

STAtOR

periodiek van de VvS+OR jaargang 14, nummer 1, februari 2013

Programma van de Dag voor Statistiek en OR 2013
'DATA VISUALIZATION'

Netwerken in de psychologie: nieuwe kennis

Eerst weekend!

Wiskunde in dienst van een sociaal leven

Gewetensconflicten van een
sociaal-wetenschappelijke onderzoeksmethodoloog

De derde Internationale Timetabling Competitie;
op zoek naar het beste schoolrooster

Revenue Management:
kansen voor de luchtvrachtindustrie

Van de President

Het is 2013, het Internationale Jaar van de Statistiek <www.statistics2013.org>. Er zijn heel wat aankondigingen en aanmoedigingen aan vooraf gegaan, maar nu is het dan zover. Ik weet niet hoeveel statistici in Nederland zich bewust zijn van het feit dat dit 'ons jaar' is. En velen denken wellicht, wat hebben we eraan (toch meestal de eerste gedachte van een Nederlander). Ik zou echter de bekende tegenvraag willen stellen: wat kunnen wij eraan bijdragen als Nederlandse statistici? Het klinkt aantrekkelijk om de Nederlandse statistiek op de wereldkaart te zetten. Maar kijk dan eens naar de volgende kaart <www.statistics2013.org/files/2012/12/STAT2013Poster.pdf>. Uit een groot aantal landen worden er verworvenheden en prestaties aan de statistiek toegeschreven. Helaas staat er niets op uit Nederland. Dat kunnen we toch niet op ons laten zitten. Daarom roep ik iedereen op om een mooie Nederlandse bijdrage aan deze kaart te leveren en deze op te sturen naar <jmeulman@math.leidenuniv.nl>. Vervolgens stuur ik (een selectie van) de bijdragen door naar de organisatie.

Ik ben me er als President van de VvS+OR zeer van bewust dat het bovenstaande alleen over de statistiektak van onze vereniging gaat, en niet over de OR, de andere tak. Helaas is het zo dat, behalve in Nederland, nergens in de rest van de wereld Statistiek en OR in een adem worden genoemd, ze samen een vereniging vormen, en dat ze dus samen zich in zouden zetten voor 'Het Jaar van...'. Is dit niet iets om over na te denken? Het bestuur van de VvS+OR dacht van wel. In december 2012 zijn we begonnen met het opstellen van een Strategisch Plan voor 2013-2017, en een van de vragen die daarbij

onmiddellijk naar voren kwam was: kan dit wel voor de VvS en OR tegelijk? Zijn de missie en de doelstelling niet te verschillend? Zijn de sterke en zwakke punten, de uitdagingen en de bedreigingen van de statistiek (jawel, ik geef het toe: we hebben onder leiding van Nick Fisher, een van de *keynote speakers* van de 2012 Annual Meeting, een SWOT analyse gedaan) niet heel anders dan die van de OR? Zoals ik in de eerste STATOR van 2012 schreef: in een aantal toepassingsgebieden wordt vaak de term *Statistics* niet meer gebruikt, maar spreekt men van *Business Analytics*. Ik denk dat dit vooral de OR-leden aanspreekt, maar zouden de de Statistiek-leden dit ook willen? Het jaar 2013 is voor de leden van de VvS+OR bij uitstek het jaar om na te denken over de gezamenlijke, maar misschien beter, afzonderlijke identiteit.

Het bestuur komt hier op de Algemene Ledenvergadering op 21 maart op terug.

En als het over 21 maart 2013 gaat, kan het vanaf dit punt alleen maar gaan over het inhoudelijke deel van onze Annual Meeting. Het onderwerp is *Data Visualization* en het bestuur hoopt dat dit grootscheeps enthousiasme op zal roepen. Bij veel leden, en natuurlijk ook bij veel niet-leden, zodat men massaal naar de Annual Meeting komt (om daarbij eventueel lid te worden). *Data Visualization* is al decennia lang een *hot topic*, en het intrigerende is dat het dat ook blijft! Getallen blijven getallen, maar de representatie van data blijft evolueren. Ook is er veel *at stake*: in de publieke opinie lijkt het wel of de reputatie van de statistiek soms afhangt van de visualisatie van de data. En werden in de kranten tot

nu toe meestal relatief eenvoudige grafieken gebruikt, in de discussie over de inkomensafhankelijke zorgbijdrage in 2012 kopte een artikel in de NCR met 'Wie naar de wolken kijkt, ziet extremen'. 'Wolken' stond voor onze 'puntenwolken' en 'extremen' voor onze 'uitbijters'. Wat willen we nog meer? Dat er niet geklungeld wordt met die data representaties natuurlijk en dat gebeurt helaas te vaak. Op onze *Data Visualization* dag hebben we tot onze grote vreugde kanonnen bereid gevonden om te spreken.

Andy Kirk heeft het aangedurfd om buiten de academische wereld om een naam in de visualisatie te worden. Afgestudeerd in de OR, 'ontdekte' hij het topic visualisatie redelijk laat, maar stortte hij zich er vervolgens voluit op. Zijn activiteiten zijn talrijk, waaronder freelance consultancy werk voor onder meer Microsoft, Disney, Intel en McKinsey. Onlang is zijn eerste boek *Data Visualization: a successful design process* verschenen. Je kunt Andy zien op <http://vimeo.com/44886980> met *The 8 Hats of Data Visualisation Design*.

Howard Wainer is zeer vooraanstaand in de psychometrie, maar we hebben hem hier als een van de meest succesvolle en productieve auteurs, zowel wetenschappelijk als populair-wetenschappelijk, over visualisatie. Het volgende citaat geeft een indruk van zijn laatste boek, *Picturing the Uncertain World*, over dit onderwerp: *In his book Picturing the Uncertain World, Howard Wainer explores how graphs can serve as maps to guide us when the information we have is ambiguous or incomplete. Using a visually diverse sampling of graphical display, from heartrending autobiographical displays of genocide in the*

Kovno ghetto to the 'Pie Chart of Mystery' in a New Yorker cartoon, Wainer illustrates the many ways graphs can be used – and misused – as we try to make sense of an uncertain world.

Lee Wilkinson is uitzonderlijk. Hij is de grondlegger van een van de beste pakketten voor statistische data-analyse (SYSTAT), en ook de ultieme expert op het gebied van de grafische representatie van resultaten van statistische analyses. Over zijn monumentale werk *The Grammar of Graphics* wordt geschreven: *'The Grammar of Graphics' presents a unique foundation for producing almost every quantitative graphic found in scientific journals, newspapers, statistical packages, and data visualization systems. While the tangible results of this work have been several visualization software libraries, this book focuses on the deep structures involved in producing quantitative graphics from data. What are the rules that underlie the production of pie charts, bar charts, scatter plots, function plots, maps, mosaics, and radar charts? Those less interested in the theoretical and mathematical foundations can still get a sense of the richness and structure of the system by examining the numerous and often unique color graphics it can produce.'*

Tot slot kan alleen nog toegevoegd worden dat Wilkinson, Wainer, en Kirk excellente sprekers zijn. We zullen gefascineerd worden, uiteraard door hun beelden, maar vooral ook door hun woorden. Voor de VvS+OR wordt dit een hoogtepunt in het Jaar van de Statistiek. *Welcome to the International Year of Statistics.*

JACQUELINE MEULMAN

Thursday, March 21th 2013

Annual meeting of the Netherlands Society of Statistics and Operations Research (VvS+OR)

DATA VISUALIZATION

LOCATION

Jaarbeurs Utrecht, Room 215,
Jaarbeursplein 6, 3521 AL Utrecht (near Central Station)

KEYNOTE SPEAKERS

Leland Wilkinson
Howard Wainer
Andy Kirk

PROGRAM

- 09.30 - 10.00 Registration and coffee/tea
- 10.00 - 10.15 Opening and introduction first speaker
- 10.15 - 11.15 Leland Wilkinson
- 11.15 - 11.30 Coffee | tea
- 11.30 - 12.30 Annual General Meeting (ALV in Dutch)
- 12.30 - 13.30 Lunch break
- 13.30 - 13.45 Introduction next speaker
- 13.45 - 14.45 Howard Wainer
- 14.45 - 15.15 VvS+OR Master's Thesis Award presentation
- 15.15 - 15.30 Coffee | tea
- 15.30 - 16.00 VvS+OR Ph.D. Thesis Award presentation
- 16.00 - 16.15 Introduction next speaker
- 16.15 - 17.15 Andy Kirk
- 17.15 Drinks



The annual meeting of the Netherlands Society of Statistics and Operations Research is the place to meet other people with an interest in statistics and operations research and people working in this field.

www.vvs-or.nl

REGISTRATION (BEFORE MARCH 14th)

Members VvS+OR

The annual meeting is free for members but registration through the society's website is required: <www.vvs-or.nl>. Register before March 14th and bring the confirmation e-mail and show it at the registration desk at the entrance of the lecture room.

Non-members

Non-members can either become a member (go to www.vvs-or.nl and click 'Become a Member' in the top menu) and then immediately also have free access after registration (see above), or pay 50 euro (bank account: 202091, VvS+OR, 'registration annual meeting') before March 14th and bring proof of payment (copy of bank statement) to show at the registration desk at the entrance of the lecture room.

GENERAL INFORMATION

- The annual meeting will be English spoken, except for the Annual General Meeting (ALV).
- The Annual General Meeting (ALV) is scheduled at the end of the morning. The relevant documents will be provided on the website two weeks before the meeting. You can also get them by e-mail if you send a request to <info@vvs-or.nl>.
- Coffee/tea during the breaks and drinks afterwards are offered by the Society.
- Lunch is not included. The restaurant of the Jaarbeurs may be busy, but at a walking distance of just a few minutes (e.g. in direction Central Station) you will find several alternatives.

ORGANIZING COMMITTEE

The annual meeting is organized by the board of the VvS+OR. For questions, contact the secretary Irene Klugkist by e-mail at <info@vvs-or.nl>.

About Andy Kirk

Andy Kirk is a UK-based freelance data visualisation architect, consultant, training provider, author and editor of the website visualisingdata.com, a popular data visualisation blog.

After graduating from Lancaster University in 1999 with a BSc (hons) in Operational Research, he held a number of business analysis and information management positions at some of the largest organisations in the UK.

Late 2006 provided Andy with his career-changing 'eureka' moment when he discovered the subject of data visualisation. He has subsequently pursued a passion in this subject, completing a research Masters MA (With Distinction) at the University of Leeds in 2009.

In February 2010 he launched visualisingdata.com with the mission of providing readers with inspiring insights into the contemporary techniques, resources, applications and best practices in this exciting subject. His consultancy work and training courses extend this ambition, helping organisations of all shapes, sizes and domains master the analysis and communication of their data. Some of his most recent clients include Microsoft, Disney, Intel and McKinsey.

In January 2013 Andy published his first book entitled *Data Visualization: a successful design process*, a handy strategy guide for anyone facing up to the challenges of designing a data visualization.

DATA VISUALISATION AND THE SCIENCE OF BETTER

Andy Kirk

www.visualisingdata.com

ABSTRACT

Exponential advances in technology have provided us with ubiquitous means for creating, recording and mobilizing incredible amounts of data. Where once data was scarce, now it is captured in abundance.

Yet, whilst we have access to fantastic tools and widespread capabilities for storage, handling and analysis, the design sensibility required to most effectively display and communicate data is still lacking. Approaches based on instinct and personal taste are no longer sufficient to cope with, let alone exploit, this data overload.

This is where art and science of Data Visualisation comes in. It is an art concerned with unleashing creativity and innovation, designing communications that appeal on an aesthetic level and survive in the mind on an emotional one. It is a science aimed at understanding and exploiting the way our eyes and brains process information most efficiently, effectively and accurately.

In this talk Andy will be assessing the multi-faceted relationship between Operational Research and Data Visualisation. We will look at how visualisation has an integral role in the classic OR pursuits of optimisation and supporting better decisions. On the other side of this relationship, we will go behind the scenes, looking through the lens of Operational Research to see how methodological approaches to visualization design can help achieve maximum effect.

About Howard Wainer

Howard Wainer received his PhD from Princeton University in 1968. After serving on the faculty of the University of Chicago, a period at the Bureau of Social Science Research during the Carter Administration, and 21 years as Principal Research Scientist in the Research Statistics Group at Educational Testing Service, he is now Distinguished Research Scientist at the National Board of Medical Examiners and Professor (adjunct) of Statistics at the Wharton School of the University of Pennsylvania.

Howard Wainer was elected a Fellow in the American Statistical Association in 1985 and a Fellow of the American Educational Research Association in 2009. He was awarded the Educational Testing Service's Senior Scientist Award in 1990 and selected for the Lady Davis Prize and was named the Schonbrun Visiting Professor at the Hebrew University in 1992. He received the 2006 National Council on Measurement in Education Award for Scientific Contribution to a Field of Educational Measurement for his development of Testlet Response Theory and given NCME's Career Achievement Award in 2007, and he received the Samuel J. Messick Award for Distinguished Scientific Contributions Award from Division 5 of the American Psychological Association in 2009.

Since 1974 when he published his first article on statistical graphics, an empirical verification of the efficacy of the suspended rootogram, Howard Wainer has been a tireless advocate for the efficacy of graphics for communicating quantitative phenomena. He is one of the principals responsible for the renewed importance of graphics in statistics, and he authored a number of books on graphical methods. Wainer's approach to the study of graphics has always shown a deep respect for the work of those who had preceded him. He also has, since 1990, written the popular column 'Visual Revelations' for *Chance* magazine.

PICTURES AT AN EXHIBITION SIXTEEN VISUAL CONVERSATIONS ABOUT ONE THING

Howard Wainer

University of Pennsylvania

ABSTRACT

In 1951 the famous graphic designer Will Burtin presented a graphic showing the efficacy of three antibiotics in treating 16 bacteria. In this talk I will examine Burtin's solution as well as 15 others. We will find that there are many paths to salvation, but that had the data been displayed differently, important discoveries could have been accelerated by decades.

Complete details are in Chapters 2-3 of: Wainer, H. (2013) *Medical Illuminations*. London: Oxford.

3

About Leland Wilkinson

Leland Wilkinson has an AB from Harvard University, an STB degree from Harvard Divinity School, and a PhD from Yale. He is currently adjunct professor of computer science at the University of Illinois at Chicago. Previously, he was adjunct professor of statistics at Northwestern University and President of SYSTAT Inc., a statistical software company he founded in 1984. Lee Wilkinson now serves as Vice President of Data Visualization at the Machine Learning Company Skytree, Inc.

Wilkinson is a fellow of the American Statistical Association, an elected member of the International Statistical Institute, and a fellow of the American Association for the Advancement of Science. He recently served on the NAS Panel on Developing Science, Technology, and Innovation Indicators for the Future. His projects have included books, journal articles, the original SYSTAT statistics package, numerous open-source software packages, and patents in visualization and distributed computing.

Lee Wilkinson wrote SYSTAT, a statistical software package, in the early 1980's. This program was noted for its comprehensive graphics, including the first software implementation of the heatmap display now widely used among biologists. After his company grew to 50 employees, he sold it to SPSS in 1995. At SPSS, he assembled a team of graphics programmers who developed the nViZn platform that produces the visualizations in SPSS, Clementine, and other analytics products. The nViZn platform was modelled after Wilkinson's book on statistical graphics, *The Grammar of Graphics*. This book also influenced the development of the R package ggplot2 and the Polaris project at Stanford. Recently Lee Wilkinson also founded Advise Analytics, a Chicago-based scientific software company, that has launched a new system for statistics called AdviseStat, an intelligent analytics advisor.

HIGH-DIMENSIONAL VISUAL ANALYTICS: EXPLORING STRUCTURE USING LOW-DIMENSIONAL PROJECTIONS

Leland Wilkinson

University of Illinois

ABSTRACT

This talk covers a research program that rests on algorithms for computing geometric graphs on low-dimensional projections. These algorithms have led to: 1. discovering canonical structures and anomalies in collections of 2D scatter plots, 2. discovering anomalies in high-dimensional point sets, 3. a random-projection ensemble classifier that outperforms decision trees, random forests, and support vector machines on widely-used test datasets, 4. a system for analyzing visually massive multivariate time series, 5. a system for clustering images, and 6. an expert system for statistical analytics that discovers and remedies statistical irregularities. Software from these projects will be briefly presented.

STAtOR is een uitgave van de Vereniging voor Statistiek en Operationele Research (VvS+OR). STAtOR wil leden, bedrijven en overige geïnteresseerden op de hoogte houden van ontwikkelingen en nieuws over toepassingen van statistiek en operationele research. Verschijnt 3 keer per jaar.

Redactie

Joaquim Gromicho (hoofdredacteur), Ana Isabel Barros, Johan van Leeuwen, Mirjam Moerbeek, Gerrit Stemerink (eindredacteur), Hilde Tobi. Vaste medewerkers: Fred Steutel, Henk Tijms

Kopij en reacties richten aan

Prof. dr. J.A.S. Gromicho (hoofdredacteur), Faculteit der Economische Wetenschappen en Bedrijfskunde, afdeling Econometrie, Vrije Universiteit, De Boelelaan 1105, 1081 HV Amsterdam, telefoon 020-5986010, mobiel 06-55886747, <j.a.dossantos.gromicho@vu.nl>.

Bestuur van de VvS+OR

Voorzitter: prof. dr. Jacqueline Meulman <president@vvs-or.nl>
Secretaris: dr. Irene Klugkist <bestuur@vvs-or.nl>
Penningmeester: dr. Ad Ridder <penningmeester@vvs-or.nl>
Studentlid: Maarten Kampert (BSc) <student@vvs-or.nl>
Overige bestuursleden: prof. dr. Fred van Eeuwijk (BMS), prof. dr. ir. Stan van Hoesel & dr. John Poppelaars (NGB), dr. Eric Cator (SMS), dr. Michel van de Velden (ECS), dr. Andries van der Ark (SWS).

Leden- en abonnementenadministratie van de VvS+OR

VVS, Postbus 244, 6700 AE Wageningen, telefoon 0317 - 419572, fax 0317 - 421364, <admin@vvs-or.nl>.
Raadpleeg onze website over hoe u lid kunt worden van de VVS of een abonnement kunt nemen op STAtOR of op een van de andere periodieken.

VvS+OR-website

www.vvs-or.nl

Sociale media

Wilt u uw vakgenoten ontmoeten en wilt u discussiëren over actuele thema's, volg dan de VvS+OR en de Young Statisticians via LinkedIn, Facebook, Twitter en Flickr.

Sluit je aan bij de LinkedIn-groep van VvS+OR of Young Statisticians; bekijk foto's op <www.flickr.com/photos/vvs-or/sets/>; Like onze Facebook-pagina; volg de President van VvS+OR op <<https://twitter.com/#!/dutchstat>>.

Advertentieacquisitie

Nikki Bisschop & Joren Brunekreef, Lange Nieuwstraat 6, 3512 PH Utrecht, 06-55874175, <adverteren.stator@vvs-or.nl>.
STAtOR verschijnt in maart, juni en oktober.

Ontwerp en opmaak

Pharos | M. van Hoetegem, Nijmegen

Uitgever

© Vereniging voor Statistiek en Operationele Research
ISSN 1567-3383

- 2 Van de president
- 4 Programma van de Dag voor Statistiek en OR 2013 'Data Visualization'
- 9 Redactioneel
- 10 Netwerken in de psychologie: nieuwe kennis
Lourens Waldorp & Denny Borsboom
- 15 Erestatistiek – column
Fred Steutel
- 16 Eerst weekend! Wiskunde in dienst van een sociaal leven
Egbert van der Veen & Erwin Hans
- 20 Gewetensconflicten van een sociaal-wetenschappelijke onderzoeksmethodoloog
Jarl Kampen
- 22 De derde Internationale Timetabling Competitie; op zoek naar het beste schoolrooster
Jeffrey H. Kingston & Gerhard Post
- 25 Young Statisticians
- 26 Revenue Management: kansen voor de luchtvrachtindustrie
Thijs Boonekamp
- 30 Facebook en moorden in New York – column
Henk Tijms
- 31 Het Peterprincipe – column
Johan van Leeuwen

NETWERKEN IN DE PSYCHOLOGIE: NIEUWE KENNIS

Netwerkanalyse vormt een methodologie voor het bestuderen van complexe systemen. Aan de Universiteit van Amsterdam ontwikkelen Lourens Waldorp en Denny Borsboom technieken om oude en nieuwe problemen in de psychologie te lijf te gaan. Een aantal daarvan lichten zij hieronder toe.

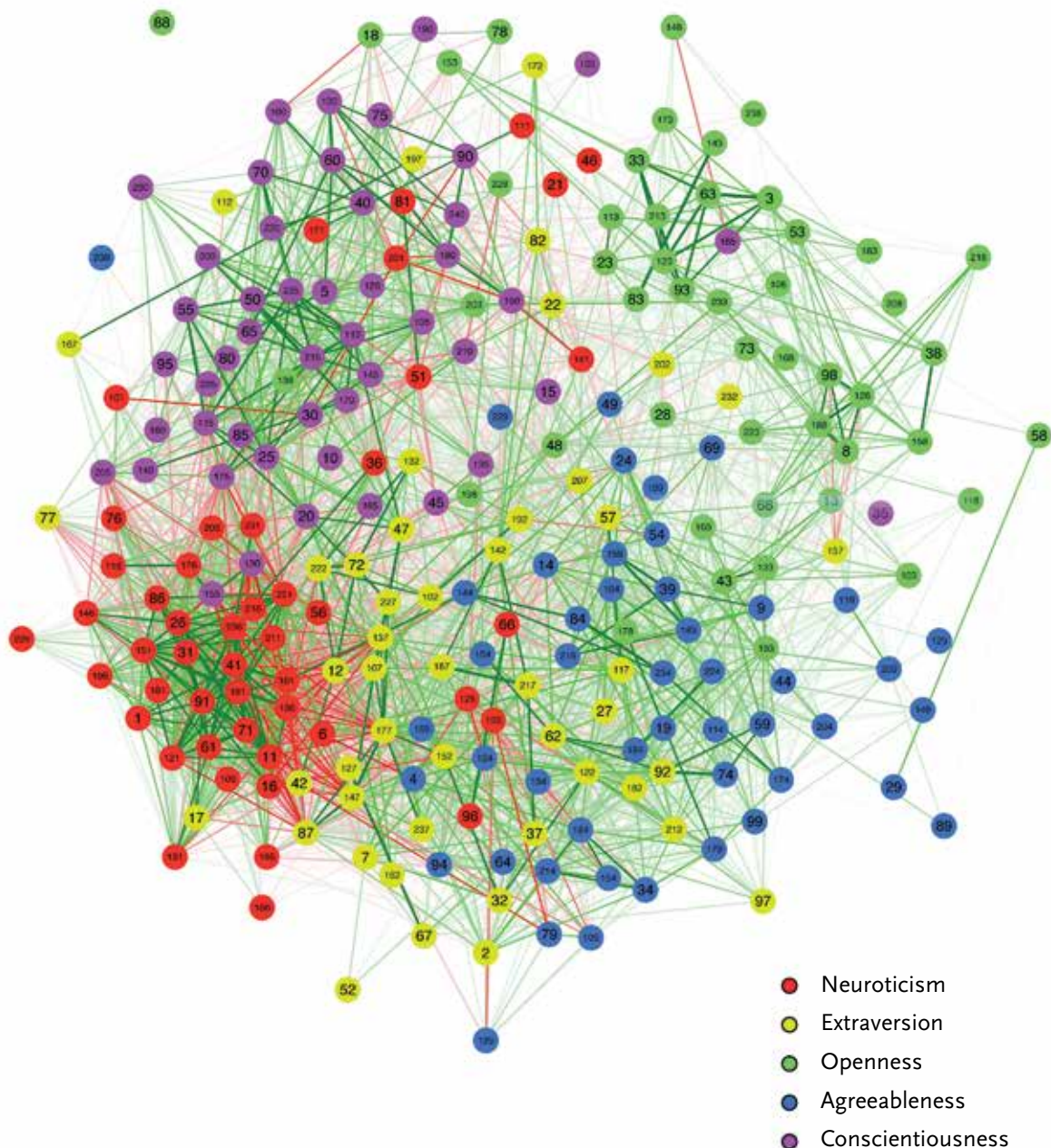
LOURENS WALDORP & DENNY BORSBOOM

Data-visualisatie

Het visualiseren van complexe patronen vormt een belangrijk onderdeel van statistische analyse. In de literatuur over netwerken is het gebruikelijk om de elementen waaruit een netwerk bestaat als knopen weer te geven, en relaties tussen die elementen als verbindingen tussen

knopen. Zo'n representatie maakt de structuur van het netwerk in één klap duidelijk.

De meeste mensen kennen plaatjes van netwerken vooral uit de sociale netwerk-literatuur, waarbij knopen staan voor individuen, en verbindingen bijvoorbeeld weergeven wie met wie bevriend is. Netwerkrepresentaties kunnen echter ook heel handig zijn bij het in kaart

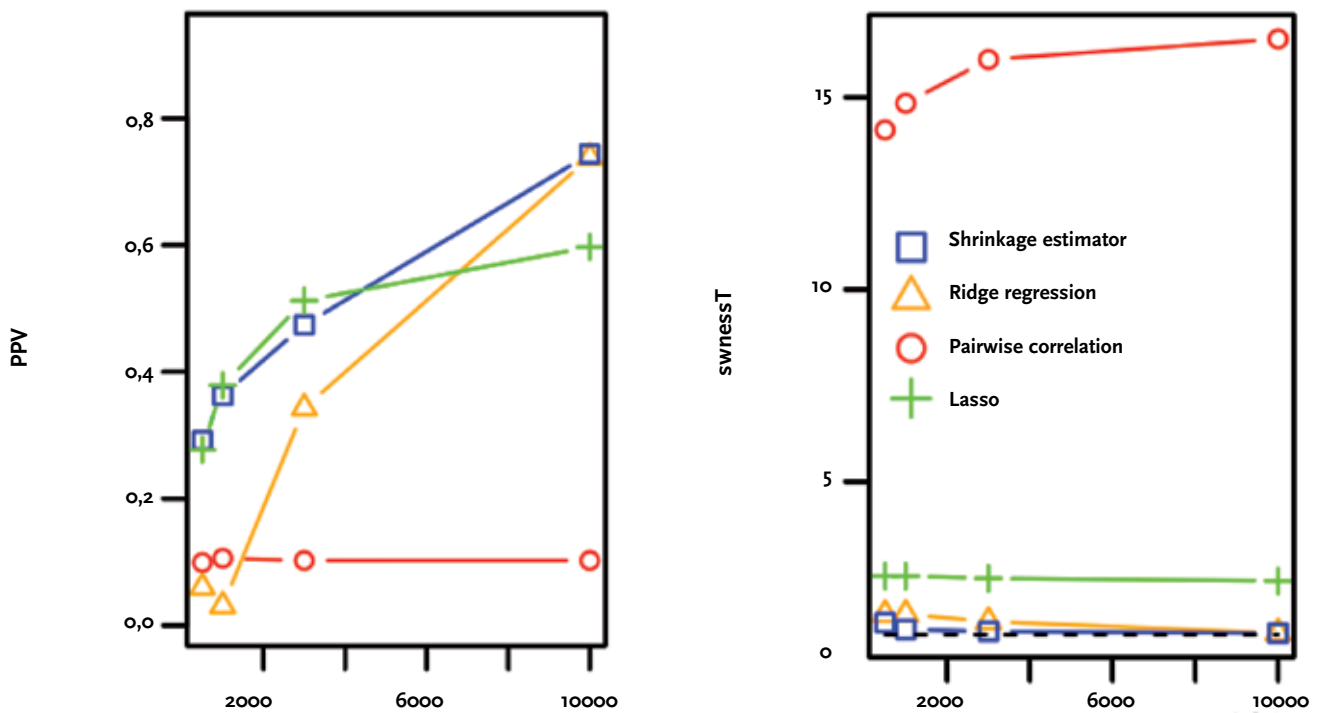


Figuur 1. Een netwerk van de 240 persoonlijkheidsitems uit de NEO-PI, zoals afgenomen bij 500 eerstejaars psychologiestudenten. Ieder item wordt weergegeven als een knoop, en (hogere) correlaties als (dikkere) verbindingen tussen knopen; negatieve correlaties zijn gerepresenteerd als rode verbindingen, en positieve correlaties als groene verbindingen.

brenge van relaties tussen grote aantallen variabelen. Daarbij worden niet individuen, maar variabelen als knoop weergegeven. Doorgaans wordt de verbinding tussen knopen gewogen met een associatiemaat, zoals bijvoorbeeld de (partiële) correlatie tussen variabelen. Zo toont figuur 1 een netwerkrepresentatie van de correlatiestructuur van 240 persoonlijkheidsitems uit

de NEO-PI, een belangrijke persoonlijkheidstest. Ons vrij beschikbare R-package qgraph (Epskamp, Cramer, Schmittmann, Waldorp & Borsboom, 2011) is speciaal geoptimaliseerd om dit type correlatieve structuren weer te geven.

De weergave van correlaties in een gewogen netwerk levert een visuele aanvulling op die van principi-



Figuur 2a & figuur 2b. De *positive predictive value* (PPV) is de verhouding van werkelijke verbindingen en significante (partële) correlaties (a) als functie van het aantal observaties. En de *small-worldness* index (b) als functie van het aantal observaties.

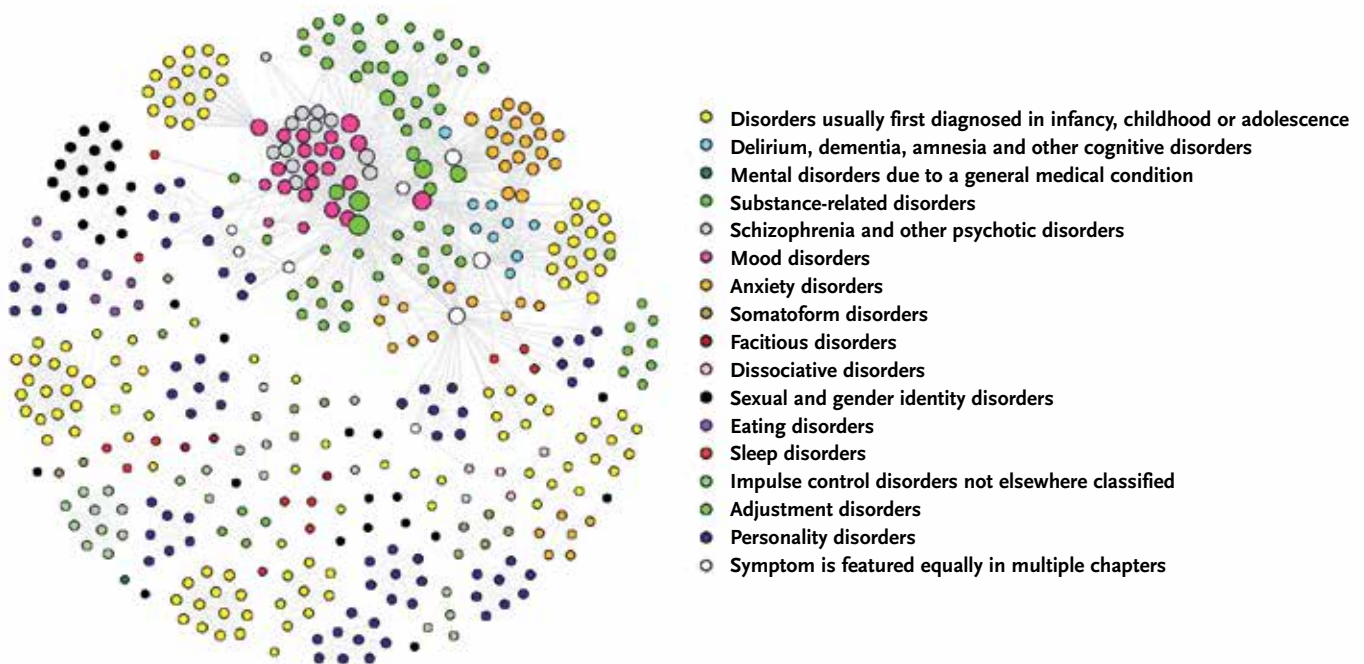
pale componenten analyse en gerelateerde technieken, maar het leidt ook tot een andere benadering van de gegevens. De representatie in figuur 1 correspondeert namelijk met een wiskundige structuur – een graaf – waaraan gerekend kan worden. Men kan bijvoorbeeld onderzoeken hoe groot de afstand tussen twee knopen is (padlengte), of wat de centraliteit van een knoop is. Een belangrijke maat voor centraliteit is *betweenness*. In een ongewogen graaf, waar afstanden berekend worden door het aantal verbindingen tussen knopen te tellen, geeft een hogere *betweenness* aan dat een knoop relatief vaak op het kortste pad tussen twee willekeurig gekozen andere knopen ligt. Voor gewogen grafen, met afstanden gewogen door bijvoorbeeld de (partiële) correlatie, zijn soortgelijke maten recent geïntroduceerd door Opsahl, Agneessens & Skoveretz (2010). Voor het netwerk in figuur 1 is het meest centrale item bijvoorbeeld ‘Ik ben een vrolijk en levendig iemand’, wat inhoudelijk ook een zeer centraal persoonlijkheidskenmerk is. Dit type analyse levert een interessante manier om de structuur van een correlatiematrix in kaart te brengen, en complementeert bekende psychometrische technieken (Schmittmann, Cramer, Waldorp, Eps-

kamp & Borsboom, 2011). Bovendien zijn de resultaten gemakkelijk te communiceren.

Het brein

De afgelopen 25 jaar is er veel geleerd over de onderliggende structuren in het brein die verantwoordelijk zouden zijn voor ons gedrag. Zo zijn er veel voorbeelden waarbij de lokalisatie van de verwerking van stimuli nauwkeurig in beeld kan worden gebracht. Maar over de vraag hoe informatie moet worden geïntegreerd om tot coherent gedrag te komen is weinig bekend. Die vraag kan worden beantwoord met netwerken.

Bij fMRI worden de hersenen ingedeeld in zo'n 200.000 tot 2 miljoen kleine kubusjes, *voxels* genaamd (combinatie van *volume* en *elements*). Om een netwerk te maken van deze voxels wordt meestal gebruik gemaakt van paarsgewijze correlaties tussen al die voxels. Laten we zeggen dat er 2.000 voxels mogelijk interessant zijn. Dat betekent bijna 2 miljoen correlaties (19.990.00 combinaties) tussen de gemeten fMRI tijdseries. Een hoge correlatie tussen twee voxels wordt geïnterpreteerd als



Figuur 3. Het DSM-IV netwerk met symptomen weergegeven als knopen die verbonden worden indien ze bij dezelfde stoornis horen volgens de DSM-IV. Iedere kleur van een knoop correspondeert met een stoornis.

evidentie voor het bestaan van een functionele relatie. Vaak wordt een drempelwaarde gebruikt voor alle correlaties, bijvoorbeeld alle absolute correlaties van tenminste 0,45, zodat het netwerk overzichtelijk wordt.

Ook hier kunnen padlengtes worden bepaald tussen gebieden in het brein. Daarbij blijkt dat de gemiddelde padlengte in het brein samenhangt met IQ: hoe hoger de gemiddelde padlengte hoe lager het IQ. De gemiddelde padlengte kan gezien worden als een maat voor efficiëntie: hoe korter de gemiddelde padlengte, hoe sneller het contact tussen neuronale systemen, en dus hoe efficiënter het brein. En daarmee betekent het resultaat dus dat efficiënte breinen slimmer zijn dan minder efficiënte breinen.

Deze resultaten staan of vallen echter met het construeren van het netwerk. Wij hebben bekeken of het berekenen van paarsgewijze correlaties een statistisch valide netwerk oplevert, in die zin dat verbindingen in een correlatieel netwerk overeenkomen met bestaande verbindingen. Daartoe hebben we verschillende netwerken gesimuleerd. Het blijkt dat paarsgewijze correlaties vaak verbindingen opleveren die er niet zijn. In figuur

2a staat de *positive predictive value* oftewel de verhouding tussen de werkelijke verbindingen en de significante verbindingen, als functie van het aantal observaties. De rode lijn representeert het resultaat van een correlatie netwerk. Het is te zien dat er weinig correcte verbindingen gevonden worden met paarsgewijze correlaties. De schatting voor de efficiëntie van een netwerk klopt daarmee ook niet helemaal. Dat is te zien in figuur 2b waarin de *small-worldness* wordt afgebeeld. De *small-worldness* refereert aan een netwerk met hoge locale clustering en korte padlengtes. De rode lijn representeert de schatting van de *small-worldness* met paarsgewijze correlaties en er is te zien dat deze de werkelijke waarde (gestreepte lijn onderin) sterk overschat.

Betere methodes maken gebruik van partiële correlaties. Hiervoor is het nodig extra informatie mee te geven in de vorm van *penalties* (zoals geïmplementeerd in *lasso*, *minimum norm*, en *shrinkage* schatters). Dit zijn de andere lijnen in figuur 2b. Ze overschatten de *small-worldness* niet of nauwelijks. Ook is te zien in figuur 2a dat wanneer een partiële correlatie significant is deze ook vaak een werkelijke verbinding is. Opvallend is dat

de paarsgewijze correlaties een beetje slechter worden met het aantal observaties. Dit is te verklaren doordat de hoge clustering (veel lokale verbindingen) bepaald wordt door het aantal driehoeksconfiguraties in een netwerk (transitiviteit). Een hoge (absolute) correlatie tussen twee variabelen kan makkelijk voorkomen doordat er een derde variabele is die de correlatie veroorzaakt.

Psychopathologie

In de klinische psychologie gebeurt het regelmatig dat twee of meer stoornissen bij dezelfde persoon voorkomen. Dit wordt comorbiditeit genoemd. In 2010 introduceerden wij de hypothese dat deze comorbiditeit het gevolg is van 'overspraak' tussen netwerken van symptomen (Cramer, van der Maas, Waldorp & Borsboom, 2010). Bijvoorbeeld angst voor verschillende gebeurtenissen (een symptoom van gegeneraliseerde angststoornis) kan leiden tot slaapgebrek (een gedeeld symptoom van gegeneraliseerde angststoornis en depressie) wat kan leiden tot sloomheid (een symptoom van depressie). Hoeveel van zulke paden tussen stoornissen bestaan er, en welk gedeelte van de empirisch gevonden comorbiditeit kan eraan worden gerelateerd? Dit hebben we onderzocht aan de hand van de *Diagnostic and Statistical Manual of Mental Disorders-IV* (DSM-IV), waarin psychopathologische stoornissen gedefinieerd worden aan de hand van symptomen (Borsboom, Cramer, Schmittmann, Epskamp & Waldorp, 2011).

In figuur 3 is elk symptoom als een knoop weergegeven, en zijn twee symptomen aan elkaar verbonden als ze bij dezelfde DSM-IV stoornis voorkomen. Het valt meteen op dat er een dominant cluster van symptomen is die direct of indirect met elkaar verbonden zijn. Zo'n cluster wordt een *giant component* genoemd. In dit geval omvat deze component bijna 50% van alle symptomen. De helft van de symptomen is dus direct of indirect met elkaar verbonden, ook al horen de meeste symptomen maar bij één stoornis. 'Prikkelbaar', 'afleidbaar', 'angstig', en 'neerslachtig' zijn de meest centrale symptomen in het netwerk. Verder blijkt dat de *giant component* een *small world* is. Dat wil zeggen dat het netwerk gekenmerkt wordt door een combinatie van hoge clustering en korte padlengtes, zodat het gemakkelijk is om van een willekeurig symptoom naar een willekeurig ander symptoom te komen.

Het netwerk verklaart ook de empirisch gevonden

comorbiditeit redelijk goed. Het blijkt dat de (tetrachorische) correlaties tussen stoornissen goed worden voorspeld door de gemiddelde padlengte tussen de stoornissen (het gemiddeld aantal stappen over de verbindingen tussen de corresponderende symptomen) in de *giant component*. Dat betekent dat comorbiditeit het inderdaad het gevolg zou kunnen zijn van overspraak tussen stoornissen via gedeelde symptomen.

Conclusie

Symptomen, persoonlijkheidskenmerken, en functioneel verschillende hersengebieden vormen netwerken. De structuur van zulke netwerken kan inzichtelijk gemaakt worden met netwerkvisualisaties en -analyses. De onderdelen van een netwerk kunnen bovendien aanleiding geven tot verrassende eigenschappen, die voortkomen uit de organisatie van het netwerk. Misschien is het geheel soms tóch meer dan de som der delen.

LITERATUUR

- Borsboom, D., Cramer, A. O. J., Schmittmann, V. D., Epskamp, S. & Waldorp, L. J. (2011). The small world of psychopathology. *PLoS ONE*, 11, e27407.
- Cramer, A. O. J., Waldorp, L. J., van der Maas, H. & Borsboom, D. (2010). Comorbidity: A network perspective. *Behavioral and Brain Sciences*, 33, 137–150.
- Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D. & Borsboom, D. (2011). *Qgraph: Network representations of relationships in data. R package version 0.4.10*. Available from <http://CRAN.R-project.org/package=qgraph>.
- Schmittmann, V. D., Cramer, A. O. J., Waldorp, L. J., Epskamp, S., Kievit, R. A. & Borsboom, D. (2011). Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*. doi: 10.1016/j.newideapsych.2011.02.007
- Opsahl, T., Agneessens, F. & Skoveretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32, 245–251.

LOURENS WALDORP is universitair docent bij de Programmagroep Psychologische Methodenleer van de Faculteit der Maatschappij- en Gedragwetenschappen van de Universiteit van Amsterdam.
E-mail: <L.J.Waldorp@uva.nl>

DENNY BORSBOOM is hoogleraar bij de Programmagroep Psychologische Methodenleer van de Faculteit der Maatschappij- en Gedragwetenschappen van de Universiteit van Amsterdam.
E-mail: <D.Borsboom@uva.nl>

ERESTATISTIEK

Er worden veel gegevens ('statistieken') verzameld over voetbal. Toch zijn sommige antwoorden moeilijk te vinden. Als je algemene informatie zoekt over voetbaluitslagen, stuurt Google je vaak door naar de site van een club.

Ik wilde weten wat de meest voorkomende voetbaluitslag is. Volgens Maarten 't Hart, die in Leiden statistiek heeft geleerd van Willem van Zwet, is dat 2-1 (of 1-2; het kan hem niet schelen wie er wint). Deze bewering, die voorkomt in zijn bundel 'Dienstreizen van een thuisblijver', vind ik niet gestaafd of weersproken op internet.

De voetbalstatistiek (eredivisie) is ingewikkelder geworden door de nieuwe regels. Laten we eenvoudig beginnen. Er zijn 18 clubs. Die spelen allemaal twee keer tegen de andere 17. Dat zijn $18 \times 17 = 306$ wedstrijden: clubs spelen weliswaar twee keer tegen elkaar: uit en thuis, maar bij iedere wedstrijd zijn twee clubs betrokken.

Vroeger waren er dan 612 punten te verdelen: elke wedstrijd leverde twee punten op: twee voor een winnaar of elk één bij een gelijkspel. Nu is het aantal punten per wedstrijd een stochastische grootheid, die de waarden twee of drie kan aannemen; gemiddeld 'dus' 2,5. Het totaal aantal behaalde punten heeft dan verwachting 765 en variantie 76,5.

Maar er zijn leukere dingen om naar te kijken. De stand in de competitie wordt weergegeven in 18 regels. In de eerste kolom staan de rangnummers, in de zesde de aantallen behaalde punten. Verder worden nog ver-

meld: het aantal gespeelde ronden en de aantallen gewonnen, gelijkgespeelde en verloren wedstrijden.

Toen ik eens naar die rijtjes keek viel me op dat, op die dag, de nummers 1 en 2 respectievelijk 18 en 17 punten hadden en de nummers 17 en 18 respectievelijk 2 en 1. Hoe het verder ging weet ik niet meer, maar ik vroeg me af of het mogelijk was dat het zo door ging: de getallen 1 t/m 18 die de rang van de clubs aangeven van hoog naar laag zijn dezelfde als de getallen die de aantallen behaalde punten aangeven van laag naar hoog.

In die situatie zijn er 171 wedstrijdpunten verdeeld. Het aantal gespeelde wedstrijden is een negenvoud, en deze 171 punten kunnen, bijvoorbeeld, behaald zijn in 45 winstpartijen en 18 gelijke spelen, dat wil zeggen in zeven ronden van negen wedstrijden. De geschetste toestand is dus niet a priori onmogelijk. Blijft de vraag of het ook echt kan; het aantal winstpartijen is natuurlijk gelijk zijn aan het aantal verliespartijen. Kan dat allemaal? Het kan, zoals hieronder blijkt. Om ruimte te sparen heb ik de kolommen horizontaal neergezet.

Rest de vraag of er bij dit resultaat ook werkelijk uitslagen passen van zeven weken eredivisievoetbal. Of is dat een domme vraag?

FRED STEUTEL is emeritus hoogleraar kansrekening aan de TU Eindhoven.

E-mail: <f.w.steutel@tue.nl>

RANG	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
PUNTEN	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
WINST	6	5	5	4	4	4	3	2	3	1	2	2	2	0	1	1	0	0
GELIJK	0	2	1	3	2	1	3	5	1	6	2	1	0	5	1	0	2	1
VERLIES	1	0	1	0	1	2	1	0	3	0	3	4	5	2	5	6	5	6

EERST WEEKEND !

wiskunde in dienst van een sociaal leven

Het sociale leven – sportactiviteiten, verjaardagsbezoeken, het ontmoeten van vrienden – vindt vooral in de weekeinden plaats. Dit maakt werken in het weekend voor veel mensen onaantrekkelijk. In sectoren waar de dienstverlening 7 dagen in de week en 24 uur per dag beschikbaar moet zijn, zoals in de gezondheidszorg en de beveiligingssector, dienen medewerkers wel in het weekend te werken. Het maken van goede dienstroosters voor het weekend is een uitdaging: medewerkers hebben vaak heel specifieke voorkeuren terwijl de werkgever moet zorgen dat er voldoende mensen worden ingezet. Het maken van goede dienstroosters wordt verder bemoeilijkt doordat er rekening moet worden gehouden met de Arbeidstijdenwetgeving (ATW) en omdat werkgevers de diensten ‘eerlijk’ over de medewerkers willen verdelen.

EGBERT VAN DER VEEN & ERWIN HANS

Weekendroosters zijn voor medewerkers van bedrijven en instellingen heel belangrijk en bepalen in hoge mate de kwaliteit van hun rooster. Daarom beginnen roostermakers in de praktijk vaak met het maken van de weekendroosters. Dienstroosteralgoritmië in geavanceerde personeelsplanningssoftware zoals ORTEC Harmony is gericht op het maken van roosters voor de hele week (weekend- én wekdagen), hetgeen niet altijd tot goede oplossingen leidt voor de weekenden. Dit was aanleiding voor ons om te onderzoeken of we een algoritme konden ontwikkelen dat specifiek gericht is op het roosteren van weekenddiensten, en dat daarmee een belangrijke stap in het natuurlijke planningsproces ondersteunt.

Geschiedenis van dienstroosteralgoritmië

Dienstroosteralgoritmen wijzen diensten toe aan medewerkers binnen een gegeven planningshorizon van bijvoorbeeld een maand. Een dienst is een tijdsinterval waarin verschillende activiteiten gepland zijn, afgewisseld met één of meerdere pauzes. Bijvoorbeeld, een ochtenddienst van 7:00 tot 15:30 uur met een lunchpauze van 12:00 tot 12:30 uur. Bezettingseisen op diensten specificeren hoeveel medewerkers op een bepaalde dienst op een bepaalde dag ingezet moeten worden. Bijvoorbeeld, 3 medewerkers moeten op maandag op de ochtenddienst worden ingezet.

Oudere wetenschappelijke literatuur (1960 – 1980) met algoritmen voor dienstroostering beschouwt veelal uitsluitend ‘harde’ roosterbeperkingen, die voortvloeien uit de Arbeidstijdenwet (ATW) en collectieve arbeidsovereenkomsten (CAO's). Zo kan de ATW bijvoorbeeld een rust van 11 uur tussen twee opeenvolgende diensten verplichten en opleggen dat in maximaal 30 van de 52 weekenden in een jaar gewerkt mag worden. In recentere literatuur (1990 – heden) spelen ook ‘zachte’ roosterbeperkingen, zoals voorkeuren van medewerkers, een belangrijke rol. Hierbij kan gedacht worden aan een voorkeur om vrije dagen zoveel mogelijk aaneengesloten te genieten, of om geen vroege dienst na een late dienst te werken.

Ook bij weekendroosters spelen ‘zachte’ beperkingen een belangrijke rol. Zo hebben medewerkers vaak heel specifieke voorkeuren voor welke weekenden ze wel en niet willen werken, vanwege het sociale leven dat sterk geconcentreerd is rondom het weekend. Medewerkers weten hierdoor vaak ver van te voren in welk weekend ze welke dag, of welk dagdeel vrij willen zijn. Daarnaast hebben medewerkers de voorkeur hele weekenden vrij te zijn in plaats van halve weekenden te werken waarin ze alleen op zaterdag of zondag vrij zijn. Door dienstroosters die slechts deels aan deze voorkeuren voldoen ontstaat vaak een levendige ruilhandel van diensten tussen medewerkers.

In de literatuur zijn geen algoritmen beschikbaar die zich specifiek richten op het roosteren van weekenddiensten. Het opstellen van een rooster voor de hele week heeft als nadeel dat weekendvoorkeuren minder belangrijk worden. Aangezien de kwaliteit van het weekendrooster voor medewerkers in grote mate de kwaliteit van het hele rooster bepaalt, hebben wij een algoritme ontwikkeld dat uitsluitend weekenddiensten roostert. Dit heeft *minstens* twee voordelen. Dit algoritme kan beter rekening houden met medewerker-voorkeuren die betrekking hebben op het roosteren van weekenddiensten, omdat het specifiek hierop gericht is. Daarnaast ondersteunt dit het natuurlijke roosterproces waarin roostermakers eerst de weekenddiensten inroosteren en hebben manuele aanpassingen op dit rooster geen

last van diensten die reeds op doordeweekse dagen ge-roosterd zijn. Dat de gevolgen voor het rooster voor de doordeweekse dagen beperkt zijn laten we verderop zien aan de hand van experimentele resultaten.

Weekenddienstroosteralgoritme

Weekenddiensten zijn de diensten die op zaterdag en zondag gewerkt moeten worden, aangevuld met vrijdagavond- en maandagochtenddiensten. De doelstelling is om zoveel mogelijk weekenddiensten aan medewerkers toe te wijzen waarbij rekening wordt gehouden met harde en zachte beperkingen.

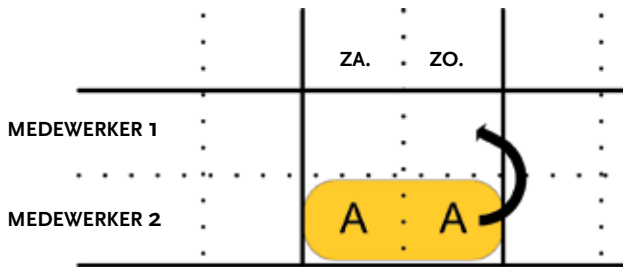
Het ontwikkelde dienstroosteralgoritme voor de weekenden bestaat uit drie stappen. Het algoritme maakt combinaties van zaterdag- en zondagdiensten (1), wijst deze combinaties aan medewerkers toe (2) en probeert deze toewijzing te verbeteren (3).

Stap 1

Voor het maken van combinaties van zaterdag- en zondagdiensten wordt een zogenaamd *minimale kosten matching* gedefinieerd. Er worden twee sets I en J van knopen gemaakt, die respectievelijk de zaterdag- en zondagdiensten vertegenwoordigen en die met lijnen met elkaar verbonden worden. Op de lijnen worden kosten gedefinieerd die bepaald worden door het aantal medewerkers dat die dienstcombinatie *niet* mag werken. Dit laatste is afhankelijk is van gedefinieerde harde beperkingen. Als een dienstcombinatie ‘kosten’ 3 heeft, betekent dit dat 3 medewerkers deze dienstcombinatie *niet* mogen werken. Als deze kosten voor alle dienstcombinaties bekend zijn wordt de minimale kostenmatching opgelost (de knopen in I en J worden aan elkaar gekoppeld, waarbij de combinatie van gebruikte lijnen minimale kosten heeft). Hiermee wordt elke zaterdagdienst gekoppeld aan een zondagdienst. Vervolgens worden deze dienstcombinaties op eenzelfde wijze uitgebreid met vrijdagavond- en maandagochtenddiensten.

MEDEWERKER	Ma	Di	Wo	Do	Vr	Za	Zo
1	V	V	L	L	L		
2			V	V	V	L	L
3	N	N				V	V
4	L	L	N	N	N		

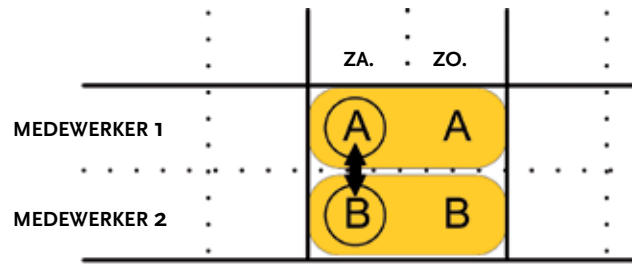
Figuur 1. Voorbeeld dienstrooster van 1 week met 4 medewerkers en 3 diensttypes: vroege (V), late (L) en nacht (N) diensten.



Figuur 2. Dienstcombinatie wordt aan andere medewerker toegewezen

Stap 2

De dienstcombinaties worden toegewezen aan medewerkers met een *hybrid ordering* heuristiek. Eerst wordt op basis van verschillende criteria een dienstcombinatie geselecteerd. Er wordt onder andere gekeken naar hoe vaak deze dienstcombinatie gepland moet worden en hoeveel medewerkers die dienstcombinatie zouden kunnen werken, afhankelijk van bijvoorbeeld kwalificaties en verlof. De dienstcombinatie die op basis van deze criteria het 'moeilijkst planbaar' is wordt aan een medewerker toegewezen. Deze medewerker wordt eveneens geselecteerd aan de hand van criteria. Deze criteria kijken onder andere gekeken naar het aantal dienstcombinaties waaraan deze medewerker is toegewezen. Nadat de dienstcombinatie aan de medewerker is toegewezen wordt bovenstaande herhaald en wordt een volgende dienstcombinatie aan een medewerker toegewezen.



Figuur 3. Diensten binnen dienstcombinatie worden verwisseld

Stap 3

Het in Stap 2 gebruikte algoritme is een *greedy* algoritme. Daarom wordt in Stap 3 het resultaat van Stap 2 verbeterd. Dit gebeurt door dienstcombinaties van de ene medewerker aan een andere medewerker toe te wijzen of door diensten binnen dienstcombinaties te verwisselen, zoals gevisualiseerd in figuur 2 en figuur 3, respectievelijk.

Experimentele set-up

Ons onderzoek is samen met een klant van logistiek consultancy en software leverancier ORTEC gedaan en heeft geresulteerd in implementatie van het hiervoor beschreven algoritme in ORTEC Harmony, de personeelsplanningssoftware van ORTEC.



Figuur 4. Percentage weekendwerk. Per instantie, per aanpak, percentages weekenden waarin medewerkers niet werken (vrij), alleen zaterdag of zondag werken (half), of het hele weekend werken (heel).

Om het effect van het algoritme op het weekend dienstrooster en op het dienstrooster als geheel te onderzoeken vergelijken we twee aanpakken: WDA+GA en GA. In de aanpak WDA+GA worden eerst de weekenddiensten ingeroosterd met het weekenddienstroosteralgoritme (WDA) en vervolgens worden de overige weekdiensten ingeroosterd met een genetisch algoritme (GA), zoals reeds beschikbaar in ORTEC Harmony (Burke, 2008; Post, 2004). In de aanpak GA worden *alle* diensten voor de hele week ingepland met het genetisch algoritme.

De beide aanpakken worden toegepast op 6 roosterinstanties afkomstig van een Belgisch politiedistrict. De roosterinstanties hebben een planningshorizon van twee maanden, 60 medewerkers en in elk rooster moeten ongeveer 2.000 diensten ingepland worden.

Resultaten

De resultaten zijn samengevat in figuur 4 en figuur 5.

Figuur 4 laat zien dat WDA+GA meer hele weekenden en minder halve weekenden roostert. Hierdoor zorgt deze aanpak voor meer vrije weekenden in de roosters. Gemiddeld bevatten de roosters die door WDA+GA worden geproduceerd 10,4% meer vrije weekenden, 21,9% minder halve weekenden en 11,4% meer hele weekenden.

Figuur 5 laat zien dat het zorgen voor meer vrije weekenden niet ten koste gaat van minder ingeplande diensten. In alle situaties plant WDA+GA meer of hetzelfde aantal weekenddiensten en in 4 van de 6 gevallen plant deze aanpak zelfs in totaal meer diensten dan GA. Gemiddeld gezien plant WDA+GA 0,1% weekenddiensten meer en ook in totaal 0,1% diensten meer dan GA. Dit komt grofweg overeen met 2 diensten per roosterinstantie.

Op roosterinstanties van een Nederlandse gehandi-

captenzorgorganisatie zien we vergelijkbare resultaten. Daar gaan we hier verder niet op in.

Conclusie

Het in dit onderzoek ontwikkelde algoritme om weekenddiensten te roosteren werkt goed binnen het genetisch algoritme van ORTEC Harmony. Op roosterinstanties uit de praktijk levert ons algoritme een beter weekend dienstrooster, zonder dat dit ten koste gaat van het hele dienstrooster. Geïnteresseerde lezers worden voor meer informatie verwezen naar een wetenschappelijk artikel over dit onderzoek.

LITERATUUR

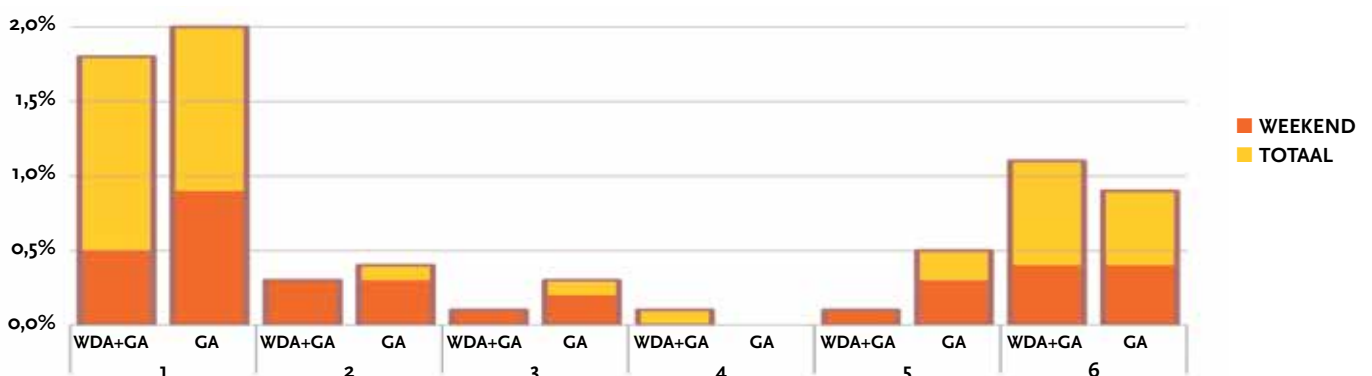
- Burke, E. K., Curtois, T., Post, G., Qu R. & Veltman B. (2008). A hybrid heuristic ordering and variable neighbourhood search for the nurse rostering problem. *European Journal of Operational Research*, 188(2), 330–341
- Post, G. & Veltman, B. (2004). Harmonious personnel scheduling. In *Proceedings of the 5th international conference on the Practice and Theory of Automated Timetabling*, 557–559
- Van der Veen, E., Hans, E. W., Post, G. & Veltman, B. (januari 2012). *Shift rostering using decomposition: assign weekend shift first*. Universiteit Twente. <<http://eprints.eemcs.utwente.nl/21419/>> (technisch rapport).

EGBERT VAN DER VEEN werkt als consultant personeelsplanning bij ORTEC en als promovendus bij het Center for Healthcare Operations Improvement and Research (CHOIR) van de Universiteit Twente.

E-mail: <Egbert.vanderVeen@ortec.com>

ERWIN HANS is als universitair hoofddocent werkzaam voor de Universiteit Twente en eveneens betrokken bij het Center for Healthcare Operations Improvement and Research (CHOIR).

E-mail: <e.w.hans@utwente.nl>



Figuur 5. Percentage niet ingeroosterde diensten. Per instantie, per aanpak, het percentage van de weekenddiensten en het percentage van het totaal aantal diensten dat niet geroosterd wordt.

Gewetensconflicten van een sociaal-wetenschappelijke onderzoeksmethodoloog

JARL K. KAMPEN

We moeten er geen doekjes om winden. Zowel in Vlaanderen als in Nederland was het in november en december 2012 niet mogelijk een dagblad open te slaan zonder dat er beschouwingen werden gegeven over slodderwetenschap en wetenschappelijke fraude. In Nederland werden in korte tijd niet minder dan vier rapporten over bedenkelijke wetenschappelijke praktijken in de sociale wetenschappen neergelegd. Vlaamse hoogleraren pleitten voor een algemeen *mea culpa* van alle leden van de wetenschappelijke gemeenschap: onderzoekers, reviewers, editors van tijdschriften, leden van beoordelingscommissies, en bestuurders van universiteiten en andere onderzoeksinstellingen. Ik behoor tot deze wetenschappelijke gemeenschap. Ik heb als onderzoeksmethodoloog een taak te vervullen. En hoewel ik niet van mening ben dat ik een *mea culpa* dien uit te spreken, wil ik wel getuigenis doen van gewetensconflicten waarmee ik in mijn beroepspraktijk worstel.

Mijn taken omvatten het geven van onderwijs, het doen van methodologisch onderzoek, en het geven van consulten over methodologische vraagstukken. Het beroep van methodologiedocent beoefen ik gewetensvol,

met grote vreugde, en bovendien tot voldoening van mijn studenten. Ik ervaar hier geen gewetensconflicten (ook al moet ik soms met een glimlach denken aan mijn natuurkundeleraar op de middelbare school, die vond dat het kindermishandeling was dat tweedeklassers natuurkunde moesten leren). Als onderzoeker heb ik zelf in de hand wat ik publiceer, en ook al zal ik in de afgelopen 15 jaar hier en daar echt wel eens een steekje hebben laten vallen, ik ben mij niet bewust van ernstige fouten en ik heb altijd getracht mijn bevindingen waarheidsgetrouw te rapporteren. Ook hier ervaar ik geen gewetensconflicten. Waar ik in deze bijdrage dan ook wat dieper op in wil gaan is mijn ervaringen met consulteren en taken die daaraan verwant zijn, zoals het becommentarieren van onderzoeksverslagen en -plannen van collega-onderzoekers met een toegepaste strekking.

Stel, iemand brengt een onderzoeksverslag of -plan bij mij met de vraag of ik dit wil beoordelen. Dit gebeurt natuurlijk regelmatig, en is een zeer aan te moedigen praktijk. Ik ben methodoloog, en ongeacht het onderwerp beoordeel ik zo'n verslag op één merite: hoe (zullen) de wetenschappelijke claims worden gerecht-

vaardigd. Allereerst onderzoek ik de centrale onderzoeksvraag. Is het eigenlijk wel een onderzoeksvraag, in de zin dat de vraag oproept tot het verwerven van kennis, of is het eigenlijk een nauw verholen ambitie om de wereld te gaan verbeteren? Vervolgens ga ik op zoek naar de wijze waarop de concepten in een onderzoeksvraag worden geoperationaliseerd (d.w.z. meetbaar worden gemaakt). Soms strand ik al op dat punt. Bijvoorbeeld, men wil het oorzakelijke verband tussen X en Y testen maar laat na om aan te geven hoe Y wordt (of zal worden) gemeten. Wanneer de operationalisering in orde is bevonden, blader ik door naar de conclusies. Sluiten die aan bij de onderzoeksvraag? En in hoeverre worden conclusies gegeneraliseerd? Daarbij zijn er twee klassen generalisaties. De eerste heeft betrekking op externe validiteit (steekproef, steekproefpopulatie, iedereen), de tweede op wat wel eens wordt aangeduid als 'ecologische' validiteit. Ik zal een voorbeeld geven.

Stel dat u de hypothese wilt testen dat kennis van klimaatverandering leidt tot milieubewuster gedrag. U beschikt over een gelegenheidssteekproef van studenten die u willekeurig verdeelt over twee groepen. In beide groepen neemt u een vragenlijst af waarin u peilt naar milieubewustzijn (pre-test). Vervolgens laat u de experimentele groep de film *An Inconvenient Truth* zien (de interventie), terwijl de controle groep twee uur over zijn eigen tijd mag beschikken. En dan neemt u een post-test af waarin u opnieuw peilt naar milieubewustzijn. Externe validiteit gaat over de vraag of u de bevindingen in uw steekproef van studenten mag generaliseren naar studenten in het algemeen, burgers in het algemeen, enz. Ecologische validiteit gaat over het vraagstuk of u uw interventie in de kennis van klimaatverandering, in dit geval het laten zien van de beroemde film met Al Gore, kunt generaliseren naar een populatie (of klasse) van interventies die het verhogen van kennis over klimaatverandering tot doel hebben. Heeft deze film net zoveel effect als een goed boek, of een aantrekkelijke docent, of een andere documentaire, enz.? Of is het *deze specifieke film* die eigenlijk verantwoordelijk is voor het verschil dat u vond? In het laatste geval kunt de stelling dat meer kennis leidt tot meer bewustzijn niet rechtvaardigen.

Dit klinkt misschien erg ingewikkeld, maar in de praktijk ben ik nu hooguit een paar minuten aan het bladeren geweest in het verslag. Dit is routine. Op het moment dat ik weet wat de onderzoeksvraag is, hoe begrippen in die vraag worden geoperationaliseerd, en tot op welk niveau de resultaten van het onderzoek worden gegeneraliseerd, kan ik systematisch uitpluizen in hoe-

verre de wetenschappelijke claims die de onderzoekers maken worden gerechtvaardigd door het onderzoeksontwerp dat men heeft gebruikt. Opnieuw ga ik een rijtje af: wat is de meetvaliditeit, wat is de interne validiteit, en wat is de externe validiteit van het onderzoek. Nu pas consulteer ik dus de delen van het verslag die gaan over methode, veldwerk en analyse. En dan is het gewoon een kwestie van strepen. Er wordt zelf-gerapporteerd gedrag geregistreerd in een interviewsetting maar de conclusies gaan over gedrag in de echte wereld? De werkzaamheid van een interventie wordt geclaimd op basis van een experiment zonder controlegroep? Er worden non-parametrische tests (bv. Mann-Whitney) uitgevoerd met als motivatie dat de steekproeven niet-random zijn? Gevonden significantieniveaus worden niet gecorrigeerd voor kapitalisatie op kans? Er wordt gebruik gemaakt van lineaire regressie op ordinale uitkomsten, maar er is geen sprake van enige modeldiagnostiek en bovendien werkten de subjecten in groepjes (zodat de waarnemingen niet onafhankelijk zijn)? Enzovoort. Ik ben hiervoor geschoold, ik doe dit graag, en wanneer ik dit doe in een *review* voor een internationaal tijdschrift of bijvoorbeeld in het kader van een *systematic review* of een literatuurstudie, dan ontstaat er ook geen enkel probleem.

Het gewetensconflict ontstaat hier. Wat als deze methodologische check (op verzoek) gebeurt op een tijdschriftartikel van collega's vlak voordat dit wordt ingediend, of zelfs heringediend? Of wanneer de check wordt gedaan op een proefschrift in de allerlaatste fase? Wat moet ik doen met mijn constatering over tekortkomingen in het onderzoek die zeker niet werden gemaakt met de bedoeling om onjuistheden te debiteren, laat staan om fraude te plegen. Wanneer gaat de grijze zone van aanvechtbare wetenschappelijke uitspraken over in de zwarte zone van de wetenschappelijke fraude? Wanneer ik zwijg (wat uiteraard vanuit professionele gronden niet mogelijk is), en de betrokken onderzoekers dus onwetend zijn en te goeder trouw hun onderzoeken publiceren, dan is er natuurlijk geen sprake van fraude. Maar wat als ik niet zwijg, en mijn bemerkingen bovendien hout snijden, is de situatie dan nog steeds dezelfde? Kan de jonge onderzoeker nog op dezelfde manier zijn onderzoek verdedigen, of moeten er consequenties worden verbonden aan het maken van onterechte generaliseringen? Ik heb geen pasklaar antwoord op deze vraag.

JARL KAMPEN is assistent professor bij de Onderzoeksmethodologiegroep (RME) van Wageningen Universiteit en Research Centre.

E-mail: <jarl.kampen@wur.nl>

DE DERDE INTERNATIONALE TIMETABLING COMPETITIE

op zoek naar het beste schoolrooster

De Internationale Timetabling Competitie 2011 vond plaats van 1 oktober 2011 tot 11 juli 2012. Het was de derde internationale competitie op het gebied van roosterproblemen, geheel gewijd aan automatisch genereren van jaarroosters voor middelbare scholen. De 17 deelnemers waren universitaire studenten, onderzoekers of onderzoeksgroepen uit 14 verschillende landen in Europa, Amerika en Azië. In dit artikel beschrijven we achtereenvolgens de vorige competities, de roosterproblemen en de opzet en resultaten van de competitie.

JEFFREY H. KINGSTON & GERHARD POST

PATAT en de eerdere competities

In 1995 werd de eerste conferentie gehouden met de naam PATAT, wat staat voor Practice And Theory of Automated Timetabling. Hoewel deze naam op de lachspieren van veel Nederlandstaligen werkt, is het een serieuze conferentie met de laatste jaren rond de 100 deelnemers. Sinds 2000 vindt deze serie van conferenties tweejaarlijks plaats in steeds een ander land. PATAT 2012 was in Son (Noorwegen), waar de prijzen van ITC 2011 werden uitgereikt.¹ De deelnemers aan de conferentie vormen een gemêleerd gezelschap: natuurlijk onderzoekers (veel PhD's), maar ook mensen uit het bedrijfs-

leven en *practitioners*, gebruikers van roostersoftware. De onderzoekers komen uit de informatica, bedrijfskunde en operations research. De nadruk ligt daarmee op de praktijk: de onderzoekers zijn vooral gericht op het oplossen van praktijkproblemen. Hiervoor zijn in de loop der jaren benchmark-instanties beschikbaar gekomen, waarbij de instanties voor Examination Timetabling (zie <<http://www.cs.nott.ac.uk/~rxq/data.htm>>) het langst in gebruik zijn. In de loop der jaren zien we dan ook onderzoek waarin de algoritmes steeds beter presteren op deze benchmarks.

Een competitie is een uitdagender en beter verifieerbare manier om algoritmes te vergelijken. De structuur

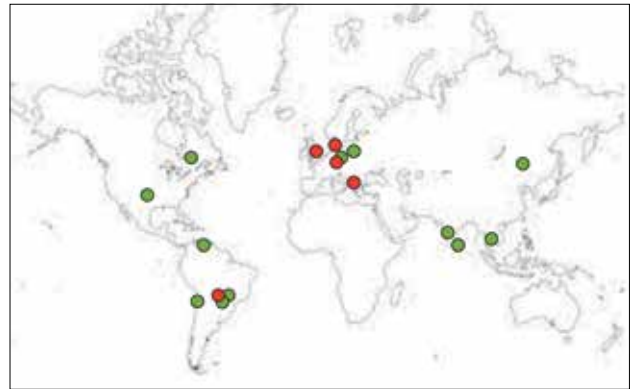
van een dergelijke competitie is dat de onderzoekers een probleembeschrijving krijgen, met een aantal voorbeeldinstanties waarop de algoritmes getest kunnen worden. Doel is het vinden van een toegelaten oplossing met de laagste kosten in een van tevoren vastgelegde (genormeerde) rekentijd. Hierbij moet de *run*, gegeven de *random seed*, reproduceerbaar zijn. Op basis van de voorbeeldinstanties worden de finalisten bepaald. Van deze finalisten wordt vervolgens de broncode gevraagd en hiermee worden door de organisatie de verborgen instanties met een aantal verschillende random seeds opgelost. Per run krijgt de winnaar 1 punt, nummer 2 krijgt 2 punten, enzovoort; deze punten noemen we de rangscore. De winnaar is het team met de laagste gemiddelde rangscore.

De eerste International Timetabling Competition (zie <www.idsia.ch/Files/ttcomp2002>) liep van 1 oktober 2002 tot 31 maart 2003. Het probleemgebied was het maken van een universitair rooster. De uitgangssituatie is dat alle vakken (met bijbehorende *events*) van de universiteit met hun studenten bekend zijn. Deze events moeten gedurende de week van 5 keer 9 uur geroosterd worden. Hierbij mogen de studenten en zalen natuurlijk niet dubbel geroosterd worden, en de zalen moeten qua capaciteit toereikend zijn. Merk op dat de docenten geen belangrijke rol spelen. De problemen waren met *reverse engineering* geconstrueerd: van elke instantie is bekend dat er een perfecte oplossing met kosten 0 bestaat. Het aantal deelnemende teams bedroeg ongeveer 10.

De tweede International Timetabling Competition (ITC 2007, zie <www.cs.qub.ac.uk/itc2007>) liep van 1 augustus 2007 tot 25 januari 2008. Dit was een zeer uitgebreide competitie met *tracks* voor Examination Timetabling, Post Enrolment based Course Timetabling en Curriculum based Course Timetabling. Vanaf een iets hoger standpunt zijn deze problemen nauw verwant: de *events* moeten op tijden gepland worden, waarbij conflicten (van studenten, curricula en zalen) vermeden moeten worden, de capaciteiten van de zalen gerespecteerd en een zekere spreiding van *events* verlangd wordt. Het aantal deelnemende teams lag boven de honderd.

High School Timetabling

Het jaarroosterprobleem op middelbare scholen is wellicht het meest tot de verbeelding sprekende roosterprobleem; iedereen heeft het aan den lijve gevoeld en mogelijk geklaagd over het aantal tussenuren in het



Figuur 1. De 17 deelnemers aan de derde Internationale Timetabling Competitie (ITC 2011) waren universitaire studenten, onderzoekers of onderzoeksgroepen uit 14 verschillende landen in Europa, Amerika en Azië.

rooster. In *STAtOR* (2010)³ is al besproken wat typisch Nederlandse problemen zijn. Internationaal is er veel onderzoek gedaan naar High School Timetabling, maar meestal beperkte dit onderzoek zich tot één land of zelfs één school; vergelijkend onderzoek tussen verschillende algoritmes op een breed scala van instanties was daarmee niet-bestaand. Een karakteristiek verschil met de probleemgebieden van ITC 2007 is dat er nu serieus rekening moet worden gehouden met de kwaliteit van docentenroosters.

Op de PATAT conferentie van 2006 in Brno (Tsjechië) besloten 4 onderzoekers om datasets bijeen te brengen in een nauwkeurig omschreven (nog te ontwikkelen) XML-formaat. Na twee jaar bleven beide auteurs van dit artikel over om dit project te trekken. Na ongeveer 20 versies werd het uiteindelijke formaat XHSTT (XML for High School Timetabling) vastgelegd en een evaluator gemaakt (zie <<http://sydney.edu.au/engineering/it/~jeff/hseval.cgi>>). Deze evaluator, gebaseerd op het eerder door de eerste auteur ontwikkelde KHE (Kingston High School Timetabling Engine) raamwerk, controleert de syntax en geeft een overzicht van de kosten van de eventueel meegegeven oplossingen. Er zijn nu bijna 50 instanties beschikbaar in XHSTT (zie www.utwente.nl/ctit/hstt), afkomstig uit elf verschillende landen. De instanties zijn voor een belangrijk deel in eerder onderzoek gebruikt.

Tussen de instanties zijn er grote verschillen in omvang en complexiteit. Zo zijn er landen (Brazilië, Kosovo, Zuid-Afrika) waar verschillende groepen op verschillende tijden van de dag les hebben. Andere landen (Australië, Groot-Brittannië) hebben juist korte roosters (van 8 uur tot 13 uur), waarbij alle uren bezet zijn voor leerling en docent. Hier kan de a priori koppeling van docenten

aan de klassen leiden tot onoplosbare conflicten. Het opsplitsen van klassen in deelgroepen komt in veel landen voor, maar op verschillende manieren. Al met al is deze variëteit een prima uitdaging voor onderzoekers om algoritmes te ontwikkelen die met heel verschillende situaties kunnen omgaan.

Het verloop van ITC 2011

In 2011 werd gezocht naar organisatoren van tracks voor een nieuwe competitie. Voor het High School Timeta-bling project was dit het juiste moment om de benchmarks onder de aandacht te brengen, en om inzicht te krijgen in de instanties en de aangeleverde oplossingen. Uiteindelijk bleek onze track de enige en werd de organisatie van onze track versterkt met de ervaren competitierotten Luca Di Gaspero, Barry McCollum en Andrea Schaerf. Wij besloten om de competitie uit drie rondes te laten bestaan.

In de eerste ronde konden deelnemers betere oplossingen op 16 instanties uit het archief XHSTT-2012 indienen.² Er was geen beperking op gebruik van programmatuur en andere hulpmiddelen. Per instantie werd de uiteindelijk beste oplossing beloond met een kleine prijs. In deze ronde stuurden vier teams in totaal 74 (steeds betere) oplossingen in. Daarmee kende de eerste ronde

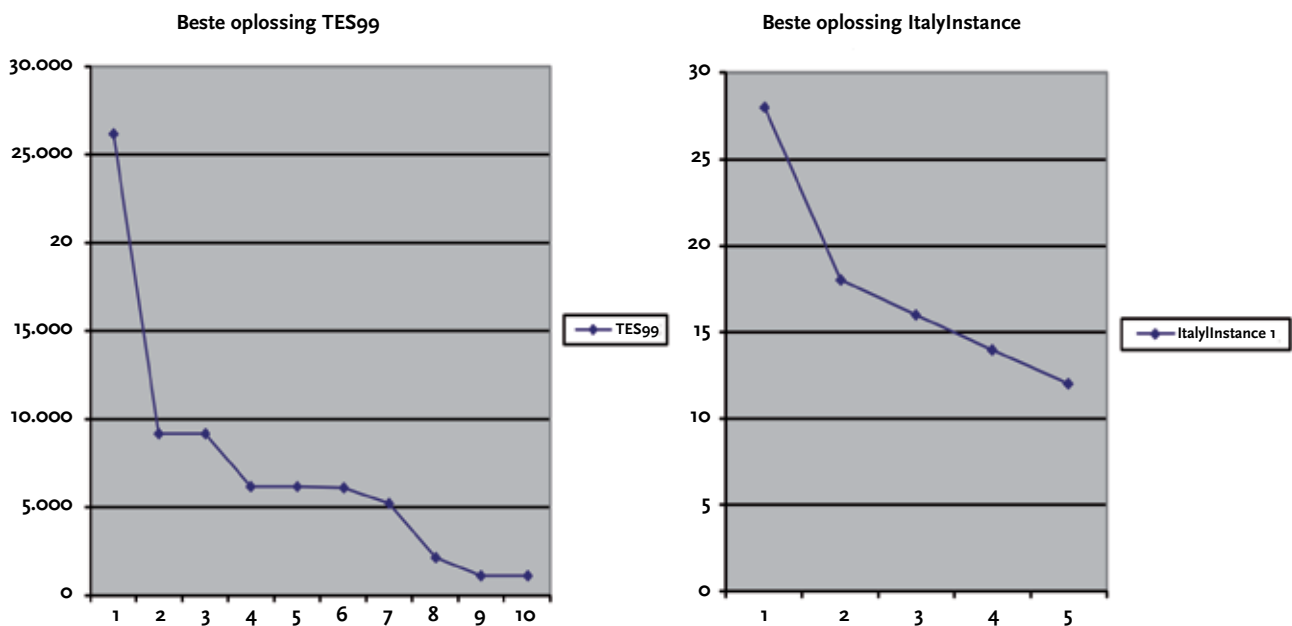
een geanimeerd verloop, zoals ook in figuur 2 te zien is. Uiteindelijk werden voor 15 (!) instanties betere oplossingen gevonden met een gemiddelde kostenreductie van 58%.

Ronde 2 was de traditionele ronde. Gebruik van commerciële externe solvers als CPLEX en Gurobi was niet toegestaan. Hier werden de programma's van de 4 finalisten getest op 18 verborgen instanties. Deze instanties waren voor een deel van dezelfde leveranciers als die uit de eerste ronde, en voor een deel van nieuwe leveranciers. De eindscore werd, als in eerdere competities, bepaald door de gemiddelde rangscore op 10 random runs van de instanties. Winnaar werd het Braziliaanse team Goal, met een zeer goede eindscore van 1,18.

In ronde 3 was, net als in de eerste ronde, de methode vrij maar werd de winnaar, net als in ronde twee, bepaald door de gemiddelde rangscore op de (nu niet meer) verborgen instanties. Ook hier was Goal de winnaar, nu met een gemiddelde rangscore van 1,64.

Conclusie

Uiteindelijk hebben slechts 5 van de 17 ingeschreven deelnemers oplossingen ingestuurd, de rode stippen in figuur 1. Een enquête onder de deelnemers liet zien dat een aantal andere teams wel behoorlijke inspanningen



Figuur 2. Kostenverloop van twee instanties in ronde 1.

(50 dagen of meer) hadden geleverd, maar toch niet het punt hadden bereikt om actief deel te kunnen nemen. Inderdaad is alleen al het opzetten van het model met de evaluatie van 13 constraint types behoorlijk veel werk. Van de vijf actieve deelnemers hebben er twee het KHE raamwerk gebruikt, waarmee een model en zelfs een aantal eenvoudige algoritmes ter beschikking stond. Een derde deelnemer gebruikte een commercieel roosterpakket als basis, en een vierde team de programma's waarmee het eerder onderzoek had gedaan.

De competitie heeft voor het eerst vergelijkend onderzoek in High School Timetabling opgeleverd. Wij hopen dat in de komende jaren onderzoekers deze lijn zullen voortzetten, zelfs als er geen geldelijke prijs mee te verdienen is. Mede op basis van de ervaringen van de competitie is een archief van instanties gemaakt (zie <www.utwente.nl/ctit/hstt/archives/XHSTT-2013>), waarvan wij denken dat het als benchmark kan dienen voor dit toekomstige onderzoek.

NOTEN

1. PATAT 2014 vindt plaats in Engeland. De opmerking dat de conferentie eerder al in Edinburgh en Belfast was, werd door Edmund Burke gepareerd met de opmerking "That's not England!".
2. Het archief XHSTT-2012 bevat 21 instanties, waarvan van 5 al een optimale oplossing bekend was.

LITERATUUR

ITC 2011 <www.utwente.nl/ctit/hstt/itc2011>. Hier is een volledig overzicht van de van de competitie en de resultaten te vinden.

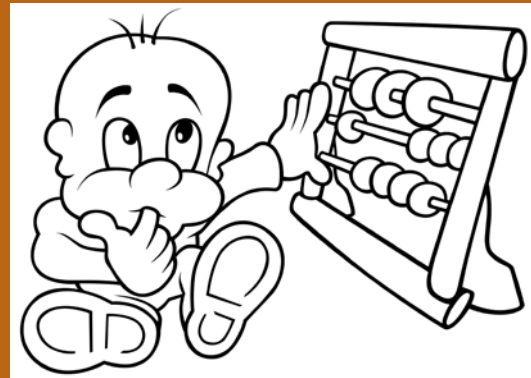
Post, G., Ruizenaar, H. & Bannink, R. (2010). Roosteren op een middelbare school. *STAtOR*, 3, 24–27.

Post, G., Di Gaspero, L., Kingston, J. H., McCollum, B. & Schaerf, A. (2012). The Third International Timetabling Competition. In *PATAT 2012; Proceedings of the Ninth International Conference on the Practice and Theory of Automated Timetabling* pp. 479–484. Trondheim: Sintef.

JEFFREY H. KINGSTON is als Honorary Associate en onderzoeker verbonden aan de School of Information Technologies van de University of Sydney. Zijn interesses zijn algoritmes en complexiteit in het bijzonder op het gebied van roosterproblemen. Ook is hij geïnteresseerd in documentopmaak en als zodanig auteur van het opmaakprogramma Lout.
E-mail: <jeff@it.usyd.edu.au>

GERHARD POST is parttime universitair docent bij de afdeling Toegepaste Wiskunde van de Universiteit Twente. Zijn onderzoek richt zich, naast de gezondheidszorg en energie, op roosterproblemen. Daarnaast is hij parttime werkzaam als OR-specialist bij ORTEC bv.
E-mail: <G.F.Post@utwente.nl>

Young Statisticians



STATISTICAL SCIENCE CAFE DECEMBER 18: SUCCESSFUL MEETING

In last mid-December I participated in the Statistical Cafe. The event was held in a cosy pub in the city center of Utrecht. The atmosphere was nice and I was rather surprised by the number of young statisticians. The topic 'Integrity or Fraud?' and was rather interesting, which of course attracted me in the first place. Prof. dr. Richard Gill gave a impressive presentation on how several cases of fabricating data were discovered, in particular the specific statistical tools (certain tests etc.) that he explained as means of detecting dishonesty in data. That methodology was rather inspiring to me, as a statistician. Later in Marcel van Assen's presentation, he elaborated on the ethic and technical issues arising in the discussions of scientific fraud. After that the young statisticians from various disciplines started asking questions, which put the discussions up to a certain heat, and there were excellent arguments and suggestions on both sides. I had several interesting interactions about the topic with other fellow statisticians. I had a nice time but still have a suggestion for the next statistical cafe: 'More free drinks'!

Fengnan Gao

A meeting is foreseen early this spring, where new activities will be planned. Please consult the Young Statisticians website regularly for up-to-date information or send an e-mail to <young@vvs-or.nl>.

REVENUE MANAGEMENT

kansen voor de luchtvrachtindustrie

Sinds de deregulering in de personenluchtvaart in de jaren tachtig van de vorige eeuw zijn in deze industrie steeds complexere Revenue Management (RM) systemen toegepast om prijzen aan te kunnen passen aan de specifieke vraag van klanten. Het succes van de toepassing van deze systemen werd al snel opgemerkt door andere sectoren die zich hier ook meer op zijn gaan richten. De bekendste voorbeelden hiervan zijn de hotel- en autoverhuursector. In deze industrieën is, net als bij luchtvaart, sprake van een vaste capaciteit en uiteenlopende wensen van klanten. Dit geldt ook voor de luchtvracht, maar ondanks vele parallellen worden RM systemen in deze branche niet of nauwelijks gebruikt. In dit artikel zal de luchtvrachtindustrie worden beschreven, RM in het algemeen worden beschouwd, worden bekeken wat RM in andere sectoren heeft opgeleverd, waarom RM tot nu toe in luchtvracht nauwelijks nog wordt toegepast en waar de kansen voor RM in luchtvracht liggen die voor alle actoren in de keten voordelig zullen zijn.

THIJS BOONEKAMP

Schets van de luchtvrachtindustrie

Luchtvracht is de duurste vorm van transport, vandaar dat het alleen wordt gebruikt als het strikt noodzakelijk is, vanwege de snelheid, betrouwbaarheid en flexibiliteit.

Hierdoor wordt slechts 1,5% van alle wereldvracht via de lucht getransporteerd. Belangrijke voorbeelden van producten die via de lucht worden vervoerd zijn:

- Producten met hoge waarde in klein bestek (medicijnen, goud, hightech)

- Spoedzendingen met onderdelen voor productielijnen (spare parts)
- Producten met korte economische waarde (groenten/bloemen/kranten)
- Andere speciale vracht zoals dieren, gekoelde producten of radioactief materiaal

Belangrijke actoren in de luchtvrachtbranche zijn de airlines, de verladers en de expediteurs.

Airlines beheren en verkopen capaciteit voor luchtvracht. Ongeveer de helft van alle vracht gaat mee met passagiersvliegtuigen. De capaciteit is dus afhankelijk van het passagiersnetwerk waarop de airline vliegt en met welke frequentie. Ook is de capaciteit afhankelijk van de ruimte die passagiers en hun bagage innemen. De andere helft van de vracht gaat mee op *full freighters*. Deze vrachtvliegtuigen hebben een veel grotere capaciteit en kunnen ook hogere lading vervoeren, die niet in personenvliegtuigen mee kan.

De verlader is de partij waar het te verzenden product vandaan komt.

Expediteurs zijn de tussenpersonen in de keten. Zij bemiddelen tussen de verlader en de airline, en zorgen voor het voor- en natransport vanaf de betrokken luchthavens. Ook regelen zij alle papieren en douaneformaliteiten voor de zending. Expediteurs hebben vaak contracten met airlines waarin zij een bepaalde hoeveelheid capaciteit reserveren tegen een gunstig tarief. Dit heeft als voordeel dat airlines zeker zijn van een bepaalde hoeveelheid verkochte capaciteit, en expediteurs kunnen capaciteit goedkoop inkopen en tegen een hogere marge doorverkopen. Een nadeel voor airlines is dat dit soort contracten vaak niet bindend zijn: er zijn geen kosten verbonden aan het afwijken van de afgesproken capaciteit door de expediteurs. In figuur 1 is de gehele keten schematisch weergegeven.

Vandaag de dag worden tarieven voor de vracht veel-

al door middel van onderhandeling bepaald. Expediteurs sluiten aan het begin van het jaar contracten met de airline op basis van hun verwachte klantvraag naar capaciteit. De overige capaciteit wordt op de vrije markt verkocht voor losse zendingen. Boeking worden over het algemeen telefonisch gedaan; er wordt dan per boeking onderhandeld over de prijs en de bijbehorende voorwaarden. Expediteurs onderhandelen aan beide kanten: zij proberen een gunstige prijs met de airline af te spreken en deze met een zo hoog mogelijke marge aan de verladers te verkopen. Prijzen worden op basis van de ervaring en kennis van de markt van het sales-personeel bepaald.

Vaak betekent dit dat alle binnenkomende vraag naar capaciteit volgens een FCFS beleid wordt geaccepteerd. Omdat zendingen die laat binnenkomen vaak meer opleveren, zou het meer winst op kunnen leveren als er capaciteit wordt gereserveerd voor vraag die op het laatste moment arriveert.

Wanneer er een duidelijke tariefstructuur wordt bepaald op basis van verschillende types zendingen met uiteenlopende eigenschappen, kunnen prijzen worden bepaald die aansluiten bij de wensen van de klant. Een RM-systeem zou kunnen helpen met het bepalen welke types vracht tegen welke prijs geaccepteerd moeten kunnen worden, zodat op de beperkte beschikbare capaciteit een zo hoog mogelijke winst kan worden bepaald.

Wat is RM en wat heeft het opgeleverd?

Bij Revenue Management gaat het erom het juiste product voor de juiste prijs op de juiste tijd aan de juiste klant te verkopen, zodat de winst wordt gemaximaliseerd. RM is met name winstgevend in industrieën waar een grote verscheidenheid aan klanten is die verschillende wensen hebben en daardoor bereid zijn een ander



Figuur 1. Schematische weergave van de gehele keten van de luchtvrachtindustrie.

bedrag voor hetzelfde goed te willen betalen.

Als er op een geschikte manier prijsdifferentiatie kan worden toegepast is het mogelijk om meer capaciteit te benutten en met dezelfde kosten een hogere omzet te bewerkstelligen. Klanten die bereid zijn veel te betalen zullen een hoger tarief betalen, maar krijgen daar bijvoorbeeld meer flexibiliteit of zekerheid voor terug. Daarentegen kunnen klanten met minder specifieke wensen tegen een lager tarief mee, terwijl deze klanten verloren waren als er maar één vast tarief gehanteerd zou worden.

In figuur 2 is weergegeven op welke manier een geschikte prijsdiscriminatie voordeel op zou kunnen leveren. Dit is niet alleen gunstig voor degene die het product verkoopt, maar ook voor de klanten: de gemiddelde prijs is gedaald.

De grootste uitdaging hierbij is het definiëren van marktsegmenten die aansluiten bij de wensen van de klant. Belangrijk is dat aan elk marktsegment duidelijke eigenschappen worden verbonden die voorkomen dat een klant van het ene naar het andere segment over kan springen. Een goed voorbeeld hiervan in het verschil tussen business class- en economy class-passagiers in personenluchtvaart. Een business class-passagier is bijvoorbeeld een stuk flexibeler om zijn vlucht om te boeken of kosteloos te annuleren.

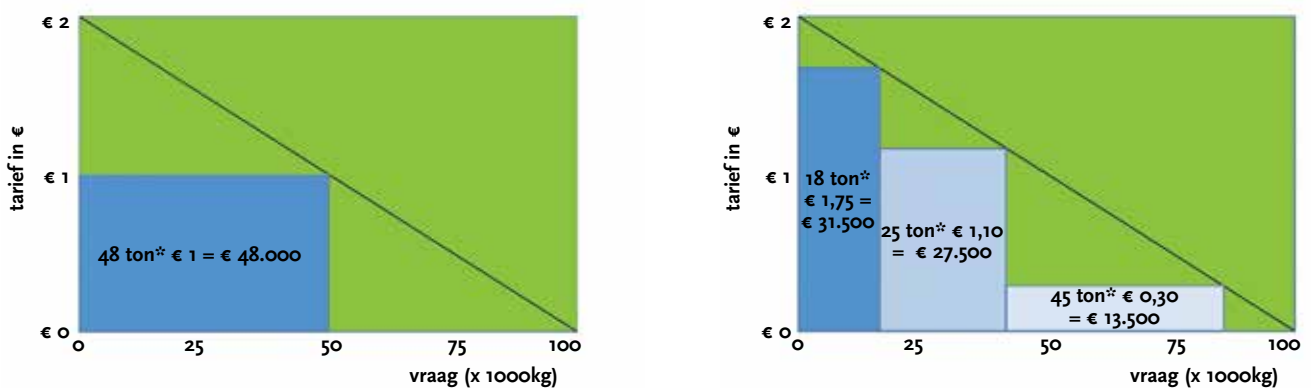
Op dit moment kunnen ingewikkelde RM-systemen in de personenluchtvaart het verschil maken tussen het schrijven van zwarte cijfers in plaats van rode. In alle industrieën waar RM succesvol toegepast is, worden winststijgingen tussen de 2% en 5% toegerekend aan het gebruik van RM (Yeoman & McMahon-Beattie, 2011). Ook bij experimenten met RM bij een grote Ne-

derlandse luchtvrachtmaatschappij is sprake van een significante winsttoename (Slager & Kapteijns, 2004).

Waarom nog geen RM in de luchtvracht?

Waarom blijft de luchtvracht ver achter in het toepassen van RM-systemen, terwijl dit in de personenluchtvaart juist zo ver ontwikkeld is? De belangrijkste reden is dat luchtvracht een veel complexer product is dan bijvoorbeeld vliegtuigstoelen of hotelkamers. Deze complexiteit wordt uitvoerig beschreven door Kasilingam (Kasilingam, 1996). Hieronder worden de belangrijkste punten samengevat.

- **Vrachtcapaciteit heeft meerdere dimensies** – Capaciteit is afhankelijk van gewicht en volume, en eventueel ook nog het aantal beschikbare containerposities in een vliegtuig.
- **Contractvormen** – Het merendeel van de capaciteit wordt verkocht in *allotments* of andere contractvormen, waar een grote hoeveelheid ruimte voor een grote klant gereserveerd wordt.
- **Vrijheden in de route** – Waar in personenluchtvaart vaak voor een bepaalde (rechtstreekse) vlucht wordt geboekt, zijn er veel meer mogelijkheden in de luchtvracht, zolang de vracht binnen de afgesproken tijd op de bestemming aankomt. Ook is een route vaak maar 1 richting uit, terwijl personen meestal ook retour vliegen.
- **Variabele capaciteit** – Omdat een groot deel van de luchtvracht met personenvliegtuigen mee gaat, is de beschikbare capaciteit voor vracht afhankelijk van de capaciteit die gebruikt wordt voor de bagage van passagiers.



Figuur 2. De vraag bij vaste prijzen en bij prijsdifferentiatie.

- **Onvoorspelbare vraag** – Veel vraag is afhankelijk van onvoorspelbare factoren. Luchtvracht wordt veel gebruikt voor bederfelijke goederen en voor spoedzendingen. Pieken in dit soort zendingen zijn afhankelijk van het weer, of onvoorziene calamiteiten.
- **Onzekere vraag** – Annuleringen of *no-shows* komen vaak voor. Ook gebeurt het regelmatig dat er meer of minder vracht verschijnt dan geboekt is.

Wat zijn de kansen voor RM in de luchtvracht?

In het boek van Talluri wordt een aantal criteria besproken waaraan industrieën zouden moeten voldoen opdat zij een goede business zijn om RM in toe te passen (Talluri & Van Ryzin, 2004).

- **Heterogeniteit van klanten** – Er zijn verschillende soorten klanten die elk een andere waarde toekennen aan gebruik van hetzelfde product. Al deze klanten hebben specifieke voorkeuren, bijv. flexibiliteit of speciale eisen.
- **Variabiliteit en onzekerheid van de vraag** – Als de vraag variabel is door seizoens- of tijdeffecten, kan er meer winst worden behaald door prijzen te veranderen.
- **Vergankelijke capaciteit** – Ongebruikte capaciteit kan niet worden opgeslagen om een andere keer te gebruiken. Ook kan de capaciteit niet snel worden aangepast aan de uiteindelijke vraag. Dus capaciteit moet zo goed mogelijk worden benut.

Klanten in de luchtvrachtbranche zijn zeer uiteenlopend. Producten die worden verzonden lopen uiteen van medicijnen en dieren tot aan grote losse machineonderdelen of bederfelijke goederen. Alle verladers van deze verschillende producten stellen andere eisen, bijvoorbeeld speciale zorg voor het product of de urgentie waarmee het verzonden dient te worden.

Er is een zeer grote onzekerheid in de vraag naar luchtvracht. Dit is iets dat de luchtvracht ook een stuk complexer maakt, maar zeker ook een mogelijkheid om effectievere prijsdiscriminatie toe te passen. Ook zijn seizoenseffecten erg belangrijk, denk bijvoorbeeld aan bloemen op Russische vrouwendag of spelcomputers voor de kerst.

Alle capaciteit die niet is benut, is misgelopen omzet. De extra kosten om een product mee te nemen zijn, als er nog ruimte is, afgezien van extra brandstofkosten minimaal.

Bovenstaande eigenschappen gelden dus allemaal voor de luchtvrachtsector. Het zal dus zeker voordelen bieden wanneer er in deze sector op een juiste manier

Revenue Management toegepast wordt. Een aantal grote luchtvaartmaatschappijen is daarom ook al begonnen met het gebruiken van RM-systemen. Uit deze recente toepassingen blijkt ook dat dit zeker winstgevend is. Gebruik van een RM-systeem zou in drie jaar voor 8% winststijging hebben gezorgd (Revenue Management: Not just a gut feeling, 2011).

Een ideaal RM-systeem zou op basis van een bepaalde tariefstructuur met verschillende categorieën vracht moeten kunnen bepalen welke types al dan niet worden geaccepteerd. Het systeem moet:

- accurate voorspellingen geven naar de vraag van elk type vracht;
- een minimale prijs bepalen die binnenkomende vraag op moet kunnen leveren;
- advies geven of de vraag wel of niet moet worden geaccepteerd.

Dit systeem zou moeten worden gebruikt door de professionals die op dit moment de onderhandelingen voor hun rekening nemen. Implementatie van een RM-model heeft als voordelen dat de prijs beter correspondeert met de wensen van de klant, er meer winst kan worden gemaakt op de beschikbare capaciteit en dat personeel zich meer kan richten op de 'ingewikkeldere' vracht, omdat veel beslissingen door het systeem kunnen worden genomen.

Er moet een aantal obstakels worden overwonnen voordat RM in luchtvracht kan worden toegepast, maar door verder onderzoek te doen naar mogelijke algoritmes om het cargo RM-probleem op te lossen kan veel winst voor deze industrie worden behaald.

LITERATUUR

- Kasilingam, R. G. (1996). Air Cargo Revenue Management: Characteristics and Complexities. *European Journal of Operational Research*, 36–44.
- Revenue Management: Not just a gut feeling. (2011). *Air Cargo World Magazine*.
- Slager, B. & Kapteijns, L. (2004). Implementation of cargo revenue management at KLM. *Journal of Revenue and Pricing Management* (3), 80–90.
- Talluri, K. & Van Ryzin, G. (2004). *The theory and practice of Revenue Management*. Kluwer Academic Publishers.
- Yeoman, I. & McMahon-Beattie, U. (2011). *Revenue management, a Practical Pricing Perspective*. Chippenham and Eastbourne: Palgrave Macmillan.

THIJS BOONEKAMP doet een afstudeeronderzoek voor zijn master Econometrics & OR aan de VU Amsterdam naar Revenue Management in luchtvracht bij ACN Air Cargo Netherlands.
E-mail: <t.boonekamp@student.vu.nl>

Facebook en moorden in New York

Het klassieke verjaardagsprobleem is alom bekend. Hoe groot moet een willekeurig samengestelde groep van personen zijn om een *fifty-fifty* kans te hebben dat twee of meer personen op eenzelfde dag jarig zijn? Het antwoord is 23. Dit onder de aanname dat elke dag even waarschijnlijk is als verjaardagsdag, waarbij tevens uitgegaan wordt van 365 dagen in het jaar. Overigens vinden geboortes verre van uniform verdeeld over het jaar plaats, maar dit maakt de kans dat twee of meer personen op eenzelfde dag jarig zijn alleen maar groter.

Een aardige variant van het verjaardagsprobleem kwam onlangs in een radioprogramma aan de orde. Als een persoon vele honderden of zelfs duizenden vrienden op Facebook heeft, wat is dan de kans dat op elke dag van het jaar wel één van zijn of haar vrienden jarig is? Hoeveel vrienden heb je nodig zodat sprake is van een *fifty-fifty* kans? Vele honderden vrienden volstaat niet. Het antwoord is dat je 2287 vrienden nodig hebt voor een *fifty-fifty* kans. De berekening van deze exacte waarde is niet direct eenvoudig, maar met een algemeen toepasbare methode kan simpel een goede benadering bepaald worden. Stel iemand heeft n vrienden op Facebook met $n > 365$. Denk eens in dat de volgende 365 experimenten worden uitgevoerd. In het i -de experiment wordt nagegaan of een vriend op dag i jarig is en het experiment wordt succesvol genoemd als dit *niet* het geval is. De succeskans van elk experiment is $(364/365)^n$ en het verwachte aantal successen is $\lambda = 365 \times (364/365)^n$. Welbekend is dat bij een groot aantal onafhankelijke experimenten elk met eenzelfde kleine succeskans, het aantal succesvolle experimenten goed beschreven wordt door een Poisson-verdeling. In ons geval zijn de 365 experimenten niet onafhankelijk, maar de afhankelijkheid is

voldoende zwak ter rechtvaardiging van de benadering $e^{-\lambda}$ voor de kans op geen enkel succes. Oftewel, voor de kans p_n dat elke dag voorkomt als verjaardagsdag onder de n vrienden geldt

$$p_n \approx e^{-365 \times (364/365)^n}.$$

Oplossen van de vergelijking $p_n = 0,5$ geeft de benadering $n = 2285$ voor de grootte van de vriendengroep zodat met een *fifty-fifty* kans elke dag van het jaar voorkomt als verjaardagsdag onder de vrienden. Overigens zal een dergelijke omvang van de vriendengroep op Facebook niet vaak voorkomen. De gemiddelde grootte van de vriendengroep op Facebook is 130.

Het Facebook verjaardagsprobleem onthult ook een journalistieke blunder in enkele grote Nederlandse kranten die eind november 2012 het volgende bericht plaatsten. 'Een bizar feit uit New York: maandag was de eerste dag sinds mensenheugenis dat er niemand vermoord werd in de Amerikaanse miljoenenstad.' Een statistisch mirakel is dit niet, want in hetzelfde bericht werd gemeld dat het aantal moorden in New York gemiddeld ongeveer 500 per jaar is over de laatste jaren. De Nederlandse kranten hadden het Amerikaanse nieuws slordig overgenomen. Die maandag 26 november 2012 was de eerste dag sinds mensenheugenis dat geen enkel ernstig misdrijf (moord, steekpartij, zware mishandeling) gemeld was in New York.

HENK TIJMS is emeritus hoogleraar operations research aan de Vrije Universiteit en auteur van diverse leerboeken over operations research en kansrekening.
E-mail: <tijms@quicknet.nl>

HET PETERPRINCIPE

Stel je voor dat je op een dag besluit je volledig te bekwamen in dat wat je het liefst doet, of het nu voetbal, koken, of wiskunde betreft. Je gaat je uitsluitend richten op dat ene ding, en laat geld hierbij geen rol spelen.

Laten we wiskunde als voorbeeld nemen. Wiskunde ontvangt jouw absolute devotie. Van kindsbeen af doe je niet aan tuinieren, want een uur tuinieren is een uur wiskunde. Muziek, literatuur, en zeker ook columns schrijven, alles laat je links leggen in naam der wiskunde. Het vergt immers jarenlange oefening om het vak in de vingers te krijgen. En zo ga je wiskunde studeren aan de universiteit. Na vier jaar is daar je bul en ook meteen je eerste baan als specialist in de wiskunde. De eerste jaren bij het bedrijf gaan voortvarend, en je krijgt al snel de reputatie van de man die alles kan. Je collega's weten je te vinden en dragen je op handen. Je krijgt lof alom. Je ademt wiskunde.

Maar met het verstrijken der jaren begin je te twijfelen aan je vroegere credo. Een combinatie van omgeving en ambitie, doet je op een dag besluiten je voor een beter salaris te laten omscholen tot manager. Vrijwel al je collega's zijn je per slot van rekening al voorgegaan. En zo word je teamleider, jij die slechts één ding echt goed kunt. Opeens leid je een team van wiskundigen, en moet je rekening houden met gevoe-

lens, relaties, en meer dingen waar je minder bedreven in bent.

Toch doe je het goed. Niet uitzonderlijk, maar goed. Het fanatisme waarmee je eens de wiskunde bedreef, wend je nu aan voor het leiding geven. Je volgt managementcursussen, en hoort van het Peterprincipe, het fenomeen dat men promotie blijft maken tot het niveau van incompetentie is bereikt. Je leert moeilijke gesprekken voeren en oefent liftpraatjes met acteurs. En zo maak je weer promotie, en geef je niet langer leiding aan een team van wiskundigen, maar aan een grote commerciële afdeling. Wiskunde speelt geen rol meer. Je salaris stijgt en je besluit dat grotere huis in het nabijgelegen dorp te kopen. Veel grond, dat wel, maar je bent inmiddels gaan tuinieren. Een week voor de verhuizing, bij het opruimen van de zolder, vind je een doos met oude schoolrapporten. Je leest terug hoe je maar moeilijk aansluiting vond bij de andere kinderen, weinig overwicht had en wat achterbleef met lezen. In hoofdletters lees je hoe geweldig goed je kon rekenen. Het zweet breekt je uit, maar je beseft dat de dozen zijn ingepakt en de loopbaan is voltrokken.

JOHAN VAN LEEUWAARDEN is hoogleeraar wiskunde aan de TU Eindhoven.

E-mail: <j.s.h.v.leeuwaarden@TUE.nl>



conferentie **IMPACT ON PEOPLE, BUSINESS AND SOCIETY**

3 – 6 september 2013 in Rotterdam

De Erasmus Universiteit Rotterdam organiseert in het kader van haar 100-jarig bestaan, samen met het Nederlands Genootschap Besliskunde (NGB) en de Duitse OR vereniging (GOR), een grote, internationale conferentie op gebied van Operations Research: OR2013. De conferentie heeft als thema *Impact on People, Business and Society*.

plenaire sprekers

Alexander Rinnooy Kan
Steve Sashihara

semi-plenaire sprekers (bevestigd)

Miguel Anjos
Antonio Conejo
Rommert Dekker
Erik Demeulemeester
Dick den Hertog
Fred Espen Benth
Nelly Litvak
Andrea Lodi
Kalyan Talluri
Horst Tempelmeier
Stein Wallace

abstracts

Het indienen van een abstract is mogelijk tot 30 april 2013.

awards

Voor jonge besliskundigen zijn er twee prijzen beschikbaar gesteld, gerelateerd aan het thema van de conferentie: Young participant with most academic impact & Young participant with most practical impact. Aanmelden voor 1 april 2013.

Voor meer informatie, zie www.or2013.org

