

Slutrapport för SICS del av projektet Uppföljning och Prediktion

Jan Ekman, Anders Holst
SICS Swedish ICT

5 juni 2015

Innehåll

1	Inledning	4
1.1	Projekttid	4
1.2	Bakgrund till SICS del av projektet UoP	4
1.3	Mål med projektet	5
1.4	Kunskapsuppbyggnad om rörelser hos tåg	6
1.5	Beslutsstöd till den operativa trafikledningen	7
1.6	Arkitektur hos ett verktyg för prediktion av trafikläge	7
1.7	Ett förbättrat utnyttjande av järnvägsinfrastrukturen	9
1.8	Projektering av järnvägsinfrastruktur och operativ drift	9
1.9	Användning av tåglägesprediktion i andra sammanhang	10
1.10	Arbetet i denna del av UoP projektet	11
2	Datakällor	12
2.1	Hittills undersökta datakällor	12
2.2	GPS-data	13
2.3	Några egenskaper att ta hänsyn till hos GPS-data	13
2.4	Värmlandsbanan i GPS-data	15
2.5	Uppdelning av tågrörelser på stationssträckor	15
2.6	Omfattningen av GPS-data	16
2.7	Tågrörelser i GPS-data	19
2.8	Användning av LUPP-data för att skilja på primär och sekundärförseningar	20
2.9	Använda data kan inte skilja på primär och sekundärförseningar	21
2.10	Jämförelse mellan GPS-data och LUPP-data	21
2.11	Jämförelse mellan GPS-data och Trainplan-data	22

3	Prediktion av framtida tågtrafikläge	24
3.1	Inledning	24
3.2	Händelsesystem	24
3.3	Övrig information	25
3.4	Olika typer av begränsningar	25
3.5	Tåginteraktioner och spridning av förseningar	25
3.6	Flera samtidigt tåginteraktioner	26
3.7	Cato	26
3.8	Om att göra relevanta analyser	28
3.9	Grundläggande idéer till trafiklägesprediktion	29
	Referenser	30

1 Inledning

Den här rapporten beskriver SICS del av projektet Uppföljning och prediktion, med Trafikverkets registrerings nummer: TRV 2014/34857 utfört under 2014 och 2015. Projektet stöds av Trafikverket som ett KAJT-projekt. Projektet UoP är ett samarbete mellan SICS Swedish ICT och Blekinge Tekniska Högskola, med Johanna Törnquist på BTH som projektledare. Projektets kontaktpersoner på Trafikverket är Magnus Wahlborg och Robin Edlund.

Projektet skall ses som en fortsättning på projektet *Koi-projektet* med det fulla namnet *Förbättringar av järnvägens informationssystem - högre transportkvalitet genom ökad tillgång till operativa data*. Koi utfördes av SICS under 2010 - 2011 som ett av Trafikverkets *FUD-projekt* med Banverkets registrerings nummer: F09-11550/AL50 och Trafikverkets registrerings nummer: TrV 2010/29759.

1.1 Projektidé

I korthet är projektidén att utnyttja all den data om järnvägstrafik som i allt större omfattning blir tillgänglig för att skapa en tydligare lägesbild och ett bättre beslutsunderlag för tågplaneprocessen i sin helhet. SICS del i projektet UoP fokuserar på prediktioner av framtida läge i trafiken att användas till att ta fram ett stöd för den operativa tågledningen.

1.2 Bakgrund till SICS del av projektet UoP

SICS har i ett tidigare mångårigt projekt åt Banverket tagit fram ett kapacitetsanalysverktyg att användas vid signalprojektering, se [2]. Verktöget består bl.a. av en implementering av alla säkerhetsregler för det svenska signalsystemet ATC2. Verktöget kan presentera alla förekommande konflikter för en sekvens av tågrörelser på en station, även för stora komplicerade stationer med ett stort antal spår. Verktöget beräknar också tiden för att avveckla en sekvens av tågrörelser samt minsta tid för upprepning för en, oändligt upprepade, rörelsesekvens.

Gångtider är en del av beräkningen i verktöget. Dessa beräkningar baseras på data om fordon och infrastruktur. Under projektets gång uppstod frågan om hur dessa *teoretiskt* beräknade gångtider förhåller sig till verkliga gångtider. Jämförelser visade att de verkliga gångtiderna ibland skiljde sig från de teoretiska. Från denna observation uppstod idén att det kanske är bättre att basera gångtidens uppskattningar på historiska data istället för på teoretiska beräkningar. I diskussioner med Banverket om gångtider angav Banverket ett starkt behov av att erhålla bättre information till den operativa driften, bland annat om förväntade ankomsttider. Baserat på detta och tillsammans med Banverket togs idén fram som ligger till grund för projekten Koi och senare UoP

1.3 Mål med projektet

Denna rapport och SICS del av projektet Uppföljning och prediktion handlar om stöd åt den operativa trafikledningen att ta rätt beslut så att, på detta sätt, ett förbättrat utnyttjande av den svenska järnvägsinfrastrukturen erhålles. Det stöd som studeras i rapporten består av information att fatta beslut på bl.a. av prediktioner om trafikläge och förseningar. Rapporten handlar inte om att ta fram förslag på beslut i den operativa driften, enbart om att bedömma olika tänkta förslag med avseende på konsekvenser i form av predicerat tågläge en kort tid framåt i tiden, c:a en timme, risker för följd-förseningar med mera.

Mål och omfång hos projektet Koi beskrivs på följande sätt i [1], slutrapporten till Koi-projektet:

The motive for this project, given in the application of it, is *the great need for available and precise information to support train control and other railway applications*. The project aim is to investigate the possibilities to provide such information and support. An underlying idea is to use GPS data as source data on time, speed and position of trains to provide information such as trains running times, estimated time of arrival, and anomalous train behaviour. During the course of the project it was discovered that GPS data from some trains are accessible and the choice was made to use those data. Moreover ETA, estimated time of arrival, was found to be a kind of information of central value. In this report the more specific aim is therefore to demonstrate the possibilities to make ETA available using GPS data as the basic data source. The ideas and methods, presented in this report, are however not limited to the use of GPS data and not limited to produce the support described in the report. Furthermore, information, such as ETA, which suits the purpose of building support tools for train control are of course not limited to support train control. The information has capability to raise the rail transport quality in many ways, for instance may estimations, if accurate, of the chance of arriving before a given time may be of interest to passengers and freight customers. Trains running times models are of interest for timetable planning and capacity analysis. The project acronym *Koi* standing for *Quality of Operative Information* intends to lead the mind to the importance of information of high quality for improving train control and other railway applications.

Denna beskrivning är gällande för SICS del av UoP-projektet som en fortsättning på Koi-projektet.

1.4 Kunskapsuppbyggnad om rörelser hos tåg

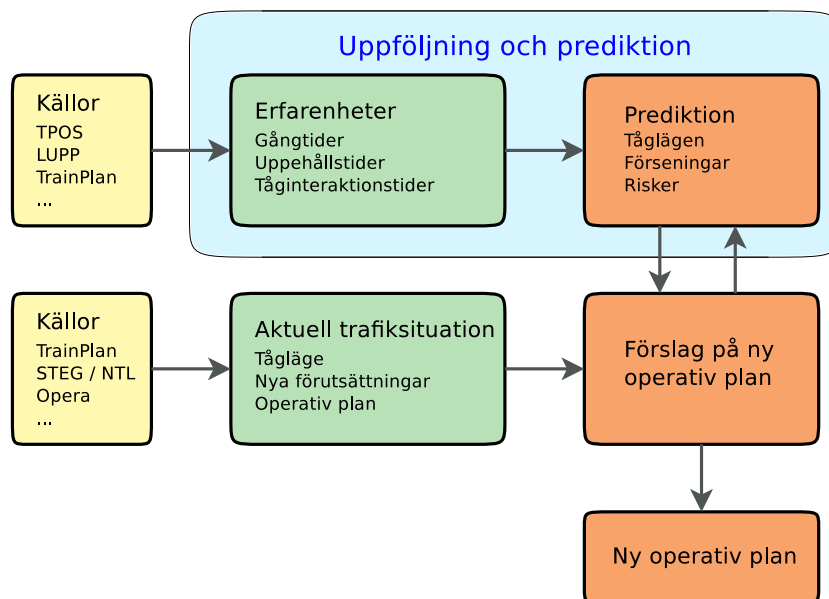
Som beskrivits ovan är idén i SICS del i UoP-projektet att predicera det framtida läget i trafiken och basera ett beslutsstöd till den operativa driften på dessa prediktioner. För att göra prediktioner kommer realtidsdata om det nuvarande tillståndet att användas tillsammans med erfarenheter erhållna från historiska data. På detta sätt används historiska data för att bygga upp kunskapen om hur tåg rör sig under olika förutsättningar. Projektbeskrivningen till projektet UoP säger att en del av SICS arbete i projektet handlar om statistisk modellering. Meningen med de statistiska modellerna är att användas för att svara på t.ex. den förväntade tiden för en rörelse, dess varians och sannoliketen att tiden för rörelsen är en visst satt gräns.

Att SICS arbete i projektet handlar om statistisk modellering skall under inga omständigheter tolkas som att kunskapsuppbyggnad om tågrörelser är en helt igenom automatiserad process där data läses in och statistiska modeller tas fram. En fas i kunskapsuppbyggnad om rörelser hos tåg består av att undersöka data på olika sätt för att få en uppfattning om vad som brukar vara fallet, vilka faktorer som har stor eller liten påverkan på trafiken och hur stark påverkan vissa händelser har på rörelserna. I denna fas av kunskapsuppbyggnaden ingår också att testa olika teorier orsakssamband som kan förklara egenskaper hos rörelserna. Utöver om uppbyggnad av kunskap om rörelserna behövs kunskap om de data och de datakällor som används för prediktionerna av tågrörelser. Denna kunskap behövs för att avläsa data på ett korrekt sätt utan att störas av de fel som är normalt i det data som används.

Undersökningarna av data för att erhålla fakta trafik och orsaker och att få klarhet i vad som gäller datat i sig självt, kan gå till på många olika sätt. Enkla sätt är att direkt titta på det data som finns eller att plotta data på olika sätt. Andra sätt kan vara att använda statistiska metoder för att undersöka om det finns stöd för olika hypoteser, åsikter och teorier. Ett exempel på en åsikt kan vara att försenade tåg har kortare gångtid. Är detta sant för alla tåg? Är det sant på alla sträckor och under alla förhållanden? Hur mycket mindre är gångtiden? Finns det andra faktorer som är viktiga att ta hänsyn till när det gäller gångtider hos försenade tåg? Skall man ta hänsyn till huruvida ett tåg har passagerare som skall vidare med ett annat tåg?

Uppgiften att ta fram passande statistiska fördelningar för tågrörelser, uppehåll m.m. förutsätter en tydlig faktabas om data, trafik och orsaker. På detta sätt kan kunskapsuppbyggnad om rörelser hos tåg betraktas som bestående av följande två faser, som kan itereras:

1. Utforskning av data med lämpliga metoder för att erhålla en grundläggande förståelse om data, trafik och orsakssamband
2. Utnyttjande av de fakta som fas 1 ger till att konstruera modeller för trafiken, t.ex. modeller för gångtider, uppehåll och tåginteraktioner.



Figur 1: Stöd till den operativa trafikledningen enligt UoP-projektet

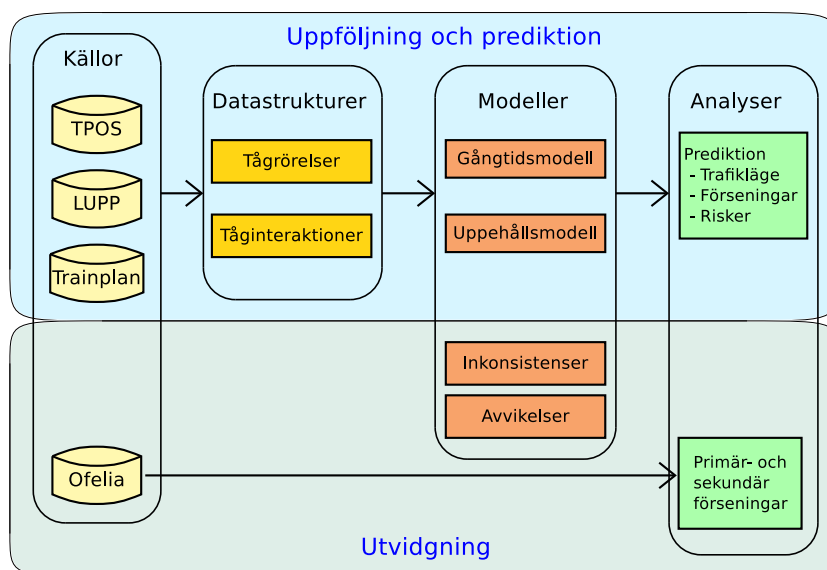
1.5 Beslutsstöd till den operativa trafikledningen

Det är UoP-projektets mål (SICS del) att ta fram information som stöd åt den operativa trafikledningen att ta rätt beslut. Denna rapport ger inte någon detaljerad bild av hur detta kommer att gå till utan bara en översiktlig idé beskriven i figur 1.

1.6 Arkitektur hos ett verktyg för prediktion av trafikläge

Figur 2 visar en bild över dataflöde och arkitektur hos ett verktyg för prediktion av trafikläge. Den övre delen av bilden visar delar som arbetet i UoP-projektet handlar om och till viss del implementerar. Den undre delen visar en tänkbar framtida utvidgning som handlar om att ta hänsyn till infrastruktur fel.

Kortfattat kan arkitekturen beskrivas på följande vis. Data från olika källor läses in och samlas i gemensamma strukturer. *Tågrörelser* består av alla data om tågpositioner, tågpassager, hastigheter både historiska data från faktiska tåg och data från planerade tåg i tidtabellen. De läses in i en enda struktur. Mycket liten behandling av data sker i avdelningen "Datastrukturer" i arkitekturen, men däremot sker en *identifiering* av data från olika källor, så att man vet vilka data från olika källor som hör ihop, som t.ex. att de rör samma tåg. Datastrukturerna har som uppgift att stödja den utläsning av data som de senare stegen



Figur 2: Arkitektur hos ett verktyg för prediktion av trafikläge

behöver. *Tåginteraktioner* består på samma sätt av alla data om interaktioner, beroenden eller konflikter mellan tåg. Detta är data om alla sätt på vilket ett tåg kan påverka ett annat tåg. Det kan vara tågvägskonflikter på stationer eller omstigning av resande på en station. I en utvidgning av verktyget skulle information om bana, signalering och spår användning på stationer kunna användas för att avgöra konflikter mellan rörelser på medelstora och stora stationer.

Baserat på datastrukturerna konstrueras statistiska modeller. Modellerna kan ofta ses som något som uttalar sig om den samlade mängden av data och som är användbart att veta för de resulterande analyserna. Som ett exempel på vad en modell är kan vi ta en modell för uppehåll. Modellen skulle säga att, för en station, vid en given tid på dygnet, för SJ:s reande tåg så är tiden för uppehåll, borträknat väntan på avgångstid enligt tidtabell och borträknat väntan på andra tåg, Weibull fördelad med vissa parametrar. Dessa parametrar som bland annat bestämmer medelvärdet och variansen för uppehållen bestäms på något sätt från indata, kanske helt och hållet från historiska data.

En utvidgning av verktyget är undersökningar av att olika datakällor överstämmer med varandra och att olika delar av en och samma datakälla också överstämmer med varandra, det vill säga att de inte innehåller några inkonsistenser. En annan utvidgning är att värdera hur normala olika värden är i data och rapportera eller ignorera värden som avviker mycket från det som är vanligt. Dessa utvidgning kan också ses som en sorts modeller, eller som baseras på införandet av nya modeller.

Det slutliga utdatat från verktyget sker genom beräkningar där modellerna används. Prediktionen av ett framtida tågläge baseras på tider för gång och uppehåll och på vilka tåg det är i konflikt med.

1.7 Ett förbättrat utnyttjande av järnvägsinfrastrukturen

Som nämns ovan avsnittet 1.3 *Mål med projektet* handlar SICS del av projektet om stöd åt den operativa trafikledningen att ta rätt beslut så att, på detta sätt, ett förbättrat utnyttjande av den svenska järnvägsinfrastrukturen erhålles. Idag existerar dock inte någon uppföljning av hur effektiv den operativa trafikledningen är. Effekten av att använda nya metoder inom trafikledningen kan därför vara svåra att mäta. Möjligen kan de resultera i förändringar i förseningarna i trafiken. Men det är inte säkert att analyser av förseningar är ett bra sätt att bedömma effektivitet hos den operativa driften. För det första kan en ökning eller minskning av förseningarna ha andra orsaker än effektivitet hos den operativa driften. För det andra kan det vara så att tidtabellen är konstruerad för att klara en viss ineffektivitet hos den operativa driften, så att en förbättring inte får någon effekt.

Istället för att studera förseningar, med koppling till tidtabellen och alla de frågor som gäller hur förseningar skall mätas, skulle trafikledningsbeslut kunna återskapas från alla de data som genereras från trafiken. Detta kan vara en möjlighet att statistiskt verifiera att en nya metoder ger möjligheter till ett bättre utnyttjande av infrastrukturen, t.ex. genom ett annat sätt att hantera bufferttider och på detta sätt öka kapaciteten.

1.8 Projektering av järnvägsinfrastruktur och operativ drift

Vid ny- och ombyggnad av järnväg genomför signalprojektörer noggranna analyser av kapacitet för olika infrastruktureförslag. De väsentliga egenskaperna hos de olika förslagen är vilka rörelser som kan ske konfliktfritt och vilken tid olika rörelser tar. Denna information är känd och ligger, tillsammans med kostnaderna, till grund för besluten om infrastrukturutbyggnaden. Rörelsekonflikter och tider för rörelser är viktig information för den operativa driften. Särskilt vid framtagande av avancerade beslutsstödssystem till operativ drift är denna information central. Byggnadsfasen av järnvägen bör därför göra sådan information tillgänglig. Järnvägen som levereras från byggnadsfasen bör inte ses som ett nät av spår, signaler och växlar utan som en möjlighet att framföra trafik. Ny- och ombyggnad av järnväg bör levereras som en handhavandebeskrivning av den levererade järnvägen till den operativa driften.

1.9 Användning av tåglägesprediktion i andra sammanhang

Såsom diskuteras i [1], skulle metoderna för tåglägesprediktion i SICS del av UoP-projektet kunna användas i många sammanhang. Låt oss här betrakta ett par exempel pågående projekt där metoderna skulle kunna användas.

Projektet FLOAT, FLeXibel Operativ omplanering Av Tåglägen, utförs av BTH på uppdrag av Trafikverket som ett KAJT-projekt. Det handlar om beslutstöd för trafikledarna. Rapporten [3] innehåller följande beskrivning av FLOAT

“De nämnda stöden [STEG/CATO] inkluderar dock inte funktioner som detekterar alla typer av relevanta potentiella konflikter i föreslagna tågplan, utreder konflikternas inbördes beroenden eller ger (del)förslag på hur konflikterna kan lösas och vad effekterna skulle bli i form av merförseningar, spårbyten, missade anslutningar, mm. Det är denna typ av funktioner som istället de tidigare projekten OAT, OAT+ och EOT har fokuserat på.”

och

“I FLOAT avser vi ta nästa steg och studera hur dessa metoder skulle fungera i ett mer praktiskt sammanhang och anpassa dem därefter.”

Precis som SICS del av UoP-projektet siktar FLOAT-projektet mot stöd till den operativa driften. Delar av UoP-projektets tåglägesprediktion, t.ex. gångtidsmodellerna, bör vara användbara i beräkningar av de effekter av potentiella konflikter som studeras i FLOAT-projektet.

Projektet PUMPS, PUnktlighet genom MålPunktsStyrning, utförs av SICS på uppdrag av Trafikverket också som ett KAJT-projekt. Rapporten [3] innehåller följande beskrivning av PUMPS.

“Utgående från erfarenheterna på Malmbanan ska detta projekt analysera potentialen av införande av CATO-liknande system mer generellt i Sverige. Projektet avser att undersöka både energi-, punktlighets- och kapacitetsmässiga aspekter.”

“Avsikten är att utvärdera målpunktsstyrning och energioptimering främst genom analyser, beräkningar och simuleringar.”

I dessa analyser, beräkningar och simuleringar skulle också delar av UoP-projektets tåglägesprediktion vara användbara.

1.10 Arbetet i denna del av UoP projektet

En av skillnaderna mellan SICS del av UoP-projektet och det avslutade projektet Koi är att UoP-projektet undersöker möjligheterna att samtidigt använda data från flera olika källor för att göra prediktioner. I Koi-projektet användes enbart GPS data från tåg. I arbetet i denna del av UoP projektet ingår att undersöka hur GPS-data och data från LUPP och Trainplan kan användas. Datastrukturer för att kunna läsa in, modellera och kombinera dessa data har tagits fram i projektet. Vissa inläsnings- och analysfunktioner har implementerats. Inledande undersökningar av egenskaper hos data har gjorts. Kapitel 2 i denna rapport beskriver resultaten från dessa inledande undersökningar. I projektet ingår diskussioner och utredningar kring metoder att göra prediktioner. Kapitel 3 sammanfattar resultaten från dessa utredningar.

Ett av resultaten från projektet är att det, för det område som undersöktes, finns relativt få tågrörelser i GPS-data av det totala antalet tågrörelser. GPS-data innehöll inte heller några resandetåg. Detta betyder att vissa undersökningar inte kan genomföras som tänkt med tillhandahållen GPS-data som huvudsaklig datakälla.

I projektet ingår en vidareutveckling av det analys och demonstrationsverktyg som togs fram i projektet Koi. Vidareutveckling är ett första steg mot en implementering av de nya typerna av prediktioner som UoP-projektet föreslår. I projektet ingår några presentationer av projektet, en presentation på ett möte arrangerat av Transrail, en presentation på Katjmöte i Borlänge och slutligen en slutpresentation på Trafikverket i Solna. Till arbetet i projektet hör en forskningsöversikt som presenteras i en separat rapport med namnet *UoP: Forskningöversikt*.

2 Datakällor

Avsikten med detta kapitel är att ge en översiktlig bild över de datakällor vi har undersökt i projektet. Omfattning och egenskaper hos data är något av det som kommer att behandlas.

2.1 Hittills undersökta datakällor

Alla sorters data om något som kan ha påverkan på trafiken är relevanta för uppgiften att predicera ett framtida läge i trafiken. I denna första fas av projektet har vi dock begränsat oss till fåtal datakällor. Längre fram i rapporten kommer vi att i olika sammanhang resonera om hur andra datakällor skulle kunna användas. Följande data har studerats i SICS del av UoP-projektet:

1. GPS-data från TPOS med tågpositioner för hela Sverige, 1 augusti till 8 september 2014. Drygt 4 miljoner datapunkter. Cirka 60 000 datapunkter från Värmlandsbanan (Laxå - Charlottenberg).
2. TrainPlan data exporterad från TrainPlan för samma period som GPS-data
3. LUPP-data med ankomster och avgångar den 27 augusti 2014 för stationer på Värmlandsbanan och mellan Laxå och Hallsberg.
4. Tågnummerförteckning
5. GPS-positioner för stationer från Eniros kartor

Eniros kartor finns på:

<http://kartor.eniro.se/>

och trafikverkets tågnummerförteckning är följande fil:

http://www.trafikverket.se/contentassets/dfcee253588e47efb036cdc427a6607e/tagnummerforteckning_t14_trv_130409.xls

Tågn:r	Datum	Klockslag	GPS-norr	GPS-öst	Hast.
19423	2014-08-04	10:09:08.460	6583086.993	414424.901	0
19423	2014-08-04	10:39:58.657	6583117.921	414425.601	0
19423	2014-08-04	10:41:18.643	6583086.993	414424.901	0
19423	2014-08-04	10:57:58.633	6583117.921	414425.601	0
19423	2014-08-04	10:58:18.603	6583086.993	414424.901	0
19423	2014-08-04	10:59:38.670	6583117.921	414425.601	0
19423	2014-08-04	11:00:08.560	6583086.993	414424.901	0
19423	2014-08-04	11:05:38.730	6583117.921	414425.601	3
19423	2014-08-04	11:07:08.547	6583086.993	414424.901	0
19423	2014-08-04	11:47:48.750	6583087.351	414409.12	6
19423	2014-08-04	11:47:58.687	6583118.635	414394.039	7
19423	2014-08-04	11:48:08.640	6583119.35	414362.477	8

Tabell 1: Exempel på GPS-data

2.2 GPS-data

Utgångspunkten för projektet är att GPS-data är källan till historiska data om tågens rörelser. Om det visar sig att GPS-data inte är tillgänglig i tillräckligt hög grad kan denna ståndpunkt komma att omprövas. Alternativet är att GPS-data tillsammans med andra data, i första hand data som genereras från spårledningspassager, utgör den bas av historiska tågpositioner som kunskapen om tågens rörelser vilar på. GPS-data kan också vara huvudkällan till realtidsdata om tågpositioner i det system som projektet har som mål att i förlängningen ta fram. GPS-data är däremot inte den enda källan till kunskap om varför tågen betar sig som de gör och därför är de inte heller den enda källan till de statistiska gångtidsmodellerna.

2.3 Några egenskaper att ta hänsyn till hos GPS-data

GPS-data består av datapunkter på formen: tågnummer, tid, GPS-koordinater och hastighet. I tabell 1 ser vi ett exempel på GPS-datapunkter, med en datapunkt per rad. Som en första inblick i naturen hos dessa data gör vi följande observationer:

1. Det är ojämna tidsintervall mellan data punkterna
2. Under den period då hastigheten är noll hoppar GPS-datapunkterna mellan två positioner: (6583086.993, 414424.901) och (6583117.921, 414425.601)
3. Hastigheten i en av datapunkterna är 3, trots att tågets position är en av de hoppande nollhastighetspositionerna och densamma som två datapunkter tidigare.

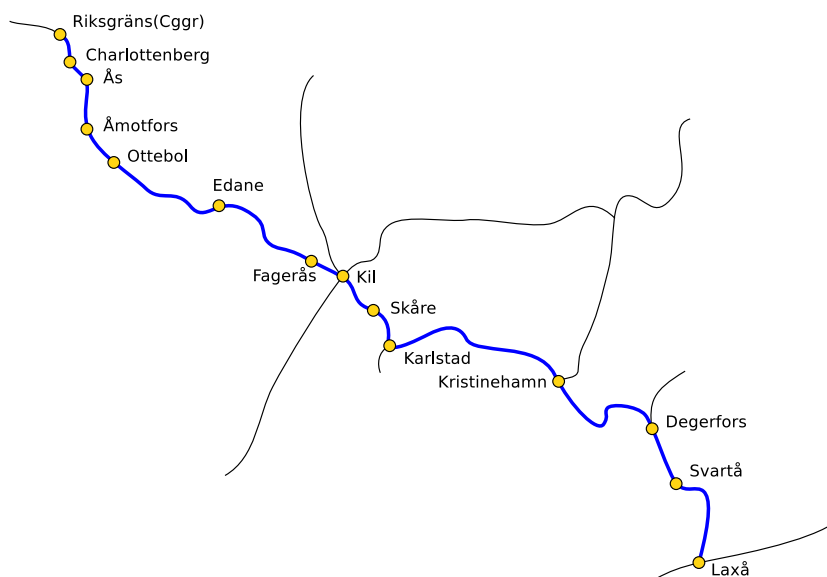
Tågn:r	Datum	Klockslag	GPS-norr	GPS-öst	Hast.
4014	2014-08-13	02:03:09.617	6628365.359	352048.983	0
4014	2014-08-13	02:03:29.290	6628396.278	352050.212	0
4014	2014-08-13	02:03:49.090	6628365.359	352048.983	0
4014	2014-08-13	02:04:19.073	6628396.278	352050.212	0
4014	2014-08-13	02:04:49.153	6628365.359	352048.983	0
4014	2014-08-13	02:05:09.233	6628396.278	352050.212	506
4014	2014-08-13	02:05:31.107	6628365.359	352048.983	0
4014	2014-08-13	02:05:39.373	6628396.278	352050.212	0
4014	2014-08-13	02:06:09.093	6628365.359	352048.983	0
4014	2014-08-13	02:06:19.033	6628365.979	352033.388	0
4014	2014-08-13	02:06:29.047	6628365.359	352048.983	0
4014	2014-08-13	02:06:59.017	6628396.278	352050.212	0

Tabell 2: Inkorrekt hastighet GPS-data

- I en datapunkt är hastigheten 6, drygt 40 minuter efter föregående datapunkt. Positionen hos datapunkten med hastigheten 6 är nära positionen hos föregående datapunkt. I den föregående datapunkten är hastigheten noll.

Vi tolkar det som att tåget står stilla så längedatapunkterna hoppar mellan de två nollhastighetspositionerna i punkt 2. Det är generellt så att GPS-datapunkterna inte är konstanta för stillastående tåg. Ofta hoppar de mellan två olika positioner, men ibland kan de hoppa mellan fler än två. Vi tolkar det också som att tåget står stilla för datapunkten med hastigheten 3, eftersom GPS-positionen är bland de hoppande positionerna för hastigheten noll. Det ser ut att vara ovanligt med helt och hållet inkorrekta värden, men de förekommer. Tabell 2 visar ett exempel på ett felaktigt värde. Här erhåller ett stillastående tåg för en enstaka datapunkt ett högt värde. Trots att datapunkten med hastigheten 6 i tabell 1 inträffar 40 minuter efter föregående datapunkt, tolkar vi denna punkt som tågets start. Orsaken till det är att hastigheten är stor nog och att dess position är nära positionen för stoppet, men inte bland de bland de hoppande positionerna för hastigheten noll.

Dessa observationer och tolkningar säger oss något om vilka avvikelser vi kan förvänta oss att hitta i data och som vi behöver ta hänsyn till vid den statistiska modelleringen.



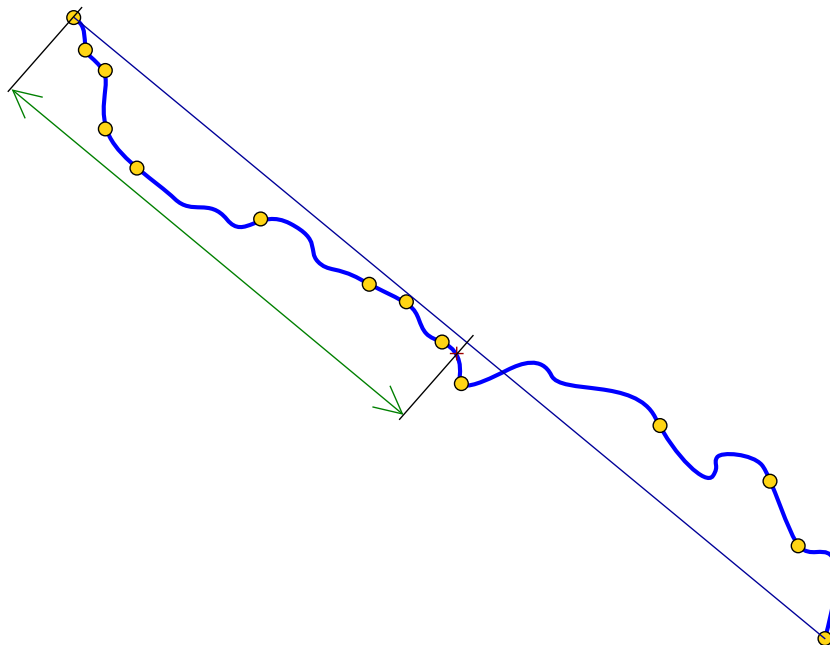
Figur 3: Värmlandsbanan

2.4 Värmlandsbanan i GPS-data

I projektet analyseras GPS-data för Värmlandsbanan, banan mellan Laxå och riksgränsen strax väster om Charlotten berg. Figur 3 visar Värmlandsbanan med de stationer där tåg startar eller slutar för de GPS-data som använts. Att ett tåg startar eller slutar på en station betyder en av två saker (1) att tåget har en hastighet nära noll på stationen (2) att stationen är den första eller sista som tåget passerar i GPS-datat.

2.5 Uppdelning av tågrörelser på stationssträckor

För att dela upp rörelserna givna av GPS-data på stationssträckor, mellan de valda stationerna, har först en grov filtrering gjorts för att plocka ut datapunkter nära Värmlandsbanan och därefter har positionerna projicerats på den raka linjen mellan Laxå och Riksgränsen som figur 4 visar. GPS-data, SWEREF 99, för samtliga stationer på Värmlandsbanan har hämtats från Eniro. Också stationernas positioner har projicerats på linjen mellan Laxå och Riksgränsen. På så sätt kan stationssträckan bestämmas för varje datapunkt i GPS-data. Eller iallafall nästan. Några tåg på anslutande linjer som kommit med i den första grova filtrering behöver plockas bort och projiceringen fungerar inte som tänkt kring stationen Björneborg, se figur 5. Det visar sig dock att Björneborg inte är start eller slutstation till något tåg, så därför bryr vi oss inte, för tillfället,



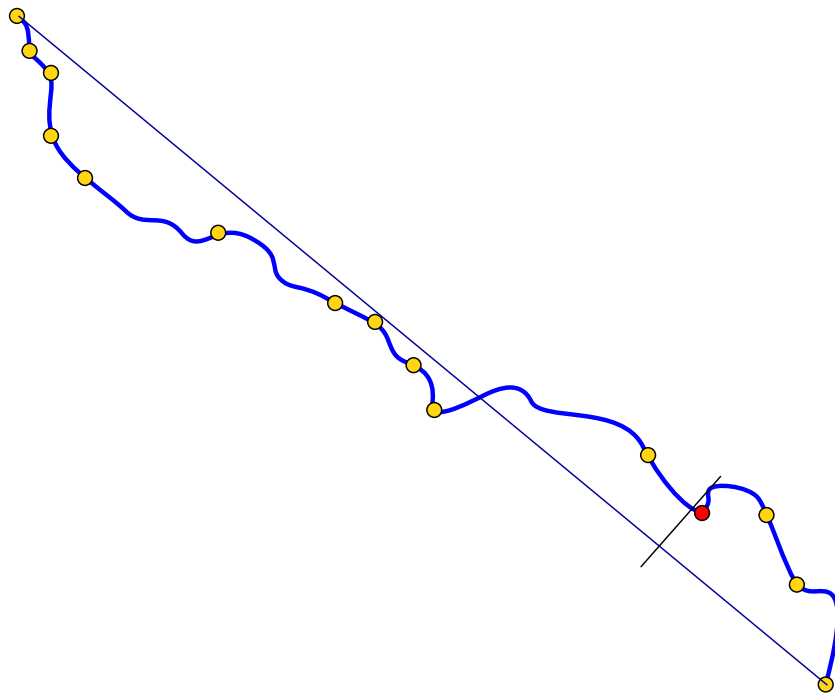
Figur 4: Projektion av position på rät linje mellan Värmlandsbanans ändpunkter

om att projiceringen falerar bara just där. Felet kan enkelt åtgärdas genom en specialprojektion för Björneborg med omnejd.

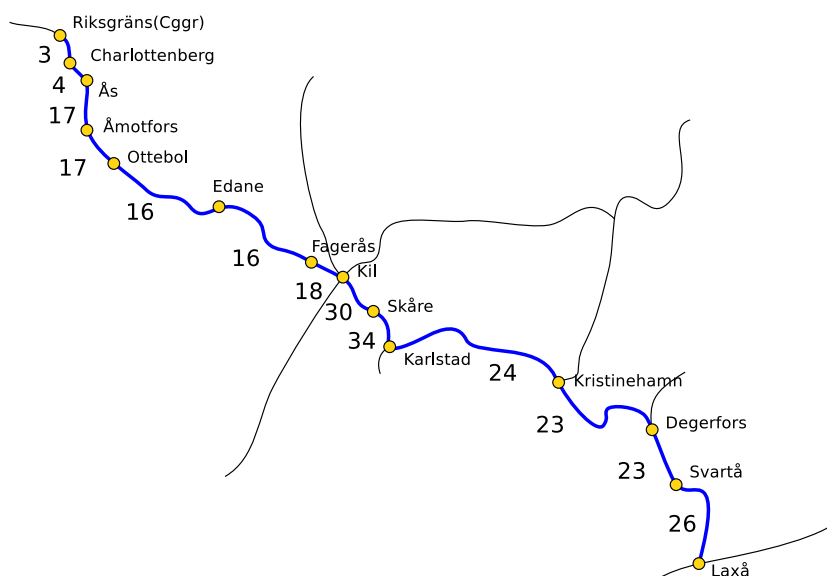
Det finns naturligtvis många sätt att bestämma GPS-datapunkternas positioner i järnvägsnätet. Ett annat sätt är att hitta en avbildning från GPS-datapunkter till punkter på spåret i BIS-data. Från BIS-data skulle man då också veta avståndet till stationerna i ändpunkterna på den stationssträcka från vilken GPS-datapunkten är hämtad. Att konstruera en sådan avbildning innebär dock betydligt mycket mer arbete än den enkla projektion som används här.

2.6 Omfattningen av GPS-data

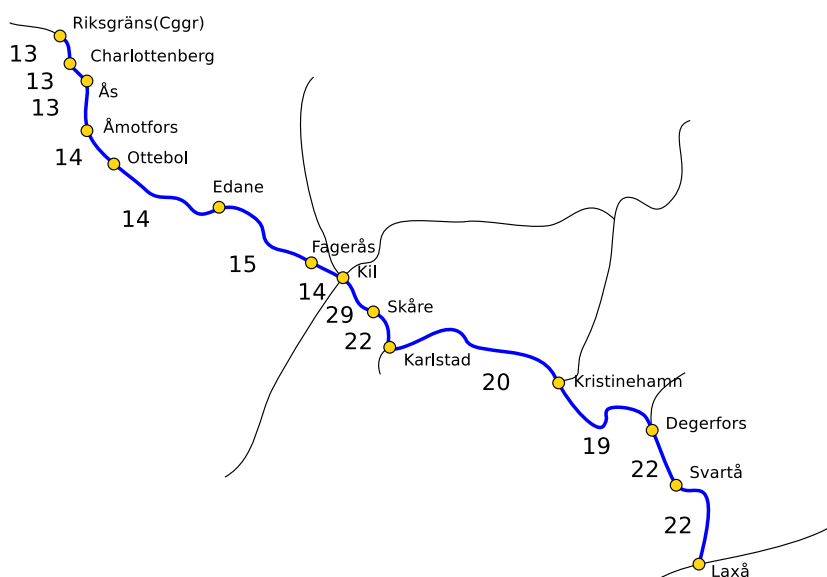
Det totala antalet tågrörelser i GPS-data på värmlandsbanan är 83. Då har vi inskränkt oss till att ta med tåg som tar sig hela vägen mellan två stationer. Att *ta sig hela vägen* mellan två stationer betyder här att tåget för båda stationerna antingen passerar stationen eller har en hastighet av högst 5 km/h i närheten av stationen. Dessa 83 tågrörelser är enbart godståg, vilket framgår av tågnumren i GPS-data tillsammans med tågnummerförteckning. De flesta är Green Cargo tåg, knappt en tredje del är Hecto Rail tåg och några få tåg anges som "TrV Södra Valfritt Adhoc" i tågnummerförteckningen.



Figur 5: Ett område där projektionen på den räta linjen inte är en ett-ett avbildning



Figur 6: Antal norrgående tågrörelser i GPS-data på Värmlandsbanan under perioden 1 augusti till 8 september 2014



Figur 7: Antal södergående tågrörelser i GPS-data på Värmlandsbanan under perioden 1 augusti till 8 september 2014

Tågn:r	Datum	Tid	GPS-norr	GPS-öst	H.	Plats
21645	2014-08-01	07:42:23	6646297	346684	89	Cg-Cggr
21645	2014-08-01	07:42:33	6646460	346504	89	Cg-Cggr
21645	2014-08-01	07:42:43	6646653	346326	89	Cg-Cggr
21645	2014-08-01	07:42:53	6646815	346162	89	Cg-Cggr
21645	2014-08-01	07:43:03	6647008	345983	89	Cg-Cggr
21645	2014-08-04	07:42:33	6597469	404431	0	Kil
21645	2014-08-13	07:19:05	6554876	472306	0	Svartå
21645	2014-08-13	07:19:15	6554876	472290	0	Svartå
21645	2014-08-13	07:19:35	6554876	472306	0	Svartå
21645	2014-08-13	07:19:45	6554876	472290	0	Svartå
21645	2014-08-13	07:20:05	6554876	472306	0	Svartå
21645	2014-08-13	07:20:45	6554876	472290	0	Svartå

Tabell 3: En enda datapunkt för ett tåg en dag

För att få en uppskattning av hur många av det totala antalet tågrörelser som finns i GPS-data räknar vi antalet rörelser per sträcka i GPS-data. Figurerna 6 och 7 visar antalet norrgående respektive antal södergående tågrörelser i GPS-data på Värmlandsbanan. Sträckan Skåre - Kil har flest tågrörelser, 59 tågrörelser totalt där 30 är norrgående och 29 är södergående. Låt oss jämföra detta med antal tågrörelser LUPP. Under en dag i Skåre, den 27 augusti 2014, är antalet ankomster plus antalet avgångar i LUPP-data 140 stycken. Det vill säga, 70 tåg bör ha passerat Skåre denna dag. Det finns inte spår som ansluter till Värmlandsbanan i Skåre, så alla dessa tåg måste vara tåg på Värmlandsbanan. Om vi antar att det har gått lika många tåg varje dag så får vi att det har gått 2730 under de 39 dagarna från 1 aug. till 8 sept. 2014. Under dessa antaganden innehåller GPS-data c:a 2% av alla tågrörelser. Detta ger en grov uppskattning av omfattningen av GPS-data.

2.7 Tågrörelser i GPS-data

Många av tågrörelserna i det urval av GPS-data som gäller Värmlandsbanan startar och slutar med antingen att tåget står still på station eller passerar gränsen till Värmlandsbanan. De data som presenteras i tabell 3 uppfyller visserligen detta, men här finns en märklig ensam data punkt en dag. Den första augusti lämnar ett tåg med tågnummer Sverige vid riksgränsen mot Norge. Tre dagar senare finns en enda datapunkt som säger att ett tåg med samma tågnummer står stilla i Kil. Det är möjligt att det är tre olika fysiska tåg med samma tågnummer som finns trepresterade i tabellen.

Tabell 4 ger ett exempel på ett tåg där GPS-data försvinner från ett som inte står still och inte lämnar Värmlandsbanan. I det ögonblick då GPS-data försvinner

Tågn:r	Datum	Tid	GPS-norr	GPS-öst	Hast.
21644	2014-08-04	10:58:52.653	6619560.625	359526.461	100
21644	2014-08-04	10:59:02.683	6619632.472	359263.063	100
21644	2014-08-04	10:59:12.683	6619703.747	359015.305	101
21644	2014-08-04	10:59:22.700	6619777.405	358705.035	101
21644	2014-08-04	10:59:32.717	6619848.125	358472.926	99
21644	2014-08-04	10:59:42.700	6619952.154	358179.481	101
21644	2014-08-04	10:59:52.733	6620085.34	357934.116	102
21644	2014-08-04	11:00:02.700	6620218.543	357688.767	102
21644	2014-08-13	08:39:35.320	6567040.193	467386.594	76
21644	2014-08-13	08:39:45.557	6567257.119	467340.913	76
21644	2014-08-13	08:39:55.523	6567474.182	467279.391	76
21644	2014-08-13	08:40:05.650	6567629.93	467153.967	80

Tabell 4: GPS-data försvinner från tåg i Ottebol

är tåget på väg in mot Ottebol från sydöst.

2.8 Användning av LUPP-data för att skilja på primär och sekundärförseningar

Vid framtagning och användning av gångtidsmodeller behöver man ta hänsyn till om tågrörelserna störs av annan trafik. Eftersom GPS-data endast innehåller en liten del av tågrörelserna kan man inte från dessa data avgöra om ett tåg kan vara stört av annan trafik på banan.

Ett sätt att försöka separera störda från ostörda tåg i GPS-data är att utnyttja andra datakällor med historiska data om trafiken. En sådan datakälla är LUPP, som står för *leverans uppföljning*. LUPP innehåller data från bl.a. följande källsystem ATL, Basun, Här & Nu, Opera, DPP, Ofelia, Bessy och Rufus. När det gäller trafikuppföljning är orsaksrapporteringen en av huvuduppgifterna för LUPP, det vill säga rapportering om punktlighet och förseningar. Trafikverket beskriver LUPP på följande sätt

Lupp sammanställer trafik- och anläggningsinformation från andra källsystem i ett datalager. Detta möjliggör uppföljning av anläggningen och dess trafik utifrån ett helhetsperspektiv.

Utgående från den uppgift LUPP har är det inte säkert att all trafik finns representerad i LUPP. Växlingsrörelser, vagnuttagningar och spärrfärder kanske är nog inte så intressanta från ett punktlighets och förseningsperspektiv och finns

Tågn:r	Datum	Tid	GPS-norr	GPS-öst	Hast.
9011	2014-08-27	00:11:01.127	6583152.422	414268.491	0
9011	2014-08-27	00:11:21.143	6583121.495	414267.79	0
9011	2014-08-27	00:11:51.207	6583152.422	414268.491	0
9011	2014-08-27	00:12:42.007	6583152.78	414252.711	2
9011	2014-08-27	00:13:01.163	6583153.138	414236.93	7
9011	2014-08-27	00:13:11.163	6583153.854	414205.368	9
9011	2014-08-27	00:13:21.213	6583185.139	414190.289	11
9011	2014-08-27	00:13:31.227	6583216.782	414159.429	13
9011	2014-08-27	00:13:41.213	6583217.499	414127.868	15
9011	2014-08-27	00:13:51.213	6583249.502	414081.228	16
9011	2014-08-27	00:14:01.120	6583281.146	414050.37	16
9011	2014-08-27	00:14:11.230	6583312.791	414019.512	17

Tabell 5: 00:14 den 27 aug. avgick tåg 9011 från Karlstad

nog därför inte i LUPP-data. Förhoppningsvis orsakas ändå få följdförseningar av trafik som inte finns i LUPP och då är LUPP potentiellt sett en bra källa till all den trafik som kan störa trafiken i GPS-data.

2.9 Använda data kan inte skilja på primär och sekundärförseningar

Det går inte att undersöka om det är möjligt att skilja på primär och sekundärförseningar i GPS-data genom att använda LUPP, data. Bara en enda tågrörelse finns i både GPS-data och LUPP data, tåg 9011 som avgick från Karlstad kl 00:14 den 27 augusti 2014. Detta tåg avgår 9 minuter sent från Karlstad, på grund av att det är sent från depån.

2.10 Jämförelse mellan GPS-data och LUPP-data

Låt oss jämföra tiderna i GPS-data och LUPP-data, för det tåg som finns i båda dessa källor för att se hur väl de överensstämmer.

Enligt LUPP-data avgick tåg 9011 från Karlstad 00:14 den 27 augusti. I GPS-data, se tabell 5, kan vi se att tåget börjar röra sig ut från stationen mellan kl. 00:11:51 och 00:12:42. Enligt LUPP-data avgår tåget från Skåre 00:21. Enligt GPS-data passeras en *för hand på kartan vald* mittpunkt på Skåre station mellan 00:21:21 och 00:21:31 och enligt LUPP-data avgår tåget från Skåre 00:21. Slutligen saktar tåget ned till en lägsta hastighet på 3 km/h på Kils station kl

	GPS	TrainPlan				
		sum	duration	performance	adjustment	recovery
Ks	0	0	0			
Ks1		226	136	90		
Skr	519	505	415	90		
SkrL3		694	544	90		60
Kud		911	740	90	21	60
Sån		1060	868	90	42	60
Kil	1200	1260	1000	90	50	120

Tabell 6: Jämförelse mellan gångtider i GPS-data och Trainplan-data

00:33:01, därefter ökar det hastigheten och försvinner bort från Värmlandsbanan, enligt GPS-data. Enligt LUPP-data avgår tåget från Kil 00:33

Vi tycks alltså ha en rätt så god överensstämmelse mellan GPS-data och LUPP-data för dessa enda tre jämförelser som går att göra med använda data. För avgången från Karlstad framgår det inte från direkt data hur LUPP bestämmer avgångstiden. I själva verket är tiderna LUPP-data beräknade tider. I Slutrapporten från projektet FUKS [3] finns en beskrivning av hur tiderna i LUPP genereras. På frågan om varför tiden för en avgång i LUPP-data från en station kan ske före tiden för ankomst till samma station har följande svar erhållits från LUPP förvaltningen:

”Tidsavvikelse och utförd tid i Lupp är inte beräknat i Lupp utan tas från källsystemet ATL som i sin tur matas av TPOS-systemet. I ATL finns justeringspunkter trafikplatser som lägger på sekunder i vissa fall och drar av sekunder i vissa fall för att kompensera mätutrustningens placering. Tidsrapporterna trunckeras sedan på sekunder ned till minuter vilket kan ge en olycklig effekt i ett fåtal fall. Exempelvis om justeringstiden är anpassad för att ett tåg skall bromsa in och stanna men tåget istället kan gå i full fart förbi den dagen. Ankomsten kan då få ett påslag av sekunder medan avgången får ett avdrag. Om det hamnar i övergången mellan minuter kan då detta fenomen uppstå.”

2.11 Jämförelse mellan GPS-data och Trainplan-data

Låt oss som en jämförelse mellan Trainplan-data och GPS-data åter betrakta tåget som avgår från Karlstad den fjärde augusti c:a kl. 11:48, se tabell 1. Tabell 6 är en jämförelse mellan gångtiden enligt GPS-data och Trainplan-data för detta tåg. I Trainplan är gångtider givna för var och en av de sträckor som

betraktas. Tiderna i tabellen är de ackumulerade tiderna från Karlstad (Ks). Som tabellen visar är Trainplans gångtider i detta fall uppdelade på *duration*, *performance*, *adjustment* och *recovery*. Den totala tiden är summan av dessa tider. Det finns även fler typer av tider än dessa i Trainplan. Vi observerar att gångtiden för tåget i Trainplan är 11 minuter och att gångtiden i GPS-data är exakt 10 minuter. Tåget var knappt en minut sent när det lämnade Karlstad.

3 Prediktion av framtida tågtrafikläge

3.1 Inledning

Som nämns i avsnittet 1.1 handlar projektet om att använda sig av all den data som idag finns tillgänglig och som rör tågtrafik för att få kunskap om hur tåg beter sig i verkligheten och att använda denna kunskap för att kunna ta bättre beslut i den operativa driften. Den mer specifika uppgiften för projektet är att göra en bra prediktion av trafikläget c:a 1 timme framåt i tiden, baserat på erfarenheter från historiska data, informationen om det aktuella trafikläget och för en given plan. Om detta är möjligt kan olika omplaneringslösningar utvärderas och en av dem väljas. Vid utvärderingen av en plan tar man också hänsyn till osäkerheter och risker och det är en del av prediktionen att uppskatta dessa. För att få en bra prediktion, med så små osäkerheter som möjligt, behöver vi så långt det går förklara varför tågen beter sig som de gör. För att använda data på bästa sätt är det också viktigt att tolka detta data på rätt sätt och förstå hur datat genererats från olika informationssystem. Detta kapitel rör prediktioner av framtida tågtrafikläge i allmänhet.

3.2 Händelsesystem

Järnvägstrafiksystemet, det vill säga alla rörelserna hos fordonen på järnvägsinfrastrukturen, kan till en del beskrivas som ett diskret händelsesystem (eng. discrete event dynamic system). Händelser i detta system är att tåg ankommer till eller avgår från stationer, att fronten hos ett tågfordon belägger en spårledning och så vidare. För vissa par av händelser finns ett beroende, så att den ena händelsen *a* har företräde framför den andra *b*, vilket betyder att för någon tid t så kan *b* inträffa tidigast tiden t efter att *a* inträffat. Eventuellt kan man tänka sig att tiden t är negativ så att *b* inträffar tidigast tiden $-t$ innan *a* inträffar. Exempel på händelser som är beroende är att ett tåg A i sin helhet måste passera en viss spårskarv innan ett annat tåg B kan få körbesked i kommande huvudsignal så att detta tåg B kan upphöra med sin inbromsning mot huvudsignalen.

Att se järnvägstrafiksystemet som ett diskret händelse system är viktigt för att dela upp problemet att predicera framtida tågtrafikläge i delar och för att kunna resonera om hur förseningar sprids. Vilka händelserna är i systemet är fixerat. De beroenden som finns mellan händelserna bestäms av planen och ändras vid en omplanering. Prediktionerna handlar främst om när händelser i händelsesystemet inträffar

En del av prediktionen är uppskattningar av tider mellan händelser i händelsesystemet. Generellt sett kommer vi att betrakta dessa tider, t.ex. uppehåll

vid station, som stokastiska variabler. Vi använder statistiska modeller för att beskriva dessa tider.

3.3 Övrig information

All den information vi har om trafiken är inte del i det diskret händelsesystemet. Fordonens positioner i ett visst givet ögonblick, fordonens egenskaper, väder och så vidare är inte del av händelsesystemet. Förutom att predicera händelser i händelsesystemet kan det vara intressant att predicera enskilda tågs positioner efter en viss tid. Sådana prediktioner görs genom att utgå från prediktionerna av händelserna i händelsesystemet. De statistiska modellerna för gångtider som kommer att användas och har använts i Koi-projektet uttalar sig om gångtid mellan godtyckliga positioner och begränsas inte enbart till att handla om de punkter i nätet som är förknippade med händelser i händelsesystemet.

3.4 Olika typer av begränsningar

Prediktionen av det framtida läget i trafiken kan ses som en konsekvens av flera olika begränsningar för trafiken där den största begränsningen avgör. En uppdelning av problemet att predicera framtida tågtrafikläge i delar erhålles genom att separat betrakta följande typer av begränsningar.

1. Gångtider för tåg som inte störs av annan trafik och uppehåll
2. Tåginteraktioner
3. Tidtabell

Exempelvis kan det vara så att ett persontåg inte kan lämna en station föränn tillräcklig tid för avstigande har passerat, inte heller före resande från ett annat tåg har gått på tåget och inte före angiven tid i tidtabellen. För att denna uppdelning skall bli meningsfull behöver vi kunna referera till den del av ett stopp på en station som inte har med tidtabell eller interaktioner andra tåg att göra. Låt oss därför använda ordet uppehåll till detta.

3.5 Tåginteraktioner och spridning av förseningar

För modellering av gångtider för ostörda tåg och för modellering av uppehåll behövs inte någon data om signalsystemet. Inte heller behövs data om spår användning på stationer. Möjligen kan isoleringsskarvarnas positioner behövas för

gångtidsmodellerna om dessa bygger på data från signalsystemet. Rörelsekonflikter kontrolleras av signalsystemet och är därmed den typ av tåginteraktion för vilken signalsystemets funktion är väsentlig. En noggrann beskrivning av signaleringen för en station, även för en mindre station, kan dock kräva ett stort arbete. Den information som behövs är huvudsakligen vilka rörelser som är i konflikt, var de är och var signaler är placerade. För att använda en sådan noggrann beskrivning av signaleringen för avgöra hur två tåg är relaterade till varandra behövs dessutom, på större stationer, data om spår användning.

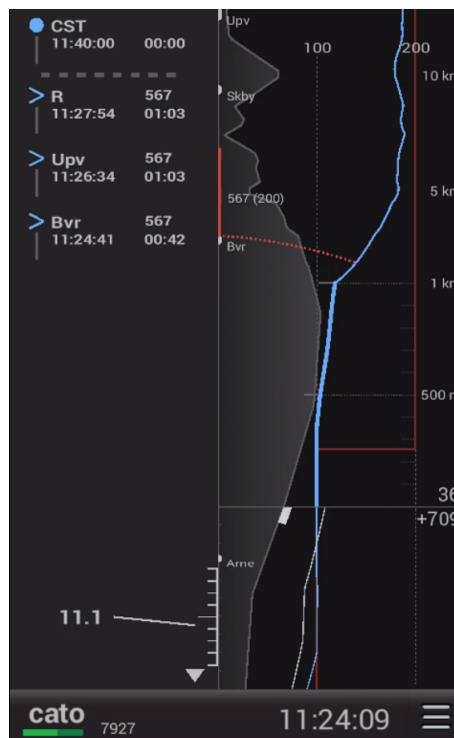
Tåg som kommer i lagom tid fram till att få tillträde till spåret är de som tar sig snabbast fram. Tåg som kommer för tidigt för att få tillträde till spåret tvingas bromsa och förlorar tid i förhållande till de tåg som kommer lite senare. Vill man ta hänsyn till dessa effekter vid modellering av spridning av förseningar behövs kunskap om signalsystemets påverkan på rörelserna.

3.6 Flera samtidiga tåginteraktioner

När det gäller tåginteraktioner kan ett tåg begränsas av flera andra tåg samtidigt. Vi förutsätter att det, för de framtida händelserna för ett tåg, kommer att finnas högst ett annat tåg som innebär den största begränsningen. Ett annat sätt att säga samma sak är att det räcker att begränsa sig till parvisa konflikter. Under detta antagande kan begränsningar som ges av andra tåg studeras separat för varje begränsande tåg för sig. Vi vet dock inte i förväg vilket av de olika möjliga tågen som kommer att innebära den största begränsningen och eftersom tiderna mellan olika händelser är stokastiska variabler måste vi ändå ta hänsyn till samtliga relevanta begränsande tåginteraktioner.

3.7 Cato

Figur 8 visar CATO systemets gränssnitt mot föraren. CATO, Computer aided train information, är ett förarstöd som tagits fram av Transrail. Avsikten är att ge stöd till förare att köra energisnålt och effektivt genom att inte i onödan tvingas till inbromsningar. CATO är ett egoistiskt system som optimerar ett enskilt tågs rörelse utan att ta hänsyn till den omgivande trafiken och därför finns en möjlighet att systemet medför en större spridning av förseningar än vad som skulle vara fallet utan att systemet används. Om vi antar att förarna bryr sig om informationen från CATO, så behöver vi undersöka vilken effekt detta har. Om det verkligen har en avgörande effekt för förseningsspridningen att förare kör enligt CATO, då behöver vi räkna ut vad lokföraren ser, för att modellera gångtiden. Troligen är detta inte fallet, det vill säga, troligen har användningen av CATO inte en avgörande betydelse för hur förseningar sprids. Under detta antagande behövs inte någon särskild behandling av CATO tåg för modellering av gångtider. Effekten på förseningsspridningen hos CATO behöver utredas.



Figur 8: Prediktion av konflikt i Cato

3.8 Om att göra relevanta analyser

Om man inte tillräckligt noga skaffar sig kunskap om hur tåg rör sig under olika förutsättningar och vilka orsakssamband som finns är det risk att man tar fram information som är oanvändbar eller till och med vilseledande och därmed förödande för beslutsfattandet i den operativa driften.

Låt oss betrakta följande situation, som ett exempel på när statistiska gångtidsmodeller inte ger information som kan användas för prediktion av framtida läge i trafiken. Antag att det som begränsar kapaciteten i en del av ett järnvägsnät är uppehåll och konflikter mellan tåg inne på stationerna. Antag också att det finns gott om bufferttid för tågen på linjen, så att om tågen håller högsta tillåtna hastighet på linjen kommer de, i de allra flesta fall, att tvingas stanna och vänta på att få köra in på stationen. Antag därutöver att förarna, efter att de lämnar en station, har tillgång till information om när spåret kommer att bli fritt så att de kan köra in på nästa station. Antag slutligen att förarna anpassar sin hastighet för att undvika att få stopp på linjen. I en sådan situation är gångtiderna att betrakta som härledbara från andra händelser. Om man i denna situation struntar i de händelser som verkligen bestämmer utfallet hos trafiken och istället försöker göra prediktioner baserat på modeller av gångtider kommer det inte att bli rätt. Det fel vi gör här är att vi studerar en sekundär effekt, gångtider, som helt styrs av en primär effekt, uppehåll och konflikter på stationer.

För att klara av att dra relevanta slutsatser från data och göra prediktioner behöver vi en grundläggande idé om vad som är primära orsaker och vad som är följdersaker. Vad som är vad kan bero på det aktuella läget i trafiken.

Låt oss betrakta följande situation, som ett ytterligare exempel på när statistiska gångtidsmodeller inte ger information som kan användas för prediktion av framtida läge i trafiken. Antag att lokförarna tar emot information från operatörernas fordonsledning om hur de skall köra. De order som utgår från fordonsledningen är antingen "kör långsamt" eller "kör fort". Dessa order är kända för den operativa tågledningen. Antag att vi sätter ett verktyg i händerna på tågledningen som räknar ut gångtider för tågen och kan predicera ankomsttider. Verkttyget använder historiska data och realtidsdata för att göra dessa prediktioner, men fordonsledningens order till lokförarna är okänd för verktyget. Det är då risk att de predicerade gångtiderna blir medelvärden av "kör långsamt" och "kör fort" beordrade tåg. Eftersom varje tåg i verkligheten kör antingen fort eller långsamt kan prediktionerna och gångtiderna som verktyget tar fram bli mycket sämre än en grov gissning baserat på att man vet vilka tåg som "kör långsamt" eller "kör fort."

Detta fiktiva exempel är inte irrelevant. Om det är så att de omgivande faktorerna och information till förarna har avgörande betydelse för hur tåg rör sig

och denna information delvis är känd av tågtrafikledningen, då kan gångtidsmodeller och prediktioner vara värdelösa om de tas fram utan hänsyn till denna information.

3.9 Grundläggande idéer till trafiklägesprediktion

Att skilja på primära och sekundära orsaker till störningar är fundamentalt för att förstå vad som händer när förseningar uppstår och därmed också en central komponent vid prediktion av ett framtida läge i trafiken.

Ett sätt att se på trafiken är att tänka på den som att det initialt inte finns några störningar och tågen går enligt tidtabellen, men rätt som det är dyker primära förseningar upp och effekterna av dem sprider sig som ringar på vattnet, via signalsystemet eller via andra informationsbärare vilka bestämmer beroendena mellan händelserna i händelsesystemet. Med kunskap om vilka beroenden som finns vet vi vad som påverkas av en primär försening. För de tider mellan händelse som påverkas av störningarna tar vi hänsyn till detta. För tider mellan händelser som inte påverkas av störningarna kan vi använda statistiska modeller från historiska data för ostörda förlopp. Från historiska data vill vi alltså urskilja de primära orsakerna till förseningar och bygga statistiska modeller för dessa. Gångtidsmodeller m.m. för ostörda rörelser kan betraktas som statistiska modeller för primära orsakerna till förseningar.

Antag att ett tåg står stilla på linjen till följd av en konflikt med ett annat tåg som är försenat. Antag att konflikten sedan helt och hållet upphör. Vi har då ett tåg som har råkat ut för en sekundär försening men inte längre är stört av något annat tåg. Eftersom rörelsen för tåget är ostörd efter att konflikten upphört bör vi kunna använda en statistisk modell för hur den beter sig. Vi kan dock kanske inte använda samma modell som för de tåg som inte har råkat ut för en försening. Gångtidsmodeller m.m. för vissa ostörda men sekundärt försenade rörelser kan då också betraktas som statistiska modeller för primära orsakerna till ytterligare förseningar.

För tåg som bara står still och väntar på att en konflikt skall upphöra behövs inte någon separat prediktion. Vi predicerar istället de ostörda förlopp som leder till att konflikten upphör. För det väntande tåget predicerar vi bara det som händer efter att konflikten upphört och tåget är ostört. Denna beskrivning innehåller grundkomponenterna i att göra prediktioner: (1) skilj på primära och sekundära orsaker till störningar (2) hantera endast de primära och (3) hantera primära förseningsorsaker hos tidigare störda rörelser separat från primära förseningsorsaker hos tidigare ostörda rörelser, om det är befogat. I vår modell är alla tider stokastiska variabler så att tågen i varje framtida ögonblick med en viss sannolikhet är störda och med en annan sannolikhet är försenade.

Referenser

- [1] K. Dogany, J. Ekman, and A. Holst. Improved train control by higher quality of operative information. Slutrapport till projektet Koi, Trafikverket, 2011.
- [2] J. Ekman and P. Kreuger. En analytisk metod for utredning av kapacitet vid signalprojektering. SICS Technical Report T2003:11, Swedish Institute of Computer Science, 2003. FoU-rapport, Banverket.
- [3] J. Törnquist Krasemann and et al. Forstudie uppföljning, kapacitetsplanering, simulering och trafikstyrning (fuks). Technical report, Trafikverket, 2014.