

PTNA: Qualitätssicherung für ÖPNV-Linien in OpenStreetMap

Toni Erdmann

Abstract

„PTNA - Public Transport Network Analysis“ [1] ermöglicht eine SOLL-IST-Analyse für ÖPNV-Linien in OpenStreetMap. Soll-Daten bestehen aus einer CSV-Liste von ÖPNV-Linien, jeweils mindestens aus der Liniennummer und dem Fahrzeugtyp. Diese Informationen werden in einer IST-Analyse mit den OSM-Daten verglichen. Zusätzlich erfolgt eine Fehleranalyse der einzelnen ÖPNV-Relationen in OSM.

PTNA sollte auf der FOSSGIS 2020 in Freiburg vorgestellt werden. Wegen Corona gibt es lediglich einen Beitrag im Tagungsband [2]. Dieser Beitrag stellt die wichtigsten Erweiterungen an PTNA seit März 2020 vor: eine Analyse von GTFS-Daten, sowie einen Vergleich von OSM-Daten mit GTFS-Daten. „GTFS“ steht für: "General Transit Feed Specification", „GTFS is a community-driven open standard for rider-facing transit information.“ [3][4][5].

Anwendung von PTNA

PTNA wird seit Februar 2017 entwickelt um einen **SOLL-IST-Vergleich** im Rahmen von **OSMs ÖPNV-Linien** zu realisieren. Die folgenden 4 Anwendungsfälle sind derzeit möglich:

1.) IST-Analyse = Bestandsaufnahme

- ✓ Definition einer Area in dem die ÖPNV-Linien gesucht werden sollen
- ✓ Einfache Auflistung der gefundenen *route_master* und *route* Relationen nach Typ des Fahrzeuges und Linien-Nummer bzw. -Bezeichnung
- ✓ Dedizierte Fehleranalyse der Relationen basierend auf der „Public Transport Version 2“ („PTv2“) -Spezifikation [7]

2.) IST-Analyse = Bestandsaufnahme für einen oder mehrere ‚network‘-Werte

- ✦ Filtern nach ‚network‘-Werten, z.B.:
 - „Münchner Verkehrs- und Tarifverbund“ und „MVV“

3.) SOLL-IST-Analyse anhand einer Liste real existierender Linien

- ✦ Erstellung und Wartung einer SOLL-Liste (CSV-Format) im OSM-Wiki
- ✦ Lesen der CSV-Liste vom OSM-Wiki, z.B. ein Eintrag:
 - `210;bus;;Brunnthal, Zusestraße;Neuperlach Süd (U/S);Verkehrsbetrieb Ettenhuber GmbH`

- + Fehlermeldung: Linien in der CSV-Liste sind nicht in OSM gemapped
- + Auflistung nicht in der CSV-Liste stehender aber in OSM existierender Linien
- + Einfache Strukturierung des Report-Layout (Abschnitte, Text, ...)
- + Erweiterte Fehleranalyse von *route_master* und *route* Relationen
 - Passen *tags* des *route_masters* zu denen der zugehörigen *routen*

4.) SOLL-IST-Analyse anhand von Vorgaben aus GTFS-feed

- + Erstellung der CSV-Liste im OSM-Wiki durch Export der GTFS-routen
- + Wartung der CSV-Liste im OSM-Wiki mit Hilfe der GTFS-Daten
- + Erweiterung des CSV-Eintrags um GTFS-spezifische Elemente
 - *feed* = Name des Datensatzes
 - „*route_id*“ = ID der Linie im GTFS-Datensatz
 - „*release_date*“ = optionales Element: die Version des Datensatzes
 - 210;bus;;Brunnthal, Zusestraße;Neuperlach Süd (U/S);Verkehrsbetrieb Ettenhuber GmbH;DE-BY-MVV;mvv-19-210-1;
- + Analyse, Prüfung und Verlinkung der GTFS-Informationen im PTNA-Report
 - GTFS-Links auf die Analyse von GTFS-route bzw. GTFS-trip
 - Vergleichs-Icons zum Start eines Vergleichs von
 - GTFS-route mit OSM-route_master
 - GTFS-trip mit OSM-route

Über die Einstiegsseite von PTNA [1] gelangt man via Klick auf eine Flagge bei „**Take me to the results ...**“ auf eine Übersicht der verfügbaren Reports für das entsprechende Land. Die wichtigsten Spalten in der Tabelle:

- Die **1. Spalte**, „Name“ führt zum **Report** der Analysen
- Die **3.-letzte Spalte**, „Letzte Änderungen“ führt zu einem Report bei der zusätzlich die **Veränderungen seit der letzten Analyse** markiert sind
- Die **letzte Spalte**, „Linien“ führt zu den **SOLL-Daten**, der CSV-Liste im OSM-Wiki

Im Folgenden wird im Wesentlichen der zuletzt genannte, 4.) Anwendungsfall beschrieben, bei dem ein Vergleich von GTFS mit OSM möglich ist. Wo keine GTFS-Daten zur Verfügung stehen kommt Anwendungsfall 3.) zum Einsatz.

Die Vorbereitung der GTFS-Daten in PTNA wird erläutert. Die Begriffe aus GTFS und OSM werden anschließend kurz gegenübergestellt.

GTFS-Daten

Import der GTFS-Daten

GTFS-Daten müssen bezüglich ihrer Lizenz mit OSM kompatibel sein. CC0 [16] ist ideal, für CC-BY-4.0 [17] gilt das zum Beispiel nur, wenn die Form der „Namensnennung“ konkreter genannt wird. Für einige GTFS-Daten liegen Angaben seitens der Eigentümer vor, die die Namensnennung im Verzeichnis der Beitragenden im OSM-Wiki [6] als ausreichend bestätigen.

PTNA prüft GTFS-Quellen regelmäßig auf neue Versionen der Daten. Es wird in der Regel nur die erste im Monat verfügbare Version importiert. Auf Anfrage kann ein Import angestoßen werden.

Aufbereitung der GTFS-Daten

Beim Import der GTFS-Daten wird aus den diversen CSV-Dateien des Datensatzes eine sqlite3-Datenbank aufgebaut. Ein zweiter, wichtiger Schritt ist die Reduktion der Datenmenge durch Aggregation, aus OSM-Sicht, redundanter Informationen. Das ist im Wesentlichen die Zusammenfassung identischer Trips mit der selben Reihenfolge von Haltestellen aber unterschiedlichen Abfahrtszeiten zu einem repräsentativen Trip.

Analyse der GTFS-Daten

Um die Verwendbarkeit der GTFS-Daten für OSM zu bestimmen, werden von PTNA diverse Analysen durchgeführt:

- 1.) Ist ein Trip eine Teilroute eines anderen Trips bezüglich der Haltestellen
- 2.) Besteht ein Trip nur aus 2 Haltestellen (Ausnahme: Seilbahnen, Fähren, ...)
- 3.) Haben die erste und zweite Haltestelle den selben Namen, die selbe *stop_id*
- 4.) Haben die vorletzte und letzte Haltestelle den selben Namen, die selbe *stop_id*
- 5.) Unterscheiden sich zwei Trips nur durch die *shape_id* der Fahrstrecke
- 6.) Unterscheiden sich zwei Trips nur durch die *stop_id* der Haltestellen
- 7.) Beträgt die Fahrzeit eines Trips 0 Sekunden

Die Ergebnisse dieser Prüfungen werden als „PTNA Kommentar“ zum jeweiligen GTFS-*trip* ausgegeben. Diese Kommentare sollen als Hilfe beim Mappen dienen, wenn entschieden werden soll, ob ein GTFS-*trip* auch in den OSM-Daten gemapped werden soll. Gerade Punkt 6.) ist häufig bei GTFS-Daten für Züge zu sehen, die in Bahnhöfen mal das eine und mal ein anderes Gleis nutzen: *stop_name* identisch, *stop_id* unterschiedlich.

Normalisierung von Namen im GTFS-Datensatz

In GTFS-Datensätzen sind die Namen von Haltestellen häufig abgekürzt. In OSM hingegen gilt die Regel: Namen werden ausgeschrieben. Dieser Unterschied kann später beim Vergleich von GTFS- mit OSM-Daten zu Fehlermeldungen führen. PTNA führt daher für den deutschsprachigen Raum eine Normalisierung/Expansion von Namen für die GTFS-Daten durch, z.B.:

- „Friedenstr.“
 - de_DE und de_AT: „Friedenstraße“
 - de_CH: „Friedenstrasse“
- „W.-Heisenberg-W.“
 - „Werner-Heisenberg-Weg“
- „Bruck Abzw.“
 - de_DE und de_CH: „Bruck Abzweig“
 - de_AT: „Bruck Abzweigung“

Die Eigentümer der GTFS-Datensätze sind beim Abkürzen sehr kreativ: PTNA kann bezüglich der Normalisierung der Namen nie vollständig sein.

Darstellung der GTFS-Daten auf der PTNA Webseite

Über die Einstiegsseite von PTNA [1] gelangt man via Klick auf eine Flagge hinter „**Take me to the GTFS ...**“ auf eine Übersicht der **verfügbaren GTFS-Daten** für das entsprechende Land. Die wichtigste Spalte in der Tabelle:

- Die **1. Spalte**, „Name“ führt zur Übersicht der *GTFS-routen* des *GTFS-feeds*

Auf der folgenden Übersichtsseite werden die **GTFS-routen des GTFS-feeds** aufgelistet, sortiert nach Fahrzeugtyp und Linien-Nummer bzw. -Bezeichnung. Die wichtigste Spalte in der Tabelle:

- Die **1. Spalte**, „Linie“ führt zur Übersicht der *GTFS-trips* der *GTFS-route*

Auf der nun folgende Übersichtsseite werden die **GTFS-trips der GTFS-route** im Abschnitt „Existierenden Linienvarianten“ aufgelistet, sortiert nach den Namen der Haltestellen (1., letzte, 2., 3., ..., letzte). Diese Art der Sortierung hält Linienvarianten zusammen, die identische Start- und End-Haltestellen haben, dazwischen aber Unterschiede aufweisen. Die wichtigste Spalte in der Tabelle:

- Die **2. Spalte**, „Trip-ID“ führt zur Analyse des *GTFS-trips*

Auf der finalen Übersichtsseite erfolgt die **Analyse des GTFS-trips** mit Karte, Taggingvorschlägen für OSM, Haltestellentabelle, Verkehrszeiten und, wenn vorhanden, dem Fahrweg (*shape*).

Begriffe: GTFS versus OSM

GTFS und OSM nutzen unterschiedliche Begriffe für Objekte des ÖPNV. Tabelle 1 zeigt einen Überblick über die von PTNA verwendeten Elemente und deren Entsprechungen. Für OSM gilt dabei die so genannte „Public Transport Version 2“ („PTv2“) [7].

GTFS Begriff	OSM Begriff	Anmerkung
<i>feed</i>	<i>network:guid</i>	Viele GTFS-Datensätze werden von den Verkehrsverbünden herausgegeben. PTNA vergibt generische <i>feed</i> -Namen [8], die in OSM als <i>network:guid</i> verwendet werden können: <ul style="list-style-type: none"> - <i>network</i> = Münchner Verkehrs- und Tarifverbund - <i>network:short</i> = MVV - <i>network:guid</i> = DE-BY-MVV
<i>route</i>	<i>route_master</i>	Fasst alle Fahrten einer ÖPNV-Linie zusammen.
<i>route_id</i>		Eindeutige ID einer <i>route</i> .
<i>route_short_name</i>	<i>ref</i>	Das OSM tag <i>ref</i> einer <i>route_master/route</i> Relation.
<i>route_color</i>	<i>colour</i>	Die Hintergrundfarbe der Linienbezeichnung.
<i>route_text_color</i>	<i>colour:text</i>	Die Textfarbe der Linienbezeichnung.
<i>trip</i>	<i>route</i>	Ein GTFS- <i>trip</i> entspricht einer einzelnen Fahrt zu einer bestimmten Uhrzeit, definiert durch die angefahrenen Stops und optionaler Fahrstrecke (<i>shape</i>). Eine OSM- <i>route</i> beschreibt alle Fahrten, die die selben Stops anfahren plus deren Fahrstrecke.
<i>trip_id</i>		Eindeutige ID eines <i>trips</i> .

GTFS Begriff	OSM Begriff (PTv2)	Anmerkung
<i>shape</i>		In GTFS: die eigentliche Fahrstrecke (optional). In OSM ist das die Kette der <i>highway/railway/...</i> member einer <i>route</i> Relation (ohne ‚role‘).
<i>shape_id</i>		Eindeutige ID eines <i>shapes</i> .
<i>stop</i>	<i>public_transport=platform</i>	Ein <i>stop</i> entspricht der neben dem Fahrweg liegenden <i>platform</i> .
<i>stop_id</i>	<i>ref:IFOPT, ...</i>	Eindeutige ID eines <i>stops</i> , nicht immer identisch mit <i>ref:IFOPT, ...</i>
<i>stop_lat</i>	<i>lat</i>	Breitengrad der Haltestelle.
<i>stop_lon</i>	<i>lon</i>	Längengrad der Haltestelle.
<i>stop_name</i>	<i>name</i>	Name einer Haltestelle.
	<i>public_transport=stop_position</i>	Den Begriff „Halteposition des Fahrzeugs“ gibt es in GTFS nicht.
<i>agency_name</i>	<i>operator</i>	Bei manchen GTFS-Daten kann der Name der <i>agency</i> als OSM- <i>operator</i> einer <i>route_master/route</i> Relation genutzt werden.

Tabelle 1: Vergleich GTFS vs. OSM Begriffe

Für den Vergleich von GTFS- und OSM-Daten können die aufbereiteten und analysierten GTFS-Daten auf verschiedene Arten genutzt werden.

Vergleich von GTFS-Daten mit OSM-Daten

Nutzung von GTFS-Daten in der CSV-Liste im OSM-Wiki

Die Definition des GTFS-feed und der GTFS-route_id (**SOLL-Daten**) erfolgt in der **CSV-Liste im OSM-Wiki**.

- Dieser Schritt sollte immer zuerst erfolgen
- Änderungen erfolgen lediglich an Daten im OSM-Wiki, nicht an OSM-Daten
- Die erstmalige Erstellung kann durch PTNA unterstützt werden
 - Export der GTFS-routen als CSV-Liste
- Wartung hält sich in Grenzen,
 - wenn die GTFS-route_id stabil [5] bleibt und weil GTFS-trip_ids nicht enthalten sind
 - Lediglich neue und eingestellte Linien müssen dann eingepflegt werden
- Im PTNA-Report erscheint jeweils ein **GTFS-Link** und ein **Vergleichs-Icon**
- Beispiel für CSV-Daten im OSM-Wiki für den Bus 210 im Münchner Verkehrs- und Tarifverbund:
 - `210;bus;;Brunnthal, Zusestraße;Neuperlach Süd (U/S);Verkehrsbetrieb Ettenhuber GmbH;DE-BY-MVV;mvv-19-210-1`

Nutzung von GTFS-Daten in OSM route_master / route Relationen

Die Definition des GTFS-feed, der GTFS-route_id bzw. GTFS-trip_id erfolgt im OSM-Datensatz der route_master- bzw. route-Relation als **IST-Daten** [9].

- Änderungen erfolgen an OSM-Daten
- Aufwändig in der Erstellung
- Gegebenenfalls aufwändig in der Wartung
 - Das hängt von der Stabilität der GTFS-id-Daten über mehrere Versionen hinweg ab (route_id, trip_id, shape_id, stop_id) [5]
- Erstellung und Wartung kann durch den vorangegangenen Abschnitt (GTFS-Daten in CSV-Liste im OSM-Wiki) und den dadurch möglichen Vergleich erleichtert werden
- Im PTNA-Report erscheint für jede so getaggte Relation ein **GTFS-Link** und ein **Vergleichs-Icon**
- Beispiel für die tags in einer OSM-route-Relation des Busses 210:
 - `gtfs:feed=DE-BY-MVV`
 - `gtfs:route_id=mvv-19-210-1`
 - `gtfs:trip_id:sample=mvv-19-210-1-1-1-H-0-MoTuWeThFrSaSu-119-344`

Bedeutung von Links im PTNA-Report

Abbildung 1 zeigt einen Auszug aus einem PTNA-Report. Der Auszug korrespondiert mit dem folgenden Eintrag in den CSV-Daten im OSM-Wiki:

210;bus;;Brunnthal, Zusestraße;Neuperlach Süd (U/S);Verkehrsbetrieb Ettenhuber GmbH;DE-BY-MVV;mvv-19-210-1




Name (name=)	Typ (type=)	Relation (id=)	PTv	Fehler	Anmerkungen
210		Von: Brunnthal, Zusestraße; Nach: Neuperlach Süd (U/S); Betreiber: Verkehrsbetrieb Ettenhuber GmbH			1
Bus 210	route_master	67811 (ID, JOSM, UNROLL, GTFS, )			2
Bus 210: Brunnthal, Zusestraße => Neuperlach Süd (S/U)	route	1549761 (ID, JOSM, Relativify, PTNA, GTFS, )	2	7	
Bus 210: Neuperlach Süd (S/U) => Brunnthal, Zusestraße	route	1549762 (ID, JOSM, Relativify, PTNA, GTFS, )	2	6	

Abbildung 1: Auszug aus PTNA-Report für den Bus 210

Für die Analyse und den Vergleich von GTFS- und OSM-Daten finden sich diverse Links:

① Ein **GTFS**-Link in der Kopfzeile außen rechts führt zu einer **Analyse** der GTFS-route mit einem Überblick über die zugehörigen GTFS-trips. Die genutzten GTFS-Daten werden der **CSV-Liste im OSM-Wiki** (= SOLL-Daten) entnommen (Abbildung 2).

Variante	Trip-ID	Anzahl Fahrten	Gültigkeit		Haltestellen		PTNA Info	
			Von	Bis	Start	Über	Ende	Kommentar
1	mvv-19-210-1-1-1-H0-MoTuWeThFrSaSu-119-344	7037 (7329)	2025-01-13	2025-03-30	Brunnthal, Zusestraße	Taufk, W.-Messerschmitt-Str. => Taufkirchen, Lilienthalstr. => Ottobrunn, Einsteinstraße => Ottobrunn, Röntgenstraße => Ottobrunn, Robert-Koch-Str. => Ottobrunn, Ottostraße => Ottobrunn, Hubertusstraße => Ottobrunn, Ortsmitte => Ottobrunn, Jahnstraße => Ottobrunn, Bahnhofstraße => Ottobrunn, Alte Landstraße => Neubiberg, W.-Heisenberg-W.	Neuperlach Süd	
2	mvv-19-210-1-1-4-R0-Su-1-1	6896	2025-01-13	2025-03-30	Neuperlach Süd	Neubiberg, W.-Heisenberg-W. => Ottobrunn, Alte Landstraße => Ottobrunn, Bahnhofstraße => Ottobrunn, Jahnstraße => Ottobrunn, Ortsmitte => Ottobrunn, Hubertusstraße => Ottobrunn, Ottostraße => Ottobrunn, Finkenstraße => Ottobrunn, Robert-Koch-Str. => Ottobrunn, Röntgenstraße => Ottobrunn, Einsteinstraße => Taufkirchen, Lilienthalstr. => Taufk, W.-Messerschmitt-Str.	Brunnthal, Zusestraße	
3	mvv-19-210-1-1-7-R1-MoTuWeThFr#2-225-268	624	2025-01-13	2025-03-28	Neuperlach Süd	Neubiberg, W.-Heisenberg-W. => Ottobrunn, Alte Landstraße => Ottobrunn, Bahnhofstraße => Ottobrunn, Jahnstraße	Ottobrunn, Ortsmitte	
4	mvv-19-210-1-1-2-H1-MoTuWeThFr#2-219-447	288	2025-01-13	2025-03-28	Ottobrunn, Jahnstraße	Ottobrunn, Bahnhofstraße => Ottobrunn, Alte Landstraße => Neubiberg, W.-Heisenberg-W.	Neuperlach Süd	Fahrt ist Teilroute von: mvv-19-210-1-1-1-H0-MoTuWeThFrSaSu-119-344

Abbildung 2: Übersicht über die trips einer GTFS-route

② Ein **Vergleichs**-Icon in der Kopfzeile außen rechts führt zu einem Vergleich der GTFS-route mit dem OSM-route_master. Die genutzten GTFS-Daten werden der **CSV-Liste im OSM-Wiki** (= SOLL-Daten) entnommen (Abbildung 7 im folgen Abschnitt: „Vergleich von GTFS-route- mit OSM-route_master-Daten“).

- ③ Ein **GTFS**-Link in der dritten Spalte einer OSM-*route_master*-Relation führt zu einer **Analyse** der GTFS-*route* mit einem Überblick über die zugehörigen GTFS-*trips*. Die genutzten GTFS-Daten werden den **gtfs:* Angaben der route_master-Relation** (= IST-Daten) entnommen (siehe Abbildung 2 oben).
- ④ Ein **Vergleichs**-Icon in der dritten Spalte einer OSM-*route_master*-Relation führt zu einem Vergleich der GTFS-*route* mit dem OSM-*route_master*. Die genutzten GTFS-Daten werden den **gtfs:* Angaben der route_master-Relation** (= IST-Daten) entnommen (siehe Abbildung 7 im folgenden Abschnitt: „Vergleich von GTFS-*route*- mit OSM-*route_master*-Daten“).
- ⑤ Ein **GTFS**-Link in der dritten Spalte einer OSM-*route*-Relation führt zu einer **Analyse** des GTFS-*trips*, mit Karte und Details (scrollen) zum Trip. Die genutzten GTFS-Daten werden den **gtfs:* Angaben der route-Relation** (= IST-Daten) entnommen (Abbildung 4).

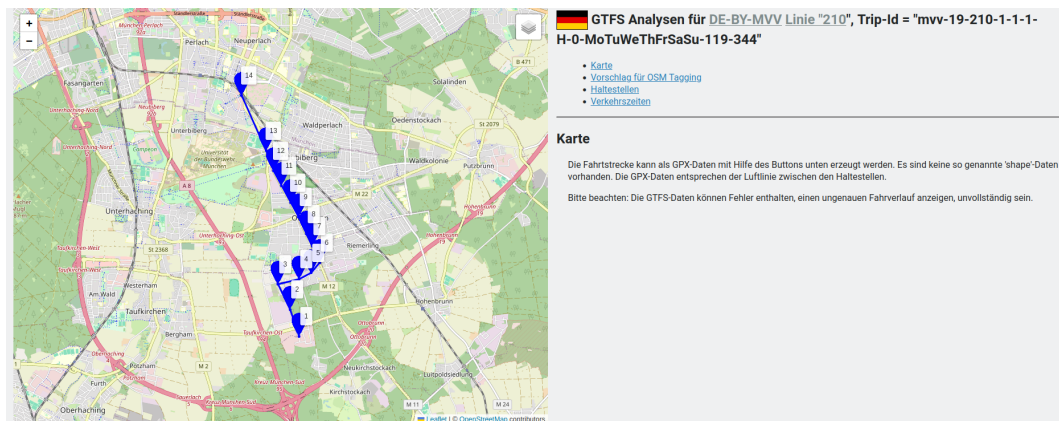


Abbildung 4: Analyse eines GTFS-*trips* durch PTNA

- ⑥ Ein **Vergleichs**-Icon in der dritten Spalte einer OSM-*route*-Relation führt zu einer Gegenüberstellung von GTFS-*trip* und OSM-*route*. Die genutzten GTFS-Daten werden den **gtfs:* Angaben der route-Relation** (= IST-Daten) entnommen (siehe Abbildungen 8 und 9 im folgenden Abschnitt: „Vergleich von GTFS-*trip*- mit OSM-*route*-Daten“).

- ⑦ Ein **PTNA-Link** in der dritten Spalte führt zu einer **Analyse** der OSM-route-Relation, mit Karte und Details der einzelnen Elemente der Relation (Abbildung 3).

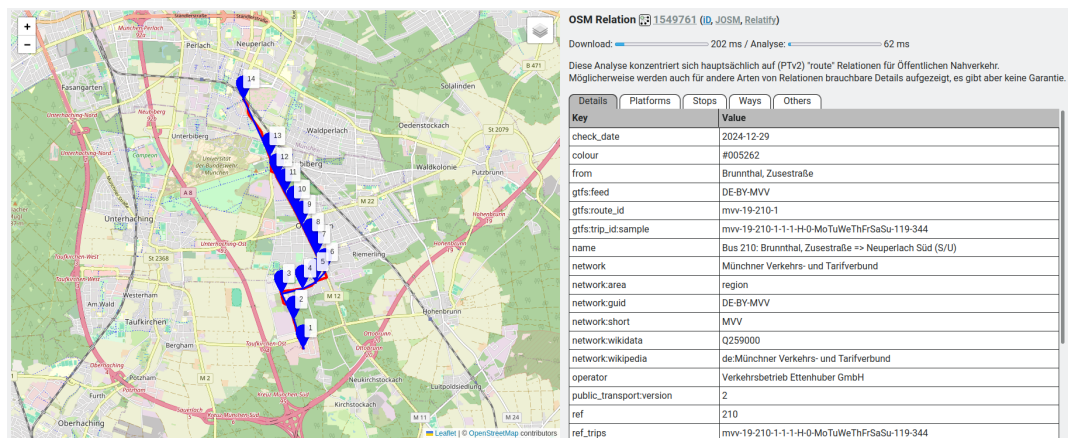


Abbildung 3: Analyse einer OSM-route-Relation durch PTNA

Abbildung 4 zeigt eine OSM-route-Relation, bei der die Reihenfolge der Haltestellen nicht korrekt erscheint: Haltestellen (PTV2-*platform*) 3 und 4 sind vertauscht.

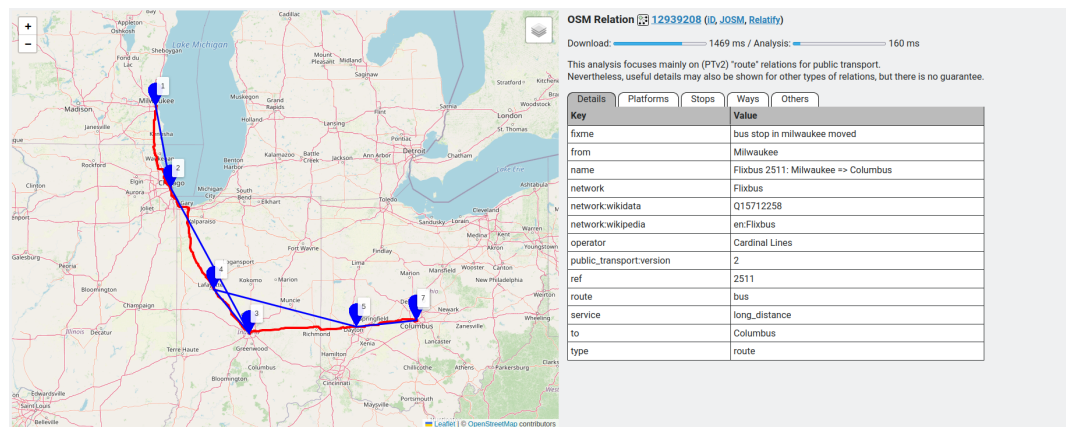


Abbildung 4: Analyse einer fehlerhaften OSM-route-Relation durch PTNA

Visueller Vergleich

Hierbei erfolgt der Vergleich eines GTFS-*trip* und einer OSM-*route* auf getrennten Karten

- Keinerlei Änderungen an bestehenden OSM-Daten notwendig
- Mühsam in der Anwendung.

Die Links zu **GTFS** (Link 5) und **PTNA** (Link 7) im PTNA-Report führen für GTFS-*trip* und OSM-*route* zu identisch aufgebauten Karten auf den Webseiten.

Ohne GTFS-Link im PTNA-Report muss die Analyse über die Länderspezifische GTFS-Übersicht, den GTFS-*feed*, die GTFS-*route* und den passenden GTFS-*trip* gesucht werden.

Ein visueller Vergleich kann mittels zweier Browserfenster erfolgen (Abbildung 5 und 6 zeigen Ausschnitte der Fenster).

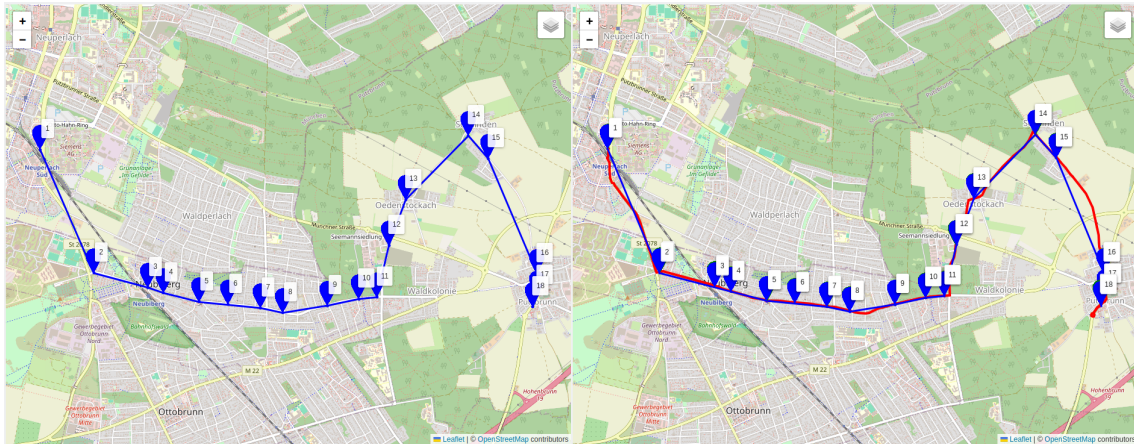


Abbildung 5: GTFS-*trip* auf Karte (ohne *shape*) Abbildung 6: OSM-*route* auf Karte (mit *shape*)

Vergleich von GTFS-route- mit OSM-route_master-Daten

Der Link bei dem **Vergleichs-Icon** für eine **OSM-route_master**-Relation führt zu dem Vergleich einer GTFS-route versus OSM-route_master. Der Vergleich erfolgt in Tabellenform mit so genannten **score**-Werten für jede Kombination GTFS-trip/OSM-route: je kleiner der **score**-Wert, desto besser die Übereinstimmung der beiden Datensätze. Ist ein **score**-Wert in fatter Schrift abgebildet, so stimmen *gtfs:feed* und *gtfs:trip_id:sample* in der OSM-route-Relation mit der GTFS-trip_id überein (Abbildung 7).

Compare GTFS route with OSM route_master

General information

	type	links	id	ref	feed	release date	members
Rows:	GTFS route		mvv-19-214-1	214	DE-BY-MVV	latest	GTFS trips of GTFS route
Columns:	OSM route_master		4280997	214			OSM routes of OSM route_master

Mismatch Score Table

Small values indicate a good match between GTFS trip and OSM route/GTFS trip.
For a more detailed comparison, click on a number.

Select rows where all scores >= x %
30

Add to selection

Replace selection

Show all			Hide selected			Clear selections			OSM routes - <input type="checkbox"/> Show OSM route 'name'			
	Num	Rides	GTFS trips		11 Eugen-Sänger-Ring => 12 platforms => Ottobrunn	11 Hubertusstraße => 18 platforms => Riemerling, Altenheim	11 Ottobrunn (S) => 12 platforms => Eugen-Sänger-Ring	11 Riemerling, Altenheim => 20 platforms => Hubertusstraße				
<input type="checkbox"/>	1	212	Brunnthal, Eugen-Sänger-Ring => 12 stops => Ottobrunn		0.22%	39.53%	42.38%	51.72%				
<input type="checkbox"/>	2	1464	Hohenbrunn, Hubertusstraße => 18 stops => Riemerling, Altenheim		44.16%	0.16%	57.50%	45.53%				
<input type="checkbox"/>	3	159	Ottobrunn => 12 stops => Brunnthal, Eugen-Sänger-Ring		42.38%	52.81%	0.66%	50.13%				
<input type="checkbox"/>	4	1581	Riemerling, Altenheim => 20 stops => Hohenbrunn, Hubertusstraße		57.19%	46.94%	55.75%	0.44%				

Abbildung 7: Vergleich von GTFS-route mit OSM-route_master mit **score**-Tabelle

Ein Klick auf einen **score**-Wert führt zum Vergleich eines GTFS-trips mit einer OSM-route (nächster Abschnitt).

Vergleich von GTFS-trip- mit OSM-route-Daten

Der Link bei dem **Vergleichs-Icon** für eine **OSM-route**-Relation führt zu dem Vergleich GTFS-trip versus OSM-route. Der Vergleich erfolgt auf einer Karte und in Tabellenform mit den **score**-Werten für einzelne Metriken: je kleiner der **score**-Wert, desto besser die Übereinstimmung.

Die Metriken basieren auf Eigenschaften von Haltestellen: Anzahl, Position, Entfernung, Namen, ID, ... Die Shapes spielen keine Rolle, sie sind häufig nicht enthalten, mitunter fehlerhaft und schwer zu vergleichen (häufig: „false-positives“).

Beim Scrollen nach unten wird ein Vergleich der Haltestellen in Tabellenform sichtbar.

Abbildungen 8 und 9 zeigen einen **positiven Fall**, wo die Übereinstimmung sehr gut ist. Lediglich die Positionen der Haltestelle 12 stimmt nicht ganz überein: OSM ist hier richtig.

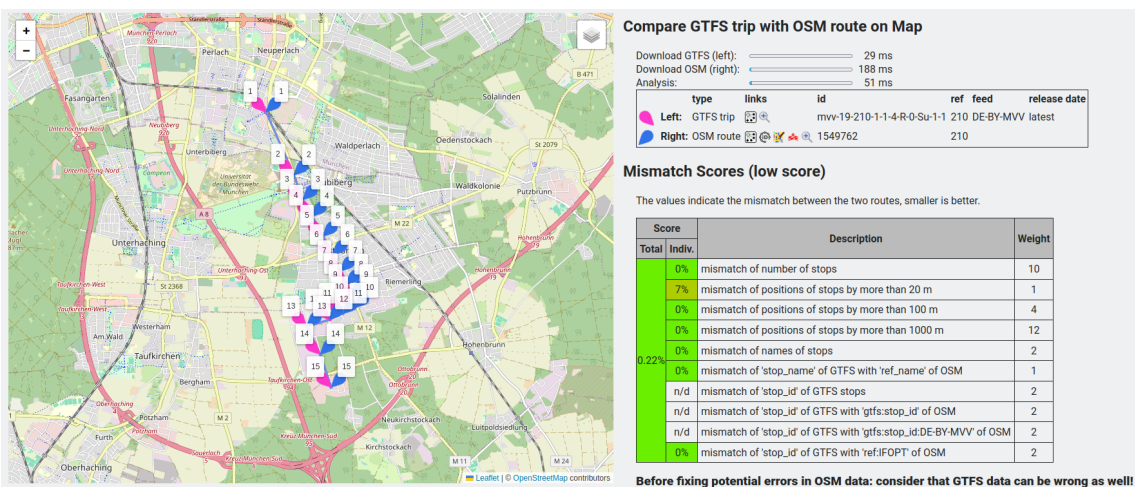


Abbildung 8: Vergleich von GTFS-trip mit OSM-route auf Karte und **score**-Tabelle: positiv
















































Compare GTFS trip with OSM route side-by-side (stops/platforms)												
Stop Number	Stop data of GTFS trip 				Distance [m]	Platform data of OSM route 					Platform Number	Edit with
	stop_id	stop_lat	stop_lon	stop_name		name	ref_name	lat	lon	ref:FOPT		
1	de:09162:1010:5:5	48.08947	11.64406	Neuperlach Süd	◀ 1 ▶	Neuperlach Süd	48.08946	11.64405	de:09162:1010:5:5	1	  	
2	de:09184:2302:0:3	48.07826	11.65148	Neubiberg, Werner-Heisenberg-Weg	◀ 4 ▶	Werner-Heisenberg-Weg	48.07824	11.65143	de:09184:2302:0:3	2	  	
3	de:09184:2189:0:2	48.07389	11.65392	Ottobrunn, Alte Landstraße	◀ 1 ▶	Alte Landstraße, Ottobrunn	48.07387	11.65391	de:09184:2189:0:2	3	  	
4	de:09184:2236:0:2	48.07095	11.65627	Ottobrunn, Bahnhofstraße	◀ 4 ▶	Bahnhofstraße, Ottobrunn	48.07091	11.65628	de:09184:2236:0:2	4	  	
5	de:09184:2235:0:2	48.06738	11.65917	Ottobrunn, Jahnstraße	◀ 10 ▶	Jahnstraße, Ottobrunn	48.06730	11.65924	de:09184:2235:0:2	5	  	
6	de:09184:2238:3:2	48.06392	11.66175	Ottobrunn, Ortsmitte	◀ 5 ▶	Ottobrunn, Ortsmitte	48.06397	11.66174	de:09184:2238:3:2	6	  	
7	de:09184:2242:0:2	48.06108	11.66386	Ottobrunn, Hubertusstraße	◀ 1 ▶	Hubertusstraße, Ottobrunn	48.06109	11.66385	de:09184:2242:0:2	7	  	
8	de:09184:2192:3:2	48.05876	11.66551	Ottobrunn, Ottostraße	◀ 2 ▶	Ottostraße, Ottobrunn	48.05874	11.66550	de:09184:2192:3:2	8	  	
9	de:09184:2196:0:1	48.05684	11.66688	Ottobrunn, Finkenstraße	◀ 0 ▶	Finkenstraße, Ottobrunn	48.05684	11.66688	de:09184:2196:0:1	9	  	
10	de:09184:2203:0:1	48.05463	11.66774	Ottobrunn, Robert-Koch-Straße	◀ 11 ▶	Robert-Koch-Straße, Ottobrunn	48.05460	11.66760	de:09184:2203:0:1	10	  	
11	de:09184:2247:0:2	48.05355	11.66458	Ottobrunn, Röntgenstraße	◀ 2 ▶	Röntgenstraße, Ottobrunn	48.05354	11.66456	de:09184:2247:0:2	11	  	
12	de:09184:2245:0:2	48.05252	11.66115	Ottobrunn, Einsteinstraße	◀ 23 ▶	Einsteinstraße, Ottobrunn	48.05250	11.66084	de:09184:2245:0:2	12	  	
13	de:09184:2389:0:3	48.05128	11.65496	Taufkirchen, Lilienthalstraße	◀ 3 ▶	Lilienthalstraße, Taufkirchen	48.05126	11.65493	de:09184:2389:0:3	13	  	
14	de:09184:2437:0:1	48.04629	11.65860	Taufkirchen, Willy-Messerschmitt-Straße	◀ 3 ▶	Willy-Messerschmitt-Straße, Taufkirchen	48.04630	11.65856	de:09184:2437:0:1	14	  	
15	de:09184:2315:0:1	48.04050	11.66141	Brunnthal, Zusestraße	◀ 8 ▶	Zusestraße, Brunnthal	48.04043	11.66144	de:09184:2315:0:1	15	  	

Abbildung 9: Vergleich von GTFS-trip mit OSM-route mit Haltestellen-Tabelle: positiv

Abbildungen 10 und 11 zeigen einen **negativen Fall**, wo die Übereinstimmung schlecht ist. Der **score** ist 22.06%. Der GTFS-*trip* stellt bezüglich Stops 6 und 7 die Werktagsvariante bis ca. 20:30 dar, die OSM-*route* stellt bezüglich Haltestelle 6 die Abend- und Sonntagsvariante dar (fährt nicht durch das Gewerbegebiet).

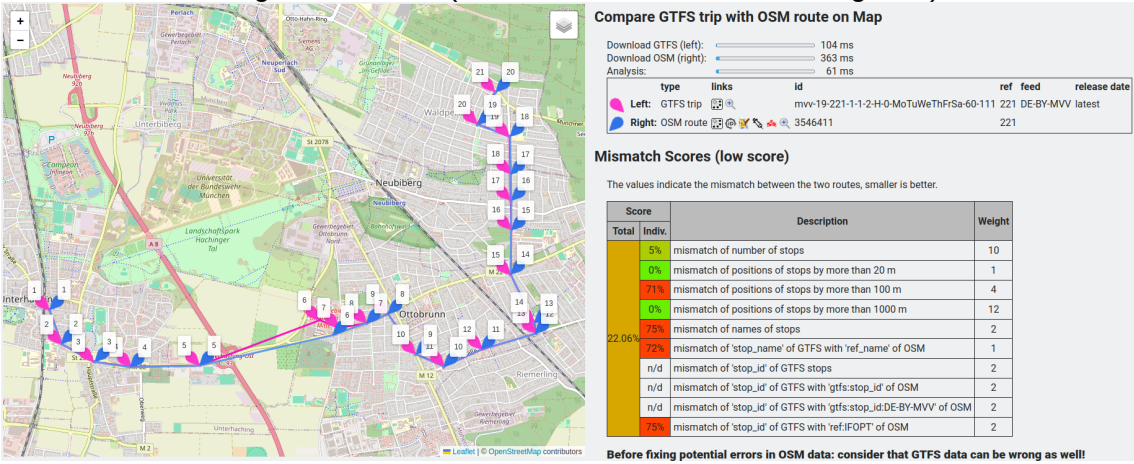


Abbildung 10: Vergleich GTFS-*trip* (werktags) mit OSM-*route* (abends, sonntags): Karte/Scores

Compare GTFS trip with OSM route side-by-side (stops/platforms)

Stop Number	Stop data of GTFS trip: mvv-19-221-1-1-2-H-0-MoTuWeThFrSa-60-111				Distance [m]	Platform data of OSM route: 3546411					Platform Number	Edit with
	stop_id	stop_lat	stop_lon	stop_name		name	ref_name	lat	lon	ref:FOPT		
1	de:09184:2310:1:1	48.06520	11.61318	Unterhaching	3	Unterhaching (S)		48.06521	11.61314	de:09184:2310:1:1	1	
2	de:09184:2366:0:1	48.06204	11.61489	Unterhaching, Jahnstraße	11	Jahnstraße		48.06209	11.61477	de:09184:2366:0:1	2	
3	de:09184:2368:0:1	48.06044	11.61926	Unterhaching, Ottobrunner Straße	11	Ottobrunner Straße	Unterhaching, Ottobrunner Straße	48.06040	11.61940	de:09184:2368:0:1	3	
4	de:09184:2378:0:1	48.05995	11.62448	Unterhaching, Friedhof	16	Friedhof	Unterhaching, Friedhof	48.05992	11.62427	de:09184:2378:0:1	4	
5	de:09184:2365:0:1	48.06008	11.63392	Unterhaching, Kirchlandweg	0	Kirchlandweg	Unterhaching, Kirchlandweg	48.06008	11.63392	de:09184:2365:0:1	5	
6	de:09184:2206:0:1	48.06428	11.65058	Ottobrunn, Siemensstraße	204	Daimlerstraße, Ottobrunn		48.06290	11.65239	de:09184:2380:2:3	6	
7	de:09184:2380:1:1	48.06355	11.65324	Ottobrunn, Daimlerstraße	285	Lindenstraße	Lindenstraße, Ottobrunn	48.06396	11.65703	de:09184:2239:0:1	7	
8	de:09184:2239:0:1	48.06399	11.65709	Ottobrunn, Lindenstraße	237	Ottobrunn, Ortsmitte		48.06484	11.66001	de:09184:2238:2:3	8	
9	de:09184:2238:2:3	48.06484	11.66005	Ottobrunn, Ortsmitte	504	Hubertusstraße	Ottobrunn, Hubertusstraße	48.06109	11.66385	de:09184:2242:0:2	9	
10	de:09184:2242:0:2	48.06108	11.66386	Ottobrunn, Hubertusstraße	313	Friedrich-Ebert-Platz	Friedrich-Ebert-Platz, Ottobrunn	48.05993	11.66770	de:09184:2281:1:1	10	
11	de:09184:2281:1:1	48.06001	11.66786	Ottobrunn, Friedrich-Ebert-Platz	409	Hans-Kandler-Weg	Hans-Kandler-Weg, Ottobrunn	48.06155	11.67285	de:09184:2190:0:1	11	
12	de:09184:2190:0:1	48.06158	11.67303	Ottobrunn, Hans-Kandler-Weg	566	Ottobrunn (S)		48.06297	11.68037	de:09184:2210:1:1	12	
13	de:09184:2210:1:1	48.06305	11.68055	Ottobrunn	106	Karl-Stieler-Straße	Karl-Stieler-Straße, Ottobrunn	48.06398	11.68030	de:09184:2188:1:1	13	
14	de:09184:2188:1:1	48.06409	11.68024	Ottobrunn, Karl-Stieler-Straße	545	Am Brunneck/Uhlandstraße	Am Brunneck/Uhlandstraße, Ottobrunn	48.06489	11.67699	de:09184:2187:0:1	14	
15	de:09184:2187:0:1	48.06487	11.67701	Ottobrunn, Am Brunneck/Uhlandstraße	456	Neubiberg, Gymnasium	Gymnasium, Neubiberg	48.07257	11.67713	de:09184:2278:0:1	15	
16	de:09184:2278:0:1	48.07261	11.67709	Neubiberg, Gymnasium	300	Walkürenstraße	Walkürenstraße, Neubiberg	48.07531	11.67708	de:09184:2275:0:1	16	
17	de:09184:2275:0:1	48.07534	11.67705	Neubiberg, Walkürenstraße	266	Brunhildenstraße	Brunhildenstraße, Neubiberg	48.07773	11.67704	de:09184:2290:0:2	17	
18	de:09184:2290:0:2	48.07781	11.67701	Neubiberg, Brunhildenstraße	382	Salzmannstraße	Salzmannstraße, München	48.08125	11.67690	de:09162:968:3:SAL	18	
19	de:09162:968:3:SAL	48.08126	11.67688	Salzmannstraße	350	Waldperlach	Waldperlach, München	48.08230	11.67243	de:09162:967:1:1	19	
20	de:09162:967:1:1	48.08237	11.67239	Waldperlach	383	Waldheimplatz	Waldheimplatz, München	48.08536	11.67496	de:09162:966:5:6	20	
21	de:09162:966:5:6	48.08538	11.67486	Waldheimplatz	8							

Abbildung 11: Vergleich GTFS-*trip* (werktags) mit OSM-*route* (abends, sonntags): Tabelle

Icons und ihre Bedeutung

Auf den Seiten des PTNA-Reports und der GTFS-/OSM-Vergleiche sind diverse Icons zu sehen. Diese führen entweder zu weiteren Informationen oder laden relevante Daten in OSM-Editoren, wo eine Bearbeitung erfolgen kann.



Das Vergleichs-Icon: ein Klick führt zum Vergleich von GTFS- mit OSM-Daten. Auf der Übersichtsseite für GTFS-*feeds* führt ein Klick zum Vergleich zweier GTFS-*feed*-Versionen.



Lade die OSM Daten des Objektes (*relation*, *way* oder *node*) in den **JOSM**-Editor.



Mit Hilfe des **JOSM**-Editors werden Daten (key, value) in die OSM-*route_master*- bzw. *route*-Relation injiziert. Mouseover zeigt an, welche Daten das sind.



Lade die OSM-Daten (*relation*, *way* oder *node*) in den **iD**-Editor.



Lade die OSM-*route*-Relation in den **Relatify**-Editor.



Der angezeigte Name (GTFS-*stop_name*) wurde „Normalisiert“. Mouseover zeigt den Originalnamen.



Zeige eine Kurzinformationen der Objektdaten in einem Pop-Up-Fenster.



Der angezeigte GTFS-*trip* stellt eine Teilroute (sub-route) mindestens eines anderen GTFS-*trips* dar. Mouseover zeigt an, welcher das ist.



Der angezeigte GTFS-*trip* hat verdächtige Merkmale. Mouseover zeigt an, welche das sind



Der angezeigte GTFS-*trip* hat nur zwei Haltestellen.



Der angezeigte GTFS-*trip* ist bezüglich *stop_name* identisch mit mindestens einem anderen GTFS-*trip* unterscheidet sich aber in *stop_id* bzw. *shape_id*. Mouseover zeigt an, welcher das ist.



Das Objekt ist ein Punkt (OSM: *node*)
Ein Klick auf das Icon bzw. die ID zeigt die Objektdaten auf <https://osm.org> an.



Das Objekt ist eine Strecke (OSM: *way*)
Ein Klick auf das Icon bzw. die ID zeigt die Objektdaten auf <https://osm.org> an.



Das Objekt ist eine Sammlung (GTFS: *route*, *trip*; OSM: *relation*)
GTFS : ein Klick führt zur Übersicht von *route* bzw. *trip*.
OSM : ein Klick auf das Icon bzw. die ID zeigt die Objektdaten auf <https://osm.org> an.



Diese Zeichen im Tabellenkopf einer Spalte erlauben eine Sortierung der Tabelle nach dieser Spalte: nicht sortiert, aufsteigend, absteigend.

Sonstiges

Auf die Angabe von Beispielen mit IDs von realen GTFS- oder OSM-Daten wird hier bewusst verzichtet. Diese können bis zum Zeitpunkt der Konferenz schon wieder veraltet sein.

Als maximale akzeptable Distanz zwischen GTFS-*stop* und zugehöriger OSM-*platform* wurden 20 Meter gewählt. Das entspricht etwa der Länge eines Gelenkbusses (18 m) und erlaubt dennoch das Erkennen von falschen Bussteigen in Busbahnhöfen. Beim Vergleich lassen sich gleichwohl alle „Distanz-Werte“ und „Gewichtungen“ GTFS-*feed*-spezifisch konfigurieren.

Über das Vergleichs-Icon in der letzten Spalte der Tabelle einer länderspezifischen GTFS-Übersicht (Beispiel: DE [10]) lässt sich ein Vergleich GTFS versus GTFS anstoßen, z.B.: was hat sich zwischen zwei GTFS-Versionen geändert?

Im PTNA-Report gibt es am Anfang unter „Datum der Daten“ einen Button „**Start einer neuen Analyse anfordern**“. Ein Cron-Job schaut alle 15 Minuten nach solchen Anforderung und startet diese. Man muss nicht auf den nächtlichen Start (zwischen 02:00-05:00 Ortszeit der zu analysierenden Region) des PTNA-Reports warten.

Die Protokolle (log) und Ergebnisse aller automatisierten Aktivitäten von PTNA sind über die „**Statistik**“-Seite [11] (vor allem für den Autor selbst) erreichbar.

Ausblick / Ideen:

- Weiterführung der Umstellung des OSM-Daten-Downloads von Overpass API [18] zu Planet-Dumps/Updates/Filter/Extracts [19][20][21]
 - 109 Analysen sind umgestellt, 244 Konfigurationen müssen noch angepasst werden
- Ausgabe der Ergebnisse (score-Werte) des Vergleichs von GTFS- und OSM-Daten im PTNA-Report
 - Basierend auf den CSV-Daten im OSM-Wiki
 - Basierend auf den *gtfs:** tags in den OSM-Relationen
 - Portieren von > 1000 Lines-of-Code von JavaScript nach Perl
 - Problem: nur wenige score-Werte sind tatsächlich 0.00%, wie behandelt man die anderen?
- Einbau einer „sidebar“ über die Fehlermeldungen (Kategorien) im PTNA-Report ein- und ausgeblendet werden können
 - Analog zu KeepRight [12] und PTSA [13]
- Erweitern von Statistiken
 - Entwicklung der Anzahl Fehler einer Kategorie über einen Zeitraum
- Re-Design des sqlite3-Datenbankschemas (PTNA-Tabellen) für GTFS-Daten
 - Man lernt nie aus
- Bei der PTNA-Analyse (Link 7) einer OSM-*route*-Relation könnte man eine Prüfung auf Lücken und falsche Sortierung bei den Ways einbauen
 - Analog zum Relations-Editor in JOSM
 - Portieren von 600-1000 Lines-of-Code in Perl nach JavaScript

Anhang

Kontakt zum Autor:

Antonius „Toni“ Erdmann

Friedenstraße 21

D-85521 Ottobrunn

osm-ToniE@web.de

Community: „PTNA: news for Public Transport Network Analysis“ [14]

PTNA-Source-Code [15]

Literatur / Links:

- [1] „PTNA - Public Transport Network Analysis“, <https://ptna.openstreetmap.de>
- [2] „PTNA: Qualitätssicherung für ÖPNV in OpenStreetMap“, Toni Erdmann, FOSSGIS 2020 Tagungsband, Seite 109, https://www.fossgis-konferenz.de/2020/data/FOSSGIS_Ta-gungsband_2020.pdf
- [3] „GTFS“, "General Transit Feed Specification", <https://gtfs.org/>
- [4] „GTFS reference“, <https://gtfs.org/documentation/schedule/reference/>
- [5] „GTFS best practices“, <https://gtfs.org/documentation/schedule/schedule-best-practices/>
- [6] „OSM contributors“, <https://wiki.openstreetmap.org/wiki/Contributors>
- [7] „Proposal:Public Transport“, approved version “PTv2“, https://wiki.openstreetmap.org/w/index.php?title=Proposed_features/Public_Transport&oldid=625726
- [8] „Codes for identifying the principal subdivisions (e.g., provinces or states) of all countries coded in ISO 3166-1“, ISO, https://en.wikipedia.org/wiki/ISO_3166-2
- [9] „GTFS“, Spezifikation der *gtfs:** tags im OSM-Wiki, <https://wiki.openstreetmap.org/wiki/GTFS>
- [10] Übersicht über die von PTNA analysierten GTFS-feeds in Deutschland, <https://pt-na.openstreetmap.de/gtfs/DE/index.php>
- [11] Statistiken und Logging von PTNA, <https://ptna.openstreetmap.de/en/statistics.php>
- [12] „KeepRight“, https://keepright.at/report_map.php?lang=de
- [13] „PTSA - Public Transport Stop Analysis“, <https://gauss.whz.de/ptsa/#3/50.72/12.50>
- [14] „PTNA: news for Public Transport Network Analysis“, OpenStreetMap Community Forum, <https://c.osm.org/t/ptna-news-for-public-transport-network-analysis/8383/>
- [15] PTNA-Source-Code, <https://github.com/osm-ToniE>
- [16] Lizenz CC0, <https://creativecommons.org/publicdomain/zero/1.0/>
- [17] Lizenz CC-BY-4.0, <https://creativecommons.org/licenses/by/4.0/>
- [18] Overpass-API, https://wiki.openstreetmap.org/wiki/Overpass_API
- [19] Planet-Dumps, <https://planet.openstreetmap.org/pbf/planet-latest.osm.pbf>
- [20] pyosmium-up-to-date, <https://github.com/osmcode/pyosmium>
- [21] Osmium-Filter, Osmium-Extract, <https://github.com/osmcode/osmium-tool>