

Artificial Intelligence for Digitalizing Industry

Mathias Schneider, M.Eng.
Seifeddine Saadani, M.Eng.
Prof. Dr.-Ing. Alfred Höß



(Verbundnr. ES2ECS18201)

Zusammenfassung

Begünstigt durch die kontinuierlichen Verbesserungen der Rechenleistung in den vergangenen Jahrzehnten erlebt die Künstliche Intelligenz (KI) in den letzten Jahren ihre Renaissance. Im Zuge der fortschreitenden Digitalisierung und den dadurch verfügbaren Datenmengen, eignen sich KI-Methoden der nächsten Generation, um u. a. Muster zu erkennen oder den Trend in Zeitreihen sehr akkurat voraussagen zu können. Im Rahmen des Forschungsprojektes "Artificial Intelligence for Digitalizing Industry" – *AI4DI* – sollen weitere Schritte unternommen werden, Methoden der Künstlichen Intelligenz in die Digitalisierung verschiedenster Industriezweige einfließen zu lassen.

Gefördert wird das Projekt von der Europäischen Union im Rahmen des H2020 ECSEL JU sowie den involvierten nationalen Förderern in Deutschland dem Bundesministerium für Bildung und Forschung (BMBF).

Die Automotive-Forschungsgruppe der Ostbayerischen Technischen Hochschule Amberg-Weiden (OTH AW) erforscht in diesem Projekt über einen Zeitraum von drei Jahren die Optimierungsmöglichkeiten seitens Latenz, Datendurchsatz und Energieeffizienz von konsekutiven Datenverarbeitungsprozessen in verteilten Systemen. Ein Anwendungsbezug besteht hierbei zum Themenkomplex von Mobility as a Service (MaaS).

Abstract

Prospering by the continuous improvements in computation performance during the last decade, Artificial Intelligence (AI) undergoes its renaissance in the previous years. Progressing digitization provides the amount of data, which are required to leverage next generation of AI methods, including algorithms for pattern recognition and prediction of trends for time-series regression problems. The European research project "Artificial Intelligence for Digitalizing Industry" (AI4DI) investigates the potential of these methods in various industries, accompanying the ongoing digitalization.

The project is funded in the program H2020 ECSEL JU by the European Union as well as the involved national authorities, including the Federal Ministry of Education and Research (BMBF) in Germany.

In the next three years, the Automotive-Research Team of the Ostbayerische Technische Hochschule Amberg-Weiden (OTH AW) investigates concepts for optimizing latency, throughput and power efficiency of data processing pipelines in distributed systems applied in the field of Mobility as a Service (MaaS).

1 Einleitung

Dass sich hinter der Künstlichen Intelligenz (KI) weit mehr als nur ein rein modisches Schlagwort versteckt, bestätigen die aufgetretenen – teils bahnbrechenden – technologischen Erfolge in diesem Feld innerhalb des noch jungen Jahrtausends. Eine Begründung hierfür

liefert die Kombination von fortlaufendem Fortschritt verfügbarer Rechenleistung gemäß des Moore'schen Gesetzes und der Zugang zu sehr großen Datenmengen. Die heutigen Ansätze in der Algorithmik unterscheiden sich dagegen nicht sehr von ihren Ursprüngen aus den letzten Jahrzehnten.

4 Verteiltes Datenverarbeitungssystem

Da diese Transportsysteme aufgrund der gegebenen Mobilität räumlich verteilt sind und gerade die Fahrzeugperipherie nur mit eingeschränkten Rechenressourcen ausgestattet ist, liegt ein Schwerpunkt der Betrachtung der OTH AW darin, arbeitsintensive Rechenprozesse (z. B. Fußgänger-detektion) dynamisch und effektiv im gegebenen Rechnernetzwerk aufzuteilen. In diese Untersuchungen werden dabei sowohl heterogene Rechenkomponenten (CPU, GPU und KI-Beschleuniger), als auch verschiedenste Kommunikationstechnologien wie Ethernet, IEEE 802.11ac und Mobilfunk eingebunden. Diese Komponenten zeichnen sich durch unterschiedliche Kapazitätseigenschaften aus. Während Recheneinheiten sich durch Merkmale, wie die maximalen Flops, verfügbarer Arbeitsspeicher und Speicherzugriffszeiten unterscheiden, differenzieren sich Kommunikationstechnologien durch Datendurchsatz und Latenz (Abbildung 4). Ziel der Betrachtung soll es sein, Datenverarbeitungsketten, welche in einem gerichteten azyklischen Graphen (engl. directed acyclic graph, kurz DAG) abbildbar sind, möglichst effizient auf dieses Netzwerk an Rechnern zu verteilen. Diese zweiteilige Optimierungsaufgabe, bestehend aus Partitioning und Scheduling, lässt sich hierbei auf das sogenannte Flow-Shop-Problem für Maschinenbelegungspläne zurückführen [7]. Diesem Problem ist nachgewiesen, NP-vollständig zu sein [8]. Da diese Problemstellung bereits seit Mitte der 1950er untersucht wird, gibt es eine Vielzahl an bekannten, rechenaufwändigen Algorithmen, darunter der Johnson [9] und der Heterogeneous Earliest Finish Time (HEFT) Algorithmus [10], welche jeweils in der Lage sind, eine nahezu optimale Lösung zu finden.

ITS und im Speziellen das "Last-Mile"-Szenario stellen dabei jedoch deutlich höhere Anforderungen an die Recheneffizienz dieser Algorithmen. Durch die Mobilität des Fahrzeugs, welches Auswirkungen auf verfügbare Re-

cheneinheiten (z.B. Edge-Computer in der Infrastruktur) sowie auf die Qualität der Kommunikationskanäle hat, wird es notwendig, das System kontinuierlich zu optimieren. Zudem soll es möglich sein, neue Verarbeitungsschritte der Datenverarbeitungskette hinzuzufügen oder bestehende zu entfernen und dementsprechend die Verteilung mit geringem Aufwand anzupassen. Hierfür sind die traditionellen und rechenaufwändigen Algorithmen nicht geeignet, da jede Veränderung im System eine neue, kostenaufwändige Berechnung der neuen Aufteilung nach sich zieht. Die Untersuchung der OTH-AW beschäftigt sich daher damit, geeignete KI-Methoden, z.B. aus dem Bereich des Bestärkenden Lernens (engl.: *Reinforcement Learning*), heranzuziehen. Durch deren Verwendung soll es nun ermöglicht werden, das gesamte System dynamisch und effizient zu optimieren. Erste Ansätze haben hierbei ein Verbesserungspotenzial bereits nachgewiesen [11]. Als finales Ziel soll die Datenverarbeitungskette dann je nach Anforderung auf verschiedene Kriterien wie Latenzzeiten, Datendurchsatz oder der Energieverbrauch auf das Rechnernetzwerk verteilt werden.

Um solche Modelle und Algorithmen zu bewerten, wird eine Simulationsumgebung erforderlich, welche es ermöglicht, aus definierten Szenarien Benchmarks zu erzeugen und so den Mehrwert der KI-Anwendung experimentell nachzuweisen. Diese Umgebung, welche z. B. auf SimDAG für Netzwerkanalysen aufbaut, soll im Lauf des Projektes um die benötigten dynamischen Aspekte erweitert werden.

5 Ausblick

Da sich das Projekt noch in einem frühen Stadium befindet, werden derzeit die frühen Phasen innerhalb des Projektablaufs gemäß dem V-Modell bearbeitet. Diese umfassen die Anforderungsanalyse und das Systemdesign. Ein regelmäßiger Austausch mit den direkten

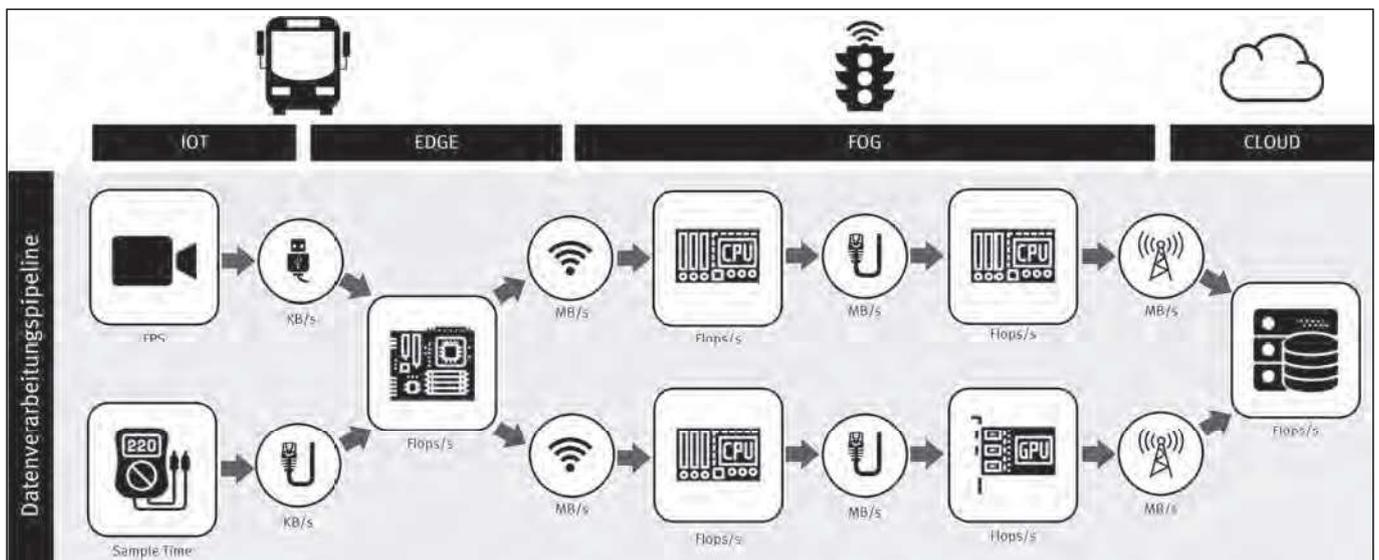


Abbildung 4: Heterogene Prozessor- und Kommunikationstechnologien in Fahrzeug, Infrastruktur und Cloud

Projektpartnern innerhalb des Demonstrators wird dabei forciert. Hierdurch soll eine einheitliche Vision der Aufgabe innerhalb der Wertschöpfungskette gefestigt und frühzeitig Unregelmäßigkeit im Projektablauf erkannt werden. Da AI4DI zudem darauf abzielt, den europäischen Fördergebern einen Leitfaden für den Einsatz von KI in der Industrie bereitzustellen, sollen konsortiumsweite

Workshops dazu dienen, die Ziele und Möglichkeiten der KI unabhängig des jeweiligen Anwendungsfalls und Industriezweigs zu definieren. Dieser Leitfaden soll dabei bestehende Standardisierungsprozesse wie die ISO Norm für KI (ISO/IEC JTC 1/SC 42) ergänzen und wird als einer der ersten Ergebnisberichte Ende 2019 an die europäische Kommission übermittelt.

Referenzen:

- [1] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, ... & S. Dieleman, "Mastering the game of Go with deep neural networks and tree search", *nature*, 529 (7587), 484, 2016
- [2] Tractica, "Artificial Intelligence Market Forecasts 2Q 2017.", Boulder, CO 80302 USA. Online verfügbar: <https://www.tractica.com/wp-content/uploads/2017/04/MD-AIMF-2Q17-Executive-Summary.pdf> [Zugegriffen am 19.9.2019].
- [3] Bundesministerium für Bildung und Forschung, „AI4DI — Elektroniksysteme für Künstliche Intelligenz in der digitalen Industrie“, 2019. Online verfügbar: <https://www.elektronikforschung.de/projekte/ai4di> [Zugegriffen am 19.9.2019].
- [4] AI4DI Consortium, "Consortium", 2019. Online verfügbar: <https://ai4di.automotive.oth-aw.de/index.php/consortium> [Zugegriffen am 19.9.2019].
- [5] ECSEL JU, "Multi-Annual Strategic Plan", 2018, ECSEL JU. Online verfügbar: https://ec.europa.eu/research/participants/data/ref/h2020/other/legal/jtis/ecsel-multi-stratplan-2018_en.pdf, [Zugegriffen am 19.9.2019].
- [6] VTT Research, "Automated Driving", 2019. Online verfügbar: <https://www.vttresearch.com/services/sustainable-and-smart-city/transport/vehicle-systems>, [Zugegriffen am 19.9.2019].
- [7] S. M. Johnson, "Optimal two- and three-stage production schedules with setup times included", *Naval Research Logistics*, vol. 1, no. 1, pp. 61–68, Mar. 1954
- [8] M. R. Garey, D. S. Johnson, and R. Sethi, "The Complexity of Flowshop and Jobshop Scheduling", *Mathematics of OR*, vol. 1, no. 2, pp. 117–129, May 1976.
- [9] H. Allaoui, S. Lamouri, and A. Artiba, "Using Johnson's algorithm to approximate some flow shop scheduling problems with unavailability periods", *IFAC Proceedings Volumes*, vol. 39, no. 3, pp. 45–50, 2006.
- [10] H. Topcuoglu, S. Hariri, and Min-You Wu, "Task scheduling algorithms for heterogeneous processors", in *Proceedings. Eighth Heterogeneous Computing Workshop (HCW'99)*, San Juan, Puerto Rico, 1999, pp. 3–14.
- [11] Q. Wu, Z. Wu, Y. Zhuang, and Y. Cheng, "Adaptive DAG Tasks Scheduling with Deep Reinforcement Learning", in *Algorithms and Architectures for Parallel Processing*, vol. 11335, J. Vaidya and J. Li, Eds. Cham: Springer International Publishing, 2018, pp. 477–490.

Projektpartner:

siehe Abbildung 1

Fördergeber:

AI4DI has received funding within the Electronic Components and Systems for European Leadership Joint Undertaking (ECSEL JU) in collaboration with the European Union's H2020 Framework Programme and National Authorities, under grant agreement n° 826060.



Kontakt:



Mathias Schneider, M.Eng.

Ostbayerische Technische Hochschule (OTH) Amberg-Weiden
Fakultät Elektrotechnik, Medien und Informatik
Kaiser-Wilhelm-Ring 23
92224 Amberg

mat.schneider@oth-aw.de



Seifeddine Saadani, M.Eng.

Ostbayerische Technische Hochschule (OTH) Amberg-Weiden
Fakultät Elektrotechnik, Medien und Informatik
Kaiser-Wilhelm-Ring 23
92224 Amberg

se.saadani@oth-aw.de



Prof. Dr.-Ing. Alfred Höß

Ostbayerische Technische Hochschule (OTH) Amberg-Weiden
Fakultät Elektrotechnik, Medien und Informatik
Vizepräsident Forschung und Technologietransfer,
wissenschaftlicher Nachwuchs
Kaiser-Wilhelm-Ring 23
92224 Amberg

a.hoess@oth-aw.de