

Datenverwaltung aus dem Digitalen Andreaskreuz durch Open-Source Big-Data-Technologien

Managing digital level crossing data using open-source big data technologies

Sangeetha Shankar | Miriam Grünhäuser

Das Forschungsprojekt „Digitales Andreaskreuz“ ist ein Teilprojekt aus dem Reallabor Hamburg (RealLabHH) und hat die Verbesserung von Mobilität und Komfort der Straßenverkehrsteilnehmenden durch die Digitalisierung von Bahnübergängen (BÜ) zum Ziel. Es wurde ein Proof-of-Concept-Demonstrator entwickelt, der während der Feldtests regelmäßige Statusmeldungen an Verkehrsteilnehmende gesendet und sie mit zusätzlichen Informationen versorgt hat. Dieser Beitrag befasst sich mit der Verwaltung der Daten, die von diesen BÜ gesendet wurden, und der Entwicklung einer Simulations-Middleware, um die Daten für alle BÜ in Hamburg zu generieren. Eine Datenverarbeitungskette mit verschiedenen Open-Source-Technologien wurde für dieses Datenmanagement entwickelt.

1 Einführung

Das Projekt RealLabHH hatte zum Ziel, ein Reallabor in Hamburg bis Ende 2021 aufzusetzen, um umfassend zu erforschen, wie sich die Digitalisierung auf die Mobilität der nahen Zukunft in ländlichen und urbanen Räumen auswirkt. Mehr als 30 Projektpartner entwickelten und testeten innovative Mobilitätslösungen in über zehn Teilprojekten. Diese Tests wurden in Hamburg durchgeführt, und die Bürger konnten sich durch Beteiligungsprozesse und Dialogveranstaltungen aktiv engagieren. Die Ergebnisse wurden bei dem ITS World Congress im Oktober 2021 in Hamburg präsentiert. Eines der zehn Teilprojekte von RealLabHH war das Forschungsprojekt Digitales Andreaskreuz (DiAK). Das Ziel von DiAK ist die Verbesserung der Wahrnehmung und dadurch die Vermeidung von gefährlichen Situationen an BÜ. Erreicht wurde dies durch die Entwicklung eines digitalen Andreaskreuzes, das durch den Einsatz der Vehicle-2-everything (V2X)-Technologie sowie über den Mobilfunkstandard 5G regelmäßige Statusmeldungen an die Verkehrsteilnehmenden in der Umgebung sendet. Es macht die Infrastruktur smarter und verbessert die Wahrnehmung der Straßenverkehrsteilnehmenden für die schienenseitige Infrastruktur. Zwei Anwendungsfälle mit zwei Arten von BÜ wurden definiert – der erste Fall ist ein technisch gesicherter BÜ, der mit Schranken gesichert ist, die verhindern, dass Fahrzeuge den BÜ passieren, während ein Zug vorbeifährt. Der zweite Fall, bei dem keine Infrastruktur das Überqueren von Fahrzeugen verhindert, ist ein nicht technisch gesicherter BÜ. Für die zwei Anwendungsfälle wurden Proof-of-Concept-Demonstratoren entwickelt und getestet. Parallel dazu wurden mit der SUMO-Software (Simulation of Urban Mobility [1]), die maßgeblich vom Deutschen Zentrum für Luft- und Raumfahrt e.V. (DLR) entwickelt wurde, Szenarien mit variierenden Parametern wie der Reichweite des digitalen Andre-

The “Digitales Andreaskreuz” project is a sub-project of “Reallabor Hamburg” (“RealLabHH”) that aims to improve mobility and road user comfort by means of the digitalisation of level crossings. A proof-of-concept demonstrator has been developed that provided regular status messages to road users during field tests. This article focusses on the management of the data generated at these level crossings and the development of a simulation middleware to generate data for all the crossings in Hamburg. The data management has been realised by developing a data processing chain consisting of several open-source technologies.

1 Introduction

The “RealLabHH” project aimed to establish a digital mobility reality test lab in Hamburg by the end of 2021 in order to extensively investigate how digitalisation will affect mobility in rural and urban areas in the near future. More than 30 project partners developed and tested innovative mobility solutions in more than ten sub-projects. These tests were carried out in Hamburg and citizens were able to get actively involved through participation and dialogue events. The results were presented at the ITS World Congress in Hamburg in October 2021. “Digitales Andreaskreuz (DiAK)” was one of the ten RealLabHH sub-projects. The goal of DiAK was to improve the perception of level crossings and thereby prevent the occurrence of dangerous situations. This was achieved through the development of digital level crossings that provided regular status updates to road users in their environs using both Vehicle-2-everything (V2X) and 5G technology. It made the infrastructure smarter and better recognisable. Two use cases involving two different types of level crossings were defined. The first type is a technically secured level crossing with barriers preventing vehicles from crossing while a train is passing through. The second type involves a technically unsecured level crossing where no infrastructure prevent vehicles from crossing. Proof-of-concept demonstrators were developed and tested for these two use cases. At the same time, Simulation of Urban Mobility (SUMO) [1] software, significantly developed by the German Aerospace Centre e.V. (DLR), was also used to create scenarios with varying parameters such as the digital level crossings’ broadcasting range and the vehicle density at the level crossings. The simulation promotes an understanding of the influence that these parameters have on the traffic flow in the vicinity of digital level crossings. In addition to the aforementioned activities, other applications and interfaces were also developed

askreuzes und der Verkehrsdichte an den BÜ erstellt. Damit wird das Konzept im gesamten Hamburger Stadtgebiet simuliert. Dadurch wurde die Möglichkeit geschaffen, den Einfluss dieser Parameter auf den Verkehrsfluss in der Umgebung von digitalen Andreaskreuzen zu erkennen und zu verstehen. Neben diesen Aktivitäten wurden Anwendungen und Schnittstellen entwickelt, um die Daten von den digitalen Andreaskreuzen und Verkehrsteilnehmenden vom Feldtest und der Simulation zu verwalten. In diesem Beitrag werden die Komponenten des Backend-Systems sowie die Schnittstellen und Datenformate, die für die Übertragung der Daten aus dem Feld und der Simulation verwendet werden, detailliert erklärt. Weiterhin wird der Datenfluss durch eine Datenverarbeitungskette im Backend sowie die Analyse und Visualisierung der Daten auf einer webbasierten interaktiven Mapping-Anwendung erläutert.

2 Backend-Architektur

Das Backend, das zur Verwaltung der Daten der digitalen Andreaskreuzen genutzt wird, besteht aus drei miteinander verbundenen Komponenten (Bild 1(a)), die vom DLR-Institut für Verkehrssystemtechnik (TS), betrieben und betreut werden. Der erste Bestandteil ist ein Big-Data-Cluster mit verschiedenen Open-Source-Big-Data-Technologien (Bild 1(b)). Es unterstützt Stream- und Batch-Prozessierung, effiziente Speicherung und Analyse von großen, komplexen Daten, einschließlich der raumzeitlichen Daten. Dieses Cluster wurde im Rahmen des Projekts "Rail2X – Smart Services" [2] im Zeitraum von 2018 bis 2019 aufgebaut. Seitdem wird es kontinuierlich weiterentwickelt und für die Aktivitäten im Bereich der Verkehrsforschung und Mobilität verwendet. Das Cluster besteht aus Apache Hadoop [3] und Apache HBase [4], die für die verteilte Datenspeicherung verwendet werden. Die Stream-Prozessierung wird durch die Verwendung von Apache Kafka [5] und Apache Flink [6] ausgeführt. Apache Kafka ist eine Streaming-Plattform zur Speicherung von Nachrichten in verschiedene Topics. Apache Flink ist ein Framework zur Verarbeitung von Datenströmen. Apache Spark [7] ist ein Cluster-Computing-Framework, das in unserem Cluster für die Batch-Prozessierung verwendet wird. Weiterhin nutzen wir GeoMesa [8], dass eine Toolkette zur

in the backend to manage the data collected from the digital level crossings and road users in the field and the simulation. This article explains in detail the backend system components, interfaces and data formats employed for the transfer of data from the field and the simulation. Furthermore, the flow of data through a data processing chain in the backend and the analysis and visualisation of the data on a web-based interactive mapping application are explained.

2 The backend architecture

Our backend for managing the data collected from the digital level crossings comprises three interconnected components (fig. 1(a)), which have been set up and are operated by the DLR Institute of Transportation Systems (TS). The first component is a big data cluster with several open-source big data technologies (fig. 1(b)). It supports stream and batch processing, efficient storage and the analysis of large, complex data, including spatiotemporal data. This cluster was setup during the "Rail2X – Smart Services" project [2] in 2018-19. Since then, it has been continually developed, operated and used for activities in the domain of transportation research and mobility. The cluster consists of Apache Hadoop [3] and Apache HBase [4], which is used for distributed data storage. Stream processing is performed using Apache Kafka [5] and Apache Flink [6]. Apache Kafka is a message-streaming platform where messages are stored under different topics, while Apache Flink is a framework used to process data streams. Apache Spark [7] is a cluster-computing framework that is used in our cluster to carry out batch processing. Furthermore, we also use GeoMesa [8], which offers a set of tools to perform an analysis of the data stored in Apache HBase and Apache Kafka. The cluster architecture has been elaborated in [9].

The second component is a data management system called "Bahnservers", which has been in operation since 2014. It mainly comprises an Oracle database where data collected for various transportation related applications and projects is stored. The system is completed with an OpenVPN [10], which connects the field units and the data management system, and an Apache

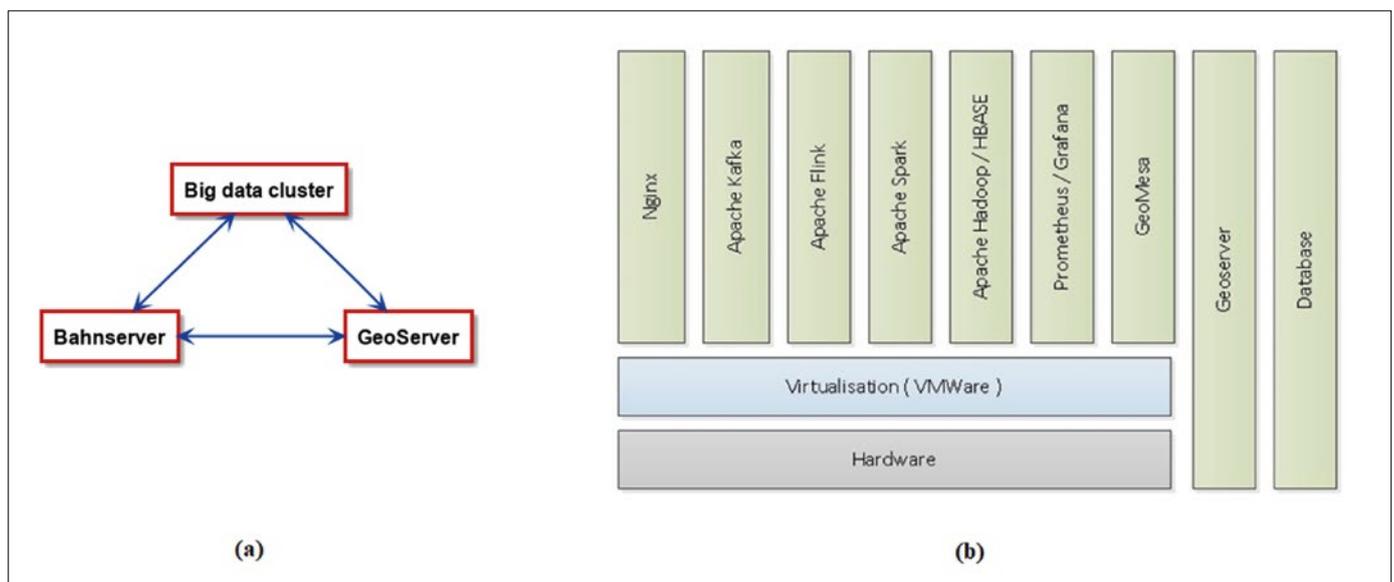


Bild 1: (a) Backend-Komponenten (b) Big-Data-Cluster

Fig. 1: (a) The backend components (b) The big data cluster

Quelle aller Bilder / Source all fig.: DLR

Analyse der in Apache HBase und Apache Kafka gespeicherten Daten bietet. Die Architektur des Clusters wird in [9] erläutert. Die zweite Komponente ist der „Bahnserver“, ein Datenmanagementsystem, das seit 2014 in Betrieb ist und hauptsächlich aus einer Oracle-Datenbank besteht. Die Speicherung der Daten aus vielen verkehrsbezogenen Anwendungen via OpenVPN [10] wird zur Verbindung zwischen Feldeinheiten und dem Datenmanagementsystem genutzt. Außerdem dient ein Apache-Web-Server [11] zur Bereitstellung vieler Schnittstellen und Web-Applikationen. Der Bahnserver wird in [12] erläutert. Außerdem betreibt das Institut zwei GeoServer [13], um Geodaten aus den DLR-TS Datenmanagementsystemen innerhalb des DLR oder zwischen Projektpartnern und mit der Community in den Open-Geospatial-Consortium (OGC)-standardisierten Diensten und Datenformaten [14] zu teilen.

3 Datenmanagement

Bild 2 stellt den im Projekt aufgebauten Datenfluss für das digitale Andreaskreuz und die Verkehrsteilnehmenden im Feld sowie in der Simulation dar. Die Daten werden über die Schnittstelle im Bahnserver empfangen und dann im Big-Data-Cluster gespeichert und analysiert. Verschiedene Aspekte der Datenverwaltung wie z. B. die Erfassung, Validierung, Speicherung, Analyse und Visualisierung der Daten werden in den nächsten Unterabschnitten erläutert.

3.1 Datenerfassung, Validierung und Speicherung

Der Prozess beginnt mit der Übertragung der Daten der zwei digitalen Andreaskreuze im Feld zu einem MQTT-(Message Queuing Telemetry Transport, ein Netzwerkprotokoll) Server [15] mittels 5G. Jedes Andreaskreuz hat Nachrichten mit der Zustandsinformation in regelmäßigen Zeitabständen gesendet. Im Anwendungsfall 2 am nicht technisch gesicherten BÜ wurde zusätzlich eine Prognoseinformation mit der voraussichtlich verbleibenden Zeit bis zur Statusänderung in den Nachrichten integriert. Eine MQTT-Client-Anwendung im Bahnserver hat diese Nachrichten vom MQTT-Server verarbeitet und sie in das OGC-Observation and Measurements [16] Datenformat (JavaScript Object Nota-

web server [11], which makes many interfaces and web applications available. The Bahnserver is explained in detail in [12]. Furthermore, the institute also administers GeoServers [13], which distribute the geospatial information stored in the DLR-TS data management systems within the DLR among the project partners and the community in the standardised Open Geospatial Consortium (OGC) service and data formats [14].

3 Data management

Fig. 2 shows the flow of data starting in the field at the digital level crossings and vehicles, as well as the simulation received at the Bahnserver interfaces and then stored and analysed in the big data cluster. Various aspects of data management such as data collection, validation, storage, analysis and visualisation are explained in the following sections.

3.1 Data collection, validation and storage

The process begins with the transfer of data from two digital level crossings to a converter in the field using V2X, which then sends the data to an MQTT (Message Queuing Telemetry Transport) Server [15] using 5G technology. Each digital level crossing sent messages with information on its open/close status at regular time intervals. In use case 2, which involved technically unsecured level crossings, a forecast for the next change of status was also included in the messages. An MQTT Client application on the Bahnserver processed the messages from the MQTT Server and transformed them into the OGC-Observation and Measurements (OM) [16] format following the JSON (JavaScript Object Notation) implementation standard. The message was then forwarded to the Apache Kafka Cluster, where the data was stored in a Kafka topic. DLR sensor boxes were used to record the positions of the two test vehicles in the field. They sent GNSS-position information to an interface on the Bahnserver, which in turn stored the data in a separate Kafka topic. Middleware was developed for the SUMO simulation in order to test the concept on a larger scale. The middleware developed for the Rail2X project was adapted and further developed for this purpose. Details on the middleware and the SUMO simulation

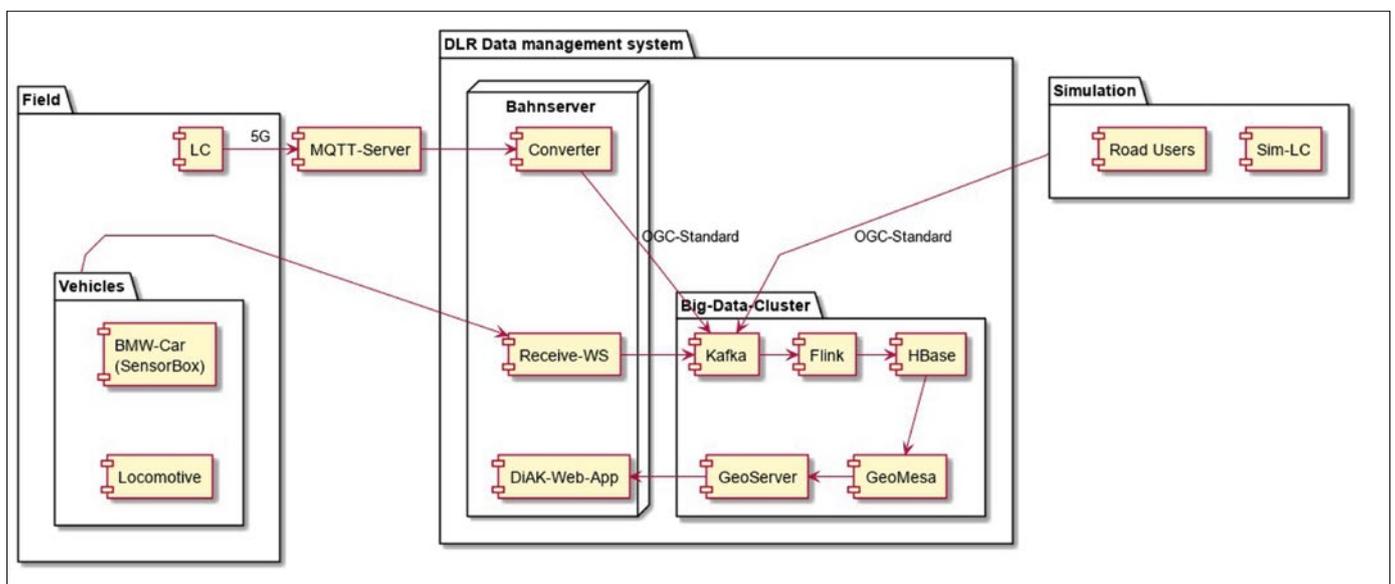


Bild 2: Datenfluss

Fig. 2: The dataflow

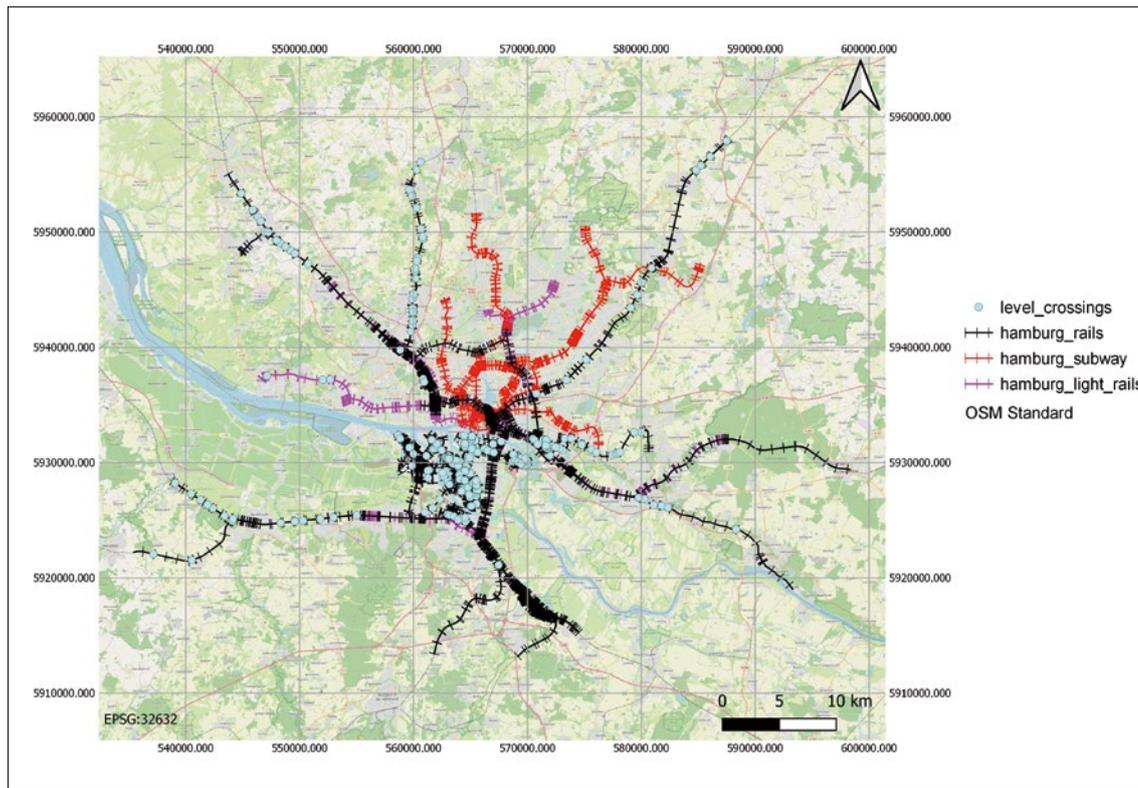


Bild 3: In der Simulation abgebildete Daten
 Fig. 3: The data included in simulation

tion (JSON) – Implementation) umgewandelt. Die Nachrichten wurden dann an das Apache Kafka Cluster weitergeleitet und in einem Kafka-Topic gespeichert. Zur Erfassung der Ortungsdaten von den Testfahrzeugen wurden DLR-Sensorboxen verwendet. Sie haben die GNSS-Ortungsdaten an eine Schnittstelle im Bahnserver gesendet, die dann die Daten in einem anderen Kafka-Topic gespeichert haben.

Um das Konzept für ein größeres Gebiet zu simulieren, wurde eine Middleware für die SUMO-Simulation entwickelt. Dazu wurde die für das Projekt Rail2X entwickelte Middleware genutzt und weiterentwickelt. In [9] wurden bereits die Middleware und die SUMO Simulations-Software erläutert. Zusätzlich wurde aus OpenStreetMap [17] das Straßen-, Bahn-, S-Bahn- und U-Bahn-Netz in Hamburg heruntergeladen und in die Simulation integriert (Bild 3). Datensätze mit BÜ wurden von verschiedenen öffentlich verfügbaren Portalen, z. B. Open-Data-Portal-Hamburg [18], Deutsche Bahn Open-Data-Portal [19], mCloud [20] sowie OpenStreetMap heruntergeladen und zu einem Datensatz zusammengefasst. BÜ mit unbekanntem Typ wurden ausgelassen. 422 BÜ in Hamburg wurden in der Simulation abgebildet. Deren Eigenschaften sind in Tab. 1 dargestellt.

Der Fahrplan für öffentliche Verkehrsmittel für das Jahr 2021 wurde im GTFS-Datenformat [21] vom Open-Data-Portal-Hamburg ebenfalls in die Simulation integriert, um den Bahnverkehr zu reproduzieren. Dieser Datensatz enthielt jedoch nicht den Fahr-

plan kann in [9] gefunden werden. Die Straßen-, Eisenbahn-, Lichtschienen- und U-Bahn-Netzwerke in Hamburg wurden von OpenStreetMap [17] heruntergeladen und in die Simulation (Fig. 3) integriert. Level crossing Datensätze von verschiedenen öffentlichen Portalen wie Open Data Portal Hamburg [18], Deutsche Bahn Open Data Portal [19], mCloud [20] und OpenStreetMap wurden heruntergeladen und zu einem Datensatz zusammengefasst. BÜ mit unbekanntem Typ wurden ausgelassen. 422 BÜ in Hamburg wurden in der Simulation abgebildet, deren Eigenschaften in Tab. 1 dargestellt sind.

Der öffentliche Verkehrsschichtplan für das Jahr 2021 im GTFS-Format [21] wurde ebenfalls in die Simulation integriert, um den Bahnverkehr zu reproduzieren. Dieser Datensatz enthielt jedoch nicht den Fahr-

Network Type / Netzwerktyp	Intersecting level crossings / Kreuzende Bahnübergänge		
	Technically secured / Technisch gesichert	Technically unsecured / Nicht technisch gesichert	Total / Summe
Light rail / S-Bahn	4	1	5
Subway / U-Bahn	0	0	0
Railway / Bahn	130	287	417
Total / Summe	134	288	422

Tab. 1: Eigenschaften der in der Simulation abgebildeten Bahnübergänge
 Tab. 1: Details of the level crossings included in the simulation

plan für alle Züge in Hamburg. Weil diese Information im GTFS-Format aus einer anderen Quelle nicht verfügbar war, wurde sie ausgelassen. Ähnlich wie im Feld erhielten alle Verkehrsteilnehmenden in der Simulation, wie z. B. Pkw, Lkw und Rettungsfahrzeuge, die Statusmeldungen von den digitalen Andreaskreuzen in ihrer Umgebung. In der Simulation bekamen sie zusätzlich die Information über die Verkehrsdichte. Es ist möglich, die Reichweite der digitalen Andreaskreuze sowie die Verkehrsdichte in der Simulation zu ändern. Die Zustandsdaten der BÜ und Ortungsdaten der Verkehrsteilnehmenden aus der Simulation im OGC-OM-Format [16] wurden direkt an die Topics im Apache Kafka gesendet.

Die Daten aus zwei Apache-Kafka-Topics wurden mittels zwei Apache-Flink-Anwendungen verarbeitet. Eine Flink-Anwendung hat die Nachrichten von einem Kafka-Topic gelesen und die Information über die BÜ-ID, den Zeitstempel und andere Eigenschaften des BÜ extrahiert. Anschließend wurden diese Werte validiert. Falls die Werte null oder unrealistisch waren, wurde eine E-Mail mit einer Fehlermeldung an den Systemverantwortlichen gesendet. Dies hilft den Nutzern, so schnell wie möglich über defekte Infrastruktur im Feld informiert zu werden. Wenn die Daten fehlerfrei waren, wurden sie in einer HBase-Tabelle für weitere Analysen und Visualisierungen gespeichert. Eine zweite Flink-Anwendung hat die Nachrichten von einem anderen Kafka-Topic verarbeitet, um die Ortungsdaten von Verkehrsteilnehmenden zu validieren und in einer HBase-Tabelle zu speichern. Beide Flink-Anwendungen liefen rund um die Uhr und belegten Task-Slots, die der Anzahl der Partitionen des Kafka-Topics entsprechen, sodass jeder Slot mit einer Partition des Kafka-Topics verbunden war, wodurch eine optimale Verarbeitungsleistung erreicht wurde.

In einem Szenario der Simulation wurden alle Verkehrsteilnehmenden darauf programmiert, eine alternative Route zu finden, wenn die Verkehrsdichte an einem sich nähernden BÜ hoch und der BÜ geschlossen war. Die Simulation hat mit wenigen Fahrzeugen gut funktioniert. Allerdings sind die Berechnungen in größeren Netzwerken mit tausenden Verkehrsteilnehmenden sehr zeitaufwendig. Dadurch wird die Live-Darstellung verzögert laufen. Weiterhin empfing die Middleware alle Daten aus dem Feld, solange Feldtests stattfanden. Wenn ein digitales Andreaskreuz ein Statusupdate geschickt hat, war es bei den simulierten BÜ sichtbar, sodass die Verkehrsteilnehmenden in der Simulation sich nach der angepassten Situation im Feld verhielten. Die Leistung einer derartigen Echtzeitsimulation wurde im Projekt erfolgreich getestet, und es wurde festgestellt, dass sie mit einigen hundert Verkehrsteilnehmenden effizient läuft.

3.2 Datenanalyse und Visualisierung

Im Projekt Digitales Andreaskreuz wurde eine webbasierte Kartenanwendung „RealLabHH DiAK Graphical User Interface (GUI)“ auf dem Bahnserver mittels Openlayer [22] und CesiumJS [23] JavaScript-Bibliotheken entwickelt, um die Daten aus dem Feld in der Simulation zu visualisieren. Diese Anwendung hilft bei der Durchführung initialer Analysen sowie der Visualisierung von historischen und Echtzeit-Datensätzen (Bild 4). Durch den Einsatz des GeoMesa-HBase-GeoServer Plugin [24] wurden HBase-Datenspeicher in dem GeoServer erstellt. Die in HBase gespeicherten Daten wurden durch OGC-standardisierte Dienste, wie z. B. Web Mapping Service (WMS) und Web Feature Service (WFS) bereitgestellt.

Im „Live Mode“ wird in der Anwendung die Position der digitalen Andreaskreuze, deren Reichweite und deren aktueller

The data present in the two Kafka topics was then processed using two Apache Flink applications. One Flink application read the messages from one Kafka topic and extracted the information on the level crossing ID, the data timestamp and any other status related properties of the level crossings. It subsequently validated these values. If any of the values were null or unrealistic, an email with an error message was automatically sent to the persons responsible. This step helps operators be aware of any defective infrastructure in the field as quickly as possible. If there were no validation errors, the data was stored in an HBase table for further analysis and visualisation. A second Flink application processed the data from another Kafka topic in order to validate and store the vehicle positions in another HBase table. These Flink jobs kept running 24/7 and took up task slots equal to the number of partitions in the Kafka topic, so that each slot connected to one partition in the Kafka topic, thereby achieving optimal processing performance.

In one simulation scenario, all the road users were programmed to find an alternate route when the traffic density at an approaching level crossing was high and the crossing was closed. This was observed to work well with fewer road users, while the computations were time consuming in larger networks with several thousands of road users and live visualization was therefore delayed. Furthermore, the simulation middleware continuously received the live data from the field whenever it was available. Whenever the level crossings in the field sent an update, it was reflected in the simulation's level crossings, so that the simulated road users behaved according to the situation in the field. The performance of such a real-time simulation was successfully tested in this project and it was observed to run efficiently with a few hundred vehicles.

3.2 Data analysis and visualisation

The web-based map application “RealLabHH DiAK Graphical User Interface (GUI)” was developed on the Bahnserver using Openlayers [22] and CesiumJS [23] JavaScript libraries in order to enable the visualisation of the data collected in the field. This application assists in the performance of preliminary data analysis, as well as in viewing live and historic data [fig. 4]. HBase datastores were created on the GeoServer using the GeoMesa-HBase-GeoServer plugin [24]. The data stored in the HBase tables was published using OGC-standardised services, such as the Web Mapping Service (WMS) and Web Feature Service (WFS).

The position of the digital level crossings, their broadcasting range and their current status (fig. 4a) were displayed on a map in the application's “Live mode”. This helps the user obtain an overview of the current status of the infrastructure in the field. However, it should be noted that the range is not a perfect sphere in practical terms, but it is affected by objects in the environment. If the Flink application found a digital level crossing to be broadcasting erroneous data or inactive, it was marked with differently on the map in order to provide the operators with another warning in addition to the email alerts. The visualisation of the broadcasting range enables the operators to find any areas that are not covered by the existing infrastructure. This can assist in identifying zones along the railway-road network, where the placement of additional infrastructure, such as hopping stations, could possibly help in enhancing the V2X coverage.

Status auf einer Karte dargestellt (Bild 4a). Es unterstützt die Nutzer, sich einen Überblick über den aktuellen Stand der Infrastruktur im Feld zu verschaffen. Es ist jedoch zu beachten, dass die Reichweite praktisch keine perfekte Sphäre ist, sondern durch Objekte in der Umgebung beeinflusst wird. Wenn ein digitales Andreaskreuz von der Flink-Anwendung als fehlerhaft oder inaktiv erkannt wurde, wird es auf der Karte markiert, um den Nutzer ergänzend zur Fehlermeldung per E-Mail darüber zu informieren. Durch die Visualisierung der Reichweite können Gebiete identifiziert werden, die von vorhandener Infrastruktur nicht abgedeckt werden. Dies kann dazu beitragen, Bereiche entlang des Eisenbahn- und Straßennetzes zu ermitteln, in denen die Errichtung zusätzlicher Infrastrukturen wie bspw. Hopping-Stationen dazu beitragen könnte, die V2X-Abdeckung zu verbessern.

Weiterhin wurden zusätzliche Verkehrsteilnehmende im Umfeld von BÜ mit digitalem Andreaskreuz in der Simulation ergänzt, um die Auswirkung einer hohen Verkehrsdichte und deren Nutzen von verkehrsdichtebasierten Alerts und Rerouting-Empfehlungen zu evaluieren. Dies könnte besonders für Rettungsfahrzeuge hilfreich sein, da sie durch ein Rerouting ihr Ziel so schneller erreichen könnten. Es ist jedoch nicht sinnvoll, alle Verkehrsteilnehmenden im Feld zu tracken. Deswegen kann eine Schätzung der Verkehrsdichte an BÜ mit anderen Methoden erfolgen wie z. B. durch Kameraaufnahme von Bildern an BÜ und Durchführung einer Objekterkennung. Mit diesen Daten könnte eine Zeitreihenanalyse der Verkehrsdichte mithilfe der GeoMesa-Analysetools ausgeführt werden.

Die in einem bestimmten Zeitfenster während der Feldtests gesammelten Daten können im „History Mode“ in der Anwendung wiedergegeben werden (Bild 4b). Die Wiedergabe der Daten wird über einen Time-Slider gesteuert. Über ein Diagramm wird die Anzahl der Fahrzeuge während des gewählten Zeitfensters angezeigt, sobald der Anwender in der Simulation auf einen BÜ klickt. Diese Anzeigen können den Nutzer unterstützen, die Verkehrsverteilung an verschiedenen BÜ im Laufe des Tages zu beobachten und möglicherweise einen Trend aufzudecken. Zuletzt wurde eine Spark-Anwendung entwickelt, die Berichte für bestimmte Zeitfenster generiert. Zusätzlich zur Kartenansicht enthält der Bericht grundlegende statistische Informationen zu den gesammelten Daten der BÜ.

4 Fazit

Der Beitrag fokussiert sich auf den Einsatz von Open-Source-Software, um die Daten aus digitalen Andreaskreuzen an zwei Arten von BÜ – technisch gesichert und nicht technisch gesichert – darzustellen. Zur Erfassung der Daten wurden Open-Source-Standards wie die des Open Geospatial Consortium (OGC) verwendet. Die Konzepte wurden durch den Einsatz der SUMO Simulations-Software für einen größeren Bereich simuliert. In der Datenverarbeitungskette wurden bewährte und weitverbreitete Technologien wie die der Apache Software Foundation verwendet. Verschiedene Aspekte der Systemleistung wie die Verarbeitungsgeschwindigkeit und die Latenz in der Datenverarbeitung wurden für die zwei genannten Anwendungsfälle evaluiert. Die Systemarchitektur besteht vollständig aus Open-Source-Technologien und ist daher kostengünstig zu reproduzieren. Das hier dargestellte Vorhaben ist ein Proof-of-Concept für derartige Systeme, die in Zukunft in großem Maßstab eingeführt werden könnten. ■

Furthermore, additional road users could also be simulated around the digital level crossings in order to study the influence of traffic density and the usefulness of density-based alerts and rerouting recommendations. This could be particularly beneficial to emergency vehicles, where rerouting could help them reach their destination as quickly as possible. In reality, however, it is not practical to track each and every road user in the field. An estimation of the traffic density at level crossings can therefore be obtained using other approaches, such as capturing images at a level crossing using a camera and performing object detection. This data can be used to perform a time-series traffic density analysis using the analysis tools offered by GeoMesa. The data collected in a specific time window can be replayed using the application’s history mode (fig. 4b). The data playback is controlled using a time slider. A graph of the vehicle count during a chosen time window is displayed by clicking on a level crossing. This supports the user’s understanding of the traffic distribution at various level crossings at different times of the day and can reveal trends. Lastly, a spark application was developed to generate reports for specific time windows. The report provides basic statistical information on the data collected from the level crossings in addition to the map view.

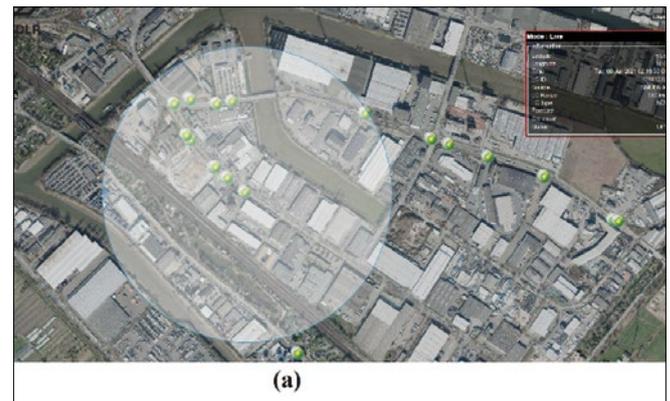


Bild 4a: RealLabHH DiAK GUI: Echtzeitdaten
 Fig. 4a: The RealLabHH DiAK GUI: Live data Quelle / Source Base map: Freie und Hansestadt Hamburg, Landesbetrieb Geoinformation und Vermessung

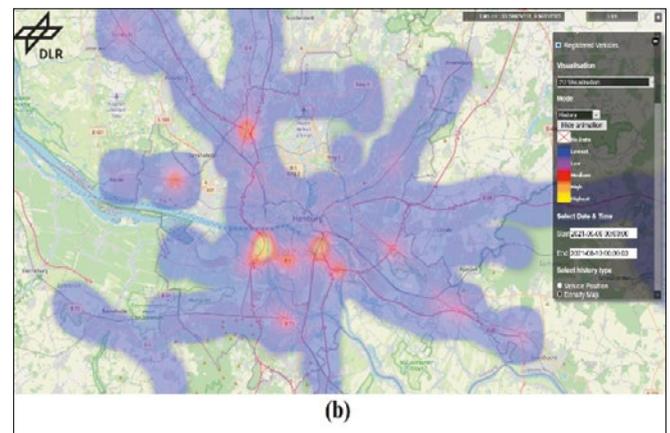


Bild 4b: RealLabHH DiAK GUI: Historische Daten für die Fahrzeugdichte
 Fig. 4b: The RealLabHH DiAK GUI: Historical vehicle density data Quelle / Source Base map: OpenStreetMap

LITERATUR | LITERATURE

- [1] Wießner, E.; Erdmann, J.; Flamm, L.; Jäger, B.: Auf Straße und Schiene mobil – intermodale Verkehrssimulation mit SUMO, SIGNAL+DRAHT (110) 11/2018, pp. 49-55. DVV Media Group. ISBN ISSN 0037-4997. <https://elib.dlr.de/124928/>
- [2] Grünhäuser, M.; Hauck, R.; Meirich, C.; Schubert, L. A.: Rail2X: Der Einsatz von V2X-Technologie im Eisenbahnbereich, SIGNAL+DRAHT (111) 7+8/2019, pp. 53-59. DVV Media Group. ISSN 0037-4997. <https://elib.dlr.de/127084/>
- [3] <https://hadoop.apache.org/>, Zugriff: 20. September 2021 um 10:00
- [4] <https://hbase.apache.org/>, Zugriff: 20. September 2021 um 10:00
- [5] <https://kafka.apache.org/>, Zugriff: 20. September 2021 um 10:00
- [6] <https://flink.apache.org/>, Zugriff: 20. September 2021 um 10:00
- [7] <https://spark.apache.org/>, Zugriff: 20. September 2021 um 10:00
- [8] <https://www.geomesa.org/>, Zugriff: 20. September 2021 um 10:00
- [9] Shankar, S.; Schubert, L. A.; Patil, A. J.; Erdmann, J.: Der Einsatz von Big Data und der Verkehrssimulation SUMO im Rahmen von Rail2X, SIGNAL+DRAHT (112) 10/2020, pp. 49-58. DVV Media Group. ISSN 0037-4997. <https://elib.dlr.de/134134/>
- [10] <https://openvpn.net/>, letzter Zugriff 20. September 2021 um 10:00
- [11] <https://httpd.apache.org/>, letzter Zugriff 20. September 2021 um 10:00
- [12] Schubert, L. A.; Rahmig, C.; Scholz, M.; Böhm, T.: Zentralisiertes Management von Geodaten im Schienenverkehr, SIGNAL+DRAHT (108) 12/2016, pp. 6-14. DVV Media Group. ISSN 0037-4997. <https://elib.dlr.de/106495/>
- [13] <http://geoserver.org/>, letzter Zugriff 20. September 2021 um 10:00
- [14] <https://www.ogc.org/standards/>, letzter Zugriff 20. September 2021 um 10:00
- [15] <https://www.ogc.org/standards/om>, letzter Zugriff 20. September 2021 um 10:00
- [16] <https://openlayers.org/>, letzter Zugriff 20. September 2021 um 10:00
- [17] <https://cesium.com/cesiumjs/>, letzter Zugriff 20. September 2021 um 10:00
- [18] <https://www.geomesa.org/documentation/stable/user/hbase/geoserver.html>, letzter Zugriff 20. September 2021 um 10:00
- [19] <https://data.deutschebahn.com/>, last accessed on 20 September 2021 at 10:00
- [20] <https://mcloud.de/>, last accessed on 20 September 2021 at 10:00
- [21] <https://gtfs.org/>, last accessed on 20 September 2021 at 10:00
- [22] <https://openlayers.org/>, last accessed on 20 September 2021 at 10:00
- [23] <https://cesium.com/cesiumjs/>, last accessed on 20 September 2021 at 10:00
- [24] <https://www.geomesa.org/documentation/stable/user/hbase/geoserver.html>, last accessed on 20 September 2021 at 10:00

4 Conclusion

This article focuses on the use of open-source software to manage the information received from two types of digital level crossings, namely technically secured and technically unsecured level crossings. Open-source standards such as those from the Open Geospatial Consortium (OGC) have been used. The concepts have been tested on a large scale using SUMO simulation software. Well-tested and widely used technologies, such as those from the Apache Software Foundation, were used in the data processing chain. Various aspects of the system's performance, such as the processing speed and the data processing latency, were evaluated for the two use cases. The system architecture adopted in this project consists completely of open-source technologies and hence, it is economical to reproduce. It stands as a proof-of-concept for such systems that could be established on a large scale in the future. ■

AUTOREN | AUTHORS

Sangeetha Shankar, M. Sc.

Wissenschaftliche Mitarbeiterin / *Research Associate*
Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR)
Institut für Verkehrssystemtechnik / *Institute of Transportation Systems*
Anschrift / *Address*: Lilienthalplatz 7, D-38108 Braunschweig
E-Mail: sangeetha.shankar@dlr.de

Dipl.-Wirtsch.-Ing. Miriam Grünhäuser

Wissenschaftliche Mitarbeiterin / *Research Associate*
Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR)
Institut für Verkehrssystemtechnik / *Institute of Transportation Systems*
Anschrift / *Address*: Lilienthalplatz 7, D-38108 Braunschweig
E-Mail: miriam.gruenhaeuser@dlr.de

**RAILWAY
DIAGNOSTIC AND
MONITORING
CONFERENCE
2022**

6th + 7th APRIL 2022
EMPIRE RIVERSIDE HOTEL, HAMBURG

KONTAKTE VOM FEINSTEN!
Nutzen Sie die **Railway Diagnostic and Monitoring Conference** wieder als Plattform und werden Sie Sponsor!
Treten Sie mit den Experten direkt in Kontakt!
Sichern Sie sich das attraktive Präsentationspaket –
Deadline ist am **21. März 2022**.
Sprechen Sie mich an!

Kontakt: Silvia SanderE-Mail: silvia.sander@dvvmedia.com

Telefon: +49/40/237 17 – 171

**Eurail
press**