

Leo Kroon

Rotterdam School of Management
Erasmus Universiteit Rotterdam, en
NS Reizigers, Proceskwaliteit en Innovatie
lkroon@rsm.nl

Lex Schrijver

Korteweg-de Vries Instituut voor Wiskunde
Universiteit van Amsterdam, en
Centrum Wiskunde & Informatica
lex@cwi.nl

Spoornetwerken

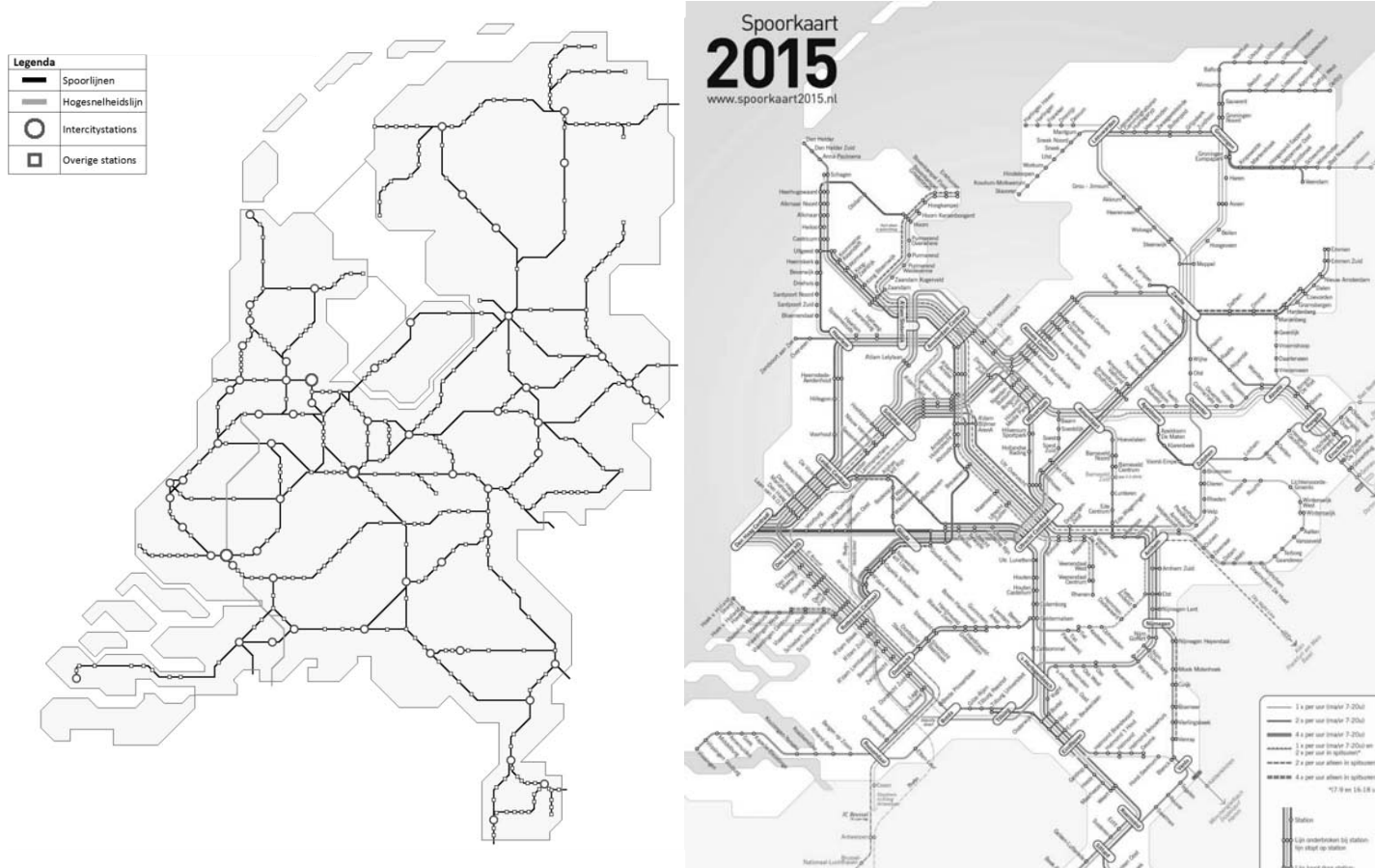
Leo Kroon en Lex Schrijver hebben samen met teams bij NS, EUR en CWI algoritmiëk ontwikkeld waarmee de dienstregeling van NS gepland kan worden, alsmede de bijbehorende routing van treinen door stations, de materieelomloop, en de personeelsdiensten. Deze algoritmen worden regelmatig door NS gebruikt, sinds hun eerste integrale toepassing ten behoeve van het ‘Spoorboekje 2007’. Hierbij ging de dienstregeling van NS eind 2006 grondig op de schop. Binnen deze algoritmen spelen netwerken, in verschillende vormen, een belangrijke rol.

Wie aan spoorwegen denkt, denkt al snel aan netwerken. Daarbij springt het netwerk van spoortrajecten als eerste in het oog, zie Fi-

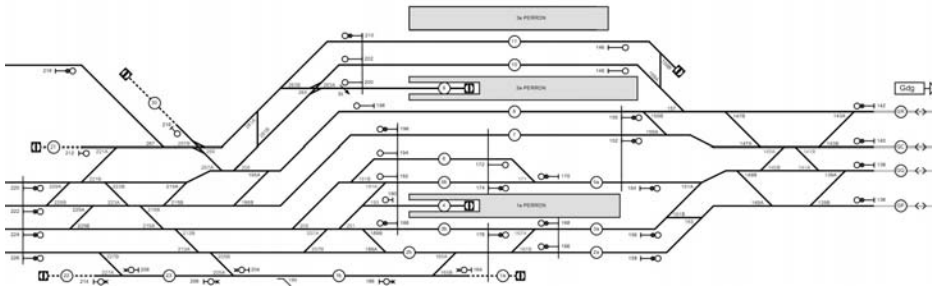
guur 1 (links): een netwerk van stations die door middel van sporen met elkaar verbonden zijn. Een meer gedetailleerd netwerk, dat

over het netwerk van spoortrajecten heen ligt, is het netwerk van spoorlijnen, zie Figuur 1 (rechts): de rechtstreekse treinverbindingen, ieder met een herkomst, een bestemming, een frequentie, en een lijst van stations waar gestopt wordt. En als je verder inzoomt, bijvoorbeeld op station Gouda in het netwerk van spoortrajecten, dan zie je het veel gedetailleerdere spoornetwerk in Figuur 2.

Minder zichtbaar zijn de virtuele wiskundige netwerken die een onmisbare rol spelen



Figuur 1 Het netwerk van spoortrajecten (links) en het netwerk van spoorlijnen (rechts).



Figuur 2 Infrastructuur van station Gouda.

Zwolle	↗ 9 56	10 26	↗ 10 56	11 26	↗ 11 56	12 26	↗ 12 56	13 26	↗ 13 56	14 26
Dalfsen	10 05	10 34	11 05	11 34	12 05	12 34	13 05	13 34	14 05	14 34
Ommen	10 15	10 44	11 15	11 44	12 15	12 44	13 15	13 44	14 15	14 44
Mariënberg		10 51		11 51		12 51		13 51		14 51
Hardenberg	10 26	10 57	11 26	11 57	12 26	12 57	13 26	13 57	14 26	14 57
Gramsbergen		11 01		12 01		13 01		14 01		15 01
Coevorden	10 36	11 07	11 36	12 07	12 36	13 07	13 36	14 07	14 36	15 07
Dalen		11 11		12 11		13 11		14 11		15 11
Nieuw Amsterdam	10 44	11 18	11 44	12 18	12 44	13 18	13 44	14 18	14 44	15 18
Emmen Barges		11 23		12 23		13 23		14 23		15 23
Emmen	A ↗ 10 51	11 28	↗ 11 51	12 28	↗ 12 51	13 28	↗ 13 51	14 28	↗ 14 51	15 28

Figuur 3 Dienstregeling Zwolle – Emmen.

bij de operationele planning van het spoorstelsel, zoals de planning van de dienstregeling, de materieelomloop en de personeelsinzet, en de routing van treinen over emplacementen.

De klassieke netwerkbegrippen *potentiaal*, *spanning* en *stroom*, bekend uit de elektriciteitsleer, blijken ook van belang in de logistiek van de spoorwegen. Deze begrippen zijn uiteraard belangrijk in de energievoorziening aan het elektrische spoorwagematerieel. Maar, zoals we zullen zien, potentialen, spanningen en stromen in netwerken vormen ook

een belangrijk gereedschap bij het *plannen* van verschillende spoorprocessen. De inzet op de volgende bladzijde geeft enkele relevante definities en eigenschappen.

We richten ons in het bijzonder op twee fasen in het planningsproces:

1. het bepalen van de dienstregeling,
2. het bepalen van de materieelomloop.

In geval 1 worden de *eisen* waaraan de dienstregeling moet voldoen voorgesteld als een netwerk, en de dienstregeling als een potentiaal op dit netwerk. De punten van dit netwerk representeren de vast te stellen vertrek- en aankomsttijden. Deze modellering gaat terug op het werk van Van der Aalst, Van Hee en Voorhoeve [12]. In geval 2 wordt de *dienstregeling zelf* als netwerk voorgesteld, en de materieelomloop als stroom in dit netwerk. Deze onderwerpen komen aan de orde in de volgende twee paragrafen. De laatste paragraaf geeft een beknopte beschrijving van nog enkele andere onderwerpen.

De dienstregeling en het netwerk van eisen

Het Basis Uur Patroon

De basis van de Nederlandse dienstregeling is het *Basis Uur Patroon* (BUP): de dienstre-

geling herhaalt zich elke 60 minuten. Dus als een bepaalde trein op tijdstip *t* van station *s* vertrekt, dan vertrekt er ook op de tijdstippen $\dots, t - 120, t - 60$ en $t + 60, t + 120, \dots$ een vergelijkbare trein van station *s*, met dezelfde bestemmingen. Hieraan worden spits-treinen toegevoegd, en in de avond en nacht en in het weekend zijn er andere afwijkingen, maar het BUP is het onderliggende patroon. Veel treinseries hebben overigens een 30-minutenfrequentie, maar dat kan eenvoudig in het model worden meegenomen.

Het betekent bijvoorbeeld dat elk uur, om 26 minuten over het hele uur, een stoptrein van Zwolle naar Emmen vertrekt, van 's morgens vroeg tot 's avonds laat. Zie Figuur 3. (De voorbeelden in dit artikel zijn gebaseerd op de situatie op het Nederlandse spoorweg-net, maar komen niet steeds overeen met de huidige (2015) dienstregeling en overige omstandigheden.)

Elke uurlijkse vertrek- en aankomsttijd kan worden gegeven door de minuut in het uur, dus door een element uit de verzameling $\mathbb{Z}_{60} = \mathbb{Z}/60\mathbb{Z}$ van getallen modulo 60. Het vaststellen van de tijden in het BUP komt dan neer op het toekennen van elementen uit \mathbb{Z}_{60} aan variabelen. Wat zijn deze variabelen en waaraan moeten ze voldoen?

De uurlijkse treinen

Voor elke uurlijkse treinserie en elke af te leggen etappe (rit) definiëren we twee punten, namelijk een punt voor de vertrektijd en een punt voor de aankomsttijd. Bekijk, als voorbeeld, de treinserie 15, een uurlijkse intercitytrein van Amsterdam naar Amersfoort. Zie Figuur 4. Deze stopt onderweg in Hilversum. Dit geeft vier punten in het netwerk:

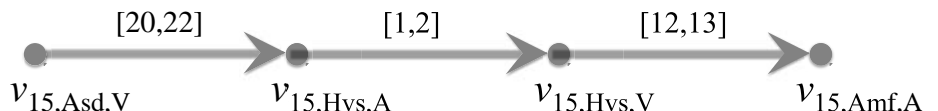
$$v_{15,Asd,V}, v_{15,Hvs,A}, v_{15,Hvs,V} \text{ en } v_{15,Amf,A} \quad (1)$$

Hierin zijn *Asd*, *Hvs* en *Amf* de afkortingen voor Amsterdam, Hilversum en Amersfoort, en staan *V* en *A* voor vertrektijd en aankomsttijd. Vervolgens voeren we ook pijlen in, overeenkomend met de verschillende etappes en stationnementen van treinserie 15:

$$(v_{15,Asd,V}, v_{15,Hvs,A}), (v_{15,Hvs,A}, v_{15,Hvs,V}) \text{ en } (v_{15,Hvs,V}, v_{15,Amf,A}) \quad (2)$$

treinnummer	4637	1539
Amsterdam Centraal	11 20	11 27
Amsterdam Muiderpoort	11 26	
Diemen	11 30	
Hooftdorp Schiphol A		
Schiphol A		
Amsterdam Zuid		
Amsterdam RAI		
Duivendrecht		
Duivendrecht		
Diemen Zuid		
Weesp	A 11 36	
Weesp		
Naarden-Bussum		
Bussum Zuid		
Hilversum Noord		
Hilversum	A 11 48	
Hilversum		
Baarn		
Amersfoort	A 11 49	12 02

Figuur 4 Treinen van serie 46 en 15 op het traject Amsterdam – Amersfoort.



Figuur 5 Eisen aan tijden in treinserie 15 Amsterdam – Hilversum – Amersfoort.

Stroom en spanning

In de context van dit artikel is een *netwerk* (meestal) hetzelfde als een *gerichte graaf*: een geordend paar $N = (V, E)$, waarbij V een eindige verzameling is en E een deelverzameling van $V \times V$ met $(v, v) \notin E$ voor elke $v \in V$. De elementen van V en E heten de *punten* respectievelijk de *pijlen* van het netwerk. Een pijl (u, v) *vertrekt uit* u en *komt aan in* v .

Als R een additieve abelse groep is, dan is een *stroom* een functie $f : E \rightarrow R$ die voldoet aan de *wet van stroombehoud*:

$$\sum_{e \in \delta^{\text{in}}(v)} f(e) = \sum_{e \in \delta^{\text{uit}}(v)} f(e) \text{ voor elke } v \in V.$$

Hierin is $\delta^{\text{in}}(v)$ de verzameling in v aankomende pijlen, en $\delta^{\text{uit}}(v)$ de verzameling uit v vertrekkende pijlen.

Een functie $g : E \rightarrow R$ heet een *spanning* als er een functie $p : V \rightarrow R$ (een *potentiaal*) bestaat zo dat

$$g(u, v) = p(v) - p(u) \text{ voor elke } (u, v) \in E.$$

De verzameling S van stromen en de verzameling T van spanningen vormen elk een deelmoduul van het R -moduul R^E .

Als R een commutatieve ring met eenheid is, dan voldoen S en T aan de volgende dualiteit. Het *inwendig product* $\langle x, y \rangle$ van elementen $x, y \in R^E$ is als gebruikelijk

$$\langle x, y \rangle := \sum_{e \in E} x(e)y(e).$$

Voor $X \subseteq R^E$ kan men dan een orthogonaal complement X^\perp definiëren:

$$X^\perp := \{y \in R^E \mid \langle x, y \rangle = 0 \text{ voor elke } x \in X\}.$$

Dan is het niet moeilijk om te bewijzen:

$$T = S^\perp \text{ en } S = T^\perp.$$

Ook geldt dat als B een opspannende boom in N is, dan is elke functie op de pijlen *in* B uniek uit te breiden op de pijlen *buiten* B zodat een spanning ontstaat. Evenzo kan elke functie op de pijlen *buiten* B uniek worden uitgebreid op de pijlen *in* B zodat een stroom ontstaat.

Een basisalgoritme uit de besliskunde, geïnitieerd door Ford en Fulkerson [3], en *sterk-polynomiaal* gemaakt door Tardos [10], geeft het volgende. Als R de ring \mathbb{Z} van gehele getallen is, en er voor elke $(u, v) \in E$ een convexe deelverzameling $I(u, v)$ van \mathbb{Z} gegeven is, dan kan snel (dat wil zeggen, binnen polynomiaal-begrensde tijd) een stroom f gevonden worden zo dat

$$f(u, v) \in I(u, v) \text{ voor elke } (u, v) \in E$$

(mits zo'n stroom bestaat). Sterker: als ook nog een functie $c : E \rightarrow \mathbb{Z}$ (de *kostenfunctie*) gegeven is, dan kan snel zo'n stroom f gevonden worden waarvoor $\langle c, f \rangle$ minimaal is. Hetzelfde geldt voor spanningen.

Hier is $[a, b]$ het interval van gehele getallen x met $a \leq x \leq b$ modulo 60. Voorgeschreven is dus dat de rijtijd van Amsterdam naar Hilversum minimaal 20 minuten bedraagt, en maximaal 22 minuten. De twee andere eisen hebben een vergelijkbare betekenis. We kunnen deze eisen aan de reijtijden grafisch weergeven zoals in Figuur 5. Een aan deze eisen voldoende dienstregeling correspondeert dus met een \mathbb{Z}_{60} -waardige potentiaal p zo dat $p(v) - p(u) \in I(u, v)$ voor elke pijl (u, v) , oftewel met een \mathbb{Z}_{60} -waardige spanning g zo dat $g(u, v) \in I(u, v)$ voor elke pijl (u, v) .

Voor de dienstregeling van treinserie 15 die weergegeven wordt in Figuur 4 geeft dit

$$\begin{aligned} p(v_{15, \text{Asd}, v}) &= 27, \\ p(v_{15, \text{Hvs}, A}) &= 48, \\ p(v_{15, \text{Hvs}, v}) &= 49, \\ p(v_{15, \text{Amf}, A}) &= 2. \end{aligned} \tag{4}$$

Deze waarden voldoen inderdaad aan de eisen. Immers:

$$\begin{aligned} g(v_{15, \text{Asd}, v}, v_{15, \text{Hvs}, A}) &= p(v_{15, \text{Hvs}, A}) - p(v_{15, \text{Asd}, v}) \\ &= 48 - 27 \equiv 21 \in [20, 22], \\ g(v_{15, \text{Hvs}, A}, v_{15, \text{Hvs}, v}) &= p(v_{15, \text{Hvs}, v}) - p(v_{15, \text{Hvs}, A}) \\ &= 49 - 48 \equiv 1 \in [1, 2], \\ g(v_{15, \text{Hvs}, v}, v_{15, \text{Amf}, A}) &= p(v_{15, \text{Amf}, A}) - p(v_{15, \text{Hvs}, v}) \\ &= 2 - 49 \equiv 13 \in [12, 13]. \end{aligned} \tag{5}$$

Dit zijn enkele eisen voor één treinserie. Alle treinseries samen geven een groot aantal los van elkaar liggende, gerichte paden, elk corresponderend met een treinserie, heen of terug.

Aansluitingen

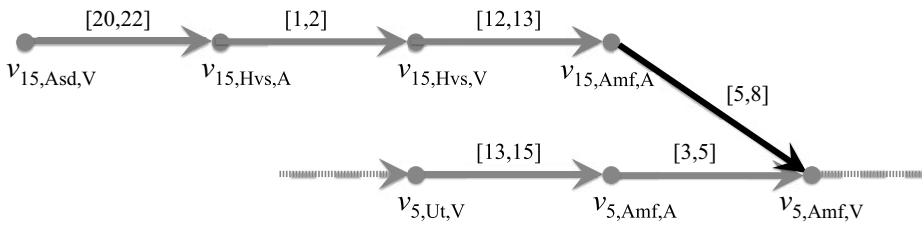
Samenhang *tussen* de treinseries ontstaat door de gewenste overstapmogelijkheden. Zo is er een uurlijkse intercity treinserie 5 van Rotterdam via Amersfoort naar Groningen, waarop de treinserie 15 uit Amsterdam in Amersfoort dient aan te sluiten. Daartoe moet de trein naar Groningen minimaal 5 en maximaal 8 minuten na aankomst van de trein uit Amsterdam van Amersfoort vertrekken. Hierdoor ontstaat een nieuwe pijl, tussen punten van twee verschillende treinseries: $(v_{15, \text{Amf}, A}, v_{5, \text{Amf}, v})$, met als interval

Voor elke etappe zijn onder- en bovengrenzen voorgeschreven voor de rijtijden, en voor elke tussenstop zijn onder- en bovengrenzen voorgeschreven voor de halteertijd.

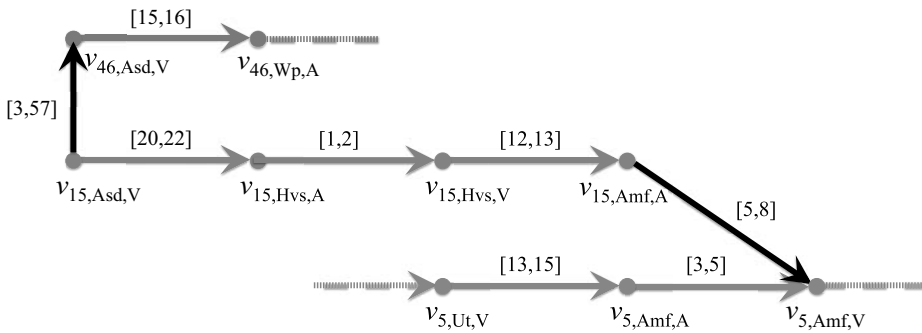
Dus voor elke pijl (u, v) is er een interval $I(u, v)$ voorgeschreven waarin het verschil van opvolgende vertrek en aankomsttij-

den dient te liggen. Zoals in het Amsterdam – Amersfoort voorbeeld:

$$\begin{aligned} I(v_{15, \text{Asd}, v}, v_{15, \text{Hvs}, A}) &= [20, 22], \\ I(v_{15, \text{Hvs}, A}, v_{15, \text{Hvs}, v}) &= [1, 2], \\ I(v_{15, \text{Hvs}, v}, v_{15, \text{Amf}, A}) &= [12, 13]. \end{aligned} \tag{3}$$



Figuur 6 Aansluiting van treinserie 15 op treinserie 5 in Amersfoort.



Figuur 7 Voorkomen van conflict tussen treinseries 15 en 46 op Amsterdam - Weesp.

$$I(v_{15,Amf,A}, v_{5,Amf,V}) = [5, 8]. \quad (6)$$

Dit is grafisch weergegeven in Figuur 6. Alle overstapvoorwaarden tezamen geven al veel meer samenhang aan het netwerk.

Voorkomen van conflicten

Daarnaast moeten zogenaamde *conflicten* worden vermeden: twee treinen mogen niet gelijktijdig van hetzelfde spoor gebruik maken, bijvoorbeeld op een kruising of op enkelspoor, maar ook langs een traject: om te zorgen dat er voldoende ruimte tussen twee treinen wordt gepland, wordt een minimale *opvolgtijd* van drie minuten geëist, dit ook om kleine vertragingen op te kunnen vangen.

Als voorbeeld de treinserie 46 van Amsterdam naar Lelystad, zie Figuur 4. Deze gebruikt op het traject Amsterdam - Weesp hetzelfde spoor als de bovengenoemde treinserie 15 van Amsterdam naar Amersfoort. Op dit traject moeten de treinen een minimale afstand van drie minuten in tijd hebben. Dit geeft dan de pijl $(v_{15,Asd,V}, v_{46,Asd,V})$ met

$$I(v_{15,Asd,V}, v_{46,Asd,V}) = [3, 57]. \quad (7)$$

Hierdoor wordt dus voorkomen dat het verschil in de vertrektijden vanuit Amsterdam gelijk is aan 58, 59, 0, 1 of 2 (modulo 60) – dan zouden de treinen te dicht op elkaar zitten. Zie Figuur 7. Deze constraints kunnen sterker

gemaakt worden door ook de minimale rijtijdsverschillen tussen de treinen mee te nemen.

Op deze manier kunnen ook de beperkingen op enkelsporige trajecten en kruisingen worden beschreven. Deze eisen zorgen voor de grootste toevoeging van pijlen aan het netwerk, en, ondanks de ruime intervallen, ook voor de grootste toevoeging van complexiteit.

Periodic Event Scheduling

Het bepalen van een dienstregeling die aan de gestelde eisen voldoet komt nu dus neer op het vinden van een potentiaal $p : V \rightarrow \mathbb{Z}_{60}$ die voldoet aan

$$p(v) - p(u) \in I(u, v) \quad (8)$$

voor elke pijl (u, v) .

Dat wil zeggen, een spanning $g : E \rightarrow \mathbb{Z}_{60}$ die voldoet aan

$$g(u, v) \in I(u, v) \text{ voor elke pijl } (u, v). \quad (9)$$

Dit *Periodic Event Scheduling Problem* (PESP) is in zijn algemeenheid een NP-volledig probleem: graafkleuring kan er eenvoudig toe worden gereduceerd. Dit in tegenstelling (als $NP \neq P$) tot het geval waarin we \mathbb{Z}_{60} zouden vervangen door \mathbb{Z} , waarmee, zoals we in de inzet meldden, het probleem snel oplosbaar zou zijn.

Je zou natuurlijk kunnen stellen dat elke \mathbb{Z} -oplossing ook een \mathbb{Z}_{60} -oplossing is, maar er

kan een \mathbb{Z}_{60} -oplossing bestaan zonder dat er een \mathbb{Z} -oplossing bestaat.

NP-volledigheid is echter niet het ultieme antwoord: NP-volledigheid heeft betrekking op de asymptotische moeilijkheid van het probleem, terwijl hier één concreet, eenduidig probleem voorligt. We beschrijven twee aanpakken om een potentiaal te vinden, als die bestaat.

Constraint propagation

We kunnen de eisen beschrijven in een $V \times V$ -matrix M waarvan de elementen deelverzamelingen van \mathbb{Z}_{60} zijn, dus als matrix in $\mathcal{P}(\mathbb{Z}_{60})^{V \times V}$. Als (u, v) een pijl is, definiëren we $M_{u,v} := I(u, v)$. Als (u, v) geen pijl is en $u \neq v$, dan $M_{u,v} := \mathbb{Z}_{60}$, en als $u = v$, dan $M_{u,v} := \{0\}$.

Dergelijke matrices kunnen vermenigvuldigd worden in de $(\cap, +)$ -algebra, als variant van de gebruikelijke $(+, \times)$ -algebra. Als A en B matrices zijn met elementen in $\mathcal{P}(\mathbb{Z}_{60})$, vervangen we $+$ door \cap en \times door $+$, en krijgen we het $(\cap, +)$ -product $A * B$ als:

$$(A * B)_{i,j} = (A_{i,1} + B_{1,j}) \cap \dots \cap (A_{i,n} + B_{n,j}), \quad (10)$$

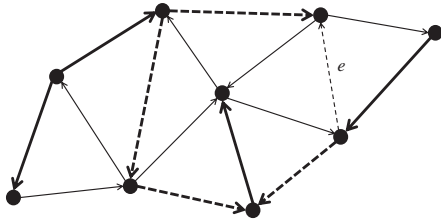
waarbij $X + Y := \{x + y \mid x \in X, y \in Y\}$ voor $X, Y \subseteq \mathbb{Z}_{60}$.

Het is niet moeilijk in te zien dat als $p(v) - p(u) \in A_{u,v}$ voor elke $u, v \in V$, dan ook $p(v) - p(u) \in (A^{*2})_{u,v}$ voor elke $u, v \in V$, waarbij A^{*2} het $(\cap, +)$ -kwadraat is van A . Bovendien geldt $(A^{*2})_{u,v} \subseteq A_{u,v}$ voor elke $u, v \in V$, doordat $A_{v,v} = \{0\}$ voor elke $v \in V$.

Daarom kunnen we, beginnend met $A := M$, en A iteratief vervangend door A^{*2} , de eisen beschreven in de initiële matrix M verscherpen, totdat geen van de elementen van de matrix A strikt kleiner wordt. Dit is een voorbeeld van *constraint propagation*.

Mocht na een aantal iteraties blijken dat een element in A gelijk is aan \emptyset , dan weten we dat er geen potentiaal (dat wil zeggen dienstregeling) bestaat die aan alle eisen voldoet. Dit is de terugmelding van het algoritme, mogelijk met een beschrijving van de stappen die tot \emptyset hebben geleid. Hierdoor wordt een minimale verzameling eisen geïdentificeerd waarvan er tenminste één verruimd moet worden om een dienstregeling mogelijk te maken.

Als alle elementen in de uiteindelijke matrix A niet-leeg zijn, geeft dit een instrument om gericht te zoeken naar een oplossing met een vertakkend zoekalgoritme. Overigens is



Figuur 8 De dikke (doorlopende of gestreepte) pijlen geven een opspannende boom weer. De gestreepte (dikke of dunne) pijlen geven de met pijl e gevormde fundamentele cykel C_e weer.

het bestaan van een toegelaten oplossing nog steeds niet gegarandeerd. Dit is de basis van het op het Centrum Wiskunde & Informatica door Schrijver en Steenbeek [9] voor NS en ProRail ontwikkelde programma CADANS, waarmee NS toekomstige dienstregelingen opstelt en ProRail toekomstige uitbreidingen van de infrastructuur doorrekent.

Zoeken in de duale ruimte

Een alternatieve manier voor het vinden van een toegelaten dienstregeling is de volgende. Daartoe verlaten we het modulo 60 model en zien we de intervallen $I(u, v)$ als intervallen van gehele getallen. Als $g : E \rightarrow \mathbb{Z}$ een toegelaten spanning is (dat wil zeggen, voldoende aan (9) mod 60), dan bestaat er een functie $k : E \rightarrow \mathbb{Z}$ zo dat

$$g(u, v) \in I(u, v) + 60k(u, v) \quad (11)$$

voor elke pijl (u, v) .

De modulo 60-voorwaarde wordt nu omzeild door er een geheel (nog vast te stellen) veelvoud van 60 bij op te tellen. Zoals we in de inzet al vermeldden, kan voor vaste k snel worden beslist of er een spanning $g : E \rightarrow \mathbb{Z}$ bestaat die aan (11) voldoet. We kunnen de zoekprocedure dan omdraaien: we zoeken naar een functie $k : E \rightarrow \mathbb{Z}$ zodanig dat er een spanning $g : E \rightarrow \mathbb{Z}$ bestaat die voldoet aan (11).

Nu geldt dat als $k, k' \in \mathbb{Z}^E$ en $k - k'$ tot T behoort (dat wil zeggen een spanning is), dan is voor elke oplossing g die hoort bij k , $g' := g + 60(k' - k)$ een oplossing die hoort bij

k' (want g' is een spanning en g' voldoet aan (11) met k vervangen door k'). Dus uit iedere klasse in de quotientruimte \mathbb{Z}^E/T hoeven we slechts één k te onderzoeken. Dit betekent dat we dan in feite zoeken in de duale ruimte S^* van \mathbb{Z} -lineaire functies $S \rightarrow \mathbb{Z}$ (doordat de functie $\mathbb{Z}^E \rightarrow S^*$ met $k \mapsto (x \mapsto \langle k, x \rangle)$ kern $S^\perp = T$ heeft, en dus $S^* \cong \mathbb{Z}^E/T$).

Nog steeds zijn er oneindig veel mogelijkheden om k te kiezen, maar de keuze kan beperkt worden tot een eindig aantal door een opspannende boom B te kiezen. We nemen daarbij aan dat het netwerk samenhangend is — dat is in Nederland inderdaad het geval.

Dan blijven er nog maar eindig veel mogelijkheden over voor de waarden van $k(e)$ als e niet in B zit. Dit kunnen we zien door de cykel C_e te beschouwen die e met B maakt (de fundamentele cykel). De intervallen langs de pijlen in C_e impliceren een onder- en een bovengrens voor $k(e)$. Zie Figuur 8.

Dan bevat elke klasse van \mathbb{Z}^E/T precies één $k : E \rightarrow \mathbb{Z}$ zo dat $k(e) = 0$ op elke e in B . Dit volgt uit het feit dat voor elke $k \in \mathbb{Z}^E$ precies één spanning g bestaat zodat $g(e) = k(e)$ voor elke e in B .

Het aantal mogelijkheden voor $k(e)$ is afhankelijk van de ruimte die deze cykel toelaat. Soms is $k(e)$ hierdoor volledig vastgelegd, soms is er keuze uit slechts twee mogelijkheden, soms ook is het aantal mogelijkheden groter. Maar in ieder geval wordt het zoekproces flink gereduceerd.

Een optimale dienstregeling

Uiteraard wil men niet alleen een toegelaten dienstregeling hebben, maar zelfs een optimale. Maar de vraag daarbij is natuurlijk ook: wat is het criterium voor een optimale dienstregeling? Er zijn verschillende, soms tegenstrijdige, criteria voor de kwaliteit van een dienstregeling.

Gegeven een toegelaten dienstregeling, kunnen de vertrek- en aankomsttijden met lineaire programmering geschoven worden, mits de volgordes der treinen op de trajecten vast zijn. In wiskundige termen: als we k in (11) vasthouden, kunnen we in polynomiale

tijd voor een willekeurige kostenfunctie op de spanning g (de verschillen van de tijden) een spanning van minimale totale kosten vinden. Dit wordt echter een stuk lastiger als we ook k vrij laten en willen optimaliseren — dan staan we weer voor een NP-volledig probleem.

Er kan bijvoorbeeld gestreefd worden naar een dienstregeling waarin de reistijden voor de reizigers zo kort mogelijk zijn. Dat wil zeggen: korte rijtijden, korte halteertijden, en korte overstaptijden. Maar een dergelijke dienstregeling is waarschijnlijk weinig robuust onder verstoringen: het dempend vermogen is beperkt. Daarom is een ander doel om een geschikt gekozen rijtjidspeeling in de dienstregeling in te bouwen. Deze kan gebruikt worden om een eventuele vertraging er weer uit te rijden. Kroon e.a. [5] beschrijven een stochastisch optimalisatiemodel om de rijtjidspeeling optimaal te verdelen over de dienstregeling. Daarnaast hebben Heidergott, Oldser en Van der Woude [4] modellen ontwikkeld op basis van (max, +)-algebra waarmee de robuustheid van een dienstregeling doorgerekend kan worden.

Bovendien kan in een onverstoorde situatie de rijtjidspeeling gebruikt worden om zo min mogelijk energie te gebruiken door zo véél mogelijk zonder tractie uit te rollen vóór iedere haltering. Met behulp van Pontryagins Maximum Principe [7] kan een snelheidsprofiel berekend worden waarbij onderweg zo min mogelijk energie gebruikt wordt. Het totale energieverbruik per rit is daarbij dalend in de geplande rijtijd: hoe langer de rijtijd, hoe meer er uitgerold kan worden. Het vinden van een geschikte rijtijd voor een rit is daardoor een trade-off tussen snelheid, robuustheid en energieverbruik.

Materieelomloop en dienstregelingsnetwerk

Als de dienstregeling is vastgesteld, zijn de bepaling van de materieelomloop en de personeelsinzet de volgende relevante logistieke problemen. We beperken ons hier in eerste instantie tot het eerste. We gaan hierbij uit van de volgende (inmiddels gewijzigde)



Figuur 9 Een koploper van type III.



Figuur 10 Een koploper van type IV.

treinnummer	2123	2127	2131	2135	2139	2143	2147	2151	2155	2159	2163	2167	2171	2175	2179	2183	2187	2191
Amsterdam V		6.48	7.55	8.56	9.56	10.56	11.56	12.56	13.56	14.56	15.56	16.56	17.56	18.56	19.56	20.56	21.56	22.56
Rotterdam A		7.55	8.58	9.58	10.58	11.58	12.58	13.58	14.58	15.58	16.58	17.58	18.58	19.58	20.58	21.58	22.58	23.58
Rotterdam V	7.00	8.01	9.02	10.03	11.02	12.03	13.02	14.02	15.02	16.00	17.01	18.01	19.02	20.02	21.02	22.02	23.02	
Roosendaal A	7.40	8.41	9.41	10.43	11.41	12.41	13.41	14.41	15.41	16.43	17.43	18.42	19.41	20.41	21.41	22.41	23.54	
Roosendaal V	7.43	8.43	9.43	10.45	11.43	12.43	13.43	14.43	15.43	16.45	17.45	18.44	19.43	20.43	21.43			
Vlissingen A	8.38	9.38	10.38	11.38	12.38	13.38	14.38	15.38	16.38	17.40	18.40	19.39	20.38	21.38	22.38			

treinnummer	2108	2112	2116	2120	2124	2128	2132	2136	2140	2144	2148	2152	2156	2160	2164	2168	2172	2176
Vlissingen V			5.30	6.54	7.56	8.56	9.56	10.56	11.56	12.56	13.56	14.56	15.56	16.56	17.56	18.56	19.55	
Roosendaal A			6.35	7.48	8.50	9.50	10.50	11.50	12.50	13.50	14.50	15.50	16.50	17.50	18.50	19.50	20.49	
Roosendaal V		5.29	6.43	7.52	8.53	9.53	10.53	11.53	12.53	13.53	14.53	15.53	16.53	17.53	18.53	19.53	20.52	21.53
Rotterdam A		6.28	7.26	8.32	9.32	10.32	11.32	12.32	13.32	14.32	15.32	16.32	17.33	18.32	19.32	20.32	21.30	22.32
Rotterdam V	5.31	6.29	7.32	8.35	9.34	10.34	11.34	12.34	13.35	14.35	15.34	16.34	17.35	18.34	19.34	20.35	21.32	22.34
Amsterdam A	6.39	7.38	8.38	9.40	10.38	11.38	12.38	13.38	14.38	15.38	16.40	17.38	18.38	19.38	20.38	21.38	22.38	23.38

Tabel 1 Dienstregeling Amsterdam – Vlissingen v.v.

dienstregeling van de treinen Amsterdam – Vlissingen v.v. Zie Tabel 1.

De treinen stoppen bij meer stations, maar Rotterdam en Roosendaal zijn de enige plaatsen waar onderweg materieel kan worden afgetrapt of bijgeplaatst. Dit kan ook aan de eindpunten Amsterdam en Vlissingen.

Ook gaan we, in eerste instantie, uit van één type twee-richtingtreinstellen, de zogenaamde koplopers type III (zie Figuur 9), die onderling gekoppeld kunnen worden tot langere treinen. Een treinstel dat afgetrapt is van een trein op een station, kan daarna weer bijgeplaatst worden aan een andere trein op dit station.

Per trein heeft NS voor elk van de deeltrajecten een schatting gemaakt van het aantal reizigers en dat vertaald in het minimale aantal gekoppelde treinstellen dat op dat deeltraject dient te rijden. Zie Tabel 2.

Met deze gegevens en condities is de vraag wat het minimale aantal koplopers is dat voor de dienst Amsterdam – Vlissingen moet worden ingezet.

Het dienstregelingsnetwerk

Een optimale materieelomloop kan worden bepaald door het probleem te modelleren als een stroomprobleem in het volgende virtuele netwerk, het dienstregelingsnetwerk [2, 13].

Figuur 11 geeft het dienstregelingsnetwerk weer voor de in Tabel 1 gegeven dienstregeling.

In het dienstregelingsnetwerk $N = (V, E)$ corresponderen de schuine pijlen met treinritten op een van de trajecten tussen Amsterdam, Rotterdam, Roosendaal en Vlissingen, terwijl de verticale pijlen corresponderen met een tijdsinterval op een van deze stations tussen twee opvolgende vertrek- of aankomsten. De pijlen lopen steeds met de tijd mee.

De punten van het netwerk zijn de eindpunten van deze treinritten — kruispunten van twee schuine pijlen ‘onderweg’ geven geen punten van het netwerk. De punten van het netwerk komen dan overeen met de 198 tijdstippen die in Tabel 1 vermeld staan. Daarnaast zijn twee extra punten toegevoegd, een beginpunt s en een eindpunt t .

Een materieelomloop is nu een stroom $f : E \rightarrow \mathbb{Z}_+$, waarbij $f(e)$ de volgende betekenis heeft. Als e correspondeert met een treinrit, dan geeft $f(e)$ het aantal gekoppelde koplopers aan dat voor deze treinrit wordt ingezet. Als e een tijdsinterval tussen twee opvolgende gebeurtenissen op station X voorstelt, dan is $f(e)$ het aantal koplopers dat op station X aanwezig is tijdens dit interval. In s en t hoeft de stroom niet aan de wet van stroombehoud te voldoen — dit kan desge-

wenst alsnog bewerkstelligd worden door s en t te identificeren.

Bij elke pijl e is de vraagfunctie $d(e)$ gezet: het minimum aantal koplopers zoals gegeven in Tabel 2. Als e niet met een treinrit correspondeert, dan is $d(e) = 0$.

Het aantal in te zetten koplopers is dan gelijk aan de minimale hoeveelheid stroom die uit s vertrekt en die op elke pijl aan de ondergrens voldoet. Klassieke netwerk- of lineaire programmeringsmethoden leveren binnen een fractie van een seconde een optimale materieelomloop op die aan alle ondergrenzen voldoet. Voor het netwerk in Figuur 11 blijkt 22 het minimale aantal in te zetten koplopers te zijn.

Deze methode kan worden aangepast en uitgebreid naar verschillende varianten van het probleem. Door de vier pijlen aankomend in t per station te koppelen aan de vier pijlen vertrekkend uit s kan bijvoorbeeld een volledig periodieke omloop worden bepaald. In plaats van het aantal koplopers te minimaliseren, kan men ook het aantal gemaakte *bakkilometers* minimaliseren, of een lineaire combinatie hiervan met het aantal koplopers. Bovendien kunnen bovengrenzen worden gesteld aan de stroom, als bovengrens aan de lengte van de trein of aan de opstelcapaciteit op een station. De methode blijft ook snel

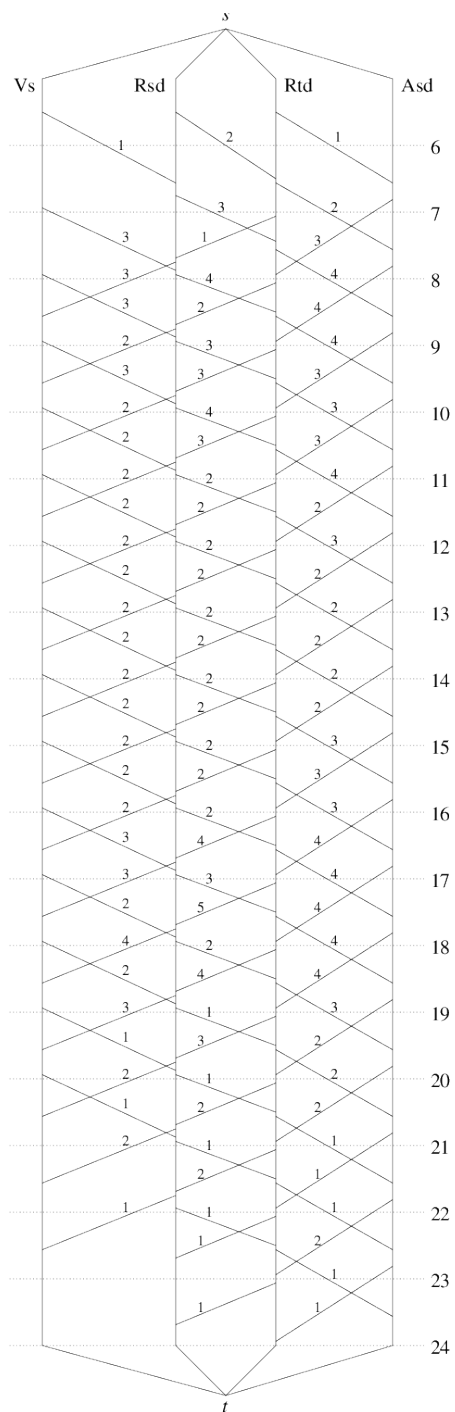
treinnummer	2123	2127	2131	2135	2139	2143	2147	2151	2155	2159	2163	2167	2171	2175	2179	2183	2187	2191
Amsterdam – Rotterdam		3	4	3	3	2	2	2	2	3	4	4	4	2	2	1	2	1
Rotterdam – Roosendaal	1	2	3	3	2	2	2	2	2	4	5	4	3	2	2	1	1	
Roosendaal – Vlissingen	3	2	2	2	2	2	2	2	2	3	4	3	2	2	1			

treinnummer	2108	2112	2116	2120	2124	2128	2132	2136	2140	2144	2148	2152	2156	2160	2164	2168	2172	2176
Vlissingen – Roosendaal			1	3	3	3	2	2	2	2	2	2	3	2	2	1	1	
Roosendaal – Rotterdam		2	3	4	3	4	2	2	2	2	2	2	3	2	1	1	1	1
Rotterdam – Amsterdam	1	2	4	4	3	4	3	2	2	3	3	4	4	3	2	1	1	1

Tabel 2 Ondergrenzen aan het aantal koplopers per deeltraject.

als de materieelomloop het hele Nederlandse intercity-netwerk omvat.

Koppelen van verschillende materieeltypes
 Het probleem wordt echter een stuk lastiger als er treinstellen van meer dan één type materieel kunnen worden gekoppeld tot een trein. Als voorbeeld geven we weer de koplopers, waar naast type III in Figuur 9, elk bestaande uit 3 bakken (rijtuigen), ook type IV



Figuur 11 Het dienstregelingsnetwerk Amsterdam – Vlissingen (Vs = Vlissingen, Rsd = Roosendaal, Rtd = Rotterdam, Asd = Amsterdam).

bestaat, elk bestaande uit 4 bakken, zie Figuur 10.

De *drietjes* en *viertjes* kunnen onderling gekoppeld worden. Met combinaties van drietjes en viertjes kan meer *op maat* worden gereden, maar het optimaliseren van de materieelomloop, en het herstellen van verstoringen in de materieelomloop wordt lastiger.

Het omloopprobleem vertaalt zich nu in een *multi-commodity* stroomprobleem op het dienstregelingsnetwerk, waarin ondergrenzen worden gesteld aan de *combinatie* van drietjes en viertjes: de drietjes vormen op zichzelf een stroom, evenals de viertjes, maar het totale aantal aangeboden zitplaatsen per trein wordt bepaald door een lineaire combinatie van de aantallen drietjes en viertjes in de trein. Dit maakt het probleem op zichzelf al lastiger – in feite NP-volledig. De constraint matrix is niet meer totaal unimodulair, en lineaire programmering levert daardoor niet meer direct een geheeltallige oplossing op.

En daar komt nog iets bij: de volgorde van de drietjes en viertjes in een trein is nu ook van belang. Als bijvoorbeeld een drietje in een station moet worden afgetrapt, dan dient dit zich aan het eind van de trein te bevinden. Bijplaatsen gebeurt steeds aan de kop van de trein. Het toegelaten zijn van een stroom

wordt dus niet alleen bepaald door de aantallen drietjes en viertjes per trein, maar ook door het *woord* gevormd door de volgorde van drietjes en viertjes, zoals 343 of 334. Door toevoeging van binaire *compositievariabelen* die de toegelaten treincomposities beschrijven, tezamen met *transitievariabelen* voor de toegelaten overgangen, kunnen dergelijke problemen met een geheeltallig optimaliseringspakket zoals CPLEX binnen korte tijd worden opgelost (zie [8]). Hierbij helpt het dat de transitievariabelen vanzelf geheeltallig zijn als de compositievariabelen dat zijn.

En er is nog meer

Reisadvies

Het dienstregelingsnetwerk kan ook gebruikt worden voor het geven van reisadviezen aan reizigers, zoals door middel van de NS Reisplanner. Een reis van een vertrekstation (inclusief een vertrektijdstop) naar een bestemming kan gerepresenteerd worden als een gericht pad in een netwerk van dezelfde structuur als het eerder aangegeven dienstregelingsnetwerk, maar waarin nu *alle* haltingen van *alle* treinen zijn meegenomen. Een zo kort mogelijke reis is een kortste pad van vertrekstation en -tijd naar de bestemming.

Reisadvies

Gilze-Rijen → Buitenpost
 Vertrek dinsdag 23 juni 2015 rond 09:00

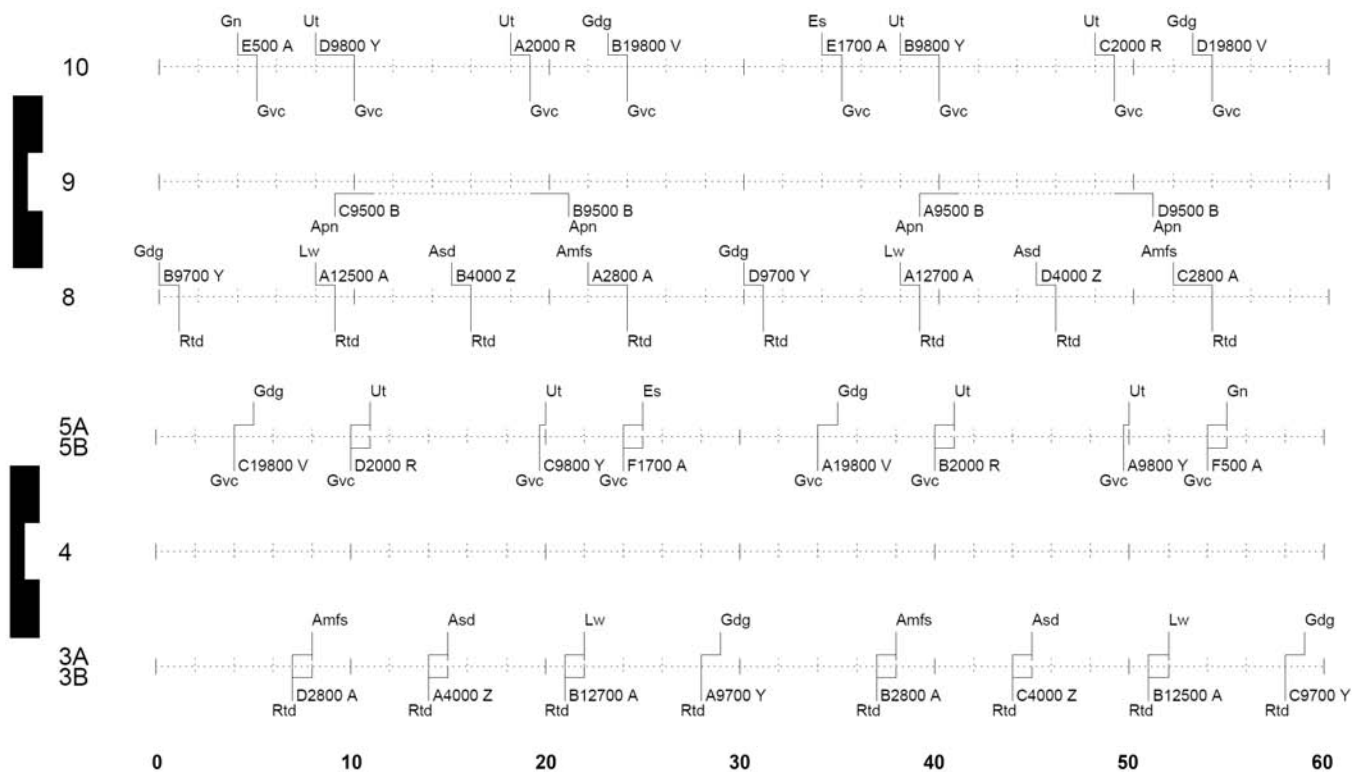
Heenreis: Vertrek 09:32 Aankomst 13:14

Dinsdag 23 juni 2015

Dit reisadvies houdt geen rekening met wijzigingen in de treindienst op de dag zelf. Plan voor een actueel reisadvies uw reis vlak voor vertrek.

	Tijd	Station / Halte	Spoor	Richting	Reisdetails
	09:32	Gilze-Rijen	4	Utrecht Centraal	Sprinter (NS)
	10:00	's-Hertogenbosch	3b		
	10:09	's-Hertogenbosch	4	Schiphol	Intercity (NS)
	10:37	Utrecht Centraal	7		
	10:50	Utrecht Centraal	11	Groningen	Intercity (NS)
	12:44	Groningen	4b		
	12:51	Groningen	6a	Leeuwarden	Stoptrein (Arriva)
	Reisdetails: Geen AVR-NS				
	13:14	Buitenpost	2		

Figuur 12 Reisadvies Gilze-Rijen – Buitenpost.



Figuur 13 Perrontoewijzing voor een deel van de sporen van station Gouda.

Wanneer de reiziger zo min mogelijk wil overstappen, dan moet het dienstregelingsnetwerk echter uitgebreid worden met extra punten en pijlen: de punten worden onderverdeeld in treinpunten en stationspunten. Treinpijlen verbinden achtereenvolgende treinpunten behorend bij dezelfde trein. Stationspijlen verbinden in de tijd opeenvolgende stationspunten. Treinpijlen representeren de verplaatsing van een reiziger in een trein van het beginpunt van de pijl (station en tijd) naar het eindpunt van de pijl (station en tijd). Stationspijlen representeren het wachten van een reiziger op een station tijdens een overstap. In- en uitstappijlen verbinden treinpunten en stationspunten met elkaar. Nu is een reis met zo weinig mogelijk overstaps te vinden als een pad van herkomst naar bestemming dat een minimaal aantal in- en uitstappijlen gebruikt.

De Reisplanner van NS is gebaseerd op het werk van Tulp [11]. De Reisplanner geeft meestal meerdere mogelijkheden om van herkomst naar bestemming te komen: de snelste mogelijkheid, en ook eventuele langere reis mogelijkheden maar met minder overstaps.

Diensten voor machinisten en conducteurs
Net zoals een reis van een reiziger een pad is in het (uitgebreide) dienstregelingsnetwerk,

is een dienst van een machinist of conducteur dat ook. Daarbij dient een dergelijk pad wel aan een aantal extra randvoorwaarden te voldoen: zo mag het pad niet te lang zijn (in tijd), het moet een pauze op een geschikt punt (in tijd en plaats) bevatten, en het moet weer terugkeren naar zijn beginpunt (in plaats). Daarnaast moet ook nog aan constraints op standplaatsniveau voldaan worden: zo mogen per standplaats het aantal diensten en de gemiddelde dienstlengte niet te groot zijn.

Het complete planningsprobleem voor de diensten van de machinisten op één dag kan dan beschreven worden als het vinden van een zo klein mogelijke set van toegelaten paden in het dienstregelingsnetwerk, die gezamenlijk alle treinpijlen in het dienstregelingsnetwerk overdekken en gezamenlijk aan de standplaatsconstraints voldoen. Het resulterende probleem is een *Set Covering*-probleem met extra randvoorwaarden. Aangezien het aantal toegelaten paden exponentieel is in het aantal ritten in de dienstregeling, worden dergelijke modellen meestal met dynamische kolomgeneratie opgelost, zie [1].

Routing in stations

Bij het zoeken naar een dienstregeling is tot nu toe buiten beschouwing gelaten of het mogelijk is de diverse treinen op een stationsem-

placement zo te routeren dat er geen conflicten ontstaan.

Gegeven de vertrek- en aankomsttijden van de treinen op een station, kan het routeren van de treinen door dat station gebaseerd worden op lijsten per trein met alle mogelijke routes die een trein door dat station kan volgen. Hierbij is een route een combinatie van een fysieke route (een reeks sporen en wissels) en een tijdstip. Het tijdstip is zo gekozen, dat de aankomst- en vertrektijd van een trein langs het perron gelijk is aan de vastgestelde aankomst- en vertrektijd. Een route kan gezien worden als een gericht pad in een *time-expanded* versie van een emplacement zoals in Figuur 2 wordt weergegeven.

Voor het oplossen van dit routeringsprobleem kan gebruik gemaakt worden van het ongerichte *conflictenetwerk* $K = (V, E)$: iedere combinatie van een trein en een mogelijke route voor die trein wordt in het conflictenetwerk weergegeven door een enkel punt. Twee punten worden met elkaar verbonden als de bijbehorende trein-route combinaties met elkaar conflicteren. Hierbij conflicteren twee routes met elkaar als ze één of meerdere infrastructuur-elementen (sporen, wissels, et cetera) tegelijk willen bezetten of reserveren.

Het probleem is nu het selecteren van precies één punt per trein, zodanig dat géén van de geselecteerde punten met elkaar verbonden zijn, en zodanig dat het totale gewicht van de geselecteerde punten maximaal is. Hierbij geeft het gewicht van een punt aan hoe gewenst het is dat de bijbehorende route aan de bijbehorende trein wordt toegewezen. Dit is een zogenaamd *Node Packing*-probleem.

Als $z_{t,r}$ een binaire beslissingsvariabele is die weergeeft of route r wel of niet aan trein t wordt toegewezen, dan worden de constraints als volgt weergegeven:

$$\begin{aligned} z_{t,r} + z_{t',r'} &\leq 1 \\ \forall \{(t,r), (t',r')\} &\in E. \end{aligned} \quad (12)$$

Voor het oplossen van dit probleem helpt het als niet slechts *paren* van verbonden punten verboden worden, maar als de *maximale cliques* in het conflictennetwerk gebruikt


worden. Daarmee worden de constraints als volgt:

$$\begin{aligned} \sum_{(t,r) \in c} z_{t,r} &\leq 1 \\ \forall c \subseteq K : c &\text{ is een maximale clique.} \end{aligned} \quad (13)$$

Als het doel is om het totale gewicht van de geselecteerde punten te maximaliseren, dan kan CPLEX instanties met een periode van één uur zelfs voor de grotere stations in korte tijd oplossen. Complexer wordt het als het doel is om de robuustheid van de routing te maximaliseren. Dan is het zaak om kruisende treimbewegingen zoveel mogelijk te vermijden, en de overblijvende kruisende treimbewegingen zoveel mogelijk in tijd te spreiden. Dit leidt tot een doelfunctie met kwadratische termen (die wel weer te lineariseren is). Figuur 13 geeft een perrontoewijzing voor een deel van de sporen van station Gouda.

Tot slot

In dit artikel hebben we slechts een topje van de ijsberg kunnen laten zien van de toepassingen van wiskundige netwerken ten behoeve van de optimalisatie van spoorwegsyste-men. Een deel van de in het voorgaande beschreven modellen wordt inmiddels succesvol toegepast in de *planningsprocessen* bij NS en ProRail, zie [6].

De volgende uitdaging is om deze modellen ook geschikt te maken voor de *real-time bijsturing* bij vertragingen of verstoringen. De grootste uitdaging daarbij is de beschikbare rekentijd: een model dat een uur nodig heeft om uit te rekenen wat er over één minuut moet gebeuren is duidelijk niet bruikbaar in de praktijk. Daarnaast is er in de bijsturing inherent sprake van een grotere onzekerheid dan in de planning. Op dit gebied zijn inmiddels significante stappen vooruit gezet, maar er zijn ook nog veel problemen op te lossen. 

Referenties

- 1 E.J.W. Abbink, M. Fischetti, L.G. Kroon, G. Timmer en M.J.C.M. Vromans, Reinventing Crew Scheduling at Netherlands Railways, *Interfaces* 35 (2005), 393–401.
- 2 T.E. Bartlett, An algorithm for the minimum number of transport units to maintain a fixed schedule, *Naval Research Logistics Quarterly* 4 (1957), 139–149.
- 3 L.R. Ford en D.R. Fulkerson, Maximal flow through a network, *Canadian Journal of Mathematics* 8 (1956), 399–404.
- 4 B. Heidergott, G.J. Olsder en J. van der Woude, *Modeling and Analysis of Synchronized Systems: A Course on Max-Plus Algebra and Its Applications*, Princeton University Press, Princeton, NJ, 2005.
- 5 L.G. Kroon, G. Maróti, M. Retel Helmrich, M.J.C.M. Vromans en R. Dekker, Stochastic Improvement of Cyclic Railway Timetables, *Transportation Research Part B* 42 (2008), 553–570.
- 6 L.G. Kroon, D. Huisman, E.J.W. Abbink, P.J. Fioole, M. Fischetti, G. Maróti, A. Schrijver, A. Steenbeek en R.J. Ybema, The New Dutch Timetable: The OR Revolution, *Interfaces* 39(1) (2009), 6–17.
- 7 L.S. Pontryagin, V.G. Boltyanskii, R.V. Gamkrelidze en E.F. Misiichenko, *The Mathematical Theory of Optimal Processes*, Interscience Publishers, New York, 1962.
- 8 A. Schrijver, Planning van materieelomlopen, Technical Report, Centrum Wiskunde & Informatica, Amsterdam, 2003, www.cwi.nl/~lex/pvm.pdf.
- 9 A. Schrijver en A. Steenbeek, Spoorwegdienstregelingontwikkeling, Technical Report, Centrum Wiskunde & Informatica, Amsterdam, 1993, www.cwi.nl/~lex/sdo.pdf.
- 10 É. Tardos, A strongly polynomial minimum cost circulation algorithm, *Combinatorica* 5 (1985), 247–255.
- 11 E. Tulp, *Searching timetable networks*, proefschrift, Vrije Universiteit Amsterdam, 1991.
- 12 W. van der Aalst, K.M. van Hee en M. Voorhoeve, The DONS rail scheduling system, *ATPN Case Studies Tutorial*, 1994, pp. 1–12.
- 13 J.W.H.M.T.S.J. van Rees, Een studie omtrent de circulatie van materieel, *Spoor- en Tramwegen* 38 (1965), 363–367.