



5

AI: Science
over Fiction

Wie kann die KI helfen, bessere Dispositionsentscheidungen zu treffen?

Faktenpapier

Aus der Serie: AI: Science over Fiction

www.bitkom.org

bitkom

Herausgeber

Bitkom
Bundesverband Informationswirtschaft,
Telekommunikation und neue Medien e. V.
Albrechtstraße 10 | 10117 Berlin
T 030 27576-0
bitkom@bitkom.org
www.bitkom.org

Verantwortliches Bitkom-Gremium

AK Artificial Intelligence

Projektleitung

Dr. Nabil Alsabah | Bitkom e. V.

Autoren

Thomas Thiele, Christine Carl, Patrick Goldschmidt, Jan-Christoph Jähne, Tobias Keller,
Georg Merz, Enno Middelberg, Anna Plazotta, Stephan Kaulbach, Rene Neuhäuser,
Sabina Jeschke | Deutsche Bahn AG

Lektorat

Anna Suhodolova | Bitkom e. V.

Satz & Layout

Katrin Krause | Bitkom e. V.

Titelbild

© Liudmila Luchkina | unsplash.com

Copyright

Bitkom 2020

Diese Publikation stellt eine allgemeine unverbindliche Information dar. Die Inhalte spiegeln die Auffassung im Bitkom zum Zeitpunkt der Veröffentlichung wider. Obwohl die Informationen mit größtmöglicher Sorgfalt erstellt wurden, besteht kein Anspruch auf sachliche Richtigkeit, Vollständigkeit und/oder Aktualität, insbesondere kann diese Publikation nicht den besonderen Umständen des Einzelfalles Rechnung tragen. Eine Verwendung liegt daher in der eigenen Verantwortung des Lesers. Jegliche Haftung wird ausgeschlossen. Alle Rechte, auch der auszugsweisen Vervielfältigung, liegen beim Bitkom.

In den letzten zehn Jahren ist die Zahl der Fahrgäste im deutschen Schienenverkehr stetig gestiegen, rund 148 Millionen Fahrgäste nutzten 2018 die Fernverkehrszüge der Deutschen Bahn. Der steigende Kapazitätsbedarf erfordert neue Methoden zur Planung und zum effizienten Management von Zügen und Infrastruktur. In diesem Artikel wird ein Ansatz vorgestellt, der darauf abzielt, diese Dispositionsentscheidungen im Schienenverkehr zukünftig durch Reinforcement Learning zu lösen.

Reinforcement Learning hat in den vergangenen Jahren große Fortschritte beim Lösen von Brett- und Computerspielen gemacht. In diesen müssen Abwägungsentscheidungen auf strategischer Ebene getroffen werden – teilweise mit widersprüchlichen Zielen. Der vorliegende Artikel beschäftigt sich mit der Herausforderung, diese Erfolge auf reale Probleme im schienengebundenen Verkehr zu übertragen. So sollen in besonders komplexen Situationen des Nahverkehrs Entscheidungen mit Blick auf eine Optimierung der Gesamtsituation getroffen werden. Es wird gezeigt, dass sich der Geschäftsprozess Disposition auf Basis historischer Daten in einen Reinforcement Learning Ansatz für S-Bahnen übersetzen lässt. Dabei wird auf die verwendete Simulation und die Modellierung eingegangen, zudem werden die Ergebnisse erster Experimente gezeigt.

1 KI unterstützt bei der Ressourcenplanung im Schienenverkehr

Der Zuwachs an Mobilität gilt als ungebrochen, besonders für die Nutzung von Rad, Bus und Bahn ergeben sich vermehrt Zuwächse in der erbrachten Fahrleistung¹. Um diesen wachsenden Bedarfen gerecht zu werden, reagieren Mobilitätsdienstleister mit dem Aufbau von Kapazitäten und Infrastrukturen. Die Steuerung dieser Ressourcen im Betrieb erfolgt in der sogenannten Verkehrsdisposition. Die Disponenten der Leitstelle werden bei dieser Aufgabe bereits heute durch Software unterstützt. Durch Ausbildung, Erfahrung und geplante Störfallprogramme beherrscht das Team der Leitstelle komplexe Situationen, in denen Störungen eintreten. Jedoch verhindert die permanente Folge von Ereignissen im Verkehr eine regelmäßige Rekapitulation und Auswertung der getroffenen Entscheidungen. Insbesondere die komplexen Folgeeffekte der möglichen Entscheidungen sind schwer einschätzen, geschweige denn die ausführliche Bewertung von Handlungsoptionen.

Ziel des vorliegenden Beitrags ist die Darstellung eines KI-basierten Ansatzes für die zukünftige Verkehrsdisposition, um angesichts wachsender Komplexität im Schienenverkehr den Disponenten eine Entscheidungsunterstützung zu geben. Der Beitrag umfasst neben der Darstellung des aktuellen Forschungsstands, eine Analyse der Ausgangssituation am Beispiel der S-Bahn Stuttgart, eine methodische Darstellung und eine Evaluation initialer Experimente.

Vielfältige Umweltfaktoren beeinflussen Dispositionsentscheidungen

Die Disponenten disponieren zwar jeweils nur für einen Teilabschnitt des Streckennetzes, müssen jedoch den gesamten Verkehr im Blick haben. Dies führt zu einem dreidimensionalen Problem. Neben den Zügen unterschiedlicher Anbieter sind Fahrzeuge mit unterschiedlichen Moden im Netz aktiv. Fernverkehrszüge fahren gemeinsam mit Regionalzügen auf derselben Infrastruktur, die zusätzlich vom Güterverkehr genutzt wird. Hinzu kommen immer individuellere Nutzungsbedürfnisse der Fahrgäste, die je nach Zeitraum (z.B. Ferien) oder Tageszeit variieren.

In der Disposition wird diesen Anforderungen meist durch Heuristiken Rechnung getragen. So wird ein schnell fahrender ICE fast immer Vorrang vor einer S-Bahn haben. Zum Verkehr gehören aber auch die Fahrgäste, die neben der S-Bahn auch U-Bahnen, Straßenbahnen und Busse benutzen können. Gerade Werktags folgen die Reiseketten der Fahrgäste bestimmten Mustern – je nachdem wo Schulen, große Arbeitgeberstandorte oder besondere Ziele, wie der Flughafen, vorzufinden sind. In den Zügen können sich leicht mehrere hundert Fahrgäste befinden. Gerade zu Stoßzeiten, wenn zu Tagesbeginn viele Bürger zur Arbeit pendeln, müssen die Disponenten über die tagesaktuellen Eigenschaften der Infrastruktur und des Fahrbetriebs Bescheid wissen.

Über das Jahr verteilt finden zudem besondere Ereignisse statt, wie z.B. Feste oder Fußballspiele. Manche dieser Ereignisse können eingeplant werden, andere entwickeln durch Starkregen oder Hitze kurzfristig eine Dynamik, die ein Abweichen vom Plan notwendig macht.

¹ Bundesministerium für Verkehr und digitale Infrastruktur: Mobilität in Deutschland – Kurzreport 2017, Online unter: https://www.bmvi.de/SharedDocs/DE/Anlage/G/mid-2017-kurzreport.pdf?__blob=publicationFile [eingesehen am 30.08.2019], 2018, S.6.

Unerwartete Ereignisse fordern Entscheidungsprozesse in der Disposition heraus

Die Leitstelle verfügt über mehrere digitale Systeme, um den Verkehr im Blick zu behalten und ggf. zu erkennen, wo steuernd eingegriffen werden muss. Die Integration dieser Informationen zu einem Lagebild ist jedoch ein manueller Prozess, den jeder Disponent selbst übernehmen muss. Disponenten handeln entweder bei Ereignissen, die unmittelbar auftreten und einer Handlung bedürfen, wie z.B. einem Notarzteinsatz im Zug, oder bei sich langsam abzeichnenden Verspätungen, die sie frühzeitig eindämmen wollen.

Disponenten haben mehrere Möglichkeiten, wie sie in den Verkehr eingreifen können. Sie können Züge ganz oder teilweise ausfallen lassen, Züge umleiten, die Anzahl der Wagen reduzieren, Fahrzeuge tauschen oder die Reihenfolge an einer Kreuzung verändern. Bei allen Entscheidungen müssen sie stets die Möglichkeiten der Infrastruktur und die Verfügbarkeit von Personal und Zügen im Blick behalten.

Besonders Anomalien, die kurzfristig auftreten, erweisen sich als herausfordernd. Hierzu gehören z.B. technische Störungen oder wetterbedingte Ausfälle. Ein weiteres Beispiel ist ein Notarzteinsatz. Muss ein Zug an einer Haltestelle aufgrund eines Notarzteinsatzes länger halten, blockiert er hinter ihm fahrende Züge. Diese Züge können dann nur mit Verspätung ihr Ziel erreichen. Bei längeren Verspätungen besteht zudem das Risiko, dass die Folgeleistung, die der Zug übernehmen soll, bereits mit Verspätung startet. In diesem Fall kann der Disponent einen Teil der Fahrtstrecke ausfallen lassen. Die Fahrgäste müssen dann einen nachfolgenden Zug nehmen, der den vom Notfall betroffenen Haltepunkt wieder passieren kann. Der aufgehaltene Zug muss nicht warten und kann – unter Berücksichtigung der entfallenen Halte – wieder planmäßig eingesetzt werden. An diesem Beispiel wird deutlich, wie schwierig es ist, Dispositionsentscheidungen unter der Maßgabe sowohl von Pünktlichkeit als auch der Reisekette zu treffen. Analog zu diesem Beispiel lassen sich je nach Art der auftretenden Anomalie andere Störungsmuster identifizieren, auf die jedoch unter Berücksichtigung wechselnder Herausforderungen (tagesaktueller Umweltfaktoren) reagiert werden muss.

2 Algorithmen des verstärkten Lernens können Entscheidungsmodelle sukzessive optimieren

In den letzten Jahren haben sich Ansätze lernender Systeme vornehmlich auf Basis von ML entwickelt. Als Methode der Künstlichen Intelligenz liegt der bedeutendste Mehrwert darin, dass sich das System auf den Output konzentriert und im Prozess des Lernens, Erkenntnisse selbstständig aus den Daten extrahier- und nutzbar macht. Es wird eine Nutzenfunktion aufgestellt, die der Algorithmus optimiert. Dies geschieht, indem einzelne, einfließende Faktoren unterschiedlich gewichtet werden. Lernende Systeme werden oftmals als *Black Boxes* bezeichnet, da für den Menschen nicht immer klar ist, welche Vorgänge und Lernprozesse darin stattfinden. Der entscheidende Vorteil liegt in der Flexibilität des Systems: Es kann im Gegensatz zu starren regelbasierten Systemen neue Umweltzustände erkennen und entsprechend darauf reagieren. Unvorhergesehene Situationen können so mit Hilfe selbsterlernter Regeln evaluiert werden.

Der ML-Prozess kann in unterschiedliche Kategorien eingeteilt werden. Die bekanntesten sind überwachtes, unüberwachtes und verstärkendes Lernen (reinforcement learning, RL). Beim überwachten Lernen wird ein Modell mit einer Datenmenge trainiert. Das System lernt selbstständig Zusammenhänge in den Daten abhängig von einer vorab definierten Ausgangsgröße (z.B. einem Label) herzustellen. Beim unüberwachten Lernen werden Daten ohne eine vorher definierte Ausgangsgröße oder Annotation betrachtet. Das Ziel ist hierbei selbstständig mit dem Input Regeln, Muster sowie Gruppierungen in den Daten zu erkennen.

Beim verstärkenden Lernen wird das Verfahren des Versuch-und-Irrtums verwendet. Der Algorithmus lernt in einer kontinuierlichen Feedback-Schleife, indem sein Verhalten in Belohnung oder Bestrafung resultiert. Grundlegend ist hierfür das Prinzip eines sequentiellen Entscheidungsprozesses, bei dem eine Umgebung (repräsentiert durch einen Status) von einem oder mehreren Agenten (als handelnde Instanzen einer KI) durch unterschiedliche vorgegebene Aktionen manipuliert wird². Hierdurch wird ein Lernprozess initiiert, bei dem die Agenten in mehreren Iterationen die Auswirkungen ihrer Aktionen und deren Verkettung auf die jeweilige Umgebung evaluieren und eine Lösungsstrategie gemäß einer vorab definierten Belohnungsfunktion entwickeln.

Die Vorteile des verstärkenden Lernens wurden in den letzten Jahren auf beeindruckende Weise präsentiert³. Auf Basis von Spielen werden die algorithmischen Ansätze stetig weiterentwickelt und optimiert, sodass Agenten nach ersten, teilweise hilflos wirkenden Aktionen, beeindruckende Strategien entwickeln konnten. In der Übertragung dieser Fortschritte aus Spielen auf ein Problem der realen Welt, wie z.B. Verkehrsdisposition, sind die zentralen Fragestellungen jedoch meist weniger auf die Optimierung des Algorithmus fokussiert. Vielmehr stellt sich die Frage nach der Modellierung, sodass ein Mehrwert in der realen Welt erzielt werden kann.

² R. S. Sutton, A. G. Barto: Reinforcement Learning: An Introduction. MIT press, 2018.

³ O. Vinyals, S. Gaffney, T. Ewalds: DeepMind and Blizzard open StarCraft II as an AI research environment. <https://deepmind.com/blog/deepmind-and-blizzard-open-starcraft-ii-ai-research-environment/> [Letzter Zugriff: 28/11/18], 2017.

Die organisatorischen Weiterentwicklungen und Verbesserungen an den bestehenden IT-Systemen in der Leitstelle, sollen durch ein datengetriebenes Entscheidungsunterstützungssystem ergänzt werden. Dadurch sollen die aktuellen und zukünftigen Herausforderungen bewältigt werden. Ein solches Unterstützungssystem soll Dispositionsentscheidungen vorschlagen und die Auswirkungen von Szenarien und Entscheidungen hinsichtlich Verspätungen und Reiseketten bewerten. Die Grundlage dafür bilden die ML-Verfahren.

Das oben beschriebene Problem soll durch einen verstärkend lernenden Algorithmus gelöst werden. Andere ML-Algorithmenfamilien sind für unsere Zwecke ungeeignet. Der Einsatz von Verfahren des überwachten Lernens würde an der nicht ausreichenden Verfügbarkeit von Daten über Entscheidungen der Vergangenheit scheitern. Zudem wäre selbst bei Vorhandensein einer solchen Datenbasis die Beurteilung der Effekte von Entscheidungen problematisch. Die vorhandenen Daten wären zwar für mehrere Jahre in der Vergangenheit verfügbar, doch ihnen würde jedes Jahr ein neuer Fahrplan bzw. geänderte Infrastruktur zu Grunde liegen, die nicht in den Daten repräsentiert sind.

Mit Blick auf die Komplexität der Dispositionsaufgaben wurden Verfahren – wie ein *Brute-Force*-Ansatz (Durchspielen aller Möglichkeiten) – verworfen. Denn allein die Berechnungszeiten würden die notwendigen Antwortzeiten von wenigen Minuten überschreiten. Hierdurch wären in den meisten Fällen geeignete Reduktionen oder Heuristiken notwendig, die jedoch wiederum zu Fehlern in der Realität führen können.

Für die erste, hier beschriebene Phase der Entwicklung des Entscheidungsunterstützungssystems sind wir in drei Schritten vorgegangen. Zunächst werden die für die S-Bahn vorhandenen Daten explorativ ausgewertet, um eine Idee von der digitalen Abbildung des Dispositionsprozesses zu bekommen. In einem zweiten Schritt wird das Konzept von Gesundheitszuständen als abstrakte bzw. dimensionsreduzierte Darstellungen des S-Bahn Verkehrs untersucht (vgl. Abschnitt 3.2). Zuletzt wird ein Prototyp eines Reinforcement-Learning-Verfahrens entwickelt, der die Geschwindigkeit eines Teilnetzes der Stuttgarter S-Bahn im Experiment signifikant optimiert.

Bei der Datenexploration werden historische Daten über den operativen Betrieb der S-Bahn Stuttgart, insbesondere über die konkreten Zugläufe, ausgewertet. Für die Zugläufe werden Meldeanlagen im Schienennetz genutzt, die zugnummerngenau die Position eines Zugs mit Zeitstempel erfassen. Für die Analyse standen 8 Millionen dieser Meldungen aus zwei Jahren zur Verfügung.

Ziel der Exploration ist es, ein besseres Verständnis für potentiell auftretende Anomalien und deren Effekte zu gewinnen. Unter einer Anomalie wird im Folgenden eine Abweichung vom Fahrplan verstanden. Nur wenn Anomalien auftreten, muss der Disponent eingreifen. Begleitend zur Datenexploration wurden Gespräche mit Domänenexperten aus der Disposition bzw. des Pünktlichkeitsmanagements geführt. Die statistischen Verteilungen und Korrelationsanaly-

sen auf Basis der Zuglaufdaten zeigen, zu welchen Wochentagen und Uhrzeiten Verspätungen auftreten und welche Haltestellen bzw. S-Bahn-Linien davon betroffen sind.

Auch die Entscheidungen der Disponenten und ihre Auswirkungen auf den Verkehr werden untersucht. So kann festgestellt werden, dass Reihenfolgenänderungen häufig zu bestimmten Uhrzeiten ausgeführt werden und welche Auswirkungen Zugausfälle auf das Verkehrsnetz haben. Durch die gezielte Prädiktion des Zeitpunkts eines Zugausfalls kann dieser zu einem früheren Zeitpunkt ausgelöst und damit der entlastende Effekt (z.B. Reduktion von Verspätung) effektiver genutzt werden.

Die Ergebnisse der Datenexploration in der Zusammenarbeit und Analyse mit Domänenexperten waren hilfreich für die nächsten Schritte. So wird unter anderem deutlich, wie das Netz und die neuralgischen Punkte strukturiert sind, die besonderer Beachtung bedürfen. Mit den Zugausfällen und Reihenfolgenänderungen zeichneten sich zwei Entscheidungstypen ab, die von besonderem Interesse sind.

3 Verstärktes Lernen eignet sich für die Modellierung von Verkehrsentscheidungen

Für den Übergang zwischen Spiel und Realität wird eine geeignete Verkehrssimulation als Umgebung für das verstärkende Lernen genutzt. Diese startet mit unterschiedlichen, echten Situationen aus der Vergangenheit der S-Bahn. Für die Verkehrssimulation wird die Open-Source-Verkehrssimulation SUMO verwendet⁴. Mittels dieser wurde ein neuralgischer Punkt im S-Bahn-Netz, die Stammstrecke im Zentrum von Stuttgart, schematisch abgebildet. Haltestellen und Signale sind als Knoten und Strecken als Kanten modelliert (vgl. Abbildung 1). Die Entfernungen und Geschwindigkeiten der Züge, die mit SUMO auf diesem vereinfachten Netz abgebildet werden, entsprechen den Zuglaufdaten der S-Bahn.



Abbildung 1: Simulierter Ausschnitt des S-Bahn Netzes Stuttgart

Der Fokus auf den S-Bahn Verkehr ergibt sich aus einer hohen Anzahl beförderter Fahrgäste in Kombination mit einer hohen Taktung des Verkehrs. Die S-Bahn Stuttgart befördert an Werktagen 435.000 Fahrgäste auf einem Streckennetz von 215 km mit 83 Haltestellen. Die 838 Züge pro Werktag fahren nach Fahrplan auf einem Schienennetz, das zu Teilen auch von Regional-, Fernverkehrs- und Güterzügen genutzt wird. Sie gehört damit zu den fünf größten S-Bahnen in Deutschland. Der Verkehr der S-Bahn wird von einer eigenen Leitstelle überwacht und geregelt, die in erster Linie die Abweichungen vom Fahrplan überwacht, um dann Maßnahmen zur Korrektur einzuleiten. Die Leitstelle verfolgt das Ziel die Qualität ihres Produkts S-Bahn hoch zu halten, indem sie Pünktlichkeit wiederherstellt und für durchgängige Reiseketten sorgt.

⁴ P. A. Lopez, M. Behrisch, L. Bieker-Walz, J. Erdmann, Y. Flötteröd, R. Hilbrich, L. Lücken, J. Rummel, P. Wagner, E. Wießner: Microscopic Traffic Simulation using SUMO. IEEE Intelligent Transportation Systems Conference (ITSC), 2018.

Die Verspätung von abfahrenden Zügen hat sich als eine zentrale Variable herauskristallisiert

Mit dem Abschluss der explorativen Datenanalyse zeigte sich, dass Anomalien in Form von Verspätungen zwar als Wirkung klar erkennbar sind, die sich andeutenden Ursachen aber immer mehrere Dimensionen betreffen. Um diesen mehrdimensionalen Zustandsräumen Rechnung zu tragen, haben wir uns von der Idee zu Gesundheitszuständen inspirieren lassen, die im Verlauf abstrakt auf das Verkehrsnetz übertragen wird.⁶ Ein Vorteil von Gesundheitszuständen ist, dass sie sich perspektivisch als Basis für eine Belohnungsfunktion des RL-Ansatzes eignen (siehe Absatz 3.1).

Bei der Modellierung der Gesundheitszustände wird deutlich, dass es sowohl lokale, auf konkrete Haltestellen beschränkte, Zustände geben muss, als auch einen globalen Zustand. Der Grund ist die Annahme, dass nicht jede Entscheidung alle Punkte im Netz positiv oder negativ beeinflussen wird, und dass es immer eine Abwägung zwischen lokalen und globalen Optima geben muss.

Für eine erste Analyse sind Gesundheitszustände pro Haltestelle an einem Tag errechnet worden. Zum Einsatz kam unter anderem t-SNE⁷ als ein grundlegendes Verfahren der Dimensionsreduktion (unüberwachtes Lernen). In Verbindung mit einer visuellen Analyse der entstehenden Cluster im zweidimensionalen Zustandsraum wurde ein Maß definiert, was ein schlechter bzw. sehr guter Gesundheitszustand in absoluten Zahlen ist (vgl. Abbildung 2).

In einem zweiten Schritt wird versucht, den Gesundheitszustand einer Haltestelle mithilfe von Modellen zu erklären. Insbesondere zeigte sich, dass Gesundheitszustände stärker mit der Verspätung von abfahrenden Zügen in Zusammenhang stehen als mit der Verspätung von ankommenden Zügen. Insgesamt wird so ein Grundstein umgelegt, um in späteren Projektstadien an jeder Haltestelle eine Prognose für den Verkehr (und evtl. Störungen) zu erzeugen und so mittels Verfahren des überwachten Lernens Voraussagen treffen zu können.

Der RL-Ansatz hat deutliche Verbesserungen in der Disposition bewirkt

Zur Evaluation des oben beschriebenen RL-Ansatzes ist ein Experiment aufgebaut worden, um die grundsätzliche Machbarkeit zu zeigen. Die Ziele des Experiments sind:

1. Eine Konfiguration der Modelle zu finden, die über den Trainingszeitraum eine positive Entwicklung der Belohnung erreicht,
2. und in ihrer Leistung einen selbstgesteckten Anfangszustand schlägt,
3. und visuell nachvollziehbare Ergebnisse in der Simulation zeigt.

⁶ Raghu, Aniruddh, et al.: Continuous state-space models for optimal sepsis treatment-a deep reinforcement learning approach. arXiv preprint arXiv:1705.08422, 2017.

⁷ Maaten, Laurens van der, and Geoffrey Hinton: Visualizing data using t-SNE. Journal of machine learning research. S. 2579-2605, 2008.

Die Umgebung ist als ein schematisches Streckennetz modelliert, das die Stammstrecke am Stuttgarter Hauptbahnhof abbildet. Aus zwei Richtungen fahren Züge auf eine Kreuzung zu (vgl. Abbildung 3). Aus einer der beiden Richtungen kommen dabei doppelt so viele Züge wie aus der anderen Richtung. Der Agent hat die Möglichkeit in jedem Simulationsschritt die Ampelschaltung zu ändern und damit festzulegen, wer zuerst fahren darf. Der Anfangszustand bediente die Züge nach dem »first come, first served«-Prinzip.

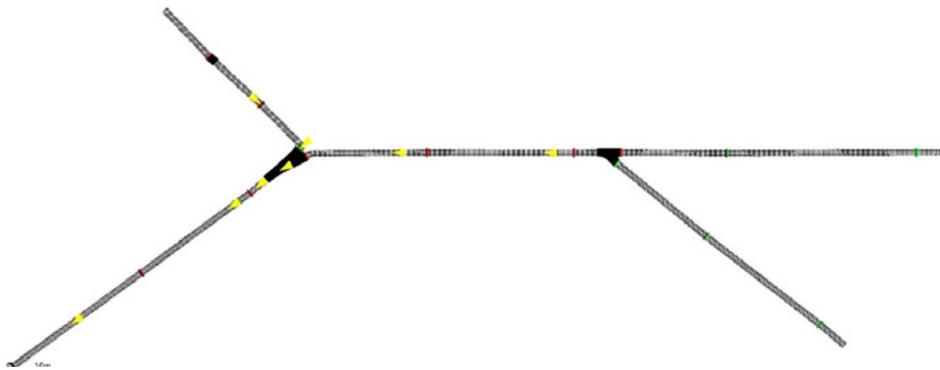


Abbildung 3: Darstellung der vereinfachten Simulation der Stammstrecke

Für die Belohnung wird die durchschnittliche Geschwindigkeit der im Netz fahrenden Züge berechnet. Der Agent hat die Aufgabe, diese Geschwindigkeit zu maximieren. Die Züge fahren in einer festgelegten Reihenfolge in das Netz ein, wodurch die Belohnung in jedem Simulationslauf mit ca. 7,6 m/s gleich ist. Die Belohnung entwickelt sich nach ca. 50 Simulationsläufen besser als der Anfangszustand und weist im weiteren Verlauf eine gewisse Stabilität auf (vgl. Abbildung 4).

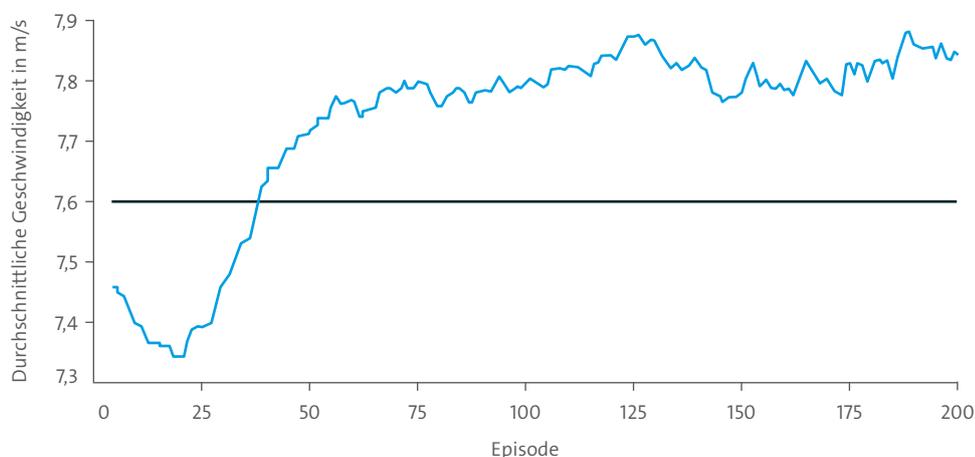


Abbildung 4: Evaluation der durchschnittlichen Geschwindigkeit des Reinforcement Learning Ansatzes (blau) gegenüber der Baseline (schwarz)

Die Betrachtung der Simulation zeigt, dass das »first come, first served«-Prinzip zu einem Stau auf der Strecke mit doppelter Zuganzahl führte. Mittels des RL-Ansatzes wurde eine Ampelschaltung gelernt, die diese Strecke begünstigte und ihr häufiger die Vorfahrt gewährte.

Die Grenzen dieses Experiments mit Blick auf die Herausforderungen im oben beschriebenen Dispositionsprozess sind vielfältig. Es kann jedoch gezeigt werden, dass das verstärkende Lernen nicht nur mit Standard-Szenarien funktioniert. Eine Übertragung auf individuelle Probleme ist durchaus möglich. Wesentliche Einflussfaktoren, die in diesem Experiment zugunsten der Machbarkeit bewusst exkludiert wurden, sind

- die Größe des Netzwerks, das derzeit auf einen wichtigen Ausschnitt beschränkt ist (Stammstrecke mit Tunnel als Engpass),
- eine größere Auswahl an Aktionen des RL-Ansatzes (z.B. direkte Steuerung von Geschwindigkeiten) und
- Einflussfaktoren der Umwelt (Wetter, Bahnhöfe, Passagiere, Fahrplan usw.).

4 In Zukunft werden RL-Agenten mehr Handlungsspielräume haben

Die Herausforderungen an die Disposition im S-Bahn-Verkehr werden in Zukunft noch steigen. Aktuell fährt auf einer Linie alle 15 Minuten ein Zug während einer Stoßzeit. Zusätzlich folgt alle 30 Minuten ein weiterer Zug, im sogenannten Nebentakt. Ab 2021 wird dann voraussichtlich durchgängig ein 15 Minuten-Takt gefahren. Diese »permanente Stoßzeit« nimmt der Disposition die Handlungsoption, bestimmte Verspätungsszenarien ohne einen Eingriff in den Verkehr aufzulösen. Mit der Einführung der digitalen Signalisierung (ETCS) wird in der Zukunft eine noch größere Kapazität für die Fahrgäste zur Verfügung stehen, da mehr Züge in kürzeren Abständen auf den Schienen fahren können. Die Komplexität in den Leitstellen der EVU steigt damit im gleichen Maße.

Im Zusammenspiel dieser beiden zentralen Randbedingungen zeigt der vorliegende Ansatz die prinzipielle Machbarkeit eines RL-basierten Ansatzes auf. Eine Indikation zum Mehrwert findet sich in der erhöhten Durchschnittsgeschwindigkeit, die trotz des Ausschnitts aus dem Stuttgarter Netz einen stabilen Zuwachs erwarten lässt.

In weiterführenden Arbeiten werden die in der Evaluation aufgezeigten Reduktionen des Experiments Schritt für Schritt abgebaut. Dies führt zu mehr Handlungsspielraum einerseits für die RL-Agenten, andererseits hinsichtlich der Verkehrsszenarien im System. Eine systematische Versuchsplanung zur Evaluation des Zusammenhangs zwischen der Komplexität im System und des erwarteten Zuwachses an Verkehrsleistung, ist daher für die weitere Entwicklung des Ansatzes wesentlich. Ein zentraler Aspekt ist die sukzessive Näherung an die Realität, um letztlich einen Mehrwert in der realen Welt zu erzielen und den bestehenden Ansatz von rein Geschwindigkeits-orientierten Parametern weiterzuentwickeln.

Bitkom vertritt mehr als 2.700 Unternehmen der digitalen Wirtschaft, davon gut 1.900 Direktmitglieder. Sie erzielen allein mit IT- und Telekommunikationsleistungen jährlich Umsätze von 190 Milliarden Euro, darunter Exporte in Höhe von 50 Milliarden Euro. Die Bitkom-Mitglieder beschäftigen in Deutschland mehr als 2 Millionen Mitarbeiterinnen und Mitarbeiter. Zu den Mitgliedern zählen mehr als 1.000 Mittelständler, über 500 Startups und nahezu alle Global Player. Sie bieten Software, IT-Services, Telekommunikations- oder Internetdiensten an, stellen Geräte und Bauteile her, sind im Bereich der digitalen Medien tätig oder in anderer Weise Teil der digitalen Wirtschaft. 80 Prozent der Unternehmen haben ihren Hauptsitz in Deutschland, jeweils 8 Prozent kommen aus Europa und den USA, 4 Prozent aus anderen Regionen. Bitkom fördert und treibt die digitale Transformation der deutschen Wirtschaft und setzt sich für eine breite gesellschaftliche Teilhabe an den digitalen Entwicklungen ein. Ziel ist es, Deutschland zu einem weltweit führenden Digitalstandort zu machen.



**Bundesverband Informationswirtschaft,
Telekommunikation und neue Medien e.V.**

Albrechtstraße 10

10117 Berlin

T 030 27576-0

F 030 27576-400

bitkom@bitkom.org

www.bitkom.org

bitkom