salvendy. Measuring consistency of web page design and its effects on performance and satisfaction. *Ergonomics*, 2000.
- [2] M. Hanefi Calp and M. Ali Akcayol. The importance of human-computer interaction in the development process of software projects. *Global Journal of Information Technology*, 2015.
- [3] Lawrence Chung et al. *Non-Functional Requirements in Software Engineering*. Springer US, 2000.
- [4] Eigene Darstellung.
- [5] Deutsches Institut für Normung. Ergonomie der Mensch-System-Interaktion - Teil 11: Gebrauchstauglichkeit: Begriffe und Konzepte. In *DIN EN ISO 9241-11: 2018-11*. Beuth Verlag GmbH, 2018.
- [6] Deutsches Institut für Normung. Ergonomie der Mensch-System-Interaktion – Teil 210: Menschzentrierte Gestaltung interaktiver Systeme. In *DIN EN ISO 9241-210: 2020-03*. Beuth Verlag GmbH, 2020.
- [7] Deutsches Institut für Normung. Ergonomie der Mensch-System-Interaktion –Teil 110: Interaktionsprinzipien. *DIN EN ISO 9241-110: 2020-10*, 2020.
- [8] Joseph Giacomini. What Is Human Centred Design? *The Design Journal*, 2014.
- [9] Johanna Amalia Robinson. End-User Feedback on a Low-Cost Portable Air Quality Sensor System—Are We There Yet? *Sensors*, 2018.
- [10] Paul Russell and Lyndon Buck. Humanity-Centered Design – Defining the Emerging Paradigm in Design Education and Practice. In *Proceedings of the 22nd International Conference on Engineering and Product Design Education (E&PDE 2020)*. The Design Society, 2020.
- [11] Government Digital Service. Using in-depth interviews - Service Manual - GOV.UK. <https://www.gov.uk/service-manual/user-research/using-in-depth-interviews>, 2017.
- [12] USU Software AG. Ueber uns. <https://www.usu.com/de-de/unternehmen/ueber-uns/>, 2023.
- [13] Andreas Sonnleitner et al. *Experimentally Manipulating Positive User Experience Based on the Fulfilment of User Needs*, volume 8120. Springer Berlin Heidelberg, 2013.

Konzeption und Umsetzung von CI/CD-Praktiken im Kontext einer webbasierten Anwendung

Julian Saile

MarkusENZweiler

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma JW Froehlich Maschinenfabrik GmbH, Leinfelden-Echterdingen

Motivation

Wenn Module einer Software unabhängig voneinander entwickelt und erst spät im Entwicklungsprozess für ein Release zusammengeführt werden, besteht die Gefahr, dass sich im Laufe der Zeit unbemerkt Fehler und Bugs ansammeln, da diese nicht frühzeitig erkannt und korrigiert werden. Dadurch kann die Integration zu einem langwierigen und unvorhersehbaren Prozess mit erhöhter Komplexität werden.

Weitere Probleme können bei der Bereitstellung moderner Anwendungen auftreten, da viele Komponenten wie Datenbanken oder Server beteiligt sind. Dies hat zur Folge, dass unterschiedliche Reihenfolgen bei der manuellen Ausführung der Schritte zur Bereitstellung der Komponenten zu Zuständen führen, die vom erwarteten Ergebnis abweichen.

Continuous Delivery (CD), das auf Continuous Integration (CI) basiert, hilft, diese Herausforderungen zu bewältigen, da es eine schnellere und verbesserte Auslieferung von Software ermöglicht. Aus diesem Grund hat sich CD in den letzten zehn Jahren von einer revolutionären Idee zum branchenweiten Standard entwickelt. Teams, die CI/CD-Praktiken anwenden, entwickeln zuverlässigere Software mit einer höheren Qualität [1].

Bestätigt wird dies durch den State of DevOps Report 2022 [5] von DevOps Research and Assessment (DORA), in dem mehr als 80 % der Fachleute angaben, dass Software-Releases in ihrem Unternehmen über eine zentrale CI/CD-Plattform erstellt werden. Der Bericht zeigt auch, dass erstklassige Entwicklungsteams, die ihre Zuverlässigkeitsziele in Bezug auf die Software erreichen, mit einer 1,4-fachen höheren Wahrscheinlichkeit CI praktizieren. Darüber hinaus zeigt der Bericht, dass die Verwendung von CD ein Indikator für eine höhere Leistungsfähigkeit bei der Software-Bereitstellung ist. Dies bedeutet, dass Code-Änderungen schneller bereitgestellt werden können und die Zeit für die Wiederherstellung von Diensten im Falle eines Fehlers kürzer ist.

Ziel dieser Arbeit

Ziel dieser Arbeit ist es, CI/CD-Praktiken zu konzipieren und implementieren und an die Anforderungen und Herausforderungen der Prüf- und Analysesoftware SOLVISO der JW Froehlich Maschinenfabrik GmbH anzupassen. Dabei sollen allgemein anerkannte Methoden und Techniken aus der Industrie berücksichtigt werden, um einen effizienten CI/CD-Prozess zu gewährleisten. Dieser soll zu einer Verbesserung der Softwarequalität und -bereitstellung führen, indem Fehler frühzeitig erkannt und behoben, Bereitstellungsprozesse automatisiert und manuelle Schritte reduziert werden.

Continuous Integration

Die Isolation durch langlebige Feature-Branches erhöht die Wahrscheinlichkeit, dass nach Abschluss der Feature-Entwicklung Integrationsprobleme auftreten, die oft sehr langwierig sind. Continuous Integration (CI) zielt darauf ab, diesen Integrationsprozess zu optimieren. Es beinhaltet das regelmäßige und automatisierte Zusammenführen von Codeänderungen, um sicherzustellen, dass diese kompatibel miteinander sind und die gesamte Anwendung ausgeführt werden kann. Darüber hinaus werden automatisierte Tests eingesetzt, um sicherzustellen, dass die Software in einem bereitstellbaren Zustand ist. Dadurch wird die Qualität der Software verbessert und das Risiko von Fehlern bei der Bereitstellung reduziert [2].

Continuous Delivery

Continuous Delivery (CD) baut auf den Grundlagen von CI auf und erweitert diese um den kontinuierlichen Auslieferungsprozess auf die Produktionsumgebung. CD stellt eine Sammlung von Techniken und Prozessen zur Verfügung, die die Zeit und Risiken bei der Auslieferung neuer Versionen einer Software minimieren, in dem die Zusammenarbeit zwischen Personen aus der Entwicklung, dem Testing und dem Betrieb einer Anwendung verbessert wird. Der

Schwerpunkt liegt darauf, Softwareänderungen durch einen hohen Automatisierungsgrad der Build-, Test- und Bereitstellungsprozesse in kurzen Zeitabständen und mit geringem Aufwand in die Produktivumgebung zu bringen. Verwendet wird dazu eine Deployment-Pipeline [3].

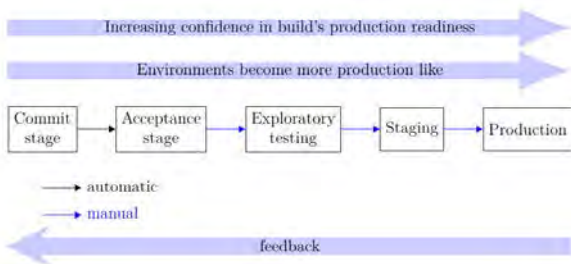


Abb. 1: Deployment-Pipeline [3]

Wichtige Voraussetzungen für die Deployment-Pipeline sind die von CI geforderten Praktiken. Dazu gehört, dass der Code kompiliert wird und mehrere automatisierte Tests dafür sorgen, dass die Anwendung in einem bereitstellbaren Zustand ist. Dies ist Teil der Commit-Phase. Die folgenden Phasen beinhalten automatisierte Tests, aber auch manuelle Tests. Der Schwerpunkt liegt hier auf der Sicherstellung der Nutzbarkeit der Anwendung und dem Auffinden von Fehlern, die nicht durch automatisierte Tests abgedeckt werden. Anschließend erfolgt die Bereitstellung auf einer Staging-Umgebung, die identisch zur Produktivumgebung ist. Ist das Deployment auf die Staging-Umgebung erfolgreich, erfolgt das Release auf die Produktivumgebung. CI ist somit Teil der Deployment-Pipeline, die häufig auch als CI/CD-Pipeline bezeichnet wird [3].

Infrastructure as Code

Infrastructure as Code (IaC) ist ein Ansatz zur Automatisierung von Infrastruktur, der auf Praktiken aus der Softwareentwicklung basiert. IaC konzentriert sich auf konsistente und wiederholbare Prozesse für die Bereitstellung und Änderung von Systemen und deren Konfiguration. Dabei werden Änderungen am Code vorgenommen und anschließend die Automatisierung verwendet, um diese Änderungen an den Systemen zu testen und anzuwenden [4].

Eines der Hauptziele von IaC ist die Vermeidung von Konfigurationsdrift. Konfigurationsdrift tritt auf, wenn manuelle Änderungen oder spontane Änderungen durch Automatisierungstools nur an bestimmten Instanzen vorgenommen werden.

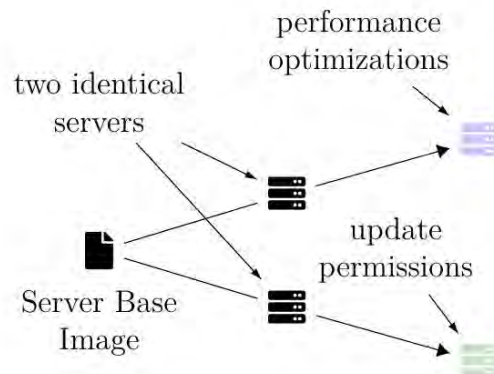


Abb. 2: Konfigurationsdrift ursprünglich gleicher Instanzen [4]

Im Laufe der Zeit wird jeder Server zu einem einzigartigen Server, der nicht automatisch reproduziert werden kann, was zu Bereitstellungsproblemen führen kann. IaC erzwingt diese Konsistenz, indem die vorhandene Infrastruktur in festen Intervallen mit der definierten Konfiguration synchronisiert wird oder durch das Verwerfen und erneute Erstellen der Infrastruktur. Ziel dieser Ansätze ist es, dass unabhängig vom Startzustand der Infrastruktur, immer der gleiche Endzustand erreicht wird.

Häufig werden dazu Infrastructure-Delivery-Pipelines verwendet. Diese dienen zur Bereitstellung und Verwaltung der Infrastruktur, ähnlich der Deployment-Pipeline für Anwendungen. Sie besteht aus einer sequenziellen Abfolge von Schritten, die zur Bereitstellung und Konfiguration der Infrastruktur dienen. Wie die Deployment-Pipeline enthält auch die Infrastructure-Delivery-Pipeline Automatisierungs-, Test- und Validierungsschritte, um sicherzustellen, dass die Infrastruktur korrekt bereitgestellt wird [4].

Der Einsatz von IaC bietet eine Reihe von Vorteilen, wie z.B. die Verringerung des Aufwands und des Risikos von Änderungen an der Infrastruktur. Darüber hinaus werden durch IaC zuverlässige und sichere Systeme geschaffen und die Geschwindigkeit der Fehlerbehebung verbessert. Auf diese Weise schafft IaC die optimalen Voraussetzungen, um die von Continuous Delivery und der Deployment-Pipeline geforderten Ziele einer zuverlässigen Anwendungsbereitstellung zu erreichen.

Konzeption und Implementierung

Die Konzeption und die Implementierung umfassen sowohl die Infrastructure-Delivery-Pipeline als auch die Deployment-Pipeline.

Um die richtigen Voraussetzungen zu schaffen, wurden zunächst die bestehenden Systeme und Prozesse, die bei der Entwicklung von SOLVISO der JW Froehlich Maschinenfabrik eingesetzt werden, analysiert.

Dabei stellte sich heraus, dass sich die Anwendung gut containerisieren lässt, was viele Vorteile beim Deployment mit sich bringt. Beispielsweise können mehrere Anwendungen auf einem Host durch Container isoliert werden. Zur Orchestrierung dieser Container wird Kubernetes eingesetzt.

Die Konzeption und Implementierung der Infrastructure-Delivery-Pipeline ist bereits abgeschlossen. Diese ermöglicht die automatisierte Erstellung eines neuen Kubernetes-Clusters bei Konfigurationsänderungen. Das neue Cluster wird zunächst parallel zum vorhandenen Cluster betrieben. Nachdem der Verkehr erfolgreich umgeleitet und das neue Cluster überprüft wurde, kann das alte Cluster automatisiert gelöscht werden. Damit folgt die bei der JW Froehlich Maschinenfabrik implementierte Infrastructure-Delivery-Pipeline dem Prinzip des Verwerfens und Neuaufbaus der Infrastruktur. Dies hat Vorteile hinsichtlich der Wiederherstellbarkeit des Clusters. Im Falle eines Ausfalls oder einer fehlerhaften Instanz kann dieses einfach verworfen und durch ein neues, vollständig vorbereitetes Cluster ersetzt werden.

Ausblick

Im weiteren Verlauf der Arbeit wird die Deployment-Pipeline konzipiert und implementiert. Es gibt grundsätzlich zwei Ansätze zur Auslieferung von Containern in einem Kubernetes-Cluster: Pull-basierte und Push-basierte Ansätze. Diese Ansätze müssen genauer untersucht werden, um den vielversprechendsten Ansatz zu identifizieren und zu implementieren.

Beim Push-basierten Ansatz erfordert die Deployment-Pipeline Zugriff auf das Cluster, um die neue Anwendung im Cluster hochzufahren. Beim Pull-basierten Ansatz synchronisiert sich ein Observer-Agent innerhalb des Clusters kontinuierlich mit dem Git-Repository. In diesem Repository ist definiert, welche Container im Cluster ausgeführt werden sollen, sowie die Konfiguration dieser Container. Dies hat den Vorteil, dass Cluster, die mit Hilfe der Infrastructure-Delivery-Pipeline erstellt werden, automatisch die Container ausführen, die auch im bereits bestehenden Cluster ausgeführt werden.

Literatur und Abbildungen

- [1] Dave Farley. *Continuous Delivery Pipelines*. Leanpub, 2021.
- [2] Martin Fowler. Continuous Integration. <https://martinfowler.com/articles/continuousIntegration.html>, 2006.
- [3] Jez Humble and David Farley. *Continuous delivery*. Addison-Wesley, 2011.
- [4] Kief Morris. *Infrastructure as Code*. O'Reilly Media, Inc, 2020.
- [5] Claire Peters, Dave Farley, Daniella Villalba, Dave Stanke, Derek DeBellis, Eric Maxwell, John Meyers, Kaiyuan Xu, Nathen Harvey, and Todd Kulesza. 2022 Accelerate - State of DevOps Report. https://services.google.com/fh/files/misc/2022_state_of_devops_report.pdf, 2022.

Konzeption und Entwicklung einer Smart Office Management Anwendung für Desk Sharing und Energieeffizienz mit Anbindung an MS-Teams

Artur Schnell

Andreas Rößler

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma jambit GmbH, Stuttgart

Motivation

Aktuell stehen wir vor einigen Herausforderungen bezüglich der Energiewende. Die steigende Nachfrage nach Energie führt zu einer Überlastung der vorhandenen Ressourcen und einem erhöhten Ausstoß von Treibhausgasen, was wiederum zu Klimaveränderungen führt. [2] Aus diesem Grund sind insbesondere Unternehmen gezwungen, Ressourcen einzusparen, um der Energiekrise entgegenzuwirken. Eine vielversprechende Option, die in den letzten Jahren an Bedeutung gewonnen hat, ist die Implementierung von Home Office in Kombination mit Desk Sharing. Räumlichkeiten werden dadurch effizient genutzt und Ressourcen wie Strom, Heizung und Kühlung eingespart. Insgesamt bietet Desk Sharing also eine Win-Win-Situation für Unternehmen und Mitarbeiter. Es hilft dabei, die Energiekrise zu bewältigen, Kosten zu senken, die Arbeitsumgebung zu optimieren und die Motivation zu steigern. [3] Auch jambit GmbH, ein 1999 gegründetes Softwareentwicklungsunternehmen, zieht in Erwägung, Desk Sharing anzubieten und hierbei ein Tool für eine optimale Organisation der Arbeitsplätze zu nutzen. Dabei spielt die Anbindung an MS-Teams sowie Energieeffizienz eine zentrale Rolle. Die Einschätzung,

ob eine Standardsoftware gekauft oder eine Eigenentwicklung gemacht werden soll, ist für Unternehmen nicht immer einfach. Auf dem Markt stehen zahlreiche Anwendungen für Desk-Sharing zur Verfügung, keine jedoch mit der Funktion zur gezielten Steuerung der Raumtemperatur, abhängig von der Raumnutzung.

Zielsetzung

Ziel ist es demnach, nicht nur die Sitzplätze zu verwalten, sondern auch Büroräume nachhaltig zu heizen bzw. zu kühlen. Hierfür soll eine Anwendung entworfen werden, welche eine maßstabsgetreue Gestaltung der Büroflächen ermöglicht, um somit die Quadratmeter eines Raumes zu berechnen. Diese werden letztlich benötigt, um eine Differenz der Heizleistung ermitteln zu können. Dies geschieht, indem die Anwendung verfügbare Tische in Räumen empfiehlt, welche bereits gebuchte Arbeitsplätze beinhalten. Der Energieverbrauch in den ungenutzten Büroräumen wird somit gesenkt, da diese weniger beheizt, gekühlt oder beleuchtet werden müssen. Zur Veranschaulichung folgt ein Diagramm der beteiligten Komponenten am Gesamtsystem.

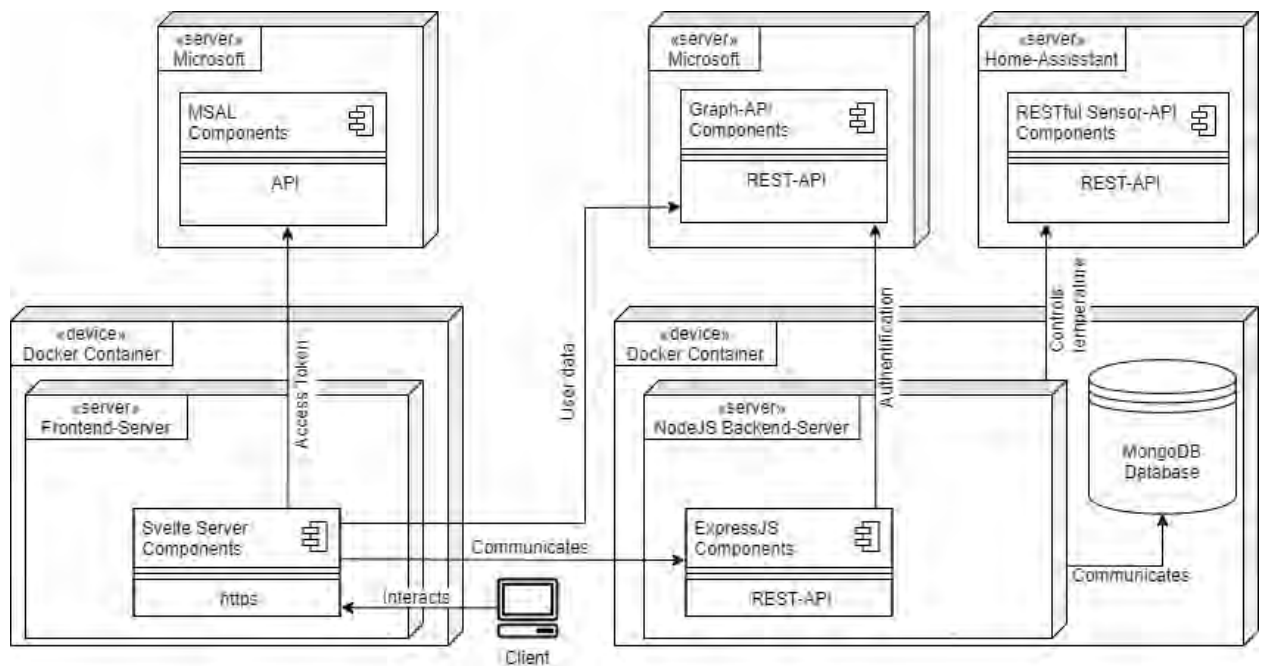


Abb. 1: Architektur der Anwendung [4]

Methodik und Vorgehen

Die Webapplikation selbst, ist eine Single Page Applikation (SPA) und basiert auf dem Svelte-Framework. Das Backend wird mit Node.js in Verbindung mit dem Webframework Express betrieben und verwendet MongoDB als Datenbank. Es bietet eine REST-Schnittstelle zur Interaktion mit den Daten. Für eine möglichst benutzerfreundliche Nutzung, wurde sich für eine Anbindung an Microsoft Teams entschieden, um somit nutzerspezifische Informationen durch die Graph-API erhalten zu können. Die Authentifizierung der Nutzer erfolgt mittels Token durch Microsofts Authentifizierungsbibliothek (MSAL). Das erhaltene Zugriffstoken wird bei jedem Request an das Backend weitergegeben, um den Nutzer zu authentifizieren, zu autorisieren sowie Nutzerdaten abzurufen. Dies

gewährleistet den geschützten Zugriff auf autorisierte Informationen.

Ist ein Benutzer als Verwalter autorisiert, erhält er die Möglichkeit, Büroflächen des eigenen Standortes maßstabsgetreu nachzubilden, sowie Tische für die Platzbuchung zu erstellen. Der erzeugte Grundriss wird zusammen mit den Tischen und der Information des Standortes an das Backend gesendet, dort verarbeitet und zuletzt in der Datenbank hinterlegt. Autorisierte Nutzer können die für ihren Standort hinterlegten Arbeitsplätze anschließend buchen. Dabei achtet die Applikation darauf, stets die Plätze vorzuschlagen, welche aus ökologischer sowie ökonomischer Sicht am effizientesten sind. Die nachfolgende Abbildung gibt einen Einblick über die Ansichten für den Benutzer, als auch Administrator.



Abb. 2: Darstellung der Anwendung mit Benutzeransicht [links] und Verwaltungsansicht [rechts] [4]

Alle Platzbuchungen werden in einer Datenbank gespeichert. Die Anwendung überprüft zu einem festgelegten Zeitpunkt T nach Räumen, die nicht genutzt werden. Die Werte zur Steuerung der vernetzten Geräte werden über das Backend an die REST-API von Home-Assistant gesendet und anschließend mithilfe des Zigbee2MQTT-Protokolls an die Thermostate oder andere Smart-Devices im Internet der Dinge (IoT) weitergeleitet. [5] Zigbee2MQTT fungiert als Software-Schnittstelle und ermöglicht die nahtlose Integration und Steuerung von Zigbee-Geräten in einem Smart-Home-System. Mithilfe des Zigbee2MQTT-Protokolls können verschiedene Zigbee-Geräte unterschiedlicher Hersteller effektiv zusammenarbeiten. Basierend auf diesen Informationen werden die Raumtemperaturen entsprechend angepasst, entweder durch Hochheizen oder Absenken der Temperatur.

Ausblick

Für die Übertragung der Temperaturen wurde die Zigbee Spezifikation gewählt, da dieses Protokoll bereits weit verbreitet ist und von vielen Smart-Home-Geräten unterstützt wird. Allerdings kommunizieren Zigbee Geräte über das lokale Netzwerk. Erst kürzlich wurde Matter als Standard für die drahtlose Kommunikation im Bereich des Smart Homes etabliert und basiert auf dem IP-Protokoll, wodurch Matter-Geräte über das Internet kommunizieren können. [1] Zum aktuellen Zeitpunkt gibt es allerdings noch vergleichsweise wenig Geräte, die vollständig mit dem Matter-Standard kompatibel sind, da die Implementierung und Zertifizierung von Geräten entsprechend Zeit benötigt. [6] Demzufolge wäre der Wechsel von Zigbee zu Matter ein vielversprechender Ausblick für die Zukunft des Smart Homes. Mit zunehmender Verfügbarkeit von Matter-kompatiblen Geräten wird Matter zum Standard für das Smart Home und bietet den Verbrauchern mehr Flexibilität und Kompatibilität.

Literatur und Abbildungen

- [1] Marius Beilhammer. Smart Home leicht gemacht: MATTER als Standard für eine einfache und nahtlose Vernetzung. <https://www.iot-mesh.de/matter-standard-connected-home-over-ip/>, 05 2023.
- [2] Stephan Bosch, Friederike Schlenker, Jochen Bohn, Simone Kupies, and Matthias Schmidt. *Energie-Weltatlas Transformation des Energiesystems in globaler Perspektive*. Springer Vieweg, Wiesbaden, 2023.
- [3] Salvador Bueno, Gonzalo Rodríguez-Baltanás, and Maria Dolores Gallego. Coworking spaces: a new way of achieving productivity. *Journal of Facilities Management*, 2018.
- [4] Eigene Darstellung.
- [5] Linda Rosencrance. Zigbee. <https://www.techtargget.com/iotagenda/definition/ZigBee>, 06 2017.
- [6] Harry Schubert. Version 1.0 von Matter definiert. <https://www.elektroniknet.de/kommunikation/wireless/version-1-0-von-matter-definiert.200218.html>, 11 2022.

Low-Code Plattformen: Vergleich am Beispiel „Monitoring von Code Compliance“

Kevin Schramm

Mirko Sonntag

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma adesso SE, Stuttgart

Einleitung

Die Digitalisierung der Wirtschaft steht vor einer Herausforderung: Der Mangel an qualifizierten Mitarbeiterinnen und Mitarbeitern. Laut Schätzungen des Branchenverbandes Bitkom fehlen im Jahr 2022 rund 137.000 IT-Fachkräfte allein in Deutschland [2]. Diese Zahl stellt eine Steigerung von 10 Prozent im Vergleich zum Vorjahr dar, trotz der negativen Auswirkungen der Corona-Pandemie. Dieser Mangel an Fachkräften hat erhebliche Auswirkungen auf die Produktivität und das Wachstum von Unternehmen. An diesem Punkt setzen Low-Code-Plattformen an, indem sie den Entwicklungsprozess vereinfachen und die Notwendigkeit tiefgreifender Programmierkenntnisse minimieren, erlauben sie sogar Mitarbeitenden mit geringerem technischen Verständnis, Anwendungen zu erstellen. Dies verschafft Unternehmen die Möglichkeit, ihren internen Entwicklungsprozess zu verbessern und die Fachkräftelücke zu überbrücken.

Motivation

Der Markt für Low-Code-Plattformen ist in den letzten Jahren massiv gewachsen, wie aus Abbildung 1 ersichtlich wird besonders im Low-Code Application Segment. Dieser Anstieg ist nicht nur auf den Fachkräftemangel zurückzuführen, sondern auch auf die Bedürfnisse von Unternehmen in einem dynamischen Marktumfeld, in dem schnelle Entwicklung entscheidend ist. Low-Code-Plattformen spielen hierbei eine wichtige Rolle, Sie ermöglichen die schnelle Entwicklung von Prototypen, wodurch die Entwicklungszeit verkürzt wird. Dies erlaubt eine schnelle Iteration, frühzeitiges Feedback und Fehlererkennung. Damit reduzieren sie Entwicklungsrisiken, verbessern die Benutzererfahrung und senken die Gesamtkosten von Projekten. Charakteristische Merkmale von Low-Code-Plattformen umfassen grafische Modellierung, Drag-and-Drop-Oberfläche, automatisiertes Deployment und Wiederverwendung von Komponenten. Diese Funktionen bieten den Unternehmen eine

nutzerfreundliche Entwicklungsumgebung, mit der Anwendungen schnell und effizient erstellt werden können.

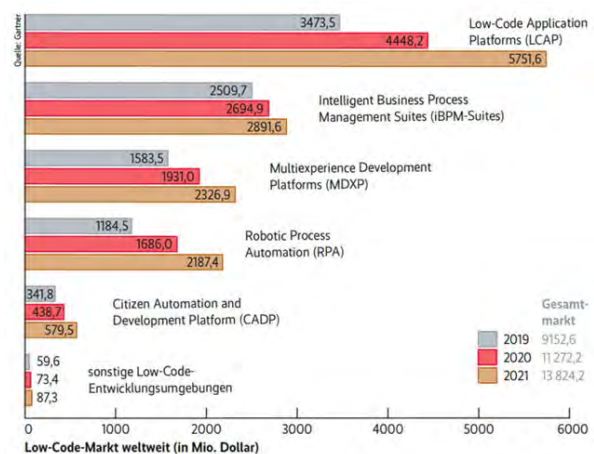


Abb. 1: Marktwachstum nach Gartner [1]

Problemstellung

Unternehmen haben jetzt eine Vielzahl von Optionen, um ihre Anwendungen schneller und effizienter zu entwickeln. Allerdings bringt diese Vielfalt auch Herausforderungen mit sich. Mit mittlerweile über 100 verschiedenen Low-Code-Plattformen wird es für Unternehmen zunehmend schwierig, den Überblick über die Vor- und Nachteile der einzelnen Plattformen zu behalten. Die Vielfalt hat bei den Low-Code-Plattformen zu unterschiedlichen Stärken und Schwächen geführt, was die Auswahl der richtigen Plattform zu einer anspruchsvollen Aufgabe macht. Jede Plattform hat ihre eigenen Eigenschaften, Funktionen und Schwerpunkte. Es ist wichtig, dass Unternehmen ihre spezifischen Anforderungen und Ziele sorgfältig analysieren, um die Plattform auszuwählen, die am besten zu ihnen passt. Die Plattformen lassen sich dabei grob in die folgenden Kategorien einteilen: Low-Code Applications

Plattformen (LCAP) decken den gesamten Lebenszyklus einer Webanwendung ab und ermöglichen eine einfache Gestaltung von Web- und mobilen Oberflächen mit integrierten Datenmanagement-Werkzeugen. Citizen Automation and Development (CADP) Plattformen sind speziell auf weniger technisch versierte Anwender ausgerichtet und bieten Funktionen zur Automatisierung und Entwicklung, ohne dass Datenmanagement-Funktionen zwingend erforderlich sind. Auf der anderen Seite bieten Multiexperience Development Plattformen (MXPDP) umfangreiche Gestaltungsmöglichkeiten und unterstützen moderne Technologien wie IoT-Geräte und Progressive Web-Apps. iPaaS (Integration Platform as a Service) Plattformen sind auf die Automatisierung von Prozessen und Workflows zugeschnitten, mit speziellen Werkzeugen für die Prozessanalyse und -steuerung. Robotic Process Automation (RPA) Tools nutzen Bots, um verschiedene Prozesse zu integrieren und zu automatisieren. Schließlich konzentrieren sich Integration Platform as a Service (iPaaS) Plattformen auf die Backend-Integration und die Koordination von Prozessabläufen, indem sie Middleware-Funktionen bereitstellen [1].

Da es sich beim Ausgewählten Anwendungsfall um eine Webanwendung handelt, die eine Schnittstelle zu externen Datenbanken hat, kommt für diesen Fall nur LCAP-Plattformen in Frage. Es ist wichtig zu beachten, dass diese Kategorien nicht starr sind und es Überschneidungen zwischen den verschiedenen Plattformen geben kann. Der Low-Code-Plattform-Markt ist dynamisch und entwickelt sich ständig weiter, da neue Anbieter auftauchen und bestehende Plattformen ihre Funktionen erweitern. Unternehmen sollten ihre Anforderungen und spezifischen Anwendungsfälle sorgfältig prüfen, um die passende Low-Code-Plattform auszuwählen, die ihren Bedürfnissen am besten entspricht.

Ziel der Arbeit

In dieser Arbeit liegt der Fokus darauf, die Wirksamkeit und Eignung von Low-Code-Application Plattformen für den Einsatz in Unternehmen zu untersuchen. Dazu wird eine eigens entwickelte Metrik verwendet, um die Leistung und Funktionalität verschiedener Plattformen zu bewerten. Um eine umfassende Analyse durchzuführen, wurden drei Plattformen aus der Vielzahl der verfügbaren Low-Code-Application-Plattformen ausgewählt. Die Auswahl erfolgte anhand verschiedener Kriterien, wie Marktrepputation, Kundeneinsatz und Funktionsumfang. Dieser sorgfältige Filterungsprozess stellt sicher, dass die ausgewählten Plattformen repräsentativ sind und eine aussagekräftige Vergleichsbasis bieten. Die Abbildung 2 zeigt die Auswertung des LCAP-Marktes durch das Marktforschungsinstitut Gart-

ner, hier positionieren sich die drei Plattformen als „Leaders“.

Durch den Vergleich dieser ausgewählten Low-Code-Plattformen können die Vor- und Nachteile jedes Systems identifiziert werden. Die Ergebnisse dieser Arbeit soll Unternehmen dabei helfen, die für ihre individuellen Anforderungen am besten geeignete Plattform auszuwählen.

Es wird erwartet, dass dieser Vergleich wichtige Erkenntnisse liefert, um die Entscheidungsfindung zu erleichtern und den Erfolg bei der Implementierung von Low-Code-Plattformen im Unternehmen zu steigern.



Abb. 2: LCAP-Markt nach Gartner [3]

Methoden

Um eine fundierte Bewertung hinsichtlich der Schnelligkeit der Implementierung der Plattformen in einem Unternehmen zu ermöglichen und dabei sowohl deren Grenzen als auch potenzielle Herausforderungen bei der Implementierung zu identifizieren, wird ein praktischer Vergleich angewendet. Der spezifische Anwendungsfall wird mittels der Plattformen Mendix, OutSystems und Microsoft Power Apps umgesetzt. Schon während der Implementierungsphase wird dokumentiert, wie sich die Plattformen hinsichtlich vorab festgelegter Metriken positionieren und wo die jeweiligen Stärken und Grenzen liegen. Es wird eine sorgfältige Betrachtung der gesamten Entwicklungsumgebung, den sogenannten Development Lifecycle, vorgenommen, um zu überprüfen, inwieweit die ausgewählten Plattformen diesen unterstützen. Dieser Ansatz ermöglicht eine detaillierte Untersuchung und Beurteilung der Implementierungsprozesse, und bietet wertvolle Einblicke in

die Effizienz und Anpassungsfähigkeit der Plattformen. Die Erfahrungen und Ergebnisse aus dieser praktischen Umsetzung dienen als Grundlage für die Bewertung der Plattformen.

Ausblick

Im weiteren Verlauf der Arbeit sollen nun genau diese Erfahrungen und Ergebnisse mit den drei Platt-

formen ausgewertet werden, um das Verständnis von Low-Code-Plattformen weiter zu vertiefen und das Unternehmen dabei zu unterstützen, fundierte Entscheidungen darüber zu treffen, welche Plattformen es für die Realisierung spezifischer Anwendungsfälle einsetzen kann.

Literatur und Abbildungen

- [1] Achim Born. Werkzeuge und plattformen für low code und no code development. *iX – Magazin für professionelle Informationstechnik*, 9:42–47, 2021.
- [2] Branchenverband der deutschen Informations- und Telekommunikationsbranche. Trotz krieg und krisen: In deutschland fehlen 137.000 it-fachkräfte. <https://www.bitkom.org/Presse/Presseinformation/Deutschland-fehlen-137000-IT-Fachkraefte>, 2022.
- [3] Paul Vincent, K. Iijima, A. Leow, M. West, and O. Matvitskyy. Magic quadrant for enterprise low-code application platforms. *Gartner Report*, 2022.

Evaluierung der IT-Sicherheitsmaßnahmen einer Software-as-a-Service-Lösung zur Umsetzung des Lieferkettensorgfaltspflichtenschutzgesetzes

Katharina Schulz

Anke Bez

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Ausgangssituation

Das Lieferkettensorgfaltspflichtenschutzgesetz (LkSG) ist ein deutsches Gesetz das am 01. Januar 2023 in Kraft getreten ist. Es stellt einen bedeutenden Meilenstein im Rahmen der unternehmerischen Verantwortung für die Einhaltung sozialer und ökologischer Standards in globalen Lieferketten dar. Das Gesetz ist eine Reaktion auf die Leitprinzipien für Wirtschaft und Menschenrechte der Vereinten Nationen sowie den Nationalen Aktionsplan für Wirtschaft und Menschenrechte von 2016 in Deutschland. Eine Studie im Jahr 2020 ergab, dass lediglich ein geringer Prozentsatz der befragten Unternehmen, nämlich zwischen 13 und 17%, die Vorgaben des Nationalen Aktionsplans erfüllen konnten. Um eine angemessene Einhaltung sicherzustellen, werden nun rechtlich verbindliche und international anerkannte Sorgfaltsstandards eingeführt [1]. Der Prozess der Entstehung des LkSG wird in Abbildung 1 bildlich verdeutlicht. Das Lieferkettensorgfaltspflichtenschutzgesetz legt fest, dass Unternehmen für die Sicherstellung der Menschenrechte in den weltweiten Lieferketten verantwortlich sind. Dies schließt unter anderem den Schutz vor Kinderarbeit, das Gewährleisten fairer Löhne sowie den Umweltschutz ein.

Gemäß diesem Gesetz sind Unternehmen seit diesem Jahr ab einer Mitarbeiterzahl von 3000 dazu verpflichtet, bestimmte Sorgfaltspflichten zu erfüllen. Im nächsten Jahr werden auch Unternehmen mit einer Größe von 1000 Mitarbeitern berichtspflichtig, was eine erweiterte Reichweite der Sorgfaltspflichten bedeutet. Als Sorgfaltspflichten werden verstanden:

1. Einrichtung eines Risikomanagements
2. Festlegung einer betriebsinternen Zuständigkeit
3. Durchführung regelmäßiger Risikoanalysen
4. Abgabe einer Grundsatzerklärung
5. Verankerung von Präventionsmaßnahmen im eigenen Geschäftsbereich
6. Ergreifen von Abhilfemaßnahmen
7. Einrichtung eines Beschwerdeverfahrens
8. Umsetzung von Sorgfaltspflichten in Bezug auf Risiken bei mittelbaren Zulieferern
9. Dokumentation und die Berichterstattung an das Bundesamt für Wirtschaft und Ausfuhrkontrolle (BAFA)

Kommen die Unternehmen dem Gesetz nicht nach, so drohen Sanktionen von bis zu 800.000 Euro, bei Unternehmen mit einem Umsatz von mehr als 400 Millionen Euro bis zu zwei Prozent des Umsatzes. Darüber hinaus kann das betreffende Unternehmen von der Vergabe öffentlicher Aufträge ausgeschlossen werden [4]. Auf dem Markt sind bereits zahlreiche Cloud-Lösungen zu finden, die die Unternehmen hinsichtlich der Umsetzung des LkSG unterstützen. Sie unterscheiden sich in Bezug auf ihre Betriebsform, ihren Funktionsumfang und somit auch den Grad der Automatisierung sowie der Kosten.

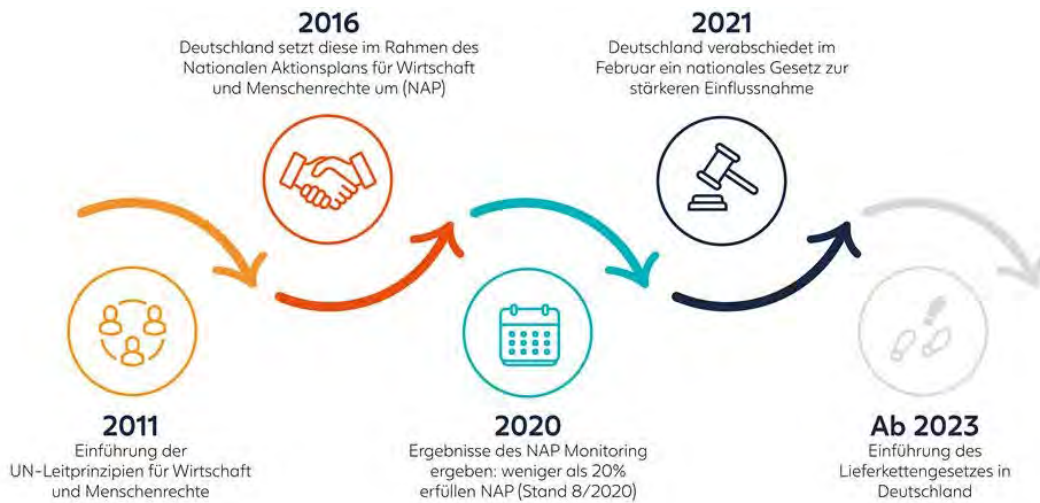


Abb. 1: Entstehung des LkSG [3]

Problemstellung

Die Nutzung der Cloud bietet zahlreiche Vorteile wie Kosteneinsparungen, eine skalierbare Anpassung an den Bedarf und den Zugriff auf Anwendungen und Daten unabhängig vom Standort. Angesichts der fortwährenden Entwicklung des digitalen Umfelds ist der Schutz der Cloud-Umgebung von entscheidender Wichtigkeit. So wird eine steigende Professionalisierung innerhalb der Cybercrime-Branche beobachtet. Das Geschäftsmodell Cybercrime-as-a-Service (CaaS)

ermöglicht mittlerweile Personen ohne Fachkenntnisse, den Zugang zu Hacking-Werkzeugen oder ganze Angriffe in Auftrag zu geben. Pro Jahr wird eine stetige Zunahme an Cyberangriffen verzeichnet [5]. Insbesondere im Cloud-Umfeld muss der Kunde den Cloud-Anbieter hinsichtlich der Sicherheitsvorkehrungen sorgsam überprüfen, um sicherzustellen, dass angemessene Schutzmaßnahmen implementiert sind und ein hohes Maß an Datensicherheit gewährleistet wird.



Abb. 2: Attacken auf die Wirtschaft werden immer professioneller [2]

Zielsetzung

Das Ziel dieser wissenschaftlichen Arbeit besteht darin, die IT-Sicherheit von SaaS-Anwendungen zur Abbildung des Lieferkettenschutzgesetzes zu untersuchen. Hierbei wird eine spezifische SaaS-Lösung eines externen Unternehmens in einer Public Cloud-Umgebung betrachtet. Um die Sicherheit dieser SaaS-Lösung zu bewerten, wird der C5-Anforderungskatalog des Bundesamts für Sicherheit in der Informationstechnik (BSI) als Leitfaden zur Bewertung von Cloud-Anwendungen herangezogen und in Einklang mit den unternehmensspezifischen Sicherheitsrichtlinien gebracht.

Ausblick

Im Verlauf dieser Arbeit werden die Bedrohungspotenziale bei der Nutzung von SaaS-Anwendungen zur Abbildung des Lieferkettenschutzgesetzes detailliert untersucht. Hierbei werden sowohl technische als auch organisatorische Aspekte betrachtet, um eine umfassende Analyse der Sicherheitsanforderungen zu gewährleisten. Die Ergebnisse dieser Arbeit sollen neben einer Einschätzung der IT-Sicherheitslage der SaaS-Anwendung, ein grundlegendes und gesammeltes Werk der IT-Sicherheitsstandards für Cloud-Lösungen liefern. Auf Basis dieses Werkes wird dem Unternehmen eine rasche Entscheidungsfindung bei der Auswahl externer SaaS-Anwendungen sowie deren Überprüfung möglich gemacht.

Literatur und Abbildungen

- [1] Auswärtiges Amt. Monitoring zum Nationalen Aktionsplan Wirtschaft und Menschenrechte. <https://www.auswaertiges-amt.de/de/aussenpolitik/aussenwirtschaft/wirtschaft-und-menschenrechte/monitoring-nap/2124010>, 2020.
- [2] Achim Berg. Wirtschaftsschutz 2022. https://www.bitkom.org/sites/main/files/2022-08/Bitkom-Charts_Wirtschaftsschutz_Cybercrime_31.08.2022.pdf, 2022.
- [3] Proactis Holdings Limited. Abbildung 1: Entstehung des LkSG. <https://www.proactis.com/de/resources/content-hubs/lksg/das-lieferkettengesetz-im-detail/was-ist-das-lksg/>, 2023.
- [4] Corporate Social Responsibility. Gesetz über die unternehmerischen Sorgfaltspflichten in Lieferketten. <https://www.csr-in-deutschland.de/DE/Wirtschaft-Menschenrechte/Gesetz-ueber-die-unternehmerischen-Sorgfaltspflichten-in-Lieferketten/gesetz-ueber-die-unternehmerischen-sorgfaltspflichten-in-lieferketten.html>, 2023.
- [5] Andreas Streim and Mann Simran. Presseinformation. <https://www.bitkom.org/Presse/Presseinformation/Wirtschaftsschutz-2022>, 2022.

Künstliche Intelligenz in der Musik

Thomas Sebast

Dirk Hesse

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Einleitung

Die rasante Entwicklung der Künstlichen Intelligenz (KI) hat in den letzten Jahren zu einer Vielzahl beeindruckender Innovationen geführt, angefangen von fortschrittlichen Large Language Modellen wie LLaMA bis hin zur Generierung realitätsnaher Bilder durch KI-Modelle wie DALLÉ2. [7] [5] Dadurch eröffnen sich neue und innovative Einsatzmöglichkeiten von KI in verschiedenen Bereichen. Doch wie sieht es mit der Anwendung von KI in der Musik aus?

Auch die Musikbranche bleibt von den Innovationen in der KI nicht unberührt. Von Songtext-Generatoren bis hin zur Musikererkennung bietet KI eine Vielzahl von Anwendungsmöglichkeiten. Modelle wie MusicLM können sogar bereits aus Texteingaben vollständige Lieder mit verschiedenen Instrumenten und Melodien generieren. [1]

Zielsetzung

Diese Arbeit bietet nicht nur einen umfassenden Überblick über den aktuellen Stand der KI in der Musik, sondern legt auch einen Schwerpunkt auf computergenerierte Instrumente wie zum Beispiel Synthesizer.

Eine Herausforderung bei der Anwendung von KI-Modellen wie GANSynth in der Musik besteht darin, dass sie oft Black Boxes sind. [4] Obwohl sie beeindruckende Ergebnisse erzielen können, ist ihre Funktionsweise oft nur schwer zu interpretieren. Aus der Perspektive von Musikern ist es jedoch wünschenswert, die Ergebnisse der Modelle in einer für sie anwendbaren Form zu erhalten.

Daher enthält diese Arbeit auch die Konzipierung und Entwicklung eines praktisch anwendbaren Systems zur

Synthetisierung isolierter Audiospuren, das Musikproduzenten bei ihrer kreativen Arbeit unterstützt.

Wie können KI-Modelle mit Klang arbeiten?

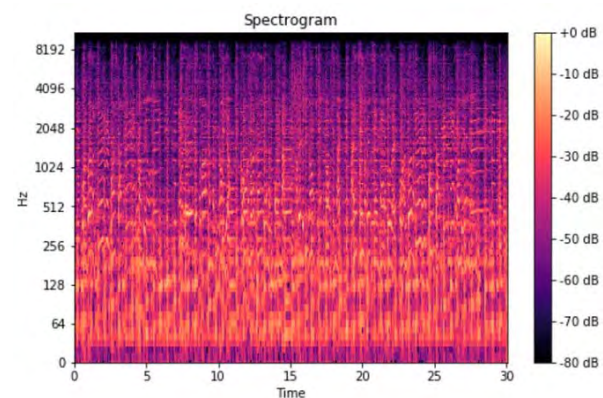


Abb. 1: Beispiel Mel-Spectrogramm [6]

KI-Modelle nutzen das Mel-Spectrogramm (siehe Abbildung 1), um mit Klang zu arbeiten. Das Mel-Spectrogramm wird durch die Anwendung von Kurzzeit-Fourier-Transformationen erstellt und ist eine Darstellung von Audiosignalen, bei der die Frequenz auf der vertikalen Achse und die Zeit auf der horizontalen Achse abgebildet werden. [6] Das Mel-Spectrogramm ermöglicht die Darstellung von Audiosignalen in bildform. Dadurch können KI-Modelle Bildverarbeitungstechniken auf den Audiosignalen anwenden, um Muster und Merkmale im Klang zu identifizieren.

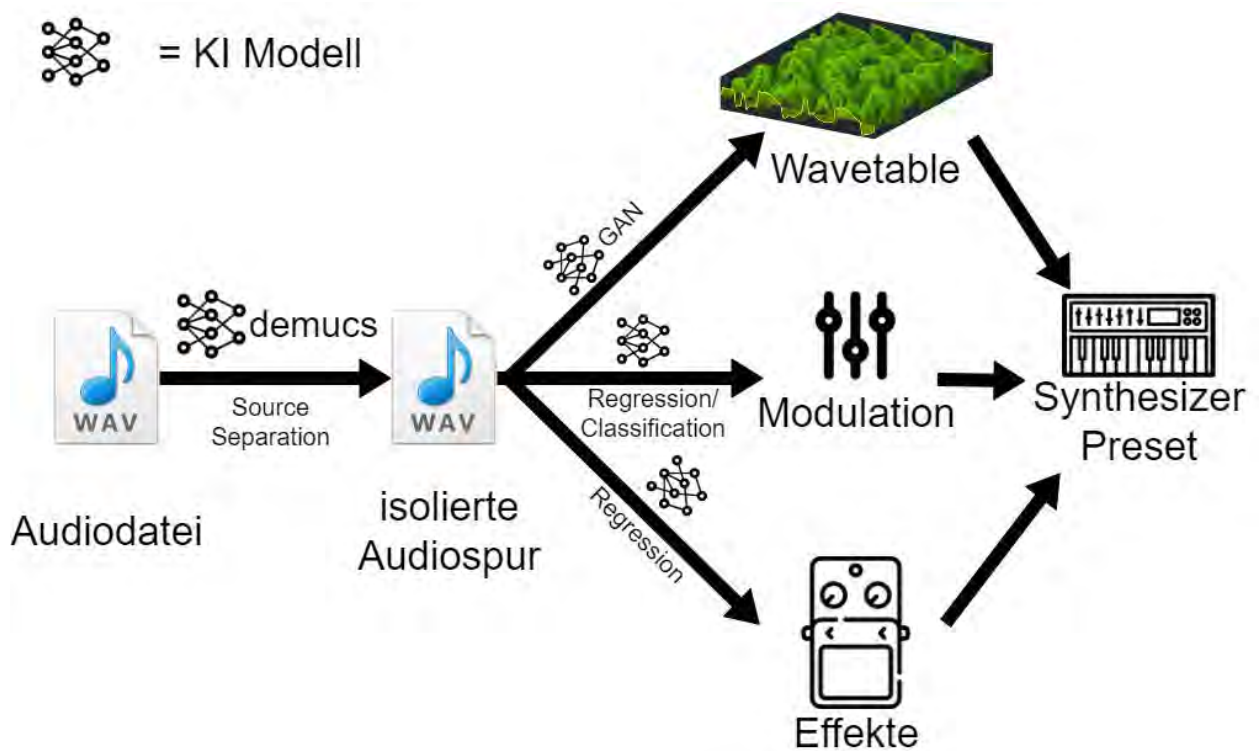


Abb. 2: Aufbau des Synthetisierungsprozesses [2]

Synthetisierung isolierter Audiospuren

Im Bereich des Sounddesigns können zwei unterschiedliche Kombinationen von Parametern den gleichen Ton erzeugen. Dadurch unterscheidet sich dieses Problem grundlegend von anderen Regressions-/Klassifizierungsproblemen. Um mit dieser Herausforderung umzugehen wird folgender Prozess vorgestellt.

Die Abbildung 2 stellt den Prozess der Synthetisierung von isolierten Audiospuren dar. Im ersten Schritt wird mithilfe von Demucs, einem KI-basierten System zur Trennung von Audioquellen, die Stimme, das Schlagzeug und der Bass aus dem ausgewählten Audioabschnitt, welches synthetisiert werden soll, entfernt. [3]

Im folgenden Schritt wird die isolierte Audiospur auf drei verschiedene Weisen weiter verarbeitet. Zuerst wird basierend auf Niklas Wantrupps TableNet ein Wavetable aus der Audiodatei generiert. [8] Die Wavetable dient als Basis für die Klangerzeugung durch den Synthesizer. Der durch den Wavetable generierte Klang kann durch den Einsatz von Filtern modifiziert werden. Dazu dient die zweite Verarbeitung der Audiodatei. Durch die Erzeugung eines Mel-Spectrogramms des Audiosignals können mittels KI definierende Features erkannt werden. Durch Regres-

sion und Klassifizierung ist es KI-Modellen möglich, Filter und dazugehörigen Filter-Modulationen aus dem Spectrogramm abzubilden. Für das Trainieren dieser Modelle wurde ein synthetischer Datensatz aus 1000 Beobachtungen generiert. Zuletzt werden noch mithilfe von Regression Effekte, die auf den generierten Ton angewendet werden, erkannt.

Im letzten Schritt werden die Ausgaben der verschiedenen Modelle in ein Preset für einen Synthesizer kombiniert.

Ausblick

Die bisherigen Ergebnisse des Systems sind vielversprechend. Wie so oft im Bereich des maschinellen Lernens gilt: Mehr Daten bedeuten in der Regel auch bessere Ergebnisse. Deshalb ist es wichtig, in Zukunft nicht nur auf synthetische Daten, sondern auch auf Realdaten zurückzugreifen.

Der vorgestellte Prozess hat vielversprechende Perspektiven, dieses komplexe Problem zu lösen. Die kommenden Monate werden dafür genutzt, die Implementierung des Systems abzuschließen.

Außerdem wird es in Zukunft von Bedeutung sein, eine benutzerfreundliches User Interface zu entwickeln, um die Anwendung des Systems zu erleichtern.

Literatur und Abbildungen

- [1] Andrea Agostinelli, Timo I. Denk, Zalan Borsos, et al. MusicLM: Generating Music From Text. <https://doi.org/10.48550/arXiv.2301.11325>, 01 2023.
- [2] Eigene Darstellung.
- [3] Alexandre Défossez, Nicolas Usunier, Léon Bottou, and Francis Bach. Demucs: Deep Extractor for Music Sources with extra unlabeled data remixed. <https://doi.org/10.48550/arXiv.1909.01174>, 09 2019.
- [4] Jesse Engel, Kumar Krishna Agrawal, and Shuo Chen. GANSynth: Adversarial Neural Audio Synthesis. <https://doi.org/10.48550/arXiv.1902.08710>, 02 2019.
- [5] Aditya Ramesh, Prafulla Dhariwal, Alex Nichol, et al. Hierarchical Text-Conditional Image Generation with CLIP Latents. <https://doi.org/10.48550/arXiv.2204.06125>, 04 2022.
- [6] Leland Roberts. Understanding the Mel Spectrogram. <https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53>, 03 2020.
- [7] Hugo Touvron, Thibaut Lavril, Gautier Izacard, et al. LLaMA: Open and Efficient Foundation Language Models. <https://doi.org/10.48550/arXiv.2302.13971>, 02 2023.
- [8] Niklas Wantrupp. A Data-Driven Approach to Wavetable-Synthesis. <https://www.creative-technologies.de/a-data-driven-approach-to-wavetable-synthesis/>, 2019.

Die Einsatzmöglichkeiten von ChatGPT im Unternehmen und die damit verbundenen Chancen und Risiken

Arasch Seddiqi

Catharina Kriegbaum-Kling

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

Einleitung

Die Bachelorarbeit untersucht die verschiedenen Einsatzmöglichkeiten von ChatGPT im Unternehmen und die Chancen und Risiken, die damit einhergehen. Im Rahmen der Arbeit werden folgende zentralen Fragestellungen erforscht:

Welche Einsatzmöglichkeiten bietet ChatGPT für Unternehmen? Welche Vorteile und Chancen ergeben sich durch den Einsatz von ChatGPT für Unternehmen? In welchen Bereichen ergeben sich Chancen und Vorteile für Unternehmen durch die Nutzung von ChatGPT? Welche Problematiken und Risiken ergeben sich durch den Einsatz von ChatGPT und welche Unternehmensbereiche gelten dabei als besonders problematisch? Verständnis von ChatGPT Zunächst einmal muss erklärt werden, was ChatGPT überhaupt ist. ChatGPT ist die Abkürzung von „Chat Generative Pre-Trained Transformer“ und bezeichnet einen Chatbot, der durch den Einsatz von Künstlicher Intelligenz Texte generiert, um mit seinen Nutzern zu kommunizieren. Dieser Chatbot baut auf einem Modell des maschinellen Lernens auf, das menschliche Eingaben interpretiert und natürlich klingende Antworten liefert.

Theoretische Grundlagen

Um ein umfassendes Verständnis über den Chatbot zu erlangen, ist es zwingend notwendig dem Leser der Bachelorarbeit grundlegende Begriffe wie Künstliche Intelligenz, Chatbots und Maschinelles Lernen zu erklären.

Künstliche Intelligenz (KI) Der Begriff Künstliche Intelligenz (englisch: „Artificial Intelligence“) wurde erstmals im Mai 1956 eingeführt und stellt eines der neuesten Forschungsgebiete in der Bereichen Wissenschaft und Technologie dar (Russell & Norvig, 2010, S. 1). [6] Die Theorie, dass Computer menschenähnliche Denkprozesse aufweisen könnten, bzw. die Idee der Künstlichen Intelligenz, wird bereits seit mehr als fünfzig Jahren in wissenschaftlichen Arbeiten untersucht und diskutiert (Kaplan und Haenlein, 2009, S.16). [4] Deshalb findet

man in der Literatur verschiedene Definitionen und Herangehensweisen zur Erläuterung des Konzepts der Künstlichen Intelligenz (Russell und Norvig (2010, S. 1ff.). [6] Kaplan und Haenlein (2019, S. 15) [4] definieren Künstliche Intelligenz als „die Fähigkeit eines Systems, externe Daten korrekt zu interpretieren, aus diesen Daten zu lernen und das Erlernte zur Erfüllung bestimmter Ziele und Aufgaben durch flexible Anpassung zu nutzen.“ Laut ibm (ibm, 2023) [3] „nutzt Künstliche Intelligenz Computer und Maschinen, um die Problemlösungs- und Entscheidungsfähigkeiten des menschlichen Verstandes nachzuahmen.“

Chatbots Chatbots sind Applikationen, die durch Nutzung von Künstlicher Intelligenz in natürlicher Sprache mit Menschen kommunizieren. Nutzer haben die Möglichkeit, Fragen an den Chatbot zu stellen, auf die der Chatbot in natürlicher Sprache antwortet. Dabei kann der Chatbot sowohl Text- als auch Audioeingaben oder eine Kombination aus beiden verarbeiten. Der Chatbot nutzt Funktionen zur Erkennung natürlicher Sprache, um zu verstehen, was der Nutzer aussagt. Es kommen Formen der Künstlichen Intelligenz zum Einsatz, um die eigentliche Absicht des Nutzers zu ermitteln. Diese Technologien verwenden maschinelles Lernen, um aus Interaktionen zu lernen und die daraus resultierenden Ratschläge und Antworten zu optimieren (vgl. ibm, 2021). [2]

Maschinelles Lernen Maschinelles Lernen stellt eine Unterkategorie der Künstlichen Intelligenz dar. Sein Fokus liegt auf der Ausbildung von Computern, die durch Daten und Erfahrungen lernen und kontinuierlich ihre Leistung steigern, anstatt durch explizite Programmierung gesteuert zu werden (vgl. SAP, 2023). [7] Bild 1 zeigt das Training des Sprachmodells mittels Reinforcement Learning. Reinforcement Learning (RL) ist ein maschinelles Lernverfahren, bei dem ein autonomer Agent durch Interaktion mit seiner Umgebung eine optimale Handlungsstrategie (Policy) entwickelt, um die kumulative Belohnung zu maximieren. Der Agent navigiert durch verschiedene Zustände, führt Aktionen aus und erhält Belohnungen.

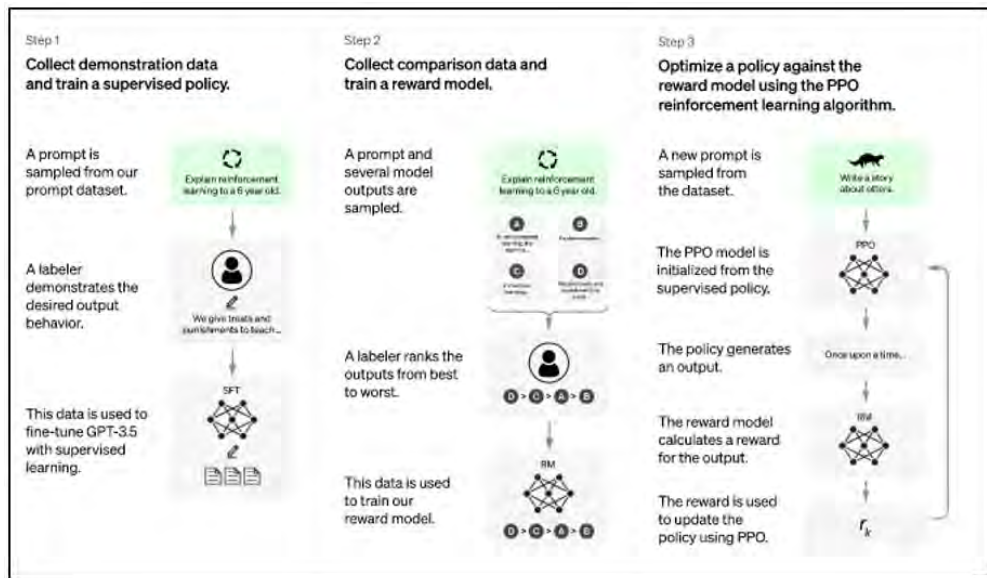


Abb. 1: Training des Sprachmodells mittels Reinforcement Learning from Human Feedback [5]

Einsatz von ChatGPT in Unternehmen

Im Hauptteil wird der aktuelle Forschungsstand hinsichtlich ChatGPT im Unternehmenskontext aufgezeigt und Chancen und Risiken mithilfe der bestehenden Literatur erläutert. Eine beispielhafte Einsatzmöglichkeit von ChatGPT im Unternehmen ist die Nutzung des Chatbots als Textgenerator:

Kommunikation ChatGPT ermöglicht es Unternehmen, sowohl die externe als auch interne Kommunikation zu vereinfachen und beschleunigen (vgl. Impulse, 2021). [1] Bild 2 zeigt die Oberfläche von ChatGBT.



Abb. 2: Oberfläche ChatGBT [5]

Content-Erstellung und Social Media Im Bereich der Content-Erstellung und Social Media bietet ChatGPT Unternehmen die Möglichkeit, menschenähnliche Texte automatisch und in Echtzeit zu generieren. Dies kann bei der Erstellung von Blog-Beiträgen, Social Media Beiträgen, Produktbeschreibungen oder Kundendienst-anfragen genutzt werden (vgl. Impulse, 2021). [1] **Kundenbetreuung und Support** ChatGPT kann als Kundensupport-Bot eingesetzt werden, um Kundenfragen rund um die Uhr zu beantworten, ohne dass ein menschlicher Operator eingreifen muss. Der Chatbot

kann Kundenfragen interpretieren und sofortige, präzise Antworten liefern, was die Kundenzufriedenheit und die Effizienz des Kundenservice verbessern kann (vgl. Impulse, 2021). [1]

Chancen und Risiken des Einsatzes von ChatGPT in Unternehmen

Zusätzlich zu den Möglichkeiten, die ChatGPT bietet, gibt es auch eine Reihe von Risiken, die bei der Implementierung und Nutzung beachtet werden müssen. Diese beinhalten sowohl ethische als auch technische Herausforderungen.

Chancen Die Chancen, die der Einsatz von ChatGPT für Unternehmen bietet, sind vielfältig. Durch den Einsatz von Künstlicher Intelligenz können Unternehmen ihren Kundenservice verbessern, ihre internen Abläufe optimieren und ihre Reichweite in den sozialen Medien erhöhen.

Risiken Die Risiken der Nutzung von ChatGPT in Unternehmen können sowohl technische als auch ethische Aspekte umfassen. Auf der technischen Seite können Herausforderungen wie Datensicherheit, Datenschutz und die Zuverlässigkeit der KI-Technologie auftreten. Auf der ethischen Seite können Probleme wie die Auswirkungen auf Arbeitsplätze und die mögliche Manipulation der öffentlichen Meinung durch KI-generierte Texte eine Rolle spielen.

Schlussfolgerung

Im Abschluss der Arbeit wird ein Fazit gezogen und ein Ausblick auf mögliche zukünftige Entwicklungen und Forschungsfragen gegeben. Es wird diskutiert, inwiefern die Erkenntnisse aus dieser Arbeit für die

Praxis relevant sind und wie Unternehmen die Chancen und Risiken von ChatGPT optimal für sich nutzen können.

Öffentlichkeitsarbeit und Marketing ChatGPT kann eine wertvolle Ressource für die PR- und Marketing-Abteilungen eines Unternehmens sein. Es kann Vorträge vorbereiten, Pressemitteilungen verfassen und sogar Marketing-Kampagnen mit erstklassigem Content unterstützen. Bei all diesen Anwendungen kann ChatGPT dabei helfen, die Konsistenz und Genauigkeit der Kommunikation zu verbessern und gleichzeitig Zeit und Ressourcen zu sparen.

Einsatz von ChatGPT in der Praxis: Eine Umfrage Anhand einer Umfrage soll ermittelt werden, inwieweit ChatGPT tatsächlich in der Praxis eingesetzt wird und welche Erfahrungen Unternehmen und ihre Mitarbeiter

damit gemacht haben. Die Umfrageergebnisse werden mit den Erkenntnissen aus der Literatur verglichen, und die daraus resultierenden Interpretationen werden im Kontext der Ziele und Fragestellungen dieser Arbeit diskutiert.

Zusammenfassung und Ausblick

Zum Abschluss der Arbeit werden die zentralen Erkenntnisse zusammengefasst und Empfehlungen für den Einsatz von ChatGPT in Unternehmen gegeben. Dabei werden auch mögliche Risiken beleuchtet und Vorschläge zur Minimierung dieser Risiken diskutiert. Darüber hinaus werden vorhandene Forschungslücken aufgezeigt und mögliche Ansätze für zukünftige Forschung in diesem Bereich vorgeschlagen.

Literatur und Abbildungen

- [1] K. Halfwassen. So hilft ChatGPT Unternehmen, Zeit und Geld zu sparen. <https://www.impulse.de/organisation/einsatzmoeglichkeiten-chatgpt-unternehmen/7612440.html>, 2023.
- [2] IBM. International Business Machines Corporation. Was ist ein Chatbot? <https://www.ibm.com/de-de/campaign/was-ist-ein-chatbot>, 2023.
- [3] IBM. International Business Machines Corporation. Was ist Künstliche Intelligenz (KI)? <https://www.ibm.com/de-de/topics/artificial-intelligence>, 2023.
- [4] Andreas KAPLAN and Michael. HAENLEIN. *Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence*. Business horizons, 2019.
- [5] Lp. OpenAI. ChatGPT. <https://chat.openai.com/>, 2023.
- [6] Stuart J. RUSSELL. *Artificial intelligence a modern approach*. Pearson Education, Inc., 2010.
- [7] SE. SAP. Was ist maschinelles Lernen? <https://www.sap.com/germany/products/artificial-intelligence/what-is-machine-learning.html>, 2023.

Exploring and Application of Large Language Models for Rule-based Tasks using the Example of Chess Playing

Felipe Sibila Valtierre

Steffen Schober

Department of Computer Science and Engineering, Esslingen University

Work carried out at Department of Computer Science and Engineering, Esslingen

Introduction

Language, more precise human language, is a valuable tool for our development as society and species. As an everyday tool, we can communicate with a precise combination of words that have a meaning, and this meaning is interpreted by another person. This human feature has been recreated over the years by artificial intelligence for text generation, answering questions, creating a summary of a given text and other language-based tasks. One of those recognized artificial intelligence is the AI chatbot ChatGPT, which today uses the newest GPT model. This research presents the idea of training a GPT-2 model in the non-conventional language of chess, meanwhile trying to explore all the different methods and settings on how the training can be optimized in terms of training time and hardware limitation. Furthermore, it will show a comparison of all the different approaches and the results when assessing the fine-tuned model against three different computers in a match of chess. Finally, the open-source chess engine Stockfish will evaluate the results of the games.

Objective

The purpose of this work is to train a model to play the game of chess through the language of SAN, an acronym that stands for Short Algebraic Notation, and to analyze all the possible approaches for an optimized resource efficient training. As a result, the model should understand an input of moves in SAN and generate the next according to the input sequence; in other words, we do not aim for a chess playing machine, we aim for a chess speaking machine.

Previous Work

Today, with the state-of-arts language models, there is the opportunity to prove that a language model can perform a simple task like answering a question, generating text, etc. The idea of this project originated

from an experiment using GTP-3 to play a simple game of chess [7]. The concept of using a language model and the transformer architecture for chess playing was proposed by DeLeon and Guve [2], in which they used BERT as a Masked Language Modelling (MLM) using the Forsyth-Edwards Notation (FEN) of chess. In fact, this notation describes the state of the chess board. This work will instead be using a small version of GTP-2 with 124M parameters [4] as a casual language model and use it to predict the next move from a given sequence of moves in SAN. For doing this, the language Python will be used together with the *Transformer* library from Hugging Face [10].

Dataset

The procedure of selecting data is one of the very crucial steps of this work, because the model learns from this dataset, and it needs to be good for the model to learn properly. The data were obtained from Kaggle, and it contains more than one million of games in the SAN notation [6]. One important distinction is that the model should learn from games that ended in checkmate, because the language model must be capable of playing a game of chess until the very end, so the games that ended with forfeit must be excluded. The so-called Elo rating, created by Arpad Elo, is a system for calculating the skills of players, and it will be used to extract the games with an average Elo between of the two players with a value above 1700. Given that the model needs to learn properly to speak chess, this was a necessary step.

As soon as the filtering has been conducted, the next procedure was to tokenize the SAN string, create a vocabulary and resize the model by adding the new embeddings to the embedding matrix [8]. An advantage of using a casual language model as GTP-2 is, that it was pretrained on raw texts, and it does not have any labelling, therefore it has an automatic process to generate labels from those texts. More specifically, if the model has the input sequence, the

targets will be the same sequence shifted only one token to the right.

Training

As soon as all the preparations were in place, the fine-tuning were performed. This was done by using a single GPU with 4.0 GB of memory, which could be seen as a limitation, but there were a few workarounds for this shortage of GPU. Hugging Face recommends different tricks to work with a single GPU, and in this work those recommendations were implemented. First, because of the lack of memory, the mixed precision of 16-bit was used, and the next issue was to check the maximum size of the data that could be used. Using trial and error, the maximum size at that moment were 300,000 games. As soon as the preparation were completed, three different approaches were implemented for the fine-tuning using a few distinguishable parameters and optimizers with the help from the Trainer class of the Hugging Face library.

1. **Adafactor optimizer:** The first approach was a simple one, but it was the initial step to work with optimizers that have the capability to save memory and to have a reliable performance. For this case, it used a batch size of 2 with a gradient accumulation step of 8.
2. **Adafactor optimizer with a scheduler:** This alternative solution was thought because the first approach did not have a determined learning rate, so it was wise and recommended by Hugging Face to use a scheduler. While this method was better than the first one, it did not make possible to use more games into the fine-tuning.
3. **8-bit Adam optimizer with hyperparameter search:** A slightly more complicated method was implemented, which required to install

the external library *bitandbytes* [3], creating a custom *Trainer* class and implementing a hyperparameter search. 8-bit Adam is an optimizer that can reduce the memory usage while keeping a reliable performance. While using this new optimizer, the next procedure was to find the values of the hyperparameters of β_1 , β_2 , and ϵ , together with the learning rate, weight decay and other parameters. Therefore, after the successful installation of 8-bit Adam, the library *optuna* [1] for hyperparameter search was used. In this case, a small part of the data was selected, only 10,000 games, because the hyperparameter search for 100 iterations takes a lot of time. After the best hyperparameters were found, the model was fine-tuned once more using the default size of 300,000 and the results were close to the second approach. The decision to switch to this optimizer was primarily made, because one can be able to train with more data and therefore a better fine-tuned model is expected.

Results

The three approaches performed as expected, and in Figure 1 indicates that Adafactor and 8-bit Adam performed similar at the end, but the most remarkable find is that this work could use 8-bit Adam further using more games for fine-tuning the language model, so it is expected that the improved model could understand better the language of chess. Another interesting fact is that after the sixth epoch the validation loss did not decrease much, therefore one conclusion is that the model could be fine-tuned with only with 6 epochs. These results are crucial data for further work because the training time could be reduced significantly while increasing the size of the training and test data.

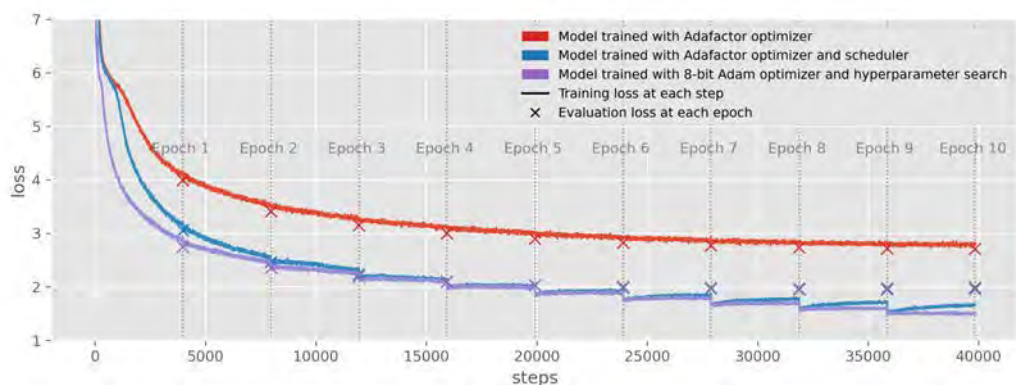


Fig. 1: Model training comparison using 300,000 chess games with different approaches. [5]

Validation

To determine whether the fine-tuned model can speak chess, it was assessed against a computer using the site *Chess.com*. It is expected that the model will not know the rules at this stage, because knowing the rules means that the model should understand the language of chess, which requires more training data. A list of the most probable next moves was generated and the Python library *chess* was used to check if the move is legal or not. If it is not, it will check for the next in the list until a legal move has been provided. Finally, the fine-tuned model was assessed against three types of computers with an Elo rating of 250 (beginner), 1000 (intermediate) and 1800 (advanced). The SAN generated from those games were processed and evaluated by Stockfish, and it shows crucial information on how the model performed. In the first game, Figure 2, the model performed exceptionally, which shows that the opponent did not have any significant score in favor. In the second and third try, Figure 3 and Figure 4 respectively, the model lost against both machines. Figure 3 shows that the model played poorly from the beginning, but in Figure 4 the model performed above average in early game, but starting middle game, it played poorly. A conclusion is that the game of chess gets more complicated with each turn and the opponents are more skillful, and it is too much for the model at this stage of this work.

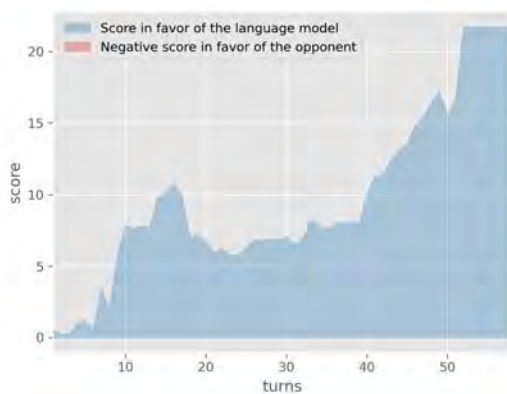


Fig. 2: Score of each turn generated by Stockfish against a computer with Elo of 250. [5]

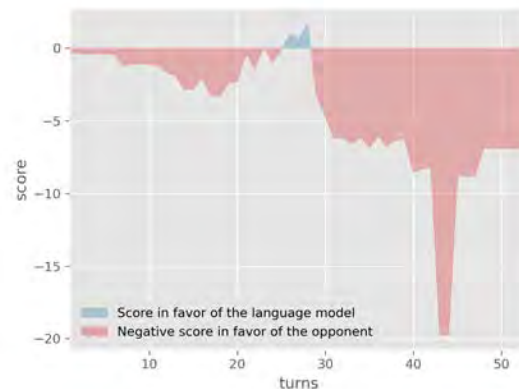


Fig. 3: Score of each turn generated by Stockfish against a computer with Elo of 1000. [5]

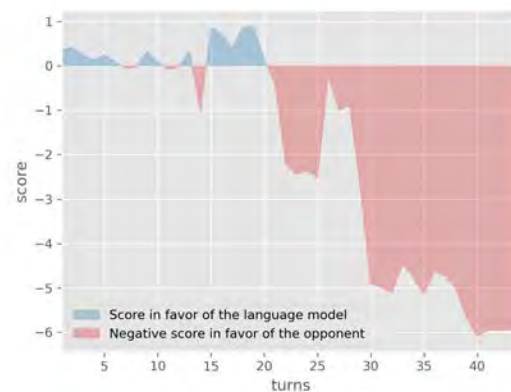


Fig. 4: Score of each turn generated by Stockfish against a computer with Elo of 1800. [5]

Further Work

Future work will focus on the 8-bit Adam optimizer and increasing the test and training data beyond the previous limitation of 300,000 games. Further research will try to find other hyperparameters to use, beside from the ones found with *optuna*. Another approach is to use the special data type of the Nvidia Ampere hardware called *tf32* [9] to optimize the implementation. Furthermore, the use of other types of schedulers will be tested.

References and figures

- [1] Takuya Akiba, Shotaro Sano, Toshihiko Yanase, Takeru Ohta, and Masanori Koyama. Optuna: A Next-generation Hyperparameter Optimization Framework. *KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019.
- [2] Michael DeLeo and Erhan Guven. Learning Chess With Language Models and Transformers. In *Data Science and Machine Learning*. Academy and Industry Research Collaboration Center (AIRCC), 2022.
- [3] Tim Dettmers, Mike Lewis, Sam Shleifer, and Luke Zettlemoyer. 8-bit Optimizers via Block-wise Quantization. *9th International Conference on Learning Representation*, 2022.
- [4] HF Canonical Model Maintainers. gpt2 (Revision 909a290). <https://huggingface.co/gpt2>, 2022.
- [5] Own representation.
- [6] A. Revel. Chess Games. <https://www.kaggle.com/datasets/arevel/chess-games>, 2021.
- [7] Shameed Sait. GPT-3, Play Chess! <https://towardsdatascience.com/gpt-3-play-chess-d123a96096a9>, 02 2022.
- [8] Tobias Sterbak. How to add new tokens to huggingface transformers vocabulary. <https://www.depends-on-the-definition.com/how-to-add-new-tokens-to-huggingface-transformers/>, 2022.
- [9] Dusan Stosic and Paulius Micikevicius. Accelerating AI Training with NVIDIA TF32 Tensor Cores. <https://developer.nvidia.com/blog/accelerating-ai-training-with-tf32-tensor-cores/>, 01 2021.
- [10] Thomas Wolf et al. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations. *Association for Computational Linguistics*, 2020.

UI und UX Design eines Alpaka Herdmanagementsystems

Manuel Staub

Astrid Beck

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt an der Fakultät Informatik und Informationstechnik, Esslingen

UI und UX Design eines Alpaka Herdmanagementsystems

In der hier vorgestellten Arbeit geht es im Kern um das UI & UX Design eines Herdmanagementsystems für Alpakas. Da diese Tiere immer beliebter werden, sowohl in der Zucht als auch im Hobbybereich, sieht der Autor die Notwendigkeit in der Konzipierung eines All-In-One Tools für Alpakahalter. Ziel ist es, eine Anwendung zu erstellen, welche Endgeräteübergreifend für jede Herdengröße und Haltungsart eine Verbesserung und Vereinfachung des Managements darstellt. Getreu dem Sprichwort von Benjamin Franklin „By Failing to prepare, you are preparing to fail.“ [2] soll diese Arbeit Grundlagen schaffen, damit ein solches System auf einer guten Basis benutzerfreundlich und erfolgreich aufgebaut werden kann.

Die Anwendung ist in erster Linie für Deutschland geplant, weswegen sich die Arbeit auf die aktuelle Alpaka-Situation in Deutschland bezieht. Hat das System Erfolg, ist eine globale Lösung in verschiedenen Sprachen und mit lokalen Anpassungen an Registrierung, Management und evtl. weiteren Marktanpassungen möglich. Zu Beginn soll ein kurzer Überblick über Alpakas und die dort anfallenden Aufgaben im Herdmanagement gegeben werden. Auf dessen Basis kann dann die Ausarbeitung eines Konzepts mittels User Centered Design Prozess nach dem dann im Anschluss das Alpaka Herdmanagementsystem entwickelt werden kann.



Abb. 1: ideales Huacaya Alpaka [1]

Alpakas wurden bereits 4000 v. Chr. aus den wild lebenden Vikunjas domestiziert. Darunter versteht man die sogenannte Haustierwerdung einer Wildform [3]. Abb. 1 zeigt ein solches Alpaka.

Der Grund dafür war, dass Vikunjas eine sehr hochwertige Faser produzieren, jedoch nur ein geringes Vliesgewicht besitzen. Das Ziel war also ein Vikunja mit höherem Vliesgewicht, bei gleichbleibender Qualität zu züchten [4]. Die daraus entstandenen Alpakas haben global verschiedene Nutzungsarten. Die Nutzung von Alpakas in Deutschland unterscheidet sich teilweise deutlich im Vergleich zum Rest der Welt. Ursprünglich wurden sie in ihrer Heimat, den peruanischen Anden, vor allem für Wollgewinnung gezüchtet, siehe Abb. 2. In der Zucht ist die Faserqualität sehr wichtig und Deutschland ist dabei global gesehen in einer sehr guten Ausgangslage. Dies zeigt sich unter anderem auch daran, dass deutsche Tiere auf internationalen Shows viele Champion-Titel gewinnen. Die Zucht ist auch laut der Umfrage, die während der Bachelorarbeit durchgeführt wurde, der häufigste Nutzungsgrund.



Abb. 2: Alpaka Vlies [1]

Zudem gaben einige Teilnehmer der Umfrage an, dass sie mit ihren Alpakas Erlebnisse, wie zum Beispiel Wanderungen, tiergestützte Therapie und Hofführungen anbieten. Diese Angebote werden aktuell sehr gut angenommen. Das ist auch der Grund, warum Tiere, die nicht für die Zucht verwendet werden in diesen Bereich Ihre Nutzung finden und Alpakas in Deutschland derzeit nicht geschlachtet werden. Anhand dieser und

weiterer Nutzungsgründe können im Anschluss auch verschiedene Nutzergruppen und Aufgaben innerhalb des Herdmanagements der Alpakas abgeleitet werden.

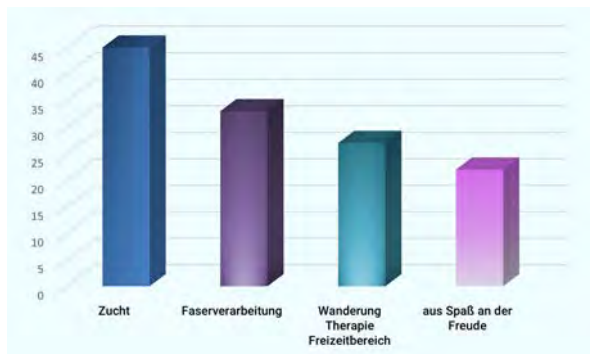


Abb. 3: Diagramm mit Umfrage-Ergebnissen der Nutzungsarten von Alpakas [1]

Ein Alpaka Herdmanagementsystem ist ein System, das alle Verwaltungsaufgaben der Alpakas vereint und übersichtliche Daten und Informationen über die Tiere bereitstellt. Für alle Nutzergruppen des Alpaka Herdmanagements bei den Aufgaben ist vor allem das Gesundheits- und Ernährungsmanagement von Bedeutung. Dies reicht von Gewichtsmonitoring, über Tuberkulose Tests bis hin zu Impfungen. Weitere Punkte sind zum Beispiel für die im vorigen Absatz beschriebenen Züchter, das Zucht- und vor allem Qualitätsmanagement von hoher Bedeutung. Genetik, Faserwerte, DNA und Showergebnisse sind bei potenziellen Nutzern von großer Wichtigkeit. Diese und viele weitere Aufgaben sollen in einem All-In-One Tool vereint werden, sodass der Endnutzer einen guten Workflow hat und auch mit bestehenden Tools zusammenarbeiten kann. Am häufigsten wurde der Wunsch in der Umfrage von potenziellen Nutzern geäußert, dass die Anwendung auf allen Geräten genutzt werden kann. Hierfür eignet sich ein responsives Design, welches sich automatisch an die jeweilige Bildschirmgröße des Smartphones, Tablets oder Computers anpasst. Gerade bei mobilen Endgeräten muss bei der Farbgebung und Schrift darauf geachtet werden, dass trotz Sonneneinstrahlung oder Dunkelheit im Stall eine optimale Nutzung möglich ist. Im Fokus soll der Prozess User Centered Design stehen, bei dem der Schwerpunkt

auf dem Endnutzer liegt und die Produktivität des Endproduktes maximiert wird. Hierfür sollen bereits früh im Planungsprozess verschiedene Nutzergruppen in den Prozess mit eingebunden werden. Dies erfolgt über gestützte und ungestützte Umfragen und sich wiederholende Tests mit Feedback, sowie Gespräche mit potenziellen Nutzern.

Beim User Experience Design soll vor allem ein schnelles und einfaches Arbeiten des Nutzers möglich sein. Die Benutzer sollen sich schnell, einfach und intuitiv zurechtfinden, falls sie das Herdmanagement auf einem anderen Gerät öffnen. Funktionen, wie Filter, Gruppierungen, übersichtliche Auflistungen und Reports sollen das Arbeiten erleichtern. Zusätzlich soll z. B. auch die Einspeicherung von Medikamenten oder Stammbäumen möglich sein, die wiederum das Nutzererlebnis steigern und vereinfachen. Das User Interface muss logisch und übersichtlich aufgebaut und darf nicht zu verschachtelt sein. Der Nutzer soll sich gut und schnell zurechtfinden. Gegebenenfalls ist es sinnvoll, dem User Anpassungen der Software anzubieten. Ein reiner Wanderbetrieb benötigt beispielsweise keine Zucht-Management-Features.

Vor dem eigentlichen User Centered Design Prozess erfolgte noch die Analyse von Mitbewerbern im Bereich Alpaka Herdmanagementsysteme und anderen Tools, wie Verkaufsplattformen, Registratur u. v. m.. Danach wurde mit dem eigentlichen UCD-Prozess, der in Abb. 4 zu sehen ist, weiterverfahren. Zu Beginn wurde der eigentliche Prozess geplant, die Punkte wie Finanzierung und Mitarbeiter enthält. Im Anschluss daran geht es um den eigentlichen Nutzungskontext mit der Festlegung auf Nutzergruppen, deren Umfeld und Ressourcen und Personas zur Veranschaulichung. Darauf folgend werden die Nutzungsanforderungen und Erfordernisse aufgeführt, um daraus Funktionen und für das spätere Projektmanagement Features ableiten zu können. Danach wird ein Konzept entwickelt, welches angefangen von einer Corporate Identity bis hin zu einem Interface Design und Click-Dummy reicht, welcher dann im Anschluss von potenziellen Nutzern getestet werden kann. Der UCD-Prozess wird so lange iteriert, bis das Konzept zufriedenstellend ist. Die Arbeit begrenzt sich aus Zeitgründen und der guten Vorkenntnisse auf einen Durchlauf.



Abb. 4: User Centered Design Prozess [1]

Zum Abschluss der Arbeit wurde sich mit einem möglichen Projektmanagement, Geschäftsmodell und möglichem Marketing befasst. Bei ersterem wurde sich für Feature Driven Development entschieden, da es gut an den UCD-Prozess anknüpft und die Vorteile von klassischen Projektmanagementmethoden mit modernen agilen Methoden vereint. Beim Geschäftsmodell hat sich ein klassisches Abo-Modell als das Beste für das Unternehmen und die Endnutzer herauskristallisiert

und beim Marketing verschiedene Werbeansätze, um möglichst viele zufriedene Nutzer aus der Alpakaszene zu generieren.

Zusammengefasst lässt sich als Ergebnis der Arbeit feststellen, dass ein gutes UI und UX Design als Basis für das Alpaka Herdmanagement erreicht wurde und bei Testpersonen und Lesern der Arbeit aus der Alpakaszene bereits hohen Anklang fand und somit das zu Beginn gesetzte Ziel erreicht wurde.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Benjamin Franklin. Forbes Quotes. <https://www.forbes.com/quotes/1107/>, 2023.
- [3] Matthias Gauly, Jane Vaughan, and Christopher Cebra. *Neuweltkameliden - Haltung, Zucht, Erkrankungen*. Enke Verlag, 3 edition, 2011.
- [4] Michael Safley. *Vicuñas - Survival of the Finest*. Quechua Benefit, 2020.

Relative Positionsbestimmung von Agrarrobotern mit visueller Odometrie

Maximilian Stock

Steffen Schober

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma AGCO GmbH, Marktoberdorf

Motivation

Mechanische Unkrautbekämpfung nimmt einen immer höheren Stellenwert ein. So hat sich die Europäische Union im Rahmen des Green Deals das Ziel gesetzt, bis 2030 den Einsatz von chemischen Pestiziden um 50% zu verringern. [3]

Ziel der Arbeit

Im Rahmen der Arbeit soll für die mechanische Unkrautbekämpfung ein visuelles Odometrieverfahren für den Einsatz im Agrarbereich implementiert und evaluiert werden. Ziel ist dabei die Positionsbestimmung detektierter Pflanzen relativ zum Hackmodul.

Grundlagen der (visuellen) Odometrie

Unter Odometrie verstehen sich Verfahren, die die Position eines Agenten wie z.B. eines Roboters relativ zu einer anderen Position schätzen. Der einfachste Ansatz hierfür ist die Radodometrie. Bei diesem Verfahren werden mit Hilfe eines Encoders die Umdrehungen der Räder gezählt. Anhand dessen kann die aktuelle Position geschätzt werden. Aufgrund eines möglichen Schlupfs der Räder ist dieses Verfahren fehleranfällig. Zudem kann sich der Fehler über die Zeit aufsummieren. Eine Alternative zur Radodometrie stellt die visuelle Odometrie (VO) dar. Als Eingangssignal für die Positionsschätzung dient bei der visuellen Odometrie eine Bildsequenz, die von einer oder mehreren Kameras bereitgestellt wird. Die Änderung der Position wird anschließend aus zwei aufeinanderfolgenden Bildern bestimmt. [1]

Visuelle Odometrieverfahren können nach Art der verwendeten Kamera klassifiziert werden. Verwendung finden folgende Kameratypen:

- Monokular
- Stereo

Eine weitere Klassifizierung von VO-Verfahren erfolgt nach dem verwendeten Ansatz für die Positionsschätzung. Es existieren drei Ansätze:

- Merkmalsbasiert
- Direkt
- Hybrid

Beim merkmalsbasierten Ansatz werden Merkmale wie Ecken und Kanten aus den Bildern extrahiert. Die Art, wie die Bewegung des Agenten bestimmt wird, unterteilt sich hierbei in zwei Klassen – Matching und Tracking. Bei Matching werden die aus zwei aufeinanderfolgenden Bildern extrahierten Merkmale einander zugeordnet und auf Grundlage dessen die Bewegung des Agenten geschätzt. Bei Tracking hingegen werden aus dem vorhergehenden Bild extrahierte Merkmale im aktuellen Bild gesucht. [1] [2]

Der direkte Ansatz hingegen verwendet keine Merkmale aus den Bildern. Stattdessen werden die Pixelintensitäten der Bilder genutzt. Unterschieden wird dabei zwischen dem Dense, Semi-Dense und Sparse Verfahren. Diese beschreiben wie viele Pixel pro Bild verwendet werden. Die Bewegungsschätzung ergibt sich aus der Minimierung eines Fehlermaßes, das auf den Intensitäten der Pixelebene basiert. [1] [4]

Der hybride Ansatz vereinigt die beiden Ansätze und kombiniert die Vorteile der einzelnen Verfahren. Der merkmalsbasierte Ansatz ist besonders für texturierte Umgebungen wie z.B. Städte geeignet, wohingegen er für texturarme Umgebungen wie z.B. Sandflächen aufgrund der fehlenden Merkmale weniger geeignet ist. In solchen Umgebungen ist der direkte Ansatz deutlich robuster. [2] [4]

Neben reinen VO-Systemen existieren sogenannte SLAM (simultaneous localization and mapping) Systeme. Ziel dieser Systeme ist es, zusätzlich zu der Positionsbestimmung eine globale Karte der Umgebung zu erstellen und die Position des Agenten darin zu bestimmen. [1]

Die Erweiterung reiner VO-Verfahren sind VIO-Verfahren (visual-inertial odometry). Für eine bessere Bewegungsbestimmung werden dabei im Rahmen einer Sensorfusion die Daten einer IMU (inertial measurement unit) miteinbezogen. Eine IMU ist eine Messeinheit, die über Beschleunigungssensoren sowie Drehratensensoren verfügt. [2] [5]

Umsetzung

Um ein geeignetes visuelles Odometrieverfahren für die spätere Implementierung auszuwählen wird im Vorfeld eine Recherche zu State of the Art VO-Verfahren durchgeführt. Für eine Übersicht zu aktuellen VO-Verfahren wird auf [2] verwiesen. In einem nächsten Schritt werden vielversprechende Kandidaten wie SVO Pro (semi-direct visual odometry) [4] und DM-VIO (delayed marginalization visual-inertial odometry) [5] näher betrachtet. Der zentrale Unterschied dieser beider Verfahren liegt im verwendeten Ansatz für die Positionsschätzung. SVO-Pro nutzt den hybriden Ansatz, wohingegen DM-VIO einen direkten Ansatz verwendet. Aufgrund der erwarteten repetitiven Texturen im Einsatzgebiet wird kein reines merkmalsbasiertes Verfahren untersucht.

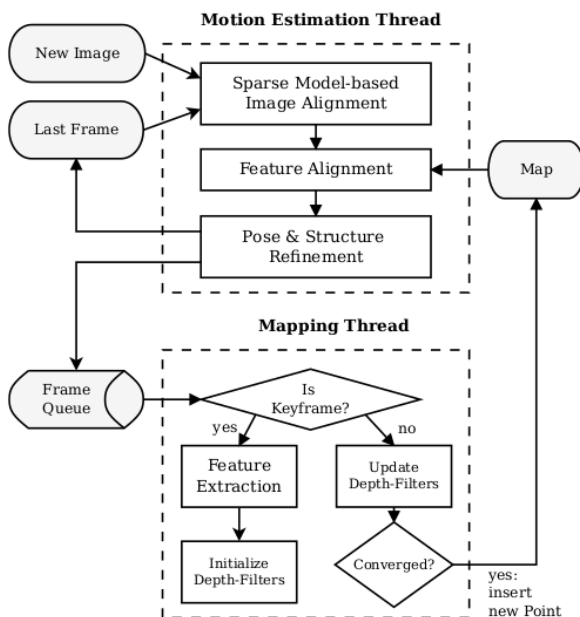


Abb. 1: SVO Pro Pipeline [4]

Als erstes Verfahren wird SVO Pro evaluiert. Dieses Verfahren zeichnet sich besonders durch seine geringe Laufzeit aus. Die SVO Pro Pipeline ist in Abbildung 1 dargestellt. Diese besteht aus zwei parallelen Threads von denen einer für die Bewegungsschätzung (Motion Estimation Thread) und einer für die Tiefenbestimmung (Mapping Thread) zuständig ist. Neben der reinen VO verfügt SVO Pro zusätzlich über eine VIO Implementierung. SVO Pro ist kompatibel mit monokularen Kameras und Stereokameras. [4]

DM-VIO ist eine direkte Methode und hat auf verschiedenen Vergleichsdatensätzen sehr gute Ergebnisse geliefert. In Abbildung 2 sind beispielhaft Punktwolken und Trajektorien für verschiedene Datensätze dargestellt. DM-VIO kann als reines VO-System betrieben werden, ist aber genau genommen ein VIO-System und erzielt unter Berücksichtigung der Daten einer IMU die besten Ergebnisse. Im Gegensatz zu SVO Pro unterstützt das Verfahren lediglich monokulare Kameras. [5]

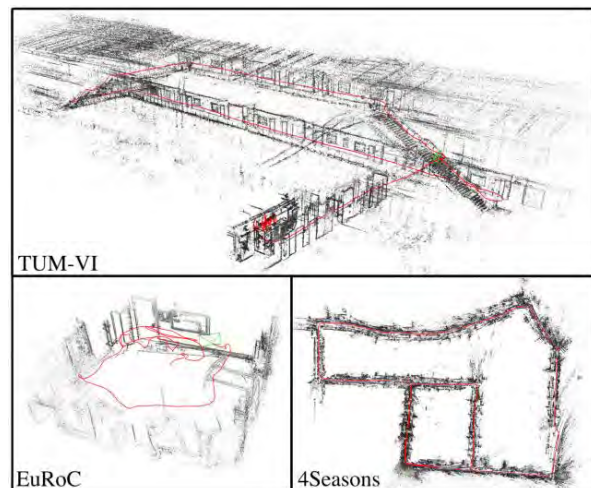


Abb. 2: Punktwolken und Trajektorien für verschiedene Datensätze [5]

Literatur und Abbildungen

- [1] M.O.A. Aqel, M.H. Marhaban, M.I. Saripan, et al. Review of visual odometry: types, approaches, challenges, and applications. *SpringerPlus* 5, 1897, 2016.
- [2] C. Campos, R. Elvira, J.J.G. Rodriguez, J.M.M. Montiel, and J.D. Tardos. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM. *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1874-1890, 2021.
- [3] Union European. About the sustainable use of pesticides. https://food.ec.europa.eu/plants/pesticides/sustainable-use-pesticides_en, 2023.
- [4] C. Forster, Z. Zhang, M. Gassner, M. Werlberger, and D. Scaramuzza. SVO: Semidirect Visual Odometry for Monocular and Multicamera Systems. *IEEE Transactions on Robotics*, vol. 33, no. 2, pp. 249-265, 2017.
- [5] L. von Stumberg and D. Cremers. DM-VIO: Delayed Marginalization Visual-Inertial Odometry. *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 1408-1415, 2022.

Untersuchung und Entwicklung verschiedener Prediction Heads in Convolutional Neural Networks für Fahrzeugerkennung

Tim Straetz

Steffen Schober

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma IT Designers, Esslingen

Im Fokus dieser Arbeit steht die Erkennung von Fahrzeugen mithilfe von CNNs (Convolutional Neural Networks), dabei wird die Art des Fahrzeugs (LKW (Last Kraft Wagen), PKW (Personen Kraft Wagen), Bus) erfasst. Des Weiteren werden verschiedene Keypoints (z. B. die Ecken eines Autos beziehungsweise sein Front-Center) erkannt und gekennzeichnet. Dazu wird YOLOv8 verwendet. YOLOv8 ist zu diesem Zeitpunkt eines der schnellsten und performantesten, frei verfügbaren Systeme.

Problemstellung

Mit klassischer Bounding Box Klassifizierung tritt das Problem auf, dass diese immer parallel zu den X und Y Koordinaten des Bildschirms sind. Weiter können die Fahrzeugausprägungen nicht hervorgehoben werden. Auch ist es nicht möglich, einen Anhänger innerhalb der Bounding Box zu identifizieren.

Für dieses Projekt werden Drohnendaten, welche eine Top-Down-Sicht auf einen Fahrzeugverkehr zeigen, verarbeitet. Auf diesen Bildern werden mithilfe von CNNs die verschiedenen Fahrzeug Keypoints eingezeichnet. Ein Keypoint können die Räder, die Ecken des Autos oder das Ende der Motorhaube (Front-Center) sein, auch der Lenkpunkt bei einem großen Bus oder die Anhängerkupplung bei einem Auto mit Anhänger können Keypoints sein.

CNNs ist eine Art Artificial Neural Network, welches häufig zur Klassifizierung von Bildern eingesetzt wird. Dieses Netzwerk wird darauf trainiert, auch mit 2D-Daten von Verkehrsflüssen umzugehen. Das ist die zentrale Frage dieser Arbeit.

Grundlagen

CNNs sind eine spezielle Art von NNs (Neural Networks), die aus mehreren Schichten, auch Hidden Layers genannt, bestehen. Sie arbeiten auf einem Rastergitter und extrahieren verschiedene Features mit einem Kernel. Eine CNNs Schicht besteht aus mindestens einem Convolutional Layer in dem die Features extrahiert werden.

Oft wird auf diesen Convolutional Layer ein Pooling Layer angewendet, das die gefundenen Merkmale zusammenfasst. Schließlich werden diese Zusammenfassungen oft an ein Fully Connected Layer weitergeben. [3]

Convolutional Layer

Ein wesentlicher Unterschied zwischen NNs und CNNs liegt darin, dass CNNs Convolutions verwenden. In seiner generellsten Form besteht eine Convolution aus zwei Funktionen, die aufeinander angewandt werden. Das erste Argument wird auch input genannt, die zweite Funktion wird oft Kernel genannt. Das Ergebnis der beiden Funktionen nennt man Feature Map. [2] Ein großer Vorteil von CNNs besteht darin, dass sie in der Lage sind einen großen Input (z. B. ein größeres Bild), effizient in Bezug auf Rechenzeit und Speicher zu verarbeiten.

Ein weiterer Vorteil von CNNs ist, dass sie nicht viele Trainingsbilder benötigen und lernen können, welche Features eines Filters am wichtigsten sind. Dieses Lernen, welche Filter wichtig sind, erfolgt während der Trainingsphase von CNNs und kann besonders hilfreich sein, wenn große Datensätze, verarbeitet werden. Grundsätzlich gilt jedoch die Aussage, dass CNNs die Kernel besser anpassen können, wenn mehr Daten zur Verfügung stehen. [4]

Funktionsweise und Aufbau von CNNs

CNNs nehmen mehrdimensionalen Arrays auch Tensoren genannt als Input entgegen und führen eine Matrizenmultiplikation auf einer Matrix aus. Diese Multiplikation kann auf verschiedene Arten von Input angewendet werden. Wie z.B. Bilder und Audiodateien verarbeitet werden. Wichtig für die Verarbeitung der Dateien sind die folgenden Punkte.

Die Dateien sind in einem multidimensionalen Array gespeichert. Es ist mindestens ein Feature vorhanden, welches geordnet werden kann (z. B. die Zeitachse einer Audiodatei). Verschiedene Schichten der Input Datei sind in Channels unterteilbar (z. B. sind die RGB Werte

von Bild-Dimensionen, die in jeweils einen Channel unterteilt werden) [7]

CNNs bauen auf Erkenntnissen der Neurowissenschaft auf, indem die Neuronen, auch Nodes genannt, in tiefen Schichten organisiert sind. Diese Nodes sind Funktionen, die eine gewichtete Summe des Inputs als Output liefern.

Oft findet nach einem CNNs-Layer auch eine Normalisierung statt.

Die Ergebnisse werden dann an die nächste Schicht von Nodes weitergeleitet. Dieser Prozess wird fortgesetzt, bis eine Klassifizierung mithilfe einer Softmax-Schicht vorgenommen wird. [4]

NNs werden als CNNs bezeichnet, wenn mindestens eine Convolution in einem der Layer durchgeführt wird. Traditionelle NNs nutzen Matrizen Multiplikation des Inputs mit einem Kernel. Die initialen Werte des Kernels können mit zufälligen Zahlen initialisiert werden. [2]

Features

Features spielen eine wichtige Rolle bei der Verarbeitung des Inputs. Dies kann zum Beispiel eine Ecke, oder eine Kante (Edge) sein. Früher wurden diese Features von Hand programmiert. Mit dem populärer werden von NNs lässt man diese Features automatisch trainieren. [2] Features Extraktion ist oft der erste Schritt eines NNs. Nachfolgende Algorithmen, die auf Features angewandt werden, können nur so gut sein wie der initiale Feature Extractor. Ein Feature Extractor können CNNs sein, oft werden diese auf jedem Pixel eines Bildes angewandt werden. Dieses Ergebnis nennt man auch Feature Map. Features können in tieferen NNs zusammengefasst werden und zu neuen komplexere Features kombiniert werden.

Kernel ist die Fensterfunktion, die mit einer Dot Multiplication auf den Input angewendet wird. Die Werte eines Kernels, mit denen multipliziert wird, sind Weights welche beim trainieren des NNs durch Backpropagation gelernt werden. Ein Kernel kann

sowohl zum Upscaling wie auch zum Downscaling des Inputs verwendet werden.

Zero Padding hier werden die Grenzen eines Inputs (in den meisten Fällen ein Bild) mit Null-Werten aufgefüllt. Somit kann die Dimension des Outputs hochskaliert werden. Damit ist es möglich, dass Input und Output Dimension nach einer Convolution oder einem Pooling gleich bleibt. Neben dem Zero Padding gibt es noch das Grayscale Padding bei welchen auf dem Padding ein Rauschen angewendet wird.

Strides ist die Schrittweite, mit der ein Kernel (Fenster) über den Input geschoben wird. Ein Stride von 1x1 bedeutet, dass jeweils mit einem Pixel nach rechts geschoben wird. Ist man am Ende der Inputreihe angekommen, wird der Kernel um ein Pixel nach unten verschoben. Dieser Algorithmus wird, solange wiederholt, bis man ganz rechts in der letzten Reihe des Inputs angekommen ist. Wird der Stride erhöht, zum Beispiel auf 2x2, wird der Kernel dementsprechend um zwei Schritte verschoben. Dies kann dazu führen, dass der Output verkleinert wird, wenn kein dementsprechendes Padding verwendet wird.

Methodik

Im Fokus dieser Arbeit steht die Erkennung verschiedener Keypoints von Fahrzeugen.

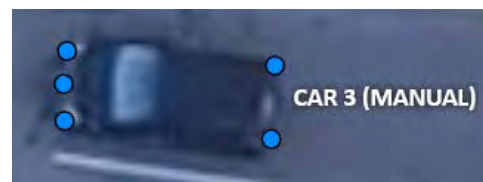


Abb. 1: Ein annotiertes Fahrzeug jeweils sind Keypoints als Front Center, Front Left, Front Right, Rear Left, Rear Right eingezeichnet [1]

Im Moment wird von YOLO nur die Pose Estimation von Menschen unterstützt. Im Laufe dieser Arbeit wird YOLO mit Top Down Bildern von Verkehrsflüssen trainiert.

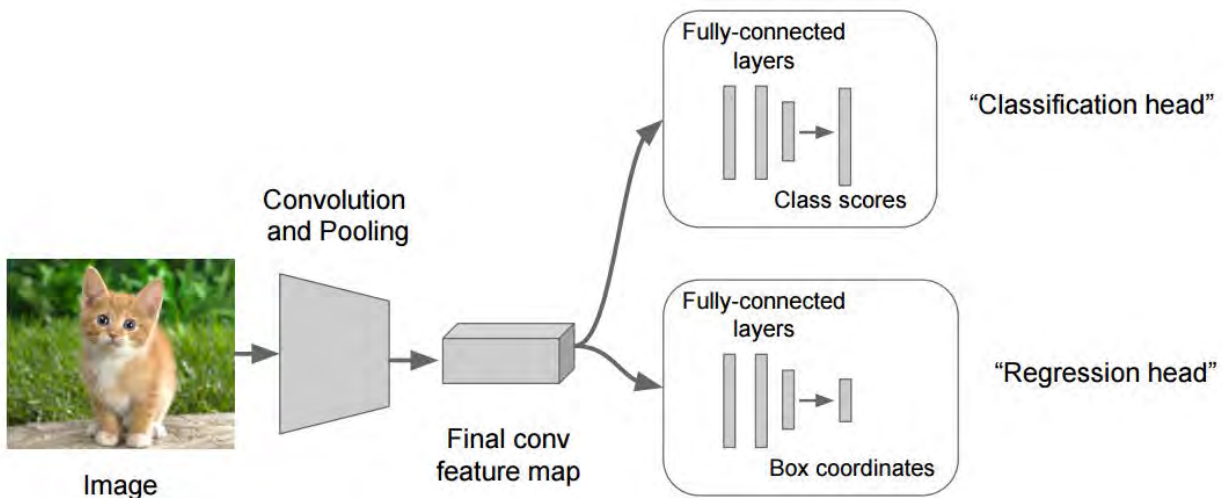


Abb. 2: Verschiedene Prediction Heads [5]

[5]

coco

Ein Ziel von Computer Vision ist das Verstehen von Visual Scenes. Diese Scene Erkennung baut auf anderen Funktionen wie z. B. Object Detection und Object Localization auf. Um diese Disziplinen zu trainieren, braucht es geeignete Datensätze. Diese Datensätze müssen eine möglichst vielschichtige Sichtweise von Views haben, ein View ist eine bestimmte Darstellung von einem Object. So kann ein Object z. B. ein Flugzeug mit verschiedenen Views dargestellt werden. Ein View kann das Flugzeug in der Luft zeigen, ein weiterer View zeigt das Flugzeug auf der Landebahn

oder ein mit Data Augmentation verändertes Flugzeug, welches um neunzig Grad gedreht wurde. In jedem Fall wird ein korrektes Labeling der Daten benötigt. [6]

Ausblick

CNNs ermöglichen eine effiziente Feature Extraction von Bildern. CNNs werden auch in dem populären YOLO Projekt verwendet und erreichen dort hohe Inferenz, Geschwindigkeit und Genauigkeit. Obwohl es nur wenige convolutional Layer benötigt, um eine effektive CNN Architektur zu entwickeln, gibt es keine eindeutige Strategie, wie diese Architektur für eine gute Performance aufgebaut sein sollte. [3]

Literatur und Abbildungen

[1] Eigene Darstellung.

[2] Goodfellow Ian, Bengio Yoshua, and Courville Aaron. *Deep Learning*. MIT Press, 2016.

[3] OShea Keiron and Nash Ryan. An Introduction to Convolutional Neural Networks. *arXiv*, 2015.

[4] McGregor Milecia. What Is a Convolutional Neural Network? A Beginner's Tutorial for Machine Learning and Deep Learning. *Internet*, 2021.

[5] bot sai. What is a multi-headed model? And what exactly is a 'head' in a model? <https://stackoverflow.com/questions/56004483/what-is-a-multi-headed-model-and-what-exactly-is-a-head-in-a-model>, 2019.

[6] Lin Tsung-Yi, Maire Michael, Belongie Serge, Bourdev Lubomir, Girshick Ross, Hays James, and Perona Pietro. Microsoft {COCO} Common Objects in Context. *arXiv*, 2014.

[7] Dumoulin Vincent and Visin Francesco. A guide to convolution arithmetic for deep learning. *arXiv*, 2018.

Evaluierung eines RISC-V Microcontrollers

Nick Suess

Clemens Klöck

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Steinbeis Embedded Systems Technologies GmbH, Esslingen

Einleitung

Das RISC-V Instruction Set wurde 2010 im Parallel Computing Laboratory der UC Berkeley von Prof. Krste Asanović, Yunsup Lee und Andrew Waterman gestartet. RISC-V ist eine öffentlich verfügbare und frei nutzbare Instruction Set Architecture (ISA). Hinter RISC-V steht die nicht-gewinnorientierte RISC-V Foundation, welche die Entwicklung des Instruction Sets leitet. Als Hauptgründe für die Notwendigkeit von RISC-V führen Krste Asanović und David Patterson an, dass eine öffentlich verfügbare ISA zu mehr Wettbewerb auf dem Prozessormarkt führt [1]. Mehr Wettbewerb sorgt für größere Innovationen, kürzere Entwicklungszeiten und niedrigere Preise. RISC-V ist keine neue Technologie, sondern vereint 40 Jahre an Erfahrung mit Instruction Set Architectures unter einem globalen, offenen Standard, mit welchem jeder eigene Hardware entwickeln kann [3]. Die Anforderungen an RISC-V können in vier Punkte aufgeteilt werden:

1. Erweiterbare ISA

Es gibt ein grundlegendes Instruktionsset, auf welchem anwendungsspezifische Erweiterungen dazu genommen werden können. Außerdem soll es möglich sein, eigene Instruktionen mit eigenem Opcode hinzuzufügen [1].

2. Kompakte ISA

Da IoT Prozessoren oft nur wenig Speicherplatz zur Verfügung gestellt bekommen, soll die ISA möglichst kurzgehalten werden [1].

3. Vierfach-, zweifach und einfache Präzision für Gleitkommazahlen

Damit die ISA auch für WSCs relevant bleibt, muss Vierfachpräzision für Gleitkommazahlen vorhanden sein [1].

4. 128-, 64 und 32-Bit Adressierung

32-Bit Adressierung ist Standard für IoT, 64-Bit ist de facto Standard für alles darüber. 128-Bit Adressierung wird aktuell nicht benötigt, aber wenn in der Zukunft 128-Bit Adressierung

genutzt wird, wäre es sehr aufwendig, eine ISA ohne 128-Bit Adressierung auf diese zu erweitern [1].

Problemstellung

Microcontroller und Mikroprozessoren werden immer häufiger in allen möglichen Anwendungen eingesetzt. Aktuell dominiert die Firma Arm Limited mit ihren Prozessordesigns den Markt für Embedded Microcontroller. Mit RISC-V existiert eine Alternative zu den marktdominierenden ARM-Prozessoren. Durch die Offenheit der ISA sollte zudem der Preis eines RISC-V Prozessors generell niedriger als der eines ARM-Prozessors sein. Jedoch ist nicht bekannt, ob RISC-V die gleiche Leistung wie ARM erbringen kann. Da die ARM-Architektur seit 1985 entwickelt wird, existieren Vorsprünge in Bezug auf Nutzer und Einsatzorte (und durch diese entstandene Erfahrung), Unterstützung und Optimierung durch Compiler und insgesamt existierender Prozessordesigns. Diese Bachelorarbeit soll herausfinden, ob RISC-V trotz dieser Hürden eine sinnvolle Alternative zu ARM-Prozessoren sein kann. Dazu ist an erster Stelle die Performance von RISC-V CPUs in verschiedenen Szenarien im Vergleich zu ARM CPUs interessant.

Benchmarking

Um die Performance von RISC-V Microcontrollern zu evaluieren, werden sie in mehreren Benchmarks mit ARM Microcontrollern verglichen. Dafür wurden drei Benchmarks durchgeführt. Im Primzahlbenchmark müssen die Microcontroller alle Primzahlen zwischen 1 und 30000 berechnen. Dieser Benchmark ist sehr rechenintensiv, benötigt aber nur sehr wenige Speicherzugriffe. Die Benchmarks Reglersimulation und Berechnung einer FFT sind durch große Datenmengen gleichzeitig rechen- und speicherintensiv. Für die Benchmarks auf ARM Microcontrollern wurden die Entwicklungsboards NUCLEO-U575ZIQ (Cortex-M33 STM32U575) und NUCLEO-H743ZI2 (Cortex-M7 STM32H743) ausgewählt. Die Bench-

marks für das NUCLEO-H743ZI2 Entwicklungsboard wurden außerdem mit aktiviertem Cache und ohne Cache durchgeführt. Für die Benchmarks mit RISC-V Microcontrollern wurden die Entwicklungsboards ESP32C3-DevKitC-02 (ESP32C3) von Espressif und Maix Bit (Kendryte K210) von Sipeed eingesetzt. Alle Benchmarks wurden mit den Optimierungsoptionen -O0, -Og, -Os und -O2 durchgeführt, es werden

jedoch nur die schnellste Optimierungsoption per Microcontroller dargestellt. Ein niedriges Ergebnis in diesen Benchmarks bedeutet eine schnelle Laufzeit und somit gute Performance. Bei 400 MHz fehlen die Ergebnisse von ESP32C3 und STM32U575, da diese Microcontroller auf eine maximale Taktfrequenz von 160 MHz begrenzt sind.

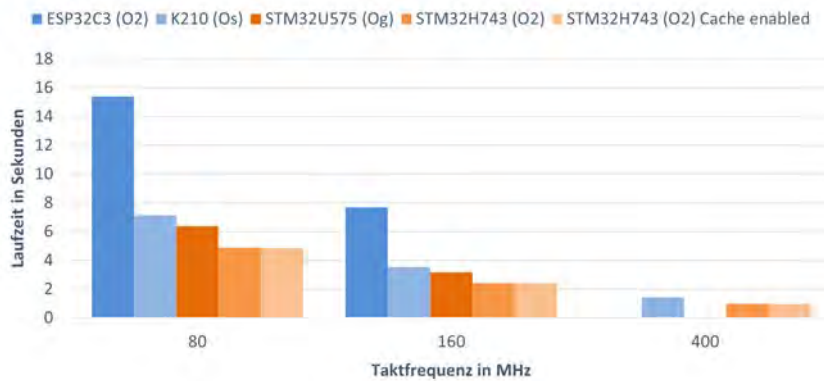


Abb. 1: Primzahlbenchmark [2]

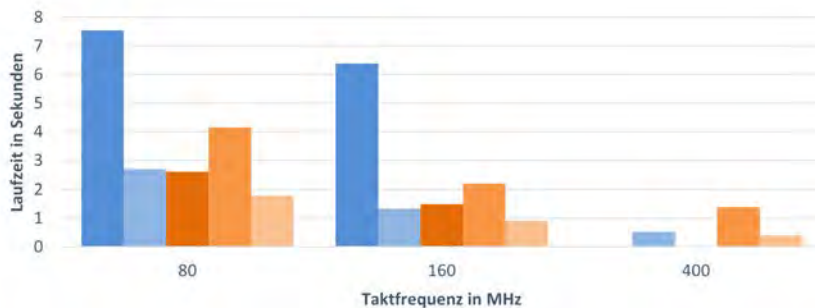


Abb. 2: Reglersimulation [2]

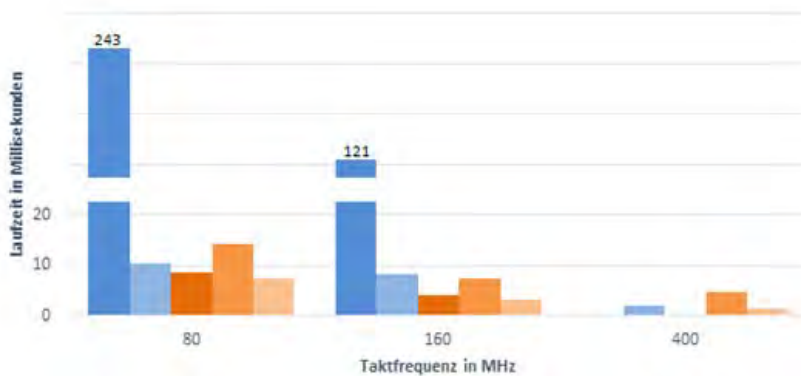


Abb. 3: FFT-Berechnung [2]

Es wurden ebenfalls Benchmarks mit RISC-V und ARM-Microprozessoren durchgeführt. Diese sind leistungsstärker als die Microcontroller und verwenden daher eine Embedded Linux Distribution als Betriebssystem. Für das Benchmarking dieser Microprozessoren mit Betriebssystem wird die Phoronix Test Suite (<https://www.phoronix-test-suite.com/>) auf dem Raspberry Pi 4B (Cortex-A72) und VisionFive2 (RISC-V JH7110) eingesetzt. In diesen Benchmarks indiziert ein höheres Ergebnis bessere Performance. Die Ergebnisse sind normalisiert, um die relative Performance in einem Diagramm darstellen zu können.



Abb. 4: Phoronix Benchmarks [2]

Ergebnis und Ausblick

Im Vergleich der Microcontroller kann der Kendryte K210 gut mit den ARM-Microcontrollern mithalten. Zwar zeigt sich im Primzahlbenchmark, dass die reine Rechenleistung des K210 schwächer ist als die der beiden ARM Vertreter. In realen Anwendungen mit Speicherzugriffen, die durch die Reglersimulation und die FFT-Benchmark besser simuliert werden, kann sich der K210 jedoch sogar dem Top-Modell der Cortex-M Serie, dem Cortex-M7, annähern. Der ESP32C3 ist in allen Benchmarks bei den Taktfrequenzen 80 MHz und 160 MHz der langsamste Microcontroller. Der Grund hierfür ist eine Kombination aus sehr kleinem

Cache, Fokus auf Energieeffizienz und eingeschränktem Befehlssatz. Der eingeschränkte Befehlssatz zeigt sich besonders negativ im FFT-Benchmark. In diesem Benchmark werden praktisch ausschließlich mathematische Operationen auf Gleitkommazahlen durchgeführt. Da dem ESP32C3 die Befehlssatzerweiterung für Gleitkommazahlen fehlt, muss bei jeder mathematischen Operation eine Bibliothek aufgerufen werden, welche die fehlende Erweiterung ersetzt. Dieser Aufruf sorgt für massive Performanceeinbuße, der durch die große Datenmenge nochmals verstärkt wird. Auch der VisionFive2 wird in allen Prozessor-Benchmarks vom Raspberry Pi 4B geschlagen. Im CoreMark Benchmark zeigt sich dies in einem Unterschied von etwa 13000 Punkten, womit der VisionFive2 eher die Performance des Vorgängers des Raspberry Pi 4s, dem Raspberry Pi 3, erreicht. Nur im Systembenchmark Redis GET kann sich der VisionFive2 durchsetzen. Insgesamt gibt RISC-V ein gemischtes Bild ab. Zwar werden der ESP32C3 und der VisionFive2 sehr eindeutig in den meisten Benchmarks geschlagen, der K210 erreicht dafür bei hohen Taktfrequenzen eine ähnliche Performance wie ein Cortex-M7 Prozessor. Die Vorteile des ESP32C3 liegen in anderen Bereichen. Er ist mit einem Stückpreis von 1€ günstiger als alle ARM-Microcontroller. Gleichzeitig bietet er viele Konnektivitätsoptionen über WIFI und Bluetooth. Zusätzlich besitzt er mehrere Energiesparoptionen und ist selbst bei voller Leistung sparsamer als alle anderen Microcontroller, mit Ausnahme des STM32U575. Zusammenfassend lässt sich sagen, dass RISC-V im Bereich Embedded Microcontroller eine sinnvolle Alternative zu ARM bildet, wenn der genutzte Befehlssatz zu den Anforderungen der Anwendung passt, oder Priorität auf Kosten, Konnektivität und Energieeffizienz gesetzt wird. Im Bereich Microprozessoren liegt RISC-V aktuell noch weit hinter ARM. Jedoch ist RISC-V auf diesem Markt aktuell nur sehr wenig vertreten, in Zukunft sind aber Smartphones und Tablets mit RISC-V Prozessoren geplant. Wenn diese neue Generation RISC-V Prozessoren auf den Markt kommt, sollte sich der Performanceunterschied im Vergleich zu ARM reduzieren.

Literatur und Abbildungen

- [1] Krste Asanović and David A. Patterson. Instruction Sets Should Be Free: The Case For RISC-V. <https://www2.eecs.berkeley.edu/Pubs/TechRpts/2014/EECS-2014-146.pdf>, 08 2014.
- [2] Eigene Darstellung.
- [3] RISC-V Foundation. History of RISC-V. <https://riscv.org/about/history/>, 2021.

Entwicklung eines künstlichen neuronalen Netzes für das Erkennen von Rissen auf Fotos von Mikrochips

Murat Tekin

MarkusENZweiler

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Robert Bosch GmbH, Reutlingen

Einleitung

Im Zeitalter der Digitalisierung werden Tag für Tag immer mehr Daten erzeugt, aus denen Erkenntnisse gewonnen werden können. Mit der steigenden Anzahl an Daten hat sich in der Informatik zunehmend auch der Bereich der Künstlichen Intelligenz weiterentwickelt, der auf große Datenmengen angewiesen ist. Ein Teilbereich der Künstlichen Intelligenz ist das Deep Learning, welches sich auf künstliche neuronale Netze fokussiert. Künstliche neuronale Netze sind dem menschlichen neuronalen Netz nachempfunden und können komplexe Problemstellungen lösen, die für den Menschen auf den ersten Blick unmöglich erscheinen [1]. Sie spielen bereits in vielen verschiedenen Bereichen eine wichtige Rolle. Ein spezifischer Aufgabenbereich von künstlichen neuronalen Netzen ist die Image Classification. Hierbei werden dem Netz Bilder übergeben, die es anschließend verschiedenen Klassen zuordnet. In der Wirtschaft kann die Image Classification zur Qualitätskontrolle verwendet werden, um bspw. Produktionsfehler automatisiert erkennen zu können.

Ziel der Arbeit

Das Ziel der Arbeit ist es ein künstliches neuronales Netz zu entwickeln, mit dessen Hilfe Anomalien auf Mikrochips erkannt werden können. Das Netz soll dazu Bilder von Mikrochips einer positiven und einer negativen Klasse zuordnen. Hierzu wird ein sogenanntes Convolutional Neural Network (CNN) zum Einsatz kommen, eine Architektur, die besonders für Bildklassifizierungen geeignet ist. Damit das CNN präzise Vorhersagen treffen kann, muss es mit einem Datensatz angelernet werden. Hierbei soll untersucht werden, wie der Datensatz vorverarbeitet werden kann, um die Genauigkeit des CNNs zu erhöhen.

Künstliche neuronale Netze

Künstliche neuronale Netze bestehen aus nacheinander geschalteten Schichten, oftmals Layer genannt. Dabei werden zwischen drei Arten von Layern unterschieden: der Eingabeschicht, den Hidden-Layern und der Ausgabeschicht. Jede dieser Schichten besteht aus einer Vielzahl an Neuronen, die miteinander verbunden sind. Daten werden über die Eingabeschicht in das neuronale Netz übergeben, wo sie nach und nach in sinnvollere Formen überführt werden. In Abbildung 1 ist ein solches klassisches neuronales Netz dargestellt. Ein neuronales Netz besitzt Parameter, die innerhalb eines Lernprozesses angepasst werden müssen, um präzise Vorhersagen treffen zu können. Dazu muss dem Netz eine große Menge an Daten übergeben werden.

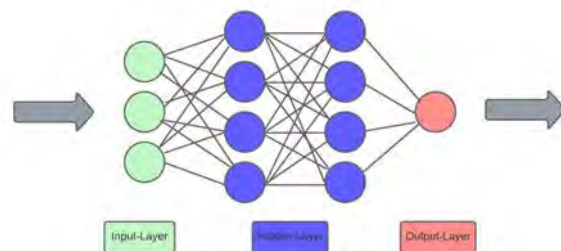


Abb. 1: Aufbau eines künstlichen neuronalen Netzes [3]

Convolutional Neural Networks

Convolutional Neural Networks sind eine spezielle Architektur eines künstlichen neuronalen Netzes. Die Architektur eines CNN ist besonders für Aufgaben geeignet, in denen Pixel verarbeitet werden. Mit ihrer Hilfe lassen sich Objekte auf Bildern identifizieren und erkennen. In Abbildung 2 ist der Aufbau eines CNN illustriert. Ein CNN kann in zwei Blöcke untergliedert werden. Hierbei ist der erste Block für die Feature Extraction zuständig, wohingegen im zweiten Block die eigentliche Klassifizierung stattfindet. Bei der Feature

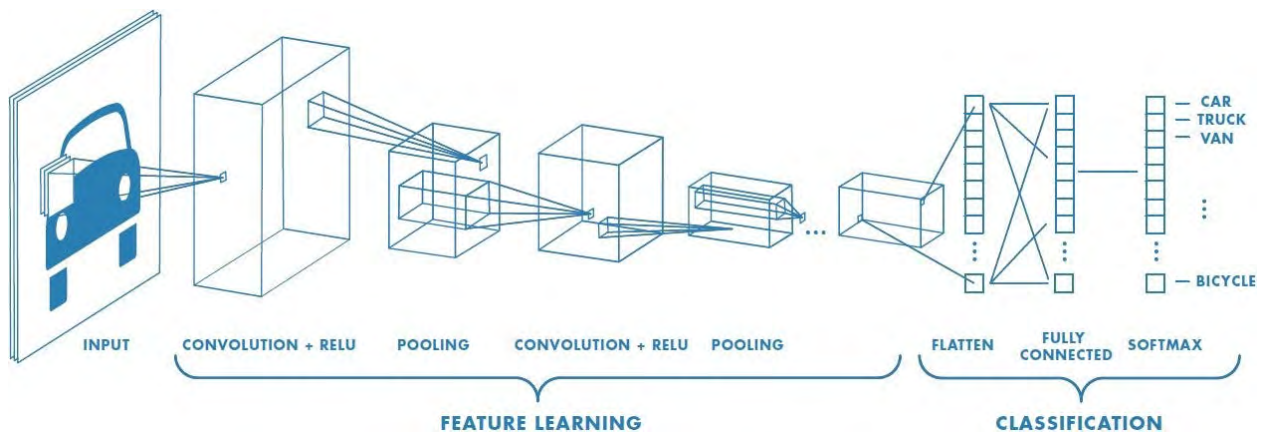


Abb. 2: Aufbau eines Convolutional Neural Network [4]

Extraction werden Merkmale aus den Bildern extrahiert, die über die Schichten hinweg zu komplexeren Strukturen kombiniert werden [2]. Mathematisch wird dazu eine Faltung der Eingabe mit sogenannten Filtern durchgeführt. Ein Filter ist eine zweidimensionale Matrix, dessen Werte als Muster interpretiert werden können. Anschließend wird in der Pooling-Schicht die Auflösung der Merkmale reduziert, wodurch die Anzahl der lernbaren Parameter verringert wird. Nachfolgend werden die Merkmale in den zweiten Block übergeben, der aus voll-verbundenen Schichten besteht. In der Ausgangsschicht wird eine Aktivierungsfunktion verwendet, die Auskunft über die Klassifizierung gibt. Für die Messung der Leistung eines CNNs werden verschiedene Metriken verwendet. Die am häufigsten verwendete Metrik ist die Genauigkeit, die den Anteil korrekter Vorhersagen misst.

Vorgehensweise

Eine exemplarische Vorgehensweise für die Entwicklung eines CNNs ist in Abbildung 3 zu erkennen. In einem ersten Schritt wird der vorhandene Datensatz vorverarbeitet. Dazu werden verschiedene Methoden der Bildverarbeitung angewandt. Dies können zum Beispiel Filter sein, um Rauschen auf den Bildern zu eliminieren und sie zu schärfen. Außerdem kann eine Data Augmentation durchgeführt werden, bei der durch das Rotieren oder das Spiegeln der Bilder, der Datensatz künstlich erweitert werden kann. Dadurch stehen mehr Daten zur Verfügung von denen das Netz lernen kann. Die Vorverarbeitung ist ein nicht zu vernachlässigender Schritt bei der Entwicklung eines neuronalen Netzes. Er hat einen großen Einfluss auf die Genauigkeit der Vorhersagen. In einem nächsten Schritt wird ein erstes Modell entworfen und mit dem vorverarbeiteten Datensatz trainiert. Für die Implementierung existieren eine Reihe an Frameworks in der Programmiersprache Python, die vordefinierte Funktionen bereitstellen. Es folgt darauf ein Kreislauf

aus Modellanpassung und Evaluierung. Ein CNN hat viele verschiedene Hyperparameter, die angepasst werden können. Dazu gehören beispielsweise die Anzahl der Convolutional Layer, die Wahl einer Aktivierungsfunktion oder das Bestimmen eines Optimierers. Das Training eines neuronalen Netzes ist rechenintensiv und kann bei schwacher Hardware sehr viel Zeit in Anspruch nehmen. Ziel ist es, diesen Prozess zu wiederholen, bis das entwickelte Modell zufriedenstellende Vorhersagen trifft.

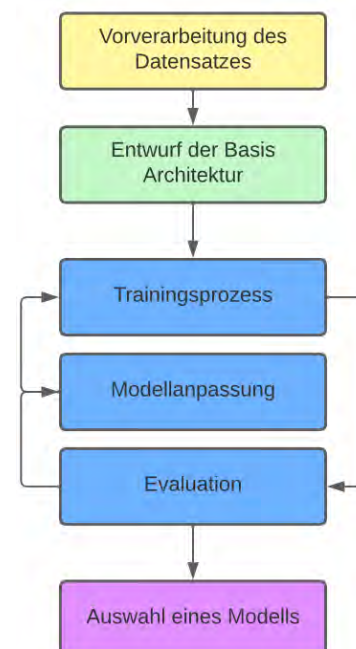


Abb. 3: Herangehensweise für die Entwicklung eines künstlichen neuronalen Netzes [3]

Ausblick

Zum Zeitpunkt der Erstellung dieses Artikels besteht bereits ein Modell zur Klassifizierung der Mikrochips. Bis zur Fertigstellung der Arbeit wird weiter an der Erhöhung der Genauigkeit des Modells gearbeitet. Auf

Basis dieses Modells könnte sich eine Folgearbeit damit beschäftigen die Anomalien auf den Mikrochips zusätzlich hervorzuheben. Desweiteren könnten die in der Arbeit vorgestellten Methoden in der Zukunft für weitere Arten von Mikrochips mit denselben Eigenschaften angewandt werden.

Literatur und Abbildungen

- [1] Charu C. Aggarwal. *Neural networks and deep learning*. Springer, 2018.
- [2] Francois Chollet. *Deep learning with Python*. Simon and Schuster, 2021.
- [3] Eigene Darstellung.
- [4] Sumit Saha. A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way. <https://saturn-cloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/>, 12 2018.

V2X Security: Analysis and Implementation

Panagiotis Theofanidis

Harald Melcher

Department of Computer Science and Engineering, Esslingen University

Work carried out at nfiniity GmbH, Esslingen

Introduction

Vehicle-to-Everything (V2X) technology is a major step towards improving safety and efficiency on our roads. Based on the exchange of information between vehicles and road infrastructure (also called C-ITS stations), such as position and speed, algorithms can be developed that can, for example, provide the optimal speed to avoid red traffic lights or give the driver a warning if there is a risk of collision with another vehicle (see Figure 1).



Fig. 1: V2X Safety Applications [1]

The benefit of V2X technology in these cases is that it can provide information about the environment that is not visible to the driver, thanks to the use of wireless communication.

Widespread V2X adoption has begun in recent years, primarily at Volkswagen [2], and will see a significant increase in the coming years as Euro NCAP has announced that V2X technology will be included in the safety rating scheme in their 2025 roadmap [5].

A critical aspect of V2X technology is security, as the exchanged information can be used to influence the behavior of drivers and infrastructure. This aspect is mostly overlooked in the current state of the art open source V2X communication stacks [4]. Closed source implementations are available, but those come with the drawbacks of blackbox behavior, vendor lock-in, lack of customizability and high costs.

The aim of this thesis is to analyze the solutions proposed by different international standards and to implement missing security features in the open *Vanetza* library.

Analysis

Integrity, sender authentication and authorization, replay protection and availability are the most important aspects of security in V2X communication. Critical threats are the denial of transmission/reception of data, modification/deletion of information, masquerading of a station and acquisition of personal information.

To ensure the effectiveness of all the following measures, system and software integrity needs to be ensured by securing the boot process and protecting cryptographic material against manipulation and theft through adequate hardware.

Confidentiality plays a secondary role for the applications that use V2X communication. One exception is the management of identities and privileges of stations.

To facilitate this process, a Public Key Infrastructure (PKI) is used to manage the certificates of all stations. The V2X specification ratified by the European Union, the ETSI-ITS suite [3], uses a PKI with components depicted in Figure 2.

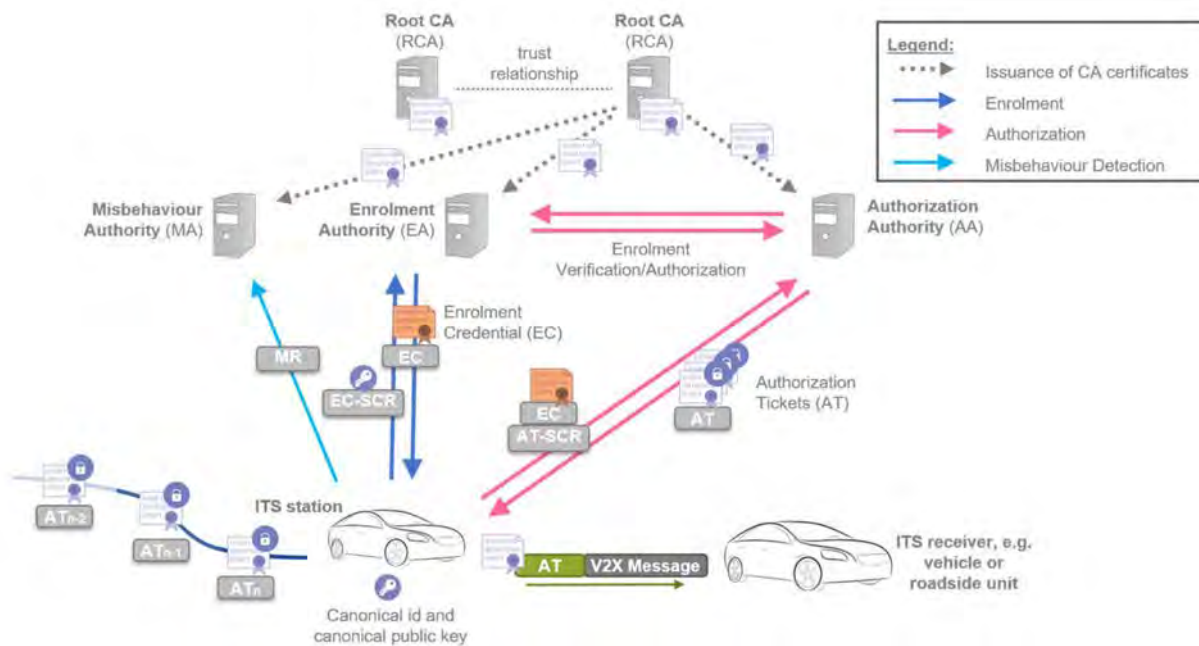


Fig. 2: ETSI PKI Architecture [3]

- The Trust List Manager (TLM) keeps track of all RCAs and combines and distributes a Certificate Trust List (CTL). This is done on a European level.
- Root CAs (RCA) form the trust anchors (highest level) in the PKI.
- Enrolment Authorities (EA) are used to register C-ITS stations, granting them a pseudonym Enrolment Credential (EC).
- Authorization Authorities (AA) issue Authorization Tickets (AT) to stations with a valid EC.

AA and EA have to be operated by separate entities to ensure privacy. AAs lack the information to link a pseudonym EC to the real identity of a C-ITS station, while EAs never get to see the Authorization Ticket. The station should change its IP address between each AT request so the AA cannot link multiple ATs to a single station.

A station should have multiple AT to rotate through, depending on time and moved distance. These should not be able to be linked back to the same owner as to not enable the tracking of a vehicle through collection of many messages over time.

When a station wants to send a message, it adds a security header and signs the packet using one of the available ATs. The receiver can then verify the signature and check the validity of the AT. The receiver can ask the sender for the AA certificate if necessary. Authorization is implemented by appending the authorized Application IDs to the certificates.

A timestamp is added to the security header to avoid replay attacks. A message can then be rejected, if the timestamp is too far in the past. For messages with longer lifetimes, hashes of previous messages can be compared to incoming ones. These methods can be combined.

By applying these security measures on network and transport layer, the radio-level technology can be independently chosen. The current industry standards are IEEE 802.11p and 3GPP LTE-V2X.

To revoke the trust in the PKI, different methods are used for each type of entity based on the longevity of their certificates:

- ITS Station: Short AT lifespan and blacklist in the EA
- Authorization Authority: Certificate Revocation List
- Root Certificate Authority: Removal from Certificate Trust List

Implementation and Outlook

The Vanetta library is a C++ implementation of the ETSI-ITS specification. Development on it initially started at the Institute of Telematics of the Karlsruhe Institute of Technology.

Vanetta already implements a lot of the security features described in the previous section. Signature generation and verification, security headers and certificate handling are already implemented.

The protocols for the generation of Enrolment Credentials and Authorization Tickets are missing and currently being implemented. This includes support for the Elliptic Curve Integrated Encryption Scheme (ECIES), a hybrid encryption scheme that uses Elliptic Curves for key exchange and AES for message encryption. The procedure for the initial registration of a station at the Enrollment Authority is not part of the standard and is handled by each PKI implementation individually.

Requests for Enrolment Credentials and Authorization Tickets are defined in ASN.1, a language for the description of data structures. This comes with the

advantage of being able to use an ASN.1 compiler to generate the necessary C++ code for the serialization and deserialization of the data structures.

To manage the automatic renewal and change of ECs and ATs, new components have to be implemented that handle the logic for the storage and rotation of the credentials.

Additionally, the download and parsing of CRLs and the European Certificate Trust List have to be implemented.

The implementation will be tested on two testbeds that support both IEEE 802.11p and LTE-V2X, which nfiniity has graciously provided for this thesis.

References and figures

- [1] Connected Motorcycle Consortium. CMC Applications Phase 1. https://www.cmc-info.net/uploads/1/2/1/4/121453783/overview05_orig.jpg, 01 2023.
- [2] Huanyu Gu. NXP, Volkswagen and Partners Continue to Accelerate the V2X Rollout. <https://www.nxp.com/company/blog/nxp-volkswagen-and-partners-continue-to-accelerate-the-v2x-rollout:BL-THE-V2X-ROLLOUT>, 07 2021.
- [3] European Telecommunications Standards Institute. Intelligent Transport Systems (ITS); Security; ITS communications security architecture and security management; Release 2. https://www.etsi.org/deliver/etsi_ts/102900_102999/102940/02.01.01_60/ts_102940v020101p.pdf, 07 2021.
- [4] Sven Laux et al. Demo: OpenC2X — An open source experimental and prototyping platform supporting ETSI ITS-G5. In *2016 IEEE Vehicular Networking Conference (VNC)*. Institute of Electrical and Electronics Engineers, 2016.
- [5] Euro NCAP. Euro NCAP 2025 Roadmap. <https://cdn.euroncap.com/media/30700/euroncap-roadmap-2025-v4.pdf>, 09 2017.

Konzeption und Realisierung eines vollautomatisierten, hochpräzisen Sensor-Kalibrator mit einem Datenbanksystem

Mladen Tramosljika

Clemens Klöck

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma comemso electronics GmbH, Ostfildern

Motivation

Elektromobilität, auch als E-Mobility bekannt, ist ein zentrales Thema des 21. Jahrhunderts, besonders in der Automobilbranche. Dieses Paradigma spiegelt sich in der mentalen Umorientierung der Bevölkerung, der medialen Präsenz und den einschlägigen EU-Gesetzen [4] wider. Der rasante und exponentielle Zuwachs bei Neuzulassungen von Elektroautos in Deutschland verdeutlicht diesen Trend.

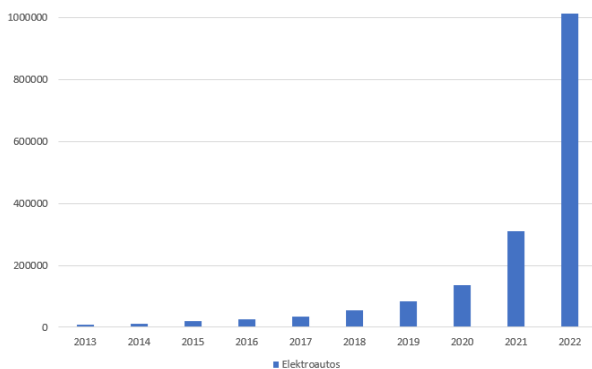


Abb. 1: Anzahl Elektroautos in Deutschland 2023 [3]

Im Kontext der Elektromobilität sind die Reichweite und die Ladeeffizienz zukünftiger Elektroautos von großer Bedeutung, ebenso wie adäquate Batteriezellen und Batterie-Management-Systeme (BMS). Das BMS als Batterie-Steuergerät hat die Aufgabe die Batteriezellen langlebig, gesund und sicher zu erhalten. Doch das Testen solcher BMS mit realen Zellen ist gefährlich, nicht reproduzierbar und nicht automatisierbar. Deswegen nutzt man sogenannte „virtuelle Batteriezellen“. Und diese sind in Batteriezellen-Simulatoren (kurz BZS) realisierbar. BZS lösen alle drei der genannten Schwierigkeiten: Ihre elektrischen Eigenschaften sind parametrierbar. Somit können alle

notwendigen Zustände und Störungen simuliert werden, aber auch Algorithmen für die Langlebigkeit mittels Ausgleichsvorgängen der Zellen (sog. Balancing) des BMS verifizieren. Und dank hochpräzisen Strommessung des BZS lassen sich Leckströme am BMS frühzeitig erkennen, welche eine Tiefentladung und eine Beschädigung der Batteriezellen vorbeugen [2]. Trotzdem bedürfen diese BZS einer hohen Präzision bei Kalibrierung und Justage. Der Werkskalibrierprozess wurde seither zwar selbst automatisiert, jedoch noch nicht mit dem hausinternen ERP verbunden und auch im Fehlerfall fehlte seither ein durchgängiger und benutzerfreundlicher Ablauf. Ziel der Arbeit Das Ziel dieser Bachelorarbeit ist es: - Die Anforderungen an das zu erstellende Modul zu realisieren - Die für die Anforderungen gemessenen Daten zu verarbeiten - Eine Konzeption für das Modul zu erstellen - Die kalibrierten Werte intern in eine Datenbank speichern

Vorgehensweise

Bei der Entwicklung eines neuen Moduls ist es essenziell, Anforderungen zu ermitteln und das Modul in die firmeninterne Kalibriersoftware zu integrieren und zu testen. Die Rolle von Recherche, Analyse und Einarbeitung ist dabei entscheidend. Ein Sensor-Kalibrator prüft die Parameter jeder BZS-Karte. Wenn die Werte außerhalb des gewünschten Bereichs liegen, werden sie entsprechend angepasst. Ursprünglich hatte der Kalibrator die Hauptsoftware für Automatisierung und Datenspeicherung, während der PC nur für Eingabe und Bedienung genutzt wurde. Doch um Benutzerfreundlichkeit, Automatisierung und Datenbankbindung zu verbessern, war eine Überarbeitung der Softwarearchitektur notwendig, um die Programmierung stärker auf die PC-Software zu verlagern (siehe Abbildung 2).

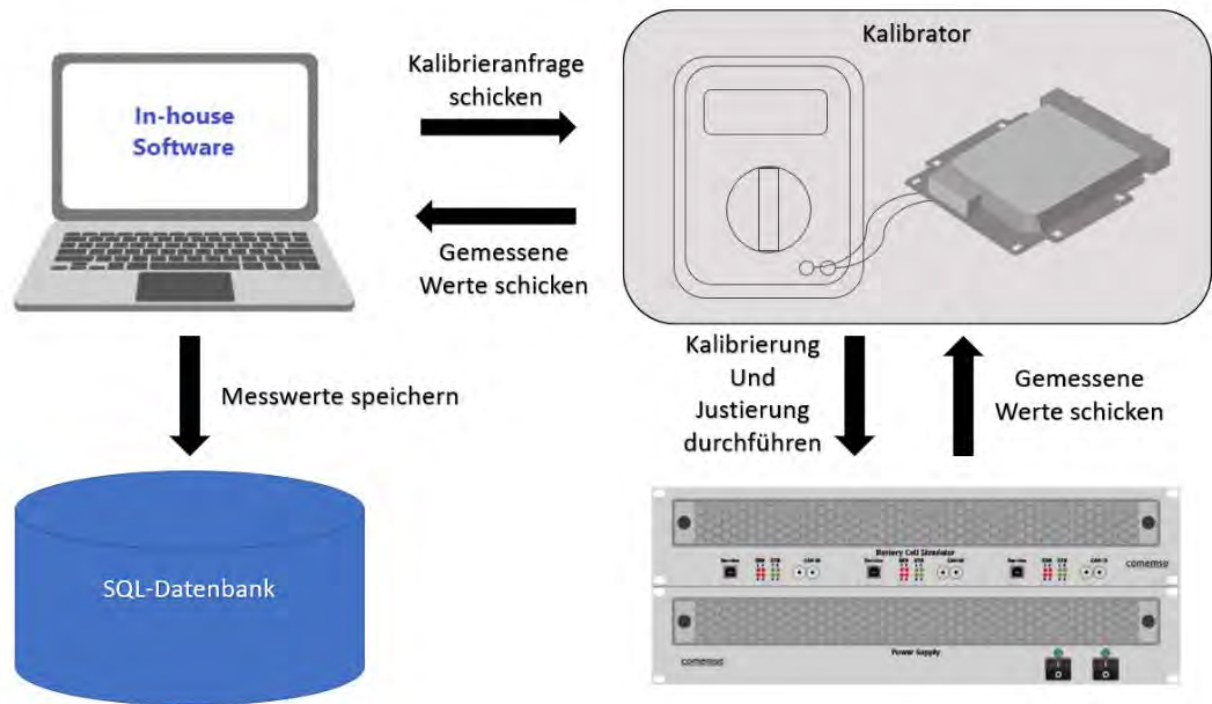


Abb. 2: Konzept der Automatisierung [1]

Die Planungsphase zur Automatisierung setzt auf objektorientierte Programmierung und die Erweiterung grafischer Nutzeroberflächen. Das Ziel ist es, die Automatisierung für Nutzer auf der Oberfläche sichtbar und einfach zu bedienen zu machen. Zusätzlich wird eine Schnittstelle, eine sogenannte API, für die Datenbank programmiert. Dies ermöglicht die unkomplizierte und schnelle Integration neuer Geräte und ihrer spezifischen Kalibrierprozesse in die Datenbank. Ein weiterer wichtiger Aspekt der Planung ist die Implementierung einer zyklischen Temperaturüberwachung. Da der Kalibrierprozess gemäß ISO 17025 eine Aufwärmphase erfordert, wird diese ebenfalls automatisiert, was zu einem optimalen Kalibrierergebnis beiträgt und den gesamten Prozess erleichtert.

Fazit und Ausblick

Mit der Realisierung der höheren Automatisierung wird nicht nur der Prozess sicherer und einfacher gestaltet, sondern auch die Effizienz des Kalibrierprozesses gesteigert. Die Daten werden in einer zentralen Datenbank gespeichert, was zu einfachen Kalibrierreports mit Darstellungsmöglichkeiten nach Kundenwunsch, sowie firmenintern zu langfristigen Auswertungen und somit weiteren Optimierungsmöglichkeiten führt. Und zu guter Letzt kann bei unvorhergesehenen Fehlern im Kalibrierprozess trotzdem der Mitarbeiter eingreifen, da der teilautomatisierte Modus weiterhin vorhanden bleibt. Wie schon in der Vorgehensweise erwähnt, können dank der API zukünftige Geräte auch die Datenbankschnittstelle nutzen, um ihre kalibrierten Messwerte im ERP-System abzuspeichern und somit jederzeit in der entsprechenden Abteilung für weitere Auswertungen nutzbar sind.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] comemso GmbH. BMS Testing with Battery Cell Simulator (BCS). <https://www.comemso.com/battery-cell-simulation-details>, 2023.
- [3] Statista GmbH. Anzahl der Elektroautos in Deutschland von 2012 bis 2022. <https://de.statista.com/statistik/daten/studie/265995/umfrage/anzahl-der-elektroautos-in-deutschland/>, 2023.
- [4] L. Gutsch. E-Mobilität in Deutschland -Einflussfaktoren auf die Marktdurchdringung. https://www.iwi.uni-hannover.de/fileadmin/iwi/Abschlussarbeiten/BA_Gutsch_K.pdf, 2020.

Einsatz von Bildverarbeitung zur Qualitätskontrolle von Fertigteilbehältern

Markus Treitz

Dirk Hesse

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Schuler Pressen GmbH, Gemmingen

Einleitung

Die Industrie der Moderne (Industrie 4.0) ist in hohem Maße auf Lagersysteme angewiesen, die eine Vielzahl von Teilen aus Blech aufnehmen müssen. Diese Lagersysteme sind häufig mit Fertigteilhaltern (im folgenden Klappen genannt) oder Fächern ausgestattet, welche die Teile in einer stabilen Position lagern und sie somit vor Beschädigung sichern. Die effiziente Erkennung und Verfolgung dieser dynamischen Klappen stellt jedoch eine erhebliche Herausforderung dar. Die manuelle Überwachung kann zu Fehlern führen und ist ineffizient, während herkömmliche automatisierte Methoden den Zustand des Lagersystems aufgrund seiner dynamischen Beschaffenheit möglicherweise nicht genau erfassen. Daher wird eine Lösung benötigt, die diese Klappen nicht nur mit hoher Genauigkeit erkennen, sondern sich auch in Echtzeit an ihre Veränderungen anpassen kann. An dieser Stelle kommt das Konzept der Bildverarbeitung ins Spiel, das eine technologiegestützte Lösung für ein weit verbreitetes Problem in der industriellen Lagerhaltung verspricht.

Bildverarbeitung

Der erste Schritt in dem Prozess der Bildverarbeitung ist die Aufnahme von Bildern der Fertigteilbehälter. Hochauflösende Kameras werden strategisch so positioniert, dass sie alle Winkel der Behälter abdecken und sicherstellen, dass keine Klappe unbeobachtet bleibt. Um die besten Ergebnisse zu erzielen, mussten gleichmäßige Lichtverhältnisse gewährleistet sein.

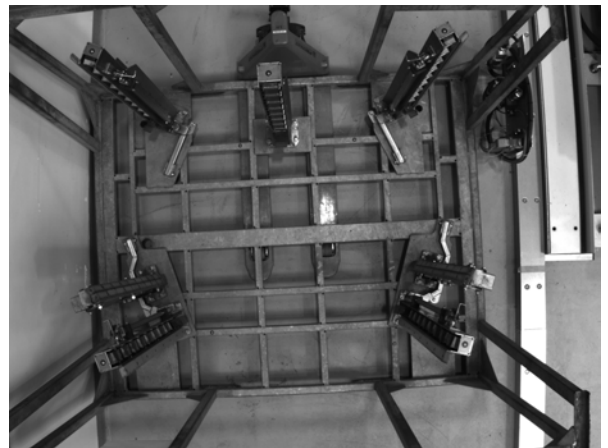


Abb. 1: Fertigteilbehälter für Motorhauben [1]

Für den Test wurde ein Behälter verwendet, der Motorhauben aufnimmt (Abbildung 1) und diese in ihrer Position sichert. Der Behälter besitzt 4 Metallsäulen, in welchen sich die Klappen befinden, die zum Sichern der Fertigteile verwendet werden.

Lösungsansatz

Mit Hilfe von normalisierter Kreuzkorrelation können die Klappen auf den Säulen erkannt werden. Im Anschluss kann verglichen werden ob auf jeder Säule die gleiche Anzahl an Klappen erkannt wurde. Sollte die Anzahl der Klappen auf den Säulen abweichen, kann von einem Fehler ausgegangen werden.

Bereich von Interesse (Region of Interest)

Das Bild besteht aus vielen Regionen, welche für die spätere Verarbeitung unwichtig sind. So kann eine Klappe nicht auf dem Boden oder in der Luft vorkommen. Deshalb wird das Bild reduziert auf die Regionen, welche die Klappen enthalten können. In dem Beispiel sind das die 4 Metallsäulen in denen

sich die Klappen befinden. Die Verwendung einer Region of Interest ist in diesem Zusammenhang von entscheidender Bedeutung, da sie die Rechenlast des Algorithmus erheblich reduziert.

Modelerstellung

Das Modell ist im Wesentlichen die Vorlage, mit der der Algorithmus Regionen des Bildes vergleicht. In diesem Fall wäre die Modellregion ein Bild einer Klappe des Behälters. Im Idealfall sollte die Modellregion eine typische Klappe so genau wie möglich darstellen. [2]

Normalisierte Kreuzkorrelation

Nachdem die Modellregion erstellt wurde, kann mit der Phase der Klappenerkennung fortgefahren werden. Es wird die Technik der normalisierten Kreuzkorrelation (NCC) verwendet, die die Ähnlichkeit zwischen der Modellregion und Teilen des Bildes misst. Der NCC-Algorithmus verschiebt das Modell über das Bild und berechnet die Korrelation für jede Position innerhalb der definierten Region of Interest. Der Wert befindet sich dabei immer zwischen $[-1;1]$ wobei ein Wert von 1 eine genaue Übereinstimmung bedeutet. [3]

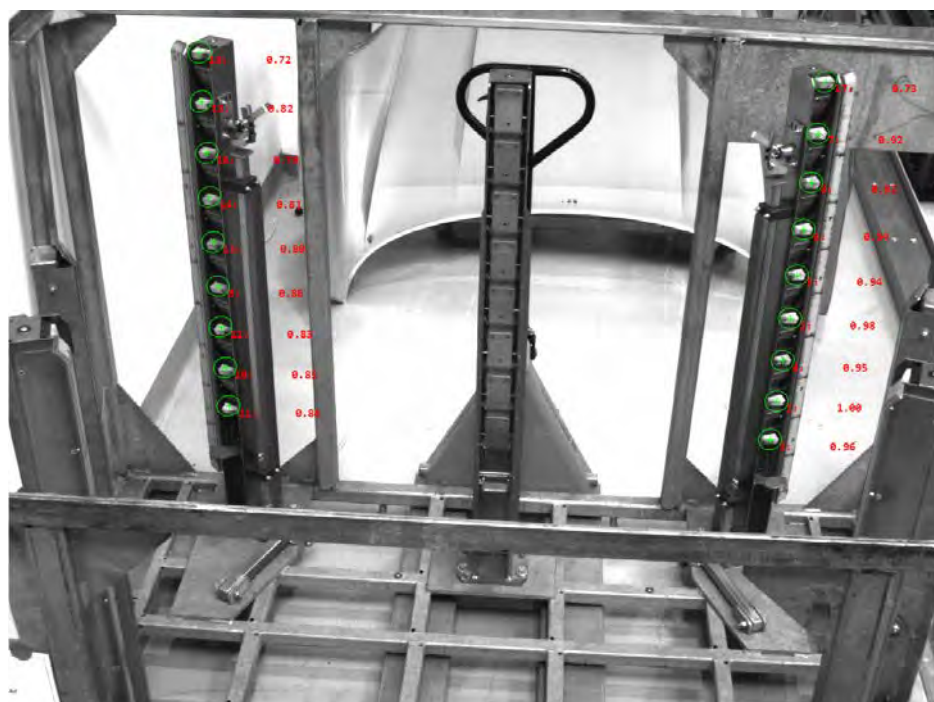


Abb. 2: Erkannte Klappen eines Behälters [1]

Ein Test (Abbildung 2) zeigt, dass alle Klappen gefunden werden können. Erste Probleme zeichnen sich jedoch bei den Korrelationswerten ab. Die Klappe mit der Nummer 18 ist sehr verschmutzt und wird deshalb lediglich mit einem Wert von 0,72 erkannt. Würde diese Klappe noch mehr verschmutzen, wird sie in der Zukunft nicht mehr erkannt.

Ausblick

Die Verwendung von Matching - Algorithmen kommt bei dem Problem der Klappenerkennung an die Grenzen. Sollten die Klappen sich überlappen, oder wie bereits erwähnt farblich verändern, können Matching-Algorithmen diesen Unterschied nicht erkennen. Mit dem Einsatz von Neuronalen Netzen könnte dieses Problem in der Zukunft gelöst werden.

Literatur und Abbildungen

[1] Eigene Darstellung.

[2] Vipin Sharma. Image-Template matching using Cross-Correlation. <https://medium.com/mllearning-ai/image-template-matching-using-cross-correlation>, 06 2022.

[3] Richard Szeliski. *Computer Vision: Algorithms and Applications*. Springer, 2010.

Evaluierung von Frameworks zur Erstellung moderner Bots im Web 2.0

Konstantinos Tsolakidis

Jörg Nitzsche

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma IT-Designers GmbH, Entennest 2 73730 Esslingen

Mit der Entwicklung des Internets in den letzten Jahrzehnten hat sich die Art und Weise, wie Menschen mit dem Web interagieren, drastisch verändert. Dieser Wandel wird oft durch den Übergang vom Web 1.0 zum Web 2.0 veranschaulicht. Während das Web 1.0 als ein statisches, überwiegend lese basiertes Medium betrachtet wird, hat das Web 2.0 die Art und Weise, wie Menschen online interagieren, durch die Förderung der Teilhabe, Zusammenarbeit und Interaktion verändert [1].

Eine der interessantesten Entwicklungen ist die verstärkte Nutzung von Bots. Diese automatisierten Programme, die eine Vielzahl von Aufgaben ausführen können, sind in der Lage, menschliche Benutzer in zahlreichen Interaktionen zu ersetzen oder zu ergänzen. Sie finden Anwendung in verschiedenen Bereichen, von Kundenservice und Datenaggregation bis hin zu Social-Media-Management und Content-Generierung [4]. Bots können die gleichen Dinge tun, die andere Arten von Software tun können – aus Dateien lesen und in Dateien schreiben, Datenbanken und APIs verwenden und die regulären Rechenaufgaben erledigen. Was Bots einzigartig macht, ist die Nutzung von Mechanismen, die in der Mensch-zu-Mensch-Kommunikation verwendet werden [3].

Diese Arbeit befasst sich mit der Entwicklung sogenannter Chatbots. Ein Grundlegender Bestandteil zur Erstellung dieser Bots, sind die Bot Frameworks. Ein Bot Framework ist eine Softwareplattform, die als Grundgerüst für die Entwicklung und Implementierung von Bots dient. Diese Frameworks bieten eine Reihe von Funktionen und Diensten die es zu Evaluieren gilt. Der Begriff "Chatbot" wird oft fälschlicherweise als Synonym zu allen Software-Bot-Technologien verwendet und umgekehrt. Obwohl viele der bekannten Mainstream-Bots tatsächlich eine gewisse Sprachfähigkeit besitzen, ist das Verständnis von Sprache keine Voraussetzung für Software-Bots. Das bedeutet, alle Chatbots sind Bots, aber nicht alle Bots sind Chatbots wie in der Abbildung von 1 ersichtlich ist. Spezifischer gesagt, zeichnen sich Chatbots in dem größeren Bereich

der Software-Bots durch ihre Fähigkeit aus, Sprache zur Kommunikation mit ihren Nutzern zu verwenden. Chatbots werden auch häufig als Chatterbots, Talkbots oder künstliche konversationelle Entitäten bezeichnet. [4]

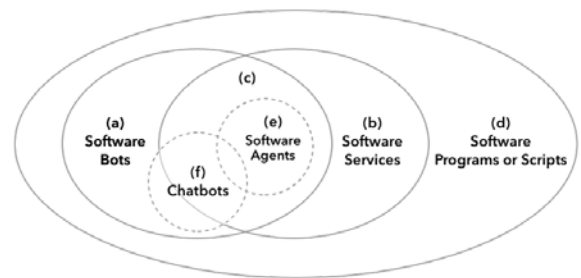


Abb. 1: Die Beziehung zwischen (a) Software-Bots, (b) Softwarediensten, (c) Software Bots mit internen Diensten, (d) Softwareprogrammen/Skripten, (e) Softwareagenten und (f) Chatbots. [4]

Problemstellung

Die Firma IT-Designers GmbH ist ein Softwareunternehmen mit einer Belegschaft von 150 Mitarbeitern. Im Zuge der Optimierung der internen Kommunikations- und Informationsflüsse strebt die Firma die Entwicklung eines maßgeschneiderten Chatbots an. Dieser soll den Mitarbeitern einen vereinfachten und effizienteren Zugang zu verschiedenen, im Unternehmensalltag relevanten Informationen ermöglichen. Dabei arbeiten die Mitarbeiter regelmäßig mit einer Vielzahl von Anwendungen und Plattformen wie Slack, Microsoft Teams, Jira, GitLab, E-Mail, Jenkins, Tomato und Confluence. Unter diesen nimmt Microsoft Teams die zentrale Rolle in der firmeninternen Kommunikation ein und wird daher als Zielplattform für den geplanten Chatbot ausgewählt.

Eine der wichtigen Informationsquellen im Unternehmen ist Confluence, auf dem alle relevanten Unternehmensinformationen gespeichert sind. Der zu entwickelnde Chatbot soll in der Lage sein, auf diese Informationen zuzugreifen und sie den Mitarbeitern auf Anfrage zur Verfügung zu stellen. Allerdings sind alle diese Anwendungen und Informationen hinter Teleport geschützt, einer Privileged Access Management (PAM)-Lösung, die den Zugriff auf interne Server und Webseiten kontrolliert. Dementsprechend muss der Chatbot so konzipiert werden, dass er diese Sicherheitsvorkehrungen berücksichtigt und dennoch die notwendigen Zugriffe ermöglicht. Er soll in der Lage sein, natürliche Sprache zu verstehen und auf Anfragen in natürlicher Sprache zu antworten. Daher muss das zu verwendende Bot-Framework die Möglichkeit bieten, diese Funktionen einzurichten bzw. durch eigenen Code erweitern zu können.

Des Weiteren, soll der Bot Zugang zur firmeneigenen Webseite Temato haben, auf der die aktuellen Schulungsangebote des Unternehmens aufgeführt sind. Die Möglichkeit, Informationen von dieser Webseite zu extrahieren, könnte durch die Implementierung eines Webcrawlers erreicht werden. Darüber hinaus soll der Bot in der Lage sein, den Mitarbeitern auf Grundlage ihrer individuellen Fähigkeiten passende Schulungsangebote vorzuschlagen wie in Abbildung 2 zu sehen ist.

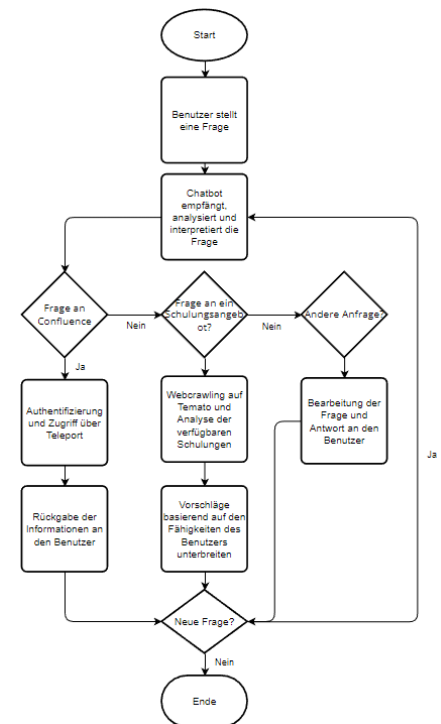


Abb. 2: Das Flussdiagramm beschreibt den Ablauf einer Interaktion mit dem Benutzer. [2]

Konzept

Zunächst wurde ein Szenario erstellt, um verschiedene Use Cases zu identifizieren, in denen der Chatbot eingesetzt werden könnte. Dies diente dazu, ein klares Bild davon zu bekommen, welche Aufgaben der Chatbot erfüllen soll und welche Interaktionen mit den Mitarbeitern und den vorhandenen IT-Systemen er zu bewältigen hat. Die Szenarien berücksichtigten auch die spezifischen Besonderheiten und Herausforderungen, wie beispielsweise die Verwendung von Teleport zur Sicherheitskontrolle und die Notwendigkeit, auf Informationen aus verschiedenen Quellen wie Confluence und Temato zuzugreifen.

Aus den erstellten Szenarien wurden dann die Anforderungen an den Chatbot abgeleitet. Diese Anforderungen spezifizierten, welche Funktionen und Fähigkeiten der Chatbot haben sollte, einschließlich der Fähigkeit, natürliche Sprache zu verstehen und darauf zu antworten, auf Informationen aus verschiedenen Quellen zuzugreifen und sich an die Sicherheitsprotokolle des Unternehmens zu halten.

Als Nächstes wurde eine umfangreiche Recherche durchgeführt, um verfügbare Frameworks und Plattformen zur Erstellung von Chatbots zu identifizieren. Diese Frameworks wurden dann anhand von zuvor festgelegten K.O.-Kriterien gefiltert. Nur die Frameworks, die alle K.O.-Kriterien erfüllten, wurden für die nächste Phase der Implementierung und Auswertung ausgewählt. Da in der Recherche eine sehr große Anzahl an Frameworks und Plattformen gefunden wurden und sich diese nicht so leicht einschränken lassen haben, mussten das Szenario und die K.O.-Kriterien mehrfach angepasst werden. Dadurch sind auch Anforderungen entstanden, wie die Eingrenzung von Programmiersprachen und die Möglichkeit den Bot On-Premise zu hosten. Dies hat dazu geführt, dass am Ende nur 3 Frameworks übrig geblieben sind.

Das Microsoft Bot Framework, das eines der Frameworks ist, was nach dem Filter übrig geblieben ist, bietet eine umfangreiche Dokumentation mit einigen Beispielen zu Implementierung. Die einfachste Form eines Chatbots ist der Echo Bot wie in 3 zu sehen ist. Er gibt einfach die Benutzereingabe wieder. Er dient als guter Einstieg in die Chatbot-Entwicklung ähnlich wie das Hello-World Programm wenn eine neue Programmiersprache erlernt werden soll.

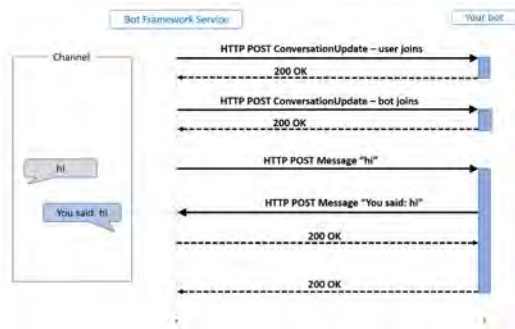


Abb. 3: Dieses Diagramm veranschaulicht zwei Aktivitätstypen, Konversationsaktualisierung und Nachricht, die möglicherweise ausgetauscht werden, wenn ein Benutzer mit einem Echo-Bot kommuniziert. [3]

Ausblick

Die übrig gebliebenen Frameworks werden dann eingehend getestet. Hierfür werden Prototypen erstellt, die die Anforderungen des Szenarios erfüllen sollen. Schließlich werden die Ergebnisse der Tests durch die Prototypen verglichen und in einer Tabelle zusammengefasst.

Literatur und Abbildungen

- [1] Paul Alpar and Steffen Blaschke. *Web 2.0 - Eine empirische Bestandsaufnahme*. Vieweg + Teubner, 1 edition, 2008.
- [2] Eigene Darstellung.
- [3] Jonathan Fingold and E. Hopf. What is the Bot Framework SDK? <https://learn.microsoft.com/en-us/azure/bot-service/bot-service-overview?view=azure-bot-service-4.0>, 11 2022.
- [4] Carlene Lebeuf. *A Taxonomy if Software Bots: Towards a Deeper Understanding of Software Bot Charachteristics*, 2018.

Leistungsfähigkeit von neuronalen Netzen zur Verbesserung von Bilddaten

Joshua Uhl

MarkusENZweiler

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma Leuze electronic GmbH + Co. KG, Owen

Motivation

Aus vielen Bereichen der Industrie ist der Einsatz von 2D-Codes nicht mehr wegzudenken. Zu den gängigsten und bekanntesten gehören der QR-Code und der DataMatrix-Code. [3] Letzterer wird beispielsweise auf Briefen und Paketen angebracht, um unter anderem die korrekte Zustellung zu gewährleisten. [4] Zum Lesen von DataMatrix-Codes kommt üblicherweise ein Kamerasensor zum Einsatz. [4] Typischerweise sind zweidimensionale Codes so strukturiert, dass sie aus verschiedenen Perspektiven gelesen werden können. Der DataMatrix-Code besitzt dafür einen gemusterten Rand, der immer gleich aussieht und zur Orientierung dient. Zudem beinhaltet die Information des Codes Redundanz zur Fehlerkorrektur, um eventuellen Lesefehlern vorzubeugen. [4]

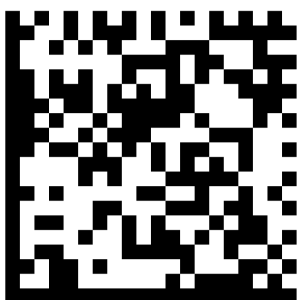


Abb. 1: Beispiel eines DataMatrix-Codes [1]

Trotz all dieser Maßnahmen, die dafür sorgen sollen, dass der Code möglichst auch unter schwierigen Bedingungen ausgelesen und interpretiert werden kann, kommt es unter bestimmten Umständen noch immer vor, dass der von der Kamera abfotografierte Code für den Dekodierer unlesbar ist. Das kann unterschiedliche Ursachen haben. Häufig liegt es daran, dass der Abstand des Sensors zum Code keine optimalen Bedingungen für die Bildaufnahme ergibt. So kann das

Bild dann unscharf oder sehr niedrigauflösend sein. Um dem nun entgegenzuwirken, soll ein neuronales Netz zum Einsatz kommen. Dessen grundlegende Aufgabe soll es sein, ein vom Dekodierer nicht interpretierbares Bild so zu verbessern, dass der darauf enthaltene DataMatrix-Code wieder lesbar wird.

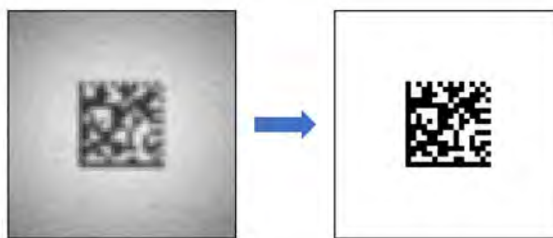


Abb. 2: Visualisierung der Aufgabe des neuronalen Netzes mit Hilfe von generierten Bildern [1]

Neuronale Netze

Neuronale Netze kommen in Bereichen zum Einsatz, in denen komplexe Zusammenhänge bestehen, die für die herkömmliche Softwareentwicklung zu anspruchsvoll sind. Anstatt wie üblich explizit programmierte Regeln anzuwenden, lernen neuronale Netze Muster aus Daten. Diese erlernten Algorithmen können anschließend übergreifend auf neue Daten angewendet werden. Typische Einsatzgebiete von neuronalen Netzen sind unter anderem statistische Modelle, medizinische Diagnostik, sowie Bildverarbeitung, zu welcher auch diese Arbeit gehört. [5]

Ein neuronales Netz besteht, wie der Name bereits vermuten lässt, aus künstlichen Neuronen, die in Schichten angeordnet sind. Die Schichten sind dabei hintereinander angeordnet. Die erste Schicht repräsentiert die Eingangsdaten, während die letzte die Ausgangsdaten enthält. Dazwischen können sich beliebig viele weitere Schichten befinden. Diese Anordnung ist in Abb. 3

dargestellt. Jedes Neuron speichert einen numerischen Wert und somit einen Teil der Informationen.

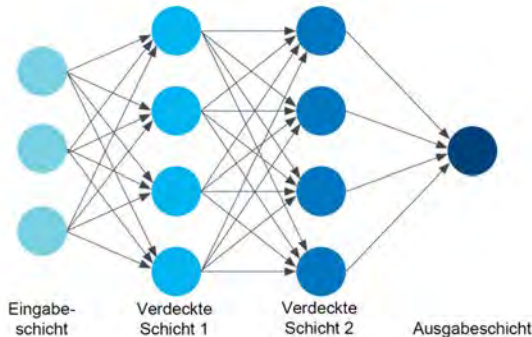


Abb. 3: Verallgemeinerter Aufbau eines neuronalen Netzes [2]

Die Art und Weise, auf die die Eingangsdaten des Netzes zu Ausgangsdaten verarbeitet werden, wird maßgeblich von den Gewichten bestimmt, welche einfache Faktoren zwischen den einzelnen Neuronen benachbarter Schichten sind. In Abb. 3 werden diese Gewichte durch die Pfeile symbolisiert. In einem Trainingsprozess werden die Gewichte optimal eingestellt. Dabei durchlaufen die Eingangsdaten zunächst das Netz, welches zu Beginn mit zufällig initialisierten Gewichten arbeitet. Die Ausgabe des Netzes wird nun mit den tatsächlichen Ausgangsdaten verglichen und ein Fehlerterm berechnet. Mit Hilfe von diesem Fehlerterm werden die Gewichte rückführend angepasst, sodass der Fehler minimiert wird. [5] Dieser Vorgang wird so lange wiederholt, bis ein ausreichend zufriedenstellendes Ergebnis erreicht wurde.

Projektdurchführung

Zum Einsatz kommt ein Convolutional Neural Network (CNN), welches sich insbesondere für Bildverarbeitung eignet, da es Informationen über räumlichen Zusammenhang beinhaltet. Die Daten, mit denen das Netz trainiert wird, werden künstlich generiert. Die Herausforderung dabei besteht darin, diese Generierung

der Bilder möglichst realistisch zu gestalten. Dafür werden unterschiedlichste Faktoren, darunter Unschärfe, Bildrauschen, inhomogene Beleuchtung und viele weitere, berücksichtigt. Zusätzlich skaliert das Netz die Bilder hoch, um die Auflösung zu erhöhen.

Ergebnis

Es hat sich gezeigt, dass das neuronale Netz in der Lage ist, generierte Bilder sehr gut zu verbessern. Darüber hinaus weist es auch eine gute Leistungsfähigkeit bei realen Bildern auf. So können viele Codes wiederhergestellt werden, die andernfalls für das System und selbst manchmal für den Menschen unlesbar sind. Trotzdem kann das Netz nicht jedes Bild "retten", da in manchen Situationen der Informationsverlust schlicht zu groß ist. Abb. 4 zeigt ein Beispiel eines aufgenommenen Codes, der vom Netz deutlich sichtbar verbessert wurde.

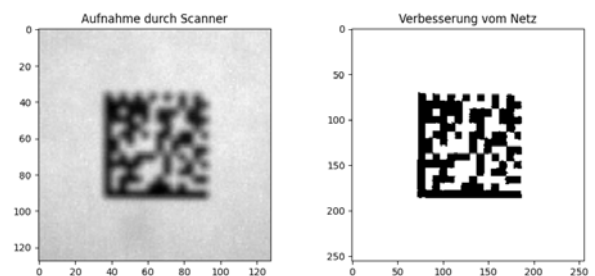


Abb. 4: Ein vom Netz verbessertes Bild [1]

Ausblick

Als nächsten großen Schritt kann nun das Netz in ein Live-System eingebunden werden. Dadurch könnte man Abläufe implementieren, um die Code-Verbesserung möglichst sinnvoll einzusetzen. Dafür würde der Code zunächst wie üblich vom System gescannt werden. Falls dieser Code nicht erkennbar oder lesbar ist, würde das neuronale Netz darauf angewandt werden, um zu versuchen, den Code wiederherzustellen. Anschließend soll der vom Netz verbesserte Code erneut dem System vorgelegt werden.

Literatur und Abbildungen

- [1] Eigene Darstellung.
- [2] Santosh Rao. Aufbau einer Daten-Pipeline für Deep Learning. *NetApp*, page 16, 2019.
- [3] . Wikipedia. 2D-Code. <https://de.wikipedia.org/wiki/2D-Code>, 2022.
- [4] . Wikipedia. DataMatrix-Code. <https://de.wikipedia.org/wiki/DataMatrix-Code>, 2023.
- [5] . Wikipedia. Künstliches neuronales Netz. https://de.wikipedia.org/wiki/K%C3%BCnstliches_neuronales_Netz, 2023.

Analyse und Konzeptionierung eines SAP-Addons zur automatisierten Verarbeitung eingehender Zahlungsavisen

Viktoria Wagner

Anke Bez

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma zur Softwareerstellung in Kooperation, Leonberg

Problemstellung

Ob Rechnungen, Mahnungen oder Zahlungseingänge: Ein Unternehmen befindet sich in ständiger Interaktion mit Dienstleistern, Lieferanten sowie Kunden und erhält im Austausch mit diesen täglich Dokumente – sei es in Papierform oder in digitaler Variante. Diese Masse an Dokumenten muss manuell bearbeitet und in entsprechende Systeme zur Weiterverarbeitung übertragen werden. Durch den Einsatz neuer Technologien können diese Prozesse automatisiert werden; dafür ist ein potenzieller Prozess die Verarbeitung von eingehenden Zahlungsavisen. Werden durch eine Zahlung mehrere Rechnungen beglichen, dienen Zahlungsavisen zur Ankündigung einer eintreffenden Zahlung und helfen bei der Zuordnung der offenen Posten. Betrachtet man das weltweit am häufigsten genutzte ERP-System SAP, mit einem Marktanteil von 22,5% (Stand 2017), wird die Zuordnung der offenen Posten zum Zahlungseingang manuell aus dem PDF-Dokument in SAP übertragen. [3]

Zielsetzung

Durch die Konzeptionierung eines SAP-Add-ons auf Basis einer vorangegangenen Analyse soll der Prozess zum Verarbeiten von eingehenden Zahlungsavisen effizienter gestaltet werden, wodurch ein Zeitersparnis in Form von freier Kapazität für andere Tätigkeiten erreicht werden soll. Zudem soll die Fehlerquote minimiert werden, da menschliche Irrtümer, wie beispielsweise Tippfehler,

reduziert werden können. Zusätzlich sollen nach der Amortisationsdauer Kosten eingespart werden. Hierfür werden im Rahmen der Thesis folgende forschungsleitenden Fragen beantwortet: Wie ist der Prozess von eingehenden Zahlungsavisen im SAP-Standard? Welche Funktionen sind im Rahmen der Automatisierung seitens des Kunden erwünscht? Welche Funktionen müssen im Rahmen einer Neueinführung abgedeckt sein? Welchen Mehrwert bietet ein solches Produkt dem Kunden?

Zahlungsavisen

Im Alltag ist der Begriff Avis bspw. durch Online-Bestellungen gebräuchlich, wenn eine Lieferung durch eine Mitteilung (ein Lieferavis) angekündigt wird und diese eine Aufzählung enthält, welche Artikel in welchem Umfang bei der Lieferung enthalten sein werden. Im Bankwesen versteht man unter einem Avis die Ankündigung einer eingehenden Zahlung. Das Zahlungsavis wird vor allem bei eingehenden Sammelüberweisungen eines Debtors verwendet, zur Zuordnung des Gesamtbetrags zu den offenen Posten der Rechnungen [5]. So wird die Sammelüberweisung in Abbildung 1 über 50€ zu den vier offenen Posten der Referenz AvisTest2 zugeordnet. Bei der Verarbeitung der Informationen aus einem Zahlungsavis, unabhängig vom Anwendungsgebiet, ist es notwendig, die Rechnungsreferenzen mit denen vom Avis manuell abzugleichen.

<input type="checkbox"/>	<input checked="" type="checkbox"/>		1400000001	DZ	17.05.2023			50,00-	EUR	1400000001
<input type="checkbox"/>	<input checked="" type="checkbox"/>	AvisTest2	1800000005	DR	26.04.2023			12,00	EUR	1400000001
<input type="checkbox"/>	<input checked="" type="checkbox"/>	AvisTest2	1800000006	DR	26.04.2023			13,00	EUR	1400000001
<input type="checkbox"/>	<input checked="" type="checkbox"/>	AvisTest2	1800000007	DR	26.04.2023			14,00	EUR	1400000001
<input type="checkbox"/>	<input checked="" type="checkbox"/>	AvisTest2	1800000008	DR	26.04.2023			11,00	EUR	1400000001

Abb. 1: Eigener Screenshot aus Transaktion FBL5N in einem SAP S4/Hana System [2]

Möglichkeiten zur Automatisierung

Finanzprozesse, einschließlich der Verarbeitung von Zahlungsavisen, sind aufgrund folgender Merkmale gut für die Automatisierung geeignet: Sie zeichnen sich durch Wiederholbarkeit aus, da häufig ähnliche wiederkehrende Aufgaben ausgeführt werden müssen. Zudem beinhalten die Dokumente strukturierte Daten. Merkmale, die sich auf strukturierte Daten beziehen, lassen sich besser extrahieren, da sie klare Muster aufweisen. Auch Schlüsselwörter wie *Avisnummer* oder *Belegnummer* helfen bei der effizienten Identifizierung. Seit der Entwicklung von Software und Systemen können durch den Einsatz von Technologien repetitive monotone Prozesse des Menschen substituiert werden. Bereits 2017 konnte in einer Studie von McKinsey festgestellt werden, dass durch derzeit eingesetzte Technologien 42% der Finanzaktivitäten vollständig automatisiert werden könnten [8]. Ein Teilbereich dieser Technologien ist die Dokumentverarbeitung, die mithilfe der optischen Zeichenerkennung die geometrische Struktur des Dokuments extrahiert und in eine logische Struktur umwandelt [9]. Dabei werden die Buchstaben und Ziffern einer Seite gescannt und in ein Bild umgewandelt. Das Bild wird anschließend in weiße und schwarze Punkte aufgeteilt und diese in eine Matrix eingeordnet, die mit einer Bibliothek abgeglichen wird, welche die Punkte einem Zeichen zuordnen kann. Dies ermöglicht, die Dokumente automatisiert zu digitalisieren und für jeden zugänglich zu machen und setzt die Voraussetzung für den weiteren Automatisierungsprozess. Durch den Einsatz von Positionserkennung kann die Genauigkeit der Dokumenterkennung zusätzlich gesteigert werden. Daten können somit auf Positionsebene erfasst werden, die Informationen somit spaltengenau extrahiert werden, was zum einen zu einer höheren Qualität der Erkennung führt und zudem neue Möglichkeiten zum Beschaffen von Informationen bietet. Die erkannten Zeichen können anschließend verarbeitet und an SAP übergeben werden. Dort kann die Workflowautomatisierung ansetzen. Ein Workflow ist eine definierte Folge von Arbeitsschritten zum Erstellen eines Ergebnisses, die von unterschiedlichen Mitarbeitern durchgeführt werden [4]. Durch eine Definition des Workflows kann festgelegt werden, wer für welche Aufgabe verantwortlich ist und welche Reihenfolge beachtet werden soll. Durch die Aufteilung des Gesamtprozesses (vgl. Abbildung 2) in die einzelnen Workflowschritte und den Einsatz von Workflow-Strategien wird die Automatisierung von Genehmigungsprozessen, Benachrichtigungen und Datenverarbeitung ermöglicht.

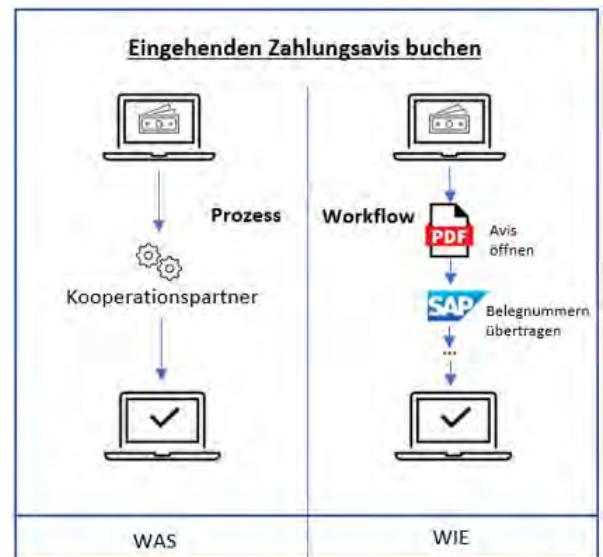


Abb. 2: Geschäftsprozess vs. Workflow in Anlehnung an Bühler et al., 2018, S.62 [1]

So ermöglicht die Workflowautomatisierung auch die Überwachung und Dokumentation der Aufgaben und Aktivitäten und bietet Einblicke in den Ablauf. Dies führt zu mehr Transparenz und bietet die Möglichkeit frühzeitig bei Verzögerungen einzugreifen.

SAP Add-ons

Durch SAP Add-ons können zusätzliche Funktionen oder Anpassungen in SAP integriert werden, welche über den SAP-Standard hinausgehen. Diese werden meist durch Drittanbieter für einen spezifischen Geschäftsprozess, wie z.B. Zahlungsavisen, entwickelt. Add-ons zeichnen sich dadurch aus, dass sie inside SAP sind, somit nahtlos integriert werden können und prozessspezifische Lösungen bieten.

Vorgehensweise

Zum Ermitteln der sich wiederholenden Teilschritte, in denen das Automatisierungspotential analysiert wird, wird der SAP-Standardprozess zum Ausgleich der offenen Posten ausgeführt und der Ist-Prozess modelliert. Auf Grundlage des Ist-Prozesses wird durch eine qualitative Befragung eine Erhebung der Kundenanforderungen durchgeführt. Hierzu wird ein semistrukturiertes Leitfadenterview zugrunde gelegt, um durch vorher festgelegte Fragen eine Vergleichbarkeit zwischen den einzelnen Antwortmöglichkeiten zu bieten, jedoch gleichzeitig die Befragten frei und in eigenen Worten antworten zu lassen. Hierdurch besteht die Möglichkeit, umfangreichere Antworten zu erhalten und die Befragung flexibler zu gestalten. Zusätzlich wird die Befragung durch einen Fragebogen der Kano-Methode ergänzt, um durch funktionale

und dysfunktionale Fragen die Basis-, Leistungs- und Begeisterungsfaktoren zu ermitteln [6]. Somit kann nach Durchführung und Interpretation der Befragung ein Kano-Modell erstellt und mithilfe eines Value-Proposition Canvas [7] können die Leistungsfaktoren den möglichen Funktionen des zukünftigen Produktes gegenübergestellt werden. Auf Grundlage dessen wird anschließend ein Soll-Prozess modelliert und im Anschluss konzeptioniert.

Ausblick

Nach Fertigstellen der Analyse wird die Konzeptionierung durchgeführt. Die jeweiligen Informationen aus dem Dokument werden den technischen Feldern in SAP gegenübergestellt und gemappt. Dabei sollen die benötigten Transaktionen und Modulabsprünge beinhaltet sein. Das Konzept beschränkt sich auf die in der Befragung erhobenen Basis- und Leistungsfaktoren. Mithilfe eines Mock-Ups sollen die zukünftigen Funktionen dargestellt, diskutiert und getestet werden.

Literatur und Abbildungen

- [1] Peter Bühler, Patrick Schlaich, Dominik Sinner, and in Anlehnung an. *Medienworkflow: Kalkulation - Projektmanagement - Workflow*. Springer, 2018.
- [2] Eigene Darstellung.
- [3] Statista GmbH. Marktanteile Der Führenden Anbieter Am Umsatz Mit Enterprise-resource-planning-anwendungen (Erp) Weltweit Im Jahr 2017. <https://de.statista.com/statistik/daten/studie/262342/umfrage/marktanteile-der-anbieter-von-erp-software-weltweit/>, 10 2018.
- [4] tinformatik GmbH. Worklfow. <https://t2informatik.de/wissen-kompakt/workflow/>, 02 2018.
- [5] Elmar Goldstein. *Belege richtig kontieren und buchen - inkl. Arbeitshilfen online: Typische Buchungsfälle nach den Kontenrahmen für DATEV, IKR, BGA*. Haufe-Lexware, 7 edition, 2018.
- [6] Jorg Holzing. *Die Kano-Theorie Der Kundenzufriedenheitsmessung: Eine theoretische und empirische Überprüfung*. Gabler Verlag, 1 edition, 2008.
- [7] Alexander Osterwalder. *Value Proposition Design: Entwickeln Sie Produkte und Services, die Ihre Kunden wirklich wollen Die Fortsetzung des Bestsellers Business Model Generation! Übersetzt von T. A. Wegberg*. Campus, 1 edition, 2015.
- [8] Frank Plaschke et al. Bots, Algorithms, and the Future of the Finance Function. <https://www.mckinsey.com/capabilities/strategy-and-corporate-finance/our-insights/bots-algorithms-and-the-future-of-the-finance-function>, 01 2018.
- [9] Yuan Yan Tang et al. Document processing for automatic knowledge acquisition“. IEEE transactions on knowledge and data engineering. In *IEEE Transactions on knowledge and data engineering*, volume 6, pages 3–21. Engineers, Institute of Electrical and Electronics, 1994.

Konzeption und prototypische Umsetzung einer Migrationsstrategie zur Überführung einer Eclipse RCP-Anwendung in eine webbasierte Lösung

Jonathan Weyl

Mirko Sonntag

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma ARS NOVA Software GmbH, Esslingen

Einleitung und Motivation

Web-Anwendungen sind aus der heutigen Software-Welt nicht mehr wegzudenken. Sowohl im Privaten als auch im Business-Bereich konkurrieren sie ernsthaft mit klassischen Desktopanwendungen. Begünstigt durch immer leistungsfähigere Webbrowser und schnell wachsende umfangreiche Entwicklungsframeworks und Bibliotheken werden Web-Anwendungen in immer mehr Anwendungsbereichen eine mögliche und oft sogar bessere Alternative. Cloud-basierte Anwendungen sind mittlerweile in der Software-Entwicklung als Standard anzusehen (siehe [1]). Hinzu kommt häufig die Anforderung, Anwendungen schnell und ohne großen Aufwand auf jedem möglichen Endgerät zu starten. Durch Web-Anwendungen ist es so möglich, dass Millionen Home-Office Arbeitenden in Deutschland und der Welt (siehe [2]) Anwendungen ohne Installationsaufwand und spezielle Systemvoraussetzungen jederzeit sogar auf ihrem privaten Endgerät verwenden können. Diese Trends und Entwicklungen motivieren dazu, auch lange bestehende Desktopanwendungen und Systeme in die Web-Welt zu überführen, um zukunftsfähig aufgestellt zu sein und den Wert langjähriger Entwicklungsarbeit zu erhalten.

Ziel

Im Auftrag eines Kunden wurde durch die ARS NOVA Software GmbH eine große Eclipse Rich Client

Platform (RCP) basierte Desktopanwendung erstellt und in mehreren Release-Zyklen stetig weiterentwickelt (System Planea, Name für diesen Beitrag geändert). Ziel dieser Arbeit ist am Beispiel von Planea die Konzeption einer Migrationsstrategie hin zu einer webbasierten Lösung sowie die Implementierung einer prototypischen Web-Anwendung zur Validierung der Strategie. Dabei sollen die Interessen von vier verschiedenen Interessensgruppen an diesem Migrationsprojekt wahrgenommen werden: die Endanwender/innen bei der Kundenfirma, die IT-Abteilung der Kundenfirma, die eigene Firma in der Rolle des Softwarelieferanten und die wissenschaftliche Community.

Konzept

Ausgangssystem Planea besteht aus einer sehr umfangreichen Desktopanwendung auf Basis von Eclipse RCP und einem Java-basierten Anwendungsserver. In diesem Projekt wird die Client-Seite des Systems betrachtet. Planea wurde über einen Zeitraum von mehr als zehn Jahren weiterentwickelt und ist sehr gewachsen. Deshalb werden hier ausschließlich die Grundkonzepte der Plattform (Eclipse RCP), eigene Grundkonzepte, die die Plattform erweitern sowie einige für eine erste Migration besonders interessante, fachliche Anwendungsanteile untersucht. Die nachfolgende Abbildung gibt einen Einblick in einen Teil des Ausgangssystems:

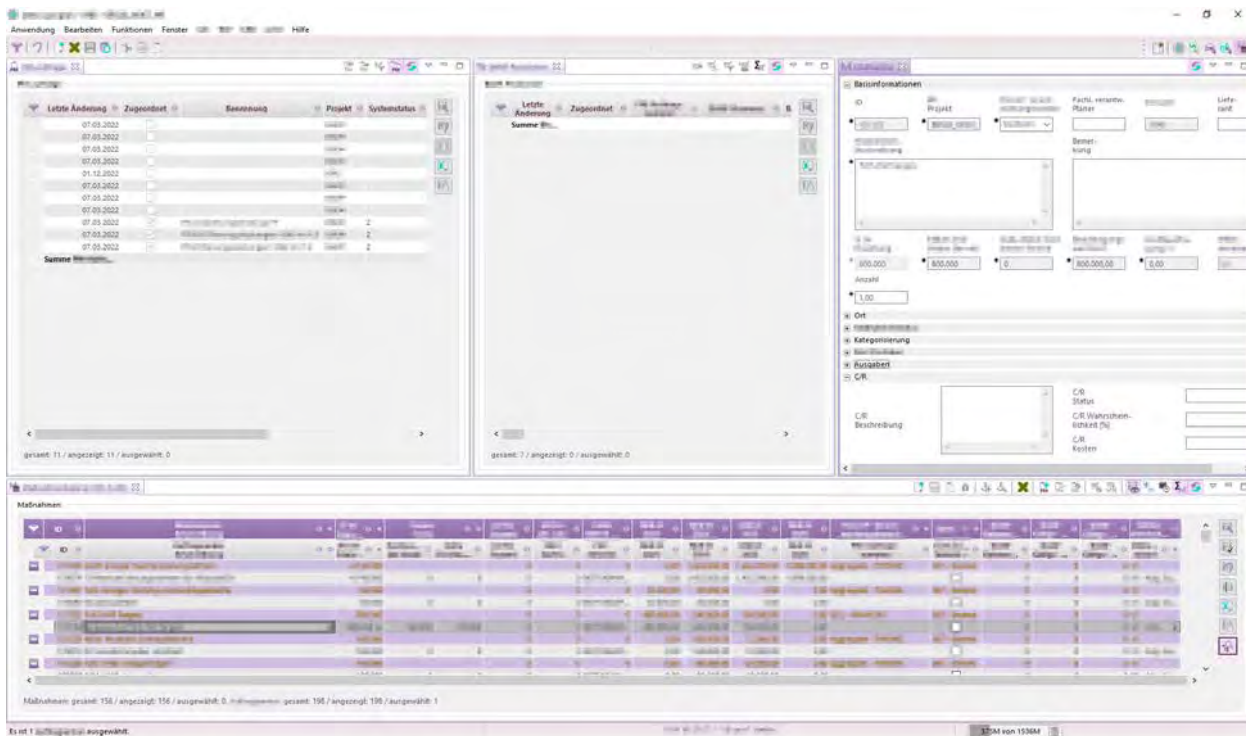


Abb. 1: Ausgangssystem Planea [3]

Die Abbildung zeigt eine Perspektive mit vier Parts, mit verschiedenen Formen der Datenabbildung. Der Part oben rechts enthält ein Formular mit verschiedenen Eingabefeldern und aufklappbaren Gruppen. Die anderen Parts zeigen Tabellen. Um den Umfang der Migration und der prototypischen Entwicklung in handhabbare Einheiten zu teilen, werden Key-Features des Ausgangssystems identifiziert. Dabei werden sowohl Features der zugrundeliegenden Plattform (Eclipse RCP) als auch eigene Konzepte berücksichtigt. Die identifizierten Key-Features sind in folgender Tabelle übersichtsartig dargestellt:

	Key-Feature	Beschreibung
1	Perspektiven	Handlung von mehreren Parts und Perspektiven
2	Dynamic Page Model	Modell zur Strukturierung von Oberflächenelementen
3	Dynamic Element Model	Modell zur Steuerung von Oberflächenelementen
4	Domain-Objekt Datenmodell	Clientseitige Datenhaltung
5	Event-Handling	Senden und Verarbeitung von Eclipse-Events
6	Command-Handling	Verarbeitung von Benutzereingaben
7	Selektions-Management	Selektionsabhängige Daten-Befüllung und Verhalten
8	Behaviour-Management	Clientseitige Anwendungslogik
9	Dirty-Management	Änderungserkennung bei bearbeiteten Daten
10	Client-Server-Schnittstelle	Kommunikation mit dem Anwendungsserver

Abb. 2: Key-Features Ausgangssystem [3]

Migrationsstrategie Die identifizierten Key-Features werden in eine sinnvolle Reihenfolge gebracht und nacheinander, aufeinander aufbauend in das Design des Zielsystems integriert. Dabei sollen nach Möglichkeit alle nicht-funktionalen (NFR) und funktionalen (FR) Eigenschaften des Ausgangssystems aus Nutzersicht auch im Zielsystem wiederzufinden sein. Die grundlegende Migrationsstrategie teilt die Migration in zwei

Phasen auf:

- **Konzept-Migration:** Migration der wichtigsten Konzepte der Plattform ohne fachliche Inhalte.
- **Fachliche Migration:** Migration ausgewählter fachlicher Inhalte des Ausgangssystems in das Zielsystem.

Zielarchitektur Das Design des Zielsystems enthält alle Key-Features des Ausgangssystems und bietet die Grundlage für die spätere (prototypische) Implementierung. Da eine Betrachtung aller Key-Features im Kontext dieses Beitrags zu umfangreich wäre, wird an dieser Stelle Einblick in das Zieldesign der Perspektiven (siehe Abbildung 2) gegeben. Perspektiven dienen der dynamischen Aufteilung der Oberfläche in verschiedene Bereiche, die mit beliebigen Anwendungsinhalten gefüllt werden können. Einige Perspektiven sind für die Benutzer/innen vorkonfiguriert, sie können aber auch individuell erstellt oder angepasst werden. Das Zieldesign einer Perspektive orientiert sich an dem Application Model in Eclipse e4 (vgl. [4]) ist aber einfacher und übersichtlicher.

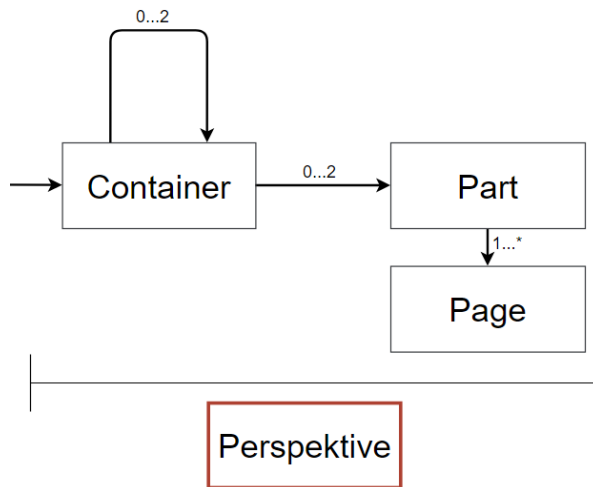


Abb. 3: Aufbau Perspektive [3]

Eine Perspektive ist als Konfigurationsobjekt mit Containern, Parts und Pages zu betrachten. Container dienen der Strukturierung und sind an der Oberfläche nur indirekt sichtbar. Sie haben verschiedene Eigenschaften, wie beispielsweise die Split-Richtung (horizontal oder vertikal). Jeder Container enthält entweder zwei Container, einen Part und einen Container oder zwei Parts. Parts sind die sichtbaren Elemente innerhalb von Containern. Jeder Part kann beliebig viele Pages enthalten. Pages enthalten die fachlichen Oberflächen, wie z.B. Formulare oder Tabellen. Auf Basis von Containern und Parts lässt sich eine Perspektive als Binärbaum darstellen:

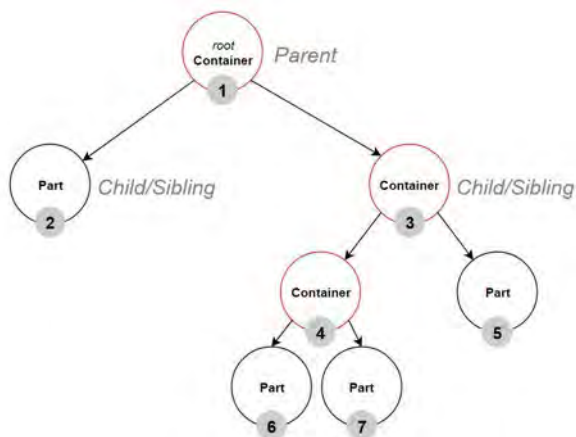


Abb. 4: Perspektive als Binärbaum [3]

Auf Grundlage dieses Binärbaums ist jede Perspektive dynamisch änderbar (Parts entfernen und hinzufügen), immer in einem konsistenten Zustand und sehr flexibel einsetzbar. Die in den Parts enthaltenen Pages stellen eine weitere Ebene im Baum dar. Sie können zur Laufzeit beliebigen Parts zugeordnet und verschoben werden.

Prototyp

Das erarbeitete Zieldesign soll durch eine prototypische Umsetzung validiert werden. Das Zielsystem ist dabei so entworfen, dass es grundsätzlich mit jedem modernen Web-Framework (Angular, React, Vue, ...) umgesetzt werden kann. Unter Berücksichtigung der verschiedenen Interessensgruppen und Bewertungsfaktoren an diesem Migrationsprojekt fällt die Wahl der Technologie auf die JavaScript Bibliotheken React (siehe [5]) und AppRun (siehe [8]) React ist dabei für das Projektsetup, das Oberflächendesign durch JSX und für das Einbinden von React-basierten Drittbibliotheken (z.B. Material UI [6]) verantwortlich. AppRun ist für die State-Verwaltung, die Steuerung des View-Renderings und das Event-Handling zuständig. Die genauen Konzepte der beiden Bibliotheken werden hier nicht weiter erläutert und können in den jeweiligen Dokumentationen (siehe [9] und [7]) nachgelesen werden. Die Entwicklung des Prototyps orientiert sich an der zweistufigen Migrationsstrategie: Prototyp I (Multi Page Platform) ist eine allgemeingültige Plattform. Sie soll alle Key-Features (siehe Abbildung 2) abdecken oder im Design berücksichtigen, an welcher Stelle diese später eingefügt werden können. Außerdem enthält sie beispielhafte Pages, um die Funktionalität der Key-Features demonstrieren zu können. Prototyp II (Planea Web) setzt dann ausgewählte fachliche Funktionalitäten und Oberflächen aus Planea auf Basis von Prototyp I (Multi Page Platform) um. Ein aktuelles Zwischenergebnis der Implementierung von Prototyp I sieht folgendermaßen aus:

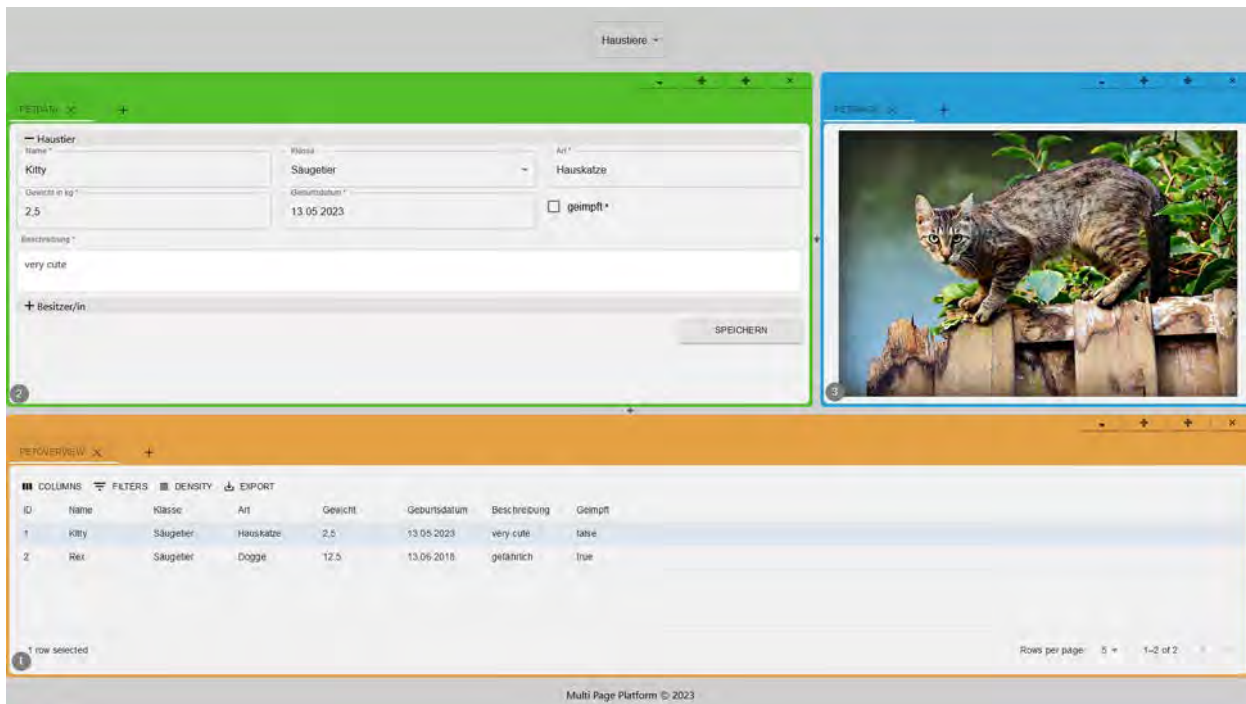


Abb. 5: Prototyp I [3]

Die Multi Page Platform ist in drei Bereiche eingeteilt: Header mit Perspektiv-Selektion (oben), Workbench mit Containern, Parts und Pages (mittig) und Statusbar (unten). In Abbildung 5 ist eine Beispiel-Perspektive mit zwei Containern, drei Parts und drei Pages zu sehen. Container 1 (root-Element) enthält Container 2 und Part 1 (orange), Container 2 umfasst Part 2 (grün) und Part 3 (blau). Jeder Part enthält genau eine Page. Selektiert man in der Tabelle in Part 1 ein Element, so wird das Formular in Part 2 entsprechend ausgefüllt und das passende Bild in Part 3 angezeigt.

Ausblick

Bisher konnten acht von zehn Key-Features entworfen, in das Design des Zielsystems integriert und im Prototyp I implementiert werden. Die weiteren beiden

Key-Features (Behaviour-Management und Client-Server-Schnittstelle) werden noch folgen, um die Konzept-Migration und die Entwicklung von Prototyp I abzuschließen. Anschließend wird Planea Web in Prototyp II konzipiert und umgesetzt, um an einem realen Beispiel zu zeigen, dass sich die Multi Page Platform als Basis für Anwendungen eignet, die bisher auf Eclipse RCP aufgebaut haben. Zum Abschluss des Projektes werden noch einmal Ausgangssystem und Zielsystem (inklusive Prototyp) gegenüber gestellt, um das Zieldesign zu validieren und den Erfolg des Migrationsprojektes zu verifizieren. Wird das Zieldesign dabei für robust genug und das gezeigte Features-Set des Prototyps für realistisch einsetzbar empfunden, kann Planea weiter iterativ in die Web-Welt migriert werden. So bleibt der Wert der langjährigen Entwicklungsarbeit erhalten und Planea ist gut für die Zukunft aufgestellt.

Literatur und Abbildungen

- [1] Statistisches Bundesamt. Unternehmen mit Nutzung von Cloud-Computing 2021. https://www.destatis.de/DE/Themen/Branchen-Unternehmen/Unternehmen/_Grafik/_Interaktiv/ikt-nutzung-cloud-computing.html, 2021.
- [2] Statistisches Bundesamt. Ein Viertel aller Erwerbstätigen arbeitete 2021 im Homeoffice. https://www.destatis.de/DE/Presse/Pressemitteilungen/Zahl-der-Woche/2022/PD22_24_p002.html, 06 2022.
- [3] Eigene Darstellung.
- [4] Jonas Helming and Maximilian Kögel. *Eclipse 4 (e4) Tutorial*. EclipseSource, 2016.
- [5] Facebook Open Source. React - A Java Script Library for building user interfaces. <https://reactjs.org>, 2023.
- [6] Material UI SAS. MUI: The React component library you always wanted. <https://mui.com>, 2023.
- [7] Yiyi Sun. *Practical Application Development with AppRun*. Apress, 2019.
- [8] Yiyi Sun. AppRun. <https://apprun.js.org>, 2022.
- [9] Marc Tefuel and Jonas Helming. *Eclipse 4 - Rich Clients mit dem Eclipse SDK 4.2*. entwickler.press, 2012.

Konzeption und Realisierung einer verlustfreien, zeitsynchronen Speicherung von Messdaten mittels moderner Echtzeitmechanismen

Jan von Scholley

Michael Scharf

Fakultät Informatik und Informationstechnik, Hochschule Esslingen

Arbeit durchgeführt bei der Firma comemso electronics GmbH, Ostfildern

Einleitung

In den letzten Jahren ist die Automobilindustrie im Umbruch. Es werden immer mehr Elektrofahrzeuge gebaut. Im Jahr 2022 waren bereits 17,7% aller Neuwagen vollelektrisch. Dieser Umschwung wird immer stärker gerade mit dem anstehenden Verbrenner-Verbot ab 2035. [1] Mit dem Wechsel von größtenteils Diesel- und Benzinmotoren zu Elektromotoren muss auch die Infrastruktur geändert werden. Neben den klassischen Zapfsäulen an Tankstellen für die Verbrenner-Motoren, benötigt man auch immer mehr Ladestationen für Elektrofahrzeuge. Das Ziel dieser Ladestationen ist, dass Laden für den Bediener so einfach wie möglich zu gestalten. Von dem Bediener verborgen, existiert jedoch zwischen den elektrischen Fahrzeugen und den Ladestationen eine Kommunikation, die sicherstellt, dass der Ladevorgang ordnungsgemäß funktioniert. Bei dieser Kommunikation kommt es, wie in vielen anderen Bereichen der Elektromobilität zu vielen Messdaten.

Problemstellung

Die Verwertung aller Messdaten unmittelbar nach ihrer Generierung ist oft nicht machbar, oder man möchte die Daten zu einem späteren Zeitpunkt erneut betrachten. Hier kommt die Speicherung ins Spiel, die es uns ermöglicht, diese Daten sicher und effizient aufzubewahren. Eine ideale Lösung besteht darin, die Messdaten digital auf der Festplatte eines Rechners zu sichern. Bei der Verarbeitung der Daten sollte es keinen Unterschied zwischen frisch gewonnenen

Messdaten und bereits aufgezeichneten Daten geben. Dieses Ziel ist erreichbar, wenn wir eine zeitsynchrone und verlustfreie Aufzeichnung sicherstellen. Darüber hinaus ist es ratsam, die Daten in ein standardisiertes Format zu überführen. Dies gewährleistet eine klare Handhabung und sorgt für Eindeutigkeit in der weiteren Verwendung und Analyse der Daten.

Speicherformat

Das ausgewählte Speicherformat für dieses Projekt, das Measurement Data Format (MDF), sorgt für eine konsistente Darstellung unserer Messdaten. Bemerkenswert ist, dass die Association for Standardization of Automation and Measuring Systems (ASAM) seit 2009 eine erneuerte Version dieses Formats als Industriestandard, bekannt als ASAM MDF, etabliert hat [3]. Im Herzen dieses Standards liegen Datenblöcke, die in einer einzigartigen Weise strukturiert sind (Abbildung 1 illustriert dies anschaulich). Der ID-Block beherbergt grundlegende Informationen, etwa den Autor des Formats. Neben dem ID-Block ist der Header-Block unverzichtbar. Er fungiert als Wegweiser, der Verknüpfungen zu den nachfolgenden Blöcken enthält. Die Data-Group-Blöcke dienen als Speicher für die Messdaten. Hier werden die rohen Daten als Bit-Werte in den Data-Blöcken gespeichert und die Interpretation der Daten wird in Channel-Group-Blöcken gespeichert. Je nach Zusammensetzung der Messdaten werden ihrer Interpretationen, innerhalb eines Channel-Group-Blocks, in ein oder mehrere Channel-Blöcke unterteilt.

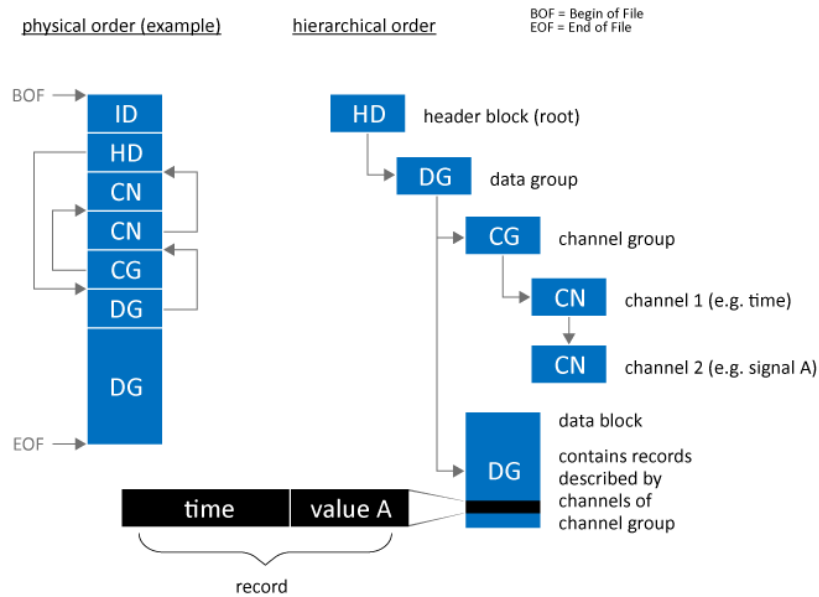


Abb. 1: Beispiel einer einfachen Blockstruktur einer MDF-Datei [4]

Lösungsansatz

Die Wahl des MDF-Formats zur Speicherung von Messdaten ist ein Meilenstein in Sachen Kompatibilität mit diversen Auswertungsprogrammen. Die Herausforderung, die durch die zeitaufwendige Datenspeicherung auf die Festplatte entsteht, insbesondere bei schnell generierten Daten, erfordert innovative Lösungen. Unsere Antwort darauf ist unter Anderem der Einsatz eines effizienten Zwischenspeichers (siehe Abbildung 2), der die Daten puffert, während sie in der richtigen Reihenfolge auf die Festplatte geschrieben werden.

Hier spielt der FIFO-Speicher (First In - First Out) eine Schlüsselrolle, indem er die Datenreihenfolge aufrechterhält und die zeitliche Synchronität sicherstellt. Die Hinzufügung von Zeitstempeln dient dabei als essentieller Mechanismus zur Sicherung der zeitlichen Abstände zwischen den Messdaten. Die Normkonformität der Datenspeicherung nach dem MDF-Format birgt ebenso seine Tücken, die beim Testen zur Interoperabilität mit verschiedenen Programmen und die unterschiedlichen Interpretationsmöglichkeiten der Norm gezeigt haben.

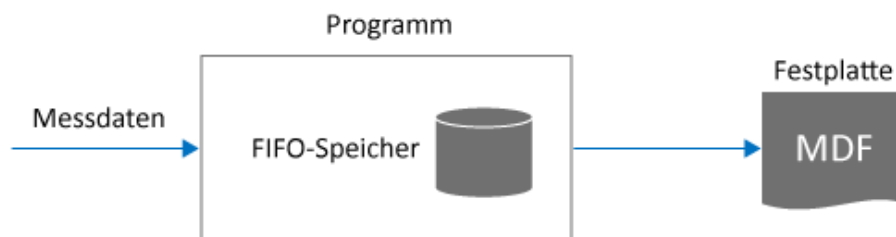


Abb. 2: Darstellung des Programms mit Zwischenspeicher [2]

Ausblick

Die Speicherung von verschiedenen generierten Testdaten funktioniert mit der aktuellen Implementierung. Auch die Interoperabilität mit anderen Programmen

wurde sichergestellt und durch verschiedene Tests bestätigt. Im weiteren Verlauf der Bachelorarbeit wird noch die finale Validierung mit verschiedenen realen Daten unter verschiedenen Bedingungen durchgeführt.

Literatur und Abbildungen

- [1] Mathias Brandt. Fast 50% der Neuwagen sind (teil)elektrisch. <https://de.statista.com/infografik/2870/neuzulassungen-von-hybrid-und-elektroautos-in-deutschland/>, 2023.
- [2] Eigene Darstellung.
- [3] ASAM eV. Standards ASAM MDF History. <https://www.asam.net/standards/detail/mdf/wiki/>, 2023.
- [4] ASAM eV. Standards ASAM MDF Technical Content. <https://www.asam.net/standards/detail/mdf/wiki/>, 2023.