

Slutrapport för delstudie 1 i projektet FLOAT

- Beslutstöd för proaktiv trafikledning på Malmbanan

Sammanfattning

I forskningsprojektet FLOAT utvecklar och utvärderar vi koncept, principer och metoder för beräkningsstöd som syftar till att underlätta trafikledarens arbete i den operativa driften.

Det är dock allmänt känt att det är svårt att snabbt, i realtid, lösa komplexa optimeringsproblem såsom operativ omplanering av tågtrafik är. Det finns således ett reellt behov av att studera olika metoder för att formulera problemet väl, att utveckla och utvärdera effektiva lösningsmetoder och framför allt att undersöka dess praktiska tillämpbarhet. Det är dessa utmaningar som är centrala i FLOAT.

Projektet FLOAT består av två delstudier, där den första fokuserar på förutsättningarna på Malmbanan medan den andra delstudien fokuserar på Södra Stambanan. I denna rapport sammanfattas arbetet inom den första delstudien. Arbetet är utfört av Blekinge Tekniska Högskola på uppdrag av Trafikverket och finansierat inom ramen för KAJTs verksamhet. Ett antal representanter från Trafikverket och forskare inom KAJT har under april 2016 erbjudits möjlighet att ge synpunkter på en remissversion av denna rapport. Trafikverkets synpunkter har inarbetats i denna slutgiltiga version av rapporten. Från forskare inom KAJT har inga synpunkter på rapporten inkommit.

Inom projektet har vi gjort en enklare analys av undervägspunktligheten av trafiken på sträckan Narvik-Kiruna för att få en bättre bild av i hur pass stor omfattning trafiken avviker från aktuell tidtabell i praktiken. Analysen baseras på utdrag av data från LUPP för perioden maj 2014 samt perioden januari-maj 2015. Analysen visar att trafiken mycket ofta avviker från aktuell tidtabell, vilket skapar merarbete för trafikledarna.

Inom projektet har vi även utvecklat och experimentellt utvärderat optimeringsmetoder för omplanering av trafiken vid störningar och andra avvikelser från tidtabeller. Experimenten har baserats på ett antal störningsscenarier av olika typer. Baserat på vald målfunktion som minimerar förseningar, så tar optimeringsmetoden beslut om:

- Tågens ankomst- och avgångstider
- Spårval på linjen och stationer (inkl. plattformsväl)
- Tågordningen (inkl. möten och förbigångar).

Vi har även diskuterat våra resultat och behoven av beslutstödande funktioner med personalen i Boden och det verkar finnas ett betydande behov av grundläggande beslutsstödande funktioner. Det som därför verkar mest relevant att arbeta vidare med i nuläget utifrån Trafikverkets behov och förutsättningar är två funktioner:

1) En leveranskvalitetssäkrande funktion som kan putsa till och trimma planen och konfliktreglera tågplanen mm när den går över till trafikledningen från korttidsplaneringen, samt

2) En konsekvensanalytisk funktion som i realtid "bedömer" aktuell plan och dess sannolika konsekvenser över tiden samt som ev. föreslår mindre justeringar av planen vid behov.

1. Inledning

Att trafikledarnas arbetsuppgifter är mycket komplexa och tidvis enormt kognitivt belastande är de allra flesta överens om sedan länge. Samtidigt är tillgången till beslutstöd och hjälpmedel mycket begränsad. Eftersom trafiken i stora delar av det svenska järnvägsnätet har blivit alltmer intensiv och oregelbunden de senaste åren och efterfrågan på transparenta, ändamålsenliga operativa prioriteringskriterier har lyfts fram, så har det blivit ännu tydligare att det finns ett behov av beslutstöd för trafikledarna.

I forskningsprojektet FLOAT utvecklar och utvärderar vi koncept, principer och metoder för beräkningsstöd som syftar till att underlätta trafikledarens arbete i den operativa driften.

Det är dock allmänt känt att det är svårt att snabbt, i realtid, lösa komplexa optimeringsproblem såsom operativ omplanering av tågtrafik är. Det finns således ett reellt behov av att studera olika metoder för att formulera problemet väl, att utveckla och utvärdera effektiva lösningsmetoder och framför allt att undersöka dess praktiska tillämpbarhet. Det är dessa utmaningar som är centrala i FLOAT.

Projektet FLOAT består av två delstudier, där den första fokuserar på förutsättningarna på Malmbanan medan den andra delstudien fokuserar på Södra Stambanan. I denna rapport sammanfattar vi arbetet inom den första delstudien. Arbetet har utförts av Johanna Törnquist Krasemann på Blekinge Tekniska Högskola och projektet finansieras av Trafikverket inom ramen för KAJTs verksamhet. Ett antal representanter från Trafikverket och forskare inom KAJT har under april 2016 erbjudits möjlighet att ge synpunkter på en remissversion av denna rapport. Trafikverkets synpunkter har inarbetats i denna slutgiltiga version av rapporten. Från forskare inom KAJT har inga synpunkter på rapporten inkommit.

2. Malmbanan: Förutsättningar, behov och möjligheter

Den enkelspåriga Malmbanan i norra Sverige tjänar som en viktig transportlänk för att transportera malm från gruvor i bl.a. Kiruna, Svappavaara och MalMBERGET till hamnarna i Narvik, Norge, samt Luleå. På banan kör även en hel del annan godstrafik samt persontrafik. Sträckan Kiruna-Riksgränsen trafikerades under 2014 av ca 30 tåg per dygn där 20% var passagerartåg, 60% malmtåg och 20% andra godståg.

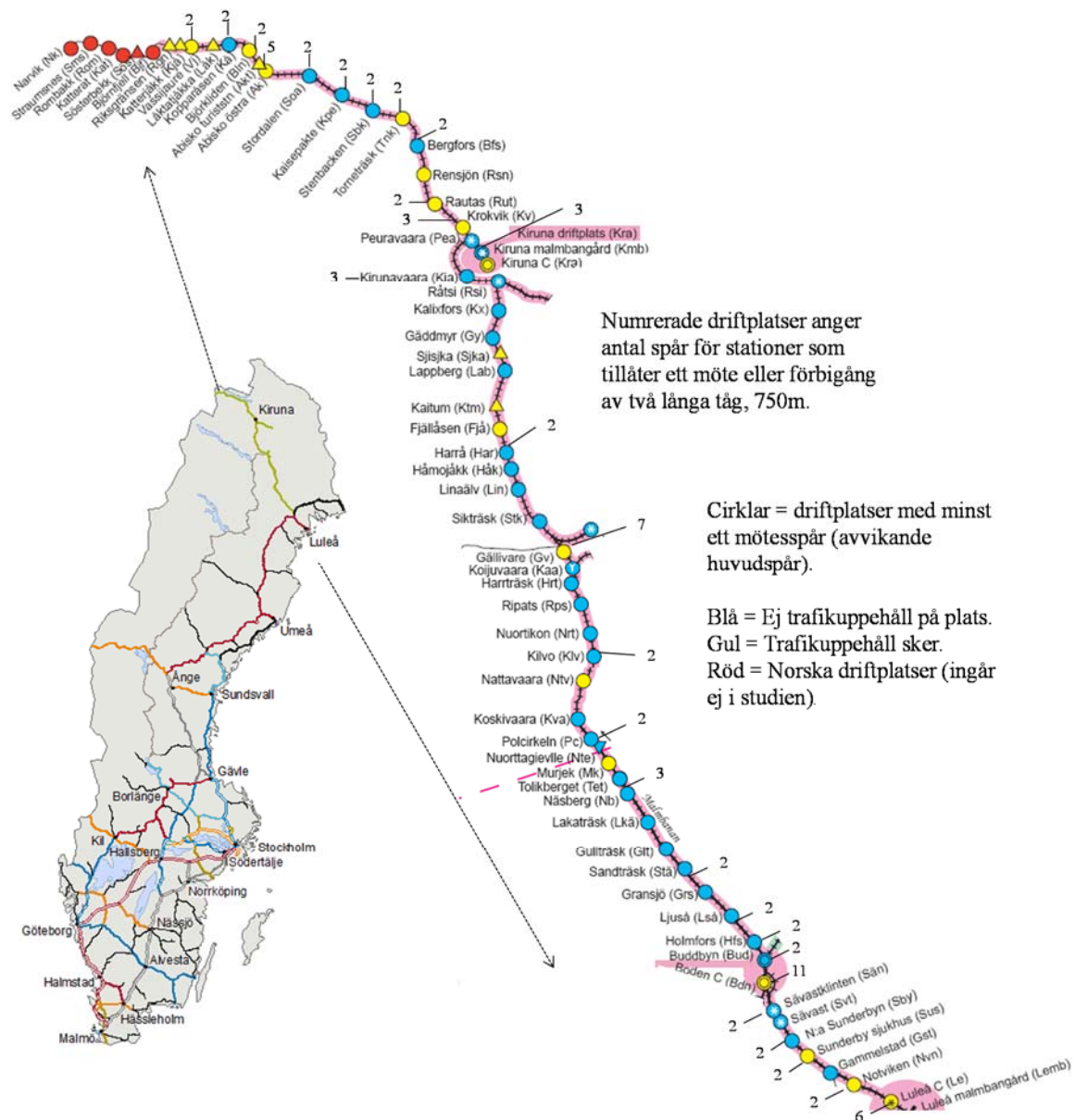
Malmbanan delas in i det norra omloppet (Narvik-Kiruna) och det södra omloppet (Kiruna-Luleå). Den norska delen mellan Riksgränsen och Narvik går även under namnet Ofotenbanan. Malmbanan är ca 500 km lång, elektrifierad och tillåter axellaster upp till 30 ton (STAX 30) och största tillåtna vikt per meter (STVM) är 12 ton. De tyngst lastade malmtågen väger ca 8600 ton, har 68 vagnar och är 750m som längst. Motsvarande olastat malmtåg väger ca 1470 ton. Andra godståg väger omkring 3400 ton och är på 430m (uppgifter från tidtabellen för 2014, T14).

Kapaciteten på Malmbanan har länge ansetts otillräcklig för att möta aktuellt och framtida behov av tågglägen. Under flera år har ett antal utredningar genomförts för att studera aktuella kapacitetshöjande åtgärder¹. Däribland förlängning av kortare mötesstationer eftersom det är långa tidsavstånden mellan de stationer som tillåter möten eller förbigångar av två långa malmtåg (750m), vilket medför onödigt långa restider på sträckan.

¹ Trafikverket, "Malmbanan bangårdsförlängning: Lakaträsk, Koskivaara, Ripats och Lappberg" Slutrapport januari 2012, Dnr TRV2010/33470.

Som exempel kan nämnas att ett södergående, olastat malmtåg hade i T14 på sträckan Björnfjell-Kiruna en planerad gångtid på 3 timmar och 11 minuter, varav ca 2h är faktisk körtid. Det innebär att mer än 1/3 av den planerade gångtiden är tidtabellteknisk tid för möten. Ett norrgående, lastat malmtåg från Kiruna har däremot en planerad körtid på 2 timmar och 33 minuter, där då endast ca 10 minuter utgörs av marginaltid och tidtabellteknisk tid.

I nedan figurer återges sträckan och de förutsättningar som gällde under tidtabellsåret 2014 och som också är de förutsättningar som gällt för de scenarier och experiment vi fokuserat på. Det vill säga kapacitetshöjande åtgärder som slutfördes under 2014 är inte inkluderade.



Figur 1. Illustration av Malmbanans sträckning och driftplatsernas kapacitet enligt förutsättningarna för T14. Egen bearbetning av figurer tillhandahållna av Trafikverket.

Under 2011-2014 genomfördes bangårdsförlängningar av Kopparåsen, Abisko, Kaisepakte, Rensjön, Ripats och Lakaträsk. Så i dagsläget (mars 2016) kan 27 av 53 mötesstationer klara av att samtidigt hantera två eller fler långa tåg på 750m. Det har inneburit att antalet malmtåg

successivt har kunnat öka. Idag går det 22 malmtåg per dygn mellan Kiruna och Narvik (i båda riktningarna) och mellan MalMBERGET och Luleå går det tio tåg per dygn.

Nu planerar Trafikverket att utöka kapaciteten ytterligare genom att bygga dubbelspår på sträckan Kiruna – Riksgränsen där den första etappen blir den 50km långa sträckan mellan Peuravaara och Bergfors². Byggstart är tidigast 2019, enligt Trafikverket. För att öka kapaciteten på den norska delen av Malmbanan, Ofotenbanan, har även Jernbanelverket genomfört åtgärdsvalstudier och föreslår dubbelspår inkl. en längre tunnel som ansluter den svenska delen vid Vassejaure. På sikt kommer även signalsystemet på Malmbanan att ersättas för att möta EUs krav och införandet av detta kommer att kräva temporära kapacitetsreduktioner.

Det är känt att framför allt godstågen frekvent avgår både betydligt för tidigt och för sent på Malmbanan och därmed följer trafiken ofta inte tidtabellen, utan avgångs- och ankomsttider behöver frekvent justeras och möten flyttas. Denna flexibilitet är naturligtvis nödvändig men medför en extra arbetsbelastning för trafikledningen.

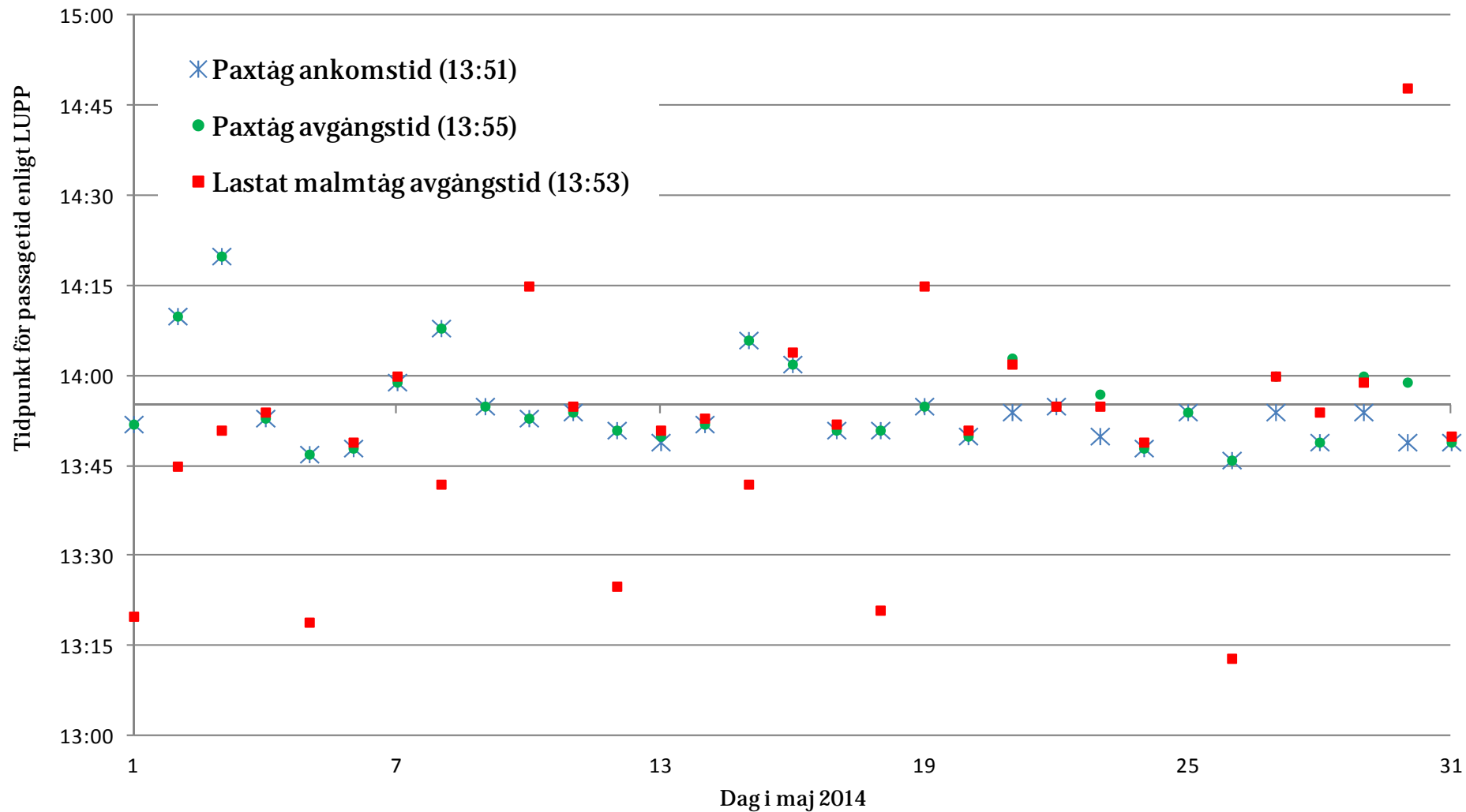
Inom projektet har vi gjort en enklare analys av undervägspunktligheten av trafiken på sträckan Narvik-Kiruna för att få en bättre bild av i hur pass stor omfattning trafiken avviker från aktuell tidtabell. Analysen baseras på utdrag av data från LUPP för perioden maj 2014 samt perioden januari-maj 2015 och visar att trafiken mycket ofta avviker från aktuell tidtabell.

Figur 2 (se nästa sida) visar faktiska passagetider under maj 2014 för ett tågpar som har ett planerat möte i Kopparåsen – ett södergående passagerartåg ("Paxtåg") och ett norrgående lastat malmtåg. I figuren kan man utläsa hur ofta planerade möten genomförts av de 29 dagar under maj då båda tågen var i trafik. Passagerartåget var planerat att anlända kl 13:51, ett par minuter före malmtåget, och göra ett kort, tidtabelltekniskt uppehåll medan malmtåget kör genom stationen utan uppehåll. Flera dagar anländer antingen malmtåget Kopparåsen en dryg halvtimme för tidigt eller så är passagerartåget försenat och i båda fallen förflyttas mötet förmodligen till en driftplats norr om Kopparåsen. I enstaka fall (3 av 29) är malmtåget försenat till Kopparåsen och persontåget passerar Kopparåsen utan uppehåll och möter malmtåget på en driftplats söder om Kopparåsen. I 16 fall av 29 genomförs det planerade mötet i Kopparåsen.

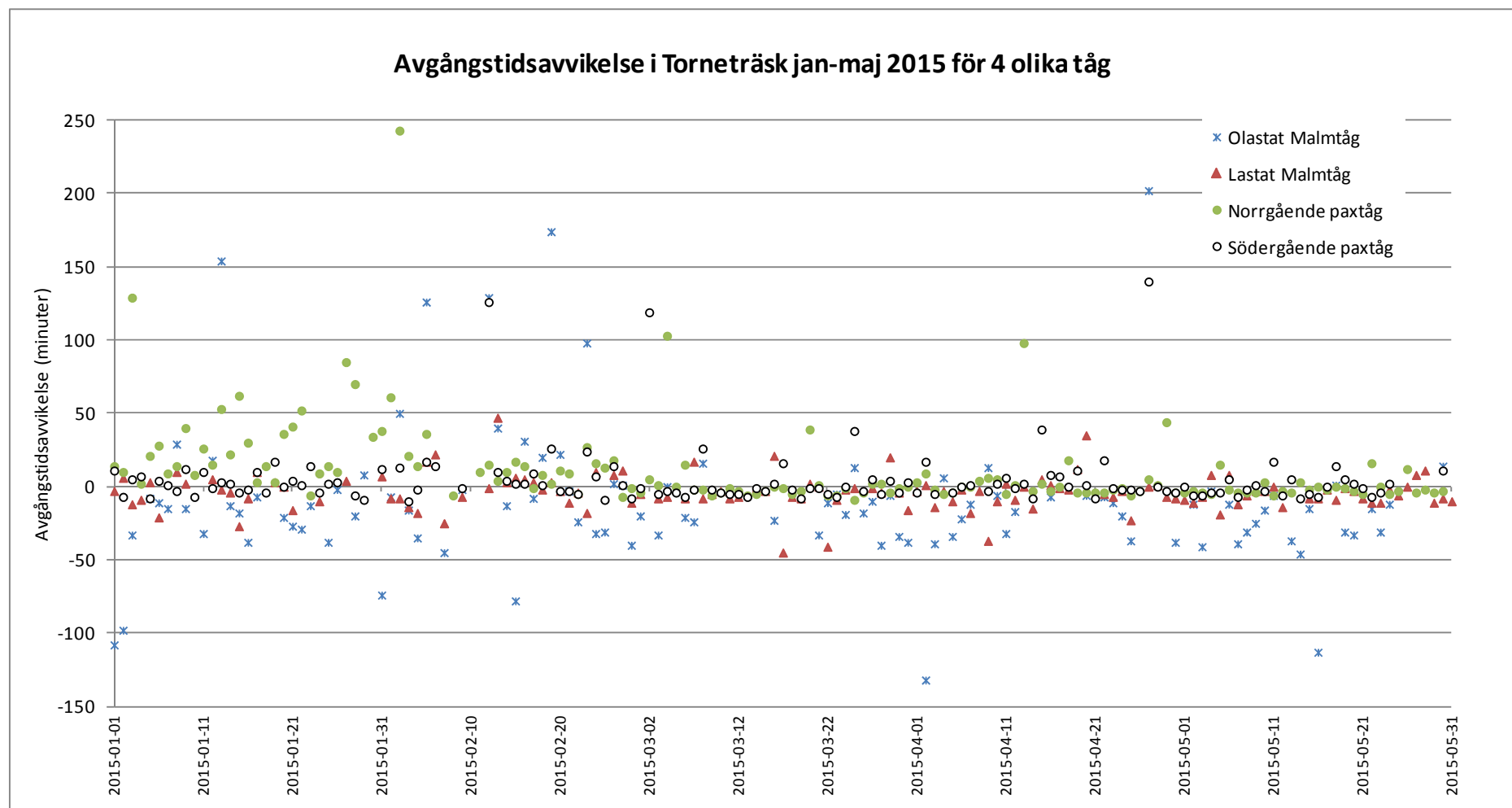
Figur 3 visar avgångstidsavvikelsen i Torneträsk för fyra olika tåg under januari-maj 2015.

² Trafikverket, "Fördjupad Åtgärdsvalsstudie Malmbanan, dubbelspår Kiruna – Riksgränsen Kiruna kommun", Norrbottens län, Beslutshandling 2015-02-27 Projektnummer: 137579.

Ankomst- & avgångstider för planerat möte i Kopparåsen under maj 2014



Figur 2. Uppföljning av ankomst- och avgångspunktighet för ett tågmöte i Kopparåsen (planerade tidpunkter inom parentes). "Paxtåg" avser resandetåg.



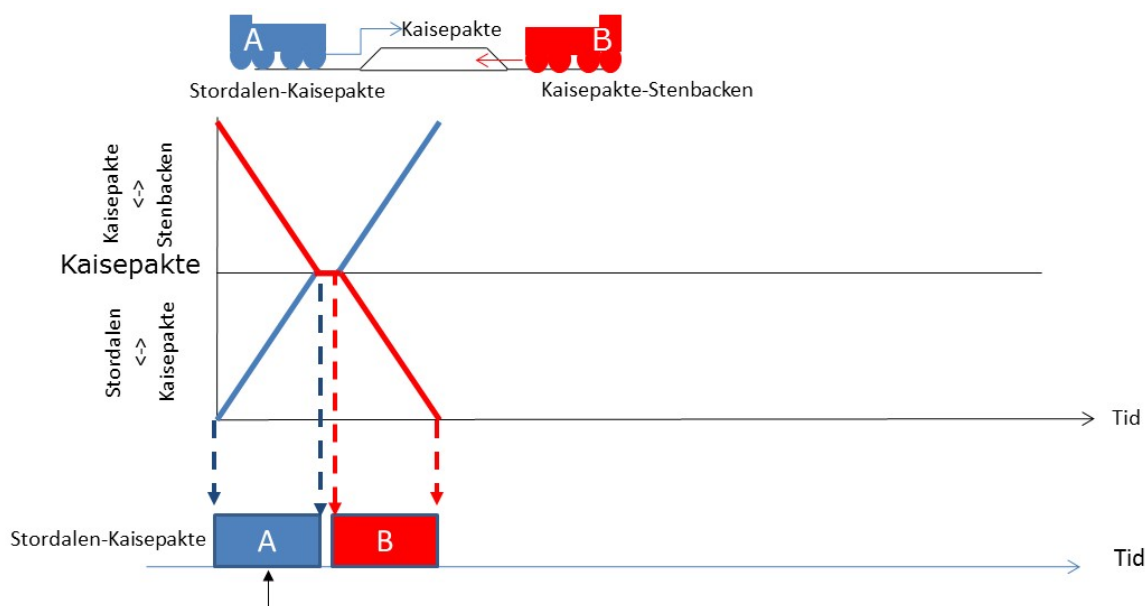
Figur 3. Avgångstidsavvikelse för ett lastat norrgående malmtåg, ett olastat södergående malmtåg, ett norrgående passagerartåg samt ett södergående passagerartåg. Under årets kalla inledande period är det som väntat fler avvikelser, i synnerhet för det norrgående passagerartåget, men det stabiliserar i mitten av februari. Det olastade malmtågets tidsavvikelser ligger stabilt i spannet -50;+50 minuter bortsett från ett mindre antal tillfällen då avvikelserna är större. "Paxtåg" avser resandetåg.

3. Trafikstyrning: Utveckling av modeller och metoder

3.1. Modeller av järnvägstrafik i tid och rum

Detta delkapitel syftar till att kort redogöra för de mest förekommande sätten att beskriva och matematiskt definiera tågrörelser och dess inbördes beroende och begränsningar baserat på infrastrukturella villkor.

Det matematiska problemet att konstruera eller revidera en tågtidtabell genom att fördela spårresurser till tåg i form av "tidsslots" beskrivs ofta som ett s.k. *job shop scheduling problem*³ (*JSS*) där jobben som ska utföras på olika maskiner istället representeras av att tågen ska utföra sina resp. deltransporter (en deltransport betraktas som *ett* jobb) på ett visst segment (en maskin). Varje tåg ska därmed utföra en mängd jobb i en given sekvens på en mängd maskiner. Vi kallar varje sådan deltransport för en *event* (eller slot/tågläge). Varje tåg har då en sekvens (dvs. en lista) av *events* som ska utföras i en logisk ordning. Den första *eventen* motsvaras då av det uppehåll tåget har på sin avgångsstation eller bangård, medan *eventen* därefter då motsvaras av den första linjesträcka från avgångsstation till nästkommande station (eller linjesträcka, om det finns en växel emellan) på just det tågets rutt på järnvägslinjen. När ett tåg lämnar ett segment (dvs. avslutar motsvarande *event*) så ansluter tåget omedelbart ett annat segment (dvs. påbörjar nästa *event*). Beroende på hur man väljer att definiera tidssepareringen mellan tåg som använder samma (eller angränsande) spårresurser så kan det krävas en viss tidsfördröjning mellan att ett tåg lämnat spåret till dess att nästa tåg får tillträde. I vissa fall kan tidsfördröjningen vara inkluderad i gångtiderna och då är tidsfördröjningen "noll".



Figur 4. Ett tågläge/"en slot" kan betraktas som ett tidsfönster inom vilket ett specifikt tåg ska använda ett specifikt spår på en specifik bansträcka. Vi benämner det en "event". Dess längd anger dess varaktighet (körtiden) och är dynamisk men med ett minimivärde (den teoretiska minsta gångtiden på sträckan). Dess reella längd blir större om det uppstår köbildning och tåget får invänta på spåret till dess att fortsatt tilldelad tågväg är lagd.

³ Shi Qiang Liu and Erhan Kozan. "Scheduling trains as a blocking parallel-machine job shop scheduling problem". In: *Computers & Operations Research* 36.10 (2009), pp. 2840–2852. ISSN: 0305-0548.

Det finns två primära aspekter att överväga vid modellering av järnvägstrafik och det är (1) hur *tid representeras* – som kontinuerlig eller diskret – och (2) hur infrastrukturen (järnvägsnätet) och dess *kapacitet modelleras*.

Huvudsakligen används två olika sätt för att representera tid. Nedan beskrivs kort de två alternativa sätten att matematiskt beskriva tidsallokeringen av individuella spårresurser.

Modell baserad på diskret tid

Tid kan betraktas som en diskret enhet och resurs så att exempelvis varje klockminut per spår är en resurs som kan tilldelas endast ett tåg. För ett tåg med en gångtid på tio minuter ska då tio sammanhängande spårminuter allokeras och starttiden är då den första av dessa minuter som tåget får tilldelat sig:

Låt $x_{i,t,m}^*$ vara en binär⁴ variabel som anger om tåg i använder spår t under minut m , dvs. $x_{i,t,m}^* = 1$ vid användning, annars 0. Låt $x_{i,t,m}^{begin}$ vara en binär variabel som anger om tåg i startar på spår t under minut m ($=1$) eller ej ($=0$). $d_{i,t}$ är minsta gångtiden för tåg i på spår t . Låt T representera mängden av tåg så att tåg i tillhör mängden T ($i \in T$), mängden spår betecknas S så att $t \in S$ och M representera mängden minuter (dvs. avsedd tidsperiod räknat från en starttidpunkt, T_0) så att $m \in M$. Låt oss även för enkelhetens skull anta att samtliga tåg ska trafikera samtliga delar av en enkelspåriga sträcka som består av $|S|$ sammanlänkande spår. Följande gäller då:

Tåg i måste starta vid en, och endast en, tidpunkt m på spår t :

$$\sum_{m \in M} x_{i,t,m}^{begin} = 1, \forall i \in T, t \in S \quad (1)$$

Tåg i måste använda spår t minst $d_{i,t}$ minuter i följd:

$$\sum_{e=0}^{(d_{i,t}-1)} x_{i,t,(m+e)}^* \geq d_{i,t} * x_{i,t,m}^{begin}, \forall i \in T, t \in S, m \in \{1, \dots, (|M| - d_{i,t})\} \quad (2)$$

Maximalt ett tåg får använda spår t under minut m :

$$\sum_{i \in T} x_{i,t,m}^* \leq 1, \forall t \in S, m \in M \quad (3)$$

Detta sätt att representera tid kan - beroende på hur man väljer att diskretisera tiden (exempelvis till en, två eller upp till 'n' minuter) - kräva en mycket stor mängd binära variabler. Å andra sidan krävs inga ytterligare explicita villkor och variabler för att hantera förbigångar och möten. Utöver villkoren ovan krävs en mängd andra villkor som t ex säkerställer att varje tågs events sker i angiven ordning och att nästkommande event i sekvensen direkt avlöser den föregående.

⁴ Binär variabel avser en variabel som bara kan anta värdet '0' eller '1'.

Modell baserad på kontinuerlig tid

Det andra vanligt förekommande sättet att modellera tid är att definiera en kontinuerlig variabel starttid per tåg och spår och en kontinuerlig variabel stopptid samt precis som innan en minsta möjliga gångtid, d . Ett tåg kan då inte få en starttid som inträffar mellan ett annat tågs start- och stopptid utan måste starta före eller efter och måste avsluta före nästa tåg startar (om tågen använder samma spår). Vidare finns ofta även en minsta separeringstid (fördröjning) mellan att ett tåg lämnar ett spår och ett annat inkommer.

Låt nu $x_{i,t}^{begin}$ vara en kontinuerlig variabel (dvs. den kan anta decimalvärden) som anger starttiden för tåg i på spår t , $d_{i,t}$ är minsta gångtid och $x_{i,t}^{end}$ är en kontinuerlig variabel som anger stopptiden, dvs. när eventen för tåg i på spår t ska avslutas. Låt mängden av tåg vara T , så $i \in T$, och låt mängden spår vara S så, $t \in S$. Följande gäller då:

Tåg i måste använda spår t minst $d_{i,t}$ minuter:

$$x_{i,t}^{end} - x_{i,t}^{begin} \geq d_{i,t}, \forall i \in T, t \in S \quad (4)$$

Tåg i måste lämna spår t innan tåg v inkommer eller tvärtom (om de använder samma spår) och det måste finnas en minsta separeringstid, Δ_t , mellan varje par av tågrörelser som använder spår t . Då gäller alltså:

$$x_{v,t}^{begin} - x_{i,t}^{end} \geq \Delta_t, \forall i, v \in T : i \neq v, t \in S \quad (5a)$$

Eller

$$x_{i,t}^{begin} - x_{v,t}^{end} \geq \Delta_t, \forall i, v \in T : i \neq v, t \in S \quad (5b)$$

Eftersom tågordningen bör vara valbar vid en omplanering behöver vi införa en ny variabel som medger att både villkor (5a) och (5b) ska gälla fast inte samtidigt. Vi inför därför en binär variabel, $y_{i,v,t}$ som får värdet '1' om tåg i använder spår t före tåg v , och värdet '0' om motsatt ordning förekommer. Villkoren (5a) och (5b) ersätts då av villkor (6) och (7) nedan, där 'M' nu istället bara är en stor konstant.

$$x_{v,t}^{begin} - x_{i,t}^{end} \geq \Delta_t * y_{i,v,t} - M * (1 - y_{i,v,t}), \quad \forall i, v \in T : i \neq v, t \in S \quad (6)$$

$$x_{i,t}^{begin} - x_{v,t}^{end} \geq \Delta_t * (1 - y_{i,v,t}) - M * y_{i,v,t}, \quad \forall i, v \in T : i \neq v, t \in S \quad (7)$$

Den typ av "disjunktiva" villkor (dvs. "antingen eller"-villkor) benämns ofta som "big-M constraints" i forskningssammanhang och dessa medför ibland att problemen blir svåra att lösa för kommersiella standardlösare såsom Cplex och Gurobi.

Utöver de ovan formulerade villkoren tillkommer givetvis en mängd andra villkor, vilka diskuteras delvis i avsnitt 3.3. och i Bilaga 1.

Modellering av järnvägsnätets resurser

Det finns en mängd olika sätt att modellera järnvägskapacitet men det är främst tre olika principer som är vanligt förekommande. Antingen är noder (stationer och andra mötespunkter) den explicita resurs som allokeras och segmenten däremellan (blocksträckorna, spåren) kontrolleras oftast implicit genom flödet till/från dessa noder.

Det andra sättet är att istället modellerar segmenten explicit och stationerna är endast brytpunkter mellan olika resurser. Kontroll av kapaciteten på stationer och mötespunkterna hanteras vanligtvis då likartat genom representation av in- och utflödet.

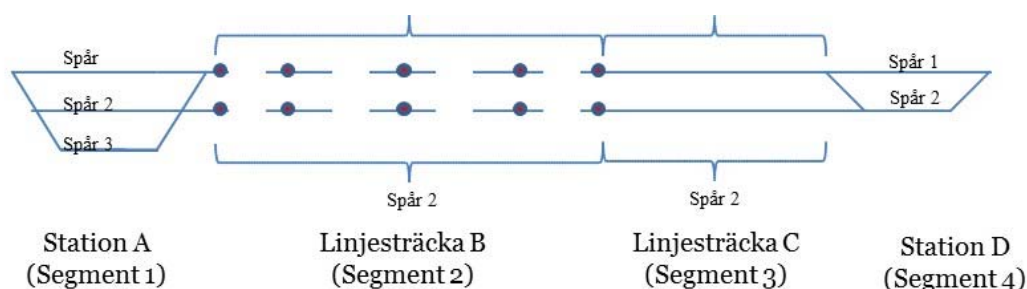
Det tredje sättet är en kombination av de första två principerna och innebär att samtliga resurser modelleras explicit, vilket vi gör i detta projekt.

Nackdelen med de två förstnämnda modelleringsprinciperna (för representation av järnvägsnätet) är att det kan bli svårhanterligt att modellera komplexa nätverk av flera sammanbundna linjer och med oriktad trafik på flerspåriga segment. Detta hanteras lättare om samtliga resurser och deras relationer modelleras explicit. Nackdelen med denna tredje princip är å andra sidan att det generellt krävs en större mängd binära och heltalsvariabler än för de två första alternativen vilken kan medföra en längre beräkningstid eftersom antalet möjliga kombinationer sannolikt blir större.

Sedan är det naturligtvis ytterst en praktisk, viktig fråga om hur detaljerat man vill, kan och behöver beskriva och särskilja de olika resurserna och beroenden givet de data och den information man har tillgång till och kan använda.

3.2 Om projektets problembeskrivning och modell

De modeller som används i projektet använder kontinuerlig tid och vi delar in infrastrukturen i olika *segment* där ett segment representerar en sträcka från en punkt till en annan och kan vara en station liksom en linjesträcka mellan två stationer. Vad som karakteriserar ett segment är att det består av ett eller flera parallella spår. Varje spår kan vara uppdelat i en eller flera blocksträckor där linjesträcka B består av flera blocksträckor och kan därmed ha tågen tätare (begränsade av ett minsta headway-avstånd) än linjesträcka C som bara har en. Vi har använt data från Trainplan där segmenten då är av typen *Network Links* (NWK) eller *Locations* (LOC).

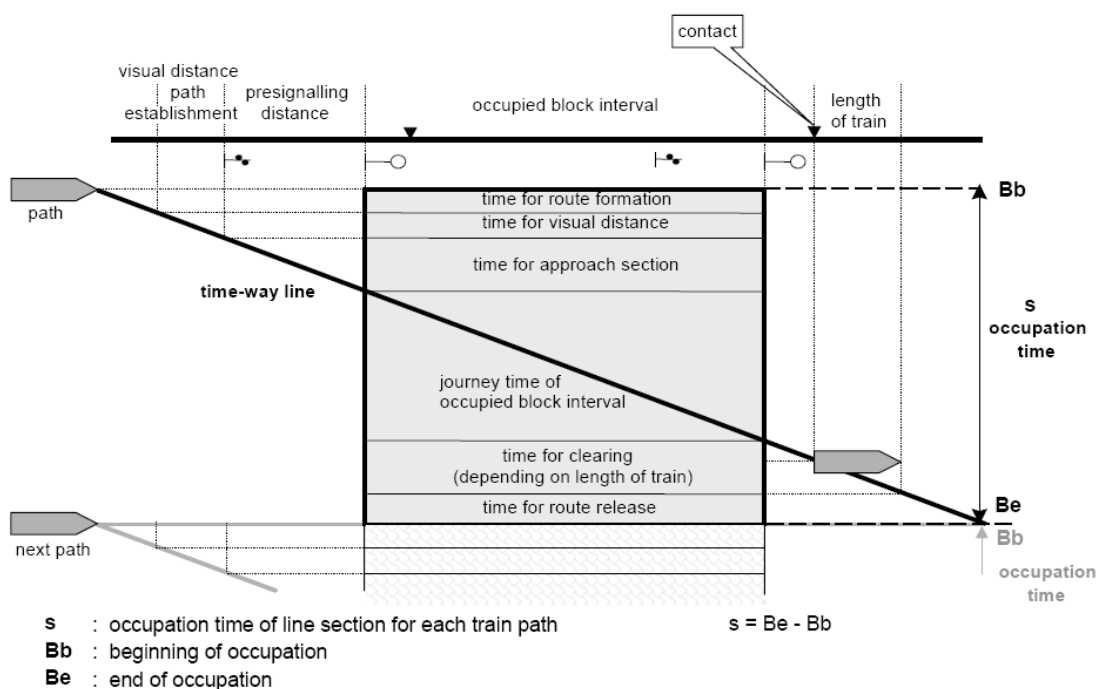


Figur 5. Modellering av infrastrukturen som enskilda resurser.

Enligt Trainplans datastruktur är då samtliga events av typen *Train Movement* (TMV, som sker på en NWK) eller *Train Stop* (TSP, som sker på en LOC). För det enkla exemplet i Figur 4 så har Tåg A en lista med tre olika events som ska ske i kronologisk ordning: {event 1 på Stordalen-Kaisepakte; event 2 på Kaisepakte; event 3 på Kaisepakte-Stenbacken}.

Motsvarande eventlista för Tåg B blir då {event 1 på Kaisepakte-Stenbacken; event 2 på Kaisepakte; event 3 på Stordalen-Kaisepakte}.

En utmaning är att baserat på befintlig tidtabellsdata om tågen, tågens interaktion sinsemellan samt med infrastrukturen definiera de olika tidskomponenter som ska gälla även för annan tågordning och ev. andra, oplanerade uppehåll än de som den fastställda tidtabellen baseras på. I Figur 6 nedan illustreras de olika tidskomponenterna som en tidsslot består av i praktiken. Om t ex tidsseparering mellan tågen redan finns inlagd som del av tågens planerade gångtid så ska ingen ytterligare tidsseparering fodras i villkoren, medan om det inte finns med så ska en explicit minsta "clear time" tillämpas i villkoren. Då vi baserat enkom på Trainplandata inte kan extrahera all information vi skulle behöva för detaljerade dynamiska gångtider, så har vi för experimenten extraherat de komponenter vi kan utifrån de tidstillägg som anges och kräver en explicit "clear time" mellan tåg som använder samma resurser samt kalibrerat tidtabellen så den är tillåten innan vi inför störningar i experimenten. I en praktisk tillämpning kan man istället behöva ett gångtidsverk som tillhandahåller bättre approximationer av tågens gångtider baserat på aktuell plan.



Figur 6. Visualisering av olika tidskomponenter i en tidsslot, vilka dock kan vara svåra att återskapa utifrån befintlig tidtabellsdata. Källa UIC406, Juni 2004.

För att underlätta beskrivningen av modellen i Bilaga 1 ges en övergripande beskrivning av notationen nedan. Notera att nedan används delvis andra variabelnamn och annan indexering än de som används i avsnitt 3.1., där en viss förenkling gjordes för att ge en lättförståelig, övergripande bild av modellens grundläggande element. Nedan notation följer den vi använt tidigare för att beskriva mer komplicerade modeller och är densamma som återfinns i de forskningsartiklar vi publicerat.

Låt T vara vår mängd tåg, B är mängden av segment, och E mängden av events där en event är tidsperioden då ett specifikt tåg planerar att köra på ett specifikt segment. Låt index i vara associerat med ett tåg, index j med ett segment och index k med en event. Varje event tillhör både ett tåg och ett segment, dvs. den representerar ett tåg på ett segment. Låt $K_i \subseteq E$ vara den

sorterade mängden av events för tåg i ($i \in T$) och $L_j \subseteq E$ vara den sorterade mängden av events för segment j ($j \in B$). Vidare, vi använder $(k+1)$ för att representera den första event som följer efter event k medan \hat{k} är någon av de event som följer event k i de sorterade mängderna, så att $k < \hat{k}$ gäller. Varje segment j har ett visst antal spår $P_j = \{1, \dots, p_j\}$. Vi använder index t för att representera ett visst spår.

Problemet handlar om att fastställa en optimal tågordning på resp. segment och givet denna även fastställa start- och sluttider (x_{ik}^{begin} resp x_{ik}^{end}) för varje event k . Tilldelningen av dessa tider ska givetvis beakta de infrastrukturella begränsningarna som finns i form av att varje event ska tilldelas ett spår. Om event k använder spår t så antar den binära variabeln q_{ikt} värdet 1, annars 0. Om två tåg planerar använda samma spår måste de dessutom separeras i tid, vilket hanteras av de binära variablerna $\gamma_{k\hat{k}}$ (som antar värdet '1' om event k ska ske före event \hat{k} , annars 0) och $\lambda_{k\hat{k}}$ (som antar värdet '1' om event \hat{k} ska ske före event k , annars 1). Vid enkelspårig sträcka behövs inte q_{ikt} eller $\lambda_{k\hat{k}}$ utan endast $\gamma_{k\hat{k}}$ eftersom det är uppenbart att tågen använder det enda spåret som finns. För stationer på enkelspåriga linjer, såsom Malmbanan, tillkommer en ytterligare binär variabel för att hantera eventuell otillåten samtidighet, men denna beskrivs inte ovan utan vi hänvisar till Bilaga 1 för ytterligare detaljer.

Vi skiljer också på två olika fall där tågen behöver separeras om de använder samma spår på en linjesträcka. Det ena fallet är (i) när tågen kör i samma riktning på ett spår som har flera blocksträckor och det andra är (ii) när två tåg kör i olika riktning eller kör på ett segment med endast en blocksträcka.

I det första fallet ska tågens starttid på sträckan separeras av ett minsta tidsavstånd (headway) och likaså dess sluttider. I (ii) måste först det ena tåget lämna segmentet, en viss "clear time" för att släppa tågvägen ska passera, och först därefter får nästa tåg starta på det segmentet. Använder tågen inte samma spår (och tågvägarna inte på annat sätt är i konflikt) så spelar tågens inbördes ordning ingen roll.

I Figur 4 tar vi inte upp events som sker på stationerna men i våra modeller behandlar vi events på stationer på i princip samma sätt som de på linjen mellan stationerna, med undantag för stationer som inte tillåter samtidig infart.

De variabler som nämns ovan ingår sedan i en mängd ytterligare villkor som begränsar vilka lösningar som blir tillåtna. Av de lösningar som är tillåtna söker lösningsmetoden (i det här fallet en kommersiell optimeringsmjukvara) efter den lösning/de lösningar som är bäst utifrån en vald målfunktion. En målfunktion kan exempelvis vara att minimera avvikelser från fastställd tidtabell och i detta arbete har vi huvudsakligen valt att 1) minimera förseningar vid slutstation som överstiger 3 minuter och b) minimera förseningar som överstiger 3 minuter vid kommersiella uppehåll och vid slutstationen. Detaljer kring optimeringsmodellen återges i Bilaga 1.

3.3 Utförda experiment och resultat

Som tidigare nämnts har arbetet fokuserat på förutsättningar under T14. Vi har huvudsakligen fokuserat på 20 olika scenarier som beskrivs i tabellen nedan och som baseras på

förutsättningarna som gällde den 28 maj 2014. Den trafik som är inkluderad är dock den som fanns planerad enligt version 1.0 av T14 (dvs. den fastställda tidtabellen när den publicerades i september 2013), vilket innebär att det finns vissa skillnader mellan den som planerades initialt och den plan som togs i drift. Banarbeten som inkluderats är dock de som framgår av den dagliga grafen som kördes ut dagen innan (27 maj 2014).

Scenario	T_0	Scenario
1	14.41	Tåg 9911 avgång Riksgränsen kl 14:20 istället för kl 14:52. Tåg 9909 avgång Bergfors 14:46 istället för kl 14:40. Tåg 9914 avgång Bergfors 14:46 istället för kl 14:10.
2	14.41	Tåg 9909 avgång Bergfors 14:46 istället för kl 14:40. Tåg 9914 avgång Bergfors 14:46 istället för kl 14:10.
3	16.05	9913 vill avgå 16.10 istället för 16:31 från RGN
4	16.05	9913 vill avgå 16.15 istället för 16:31 från RGN
5	16.25	9913 avgår lite sent 16.46 istället för 16:31 från RGN
6	13.20	9914 blir minst 10 min sent fr. KMB.
7	13.20	9914 blir minst 15 min sent fr. KMB.
8	13.20	9914 blir minst 20 min sent fr. KMB.
9	13.20	Tåg 93 är försenad ankomst till RGN kl 13.45 (kl istället för 13:29)
10	13.20	Tåg 93 är försenad ankomst till RGN kl 13.55 (kl istället för 13:29)
11	10.30	Tåg 96 avgår sent kl 10.45 från Luleå istället för kl 10.36
12	10.55	Tåg 96 avgår sent 11.12 från Boden istället för kl 11.02
13	09.34	Tåg 7155 avgår sent kl 9.50 från KMB (9.41)
14	10.25	Tåg 7155 ankommer sent till LIN kl 10.45 istället för 10:33
15	13.30	Tåg 93 får stopp 45min mellan stationerna Riksgränsen-Katterjåk pga ren på spår
16	12.42	Tåg 9909 får stopp 45min mellan stationerna Riksgränsen-Katterjåk pga ren på spår
17	13.35	Tåg 93 får stopp 45min mellan stationerna Katterjåk-Vasserjaure pga ren på spår
18	12.44	Tåg 9909 får stopp 45min mellan stationerna Katterjåk-Vasserjaure pga ren på spår
19	12.10	Tåg 7155 får stopp mellan Lakaträsk och Gullträsk på 45min
20	13.30	Tåg 9231 får stopp mellan Lakaträsk och Gullträsk på 45min

Tabell 1. Beskrivning av scenarier.

Varje scenario utgör ett experiment där planerad trafik och banarbeten inom ett tidsfönster på 4 timmar inkluderas och definierade tåg störs initialt. Omplaneringsmetoden syftar till att konfliktreglera trafiken och minimera förseningarna som uppstår pga. störningen och hur den eventuellt sprider sig till andra tåg, beroende på hur trafiken planeras om. Detta problem modelleras matematiskt via modellspråket AMPL och löses sedan med hjälp av den kommersiella optimeringsmjukvaran CPLEX 12.5 från IBM med upp till 8 trådar på en server med 4 processors (2 GHz), 24 GB of RAM och GNU/Linux 3.2.0-x86-64.

Den optimerande omplaneringsmetoden kan modifiera och ta beslut om:

- Tågens ankomst- och avgångstider
- Spårval på linjen och stationer (inkl. plattformsväl)
- Tågordningen (inkl. möten och förbigångar).

Metoden tar dock inga beslut om omledning, inställelse och tidig vändning av tåg utan dessa typer av eventuella beslut hanteras endast om det ges som input (från trafikledaren) till problemformuleringen. Eftersom just denna fallstudie baseras på en enkelspårig bana sker naturligtvis inget spårval på linjen.

Vi har ett antal huvudsakliga antaganden:

- 1) Tidtabellstekniska uppehåll kan förläggas på annan (tillåten) driftplats eller utebli och gångtiderna justeras då för de uteblivna behoven av tidstillägg som ges av planen.
- 2) Tidstillägg för oplanerade uppehåll: Malmtågen får ett tidspåslag med 2+2 minuter för att kompensera inbromsning samt acceleration. Detta tidspåslag läggs på minimigångtiden för den specifika linjesträckan och det specifika tåget.
- 3) Lastade malmtåg tillåts stanna på driftplatserna (om spårets längd tillåter detta vid möten/förbigångar) men det är förenat med de tidspåslag som framgår av 2).
- 4) Tidseparering vid användning av samma spår på driftplats resp. linjen: 30s mellan utfart av tåg A och infart av tåg B.
- 5) Förbjuden samtidig infart på samtliga driftplats: Infarterna ska tidsepareras med 2 minuter om det är ett persontåg som ska stanna och 4 minuter om det är ett godståg som ska stanna för möte⁵.

Antagandena ovan och gångtider som används baseras på observationer av hur tågen framförts via Malmbanans motsvarighet till "Trafikbilder" samt baserat på data från Trainplan, LUPP och OPERA. Det finns ytterligare praktiska kriterier för t ex när, var och hur vissa tåg inte bör stanna men dessa har inte inkluderats i de modeller vi tillämpat här och detta pga. att vi inte hade tillgång till den information som krävs för att definiera dessa riktlinjer.

Vi har som sagt ett tidsperspektiv och en planeringshorisont på fyra timmar från tiden T_0 , som vi betraktar som den tidpunkt då trafikledaren/planeringsmetoden får information om tillståndet i trafiken och ska omplanera trafiken vid behov. Vi planerar för hela sträckan Björnfjell/Riksgränsen-Kiruna-Luleå (inkl. varje driftplats på denna sträcka) men ej bortom dessa gränser. Huvudfokus för studien har varit att dels analysera hur lösningarna som beräkningsmetoden genererar ser ut för att identifiera metodens svagheter, dels analysera metodens prestanda avseende optimalitet och svarstid (dvs. hur snabbt metoden levererar en optimal lösning).

Utifrån experimenten kunde flera intressanta observationer göras som gav bra underlag till diskussioner med trafikledningen i Boden om vad som karakteriserar en bra resp. mindre bra lösning och vilka prioriteringar som är acceptabla, eller inte, under olika omständigheter.

Ett urval av resultat presenteras i tabellerna och figurerna nedan. De gråmarkerade raderna i Tabell 2 anger de scenarier där de två alternativa målfunktionerna står i viss konflikt, medan de vita är de fall där det är tydligt att det är mest lämpligt att använda TDC+3 istället för TFD+3 baserat på dessa två mått.

⁵ Den tid som ett visst ställverk minst kräver för att separera två tågs infarter (räknat utifrån passagen av en viss infartssignal) på en enkelspårig sträcka där stationerna inte tillåter samtidig infart är i praktiken oberoende av vilka tåg som möts. Däremot kan ett visst tåg behöva bromsa in tidigare än ett annat, vilket medför ett extra tidspåslag för ett visst tåg i förhållande till ett annat. Detta tidspåslag ingår ofta i tågens gångtider som är baserade på hur tågläget är planerat och ses därför inte som ett separat mötestidspåslag. Eftersom våra modeller kräver indata om gångtider som gäller med resp. utan genomfört planerat möte behövs minimigångtid, tidstillägg och krav på tidseparering för möten definieras separat och därför är kraven på tid mellan infarter olika för persontåg resp. malmtåg i våra antaganden. Se även förklaring i kapitel 3.2.

Scenario	Minimering av försening > 3 minuter, dvs. målfunktion Σ TFD+3					Minimering av försening > 3 minuter vid kommersiella uppehåll och slutstation, dvs. målfunktion Σ TDC+3				
	Lösningstid(s) med Cplex	Optimalvärde Σ TFD+3 (s)	Värde på Σ TDC+3 (s)	Specifika tåg försenade +3min (TFD+3)	Antal extrastopp för lastade malmtåg	Lösningstid(s) med Cplex	Värde på Σ TFD+3 (s)	Optimalvärde Σ TDC+3	Specifika tåg försenade +3min (TFD+3)	Antal extrastopp för lastade malmtåg
1	45,86	1545	4449	9914	2	34,8	1545	1545	9914	1
2	34,42	1245	5220	9914	2	38,71	1245	1459	9914	0
3	37,54	0	880	-	0	34,68	0	0	-	1
4	33,72	0	4065	-	0	19,9	0	0	-	0
5	29,78	192	3486	96	-1	21,58	192	576	96	-1
6	17,35	0	2808	-	-1	25,43	0	541	-	1
7	2,69	150	4775	9914	-1	22,87	150	323	9914	0
8	27,3	450	4277	9914	0	16,14	545	545	9914	0
9	21,56	0	6084	-	3	23,25	763	3313	9912 (936), 9914(187)	4
10	22,3	0	11375	-	-1	36,63	28	7343	9912	1
11	1,6	0	656	-	-3	6,53	0	458	-	3
12	2,06	0	645	-	1	9,55	0	420	-	0
13	1,21	0	220	-	0	1,14	0	30	-	0
14	14,55	0	1773	-	-3	1,85	0	948	-	-4
15	60,5	1593	16006	93 (793), 7156 (208), 9912(1132)	2	43,07	1875	13315	93 (667) 9912 (1132) 9914 (616)	2
16	11,34	363	2257	9909	1	22,9	363	363	9909	-1
17	15,75	3225	16634	93(2579) 9912 (1006)	2	34,15	3595	12152	93 (2579) 9912 (1036) 9914 (520)	2
18	24,63	363	7207	9909	0	16,24	363	2883	9909	1
19	2,74	2301	11774	7155	-1	16,86	2315	9262	7155	1
20	21,79	1918	2697	9231	0	22,11	1918	1918	9231	0

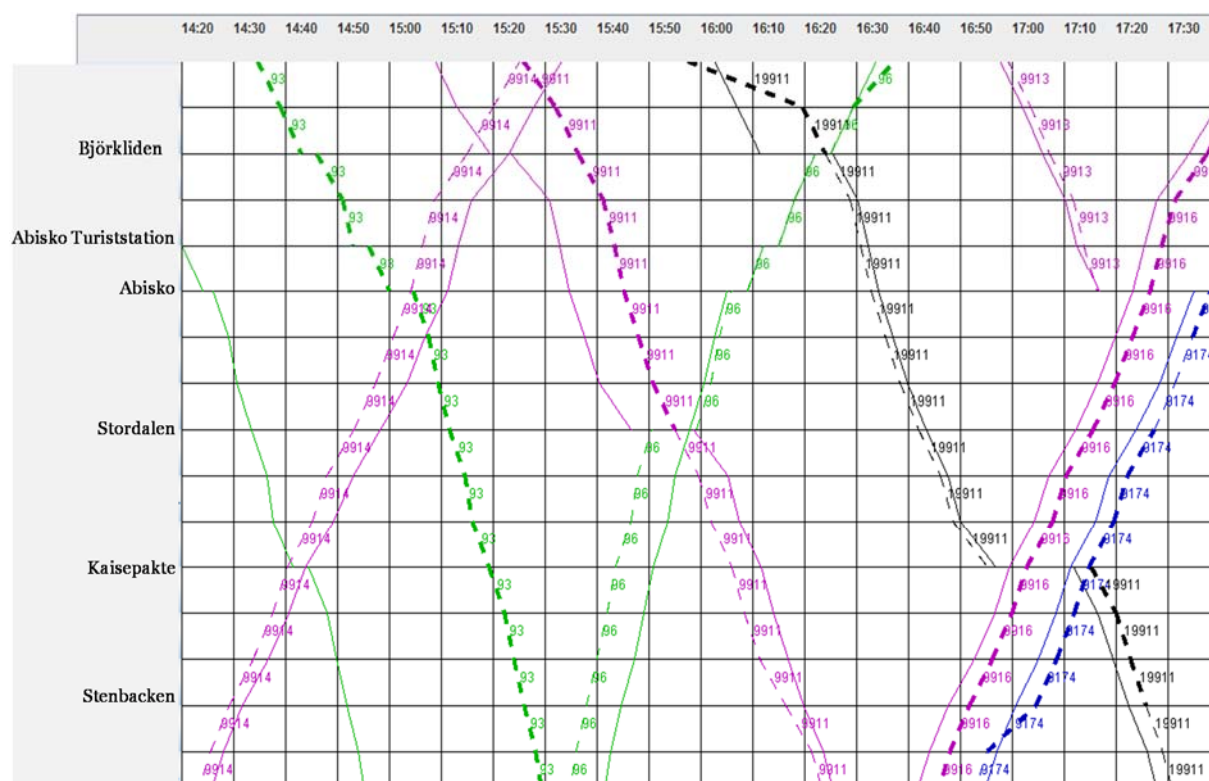
Tabell 2. Den andra kolumnen anger beräkningstiden för att finna optimallösningen. TFD+3 innebär att endast förseningar vid slutstation och som är över 3 minuter inkluderas. TDC+3 inkluderar även förseningar över 3 minuter vid kommersiella stopp (dvs. uppehåll).

För scenario 3 markerar det röda fältet i Tabell 2 att när målfunktionen endast straffar förseningar vid slutstation (TDF+3) så blir det en hel del förseningar vid de kommersiella stopp som finns under vägen.

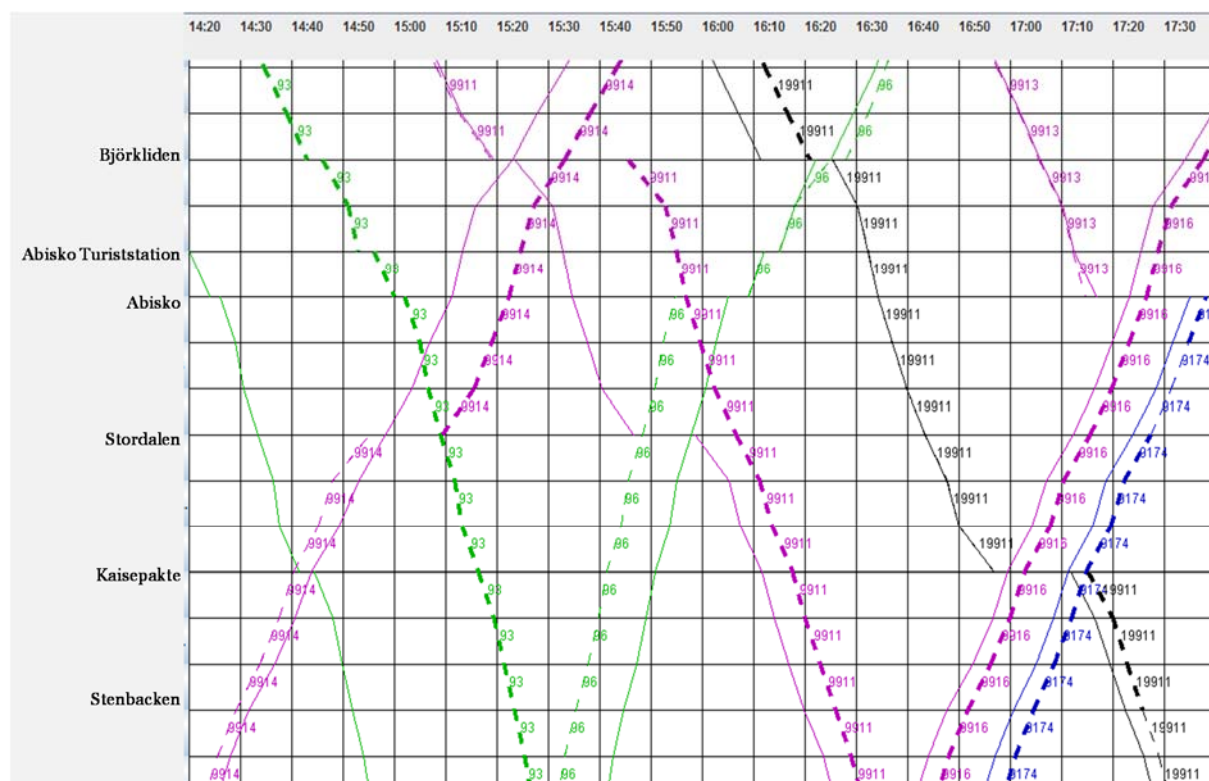
För scenario 9, kan vi observera att efter omplaneringen mhja målfunktionen TFD+3 så får vi inga tåg som anländer till slutstation senare än 3 minuter efter planerad ankomsttid, men det medför istället 3 extra stopp för lastade malmtåg och en hel del förseningar vid de kommersiella stopp som finns under vägen. När vi istället tillämpar målfunktionen TDC+3 som tar hänsyn till förseningar även vid kommersiella stopp, så reduceras förseningarna TDC+3 till hälften men lösningen blir mycket sämre om man ser till antalet tåg som får en slutförsening på mer än 3 minuter och antalet extrastopp för lastade malmtåg.

Det är även relevant att visualisera och studera lösningarna och se vilka tåg som har prioriterats och vilka möten som har flyttats. I figurerna nedan avser heldragna linjer planerat tågläge och streckade linjer avser motsvarande omplanerade tågläge (och en tjockare, streckad linjer indikerar en försening).

I Figur 7 och 8 nedan återges två alternativa lösningar från scenario 15, där vi har ett passagerartåg, 93, som får ett tågstopp på minimum 45 minuter mellan Riksgränsen och Katterjåk. Driftplatserna visas i graferna endast som en linje, dvs. en brytpunkt mellan sekvensen av signalsträckor.

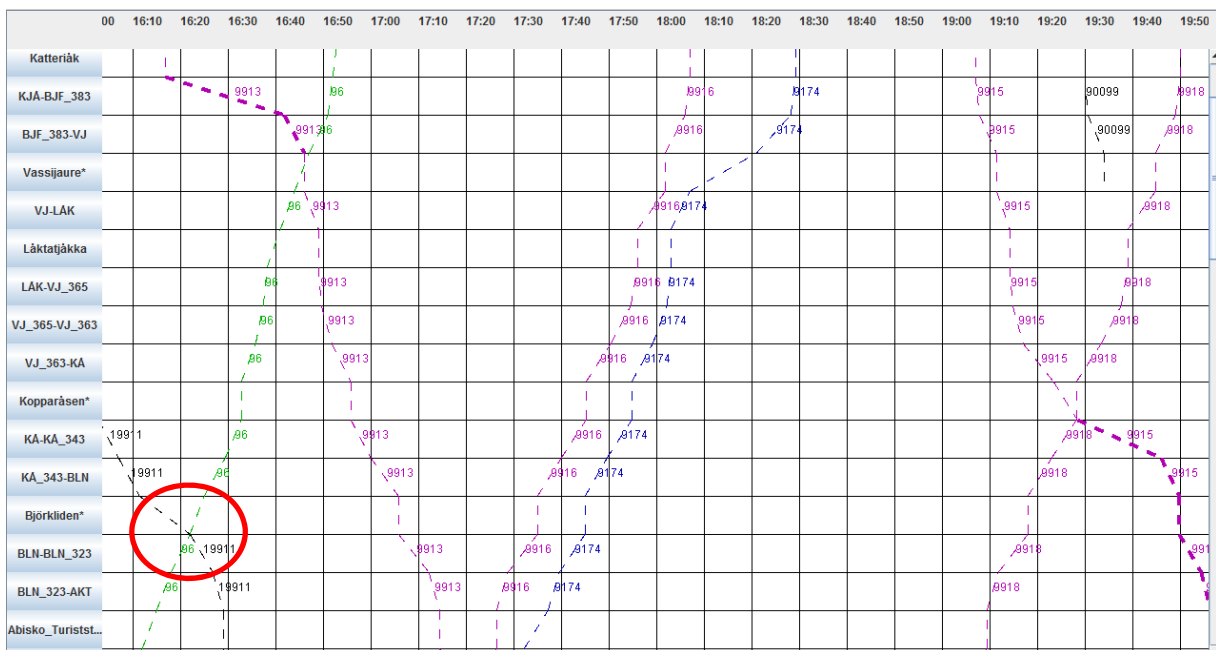


Figur 7. Lösningen som fås med tillämpningen av målfunktionen TFD+3 som minimerar förseningar till slutstation som överstiger 3 minuter. Mötet flyttas från Kaisepakte till Abisko.



Figur 8. Lösningen som fås med tillämpningen av målfunktionen TDC+3 som minimerar förseningar till slutstation som överstiger 3 minuter samt vid kommersiella stopp. Mötet flyttas från Kaisepakte till Stordalen. Tåg 93 hålls fram på bekostnad av 9914 eftersom tåg 93 har flera kommersiella stopp medan 9914 blir totalt 10 minuter sen till slutstation Tåg 9914 påverkar/påverkas av i sin tur 9911 som interagerar med 96.

En annan observation som gjordes var i scenario 3, se Figur 9 nedan där den omplanerade tidtabellen kan ses och där driftplatserna visas explicit. Ett olastat malmtåg, 19911, ska möta ett passagerartåg, 96 i Björkliden där tåg 96 har ett uppehåll. Detta innebär att malmtåget inte behöver stanna utan kan hålla en lägre hastighet och köra genom stationen. Det förutsätter att passagerartåget anländer minst 2 minuter före malmtåget och stannar på sidospåret. Men detta behov – att malmtåget inte bör stanna - speglas inte av den övergripande målfunktionen och därför stannar malmtåget i Björkliden och väntar in passagerartåget, innan det kör vidare (eftersom det inte medför någon betydande tidsförlust). Därför hade det varit relevant att ha en eco-driving funktion som "putsar till" planen för resp. tåg när den övergripande planen för systemet är föreslagen.



Figur 9. Ett snitt från en omplaneringslösning i scenario 3 där planen för det olastade malmtåget 19911 genom stationen Björkliden inte är optimal sett utifrån ett eco-driving perspektiv.

Eftersom experimenten visade att lösningstiderna varierar kraftigt beroende på val av scenario när Cplex användes som lösare, så valde vi att även använda en annan lämplig, kommersiell lösare, nämligen Gurobi version 6.5.1. Vi valde att modellera optimeringsproblemet direkt i java istället för i AMPL eftersom Gurobi tillhandahåller ett användarvänligt java API. Vi använde en fullvärdig akademisk licens installerad på en HP Zbook 15 G2 med en 4-kärnig Intel Core i7 som möjliggör 8 trådar. Detta har medfört en signifikant uppsnabbning men framför allt medför detta en förenklad datahantering eftersom vi då använder samma programmeringsspråk för alla programdelar (filtrering av tidtabellsdata, scenariogenerering, problemlösning och visualisering av resultat) och blir dessutom inte beroende av en mjukvarulicens för modelleringsspråket AMPL.

Nedan återges resultaten från experimenten där Gurobi användes i jämförelse med Cplex och där målfunktionen som minimerar förseningar som överstiger 3 minuter vid slutstation användes. Lösningstiderna för resp. mjukvara anges i kolumn 3 resp. 5. Eftersom det är samma problem som Cplex och Gurobi löser är det samma optimalvärde, vilket anges i kolumn 2. Vi mäter även separat förseningar som överstiger 3 minuter vid kommersiella uppehåll samt vid

slutstation (kolumn 4 resp. 6), och då dessa värden skiljer sig åt markant kan vi se att Cplex och Gurobi hittar olika optimallösningar, dvs. lösningarna är strukturellt olika. Detta är inte speciellt överraskande eftersom det finns en stor redundans i lösningsmängden. Detta betyder att det finns många lika bra lösningar, vilket tydligt visar behovet av att dels ytterligare studera lösningars egenskaper, dels skärpa till problemformuleringen inkl. målfunktionen och mjukare preferensvillkor för att snabba upp lösningsförfarandet ytterligare.

Scenario	Minimering av försening > 3 minuter, dvs. målfunktion $\Sigma TFD+3$				
	Optimalvärde $\Sigma TFD+3$ (s)	Lösningstid (s) med CPLEX	Värde på $\Sigma TDC+3$ (s) med CPLEX	Lösningstid (s) med Gurobi	Värde på $\Sigma TDC+3$ (s) med Gurobi
1	1545	45,86	4449	9,52	6414
2	1245	34,42	5220	3,87	4259
3	0	37,54	880	5,46	1745
4	0	33,72	4065	5,09	2610
5	192	29,78	3486	3,81	3885
6	0	17,35	2808	3,97	3049
7	150	2,69	4775	4,17	4724
8	450	27,3	4277	6,39	5363
9	0	21,56	6084	3,59	8100
10	0	22,3	11375	6,07	8002
11	0	1,6	656	1,61	1442
12	0	2,06	645	1,78	1064
13	0	1,21	220	1,46	1781
14	0	14,55	1773	2,6	1530
15	1593	60,5	16006	4,92	14606
16	363	11,34	2257	3,12	5703
17	3225	15,75	16634	5,78	16558
18	363	24,63	7207	2,68	4734
19	2301	2,74	11774	2,76	14114
20	1918	21,79	2697	3,49	9485

Tabell 3. En jämförelse mellan resultatet från tillämpningen av Cplex resp. Gurobi för att lösa de 20 problemscenarier som definierats. Olika optimallösningar hittades och grönmärkade celler visar vilken lösning som är signifikant bättre baserat på måttet TDC+3, som bara var delvis inkluderat i målfunktionen och därmed inte nämnvärt påverkar lösarens val av lösning.

3.4 Slutsatser

De huvudsakliga slutsatserna från den experimentella studien är följande:

- Med hjälp av kommersiella lösare såsom Cplex eller Gurobi och en matematisk modell så går det att relativt snabbt (2-10 s) att omplanera trafiken på banan Riksgränsen-Kiruna-Boden-Luleå fyra timmar framåt. Resultaten från experimenten visar dock att lösningstiden varierar betydligt beroende på scenariernas egenskaper, vilket skapar en viss osäkerhet.

- Eftersom det är en enkelspårsträcka där tidsavstånden mellan möjliga mötes-/förbigångsstationer är ganska stora finns det mycket marginal i tågens tidtabeller, i synnerhet de olastade malmtågerna. Tidtabellen är därför mycket störningstålig i det avseendet och mjuka effektiv omplanering så kan tidtabellen absorbera ganska stora temporära avvikelser.
- Vilken målfunktion som används har stor betydelse för framför allt de tåg som avgår före sin tidtabell på denna bana. Traditionellt sett så "straffar" man bara förseningar i denna typ av metoder medan det alltså kan finnas behov av att "fösa fram" tåget så länge det inte är i konflikt med annan trafik och orsakar förseningar. Det innebär att man även bör straffa en ökad körtid fast då mycket mindre än förseningar. Vi införde detta tillägg i våra målfunktioner vilket gav bättre lösningar utifrån ett praktiskt perspektiv. Det är dock sannolikt svårt att hitta en målfunktion som fungerar för alla tillfällen utan ett antal parallella processer kan krävas med olika målfunktioner för att på så sätt få ett stabilare system som också kan ta hänsyn till flera olika perspektiv samtidigt. Varje lösnings konsekvenser och egenskaper bör också mätas baserat ett antal relevanta nyckeltal för att enklare kunna jämföra och bedöma lösningarna.

För detaljer kring optimeringsmetoden och experimenten hänvisas läsaren till följande forskningsartikel (som återfinns i Bilaga 1):

Johanna Törnquist Krasemann, (2015), "Computational decision-support for railway traffic management and associated configuration challenges: An experimental study", Journal of Rail Transport Planning & Management, Elsevier, Volume 5, Issue 3, November 2015, pp.95–109.

I artikeln finns även en sammanfattning av forskningsfronten.

4. Fortsatt arbete

För att diskutera dels det arbete som utförts inom delstudien, dels behovet av beslutsstödjande funktioner med trafikledare som har ansvar för Malmbanan, så gjordes ett studiebesök i september 2015. Hans Lahti (erfaren trafikledare och STEG-användare) och Åsa Jonsson (tidigare trafikledare samt nu sektionschef Trafikledning Boden) med kollegor svarade på frågor under en rundvandring och demonstrerade STEGs styrkor och svagheter med nuvarande konfiguration. Även Arne Andersson från Uppsala Universitet deltog. Nedan är en summering av observationer och reflektioner från studiebesöket. Denna summering skickades kort efter studiebesöket för genomläsning till Hans, Åsa och Arne som inte hade något att tillägga eller justera.

Under studiebesöket diskuterades bl.a. på ett övergripande plan vilken typ av funktionalitet STEG erbjuder och hur/när AEF fungerar, dess fördelar och nackdelar.

Vid rundvandring var det tydligt att när tidtabellen importerar i STEG innehåller tidtabellsdatan en hel del felaktiga spårangivelser mm, vilka markerades i gränssnittet som konflikter vilka trafikledaren då måste hantera för att AEFen ska fungera. AEF-en utgår från att alla tåg verkligen följer planen och eftersom det inte finns någon justeringsmekanism eller logiskt regelverk i systemet som skulle kunna upptäcka och korrigera uppenbara behov av justeringar i planen, så kan AEF ibland uppfattas som ett trubbigt verktyg av vissa medarbetare och stängs av om trafiken inte flyter på enligt plan.

STEGs konfliktdetektering på linjen baseras på en grafisk kontroll av att två körtidslinjer inte korsas eller ligger för nära varandra (om tåg kör i kolonn). Konflikthanteringen på stationerna upptäcks på snarligt sätt om två tåg är planerade att använda samma tåg under

samma/överlappande tidsperiod. Däremot upptäcker STEG inte konflikter med avseende på förbjuden samtidighet utan "tillåter" att tåg kör in i princip samtidigt. Det vill säga, det finns ingen kontrollfunktion/logiskt regelverk som säkerställer att tågens infarter måste tidssepareras. Dessa tidssepareringsvärden finns heller inte definierade någonstans utan det är upp till varje trafikledare att ha kännedom om och tillämpa dessa. Vidare, när ett tåg gör ett oplanerat uppehåll eller inte genomför ett planerat uppehåll så ligger gångtidslinjen från den ordinarie tidtabellen kvar och tar därmed inte hänsyn till förändringar i planen m a p förändrad acceleration och retardation. Återigen är det upp till trafikledaren att ha kännedom om dessa tidspåslag/tidsavdrag och manuellt justera detta genom drag-and-drop funktionen eller genom att ändra tågets snitthastighet på det berörda banavsnittet. I nuläget finns det heller inget direkt stöd för att ha koll på hur terrängen i kombination med tågets egenskaper påverkar tågföringen och att exempelvis vissa tyngre tåg inte bör planeras in på ett oplanerat uppehåll där det är kraftigt uppförlutning efteråt (t ex Stenbacken). Arne Andersson nämnde att det finns/fanns en handbok av Per Leander (Transrail) som beskriver några tumregler för tidspåslag mm, men denna verkar inte finnas tillgänglig.

Enligt diskussionen med Hans, Åsa och Arne finns det behov av funktionalitet som mer dynamiskt justerar/uppdaterar planen efter gällande förhållanden, dock med full insyn, medvetenhet och godkännande av trafikledaren. Ett annat exempel på behov som diskuterades var att enkelt, dynamiskt och semi-automatiskt uppdatera ett tågs plan och gångtider mm om det framkommer i driften att tåget har en annan körprofil (t ex lok med sämre dragkraft än planerat som medför ökade gångtider eller lättare last, kortare tågset etc)⁶. Motsvarande ändringar skulle behövas göras för samtliga tåg som passerar ett banavsnitt som akut av trafikledaren markerats med en hastighetsnedsättning till t ex 40 km/h. Då borde automatiskt alla tåg som berörs få en annan gångtid på just det banavsnittet som berörs. Behovet av att göra prel. ändringar för flera tåg eller flera avsnitt samtidigt kan också uppstå pga. väderrelaterade aspekter såsom lövhalka (som är ett problem även för persontågen) eller vid HAU (Hastighetsnedsättning Utan Signalering, max 30 km/h och ordergivning gäller då). Dessa åtgärder skulle dessutom stödja rapporteringen till BASUN och möjligheterna att följa upp tågföringen bättre.

Trafikledarens möjlighet att agera beror mycket på vilken information hen har att tillgå. Exempelvis är det ibland stor osäkerhet gällande vissa tågs faktiska köregenskaper, när tåget/lokföraren planerar att avgå (tidigt, i tid, sent) mm. Det medför dels att trafikledaren behöver planera in en viss flexibilitet i sin plan för att kunna ändra vid behov (om osäkerhet finns), dels att det blir fler omplaneringar än det kanske behövt bli och därmed är det inte lika lätt att dra nytta av AEFens fulla potential. Järnvägsföretagens varierande framförhållning och förmåga att kommunicera status till trafikledaren verkar ha en väsentlig betydelse för trafikledarens möjlighet att agera proaktivt.

En aspekt som kort även togs upp ang. STEG, var att tåg som inte rör sig (de står stilla på en station pga t ex en störning som leder till trafikstopp eller de kommer inte iväg av annan anledning) kan falla under nutidslinjen och i värsta fall glömmas bort om det är mycket för trafikledningen att hålla reda på, vilket motverkar STEGs syfte att stödja trafikledningen. Tåg som rent fysiskt gör anspråk på en resurs borde ha sin tågplan förskjutet i tid och aktuell

⁶ Enligt en kommentar från Trafikverket, på remissutgåvan av rapporten så "är det tänkt att användaren ska använda rotationspunkt för att anpassa till exempelvis ny medelhastighet på hel eller del av sträcka."

position plottas på nutidslinjen ordentligt markerad. Tåg som har en planerad (sen, osäker) ankomst in i området (dvs. tågen uppehåller sig fysiskt utanför grafen) borde också markeras på något snarlikt sätt och hållas kvar på nutidslinjen.

I forskningssammanhang finns en viss förkärlek för att kategorisera störningar i kategorin "Mindre störningar" (Eng. *Disturbances*) resp. "Större störningar" (Eng. *Disruptions*). Vad trafikledarna anser karakteriserar en mindre resp. större försening är därför intressant. Hans ansåg att en mindre, vanligt förekommande störning är exempelvis när det förekommer växelfel, spår på station ska stängas av eller andra fel som dock tillåter att trafiken flyter på till viss del – om än med begränsad framkomlighet. HAU är ett exempel på en sådan störning (dvs. liten störning med stor påverkan). Större fel är när t ex en kontaktledning är nedriven eller rälsbrott uppstått och det blir totalstopp på linjen. Då kan det vara viktigt att det finns framkomlighet för t ex en elbuss som ska avhjälpa felet. Vid störningar har ofta resandetågen hög prioritet och vid uppstart så är det ofta effektivt att tillämpa kolonnkörning ett tag i den ena riktningen, och sedan i den andra riktningen. Framförhållningen, dvs. planeringshorisonten, vid trafikstörningar varierar men när det är större avvikelser/förseningar i Bodenområdet - och därmed viss osäkerhet i informationsunderlaget - så är framförhållningen ca 30-60 min. När det gäller hur nära i tiden större omplaneringar görs (flytta möten/byta tågordningen) så görs de med minst ca 30 min framförhållning. Därmed kan spårbyten göras enklare och senare.

Vi diskuterade även vad som karakteriserar en bra resp. mindre bra plan och vad som är kvalitet i tågföringen. Ett sätt att se på det, enligt Hans, är att tåg som inkommer området ska inte fått en merförsening efter att det lämnat området/nått slutstation. Dock gör trafikledningen ibland temporära försämringar (som medför merförseningar, vilka kodas "driftledningen prioriterar" i BASUN) för att förbättra helheten inom området eller över ett större område. Vidare kommuniceras och bibehålls de prioriteringar som görs för genompasserande tåg av annan trafikledare av olika skäl.

Avslutningsvis; baserat på diskussionerna med personalen i Boden och den enklare analys av trafikens punktlighet som gjorts och summeras i slutet av kapitel 2, så finns det ett betydande behov av beslutsstödjande funktioner för trafikledningen i Boden. Det är dock mer grundläggande stöd som behövs inledningsvis, snarare än mycket avancerade optimeringsmetoder.

Det som därför verkar mest relevant att arbeta vidare med i nuläget utifrån Trafikverkets behov och förutsättningar är två funktioner:

- 1) En leverans kvalitetssäkrande funktion som kan putsa till och trimma planen och konfliktreglera tågplanen mm när den går över till TL från korttidsplaneringen, samt
- 2) En konsekvensanalytisk funktion som i realtid "bedömer" aktuell plan och dess sannolika konsekvenser över tiden samt som ev. föreslår mindre justeringar av planen vid behov.

Det pågår dock just nu flera utvecklingsprojekt internt på Trafikverket där nya IT-verktyg och system utvecklas för kommande implementering, däribland MPK (Marknadsanpassad planering av kapacitet) och NTL (Nationellt tågledningssystem). I dessa projekt förs också diskussioner om aktuella och framtida behov av den typ som diskuterats här.

5. Övrigt material

Under vårterminen 2015 utfördes även två examensarbeten på BTH där handledare var Johanna Törnquist Krasemann och Håkan Grahn. Båda examensarbetena fokuserade på utveckling, tillämpning och utvärdering av effektiva omplaneringsmetoder av tågtrafik på en enkelspärig bana med Malmbanan som utgångspunkt:

- Anton Peterson, (2015) "Train Re-scheduling: A Massively Parallel Approach Using CUDA⁷", Masteruppsats i Datavetenskap, Blekinge Tekniska Högskola.
- *Petrit Gerdovci och Sebastian Boman*, (2015) "Re-scheduling the Railway Traffic using Parallel Simulated Annealing and Tabu Search: A comparative study⁸", Examensarbete i Datavetenskap för civilingenjörer, Blekinge Tekniska Högskola.

⁷ <http://bth.diva-portal.org/smash/record.jsf?pid=diva2%3A868498&dswid=-9522>

⁸ <http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A839041&dswid=-3127>

Computational decision-support for railway traffic management and associated configuration challenges: An experimental study

Johanna Törnquist Krasemann

Department of Computer Science and Engineering, Blekinge Institute of Technology

Postal address: SE-371 79 Karlskrona, Sweden

E-mail: Johanna.tornquist@bth.se, Phone: +46 (0) 455 38 5881

Abstract

This paper investigates potential configuration challenges in the development of optimization-based computational re-scheduling support for railway traffic networks. The paper presents results from an experimental study on how the characteristics of different situations influence the problem formulation and the resulting re-scheduling solutions. Two alternative objective functions are applied: Minimization of the delays at the end stations which exceed three minutes and minimization of delays larger than three minutes at intermediary commercial stops and at end stations. The study focuses on the congested, single-tracked Iron Ore line located in Northern Sweden. A combinatorial optimization model adapted to the special restrictions of this line is applied on 20 different disturbance scenarios and solved using commercial optimization software. The resulting re-scheduling solutions are analyzed numerically and visually in order to better understand the practical impact of using the suggested problem formulations in this context. The results show that the two alternative, objective functions result in structurally, quite different re-scheduling solutions. All scenarios were solved to optimality within 1 minute or less, which indicates that commercial solvers can handle practical problems of a relevant size for this type of setting, but the type of scenario has also a significant impact on the computation time.

Keywords

Railway traffic management, Real-time scheduling, Decision Support, Optimization, Job Shop Scheduling.

1 Introduction

A common challenge for many national railway traffic administrations is to achieve a high punctuality and ensure that the traffic system can provide reliable, attractive freight and passenger transport services. This means that the network managers need to balance the intended traffic load in the network and operator preferences with the desired network stability and timetable robustness. It also means that the occurrence of primary disturbances needs to be avoided as well as consequential knock-on delays that may occur due to congestion in the network. To what extent knock-on delays spread when a disturbance has occurred depends significantly on the ability of the timetable to absorb and recover from delays and how effectively the trains can be re-scheduled. In many railway networks, the re-scheduling is still today done manually without any computational decision-support despite that the potential benefits are evident and that the research efforts in academia as well as in industry have been intensified during the past 15

years. The challenges associated with developing, implementing and applying computational train traffic management support for different levels of decision-making are, however, extensive.

The development of computational real-time railway traffic re-scheduling support is composed of three main challenges:

- a) Modeling and solving the specified re-scheduling problem for various types of scenarios and contexts.
- b) Requirements engineering concerning how to configure the computational support as functionalities that are perceived useful by the traffic managers and in line with the traffic management prioritization regulations.
- c) Handling technical and administrative questions regarding how to incorporate the existing IT-systems and ensuring data availability.

This paper focuses on the first and second aspect and investigates potential configuration challenges in the development of a computational re-scheduling support for a larger, heterogeneous railway traffic network. More specifically, this paper presents results from an experimental study on how the characteristics of different situations and networks influence the problem formulation and the resulting re-scheduling solutions.

The study is based on the current prerequisites of the congested, single-tracked Iron Ore line located in Northern Sweden and partially Norway. A combinatorial optimization model presented in (Törnquist and Persson, 2007) and used earlier for a quite different context and has here been extended and adapted to the special restrictions on this line and solved using the commercial optimization software IBM ILOG CPLEX version 12.5.

The intended contributions of this paper are the proposed model extensions as well as the study of the practical implications of using the selected problem formulation for this particular context. In the study, the resulting re-scheduling solutions are analyzed both visually and numerically in detail - beyond aggregated, average numbers - in order to identify potential weaknesses and issues which cannot be identified merely by considering traditional key performance indicators related to the train delays and their recoverability.

2 Scope and Related Work

The research in the area of computational support for railway traffic scheduling and re-scheduling is significant. A recently published overview can be found in (Cacchiani et.al., 2014). Computational decision-support for re-scheduling of trains encompasses decision-making at various levels in the traffic system. It may concern the computation of the optimal train trajectory from the current train position to the next pre-determined target point in time and space for an individual train. This type of decision-support system is often referred to as Driver Advisory System (DAS), see e.g. Yang et. al. (2013).

The re-scheduling may also, or instead, concern the management of the traffic system and all trains that are planned to interact within a specified time frame and part of the network. The re-scheduling decisions can be divided into re-timing, re-ordering, local re-routing and global re-routing (Hansen and Pachl, 2008). Local re-routing refers to that there are alternative paths (i.e. tracks) for the trains to use on the line between two stations or, through the stations, while global re-routing refers to that the trains can take a completely, or partly, alternative line stretch from their origin to their destination.

One sub-topic related to the re-scheduling in focus here concerns how to conceptually model and mathematically formulate the re-scheduling problem. A common way is to model the train occupation in the network in terms of train events and assign the events a set of time slots for the associated network resources. The problem of deciding i) which resource to assign to each event, ii) in what order different events should be allocated the resources and iii) during which time period, is then commonly formulated as a Mixed Integer Linear Problem (MILP) with continuous time. See e.g. one of the earlier models proposed by Carey (1994). Several extensions and other MILP formulations have later been proposed to represent more complex networks.

The re-scheduling problem is often denoted as a Job Shop Scheduling Problem with "no-wait constraints" and this has been modeled as an Alternative Graph (Mascis and Pacchiarelli, 2002) and formulated as a MILP assuming that the available alternative train paths are pre-generated and that the make span is minimized (i.e. the approach minimizes the finishing time of the latest event). This approach has iteratively been extended and improved in various ways, see e.g. (D'Ariano et. al., 2008), (Mannino and Mascis, 2009) and (Kecman et. al., 2013). In contrast to these approaches where there is a minimization of the worse case (i.e. the maximum experienced consecutive delay) implying that the number of delays below that threshold is unregulated and not penalized, several other objectives are considered. Commonly used objectives are for example minimization of all train delays with possibilities to weight them differently and minimization of passenger delays in different ways, see e.g. (Törnquist, 2007), Corman et. al., 2012), (Sato et. al., 2013) and (Dollevoet et. al., 2014). There are also some recent studies investigating the application of multiple objectives and criteria, see e.g. (Sama et. al., 2015).

Even though the majority of the proposed formulations use a continuous time representation, there are also several researchers who use discrete time, e.g. (Caimi et. al., 2012), and (Meng and Zhou, 2014).

Many researchers use commercial solvers to solve the formulated problems. The capabilities of commercial solvers to handle these types of problems depend heavily on the problem structure, the mathematical problem formulation and the size of the problem instances. See e.g. results from comparative studies presented in (Törnquist, 2007) and (Harrod and Schlechte, 2013). When the commercial solvers do not provide solutions sufficiently fast, which is often the case for larger networks and time frames of 60 minutes or more, several researchers resort to a rolling time-horizon approach (see e.g. (Törnquist, 2007); (Meng and Zhou, 2011); (Quaglietta et. al., 2013); (Pelligrini et. al., 2014)), various heuristics (see e.g. (Corman et.al., 2010)), parallelization approaches (see e.g. Iqbal et.al. (2013)), or decomposition schemes (see e.g. (Lamorgese and Mannino, 2015)). Another important aspect of real-time train traffic re-scheduling is how to handle the transition and coordination between multiple dispatching areas, see e.g. (Corman, et. al., 2014), but that is beyond the scope of this paper.

This study focuses on a single-tracked line with heterogeneous traffic. Related approaches applied on single-tracked lines are proposed by e.g. Zhou and Zhong (2007), Zaninotto et.al. (2013), Mannino (2011) and Lamorgese and Mannino (2015).

Zhou and Zhong (2007) propose a job shop scheduling formulation where the total completion time under different release times is minimized. The corresponding integer programming problem minimizes the total travel time of the trains and is solved by a branch-and-bound solution procedure where lower bounds and upper bounds are computed by Lagrange relaxations and constructive heuristics, respectively.

In (Zaninotto et. al., 2013), the re-scheduling problem for an Italian single-tracked railway line is formulated as an Alternative Graph formulation as in e.g. (D'Ariano et. al.,

2008). The running times on the lines are fixed, while the dwell times are variable, and the objective is to minimize the maximum consecutive delay.

In (Mannino, 2011), a MILP formulation for the single-track re-scheduling problem is proposed based on a previous formulation presented in (Mannino and Mascis, 2009) and developed initially for the Milan metro stations. Two alternative formulations to model the capacity restrictions on stations were proposed and benchmarked: A “non-compact” formulation which counts the number of pairs of trains that simultaneously meet at a specific station and then ensures this respects the capacity limit, and a “compact” formulation where the station track occupancy by each train is modeled explicitly. The objective is to minimize costs associated with deviations from the master timetable (earliness and tardiness) as well service irregularities. The problem formulations were solved using CPLEX 12.2. The approach has later been refined and put into operation for several single- and double-tracked regional railway lines in Italy. The problem formulation has been even further extended, see (Lamorgese and Mannino, 2015), to support the use of a decomposition scheme based on Benders’ decomposition and solved using CPLEX combined with self-developed heuristics. This approach was put into operation in the Stavanger control center in Norway in February, 2014.

The type of re-scheduling performed in this study does not include any DAS component and it does not include global re-routing since it is single-tracked line. That is, the sequence of stations that each train is planned to visit is pre-determined (but it differs between the trains) but the selection of track and path through the stations is variable. The problem formulation uses a continuous time representation. Running times are variable (in contrast to e.g. Zaninotto et. al. (2013) where only dwell times at stations are variable) and those are based on predefined minimum run time values and time supplements that are enforced depending on if trains are re-scheduled to stop, or not. The dynamically added time supplements are approximations based on empirical data rather than computed during the re-scheduling process iteratively based on detailed train-specific speed profiles and infrastructure characteristics. The variable running times and dwell times are naturally also affected by the availability of the tracks, and will increase as a result of any congestion on the line, or within the stations during the re-scheduling. The objective is to minimize the total final delay exceeding three minutes when trains arrive at their final destination, or to minimize the delays exceeding three minutes at intermediary, scheduled commercial stops as well as the final destinations. The problem formulation is then solved using CPLEX 12.5. One advantage of using commercial software instead of tailored algorithms is that the flexibility to dynamically modify the problem formulation in terms of objective function, constraints and critical parameter settings increases. In contexts, where the applicable goal function and preferences vary depending on the traffic situation it may be useful, or even necessary, to allow certain modifications of the formulated model. However, certain modifications may lead to that the problem becomes so different that the solver has difficulties solving the re-formulated problem. In this study, we therefore also analyze the capability of the applied commercial solver to handle a variety of problems and problem formulations.

3 The Iron Ore Line

The Iron Ore line in the Northern part of Sweden serves as a critical transport link between the mines around Kiruna (Sweden) and the ports in Narvik (Norway) and Luleå (Sweden). This single-tracked line is 500 km long, electrified with line blocking and has a maximum permitted axle load (STAX) of 30 tonnes and a maximum weight (STVM) of

12 tonnes/meter. A large share of the traffic consists therefore of very heavy iron ore trains. These trains are approximately 750m long and generally weigh 8160 tonnes when loaded (1470 tonnes unloaded) compared to the other freight trains of e.g. 430m length and 3400 tonnes weight when loaded that run on that stretch.

The line is divided in two parts: The Northern part runs between Kiruna (via the Swedish border station Riksgränsen) and Narvik, and transports around 15 million tonnes net of iron ore per year. The Southern part runs between Kiruna- Gällivare-Boden-Luleå and transports around 7 million tonnes net of iron ore per year, according to Trafikverket. The railway traffic on that line is also an important transport service for people living in the region and to accommodate the significant number of hiking tourists that visit Abisko and its surroundings. The traffic load on this line is significant and it is classified as "red" with a capacity utilization in the interval 81-100% according to the UIC406 capacity assessment by Trafikverket. The most congested parts are the stretches Gällivare-Kiruna and Kiruna-Riksgränsen, where the later has around 30 trains per 24h-day where 1/5 are passenger trains, 3/5 are iron ore trains and 1/5 are other freight trains. An illustration of the Iron Ore Line is depicted in Figure 1.

Since this railway line is single-tracked with large time distances between meeting locations that enable longer freight trains to meet and it serves very heterogeneous traffic, the ability to re-schedule and manage the traffic when deviations from the initial timetable occurs is critical. Today this re-scheduling is done manually with no computational support to find alternatives solutions or to assess the impact of planned actions. Currently the traffic management that serves the Iron Ore Line from the dispatching center in Boden has a visual, interactive digital timetable graph that is connected to the driver assistant system (CATO) which is installed at the Iron Ore trains operated by LKAB. The CATO system provides the dispatchers with the actual position of the trains and the dispatchers provide the CATO system with target points for the trains. This information is exchanged via the digital graph. The dispatchers have, however, no computational support to re-schedule the traffic and compute the corresponding target points, or to assess the implications of the intended re-scheduling actions or any other change of the traffic conditions.

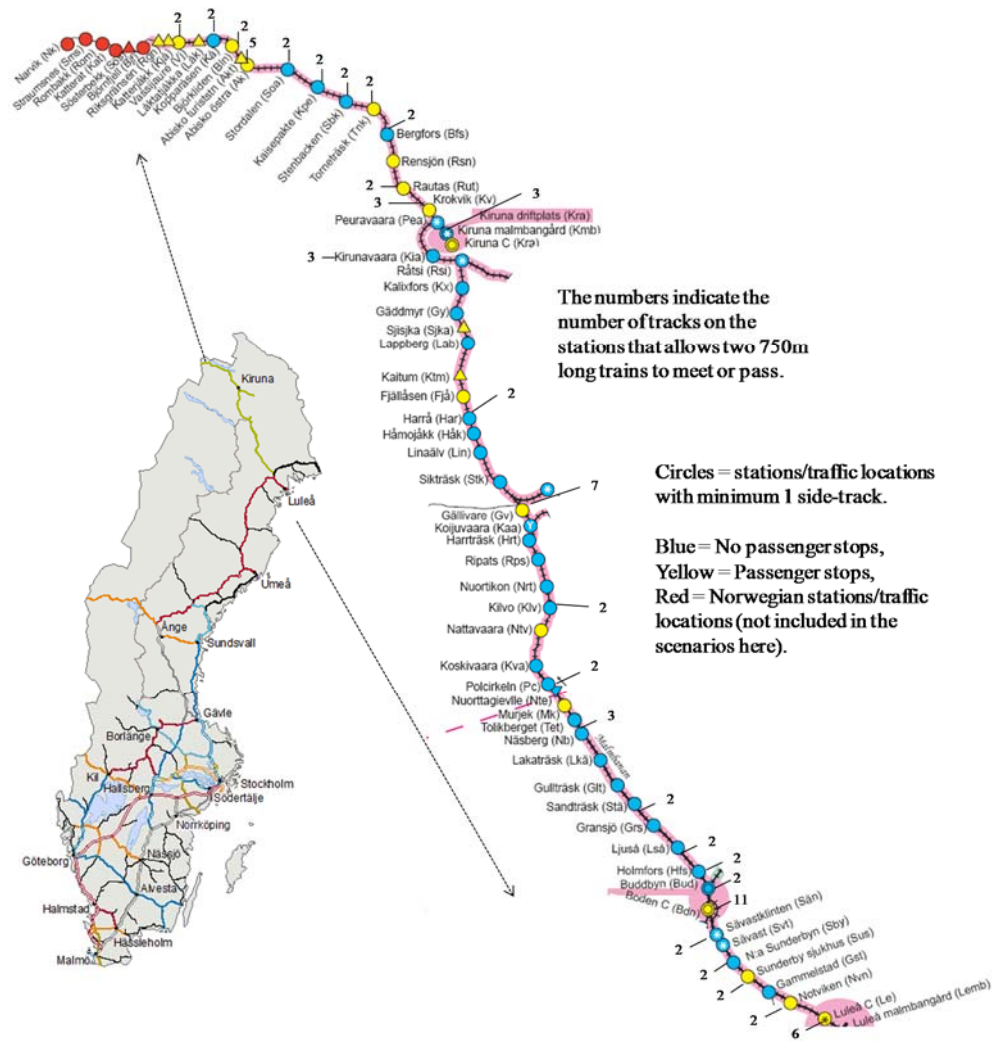


Figure 1. Overview of the studied railway line Björnfjell-Riksgränsen-Kiruna-Boden-Luleå and its characteristics. The stations indicated with a number, are stations that can accommodate two long (750m) freight trains and the number specifies the number of tracks. Source: Compilation of illustrations and data provided by Trafikverket.

4 Scenarios and experimental framework

4.1 Selection of scenarios

The selection of scenarios was based on an empirical analysis of common and relevant disturbances on the line. Freight trains often depart significantly ahead or behind of schedule on this line, but also passenger trains sometimes join the line delayed from south of Sweden, or from Narvik, Norway. Infrastructure problems such as signalling issues occur everywhere in the network and are also relevant to consider in this context. Animals

on the track (particularly reindeers this far north) sometimes affect the traffic and can cause a temporary blockage on the line. An overview of the defined scenarios can be found in Table 1.

Table 1. Overview of the 20 main scenarios that were experimentally analyzed in detail. These are based on the planned traffic and maintenance during Wednesday May 28th, 2014. In each scenario 113 trains are scheduled while the numbers in the parentheses specify the number of trains that run more than just through any of the junction stations. Even-numbered trains are north-bound and odd-numbered trains are south-bound. The stretch is composed of 162 line and station sections.

Scenario	T_0	Disturbance description	Problem size (#trains/#events)
1	14.41	Early and delayed iron ore trains: 9911 departs early (14.20 instead of 14.52 from Riksgränsen, 9914 departs delayed (14.46 instead of 14.10) from Bergfors, 9909 departs delayed from Bergfors (14.46 instead of 14:40)	113(46)/1234
2	14.41	Delayed iron ore trains: 9914 departs delayed (14.46 instead of 14.10) from Bergfors, 9909 departs delayed from Bergfors (14:46 instead of 14:40)	113(46)/1235
3	16.05	Iron ore train 9913 wants to depart 16.10 instead of 16:31 from Riksgränsen	113(41)/1201
4	16.05	9913 wants to depart 16.15 instead of 16:31 from Riksgränsen	113(41)/1201
5	16.25	9913 departs late, 16.36 instead of 16:31, from Riksgränsen	113(42)/1161
6	13.20	9914 departs minimum 10 min late from Kiruna	113(42)/1136
7	13.20	9914 departs minimum 15 min late from Kiruna	113(42)/1136
8	13.20	9914 departs minimum 20 min late from Kiruna	113(42)/1136
9	13.20	Paxtrain 93 arrives late (13.45 instead of 13:29) to Riksgränsen	113(42)/1136
10	13.20	Paxtrain 93 arrives late (13.55 instead of 13:29) to Riksgränsen	113(42)/1136
11	10.30	Paxtrains 96 departs late (10.45 instead of 10.36) from Luleå.	113(34)/921
12	10.55	Paxtrains 96 departs late (11.12 instead of 11.02) from Boden.	113(36)/954
13	09.34	Pax train 7155 departs late (09.50 instead of 9.41) from Kiruna	113(34)/831
14	10.25	Pax train 7155 arrives late to Linträsk (10.45 instead of 10:33)	113(34)/932
15	13.30	Pax train 93 get a suddden 45min stop between Riksgränsen-Katterjäk.	113(45)/1160
16	12.42	Iron ore train 9909 get a suddden 45min stop between Riksgränsen-Katterjäk.	113(46)/1223
17	13.35	Pax train 93 get a suddden 45min stop between Katterjäk-Vasserjaure.	113(47)/1176
18	12.44	Iron ore train 9909 get a suddden 45min stop between Katterjäk-Vasserjaure	113(44)/1218
19	12.10	Pax train 7155 gets a 45 min stop between Lakaträsk and Gullträsk.	113(44)/1137
20	13.30	Freight train 9231 gets a 45 min stop between Lakaträsk and Gullträsk.	113(46)/1240

The planning time horizon is set to four hours based on the scheduled travel times for the different trains on the different stretches. For passenger trains it takes approximately two hours between Riksgränsen and Kiruna, ca 3.5 hours between Kiruna and Boden and approximately 25 min between Boden and Luleå. Very few trains run the complete stretch. For the freight trains it takes approximately three hours from Kiruna to Riksgränsen and four hours from Kiruna to Luleå. The line is connected to other single-tracked lines and there are several trains that join/leave the line at e.g. Råtsi, Gällivare and Boden. In the experiments, we include the first line section that connects each line with another to capture also the train movements that just start at, or run through, the modeled stations on the line depicted in Figure 1.

The iron ore trains (usually numbered 99XY, or 199XY) that are running north towards Narvik from Kiruna are loaded, i.e. heavier than the other trains, while the south-bound trains towards Kiruna are unloaded. Furthermore, south-bound iron ore trains that run from Kiruna to Luleå are loaded while the trains in the reversed direction are

unloaded. This means that in the timetable, the loaded trains have no scheduled stops (apart from a few exceptions e.g. train 9912) for meetings and no commercial stops at all, but are running through the stations on the main track while other trains wait on the side-tracks (see a snap shot of the timetable in Figure 2 below). The timing of the train entrances during meetings is very important since the stations require a rather long time separation between the entrances of the two different trains. This time separation differs between train pairs and the order of entering the station. A passenger train that enters a station prior to any other train needs to arrive minimum two minutes before, while if it is a long freight train that enters first it needs to arrive minimum four minutes before. The trains that have scheduled stops consequently have more time margins (see Table 2) and due to the length restrictions of certain stations, the time distances between alternative meeting locations for two iron ore trains are large, see Table 3.

Table 2. Runtimes and margins for a number of train types.

Train	Stretch	Scheduled travel time (hh:mm:ss)	Minimum travel time between stations (hh:mm:ss)	Margin (hh:mm:ss)
SB Iron Ore 9913	Björn fjell-Kiruna	03:11:39	01:57:09	01:14:30
SB Iron Ore 9915	Björn fjell-Kiruna	02:57:21	01:39:21	01:18:00
NB Iron Ore 9914	Kiruna-Björn fjell	02:33:01	02:23:31	00:09:30
NB Iron Ore 9916	Kiruna-Björn fjell	02:29:01	02:23:31	00:05:30
SB pax train 93	Björn fjell-Kiruna	02:22:00	02:00:32	00:21:28
SB pax train 93	Kiruna-Luleå	04:14:00	03:41:09	00:32:51
SB pax train 7155	Kiruna-Luleå	03:29:00	03:05:29	00:23:31

Table 3. The time distance between the locations between Kiruna and Björn fjell which permit two iron ore trains to meet.

Stretch		Time distance (hh:mm:ss) based on 9916
Kiruna	Krokvik	00:09:40
Krokvik	Rautas	00:10:52
Rautas	Bergfors	00:23:03
Bergfors	Torneträsk	00:10:51
Torneträsk	Stenbacken	00:10:11
Stenbacken	Kaisepakte	00:11:38
Kaisepakte	Stordalen	00:12:48
Stordalen	Abisko Östra	00:11:00
Abisko Östra	Björkliden	00:10:50
Björkliden	Kopparåsen	00:10:43
Kopparåsen	Vassijaure	00:17:02
Vassijaure	Björn fjell	00:07:13

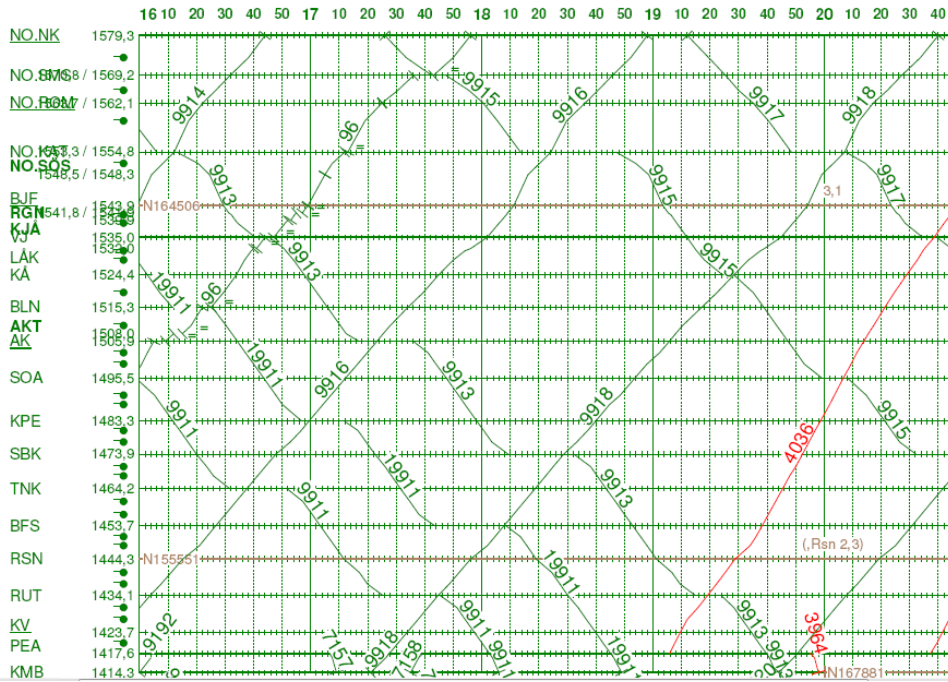


Figure 2. A snap shot of the timetable for the stretch between Kiruna (KMB), Riksgränsen (RGN) and Narvik (NO.NK) during May 28th, 2014.

4.2 Problem formulation and solution methodology

A previous survey of the characteristics at the Iron Ore Line and observations of the traffic situation indicated early on that the objective of the re-scheduling is very situation-dependent and primarily because it is a congested, single-tracked line with a large share of freight trains and significant limitations on available meeting locations. Many of the real-time conflicts also occur between trains that belong to the same operator. The operators may therefore have a stronger opinion on the proposed re-scheduling plan since it is under certain circumstances allowed to prioritize between its own trains. The need to study the relevance and effect of alternative objective functions was therefore obvious as well as the need to investigate the capability of current commercial solvers that could enable a flexible, dynamic configuration of the problem formulation.

The optimization approach is based on the mixed integer linear program (MILP) formulation outlined in (Törnquist Krasemann, 2012) where it was applied to a double-tracked Swedish sub-network. Since the case study presented here involves a railway line which is frequently used by trains that are significantly longer and heavier than what is normally permitted in Sweden, additional constraints are proposed and applied and some constraints have been modified.

The problem formulation for this line may seem less complex and more straightforward than the one for the Swedish double-tracked, busy Southern Main Line studied in (Törnquist Krasemann, 2012). However, due to the relatively large differences in the characteristics of the trains running on this line and the significant time distances between possible meeting points which give rise to significant difference between two

alternative solutions where one train is prioritized over another, the choice of soft side-constraints and objective function is non-trivial.

We have evaluated the effect of several alternative objective functions and how to incorporate soft constraints related to when and how a loaded iron train should stop in favor of a another train. We have chosen to focus on two main objective functions: 1) Minimizing the total final delay exceeding three minutes when trains arrive at their final destination (within the problem instance) and 2) minimizing the delays exceeding three minutes at intermediary, scheduled commercial stops as well as the final destinations. There are some obvious pros and cons with these two objective functions. The first one does not consider what happens to the trains “en-route” although it is often reflected partly at the final destination, but connections and so forth may then be overlooked. The second objective, which attempts to include the en-route punctuality by minimizing the delay at commercial stops, can be interpreted as if trains with several stops are given priority. One way to partly handle this is to use weights and normalize the weights in line with the number of commercial stops an operator may have, but it is far from trivial how such weights should be set in a practical setting.

A soft restriction for the dispatchers managing the Iron Ore line is that loaded iron ore trains should not be stopped since it is associated with a substantial time loss, energy loss and it wears on the tracks and brakes of the train. This preference is possible to accommodate in many cases since these trains cannot stop everywhere and the passenger trains have already some scheduled stops, which then motivates prioritization of the loaded train. When there is a conflict between two iron ore trains (most commonly one in each direction), the operator (LKAB) have the possibility to advice the dispatchers which should be given priority. In case of delays that would have a significant impact on the rolling stock, the unloaded train may be considered more important by LKAB, however. Consequently, this prioritization is very dynamic and context-dependent. In this experiment, we have some time penalties for making unscheduled stops. These penalties are associated with the run time extensions (specified by the parameter d_{k+1}^+) that arise when a train needs to brake prior to the stop and accelerate after. These values are good estimations based on empirical data and data from the run time profiles included in Trafikverket’s timetable.

The adapted optimization model is outlined below. A list of the symbols and notation used is also provided in the appendix.

T is defined as the set of trains, B is the set of sections, and E is the set of events where an event is a resource request by a certain train for a specific section. We let index i be associated with a train, j with a section and index k with an event. Each event is connected to both a train and a section. Let $K_i \subseteq E$ be the ordered set of events for train i ($i \in T$) and $L_j \subseteq E$ be the ordered set of events of section j ($j \in B$). Events in K_i and L_j are ordered according to the nominal timetable. We use $(k+1)$ to denote the first proceeding event of event k (in K_i and L_j) and $k < \hat{k}$ to denote that \hat{k} is any event proceeding event k with respect to the order in the sets. Furthermore, let n_i and m_j denote the last event of K_i and L_j , respectively where n_0 is used to denote the first event in K_i . Each event k has an initial scheduled starting time, b_k^{initial} , and end time, e_k^{initial} where d_k specifies the minimum occupation time and h_k specifies if the train has a scheduled commercial stop or not.

Each section has a set of parallel tracks $P_j = \{1, \dots, p_j\}$ which in this context only concerns the stations since we focus on a single-tracked line. If two trains are to be scheduled on the same track t within a station section j , a minimum of Δ_j time units separation time is required, where $\Delta_j > 0$. We have used 30 seconds in our experiments.

The formulation contains eight types of decision variables. The continuous variables x_k^{begin} and x_k^{end} represent the re-assigned start and end time of event k . z_k represents the delay train i experiences when finishing event k . We have five binary variables:

$$q_{kt} = \begin{cases} 1, & \text{if event } k \text{ uses track } t, \text{ where } k \in L_j, t \in P_j, j \in B. \\ 0, & \text{otherwise.} \end{cases}$$

$$h_k = \begin{cases} 1, & \text{if event } k \text{ occurs on a station and the corresponding train } i \text{ stopped} \\ & \text{where } k \in K_i, i \in T: S_k = 0. \\ 0, & \text{if event } k \text{ passed through the station without stopping.} \end{cases}$$

$$\gamma_{k\hat{k}} = \begin{cases} 1, & \text{if event } k \text{ occurs before event } \hat{k} \text{ (as in the nominal timetable),} \\ & \text{where } k, \hat{k} \in L_j, j \in B: k < \hat{k}. \\ 0, & \text{otherwise.} \end{cases}$$

$$\lambda_{k\hat{k}} = \begin{cases} 1, & \text{if event } k \text{ is re-scheduled to occur after event } \hat{k}, \\ & \text{where } k, \hat{k} \in L_j, j \in B: k < \hat{k} \ \& \ |P_j| > 1. \\ 0, & \text{otherwise.} \end{cases}$$

$$\omega_{k\hat{k}} = \begin{cases} 1, & \text{if event } k \text{ occurs on a station and is re-scheduled to start before event } \hat{k}, \\ & \text{where } k, \hat{k} \in L_j, j \in B: k < \hat{k} \ \& \ S_k = 0 \ \& \ |P_j| > 1. \\ 0, & \text{if event } k \text{ is re-scheduled to start after event } \hat{k}. \end{cases}$$

S_k specifies if event k takes place at a line section ($S_k = 1$) or a station section ($S_k = 0$). $\omega_{k\hat{k}}$ is used to capture which of any two trains that enters a certain station first. The associated constraint is included to ensure that trains enter the stations with a sufficient time separation (i.e. no station on this line permits simultaneous entrance).

In this context, we have no line sections that contains more than one block section and therefore we can omit the constraints in the previous MILP formulation that maintain headway between trains since the safety distances are taken care of by the other constraints. One could argue that with the presence of many long, freight trains, a minimum headway distance between trains in the same direction could be useful in order to ensure increased robustness, but in these experiments we have not considered that. The stretches between two stations which have more than one block section are represented by the same number of consecutive line sections.

Based on results from some previous, related experiments, the selection of a relevant objective function depends on the actual type of scenario and context. In this paper, we

primarily study the effect of two alternative objective functions. We applied them separately in order to compare and analyze their impact on the re-scheduling solutions generated. Objective function (1a) minimizes the delay of the trains at their end station (within the problem instance) compared to the nominal timetable, while objective function (1b) minimizes the delay from the nominal timetable for all scheduled, intermediary commercial stops as well as the delay at the end station. Constraint (2) ensures that each event is directly followed by the next event within the sorted event list for each train. Constraint (3) ensures that a train that has a scheduled commercial stop does not leave the corresponding station before the scheduled departure time. Constraints (4)-(5) ensure that events that are already active in different ways when the re-scheduling process is initiated, are allocated valid start, and, in some cases, end times. Constraint (6) captures if a train is passing through a station or making a stop, depending on the actual duration time where there is a threshold value, ψ , used to classify if the duration is long enough to be considered a stop, or not. We have used 30s here. Constraint (7) is used to set the binary stop variable to one once there are scheduled commercial stops (i.e. not stops that are scheduled to enable only a meeting, which could be shifted or cancelled). Constraints (8)-(10) are used to ensure that the minimum run time requirement is enforced depending on if station stops are made or not.

Constraint (11) computes the delay exceeding a certain threshold value, ε , and in this study we have set the value to three minutes.

$$\text{Minimize } \sum_{i \in T} z_{n_i} \quad (1a)$$

$$\text{Minimize } \sum_{i \in T} \sum_{k \in K_i} z_k + \sum_{i \in T} z_{n_i} \quad (1b)$$

Train restrictions

$$x_k^{end} = x_{k+1}^{begin} \quad i \in T, k \in K_i : k \neq n_i \quad (2)$$

$$x_k^{begin} \geq b_k^{initial} \quad k \in E : h_k = 1 \quad (3)$$

$$x_k^{begin} = b_k^{static} \quad k \in E : b_k^{static} > 0 \quad (4)$$

$$x_k^{end} = e_k^{static} \quad k \in E : e_k^{static} > 0 \quad (5)$$

$$M * \hat{h}_k \geq x_k^{end} - x_k^{begin} - \psi \quad k \in E : S_k = 0 \quad (6)$$

$$\hat{h}_k \geq h_k \quad k \in E : S_k = 0 \quad (7)$$

$$x_{k+1}^{end} \geq x_{k+1}^{begin} + d_{k+1} + \hat{h}_k * d_{k+1}^+ \quad i \in T, k \in K_i : S_k = 0, k \neq n_i \quad (8)$$

$$x_{k-1}^{end} \geq x_{k-1}^{begin} + d_{k-1} + \hat{h}_k * d_{k-1}^+ \quad i \in T, k \in K_i : S_k = 0, k \neq n_0 \quad (9)$$

$$x_k^{end} \geq x_k^{begin} + d_k \quad k \in E \quad (10)$$

$$z_k \geq x_k^{begin} - b_k^{initial} - \varepsilon \quad k \in E \quad (11)$$

Infrastructure restrictions

Constraints (12)-(14) ensure that the train is allocated exactly one track per section it occupies and that in line with the length requirement on the assigned track with respect to the length of the train. Constraint (14) combined with constraints (20)-(21) below are then enforcing a restriction for meetings/passes involving two or more long (750m) freight trains.

$$\sum_{t \in P_j} q_{kt} = 1 \quad j \in B, k \in L_j : |P_j| > 1 \quad (12)$$

$$\sum_{t \in P_j} t^* q_{kt} = r_k^{track} \quad j \in B, k \in L_j : |P_j| > 1 \& r_k^{fixed} = 1 \text{ (i.e. when } b_k^{static} > 0 \text{)} \quad (13)$$

$$q_{kt} = 0 \quad j \in B, k \in L_j, i \in T, k \in K_i : |P_j| > 1 \& l_{jt}^{tracklength} < l_i^{trainlength} \quad (14)$$

$$q_{\hat{k}t} + q_{kt} - 1 \leq \lambda_{k\hat{k}} + \gamma_{k\hat{k}} \quad j \in B, k, \hat{k} \in L_j, t \in P_j : k < \hat{k} \& |P_j| > 1 \quad (15)$$

Constraints (16)-(19) serve to ensure required safety regulations associated with line blocking on stations as well as line sections. Constraint (17) is for single-tracked sections and constraint (18) is for sections with multiple tracks (i.e. stations in this context). M is a large constant with the value corresponding to eight hours in these experiments.

$$x_{\hat{k}}^{begin} - x_k^{end} \geq \Delta_j \gamma_{k\hat{k}} - M(1 - \gamma_{k\hat{k}}) \quad j \in B, k, \hat{k} \in L_j : k < \hat{k} \quad (16)$$

$$x_k^{begin} - x_{\hat{k}}^{end} \geq \Delta_j (1 - \gamma_{k\hat{k}}) - M\gamma_{k\hat{k}} \quad j \in B, k, \hat{k} \in L_j : k < \hat{k} \& |P_j| = 1 \quad (17)$$

$$x_k^{begin} - x_{\hat{k}}^{end} \geq \Delta_j \lambda_{k\hat{k}} - M(1 - \lambda_{k\hat{k}}) \quad j \in B, k, \hat{k} \in L_j : k < \hat{k} \& |P_j| > 1 \quad (18)$$

$$\lambda_{k\hat{k}} + \gamma_{k\hat{k}} \leq 1 \quad j \in B, k, \hat{k} \in L_j : k < \hat{k} \& |P_j| > 1 \quad (19)$$

Constraints (20)-(21) serve to ensure that the entrance of trains to stations is separated by a minimum time interval if the station does not permit simultaneous entrances. This is a specific property related to traffic involving long and/or heavy trains and station tracks which do not cater for a sufficient safety distance that would ensure that all trains can make a full stop before or within the station without risking running into the train in the opposite direction when it enters.

Constraints (22)-(23) explicitly formulate the implicit relationship between the train order variables and the entrance variables for the stations.

$$x_{\hat{k}}^{begin} - x_k^{begin} \geq \Delta_{k\hat{k}} \overline{\omega}_{k\hat{k}} - M(1 - \overline{\omega}_{k\hat{k}}) \quad j \in B, k, \hat{k} \in L_j : k < \hat{k} \& |P_j| > 1 \& S_k = 0 \quad (20)$$

$$x_k^{begin} - x_{\hat{k}}^{begin} \geq \Delta_{k\hat{k}} (1 - \overline{\omega}_{k\hat{k}}) - M\overline{\omega}_{k\hat{k}} \quad j \in B, k, \hat{k} \in L_j : k < \hat{k} \& |P_j| > 1 \& S_k = 0 \quad (21)$$

$$\gamma_{k\hat{k}} + (1 - \varpi_{k\hat{k}}) \leq 1 \quad j \in B, k, \hat{k} \in L_j : k < \hat{k} \& |P_j| > 1 \& S_k = 0 \quad (22)$$

$$\lambda_{k\hat{k}} + \varpi_{k\hat{k}} \leq 1 \quad j \in B, k, \hat{k} \in L_j : k < \hat{k} \& |P_j| > 1 \& S_k = 0 \quad (23)$$

$$x_k^{begin}, x_k^{end} \geq 0 \quad k \in E \quad (24)$$

$$z_k \geq 0 \quad k \in E \quad (25)$$

$$\gamma_{k\hat{k}} \in \{0,1\} \quad j \in B, k, \hat{k} \in L_j : k < \hat{k} \quad (26)$$

$$\lambda_{k\hat{k}} \in \{0,1\} \quad j \in B, k, \hat{k} \in L_j : k < \hat{k} \& |P_j| > 1 \quad (27)$$

$$\varpi_{k\hat{k}} \in \{0,1\} \quad j \in B, k, \hat{k} \in L_j : k < \hat{k} \& |P_j| > 1 \& S_k = 0 \quad (28)$$

$$\hbar_k \in \{0,1\} \quad k \in E : S_k = 0 \quad (29)$$

$$q_{kt} \in \{0,1\} \quad j \in B, k \in L_j, t \in P_j : |P_j| > 1 \quad (30)$$

On the line there are also some maintenance scheduled in different locations and during certain time periods. Rensjön, for example, were subject to maintenance on all side-tracks during May 28th, 2014 and thus it was not possible to schedule trains to meet or pass there. The maintenance have been modeled as “ghost train events” with a fixed start and end time to make sure the corresponding line and station track sections are not available for the trains during the specific time periods. That is, these ghost trains are included in a set of ghost trains, G , where G is a subset of T . Each ghost train i has only one event, which belongs to the sets of events, K_i and L_j respectively. These ghost train events have a static start time, b_k^{static} , and a static end time, e_k^{static} , corresponding to their initial start and end time and a fixed track to use, in line with the scheduled track blockage and time duration. Consequently, when there is a longer stretch of maintenance, or multiple tracks on a station blocked due to maintenance, the corresponding number of ghost trains and ghost train events need to be defined. These ghost train events are naturally not included in any of the evaluation metrics used since they do not experience any delays. To formally state the exclusion of those ghost train events in the delay calculation, the following constraint is included:

$$z_k = 0 \quad i \in G, k \in K_i \quad (31)$$

All scenarios were solved with a planning horizon of 4 hours, using CPLEX 12.5 in parallel, deterministic mode using up to 8 threads on a server with 4 processors at 2 GHz, 24 GB of RAM, running with GNU/Linux 3.2.0-x86-64.

5 Results and discussion

Table 4 below presents some results from the experiments based on the 20 defined disturbance scenarios described in Table 1 using objective functions defined by Eq. 1a) and 1b) to solve the corresponding re-scheduling problems. It took the solver one minute, or less, to solve all scenarios to optimality. No pattern, regarding which of the objective

functions that generated a problem that was easier (i.e. faster) to solve than the other, could be found. There was also no obvious indication that a certain type of disturbance scenario was harder to solve than another. However, the number of different scenarios was quite limited.

Comparing the two delay metrics, $\Sigma TFD+3$ and $\Sigma TDC+3$, corresponding to the two alternative objective functions, the second objective function often results in a solution that is equally good or better considering both metrics. In six scenarios, marked grey in Table 4, the solutions differ in both metrics.

The 5th and 10th column of Table 4 shows the number of trains that were delayed more than three minutes at the end stations for the problem formulation using either of the objective functions. The minimization of delays at commercial stops and end station resulted in only a few more delayed trains (in scenario 9 and 17). In the 6th and 11th column, the number of extra stops by loaded iron ore trains is specified. There are very few scheduled stops for loaded iron ore trains, mainly for 9912, 9957 and 9993. In certain scenarios, several scheduled meetings are cancelled which is shown with a negative number. A reduction of meetings can also arise if a train is stopped to meet two trains in one and the same location instead of stopping twice at different locations.

Table. 4. A delay is classified as a positive deviation larger than three minutes from the initial timetable. Consequently, TFD+3 specifies for a specific train, the delay exceeding three minutes at the end station while TDC+3 measure the same thing but at the intermediary commercial stops and the end station. Some trains, such as the iron ore trains, have no intermediary commercial stops while passenger trains typically have a few. The delays are given in seconds.

Scenario	Objective function $\Sigma TFD+3$					Objective $\Sigma TDC+3$				
	Computation time (s)	$\Sigma TFD+3$ (s)	$\Sigma TDC+3$ (s)	# trains TFD+3	# extra stops for loaded iron ore trains	Computation time (s)	$\Sigma TFD+3$ (s)	$\Sigma TDC+3$	# trains TFD+3	# extra stops for loaded iron ore trains
1	45,86	1545	4449	1	2	34,8	1545	1545	1	1
2	34,42	1245	5220	1	2	38,71	1245	1459	1	0
3	37,54	0	880	0	0	34,68	0	0	0	1
4	33,72	0	4065	0	0	19,9	0	0	0	0
5	29,78	192	3486	1	-1	21,58	192	576	1	-1
6	17,35	0	2808	0	-1	25,43	0	541	0	1
7	2,69	150	4775	1	-1	22,87	150	323	1	0
8	27,3	450	4277	1	0	16,14	545	545	1	0
9	21,56	0	6084	0	3	23,25	763	3313	2	4
10	22,3	0	11375	0	-1	36,63	28	7343	1	1
11	1,6	0	656	0	-3	6,53	0	458	0	3
12	2,06	0	645	0	1	9,55	0	420	0	0
13	1,21	0	220	0	0	1,14	0	30	0	0
14	14,55	0	1773	0	-3	1,85	0	948	0	-4
15	60,5	1593	16006	3	2	43,07	1875	13315	3	2
16	11,34	363	2257	1	1	22,9	363	363	1	-1
17	15,75	3225	16634	2	2	34,15	3595	12152	3	2
18	24,63	363	7207	1	0	16,24	363	2883	1	1
19	2,74	2301	11774	1	-1	16,86	2315	9262	1	1
20	21,79	1918	2697	1	0	22,11	1918	1918	1	0

In order to better understand how different solutions the two objective functions generate, we further compare and analyze them more in detail. It can be observed that in scenario 3 and 4, where we have trains that request to depart ahead of schedule (which is a common

phenomenon at the Iron Ore Line), the objective function that minimizes delays at end station, does not surprisingly produce solutions that have delays en-route. When the objective function instead is set to minimize the delays at intermediary stations with commercial, scheduled stops and at the end station, the solutions converge more towards the initial timetable. A phenomenon that may appear when using either of the objective functions and which was observed when visually analyzing the generated solutions in scenario 3 and 4, is that early trains are not "pushed forward" as intended. This is obviously, and not surprisingly, the effect of using an objective that does not promote a good traffic flow on the line. The objective function was consequently modified to also include a minimization of the actual travel time for all trains, but with a much lower weight per time unit than for the delays. This modification generated satisfactory solutions but naturally required that the early trains were allocated a preliminary, early starting time which their actual travel time computation was based on. This secondary objective only intended to distinguish between two similar solutions when there was room for improvement with respect to other aspects such as giving the early freight trains priority as long as this did not create delays at commercial stops for other trains. The idea of minimizing travel time for heavy freight trains could also be questioned based on the principles of eco-driving. This is of course an important aspect but needs to be compared with the benefits of making sure the trains are able to run ahead of schedule and create extra safety margins in the rolling-stock and transport system schedule that the operator may want to achieve.

Another interesting observation which was made in e.g. scenario 3, where the unloaded iron ore train, 19911, enters the station Björkliden for a meeting with passenger train 96. Train 96 has a scheduled commercial stop there but both trains stop since it means no time loss for the iron ore train. Ideally, train 96 should have entered first and 19911 would have passed through the station at a low speed. See Figure 3 below. In situations like this one, a DAS system would be beneficial to "polish" the solution.

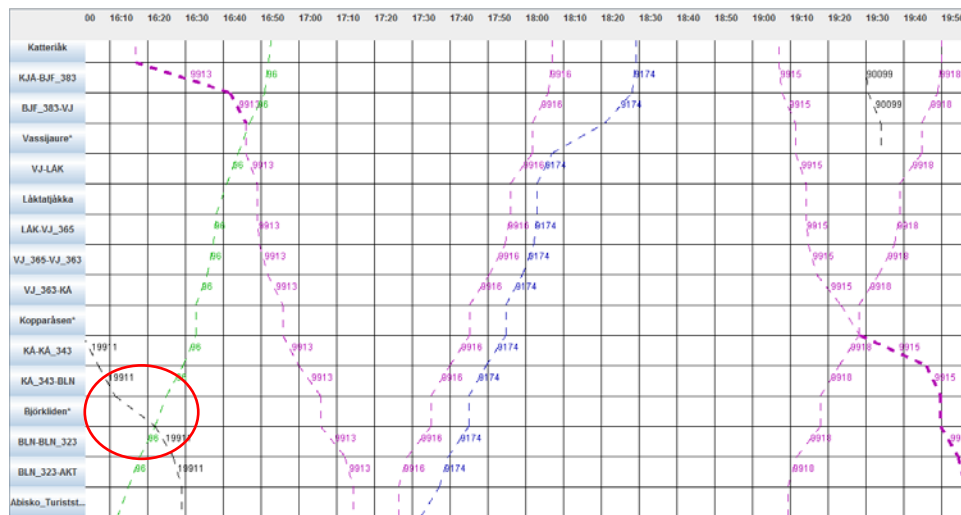


Figure 3. A snapshot of the re-schedule timetable in scenario 3, where the trajectory of iron ore train 19911 through station Björkliden is not optimal from an eco-driving perspective.

In several of the scenarios we also exposed different trains to smaller initial delays in the interval 5-12 minutes. Because of the large amount of margin time allocated in the meeting points, the timetable could easily absorb these delays by shifting meetings. This pattern can also be observed in the statistics from Trafikverket that shows the actual on-time performance of these trains.

In scenarios 15-20, the trains are exposed to larger, temporary disturbances of 45 minutes and the structural differences between the alternative solutions provided by the two objective functions are larger. In e.g. scenario 15, it is interesting to note that partly different trains are delayed. Similar to other scenarios, objective function 1a) has a tendency to assign trains with a lot of margin time and no intermediate commercial stops quite freely and not drive them forward as long as it is not in conflict with the on-time performance of other trains. In that sense, objective function 1b) generates more reasonable solutions for a single-tracked line with multiple dependencies. On the other hand, a delayed train with multiple commercial stops may be prioritized over several trains with smaller arrival delays. In Figure 4 below, such an example from scenario 15 is illustrated, where the second solution generated by objective function 1b) clearly promotes the passenger train 93.

For larger and more severe disruptions than the ones represented by scenario 15-20, the objective probably varies. The situation can be divided into three phases which is initiated by and dependent on the prognosis and assessment of the traffic situation done by the dispatchers:

- Traffic reduction phase
- Intermediary traffic management phase
- Recovery and reset phase

The choice of objective function in these different phases is not obvious and it probably depends on a number of factors and the complications that a traffic reduction is associated with. The dispatchers (in cooperation with the operators) probably need to define a basis for the re-scheduling solution in these situations and determine where certain trains should meet. The decisions taken may also involve cancellation of trains and turning trains earlier than planned, which are decisions not included in the proposed problem formulation.

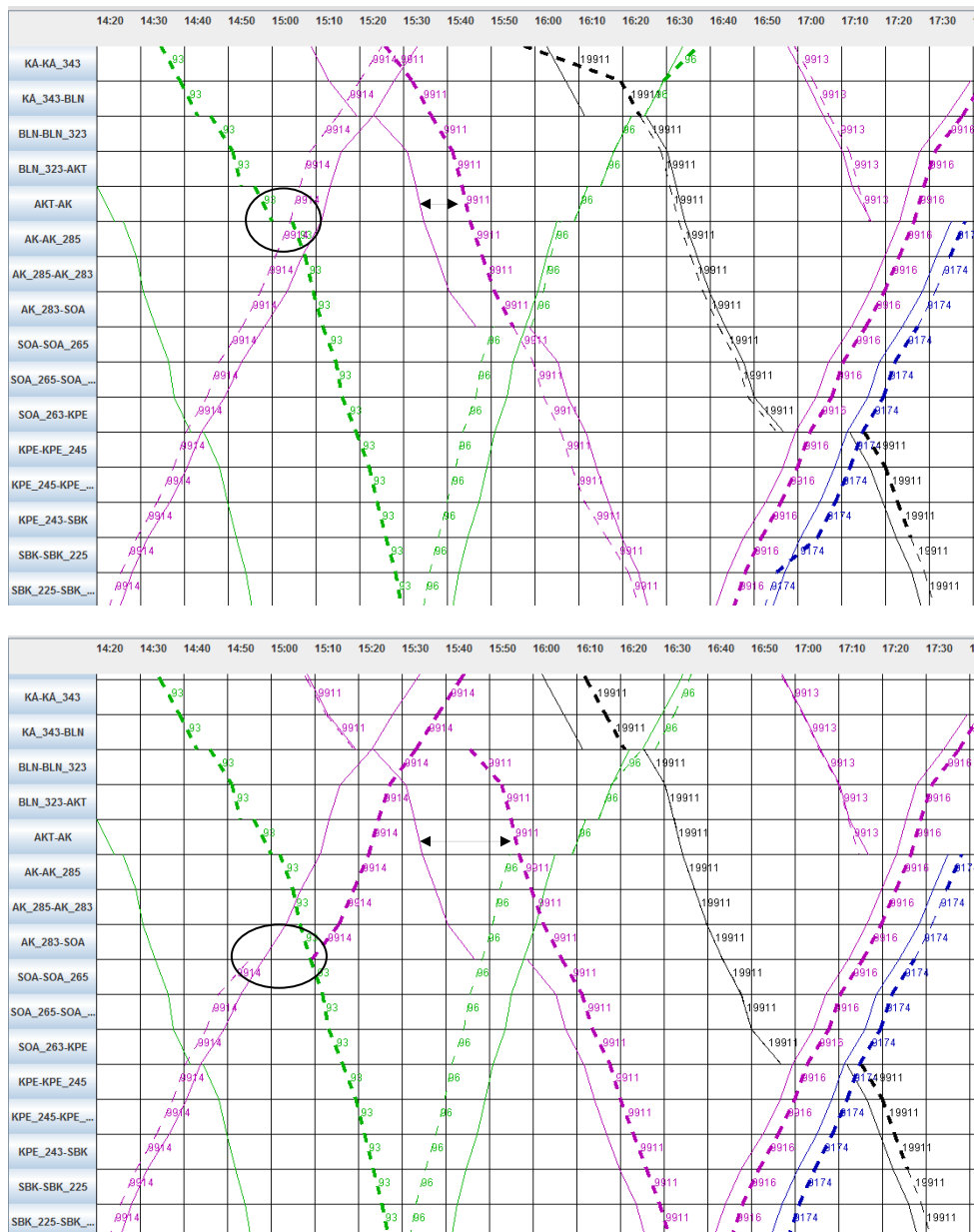


Figure 4. The top image illustrates a snap shot of the solution from objective function 1a) and below the corresponding one from objective function 1b) in scenario 15. Note how different the trajectories of train 9914 and 9911 are in the alternative solutions. In the upper graph, 9914 is prioritized over 93 while in the second case it is the reversed since the accumulated delay by several commercials stops of 93 weighs more than an arrival delay of 9914. This in turn gives ripple effects on train 9911 and 96 as can be observed in the lower graph. The dotted lines are the re-scheduled timetable slots and the solid lines the nominal timetable slots.

During the experiments, we feared that the re-scheduling solutions generated from the problem formulation frequently would lead to multiple stops for the loaded iron ore trains. We did choose to not forbid them from stopping because there are in practice situations where it makes sense to let them stand aside and prioritize other trains. This did, however, never become a large problem and that is probably due to that these trains have relatively little time margin in their schedule and the time required to let a freight train enter a meeting station before a passenger train is larger than the reversed order. It is, however, not a guarantee that it will not become a problem in other scenarios and this aspect needs to be studied further. Furthermore, in Table 4 we can observe that the number of extra stops for the loaded iron ore trains increased by 3 and 4 (scenario 9 and 11) and the solutions were investigated in more detail. The extra stops were situations where either a loaded train stopped for an unloaded train (when the initial, reversed, solution would have been just as good w.r.t. the objective function value) or - even more unmotivated - stopped at an intermediary station rather than cruising until the next relevant meeting point with another train. When the objective functions were fine-tuned to also penalize (with a low weight) extra stopping time, this poor behavior was eliminated in the two mentioned scenarios.

A final, relevant aspect to consider in this context is how the problem formulations and resulting re-scheduling solutions comply with the nation-wide traffic management guidelines that dictate how dispatchers may prioritize between trains in real-time when deviations occur. The Swedish guidelines basically state that the conflict resolution should be fair and a train on-time should not have to be delayed by an already delayed train. Today, these guidelines are under investigation and under revision since they are not always found effective and the dispatchers therefore often follow their gut feeling based on their long experience and good skills. However, the principle of being fair will always be a key aspect in the train traffic management of the Swedish national railway network and therefore relevant to reflect upon in a study like this. The results presented in Table 4 above indicate that the behavior of the solution method do not provide solutions that deviate significantly from the principle of fairness. That is, the affected trains (when considering the delays) are primarily the ones that were involved in the initial disturbances and it may be reasonable to not disturb also other trains if avoidable.

6 Conclusions and Future Work

This paper presents results from an experimental study investigating potential configuration challenges in the development of a computational re-scheduling support for Swedish train traffic dispatchers controlling the congested, single-tracked Iron Ore railway line located in Northern Sweden and partially Norway. The intended contributions are the proposed model extensions as well as the study of the practical implications of using the proposed problem formulation in this particular context. The resulting re-scheduling solutions were analyzed both visually and numerically in detail - beyond aggregated, average numbers - in order to identify potential weaknesses and issues.

From the results, we can conclude that commercial solvers can handle practical problems of a practical, relevant size and for this type of setting (i.e. a single-tracked network) and problem formulation. However, the type of scenario has a significant impact on the computation time which is not surprising. Furthermore, alternative, but similar, objective functions seem to provide solutions that differ more from each other in this single-tracked context than in a double-tracked network with bi-directional tracks. If we compare the conclusions from the study here which focuses on a single-tracked line with

very heterogeneous traffic and a large share of freight trains with an earlier study presented in (Törnquist, 2007) there are some interesting differences with respect to how much the alternative solutions (i.e. provided by the two objective functions in question) differ. The earlier study from 2007 was focusing on a part of the Swedish core network which is mainly double-tracked, very congested and a larger share of passenger trains than freight trains. An explanation to the observations found in this Iron Ore Line study may be that for a single-tracked line the share of margin often is larger than for a double-tracked line since trains usually do not meet. These margins, combined with the small amount of trains with intermediary commercial stops, may result in a problem space with more heterogeneous optimal solutions. This needs to be studied in more detail.

While analyzing the re-scheduling solutions both numerically and visually, we could observe that trains that are ahead of their schedule by early departure or by having a lot of margin time due to waiting time in meeting/passing locations, are not always "pushed" forward unless the objective functions promotes that in some way. This concerned primarily the freight trains.

The results also indicated the need to consider additional modifications of the applied objective function. A secondary objective for heavy freight trains is often to achieve smooth passages through meeting locations and avoid braking in front of a red light due to that the other train did not time its arrival. This objective, or preference, was also not represented in the objective functions used. The use of a DAS to "polish" the solutions provided by the dispatchers, or to use it as a secondary optimization criteria when there are multiple optimal solutions, is probably to recommend in this case rather than to include "eco-driving" in the main objective function. For further information concerning approaches incorporating a so-called "green wave policy", we refer to e.g. Corman et. al. (2009).

This study needs to be and will be continued with a detailed analysis of what characteristics in a re-scheduling solution that actually matters, when, why and how. The study will continue to focus on the Iron Ore Line but also the double-tracked Swedish Southern Mainline since a comparative analysis provides insight in which conclusions that seem to hold for different contexts and which observations and conclusions that seems to be very context-dependent.

Acknowledgements

This research has been conducted within the research project FLOAT at Blekinge Institute of Technology, Sweden which is funded by mainly Trafikverket (the Swedish Rail Administration) via the research programme KAJT as well as the municipality of Karlshamn. The author would like to thank foremost Peter Hammarberg and his colleagues at Trafikverket for their encouragement, time and support as well as research colleagues within KAJT for valuable feedback and discussions.

References

- Cacchiani, V., Huisman, D., Kidd, M., Kroon, L., Toth, P., Veelenturf, L., Wagenaar, J., (2014). An overview of recovery models and algorithms for real-time railway rescheduling. *Transportation Research Part B* 63(1), pp. 15-37.
- Caimi, G., Fuchsberger, M., Laumanns, M., Luethi, M. (2012) A model predictive control approach for discrete-time rescheduling in complex central railway station areas, *Computers & OR* 39, pp. 2578–2593.
- Carey, M., (1994). Extending a train pathing model from one-way to two-way track. *Transportation Research B* 28(5), pp. 395–400.
- Corman, F., D'Ariano, A., Pacciarelli, D., Pranzo, M. (2009) Evaluation of green wave policy in real-time railway traffic management, *Transportation Research Part C*, 17 (6), pp. 607-616.
- Corman, F., D'Ariano, A., Pacciarelli, D., Pranzo, M. (2010). A tabu search algorithm for rerouting trains during rail operations. *Transportation Research Part B*, 44 (1), pp. 175–192.
- Corman, F., D'Ariano, A., Pacciarelli, D., Pranzo, M. (2012). Bi-objective conflict detection and resolution in railway traffic management, *Transportation Research Part C* 20(1), pp. 79-94.
- Corman, F., D'Ariano, A., Pacciarelli, D., Pranzo, M. (2014) Dispatching and coordination in multi-area railway traffic management, *Computers and Operations*, 44, pp. 146-160.
- D'Ariano, A., Corman, F., Pacciarelli, D., Pranzo, M. (2008). Reordering and local rerouting strategies to manage train traffic in real time. *Transportation Science*, 42(4): pp. 405–419.
- Dollovoet, T., Huisman, D., Kroon, L., Schmidt, M., Schöbel, A. (2014), Delay Management including capacities at stations, *Transportation Science*, Vol. 24 (2), pp. 185-203.
- Hansen, I.A., Pachel, J. (Eds.) (2008). *Railway Timetable and Traffic: Analysis, Modelling and Simulation*. Eurailpress, Hamburg, Germany.
- Harrod, S., Schlechte, T. (2013), A direct comparison of physical block occupancy versus timed block occupancy in train timetabling formulations, *Transportation Research Part E* 54, pp. 50–66.
- Iqbal, S.M.Z., Grahn, H., Törnquist Krasemann, J. (2013), Multi-strategy-based train scheduling during railway traffic disturbances, In proceedings of RailCopenhagen 2013, Copenhagen Denmark, May 13-15.
- Kecman, P., Corman, F., A D'Ariano, RMP Goverde (2013) Rescheduling models for railway traffic management in large-scale networks, *Public Transport: Planning and Operations* 5 (1-2), pp. 95-123

Lamorgese, L., Mannino, C., (2015). An exact decomposition approach for the real-time Train Dispatching problem, *Operations Research*, 63 (1), pp. 48-64.

Mannino, C., Mascis, A. (2009). Optimal real-time traffic control in metro stations. *Operations Research* 57 (4), 1026–1039.

Mannino, C., (2011), Real-time traffic control in railway systems, in (eds.) Caprara, A., Kontogiannis, S., Proceedings of the 11th Workshop on Algorithmic Approaches for Transportation Modelling, Optimization, and Systems (ATMOS2011), DROPS, pp. 1-14.

Mascis, A., D. Pacciarelli. (2002). Job shop scheduling with blocking and no-wait constraints. *European Journal of Operational Research*, 143 (3), 498-517.

Meng, L., Zhou, X. (2011) Robust single-track train dispatching model under a dynamic and stochastic environment: A scenario-based rolling horizon solution approach, *Transportation Research Part B*, 45 (7), Pages 1080-1102.

Meng, L., Zhou, X., (2014) "Simultaneous train rerouting and rescheduling on an N-track network: A model reformulation with network-based cumulative flow variables", *Transportation Research Part B: Methodological*, Volume 67, September 2014, pp. 208-234.

Pellegrini, P., Marlière, G., Rodriguez, J., (2014). Optimal train routing and scheduling for managing traffic perturbations in complex junctions. *Transportation Research Part B: Methodological*, 59C:58-80.

Quaglietta, E., Corman, F., and Goverde, R. (2013) Impact of a Stochastic and Dynamic Setting on the Stability of Railway Dispatching Solutions. In Proceedings of the 16th International IEEE Annual Conference on Intelligent Transportation Systems, the Hague, Netherlands, 2013.

Sato, K., Tamura, K., Tomi, N., (2013) A MIP-based timetable rescheduling formulation and algorithm minimizing further inconvenience to passengers, *Journal of Rail Transport Planning & Management*, Volume 3, Issue 3, August 2013, Pages 38–53.

Törnquist Krasemann, J. (2012), Design of an effective algorithm for fast response to the re-scheduling of railway traffic during disturbances, *Transportation Research Part C* (20), pp. 62-78, Elsevier.

Törnquist, J. (2007), Railway traffic disturbance management-An experimental analysis of disturbance complexity, management objectives and limitations in planning horizon, *Transportation Research Part A: Policy and Practice*, Vol. 41, No. 3, pp. 249-266, Elsevier.

Törnquist, J., Persson, J.A. (2007) "N-tracked railway re-scheduling during disturbances", *Transportation Research Part B*, vol. 41, pp. 342-362.

Samà, M., Meloni, C., D'Ariano, A., Corman, F., (2015), A multi-criteria decision support system for real-time train rescheduling, In proceedings of RailTokyo, March 2015, Tokyo, Japan.

Yang, L., Liden, T., Leander, P., (2013), "Achieving energy-efficiency and on-time performance with Driver Advisory Systems.", In Proceedings of the 2013 IEEE International Conference on Intelligent Rail Transportation (ICIRT), pp.13-18, Aug. 30 2013-Sept. 1, 2013, ISBN 978-1-4673-5278-9.

Zaninotto, G., D'Ariano, A., Pacciarelli, D., Pranzo, M. (2013) Intelligent decision support for scheduling and rerouting trains on an Italian railway line, IEEE Conference on Intelligent Transportation Systems, Proceedings, Article number 6728326, pp. 779-784.

Zhou, X., Zhong, M., (2007), Single-track train timetabling with guaranteed optimality: Branch-and-bound algorithms with enhanced lower bounds, *Transportation Research Part B* 41(3), Elsevier Science Ltd, pp. 320-342.

Appendix

Table A1. Subscripts and parameters used in the mathematical problem formulation in section 4.2 .

Symbol	Definition
i	= train index.
j	= segment index.
t	= track index
T	= set of trains.
B	= set of segments.
G	= set of ghost trains, i.e. scheduled track maintenance periods. $G \subseteq T$.
E	= set of events.
K_i	= the ordered set of events for train i , $K_i \subseteq E$.
L_j	= the ordered set of events for train i , $K_i \subseteq E$.
P_j	= set of parallel tracks $P_j = \{1, \dots, p_j\}$.
$\Delta_{\hat{k}\hat{k}}$	= minimum time separation required between pair of trains using the same track resource, where $\Delta_{\hat{k}\hat{k}} > 0$.
n_i	= number of events for train i .
b_k^{initial}	= initial scheduled start time of event k .
e_k^{initial}	= initial scheduled end time of event k .
d_k	= minimum running time, or dwell time, of event k .
d_k^+	= run time supplement associated with actual stop at the previous station
b_k^{static}	= the fixed (if any) start time of event k
e_k^{static}	= the fixed (if any) end time of event k
r_k^{track}	= the planned track for event k
r_k^{fixed}	= the fixed (if any) track allocated to event k
S_k	= specifies if event k is a station event or line segment event.
Ψ	= threshold value to classify if a train has made a stop at a station, or not.
h_k	= specifies if event k has a scheduled earliest departure time (i.e. end time).
$l_{jt}^{\text{tracklength}}$	= the length of track t on segment $.j$
$l_i^{\text{trainlength}}$	= the length of train i .
ε	= delay tolerance threshold value.
M	= large constant.