

Prediktion av ett framtida läge i tågtrafiken
- en forskningsöversikt

Jan Ekman,
SICS Swedish ICT

5 juni 2015

Innehåll

1	Inledning	3
2	Arbeten utförda på Delfts tekniska högskola om prediktion av tågtrafikläge	4
2.1	Inledning	4
2.2	TNV-prepare	5
2.3	TNV conflict	5
2.4	Prediktion baserat på uppskattning av minimala processtider . . .	6
2.5	Prediktion baserat på medelavvikelse från plan hos realtids data	7
2.6	Mer om faktorer och processtider	8
2.7	Prediktion baserat på statistiska förseningsmodeller	9
3	Exempel på arbeten utförda i Sverige	9
3.1	Samband mellan infrastrukturfel och tågförseningar	9
3.2	Prediktion av trafikläge i CATO och i On-Time projektet	10
	Referenser	12

1 Inledning

Projektet *Uppföljning och prediktion*, med Trafikverkets registrerings nummer TRV 2014/34857, utförs under 2014 och 2015 som 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. Det är en del av SICS arbete i projektet att ta fram denna forskningsöversikt. Det arbete som SICS i övrigt har utfört i projektet beskrivs i rapporten *Slutrapport för SICS del av projektet Uppföljning och Prediktion*.

SICS del av projektet *Uppföljning och prediktion* skall ses som en fortsättning på projektet *Koi* 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.

2 Arbeten utförda på Delfts tekniska högskola om prediktion av tågtrafikläge

2.1 Inledning

Detta kapitel ger en översikt över den forskning kring prediktion av ett framtida läge i trafiken som utförts av forskare på Delfts tekniska högskola. Med ett undantag rör denna forskning Holländsk järnväg. Den väsentliga innebörden i detta är att, bortsett från det enda undantaget, forskningsresultatens fallstudier handlar om Holländsk infrastruktur och tågtrafik och att data från Holländska infrastrukturförvaltare och operatörer har använts. Vi kommer att gå igenom delar den omfattande mängden av publicationer från Delfts tekniska högskola.

För att göra en begriplig genomgång av denna forskning kommer vi att för var och en av de olika arbetena försöka svara på ett antal frågor. Detta är exempel på frågor vi kommer att försöka svara på:

1. Vilken är avsikten i det längre perspektivet, är det ett stöd till den operativa driften eller något annat.
2. Vilket delproblem löses.
3. Hur skall lösningen användas, d.v.s. vilket är nästa steg för att komma en bit på väg mot det målet i det längre perspektivet
4. Vilka data har använts och var kommer de ifrån
5. Vilka faktorer tar man hänsyn till, t.ex. tågtyper, rusningstrafik, årstid, för tidiga respektive försenade tåg.
6. Vilka metoder har använts i lösningen, har t.ex. statistiska fördelningar över gångtider använts.
7. Var ligger fokus och tyngdpunkt i presentationen av arbetet.
8. Vilka fallstudier har gjorts och
9. Vilket resultat har uppnåtts.
10. Är de presenterade metoderna och verktygen användbara för Trafikverket och svensk järnväg

Särskilt fokus ligger på metoderna som används för olika delproblemen till att predicera ett framtida tågtrafikläget. Det finns säkert många andra intressanta frågor att ställa men låt oss, i denna studie, nöja oss med att undersöka de olika arbetena från det perspektiv som dessa frågor ger.

2.2 TNV-prepare

För järnvägsforskningen vid Delfts Tekniska högskola var ett viktigt steg, för förbehandling av data, verktyget *TNV-Prepare* beskrivet i [5] från år 2000. Detta verktyg lägger grunden för forskning som senare utförs om prediktioner av tågtrafikläget.

TNV-Prepare läser data som erhållits från signalsystemet. Denna källa, signalsystemdata, refereras i [5] som “the Dutch train describer systems.” Tidtabell och infrastruktur är också indata till TNV-Prepare. Det data som TNV-Prepare producerar är data per tåg och linje om tider för olika händelser, som att tåget når en viss blocksträcka och att en tågväg låses upp. TNV-Prepare anger också tågnummer till det tåg som en händelse rör. Den avsikt med TNV-Prepare som [5] anger är att vara en viktig *möjliggörare* i tågplaneprocessen. I introduktionen till [5] sägs att:

“Existing data collection methods used by NS Verkeersleiding and the railway customer organization ROVER only give rough estimates in minutes of arrival and departure times. These figures give some indication about punctuality but lack the precision, accuracy, and detail for research purposes”

Det vill säga en av avsikterna med *TNV-Prepare* är att skapa data som ger bättre förutsättningar för järnvägsforskning än befintliga data med grova uppskattningar av ankomster och avgångar på minutnivå.

Större delen av [5] ägnas åt att beskriva dataflöden, arkitektur och gränssnitt hos verktyget *TNV-Prepare*, men även noggrannheten diskuteras. I slutet på [5] diskuteras också vilka möjligheter som ges med det nya verktyget: analys av kapacitet, punktlighet och stokastiska processer. En av slutsatserna från är [5] att det initiala arbetet att mata in infrastrukturdata för att kunna använda *TNV-Prepare* är aldeles för mödosamt och att verktyget bör kopplas till ett annat verktyg “CARE” som beskrivs i [4].

Verktyget *TNV-Prepare* är anpassat till Holländska system och data och är förmodligen inte lämpligt att använda med indata från Trafikverkets system. Idén att ta fram allmänt användbara verktyg till att skapa precis information att använda i järnvägsforskningen och i nya tillämpningar i tågplaneprocessen är relevant också i Sverige.

2.3 TNV conflict

Verktyget *TNV conflict*, beskrivet i [6] och [2], använder historiska data för att avgöra vilka tåg som har varit i konflikt. Det vill säga *TNV conflict* avgör på

detta sätt vilka förseningar som är primära och sekundära. Den uppdelning av förseningar i primära och sekundära som *TNV conflict* gör används för prediktion av det framtida läget i efterföljande arbeten. Det anses i artikeln att konflikter ofta inte identifieras i den operativa driften och att man därför inte vet vad som är primära och sekundära förseningar. Avsikten i det längre perspektivet med *TNV conflict* är att bättre förstå orsakerna till förseningar och använda detta till att förbättra kapacitetstilldelning och tidtabellkonstruktion.

Indata till *TNV conflict* är infrastruktur meddelanden, infrastrukturdata från PRO Rail, tidtabell, och övrig infrastrukturdata. *TNV conflict* använder inte data producerat av *TNV-Prepare* utan betraktas som en utvidgning av detta system. Huvudidén till *TNV conflict* är att använda data om när tågvägar ställs för att upptäcka tåg som tvingas bromsa för att spåret inte är fritt. För att kunna avgöra vilka tåg som är i konflikt gör man en modell av signaleringen “the signalling logic” i form av ett *färgat petrinät*, som beskriver signalsystemets tillstånd för de aktuella tågen och transitioner mellan sådana tillstånd.

På samma sätt som för *TNV-Prepare* är *TNV conflict* anpassat till Holländska system och data och är förmodligen inte lämpligt använda med indata från Trafikverkets system. Idén att ta fram allmänt användbara verktyg till att skilja primära förseningar från sekundära är relevant också i Sverige, för t.ex. att kunna göra körtidsmodeller och predicera framtida läge i tågtrafiken.

2.4 Prediktion baserat på uppskattning av minimala processtider

I [12] från 2010 beskrivs en metod att predicera det framtida läget i tågtrafiken. Metoden baseras på data om tågkonflikter som *TNV conflict* producerar. Avsikten är att predicera det framtida läget en timme framåt i tiden som ett stöd till den operativa trafikledningen. Det vill säga i detta fall används *TNV conflict* på realtidsdata och inte på historiska data som beskrivet i [6] och [2].

Arbetet är inriktat mot att stödja den operativa driften genom att förutsäga konsekvenserna för olika operativa beslut. Avsikten i det längre perspektivet är alltså det samma som för Koi-projektet [3]. Målet för arbetet i [12] sägs vara *att ta fram och utvärdera en modell för förseningsspridning som erhåller sina parametrar från statistiska analyser på historiska data och som tar hänsyn till konflikter nedströms*.

Det delproblem till att predicera det framtida läget som studeras i [12] är att bestämma *tiden* för ett tåg mellan två större stationer. Denna tid, som kallas gångtid [“**train running time**”] i [12], är inte bara tiden för tågets rörelse utan den totala tiden för tåget att ta sig mellan två stationer, inklusive stopp på mellanliggande stationer. Det handlar alltså om prediktioner på *makronivå*.

För att bestämma tiden för ett tåg mellan två stationer konstrueras en händelsegraf [“**timed event graph**”] för trafiken som helhet. Händelserna i denna händelsegraf är på en finare nivå än den som gäller tiden för ett tåg mellan två stationer. Konstruktionen av händelse grafen består av att (1) bestämma vilka händelser som är beroende och (2) bestämma minsta tiden mellan beroende händelser. Dessa minsta tider refereras till som kantvikter [“**arc weights**”] och förloppet mellan ett par av beroende händelser kallas process. Kantvikterna är inte stokastiska variabler utan fixerade tal. Det finns ett antal olika typer av beroenden (kanter) att bestämma en minsta tid för: gångtid, uppehåll, försprång, möten och andra tåginteraktioner. Gångtid som kantvikt i händelsegrafens kallas också för [“**train running time**”] men är inte av samma typ som de gångtider på makronivå som inkluderar uppehåll. Tider för alla typer av tåginteraktioner benämns [“**headway**”] i [12].

För att bestämma kantvikterna, det vill säga minimumtiderna för processerna, används historiska data. I den fallstudie som utförts, på sträckan Rotterdam-Haag, används historiska data på följande sätt: för tåginteraktioner väljs som minsta tid, d.v.s. kantvikt, den tionde percentilen av en utvald mängd tåginteraktioner, så att det är c:a 90% chans att en tåginteraktion tar längre tid än minsta tiden. På samma sätt väljs den tjugonde percentilen av utvalda mängder stopp och gångtider, som kantvikter till processerna stopp respektive gång. Det finns i [12] inte någon förklaring till dessa val.

En stor del av arbetet som presenteras i [12] är diskussioner om faktorer som har påverkan på processtiderna. Två faktorer som diskuteras är rusningstid och försenade tåg. I [12] hävdas att man från tillgängliga data inte kan se att förseningar har någon påverkan på gångtider. Det hävdas att det statistiska underlaget är för litet för att dra slutsatser om vilken påverkan rusningstid har. Vi kommer att fortsätta genomgången av diskussionerna kring dessa faktorer i avsnittet 2.6.

2.5 Prediktion baserat på medelavvikelse från plan hos realtids data

I [9] från 2013 presenteras en metod att predicera ett tågs framtida försening. Metoden beskrivs också i [10] och som kapitel 5 i avhandlingen [7].

Huvudidén i detta arbete är att ta hänsyn hur tåg som tidigare, flera blocksträckor tillbaka i tiden, underskridit eller övererskridit den planerade tiden att passera en blocksträcka. Metoden som används är att beräkna avvikelsen i medel från den planerade tiden och anta att avvikelsen även i fortsättningen kommer att vara lika stor. Avsikten i det längre perspektivet är en “effektivare tågstyrning, omplanering och passagerarinformation.” Det fallstudien i [9] demonstrerar och som verkar vara det outtalade målet i det kortare perspektivet är att predicera ankomsttider för tåg, med hänsyn till framtida konflikter.

Indata till metoden är tidtabell, operativ plan, tågens aktuella positioner och infrastrukturmeddelanden [“data from the train describer system”]. Det finns inte någon referens till *TNV prepare* eller *TNV conflict* i [9] men det ser ut som tågkonflikterna är indata till metoden. I [7] sägs att:

“ The graph is constructed based on the actual process plan that includes the given train routes, scheduled event times (actual timetable) and connection plan (§2.2.1). The connection plan contains all planned synchronisation constraints.”

Baserat på indata skapas en händelsegraf [“timed event graph”], där kantvikterna för gångtider beror på den uppskattade medel-avvikelsen från planen. En faktor som tas hänsyn till vid beräkning av kantvikter är försenade tåg. [9] refererar till en fallstudie för sträckan Leiden – Dordrecht och rapporterar ett prediktionsfel i medel på mindre än en minut, för en 30 minuters prediktionshorisont.

2.6 Mer om faktorer och processtider

En central del i abretena i som presenteras i avsnitten 2.4 och 2.5 i denna rapport är uppskattningen av processtider, det vill säga tider för gång på mikronivå, uppehåll och tåginteraktioner. I [12] ligger tyngdpunkten på uppskattningen av de minimala processtiderna. En utförligare beskrivning av arbetet som presenteras i [12] finner vi i [11] som också innehåller en diskussion om vilka faktorer som det är väsentligt att ta hänsyn till vid skattning av processtider. Några faktorer som diskuteras och analyseras är: försening, rusningstid, tågtyp [“rolling stock”] och väder. Det framgår av [11] att det kan vara svårt att ta hänsyn till dessa faktorer och att de i sin tur kan bero av andra faktorer. Om vi tar lövhalka som ett exempel, som betraktas som ett specialfall av väder, så beror denna på förekomsten av träd vid spåret. Ett annat exempel är regnets inverkan på adhesionen och därmed på gångtiderna. Ett lätt regn kan göra spåret halt medan ett kraftigt regn rengör spåret och kan göra spåret mindre halt.

Kapitel 4 i avhandlingen [7] ägnas åt processtider. Metoden som används att bestämma gångtider delas upp i en *global* del och en *lokal* del. Den globala delen rör skattning av gångtider från historiska data med gångtider från samtliga sträckor och inte bara från den sträcka för vilken man vill uppskatta gångtiden. För den lokala delen anges att dess uppgift är att *undersöka påverkan från en avgångsförsening*. Den globala delen av gångtidsuppskattningen tar hänsyn till följande faktorer: rusningstid, avgångsförsening, försprång [“headway”], avstånd från föregående stopp, avstånd till kommande stopp, blocksträckornas längd och tågtypen. Eftersom det är komplicerat att avgöra från data huruvida det råder rusningstid eller inte så används NS definition av rusningstid. NS (Nederlandse Spoorwegen) är den dominerande holländska tågoperatören på persontågtrafik sidan. Vi noterar att lutningar inte tas hänsyn till i den globala delen, något som kanske behöver göras om metoden skall anpassas till svenska förhållanden.

2.7 Prediktion baserat på statistiska förseningsmodeller

I [8] presenteras en metod att skatta risken för en framtida försening för tåg på höghastighets banan mellan Shanghai och Beijing. Avsikten i ett längre perspektiv med detta arbete är att kunna göra bättre prediktioner av läget i tågtrafiken. Arbetet är en del av Capacity4Rail projektet.

Idén som arbetet bygger på är att det finns ett starkt samband mellan aktuell försening och framtida försening. I den statistiska modell som används reduceras förseningsstatusen hos tåg till tre olika tillstånd: (1) tåget är före tidtabell eller i tid, (2) tåget är försenat med en försening på högst fem minuter (3) tåget är mer än fem minuter försenat. Den statistiska modellen består av sannolikheter för transitioner mellan dessa tillstånd, mellan valda händelser för tågen, där händelserna är ankomster till och avgångar från stationer. Ett exempel på en sådan modell är den följande

$$\begin{array}{l} \text{early} \\ \text{small} \\ \text{large} \end{array} \begin{pmatrix} \text{early} & \text{small} & \text{large} \\ 0.97 & 0.02 & 0.01 \\ 0.45 & 0.45 & 0.10 \\ 0.00 & 0.14 & 0.86 \end{pmatrix}$$

som beskriver sannolikheterna för transitionerna för tåg från det att de avgår från Beijing till att de ankommer Tianjin. Modellen säger t.ex. att om tåget avgår med en stor försening, d.v.s. på mer än fem minuter, från Beijing så är sannolikheten 0 att det ankommer före tidtabelltiden till Tianjin. För att skatta transitionssannolikheterna används historiska data. Metoden baserar sig uteslutande på sambanden mellan aktuell försening och framtida försening och ignorerar alla andra orsaker till förseningar.

[8] rapporterar ett resultat med ett klassificeringsfel på knappt 0.2 på en timmes horisont. Med en tillståndsrymd med förseningar i hela minuter, istället för bara de tre beskrivna tillstånden, erhålles ett prediktionsfel på c:a 1.5 minuter på en timmes horisont.

3 Exempel på arbeten utförda i Sverige

3.1 Samband mellan infrastruktur fel och tågförseningar

Projektet *Utvecklad fel- och förseningsmodell* är ett CTS-projekt, där CTS står för *Centrum för Transportstudier*. CTS har syftet “att skapa en internationellt framstående forskningsmiljö inom transportforskningsområdet.” CTS erhåller

basfinansiering från VINNOVA och Trafikverket, som främst används för att finansiera forskning på KTH - Transportvetenskap och VTI - TEK, och i lägre grad på JIBS. Projektet *Utvecklad fel- och förseningsmodell* utförs av WSP och KTH.

Som namnet antyder handlar projektet *Utvecklad fel- och förseningsmodell* om förbättringar i beräkningarna som ligger till grund för den Fel- och förseningsmodell som Trafikverket äger och förvaltar. Trafikverket beskriver modellen följande sätt:

Fel- och förseningsmodell är en övergripande modell som används för samhällsekonomiska effekter av förändrade antal fel i järnvägsanläggningen.

Felen i detta fall är samtliga fel registrerade i järnvägsanläggningen. Beräkningsmetoden är tillämpningen av ett effektsamband, metoden innefattar endast konsumenteffekter – förseningstid för persontåg.

En del av Fel- och förseningsmodellen är att beräkna felelastisiteter, vilket betyder förändringen i andelen förseningar relativt en förändring i andelen fel. Som ett exempel, om en minskning av infrastrukturfelen på 10% leder till en minskning av förseningarna på 8% då är felelastisiteteten $0.08/0.1 = 0.8$. För precisa beräkningar av felelastisiteter behövs en analys av förseningsspridningen för de primärförseningar som orsakas av infrastrukturfelen. Problemet att beräkna hur förseningar sprids från några givna förseningar är i stor utsträckning samma problem som att predicera ett framtida läge i trafiken från ett givet läge.

Den metod som används i projektet *Utvecklad fel- och förseningsmodell* beskrivs i [1]. Metoden består av följande två delar. (1) Sannolikhetsfördelningar för gång och uppehållstider konstrueras från historiska data från Trafikverkets Banstat system. (2) Dessa sannolikhetsfördelningar används för att uppskatta sekundärförseningar, via simulering.

Tåginteraktioner och tider för tåginteraktioner är indata till modellen. Rapportens tyngdpunkt ligger på konstruktionen av statistiska fördelningar för gångtider och väntetider på stationer, där gångtider antas vara invers normalfördelade och väntetider weibull fördelade.

3.2 Prediktion av trafikläge i CATO och i On-Time projektet

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. Det som möjliggör detta är prediktioner av när spåret framför ett tåg blir fritt.

EU-projektet ONTIME (Optimal Networks for Train Integration Management across Europe) har som långsiktigt mål att reducera förseningar och förbättra trafikflödet, enligt hemsidan

<http://www.ontime-project.eu/>

Transrail deltar i detta projekt med en delaktivitet som handlar om predicering "Predicting train movements". Arbetet verkar inte vara publicerat, men idén beskrivs i följande presentation

http://www.trafikverket.se/contentassets/e40057ec1af54a6583bf1a7a41ddf93b/8_t_liden_transrail_research_ontime.pdf

Det framgår att historiska data används och att en graf används, förmodligen för att modellera konflikter. Det går inte att svara på frågor om metoden och vad den skall användas till från detta material.

Referenser

- [1] Utvecklad modell för effektsamband mellan fel i infrastruktur och tåg förseningar. Wsp, 2014.
- [2] Winnie Daamen, Rob MP Goverde, and Ingo A Hansen. Non-discriminatory automatic registration of knock-on train delays. *Networks and Spatial Economics*, 9(1):47–61, 2009.
- [3] K. Dogany, J. Ekman, and A. Holst. Improved train control by higher quality of operative information. Slutrapport till projektet Koi, Trafikverket, 2011.
- [4] G Fries. Computer aided railway engineering. In *The 1998 6 th International Conference on Computer Aided Design, Manufacture and Operation in the Railway and Other Advanced Mass Transit Systems*, pages 53–59, 1998.
- [5] RMP Goverde and IA Hansen. Tnv-prepare: analysis of dutch railway operations based on train detection data. In *International conference on computers in railways*, pages 779–788, 2000.
- [6] Rob MP Goverde, Winnie Daamen, and Ingo A Hansen. Automatic identification of route conflict occurrences and their consequences. *Computers in Railways XI*, pages 473–482, 2008.
- [7] P Kecman. *Models for Predictive Railway Traffic Management*. TU Delft, Delft University of Technology, 2014.
- [8] P Kecman, F Corman, and L Meng. Train delay evolution as a stochastic process. In *The 66th Conference on on Railway Operations Modelling and analysis - RailTokyo2015*, 2015.
- [9] Pavle Kecman and Rob MP Goverde. Adaptive, data-driven, online prediction of train event times. In *16th IEEE Annual Conference on Intelligent Transportation Systems (ITCS 2013)*, 2013.
- [10] Pavle Kecman and Rob MP Goverde. Online data-driven adaptive prediction of train event times. 2014.
- [11] D.J. Van der Meer, R.M.P. Goverde, and I.A. Hansen. Prediction of train running times using historical track occupation data. Technical report, Department of Transport and Planning, Delft University of Technology, Delft, 2009.
- [12] D.J. Van der Meer, R.M.P. Goverde, and I.A. Hansen. Prediction of train running times and conflicts using historical track occupation data. In *The 12th World Conference on Transport Research*, 2010.