





Vorgängen wird dann die sogenannte „Engpasschwere“ als dimensionslose Kennzahl berechnet. Die Engpasschwere setzt sich zusammen aus Kategoriewechseln sowie Verspätungszuwächsen. Es werden die Anzahl der Wechsel in schlechtere Kategorien (ab Wechsel in Kategorie 2, also Abweichung von mehr als 30 s von der Planlage, inkl. der Anzahl der „übersprungenen“ Kategorien) und die Anzahl der „Verspätungsereignisse“ ermittelt. Verspätungsereignisse sind Verspätungszuwächse größer eines definierten Schwellwertes (z. B. 90 s). Die Engpasschwere ergibt sich aus der gewichteten Summe der Anzahl der Kategoriewechsel und Verspätungsereignisse.

Durch die Auswertung von Verspätungszuwächsen anstelle absoluter Verspätungen werden die Stellen im Netz identifiziert, an denen Verspätung entsteht. Gegenüber der Betrachtung von absoluten Verspätungswerten bietet die Auswertung der Kategorien verschiedene Vorteile:

- Der Fokus der Identifikation von Engpässen liegt auf kritischen, systematisch auftretenden Verspätungsentwicklungen. Durch die Kategorisierung sind Einzelergebnisse mit großem Verspätungszuwachs nicht ausschlaggebend.
- Schleichend aufgebaute Verspätungen werden durch die Kategoriewechsel erfasst.
- Die Definition der Kategoriengrenzen (Breite der Kategorien verdoppelt sich mit steigender Verspätung) gewichtet die Abweichung von der Planlage höher als den Verspätungsaufbau bereits (stark) verspäteter Zugfahrten.

Die zusätzliche Auswertung von Verspätungsereignissen ist sinnvoll, um Verspätungszuwächse innerhalb der Kategorien zu berücksichtigen. Dabei werden Verspätungsereignisse in Kategorie 0 (Vorplanfahrt) nicht gezählt, da diese zur Herstellung der Planlage dienen, jedoch keine Verspätung verursachen. Verspätungsereignisse innerhalb der Verspätungskategorie 5 werden ebenfalls nicht in der Engpasschwere berücksichtigt, da die Abweichung von der Planlage hier so groß ist, dass nicht mehr auf systematische Probleme, sondern auf Einzelergebnisse zu schließen ist.

Die Engpasschwere je Zugnummer wird anhand der jeweiligen Verkehrsart gewichtet, wobei die Gewichtungsfaktoren auf den in der Kapazitätsermittlung von Eisenbahninfrastrukturen gebräuchlichen Rangziffern [3] basieren. Die so berech-

nete zugscharfe Engpasschwere bildet die Grundlage, um Engpässe auf verschiedenen Aggregationsebenen ermitteln zu können. So können Engpasschweren sowohl für Bahnhöfe als auch für Streckenabschnitte ausgegeben werden. Zudem ist eine Auswertung getrennt nach Zugklassen oder Linien bis zur Auswertung einzelner Zugnummern möglich.

Im bisherigen Projektverlauf wurden für die Identifikation von Engpässen zwei Aggregationsebenen ausgewertet:

- Bahnhöfe: Die Engpasschwere der Vorgänge aller Zugfahrten innerhalb eines Bahnhofes wird für alle Strecken und Richtungen zusammengefasst
- Streckenabschnitte: Die Engpasschwere der Vorgänge aller Zugfahrten zwischen zwei Betriebsstellen auf einer Strecke wird für beide Richtungen zusammengefasst

Die so identifizierten Infrastrukturabschnitte mit der größten Engpasschwere werden als Engpässe anschließend detailliert analysiert.

### Analyse

Zur Analyse eines identifizierten Engpasses werden relevante Informationen aus verschiedenen Datenquellen zusammengestellt. Darüber hinaus werden Mechanismen der Verspätungsübertragung zwischen konkreten Zugfahrten mittels Methoden der Datenanalyse („Episode Mining“) ermittelt. Die Ergebnisse der Datenaufbereitung und -analyse werden übersichtlich in einem Steckbrief zusammengefasst.

### Datenaufbereitung in Form von Steckbriefen

Ursächlich für Engpässe können infrastrukturelle sowie betriebliche Gründe sein. Um einen umfassenden Überblick über die Rahmenbedingungen des identifizierten Engpasses zu geben, werden verschiedene Datenquellen je Engpass in Form von Steckbriefen (vgl. Bild 2) aufbereitet und gegenübergestellt. Die Daten werden im Folgenden kurz beschrieben:

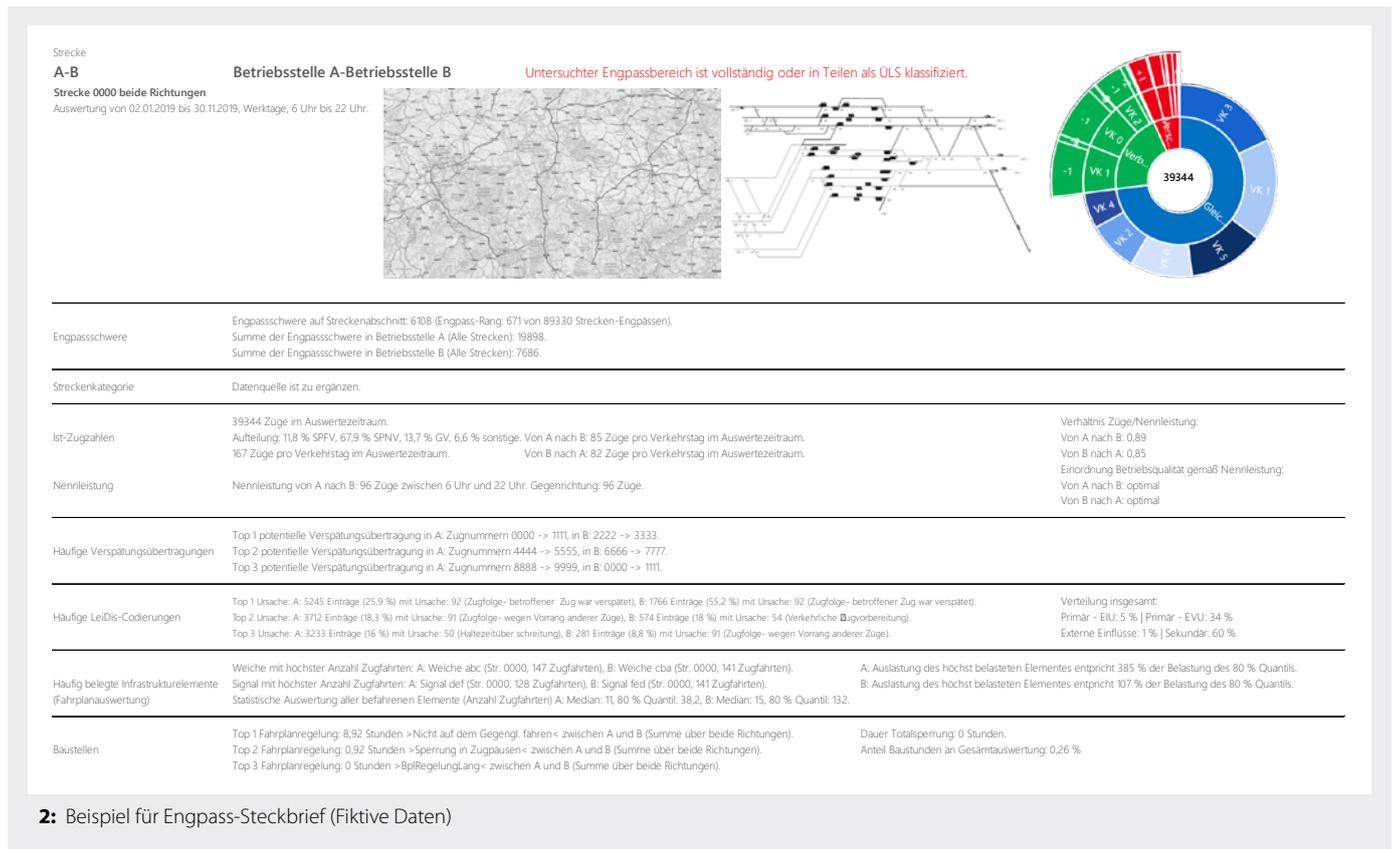
- Die Ist-Zugzahlen werden erfasst und nach Verkehrsarten aufgeschlüsselt angegeben. Zudem wird (bei Streckenabschnitten) die Nennleistung des Abschnittes ausgegeben. Aus der

Kombination der beiden Angaben kann direkt entnommen werden, in welcher Qualität der untersuchte Engpass gemäß Leistungsfähigkeitsberechnung betrieben wird.

- Die Zugnummern, zwischen denen gemäß der unten beschriebenen Datenanalyse die signifikantesten Verspätungsübertragungen ermittelt wurden, werden im Steckbrief aufgenommen.
- Treten im Betrieb Verspätungszuwächse von mehr als 90 s auf, wird durch das Betriebspersonal eine Verspätungsursache codiert. Die häufigsten Verspätungsursachen sowie die Gesamtverteilung der Verspätungen in primäre (unterteilt in EIU und EVU), sekundäre und weitere Ursachen werden ausgegeben.
- Aus mikroskopischen Infrastruktur- und Fahrplandaten können mittels der Software LUKS® die Belastungen (Anzahl Überfahrten, Belegungszeit, maximal auftretende Geschwindigkeit, ...) der einzelnen Infrastrukturelemente berechnet werden. Die am stärksten belasteten Elemente werden ausgegeben, außerdem das Verhältnis zwischen der Anzahl der Belegungen des am häufigsten belegten Elementes und eines repräsentativen Elementes (z. B. 80% Quantil). Das Verhältnis gibt einen Anhaltspunkt darüber, ob die Infrastruktur einer Betriebsstelle verteilt genutzt wird, oder ob einzelne Elemente deutlich erhöhte Belegungszahlen aufweisen. Letzteres kann auf einzelne Elemente mit Engpasswirkung hindeuten.
- Informationen zum Baugeschehen werden dargestellt. Ausgegeben werden zum einen die Zeiträume mit Totalsperungen, zum anderen die Dauer der längsten Bauplanregelungen (z. B. „Gleiswechselbetrieb“).

### Ermittlung von Verspätungsübertragungen

Zur Analyse von Engpässen sind die im Betrieb auftretenden Mechanismen der Verspätungsübertragungen zwischen konkret auftretenden Zugfahrten elementar. Auch wenn geplante Zugfolgefälle und Mindestzugfolgezeiten aus dem Fahrplan entnommen werden können, ist die Analyse realer Betriebsdaten notwendig, um systematische Abweichungen vom Fahrplan zu detektieren. In den Betriebsdaten zeigen sich unter anderem Ausschlüsse aufgrund in der Belegungsrechnung nicht sichtbarer Abhängigkeiten, Anschlüsse und ähn-



lichem. Hier setzt die im Projekt auf Basis von [6] weiterentwickelte Methodik des „Episode Mining“ an, welche Betriebsdaten auswertet. Vereinfacht gesagt wird ein fixes Zeitfenster („Sliding Window“) schrittweise mit einer definierten Schrittweite (z.B. 1s) über eine Liste der Zeitstempel aller verspäteten Zugfahrten eines Messpunktes verschoben. Die Anzahl der Schritte, in denen die Zeitstempel zweier Zugfahrten gemeinsam innerhalb des Sliding Windows liegen, wird über alle Verkehrstage summiert. Dadurch werden Zugfahrten detektiert, die oftmals gemeinsam verspätet verkehren bzw. verspätet werden. Diese sogenannten „Episoden“ können sowohl zwischen zwei als auch zwischen mehreren Zugfahrten identifiziert werden. So werden systematische Reihenfolgen („Zug A verkehrt immer vor Zug B“) ermittelt, bei denen die nachfolgenden Züge Verspätung erleiden. Ziel dabei ist neben der Identifikation von Zugfolgekonflikten die Analyse der Verspätungsübertragungen beispielsweise von kreuzenden Zugfahrten in Bahnhofsköpfen. Durch die Auswertung von Betriebs- anstelle von Fahrplandaten sollen sowohl nicht geplante, aber im realen Zugbetrieb oftmals auftretende Konstellationen der Verspätungsübertragung als

auch relevante Konstellationen zwischen mehr als zwei Zügen identifiziert werden. Die Ergebnisse des Episode Mining werden unter Berücksichtigung der weiteren Informationen des Steckbriefs ausgewertet und geben einen vertieften Einblick in den Engpass. Zum aktuellen Zeitpunkt erfolgt die Validierung vor allem auf Basis von Zugnummern. Die Methodik kann darüber hinaus Linien auswerten, um wiederkehrende Konstellationen in einem vertakteten Fahrplan zu finden.

**Einblick ins Projekt**

Auf dem Weg zur Erreichung der zuvor beschriebenen Ergebnisse sind verschiedene Besonderheiten im Umgang mit den Eingangsdaten und der entwickelten Methodik berücksichtigt worden:

Verspätungen entstehen nicht unbedingt am Engpass selbst, sondern insbesondere im Warteraum vor Engpässen. Bei der Analyse sind also stets die Vorgänger- sowie Nachfolger-Betriebsstellen mit zu berücksichtigen.

Die große Datenmenge (alle Zugfahrten auf dem Netz der DB Netz AG eines Jahres) erfordert effektive Algorithmen der Datenaufbereitung und -analyse. Teils

sind Vereinfachungen notwendig, um die Datenmenge verarbeiten zu können. Zudem liegen Daten verschiedener Quellen in unterschiedlichen Aggregierungsebenen vor. Zur validen Auswertung muss die Konsistenz der auszuwertenden Daten gewährleistet sein. Dabei sind Fehler in den Eingangsdaten als solche zu erkennen und von der Analyse auszuschließen.

Die Methode des Episode Mining wurde an die Besonderheiten des Eisenbahnbetriebes angepasst. Zugfolgen werden fahrplanunabhängig analysiert. Neben den o.g. Vorteilen besteht dabei die Gefahr, tatsächlich unabhängige Vorgänge fälschlicherweise als Kausalzusammenhänge zu identifizieren. Um dies zu vermeiden, wird u.a. neben der absoluten Verspätung zusätzlich der Verspätungszuwachs von Zugfahrten berücksichtigt. D.h. Episoden werden nur dann als solche identifiziert, wenn die zweite/dritte/... Zugfahrt einer Episode einen Verspätungsaufbau erleidet (andernfalls verkehren beide Zugfahrten zwar verspätet, es ist jedoch davon auszugehen, dass keine Verspätung übertragen wird). Weitere Informationen zur Identifikation von Zugfolgen mittels der beschriebenen Methodik werden voraussichtlich im zweiten Halbjahr 2022 veröffentlicht.

## Fazit und Ausblick

Die beschriebene Methode zur Engpassidentifikation bietet einen Ansatz zur Detektion problematischer Netzbereiche basierend auf Betriebsdaten. Sie kann dadurch anders als bestehende analytische Verfahren, die auf der Auswertung von Fahrplandaten basieren, tatsächliche Abweichungen im Betrieb erkennen. Im Betrieb treten Verspätungen nicht nur aufgrund von Belegungskonflikten, sondern auch durch weitere Abhängigkeiten, wie z.B. das Warten auf Anschlüsse auf. Es werden somit systematisch auftretende Abhängigkeiten und Verspätungsquellen identifiziert.

Zur Analyse der identifizierten Engpässe werden relevante Informationen aus verschiedenen Datenquellen in Steckbriefen übersichtlich zusammengestellt. Dabei werden unter anderem mittels Methoden der Datenanalyse („Episode Mining“) erkannte, häufig auftretende Verspätungsübertragungen zwischen einzelnen Zugfahrten angegeben.

Basierend auf der automatisierten Analyse der identifizierten Engpässe werden in den nächsten Schritten Lösungsmöglichkeiten entwickelt. Darüber hinaus sollen weitergehende Möglichkeiten der manuellen Analyse geschaffen werden, um die Wirkmechanismen und Ursachen der Engpässe aufzuzeigen und diese auflösen zu können.

## Summary

### Insight into DZSF project „Identification of capacity bottlenecks“

Capacity bottlenecks significantly determine capacity and operation quality of the entire railway network. Within the framework of a project initiated by DZSF, a method is being developed and implemented as prototype which identifies bottlenecks in railway networks and provides proposals for solutions. Hereinafter, there is an insight into the current work status.

## Literatur

- [1] Bundesministerium für Verkehr und digitale Infrastruktur, Masterplan Schienenverkehr, Juni 2020.
- [2] SPD, BÜNDNIS 90 / DIE GRÜNEN, FDP, „Mehr Fortschritt wagen - Bündnis für Freiheit, Gerechtigkeit und Nachhaltigkeit - Koalitionsvertrag 2021 – 2025 zwischen der Sozialdemokratischen Partei Deutschlands (SPD), BÜNDNIS 90 / DIE GRÜNEN und den Freien Demokraten (FDP)“, 2021.
- [3] DB Netz AG, Richtlinie Fahrwegkapazität (Ril 405), 2008.
- [4] A. Pfeifer, B. Kogel, N. Nießen, S. Kurby und U. Steinborn, „Zielgerichtete Infrastrukturdimensionierung durch eine netzweite Ermittlung der Kapazität,“ Eisenbahntechnische Rundschau (ETR), pp. 29-33, Oktober 2018.
- [5] Deutsches Zentrum für Schienenverkehrsforschung beim Eisenbahn-Bundesamt, „Methodik der Kapazitätsbewertung des Gesamtsystems und Knotenberechnung,“ 21 Juni 2021. [Online]. Available: [https://www.dzsf.bund.de/SharedDocs/Standardartikel/DZSF/Projekte/Projekt\\_41\\_Kapazitaetsbewertung\\_Knotenberechnung.html](https://www.dzsf.bund.de/SharedDocs/Standardartikel/DZSF/Projekte/Projekt_41_Kapazitaetsbewertung_Knotenberechnung.html). [Zugriff am 6 Februar 2022].
- [6] B. Cule, B. Goethals, S. Tassenoy und S. Verboven, „Mining Train Delays“ [https://doi.org/10.1007/978-3-642-24800-9\\_13](https://doi.org/10.1007/978-3-642-24800-9_13), in Advances in Intelligent Data Analysis X, Springer, 2011, pp. 113-124.

# WISSEN, WAS BAHNEN BEWEGT



[www.eurailpress.de](http://www.eurailpress.de)