CWI Syllabi

CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands
Telephone + 31 - 20 592 9333
Telefax + 31 - 20 592 4199
WWW page http://www.cwi.nl/publications/

CWI is the nationally funded Dutch institute for research in Mathematics and Computer Science.

Vakantiecursus 2003
Wiskunde in het dagelijks leven

Centrum voor Wiskunde en Informatica
CWI SYLLABUS
52

De Vakantiecursus Wiskunde voor leraren in de exacte vakken in VWO, HAVO en HBO en andere belangstellenden is een initiatief van de Nederlandse Vereniging van Wiskundeleraren. De cursus wordt sinds 1946 jaarlijks gegeven op het Centrum voor Wiskunde en Informatica en aan de Technische Universiteit Eindhoven.

Deze cursus is mede mogelijk gemaakt door een subsidie van de Nederlandse Organisatie voor Wetenschappelijk Onderzoek.

ISBN 90 6196 520 9
NUGI-code: 811

Inhoud

Medewerkers

Drs. R. Bosch, Koninklijke Militaire Academie, Postbus 90002, 4800 PA Breda, 076-5273267, R.Bosch2@mindef.nl

Prof.dr. J. van de Craats, Koninklijke Militaire Academie, Postbus 90154, 4800 RG Breda, 076-5273816, J.vd.Craats@mindef.nl

Prof.dr. G.B. Huitema, TNO Telecom, PAV B16, Postbus 15.000 9700 CD Groningen, 050-5821024, G.B.Huitema@telecom.tno.nl

Drs. M.F.M. Nuyens, Korteweg de Vries Instituut, Universiteit van Amsterdam, Plantage Muidergracht 24, 1018 TV Amsterdam, 020-5256070, mnuyens@science.uva.nl

Drs. R. Planqué, CWI, Kruislaan 413, Postbus 94079, 1090 GB Amsterdam, 020-5924234, R.Planque@cwi.nl

Mw.dr. V. Rottschäfer, Mathematisch Instituut, Universiteit Leiden, Postbus 9512, 2300 RA Leiden, 071-5277113, vivi@math.leidenuniv.nl

Prof.dr.ir. P.J.G.Teunissen, Technische Universiteit Delft, Afdeling Geodesie, Thijsseweg 11, 2629 JA Delft, 015-2782558, P.J.G.Teunissen@geo.tudelft.nl

Ir. W.A. Timmer, Technische Universiteit Delft, Civiele Techniek en Geowetenschappen, Stevinweg 1, 2628 CN Delft, 015-2788279, W.A.Timmer@citg.tudelft.nl

Prof. dr. H.J. van Zuylen, Technische Universiteit Delft, Civiele Techniek en Geowetenschappen, Postbus 5048, 2600 GA Delft, 015-2782761, H.J.vanZuylen@ct.tudelft.nl

Dr. E. van Zwet, Universiteit Leiden, Postbus 9500, 2300 RA Leiden, vanzwet@math.leidenuniv.nl

Contacten Centrum voor Wiskunde en Informatica

Dr. M. Bakker
Centrum voor Wiskunde en Informatica, Kruislaan 413, Postbus 94079, 1090 GB Amsterdam, 020 592 4172, Miente.Bakker@cwi.nl

Wilmy van Ojik
Centrum voor Wiskunde en Informatica, Kruislaan 413, Postbus 94079, 1090 GB Amsterdam, 020 592 4200, Wilmy.van.Ojik@cwi.nl

Ten geleide

Jan van de Craats
Koninklijke Militaire Academie, Breda
e-mail: J.vd.Craats@mindef.nl

De wiskunde van vandaag staat met beide benen stevig in de maatschappij. Ook de CWI-Vacantiecursussen van de afgelopen jaren hebben laten zien dat wiskunde de meest uiteenlopende toepassingsmogelijkheden in de moderne samenleving heeft, en dat die toepassingen ook bijna allemaal wel aanrakingspunten hebben met de wiskunde op school en de belevingswereld van scholieren. De positieve reacties van de deelnemers tonen aan dat dit soort onderwerpen zeer gewaardeerd worden.

Ook dit jaar zetten we die trend daarom voort met een programma dat als thema gekregen heeft *Wiskunde in het dagelijks leven*. De sprekers behandelen ieder vanuit hun eigen deskundigheid onderwerpen waarbij iedere Nederlander zich wel iets kan voorstellen. Dat er ook wiskunde aan te pas komt, zal soms een verrassing zijn. De onderwerpen laten een bont scala zien: plaatsbepaling op aarde met behulp van satellietverbindingen, de wiskunde van de rozenkweker, de eurodiffusie, paradoxen van machtsverhoudingen en stemsystemen, de aerodynamica van schaats- en fietspakken. Ook zijn er twee voordrachten over verkeersproblemen op autosnelwegen en een voordracht over verkeersproblemen op de digitale snelweg. De rode draad bij al die voordrachten is de rol van de wiskunde.

Gaarne wil ik hier allen bedanken die in 2003 opnieuw een Vacantiecursus mogelijk hebben gemaakt. In de eerste plaats natuurlijk de sprekers, die naast hun lezing ook een tekst voor deze Syllabus hebben geleverd. Daarmee wordt opnieuw een aantrekkelijk deel toegevoegd aan een serie syllabi met voor leraren en andere belangstellenden uiterst waardevol materiaal. Het Centrum voor Wiskunde en Informatica te Amsterdam en de Technische Universiteit Eindhoven stelden zaalruimte beschikbaar, de administratieve en praktische organisatie van de cursus was in handen van mw. Wilmy van Ojik en dr. Miente Bakker, die ook samen met mevrouw Minnie Middelberg de inhoudelijke coördinatie van deze syllabus verzorgde.

Allen hartelijk dank!

CWI Vakantiecursus 2003

pp. 3 – 14

GPS and integer estimation

Peter Teunissen
Technische Universiteit Delft
e-mail: P.J.G.Teunissen@citg.tudelft.nl

1. Introduction

Precise ranges for positioning with the Global Positioning System (GPS) are obtained from carrier phase measurements. These measurements of range inherently contain unknown integer ambiguities to account for the mismatch of a whole number of wavelengths or cycles. This contribution describes the problem of GPS carrier phase ambiguity resolution, discusses some relevant elements of integer estimation theory and reviews some of the high precision positioning applications that come into reach when the integer carrier phase ambiguities can be resolved quickly and correctly.

2. Redundant measurements

As in other physical sciences, empirical data are used in geodesy to make inferences so as to describe the physical reality. Many such problems involve the determination of a number of unknown parameters which bear a linear or linearized relationship to the set of data. In order to be able to check for errors or to reduce for the effect these errors have on the final result, the collected data often exceed the minimum necessary for a unique solution (redundant data). As a consequence of measurement uncertainty, redundant data are usually inconsistent in the sense that each sufficient subset yields different results from another subset. Hence, redundancy generally leads to an inconsistent system of linear(ized) equations, say

$$y \cong Ax \tag{1}$$

where vector y contains the m observations, vector x the n unknown parameters. The $m \times n$ matrix A relates the observations to the parameters. Redundancy of the above system is defined as $m - \mathrm{rank}A$, which in case of a full rank matrix simplifies to $m - n$, the difference between the number of observations and the number of unknown parameters.

The above inconsistent system is without additional criteria not uniquely solvable. The problem of solving an inconsistent system of equations has attracted the attention of leading scientists in the middle of the 18th century. Historically, the first methods of combining redundant measurements originate from studies in geodesy and astronomy, namely from the problem of determining the size and shape of the Earth, and the problem of finding a mathematical representation of the motions of the Moon. Since its discovery almost 200 years ago, the method of least-squares has been and still is too a large extent one

of the most popular methods of solving an inconsistent system of equations. Although the method of least-squares may seem 'natural' for a student in modern times, its discovery evolved only slowly from earlier methods of combining redundant observations [1]

GPS positioning basically is determining the location of a (user) receiver with respect to satellites of which the locations (orbits) are known. This determination takes place by measuring distances, and from a geometric point of view three measurements would suffice to determine the three coordinates of the user (fortunately we know on which side of the satellite configuration the Earth is located). The simplest example of (1) in case of GPS is therefore when distances are measured from an unknown GPS receiver position to more than three GPS satellites of which the positions are known. Since the distance from the unknown receiver position r to the known position of satellite s is a nonlinear function of the unknown position coordinates,

$$l_r^s = \sqrt{(x^s - x_r)^2 + (y^s - y_r)^2 + (z^s - z_r)^2} \tag{2}$$

the common approach is to approximate this relation by a linearized version, i.e. developing the nonlinear relation in a Taylor series with zeroth and first order terms only, using good approximate values for the unknown parameters. As a result the (increments of the) observed distances are collected in vector y, the (increments of the) three unknown coordinates in vector x and the partial derivatives in matrix A. In reality the equations are far more complicated than (2) due to the fact that one also has to account for clock errors, atmospheric delays and orbital errors.

3. Least-squares

Around 1800 Legendre and Gauss, see figure 1, at the same time (most likely independently), invented the method of least-squares for solving an inconsistent system of equations. The least-squares solution to (1) reads

$$\hat{x} = (A^T Q_y^{-1} A)^{-1} A^T Q_y^{-1} y \tag{3}$$

with Q_y^{-1} being the weight matrix. This solution is obtained by first adding an unknown error vector e to (1), giving the consistent but undetermined system $y = Ax + e$, and then minimizing the weighted norm of e, $\| e \|_{Q_y}$, as function of x. The least-squares estimator has various desirable properties. When the positive definite matrix Q_y is chosen as the variance-covariance matrix of the observations, the least-squares estimator has the smallest variance (best possible precision) of all linear unbiased estimators.

The geometric interpretation of what least-squares does to the observations is shown in figure 2. The inconsistency between observations on one hand and model (with unknown parameters) on the other is removed by orthogonal projection. Vector $\hat{y} = A\hat{x}$ eventually lies in the plane or linear manifold spanned by the columns of matrix A (indicated by $R(A)$). The orthogonal projection realizes shortest distance between the original observation values y and the adjusted ones $\hat{y}$; the observation values are modified as little as possible,

Figure 1. *Carl Friedrich Gauss (1777–1855) is portrayed on the former 10 Mark banknote in Germany. The banknote also shows the town of Göttingen and the Gaussian or normal probability density function*

though satisfying the assumed model afterwards. This follows from interpreting the least-squares estimation principle as the principle of least distance

$$\min_x \|y - Ax\|^2_{Q_y} \tag{4}$$

The (squared and weighted) distance between y and $\hat{y} = A\hat{x}$ is minimized.

In order to evaluate the quality of the least-squares solution in a probabilistic sense, we need the probability density function (pdf) of $\hat{x}$. Since $\hat{x}$ of (3) is a linear function of y, the least-squares estimator has a Gaussian distribution whenever y is Gaussian distributed. The pdf of the unbiased least-squares estimator $\hat{x}$ can therefore be uniquely characterized by means of the variance-covariance matrix of $\hat{x}$. With Q_y being the variance-covariance matrix of the observations, application of the error propagation law to (3) gives the variance-covariance matrix of the estimated parameters as

$$Q_{\hat{x}} = (A^T Q_y^{-1} A)^{-1} \tag{5}$$

This matrix can be used to evaluate the precision of the parameter estimators, as for instance the position coordinates.

4. GPS carrier phase observable

GPS observations of distance or range are obtained by measuring signal traveltimes (from satellite to receiver) and multiplying these by the speed of light. Two types of distance measurements are employed: pseudo range code and carrier phase. The code observation is based on the (binary) code the satellite modulates onto the signal carrier; the distance can be measured virtually unambiguously. For the carrier phase, the receiver measures the difference in phase between the carrier wave received from the satellite and the reference carrier wave it generated itself. The (physical) phase difference reads

$$\psi_r^s = \phi_r - \phi^s \tag{6}$$

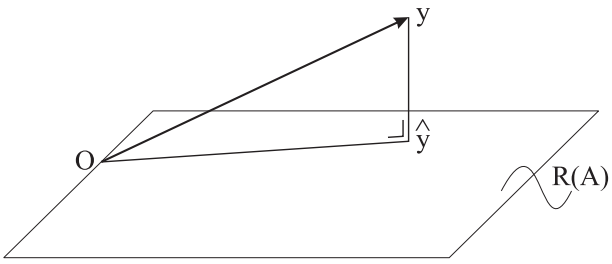


Figure 2. *Least-squares estimation implies a (n orthogonal) projection of the observation vector y onto the plane spanned by the columns of matrix A. Example with three observations and two unknown parameters*

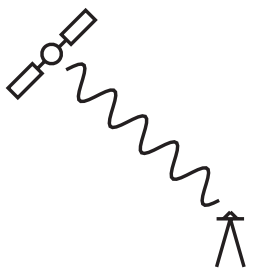


Figure 3. *Measurement of phase on the continuous carrier wave transmitted by the satellite*

With some simplifying assumptions, the phase of a carrier wave at some epoch t equals frequency f multiplied by time t: $\phi = ft$. The receiver compares the reference carrier at time of observation t_r with the carrier received from the satellite, which was generated a little earlier in order to be 'in time' at the receiver, namely at $t_r - \tau_r^s$, where τ_r^s is the signal travel time from satellite to receiver.

The above phase difference becomes

$$\psi_r^s = f\tau_r^s \tag{7}$$

and when multiplied by wavelength $\lambda = \frac{c}{f}$, $\lambda\psi_r^s = c\tau_r^s = l_r^s$, the distance in meters is obtained; it equals the travel time pre-multiplied by the speed of light c, exactly as with the code observation.

As a consequence of carrying out measurements on a (monotone) continuous carrier wave, the receiver can not distinguish one cycle from another. The satellite keeps on transmitting the carrier wave, in principle cycle after identical cycle, see figure 3.

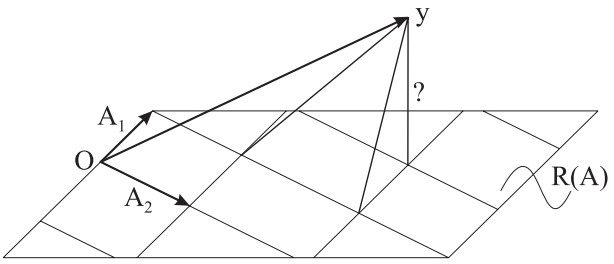


Figure 4. *Least-squares with integer parameters: possible solutions for the vector of observations form a grid in the column-space of matrix A (A_1 and A_2 are two columns of matrix A); the solution is no longer allowed to lie anywhere in the plane $R(A)$*

At some epoch in time the receiver simply starts outputing the measured *fractional* difference in phase: $\mathrm{frac}(\psi_r^s) \in [0, 1\rangle$ cycle. The full (physical) phase difference is then decomposed into

$$\psi_r^s = \underbrace{\mathrm{int}(\psi_r^s)}_{N_r^s} + \underbrace{\mathrm{frac}(\psi_r^s)}_{\phi_r^s} \tag{8}$$

The observed (fractional) phase difference ϕ_r^s (times the wavelength) does thereby not equal the distance from satellite to receiver l_r^s, but equals this distance apart from an integer number of wavelengths

$$\lambda\phi_r^s = l_r^s - \lambda N_r^s \tag{9}$$

As a consequence the vector x in (1) will, next to the unknown receiver coordinates, now also contain unknown *integer* cycle ambiguities N_r^s.

5. Integer least-squares

The least-squares solution (3) is obtained from solving (4), where x is allowed to vary over the whole n- dimensional space of *real* numbers. In case of GPS however, when use is made of the carrier phase observations, the vector of unknown parameters x consists of both real-valued and integer valued parameters (real-valued coordinates and integer-valued carrier phase ambiguities). We therefore need to modify the solution (3) so as to take the integerness of some of the parameters into account. To keep the discussion simple, it will be assumed here that *all* parameters in vector x are integer-valued. Due to the integerness of the parameters, orthogonal projection of y will now not do the job properly, see figure 4. Nevertheless one can start with 'ordinary' least-squares as a first step, see figure 5. The solution so obtained for the unknown parameters will be real-valued and is usually referred to as the 'float' solution.

To apply the least-squares principle (4), but now under the condition that the parameters in x are all integers, a second step has to be carried out. Since

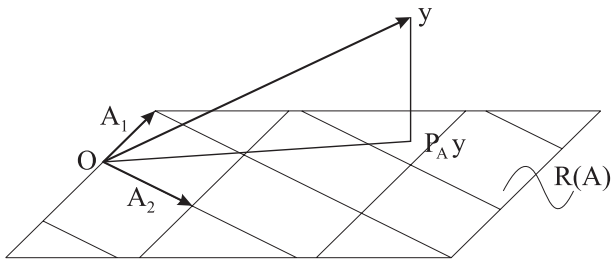


Figure 5. *Least-squares with integer parameters: the first step consists of 'ordinary' least-squares (orthogonal projection); the solution $\hat{x}$ for the parameters will consist of real-valued numbers*

the first step projects orthogonally to the plane $R(A)$, the second step takes place *in* the plane. From the orthogonal decomposition

$$\|y - Ax\|^2_{Q_y} = \|y - \hat{y}\|^2_{Q_y} + \|\hat{y} - Ax\|^2_{Q_y} \tag{10}$$

it follows that the second step amounts to solving the minimization problem

$$\min_x (\hat{y} - Ax)^T Q_y^{-1} (\hat{y} - Ax) =$$

$$\min_x (\hat{x} - x)^T A^T Q_y^{-1} A (\hat{x} - x) = \min_x (\hat{x} - x)^T Q_{\hat{x}}^{-1} (\hat{x} - x) \tag{11}$$

for x being integer, where in the last equation (5) has been used. This minimization can also be visualized in the parameter space, see figure 6, instead of in the observation space as in figures 2 and 4.

The integer least-squares principle has been applied very successfully to GPS ambiguity resolution. By the presence of the variance-covariance matrix $Q_{\hat{x}}$ in (11), the precision and correlation of the individual real-valued ambiguity estimates is properly and fully exploited. In contrast to the 'ordinary' least-squares solution (3), there does not exist an analytical solution to (11). In practice a search over possible integer solutions has to be carried out. The space of integer solutions is restricted by limiting the squared and weighted distance in (11) to a convenient value. As a result, the volume of the corresponding ellipse (or hyper-ellipsoid in higher dimensions) has to be searched through in order to find the integer least-squares solution of x.

When the ambiguities of the first step are of poor precision and at the same time highly correlated, the ellipse or ellipsoid gets very elongated and narrow. As a consequence the discrete search may get computationally inefficient. For computational efficiency the quadratic form (11) can be integer transformed, so that the resulting ellipsoid becomes more sphere-like and the transformed ambiguities become less correlated [2], [3].

6. Alternative integer estimators

Instead of the integer least-squares estimator one can also think of alternative integer estimators. Starting from the 'float' solution, such an estimator $\breve{x} =$

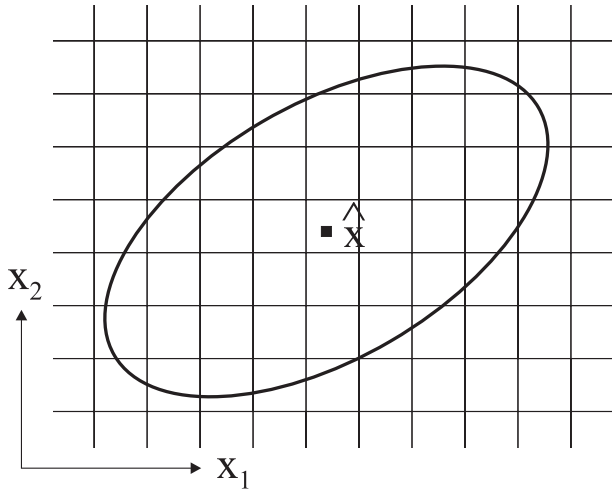


Figure 6. *Least-squares with integer parameters: in the second step the integer solution for x is sought that is* closest *to the real-valued solution $\hat{x}$ of the first step; 'closest' is to be measured in the metric of the variance-covariance matrix $Q_{\hat{x}}$; the quadratic form (11), set equal to a constant, is represented by the ellipse in this example with two ambiguities x_1 and x_2*

$F(\hat{x})$ will consist of a mapping $F : R^n \mapsto Z^n$ from the n-dimensional space of real numbers to the n-dimensional space of integers. Due to the discrete nature of Z^n, the map F will not be one-to-one. This implies that different real-valued ambiguity vectors will be mapped to the same integer vector. One can therefore assign a subset $S_z \subset R^n$ to each integer vector $z \in Z^n$:

$$S_z = \{x \in R^n \mid z = F(x)\}, \quad z \in Z^n \tag{12}$$

The subset S_z contains all real-valued ambiguity vectors that will be mapped by F to the same integer vector $z \in Z^n$. This subset is referred to as the *pull-in-region* of z. It is the region in which all ambiguity 'float' solutions are pulled to the same 'fixed' ambiguity vector z.

Since the pull-in-regions define the integer estimator completely, one can define classes of integer estimators by imposing various conditions on the pull-in-regions. One such class is given as follows [4].

An integer estimator is said to be *admissible* if

$$\begin{array}{ll} (i) & \bigcup_{z\in Z^n} S_z = R^n \\ (ii) & S_{z_1} \bigcap S_{z_2} = \{0\}, \ \ \forall z_1, z_2 \in Z^n, z_1 \neq z_2 \\ (iii) & S_z = z + S_0, \ \ \forall z \in Z^n \end{array} \tag{13}$$

This definition is motivated as follows. Each one of the above three conditions describe a property of which it seems reasonable that it is possessed by

an arbitrary integer ambiguity estimator. The first condition states that the pull-in-regions should not leave any gaps and the second that they should not overlap. The absence of gaps is needed in order to be able to map any 'float' solution $\hat{x} \in R^n$ to Z^n, while the absence of overlaps is needed to guarantee that the 'float' solution is mapped to just one integer vector. Note that the pull-in-regions are allowed to have common boundaries. This is permitted if we assume to have zero probability that $\hat{x}$ lies on one of the boundaries. This will be the case when the probability density function (pdf) of $\hat{x}$ is continuous.

The third and last condition follows from the requirement that $F(x+z) = F(x)+z, \forall x \in R^n, z \in Z^n$. Also this condition is a reasonable one to ask for. It states that when the 'float' solution is perturbed by $z \in Z^n$, the corresponding integer solution is perturbed by the same amount. This property allows one to apply the *integer remove-restore* technique: $F(\hat{x}-z)+z = F(\hat{x})$. It therefore allows one to work with the fractional parts of the entries of $\hat{x}$, instead of with its complete entries.

There exist various admissible integer estimators. The simplest integer map is the one that corresponds to integer rounding. In this case the integer vector is obtained from a rounding of each of the entries of $\hat{x}$ to its nearest integer. Since componentwise rounding implies that each real-valued ambiguity estimate $\hat{x}_i, i = 1, \ldots, n$, is mapped to its nearest integer, the absolute value of the difference between the two is at most $\frac{1}{2}$. The subsets $S_{R,z}$ that belong to this integer estimator are therefore given as

$$S_{R,z} = \cap_{i=1}^{n} \{\hat{x} \in R^n \mid \mid \hat{x}_i - z_i \mid \leq \frac{1}{2}\} \ , \ \forall z \in Z^n \tag{14}$$

The subset $S_{R,z}$ is an n-dimensional cube, with sides of length 1 and centred at the grid point z.

Another relatively simple integer ambiguity estimator is the integer bootstrapped estimator. This estimator can be seen as a generalization of the previous one. It still makes use of integer rounding, but it also takes some of the correlation between the ambiguities into account. The bootstrapped estimator results from a sequential conditional least- squares adjustment and is computed as follows. If n ambiguities are available, one starts with the first ambiguity $\hat{x}_1$, and rounds its value to the nearest integer. Having obtained the integer value of this first ambiguity, the real-valued estimates of all remaining ambiguities are then corrected on the basis of their correlation with the first ambiguity. Subsequently the second, but now corrected, real-valued ambiguity estimate is rounded to its nearest integer. Having obtained the integer value of the second ambiguity, the real-valued estimates of all remaining $n-2$ ambiguities are again corrected, but now on the basis of their correlation with the second ambiguity. This process of rounding and correcting is continued until all ambiguities are taken care of.

With c_i denoting the ith canonical unit vector having a 1 as its ith entry, the pull-in-regions $S_{B,z}$ that belong to the bootstrapped estimator can be shown to be given as

$$S_{B,z} = \cap_{i=1}^{n} \{\hat{x} \in R^n \mid \mid c_i^T L^{-1}(\hat{x}-z) \mid \leq \frac{1}{2}\} \ , \ \forall z \in Z^n \tag{15}$$

with matrix L being the unit lower triangular matrix of the triangular decomposition of $Q_{\hat{x}}$. Note that these pull-in-regions reduce to the ones of (14) when L becomes diagonal. This is the case when the ambiguity variance- variance-covariance matrix is diagonal. In that case the two integer estimators $\check{x}_R$ and $\check{x}_B$ are identical.

The third admissible estimator of which the pull-in-region will be given is the integer least-squares estimator. By again using the LDL^T-decomposition of $Q_{\hat{x}}$ the least-squares' pull-in-region reads

$$S_{LS,z} = \cap_{c_i \in L^{-1}(Z^n)} \{\hat{x} \in R^n \mid \mid c_i^T D^{-1} L^{-1} (\hat{x} - z) \mid \leq \frac{1}{2} c_i^T D^{-1} c_i\} \qquad (16)$$

Note that (16) and (15) become identical when the matrix entries of L^{-1} are all integer. This is the case when L is an admissible ambiguity transformation.

7. The ambiguity success rate

The quality of the integer ambiguity estimator is particularly of interest in case of GPS. One therefore needs the probability mass function (pmf) of $\check{x}$. It can be obtained as follows. Using the concept of the pull-in-region, the integer estimator is defined as $\check{x} = z \quad \Leftrightarrow \quad \hat{x} \in S_z$. Hence, for the probability masses one has $P(\check{x} = z) = P(\hat{x} \in S_z)$. With the pdf of $\hat{x}$ given as $p_{\hat{x}}(x)$, the pmf of $\check{x}$ follows as

$$P(\check{x} = z) = \int_{S_z} p_{\hat{x}}(x) dx \ , \ \forall z \in Z^n \qquad (17)$$

The *ambiguity success rate* is defined as the probability of correct integer estimation $P(\check{x} = x)$. Note that the pmf (17) as well as the success rate still depend on the type of pull-in-region and thus on the type of integer estimator chosen. Changing the geometry of the pull-in-region will change both the pmf and the ambiguity success rate. It is therefore of interest to know which integer estimator maximizes the ambiguity success rate. The answer is given by the following theorem [4]:

Let the pdf of $\hat{x}$ be elliptically contoured and the integer least-squares estimator be given as

$$\check{x}_{ILS} = \arg \min_{z \in Z^n} \parallel \hat{x} - z \parallel^2_{Q_{\hat{x}}}$$

Then

$$P(\check{x}_{ILS} = x) \geq P(\check{x} = x) \qquad (18)$$

for any admissible estimator $\check{x}$.

This theorem gives a probabilistic justification for using the *integer* least-squares estimator. It applies to GPS ambiguity resolution for which the pdf $p_{\hat{x}}(x)$ is often assumed to be a multivariate normal distribution. For GPS ambiguity resolution one is thus better off using the integer least-squares estimator than any other admissible integer estimator, such as integer rounding or integer bootstrapping.

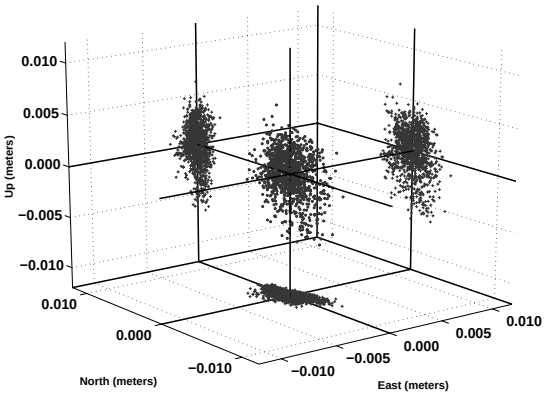


Figure 7. *Example of the repeatability of GPS positions after resolving the ambiguities by means of integer least-squares. The three-dimensional position is obtained from a single epoch of observations (so-called instantaneous positioning). The experiment has been carried out 1200 times, and shown are all 1200 ambiguities-fixed position solutions. The measurement noise in the carrier phase observation is at the few millimeter level and the consequent spread in position is clearly below 1 centimeter*

8. Applications

Once the integer carrier phase cycle ambiguity has been resolved, the phase observation turns into a direct measurement of distance. These phase observations possess millimeter precision and consequently the user receiver position can be determined with a similar level of precision, see figure 7.

Already early in the history of GPS positioning, the application of surveying topography emerged. By taking the GPS receiver to sites and features on the Earth's surface, their locations can be determined and consequently be mapped. Today, GPS positioning is an important tool in producing and maintaining road-maps, town-plans and precise cadastral maps (and databases).

In the early days, precise positions got available only after considerable time spans (of one or several hours). By including the integer constraints on the ambiguities and developing efficient ways of solving the integer least-squares problem, high precision positions become available virtually immediately, see also figure 7. The ambiguities have been demonstrated to be resolved correctly using just one epoch (second) of observations, thus greatly improving surveying productivity. At present the position can be determined directly in the field, by Real-Time Kinematic (RTK) GPS, see figure 8.

Similar equipment and algorithms can be used for high precision navigation of moving vehicles on land, vessels at sea and aircraft in the air. Challenging applications are vessel guidance through narrow straights with critical clearance and landing aircraft in conditions of poor visibility.

Precise GPS positioning anywhere on Earth is of great benefit also to Earth

Figure 8. *Real-Time Kinematic (RTK) GPS surveying: the surveyor directly 'digitizes' the points of interest in the field, by holding the antenna accurately in place for just a few seconds*

sciences. Tectonic plates may move by several centimeters a year with respect to each other. Such motions of the Earth's crust can be monitored with GPS at the required level of precision. This is of particular interest in areas with considerable seismic activity. For instance in California in the United States, with emphasis on the greater Los Angeles metropolitan region, an array of GPS receivers has been installed — under the name of Southern California Integrated GPS Network (SCIGN) — to study geodynamical phenomena. Over 200 locations are covered and GPS receivers are in operation 24 hours a day, 7 days a week. Figure 9 shows an example of a station of the SCIGN.

9. Conclusion

In this paper the problem of the integer cycle ambiguity of the GPS carrier phase observations for ranging has been addressed. The ambiguities are resolved using the integer least-squares principle thus allowing very precise and fast GPS positioning. Since various details were skipped in the above presentation, the interested reader is referred to the many textbooks available on GPS positioning [5-9].

References

1. Teunissen, P.J.G. (2000): A brief account on the early history of adjusting geodetic and astronomical observations. *De Hollandse Cirkel*, 2(1/2), pp. 12-17.

2. Teunissen, P.J.G. (1993): *Least-squares estimation of the integer GPS ambiguities.* Invited lecture. Section IV Theory and Methodology, IAG Ge-

Figure 9. *A GPS receiver and antenna permanently installed for precisely monitoring motions of the Earth's crust. Site Ranchita in California in the US. Photo taken from album at http://www.scign.org/*

neral Meeting. Beijing, China. August. Also in LGR-series No. 6, Delft.

3. Lenstra, H.W. (1981): Integer programming with a fixed number of variables. University of Amsterdam, Dept. of Mathematics, Report 81-3.

4. Teunissen, P.J.G. (1999): An optimality property of the integer least-squares estimator. *Journal of Geodesy*, 73:587-593.

5. Leick, A. (1995): *GPS Satellite Surveying.* 2nd ed. John Wiley, New York.

6. Strang, G. and K. Borre (1997). *Linear algebra, geodesy, and GPS.* Wellesley-Cambridge Press.

7. Teunissen, P.J.G. and A. Kleusberg (1998): *GPS for Geodesy.* 2nd enlarged edition, Springer Verlag.

8. Hofmann-Wellenhof, B., H. Lichtenegger, J. Collins (2001): *Global Positioning System: Theory and Practice.* 5th ed. Springer Verlag.

9. Misra, P. and P. Enge (2001): *Global Positioning System: Signals, Measurements and Performance*, Ganga-Jamuna Press.

CWI Vakantiecursus 2003

pp. 15 – 25

Wiskunde voor de rozenkweker

Vivi Rottschäfer
Universiteit Leiden
e-mail: vivi@math.leidenuniv.nl

1. Introductie

In februari 2002 vond de tweeënveertigste Studiegroep Wiskunde met de Industrie plaats. Een van de problemen betrof het modelleren van de groei van rozen in een kas en ik maakte deel uit van de groep die daaraan heeft gewerkt.

De productie van rozen is steeds commerciëler geworden waarbij ook de concurrentie groot is. Terwijl de ervaring van de rozenkweker nog steeds een grote rol speelt, is het belang van modelleren van de biochemische processen die plaatsvinden toegenomen. De wens van een rozenkweker is natuurlijk om de productie van zijn rozen te optimaliseren bij beperkte kosten. De groei van rozen en daarmee de productie hiervan wordt beïnvloed door het klimaat in de kas. Dit wordt tegenwoordig volledig door de computer gestuurd. Tot nu toe is het programmeren van zo'n klimaatcomputer gebaseerd op de ervaring van de kweker, maar hoe een maximale productie gerealiseerd kan worden is nog niet bekend. Het is namelijk niet voldoende om het interne klimaat constant te houden omdat plotselinge veranderingen in het weer (buiten de kas) en in de seizoenen invloed zal hebben op de temperatuur en andere condities in de kas. Door een plotselinge regenbui, bijvoorbeeld, kan de temperatuur verlagen wat de rozenoogst een aantal dagen of weken later zal beïnvloeden.

Als eerste stap naar het aansturen van het klimaat voor het verkrijgen van een maximale rozenproductie hebben wij een wiskundig model ontwikkeld. Dit beschrijft, gegeven de klimaatcondities, de productie van de totale massa van de rozen in de kas. Uiteindelijk hopen we dat het model gebruikt kan worden om het klimaat in de kas dynamisch in te stellen zodat de rozenproductie maximaal wordt.

Rozen groeien door assimilatie van CO_2, dit proces vindt plaats in de bladeren en heet fotosynthese. De CO_2-assimilatie, en daarom ook de groei van rozen, wordt beïnvloed door verschillende omgevingsfactoren. Over sommige van deze factoren heeft de kweker (enige mate van) controle door bijvoorbeeld het gebruik van verwarming en lampen, het openen of sluiten van ramen en door blindering voor de ramen om schaduw te creëren. Hiermee kan hij de **CO_2-concentratie in de lucht $\mathbf{C_a}$** (door ventilatie), de **relatieve luchtvochtigheid $\mathbf{R_H}$**, de **temperatuur in de kas $\mathbf{T_a}$** en de **lichtintensiteit $\mathbf{I_0}$** veranderen.

2. Het lokale model

Wij zijn gestart met een model dat we hebben afgeleid uit [2, 3]. Dit model beschrijft de snelheid van fotosynthese van één enkel blad met een bepaalde leeftijd afhankelijk van de temperatuur T_a, de luchtvochtigheid R_H en de concentratie CO_2 in de kas C_a en van de hoeveelheid licht wat op het blad valt. Omdat dit slechts de fotosynthese van een blad met één bepaalde leeftijd geeft noemen dit het 'lokale' model. Voor het lokale model blijkt dat de **snelheid van fotosynthese P** per eenheid bladoppervlakte geschreven kan worden als een vergelijking van de vorm

$$P = f(P, a, T_a, R_H, C_a, I). \tag{1}$$

waarbij a de leeftijd van het blad is en I de hoeveelheid licht wat er op valt. De functie f is niet-lineair en expliciet bekend. Aangezien de uitdrukking voor f tamelijk gecompliceerd is in termen van een groot aantal (bekende) constanten geven we deze hier niet, maar verwijzen we voor de formule en verdere details hiervan naar [1].

Aan ons de uitdaging om dit lokale model van de fotosynthese van een enkel blad uit te breiden naar een model dat de gehele rozenproductie in de kas beschrijft. Het mag duidelijk zijn dat dit geen eenvoudige taak is omdat de fotosynthese afhangt van de leeftijd van een blad en de hoeveelheid licht die er op valt. Allebei deze grootheden veranderen dynamisch met de tijd in de kas omdat de rozenplanten groeien. Hierbij ontstaan nieuwe jonge bladeren waardoor de leeftijdsverdeling in de kas verandert en ook zullen de lager gelegen bladeren minder licht ontvangen.

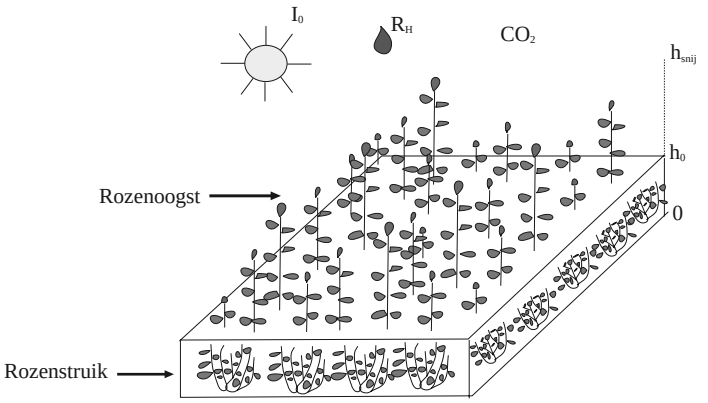
3. Aannames in het model

Om de rozen te modelleren hebben we verschillende aannames gemaakt over de structuur van de rozenplanten en de groei van rozen. Deze aannames zijn in overleg met een adviseur van rozenkwekers (mede bioloog) tot stand gekomen en zijn (redelijk) representatief voor rozengroei in een kas. Allereerst nemen we aan dat elke plant verdeeld kan worden in twee delen:

1. de **rozenoogst**, bovenin, welke bestaat uit 1 of meer **rozenstammen** die afgesneden worden; dit is de uiteindelijke rozenproductie.
2. de **rozenstruik**, onderop, die de stammen ondersteunt en niet geoogst of afgesneden wordt.

Zie figuur 1 voor een schets van onze geïdealiseerde kas. De struik heeft een

constante hoogte $\mathbf{h_0}$ en bevat bladeren die CO_2 assimileren en dus bijdragen aan de totale hoeveelheid energie die wordt geproduceerd in de planten. De rozenstammen in de oogst groeien verticaal uit de rozenstruik en worden geplukt op het moment dat ze de hoogte $\mathbf{h_{snij}}$ hebben bereikt. Op dat moment worden



Figuur 1. *Een schets van onze geïdealiseerde kas. Hier is R_H de relatieve luchtvochtigheid, I_0 de totale lichtintensiteit en h_0 het niveau waarop de rozen afgeknipt worden op het moment dat ze de hoogte h_{snij} hebben bereikt*

de rozen afgesneden op hoogte h_0 zodat ze allemaal dezelfde lengte $h_{snij} - h_0$ hebben. De hoogte $h = 0$ wordt gedefinieerd op de grond van de kas dus bij de onderkant van de struik. We nemen ook aan dat nieuwe stammen uit de bovenkant van de struik groeien. Dus, nieuwe scheuten ontstaan allemaal op hoogte h_0 en verder veronderstellen we dat ze ontstaan met een snelheid die recht evenredig is met de totale fotosynthese in de kas.

Als verdere simplificatie verwaarlozen we dat deel van de door de rozen geproduceerde energie die gebruikt wordt voor onderhoud, opslag en het ontstaan van bloemen in de planten. We nemen aan dat de energie die verkregen wordt door de fotosynthese in de struik én in de stammen *volledig* gebruikt wordt om de massa van de stammen in de oogst te vergroten.

We weten uit het lokale model, sectie 2, dat de fotosynthese in één enkel blad afhangt van de leeftijd van het blad en daarom moeten we weten waar jongere en oudere bladeren gepositioneerd zijn op een rozenplant. Dit is de reden dat in ons model alle nieuwe bladeren aan de top van de rozenstruik ontstaan, wat relatief goed overeen komt met de realiteit. Ook nemen we aan dat de bladeren, en daarom het bladoppervlakte, uniform verdeeld zijn langs de stam. Met andere woorden, de oppervlakte van de bladeren aan een stam is recht evenredig met de lengte van de stam. De rozenkweker is uiteindelijk geïnteresseerd in de massa van de oogst en daarom veronderstellen we dat de massa van elke stam (inclusief de bladeren) recht evenredig is met zijn lengte.

3.1. Rozen zijn niet egoïstisch

Een verdere belangrijke, maar ook realistische, aanname die het model van de rozen in de kas vereenvoudigd is het zogenaamde **principe van niet egoïstisch zijn**. Dit principe zegt dat energie verkregen uit fotosynthese van een blad, aan een stam of in de struik, gelijkmatig bijdraagt aan de groei van alle rozenstammen of deze nou kort of lang zijn. Vrij vertaald "rozen zijn niet egoïstisch". Hieruit volgt dat, alhoewel een langere stam meer bladeren heeft en meer CO_2 zal assimileren dan een kortere stam, de totale geproduceerde energie gelijkelijk tussen hen verdeeld zal worden. Als resultaat hiervan groeit elke stam met dezelfde snelheid onafhankelijk van zijn eigen fotosynthese-snelheid.

Dit 'principe van niet egoïstisch zijn' komt naar voren in gegevens gemeten in kassen en ook in de observatie dat een enkele rozenplant met een aantal rozenstammen van verschillende lengtes zich gedraagt als een geheel. Op deze manier kunnen nieuwe jongere (kortere) stammen zich snel ontwikkelen ook al bezitten ze niet zo'n groot bladoppervlakte (en assimileren ze dus niet zoveel CO_2) als oudere (langere) stammen.

4. Het globale model van de kas

In deze sectie zullen we het totale model van de groei van rozen in een kas in een aantal stappen beschrijven. Uiteindelijk zal dit model in termen van het lokale model (1) worden gegeven.

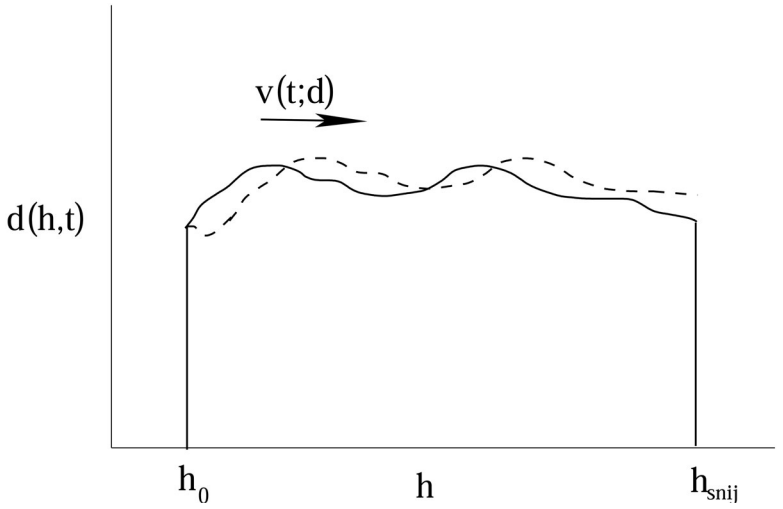
4.1. De groei-vergelijking

We beschrijven de toestand van de rozen op een gegeven tijdstip t door de **stamdichtheidsfunctie** $\mathbf{d}(\mathbf{h}, \mathbf{t})$ die de verdeling van stammen van verschillende hoogtes representeert. De functie $d(h,t)$ wordt gedefiniëerd als het *aantal* stammen met hoogte h per vierkante meter kas op tijdstip t. De dynamica van d kan verkregen worden met behulp van het principe van niet egoïstisch zijn wat geïntroduceerd is in de vorige sectie. Uit dit principe volgt namelijk dat elke rozenstam met dezelfde snelheid groeit. Daarom is er sprake van advectie van de dichtheidsfunctie $d(h,t)$ met een **groeisnelheid** $\mathbf{v} = \mathbf{v}(\mathbf{t}; \mathbf{d})$ die *onafhankelijk* is van h, zie figuur 2. De dynamica van d wordt beschreven door de volgende advectie-vergelijking

$$\frac{\partial d}{\partial t} + v\,\frac{\partial d}{\partial h} = 0. \tag{2}$$

Om deze vergelijking te vervolledigen moeten ook een begintoestand van de rozen en een randvoorwaarde gedefinieerd worden. De randvoorwaarde op $h = h_0$ die het ontstaan van nieuwe stammen uit de rozenstruik weergeeft, volgt opnieuw uit een van de aannames gedaan in de vorige sectie. We weten namelijk dat nieuwe stammen ontstaan op $h = h_0$ met een snelheid recht evenredig met de totale fotosynthese in de kas.

Met behulp van de functie d kan ook de **oogstsnelheid** $\mathbf{H}(\mathbf{t})$ per vierkante



Figuur 2. *Elke roos groeit met dezelfde snelheid. Hier is* $d(h,t)$ *het aantal stammen per vierkante meter kas als functie van de hoogte* h *en tijd* t*. Er vindt advectie van deze dichtheidsfunctie* d*plaats met een groeisnelheid* $v = v(t;d)$

meter kas worden bepaald. Namelijk, omdat rozen geoogst worden wanneer ze de hoogte h_{snij} hebben bereikt met een lengte $h_{snij} - h_0$ is de oogstsnelheid

$$H(t) \propto v(t;d)\,(h_{snij} - h_0)\,d(h_{snij}, t).$$

Van nu af aan betekent $\propto$ 'recht evenredig met'. Merk op dat de rozenkweker deze oogstsnelheid wil maximaliseren !

4.2. Bepalen van de groeisnelheid

De groeisnelheid v kan bepaald worden met behulp van een massa balans. Hiertoe bekijken we de **netto fotosynthese-snelheid** $\mathbf{P_{net}}$ die de uitstoot of opname van CO_2 per vierkant meter weergeeft in zowel de struik als de stammen. De netto fotosynthese-snelheid P_{net} verandert niet alleen door wijzigingen in het klimaat in de kas maar ook door de groei van de rozen (struik en oogst) én door het afsnijden van de rozenstammen. Specifieker gezegd is P_{net} recht evenredig met de verandering in de massa van de rozenoogst **plus** de oogstsnelheid $H(t)$. Als we de **massa van de rozenoogst** aangeven met $\mathbf{M(t)}$ volgt er dus dat

$$P_{net}(t;d) \propto \frac{dM}{dt} + H(t). \tag{3}$$

Hierbij hoeven we in $M(t)$ slechts de massa van de oogst mee te nemen omdat we aangenomen hebben, zie sectie 3, dat de geproduceerde energie alleen gebruikt wordt om de oogst te laten groeien en er geen energie naar de struik gaat.

De totale massa van de rozenoogst wordt beschreven door

$$M(t) \propto \int_{h_0}^{h_{snij}} (h - h_0)\, d(h,t)\, \mathrm{d}h. \tag{4}$$

Hier wordt, voor een stam met lengte h, de stamdichtheidsfunctie $d(h,t)$ gewogen wordt met de stamlengte $h - h_0$ voor alle hoogtes tussen h_0 en h_{snij}, om de massa te krijgen. Differentiëren van de uitdrukking (4) voor $M(t)$ geeft, met substitutie van de advectievergelijking (2) voor d, dat

$$\begin{aligned} \frac{dM}{dt} &\propto -v(t;d) \int_{h_0}^{h_{snij}} (h - h_0)\, \frac{\partial d}{\partial h}\, \mathrm{d}h \\ &= v(t;d) \int_{h_0}^{h_{snij}} d(h,t)\, \mathrm{d}h - H(t), \end{aligned} \tag{5}$$

na partieel integreren van de rechterkant van de vergelijking. Uiteindelijk geeft substitutie van (5) in de uitdrukking voor P_{net} (3) de groeisnelheid in termen van P_{net} als

$$v(t;d) \propto \frac{1}{\int_{h_0}^{h_{snij}} d(h,t)\, \mathrm{d}h}\, P_{net}(t;d). \tag{6}$$

4.3. De netto fotosynthese-snelheid
Onze volgende stap is om een uitdrukking voor de netto fotosynthese-snelheid P_{net} in (6) te bepalen. Aangezien dit de totale fotosynthese in de kas is, wordt deze bepaald door de fotosynthese van de oogst (de stammen) en de struik samen te nemen. Ofwel,

$$P_{net} = P_{oogst} + P_{struik},$$

waarin de bijdragen van de oogst en van de struik apart bepaald moeten worden.

Uit het lokale model, geïntroduceerd in sectie 2, volgt dat de fotosynthesesnelheid van een blad afhankelijk is van zowel zijn leeftijd als de hoeveelheid licht dat erop valt. Daarom is het heel erg belangrijk om de leeftijd- en hoogteverdeling van de bladeren goed te kunnen modelleren. Van nu af aan zullen we ons concentreren op het modelleren van deze dichtheden in de oogst, op de bepaling van de fotosynthese in de struik komen we later terug.

Om te beginnen definiëren we hiertoe een **bladdichtheidsfunctie** $\rho(\mathbf{h}, \mathbf{t})$ waarbij $\rho(h,t)\, \mathrm{d}h$ de *oppervlakte* van de bladeren geeft op hoogtes tussen h en $h + \mathrm{d}h$ per vierkante meter kas. Zolang $h_0 < h < h_{snij}$ is $\rho(h,t)$ gerelateerd aan de stamdichtheidsfunctie $d(h,t)$ doordat alleen rozenstammen met een totale hoogte groter dan h bijdragen aan de bladoppervlakte op hoogte h. Stammen

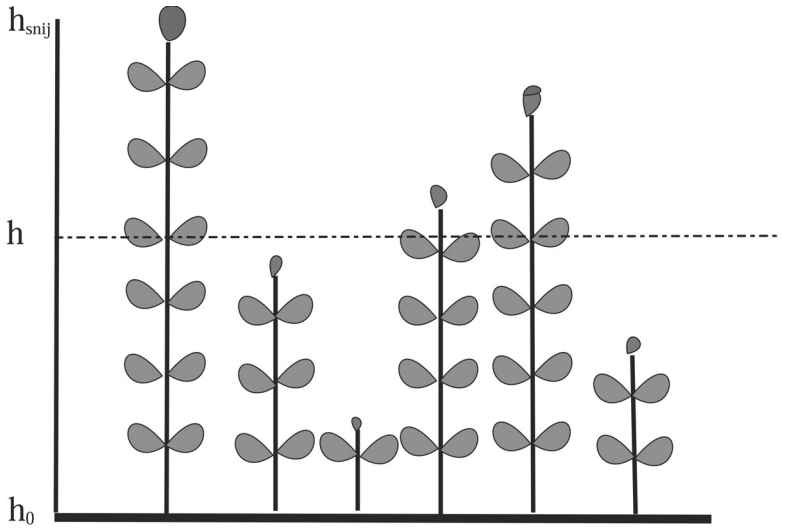
kleiner dan hoogte h dragen niet bij aan $\rho(h,t)$. Zie voor een schets hiervan figuur 3. Dit geeft dat

$$\rho(h,t) = k_\rho \int_h^{h_{snij}} d(\zeta,t)\,\mathrm{d}\zeta,$$

waar k_ρ een evenredigheids-constante is.

We moeten ook bepalen hoe de leeftijden verdeeld zijn op een bepaalde hoogte en daarom introduceren we de **leeftijddichtheidsfunctie $\mathbf{q(h,t)}$**. Deze is zo gedefinieerd dat $q(t,a,h)\,\mathrm{d}h\,\mathrm{d}a$ de bladoppervlakte van leeftijden tussen a en $a+\mathrm{d}a$ op hoogtes tussen h en $h+\mathrm{d}h$ per vierkante meter kas is. We hebben aangenomen dat de jongste bladeren bovenaan de stam zitten en de oudste onderaan, daarom is het duidelijk dat de leeftijd van een blad gerelateerd is aan zijn afstand tot de top van de stam. Dit is geschetst in figuur 4 waar we de rozenstammen geordend hebben op lengte. Om de relatie te versimpelen nemen we aan dat de leeftijd van een blad recht evenredig is met zijn afstand tot de top van de stam. Hieruit volgt dat de hoogte van de stammen waaraan bladeren op hoogte h met leeftijd a zich bevinden, $h+\frac{a}{k}$ is, waarbij k de gemiddelde inverse groeisnelheid is (zie figuur 4). Uiteindelijk kunnen we dus afleiden dat

$$q(t,a,h) \propto d\left(h+\frac{a}{k},t\right).$$



Figuur 3. *Een schets van de kas op een bepaald tijdstip. Op een bepaalde hoogte h dragen alleen de rozenstammen met hoogtes groter dan h bij aan de bladoppervlakte en dus aan de bladdichtheidsfunctie $\rho(h,t)$. Kleinere stammen dragen niet bij*

Vervolgens moet de hoeveelheid licht wat een blad bereikt nog bepaald worden. Deze lokale licht intensiteit op een blad hangt natuurlijk af van hoeveel het blad in de schaduw ligt, met andere woorden, van de blad bedekking (bladoppervlakte) boven het blad. De observatie dat alle stamhoogtes gelijkmatig door de kas verdeeld zijn suggereert dat bladeren op dezelfde hoogte in ongeveer dezelfde hoeveelheid schaduw liggen. Daarom is de verandering in de lichtintensiteit $\frac{\mathrm{d}I}{\mathrm{d}h}$ voor hoogtes tussen h_0 en h_{snij} ook een functie van h. We nemen $\frac{\mathrm{d}I}{\mathrm{d}h}$ recht evenredig met $\rho(h,t)$ en met de **lichtintensiteit op hoogte h, I(h, t),** met evenredigheids-constante k_I:

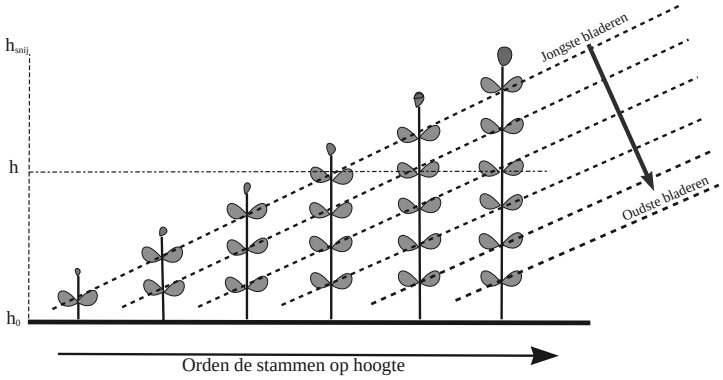
$$\frac{dI(h,t)}{dh} = k_I\,\rho(h,t)\,I(h,t), \quad I(h_{snij}) = I_0(t).$$

Merk op dat $I_0(t)$ gedefiniëerd was als de totale hoeveelheid lichtintensiteit die in de kas binnenkomt. Ofwel, $I_0(t)$ is de lichtintensiteit die de bovenkant van de rozen bereikt. Bovenstaande vergelijking oplossen levert

$$I(h,t) = I_0(t)\,e^{-k_I \int_h^{h_{snij}} \rho(\zeta,t)\,\mathrm{d}\zeta}.$$

Met behulp van alle bovenstaande dichtheidsfuncties kunnen we de totale fotosynthese van de oogst per vierkante meter kas bepalen. Deze volgt uit de lokale fotosynthese-snelheid $P(t,a,h)$ van één blad met leeftijd a waarop een bepaalde lichtintensiteit I valt. De fotosynthese-snelheid van bladeren met leeftijden tussen a en $a + \mathrm{d}a$ op hoogtes tussen h en $h + \mathrm{d}h$ kan namelijk bepaald worden door $P(t,a,h)$ te wegen met de leeftijddichtheidsfunctie $q(t,a,h)$ en wordt gegeven door

$$q(t,a,h)\,P(t,a,h)\,\mathrm{d}a\,\mathrm{d}h.$$



***Figuur 4**. Door de stammen op hoogte te ordenen wordt het eenvoudiger om te zien wat de leeftijdsverdeling van bladeren op hoogte h is. De schets geeft de situatie weer waar de groeisnelheid benaderd is door een constante om de leeftijddichtheidsverdeling te vereenvoudigen*

Dit integreren over alle leeftijden en hoogtes in de oogst levert uiteindelijk de fotosynthese-snelheid in de oogst:

$$P_{oogst}(t;d) = \int_{h_0}^{h_{snij}} \int_{0}^{T_{max}} q(t,a,h)\, P(t,a,h)\, \mathrm{d}a\, \mathrm{d}h.$$

Hier is T_{max} de leeftijd van het oudste blad in de huidige rozenoogst; T_{max} is verschillend voor elk type rozen en hangt ook af van het seizoen.

4.4. De fotosynthese van de struik

Om de totale netto fotosynthese-snelheid P_{net} in (6) te bepalen moeten we ook een uitdrukking voor P_{struik} hebben. Deze kan op een soortgelijke manier verkregen worden als waarop P_{oogst} bepaald is, maar hiervoor is dan wel enige kennis nodig van de bladverdeling in de struik. Zoals aangenomen, wordt dit deel van de planten niet afgesneden. Wat we eigenlijk nodig hebben zijn uitdrukkingen voor de bladdichtheidsfunctie $\rho(h,t)$ en de leeftijddichtheidsfunctie $q(t,a,h)$ in de struik. Ook moeten we de positie van de struik ten opzichte van de stammen weten. Of de struik bijvoorbeeld direct onder de stammen of ook gedeeltelijk naar de zijkanten overhangt is namelijk van invloed op de hoeveelheid licht dat de bladeren in de struik bereikt.

Relatief eenvoudige modellen voor de struik ontstaan door aan te nemen dat de struik direct onder de stammen ligt en daarom goed in de schaduw. Een simpel maar bruikbaar model wordt verkregen als we daarnaast ook nog aannemen dat bladeren van verschillende leeftijden uniform verdeeld zijn door de struik tussen $h = 0$ en $h = h_0$. Dit impliceert dat de gemiddelde hoogte in de struik $\frac{h_0}{2}$ en de gemiddelde leeftijd $\frac{\tau}{2}$ is, waarbij τ de lengte is van het groeiseizoen. Deze aannames leiden tot

$$P_{struik} \propto P(t, \frac{\tau}{2}, \frac{h_0}{2}).$$

Het is ook mogelijk om andere modellen voor de struik te combineren met ons model voor de kas. Dit is belangrijk voor de kweker omdat hiermee de vragen hoe en waar de struik moet groeien en hoe deze onderhouden moet worden voor een maximale rozenoogst, hopelijk beantwoord kunnen worden.

5. Schatten van de evenredigheidsconstanten

Tot nu toe hebben we een model beschreven voor de rozenproductie in een kas. In dit model komen een aantal evenredigheidsconstanten voor die nog onbekend zijn. De meeste van deze constanten kunnen bepaald worden door de rozenkweker door middel van metingen aan de rozenplanten. Dit zijn, bijvoorbeeld, de constante die de oppervlakte van de bladeren aan een stam per eenheid lengte beschrijft en de constante die de massa van de stam per eenheid lengte representeert. Er zijn twee evenredigheidsconstanten die niet door directe metingen aan de planten bepaald kunnen worden. Deze overgebleven

constanten worden verkregen door het model te fitten aan data uit de kas. De data komt uit bestaande rozenkassen en geeft de massa van de rozenproductie per week met gemeten klimaatcondities weer. Voor meer details hoe dit in zijn werk gaat, zie [1]. Deze schatting van de constanten uit de meetgegevens is nog niet voltooid.

Na bepaling van alle evenredigheidsconstanten zou het in principe met behulp van het model mogelijk moeten zijn om de rozenkweker te helpen om de rozenproductie te maximaliseren. Dit kan door in de simulatie van het model de rozenproductie te optimaliseren afhankelijk van de klimaatcondities in de kas.

6. De toekomst

Natuurlijk zijn sommige van onze aannames om tot het model te komen een vereenvoudiging van de werkelijkheid. Een van de nadelen van onze aanpak lijkt het feit dat we veronderstellen dat de totale energie die voorkomt uit de fotosynthese van alle rozenplanten alleen gebruikt wordt voor de groei van de stammen. Dit is niet erg realistisch aangezien er bijvoorbeeld ook seizoensverschillen in de dikte van de geoogste rozenstammen voorkomen. In het bijzonder is de energie die nodig is voor het bloeien, wat een cruciaal punt is voor het tijdstip waarop een roos geoogst moet worden, niet bekeken. Ook is de energie die de planten gebruiken voor onderhoud en opslag niet meegenomen in het model. Om de rozen realistischer te kunnen modelleren hebben we gedetailleerdere gegevens nodig over hoe de totale fotosynthese verdeeld wordt over het groeien, de productie van meer bladeren, het ontstaan van nieuwe stammen (waarvoor we hier slechts een basis model gebruiken) en het bloeien. De eerste modellen van deze processen in een enkele rozenplant zijn aan het verschijnen in de literatuur en zouden verder onderzocht moeten worden.

Na een bezoek aan een rozenkas weten we ook dat onze aanname over de positie van de rozenstruik ten opzichte van de stammen niet altijd overeenkomt met de werkelijkheid. Dit is te zien op de volgende foto van de kas.
In deze kas wordt een deel van de struik naast de planten gebogen waardoor dit aanzienlijk meer licht ontvangt dan wanneer de struik onder de stammen zou zitten. Door een ander model voor de struik te ontwerpen zou ook dit geïmplementeerd kunnen worden in ons model voor de gehele kas.

Ondanks deze (en andere) tekortkomingen hopen we dat onze aanpak via stam-, blad- en leeftijddichtheidsfuncties flexibel genoeg zal blijken om te koppelen aan complexere en preciezere groei-modellen en dat dit zal leiden tot een accuraat, interactief model van een gehele rozenkas. Een voordeel van ons model is dat het vergeleken kan worden met echte data gemeten in kassen, dit is iets wat niet gedaan kon worden met het enkele blad model (het lokale model). Verdere verbeteringen aan het model en het schatten van de constanten is een interessante uitdaging voor een vervolgonderzoek in het optimaliseren van de rozenproductie.

Dank aan Onno Bokhove (Twente), Johan Dubbeldam (Eindhoven), Philipp Getto (Utrecht), Bas van 't Hof (Vortech Computing), Nick Ovenden (Eindhoven), Derk Pik (Leiden) and Georg Prokert (Eindhoven) in samenwerking met wie dit werk tot stand is gekomen. Ook dank aan Dick van der Sar van Phytocare, hij was degene die het probleem heeft voorgesteld.

Literatuur

1. Onno Bokhove, Johan Dubbeldam, Philipp Getto, Bas van 't Hof , Nick Ovenden, Derk Pik, Georg Prokert, Vivi Rottschäfer, Dick van der Sar, Roses are unselfish: a greenhouse growth model to predict harvest rates, *Proceedings 42nd Eur. Study Group Ind.* pp. 59–76, 2002.
2. P.C. Harley, R.B. Thomas, J.F. Reynolds, & B.R. Strain, *Plant, Cell, and Environment* **15**, 271–282, 1992.
3. S.-H. Kim & J.H. Lieth, *Proc. III Rose Research*, 111–119, 2001.

CWI Vakantiecursus 2003

pp. 27 – 40

Politieke Macht en Onmacht

Rob Bosch
Koninklijke Militaire Academie, Breda
e-mail: r.bosch2@mindef.nl

1. Inleiding

In de politiek gaat het om *macht*. Politieke macht kan worden uitgeoefend op basis van bijvoorbeeld het aantal zetels in een parlement, het aantal stemmen in een raad met een gewogen stemsysteem of door het gebruiken van een vetorecht. Het ligt voor de hand te denken dat een groter aantal zetels of een groter aantal stemmen ook een toename van de politieke macht betekent. Dit hoeft echter niet altijd het geval te zijn. Ja sterker nog, een toename van het aantal zetels of stemmen kan zelfs gepaard gaan met een dalende politiek invloed.

In dit artikel zullen we de relatie tussen zetelaantal of stemmenaantal en de politieke macht die op basis van dit aantal kan worden uitgeoefend, bespreken. Hiertoe introduceren we voor gewogen stemsystemen de zogenaamde *Shapley-Shubik-index* die als een maat voor de politieke macht kan worden opgevat. Tevens zullen we laten zien dat een systeem waarbij één of meerdere partijen over een vetorecht beschikken altijd kan worden opgevat als een gewogen stemsysteem zonder vetorecht. We zullen het bovenstaande illustreren aan de hand van de situatie in de Europeese Unie en de Veiligheidsraad van de Verenigde Naties.

2. Een Unie van drie landen

In deze paragraaf bekijken we de machtstructuur binnen een unie van drie landen. In deze unie geven we de landen aan met A, B en C. In de Unieraad bestaande uit drie vertegenwoordigers van de aangesloten landen wordt over zaken van gemeenschappelijke belang gestemd. Op basis van het aantal inwoners krijgt land A in deze raad 6 stemmen toegewezen, voor land B zijn dat er 4 en het kleinste land C krijgt 1 stem. In het totaal worden er dus 11 stemmen uigebracht. In de raad worden de besluiten met meerderheid van stemmen genomen. Omdat het aantal uitgebrachte stemmen oneven is hoeft er geen extra regel te worden opgesteld voor het geval de stemmen staken. De vraag die we hier zullen beantwoorden is hoe op basis van deze stemmenverhouding de machtsverhoudingen liggen binnen de unie.

Een oppervlakkige beschouwing zou kunnen leiden tot het idee dat land A anderhalf keer zoveel invloed heeft als land B, immers A heeft anderhalf keer zoveel stemmen als B. Land B heeft vier keer zoveel stemmen als land C en zou derhalve vier keer zoveel invloed hebben als C. Dit is inderdaad een nogal oppervlakkige beschouwing want de besluiten worden immers met

meerderheid van stemmen genomen. Er zijn derhalve minstens 6 stemmen nodig om een voorstel aangenomen te krijgen en 6 stemmen zijn ook voldoende om een voorstel te blokkeren. Aangezien A over 6 stemmen beschikt kan hij ieder voorstel dat hem niet zint, blokkeren. De 6 stemmen van A zijn ook voldoende om een voorstel aan te nemen. De landen B en C hebben samen niet genoeg stemmen om een voorstel aan te nemen noch is hun gezamenlijk stemmenaantal voldoende om voorstellen te blokkeren. Land A bepaalt dus welke voorstellen wel of niet worden aangenomen. Met andere woorden A heeft met zijn absolute meerderheid in de raad alle macht.

Als we de macht van een land een een getal willen uitdrukken dan ligt het voor de hand land A 100% van de macht toe te kennen. Het is gebruikelijk om als maat voor de macht een index te kiezen, dwz. een getal tussen 0 en 1. Waarbij 0 aangeeft dat een partij geen enkele macht kan uitoefefen en een 1 betekent dat een partij alle macht heeft. Een dergelijke index wordt *machtsindex* (power index) genoemd. In het bovenstaande voorbeeld wordt de verdeling van de macht dan als volgt:

	A	B	C	totaal
aantal stemmen	6	4	1	11
machtsindex	1	0	0	1

Zolang land A 6 stemmen houdt, is de verdeling van de overblijvende 5 stemmen tussen B en C voor de verdeling van de macht niet interessant. Land A houdt alle macht en de machtsindices van B en C blijven 0. Een land met een machtsindex van 0 wordt ook wel een *dummy* genoemd.

Het is duidelijk dat de stemmenverdeling $(6, 4, 1)$ voor de landen B en C niet aanvaardbaar is. Het is in dit geval dus onwenselijk de stemmen te vertellen naar evenredigheid van de inwoneraantallen. In de volgende paragraaf bekijken we daarom andere stemmenverhoudingen.

3. Verschuiving van de macht

De landen van de Unie besluiten nu tot de volgende verdeling van de stemmen: land A en land B beide 5 stemmen en land C 1 stem. Bij deze verdeling wordt nog enigszins rekening gehouden met de grootte van de lidstaten. Hoe veranderen de machtsindices van de landen in dit geval? Het is duidelijk dat land A die zijn absolute meerderheid kwijt is, macht zal inleveren. Land B zal met de extra stem aan invloed winnen en dezelfde invloed krijgen als land A. Voor land C dat nog steeds maar 1 stem heeft, verandert er ogenschijnlijk weinig of niets.

Nu geen enkel land meer een absolute meerderheid heeft, kan een voorstel slechts worden aangenomen of verworpen door een coalitie van tenminste twee landen. We merken op dat elk tweetal landen over een meerderheid van minstens 6 stemmen beschikt. Ieder tweetal landen kan derhalve een voorstel aannemen of de aanname ervan blokkeren. De meerderheidscoalities of winnende coalities zijn bij deze stemverdeling:

$$AB \qquad AC \qquad BC$$
$$ABC$$

Uit de bovenstaande lijst van winnende coaliteis blijkt dat ieder land dezelfde rol speelt. Op grond van deze symmetrie kennen we dan ook aan ieder land dezelfde macht toe. De machtsverdeling binnen de Unie wordt bij deze stemmenverhouding gegeven door

	A	B	C	totaal
aantal stemmen	5	5	1	11
machtsindex	1/3	1/3	1/3	1

Merk op dat alhoewel aan C geen extra stem is toegekend dit land toch aanzienlijk aan invloed gewonnen heeft. Als kleinste lidstaat met slechts 1 stem heeft het dezelfde invloed als de twee grote lidstaten! Een dergelijke verdeling zal waarschijnlijk op verzet stuiten van de grote landen. Immers, hoe maakt men het de eigen bevolking duidelijk dat de 12 miljoen inwoners van land A dezelfde invloed hebben als de 2 miljoen inwoners van land C? We proberen daarom nog enige andere stemmenverdelingen om zo tot een voor alle partijen aanvaardbare machtsverhouding te komen.

4. Een paradox

Om tot een aanvaardbare machtsverhouding te komen, beginnen we met de verdeling van de macht om er vervolgens de stemmenverhouding bij te vinden. Stel dat de landen een machtsverhouding van $m(A) = 1/2, m(B) = 1/3$ en $m(C) = 1/6$ allerzins redelijk vinden. Welke stemmenverhouding bewerkstelligt dit? Met andere woorden welke getallen moeten we in de onderstaande tabel invullen.

	A	B	C	totaal
aantal stemmen	?	?	?	11
machtsindex	1/2	1/3	1/6	1

Al snel blijkt dat het niet lukt de stemmen zo te verdelen dat de gevraagde machtsstructuur verkregen wordt. De reden hiervoor is de volgende. Bij 3 landen heeft of een land de absolute meerderheid of ieder tweetal landen vormt een winnende coaltie. In het eerste geval levert dat een machtsvector $[1, 0, 0]$ op en het tweede geval geeft weer de egalitaire verdeling $[1/3, 1/3.1/3]$. In de Unie moet men dus kiezen voor een dictatoriaal systeem waarbij een land het voor het zeggen heeft of voor een gelijkwaardigheid van de landen ongeacht de grootte van het land. Zolang men een oneven aantal stemmen verdeelt en bij meerderheid besluiten neemt geldt het bovengenoemde argument. Voor een dergelijke Unie geldt dan dus

> *In een Unie van 3 landen waarin bij meerderheid wordt beslist, kan een oneven aantal stemmen alleen zo worden verdeeld dat er of een dictatoriale machtsstructuur of een egalitaire machtsstructuur ontstaat.*

Aangezien de twee bovengenoemde machtsstructuren onbevredigend zijn, gaan we in de volgende paragraaf op zoek naar mogelijkheden om een structuur te vinden die meer recht doet aan de grootte van de lidstaten.

5. Het quotum

Daar het bij meerderheid besluiten nemen tot de twee onbevredigende structuren aanleiding geeft, verlaten we dit principe. Behalve een stemmenverdeling stellen we een aantal stemmen vast dat nodig is voor aaname van een voorstel. Dit aantal stemmen dat we aangeven met q heet het *quotum*. Uitgaande van de oorspronkelijke verdeling van $(6, 4, 1)$ kunnen we dit quotum op 6 t/m 11 stemmen vaststellen. Een quotum van minder dan 6 stemmen levert meerdere disjuncte winnende coalities op bv. A en BC hetgeen een tegenstrijdigheid in de besluitvorming geeft. We gaan nu de consequenties van de verschillende quota na.

1. Quotum $q = 11$.

In dit geval wordt er besloten bij unanimiteit. Elke lidstaat kan hier een voorstel blokkeren en alleen de *grand coalition* ABC is winnend. Dit geeft uiteraard weer de egalitaire machtsverdeling.

quotum=11

	A	B	C
aantal stemmen	6	4	1
machtsindex	1/3	1/3	1/3

2. Quotum $q = 8, 9, 10$

In deze gevallen hebben A en B het voor het zeggen. Behalve de grand coalition is alleen de coalitie AB winnend. Voor de aaname van een voorstel heeft A de steun van B nodig en andersom. De eventuele steun van C is hierbij niet van belang. Beide landen kunnen ook op eigen kracht een voorstel blokkeren. Ook in dit geval is het irrelevant wat C doet. De macht wordt hier dus verdeeld tussen A en B. Land C is bij deze quota een dummy.

quotum=8, 9, 10

	A	B	C
aantal stemmen	6	4	1
machtsindex	1/2	1/2	0

3. Quotum $q = 6$

Dit is het geval van stemmen bij meerderheid waarbij land A alle macht bezat.

quotum=6

	A	B	C
aantal stemmen	6	4	1
machtsindex	1	0	0

4. Quotum $q = 7$

Dit is het meest interessante geval. Land A kan met behulp van B een voorstel doen annemen. Dit is ook mogelijk als A zich verzekerd weet van de steun van C. A heeft dus de keuze uit twee mogelijke coalitiepartners. Aangezien B en C samen geen meerderheid hebben, is voor hen de enige mogelijkheid A als

coalitiepartner te strikken. De landen B en C zijn in deze situatie vergelijkbaar en hebben derhalve dezelfde macht. Op grond van het feit dat A een twee keer zo grote keuze heeft in coalitiepartners als B en C zou men misschien de machtsverhouding $[1/2, 1/4, 1/4]$ verwachten. De werkelijke machtsverhoudingen blijken echter als volgt te zijn.

quotum=7

	A	B	C
aantal stemmen	6	4	1
machtsindex	2/3	1/6	1/6

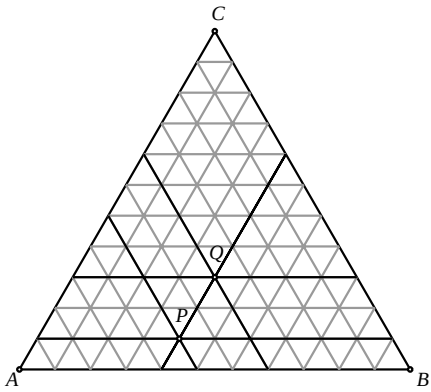
Land A blijkt hier dus over vier keer zoveel macht te beschikken als B en C. Vanwaar dit grote verschil? Wel, men kan twee soorten macht onderscheiden; ten eerste de macht om alleen of met anderen een voorstel aangenomen te krijgen en ten tweede de macht om een onwelgevallig voorstel te blokkeren. Alhoewel land A niet op eigen kracht een voorstel kan doen annnemen, beschikt het wel over de macht om voorstellen te blokkeren de zgn. *blocking power*. Geen enkel voorstel dat A niet zint, wordt aangenomen. Anders gezegd: land A heeft hier een vetorecht hetgeen A een sterke machtsbasis garandeerd. De landen B en C missen deze blocking power. In het vervolg zullen we op een formele wijze de bovenstaande machtsstructuur berekenen.

Bij de stemmenverdeling $(6, 4, 1)$ geldt voor $q = 6, 8, 9, 10$ dat een of meer landen een dummy zijn. Dit is zeker geen goede basis voor het vormen van een Unie. Blijft over de egalitaire machtsverdeling bij het besluiten bij meerderheid of unanimiteit en de verdeling $[2/3, 1/6, 1/6]$ bij $q = 7$. Geen van deze verdelingen doet volledig recht aan de grootte van de landen zodat we kunnen concluderen dat er geen machtsstructuur bestaat die een goede afspiegeling is van de relatieve grootte van de lidstaten.

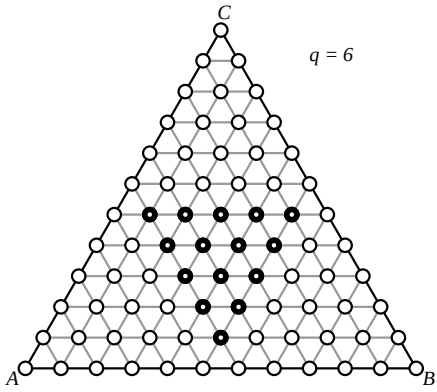
6. Machtsdriehoeken

De machtsstructuren die voortvloeien uit de stemmenverdelingen en het quotum kunnen overzichtelijk in een driehoek worden weergegeven. In de driehoek ABC in figuur 1 is de situatie weergegeven van een verdeling van 11 stemmen over drie landen A, B en C. Ieder roosterpunt in de driehoek heeft hier drie coördinaten. Zo zijn de coördinaten van punt $A = (11, 0, 0)$ hetgeen hoort bij een verdeling van 11 stemmen voor A en geen voor B en C. Bij de punten B en C horen respectievelijk de coördinaten $(0, 11, 0)$ en $(0, 0, 11)$. De punten P en Q hebben respectievelijk de coördinaten $(6, 4, 1)$ en $(4, 4, 3)$. De coördinaten van deze punten geven ook hier weer de stemmenverdeling aan. Iedere stemmenverdeling correspondeert met een roosterpunt in de driehoek en ieder roosterpunt in de driehoek geeft een mogelijke stemmenverdeling aan. Merk op dat op lijnen evenwijdig aan de basis AB de C-coördinaat constant is. Evenzo is op lijnen evenwijdig aan de lijn AC de B-coördinaat constant.

Bij ieder punt hoort ook een machtsstructuur. In het geval van stemmen bij meerderheid (quotum=6) krijgen we de machstructuur uit figuur 2. De witte punten zijn de punten waarbij de stemmenverdeling resulteert in de dictatoriale



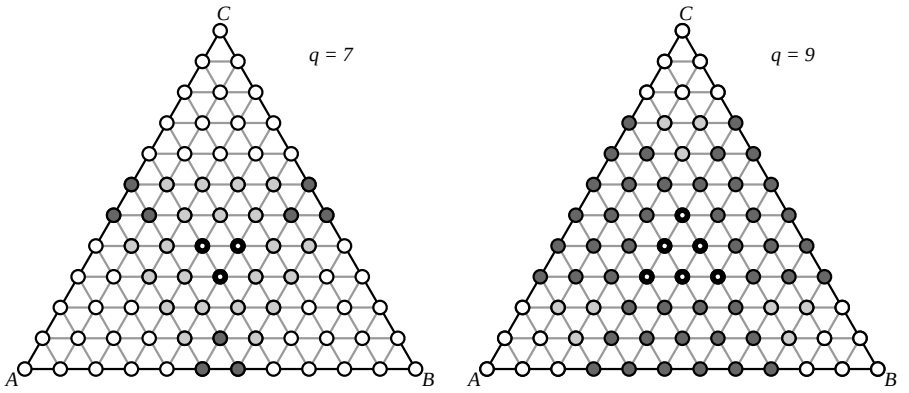
Figuur 1. $\triangle ABC$ *met* $P(6,4,1)$ *en* $Q(4,4,3)$



Figuur 2. *Machtsverdeling bij* $q = 6$

machtsstructuur $[1,0,0],[0,1,0]$ of $[0,0,1]$, terwijl de zwarte punten horen bij stemmenverdelingen die leiden tot de egalitaire machtstructuur $[1/3,1/3,1/3]$. Kortom, in deze driehoek geven punten van gelijke kleur eenzelfde machtsstructuur aan.

In figuur 3 zijn voor diverse quota de machtstructuren weergegeven. De witte punten geven weer een dictatoriale structuur aan, terwijl de zwarte punten horen bij een egalitaire machtsstructuur. De donkergrijze punten geven de machtsstructuur $[1/2,1/2,0],[1/2,0,1/2]$ of $[0,1/2,1/2,0]$ aan. Ten slotte hoort bij de lichtgrijze punten de structuur $[2/3,1/6,1/6],[1/6,2/3,1/6]$ of $[1/6,1/6,2/3]$.



Figuur 3. *Machtsverdeling bij* $q = 7$ *en* $q = 9$

7. Twee paradoxale voorbeelden

Bij een Unie van drie landen zijn de in de vorige paragraaf besproken vier machtsverdelingen de enig mogelijke verdelingen. De volgende twee voorbeelden laten zien dat de combinatie van de stemmenverdeling en het quotum tot paradoxale situaties kan leiden.

quotum=50

	A	B	C	totaal
aantal stemmen	49	49	1	99
machtsindex	1/3	1/3	1/3	1

quotum=16

	A	B	C	totaal
aantal stemmen	8	8	7	23
machtsindex	1/2	1/2	0	1

In het eerste voorbeeld hebben A en B bijna 50 keer zoveel stemmen als C. Bij een besluitvorming bij meerderheid hebben de drie parijen echter alle evenveel macht.

In het tweede voorbeeld hebben de drie landen alle bijna evenveel stemmen. Het quotum van 16 zorgt er echter voor dat C een dummy wordt.

Alvorens de situatie te bekijken voor 4 of meer landen besteden nog even aandacht aan de blocking power of vetomacht. Wellicht is het mogelijk andere machtstructuren te krijgen door een of meer landen een vetorecht te geven.

8. Vetomacht

Het toekennen van een vetorecht aan een of meer leden kan wellicht tot machtsstructuren aanleiding geven die bij een stemmenverdeling met quotum niet kunnen worden verkregen. Stel we geven de drie landen alle 5 stemmen en kennen

bovendien aan het grootste land A een vetorecht toe. Tot welke machtsstructuur leidt dit bij een quotum van 9?

quotum=9

	A	B	C	totaal
	veto			
aantal stemmen	5	5	5	15
machtsindex	?	?	?	1

Vanwege het vetorecht is voor de aanname van een voorstel altijd de steun van A nodig. De winnende coalities zijn hier derhalve AB, AC en ABC. Alhoewel B en C samen meer stemmen hebben dan het vereiste quotum is hun samenwerking vanwege het vetorecht van A niet voldoende om een voorstel te doen aannemen. Wel kunnen zij samen de aanname van een voorstel blokkeren. Bovendien heeft A de steun van minstens een van de andere partijen nodig. Deze situatie is vergelijkbaar met de volgende stemmenverdeling zonder vetorecht.

quotum=4

	A	B	C	totaal
aantal stemmen	3	1	1	5
machtsindex	2/3	1/6	1/6	1

Men gaat gemakkelijk na dat ook hier de winnende coalities AB, AC en ABC zijn. Bovendien zijn B en C samen in staat een voorstel te blokkeren en heeft A minstens een van de andere partijen nodig om een voorstel aangenomen te krijgen. Het systeem met vetorecht geeft dus dezelfde machtsstructuur als de bovenstaande verdeling. We vinden derhalve niets nieuws. De bovenstaande stemmenverdeling kunnen we op de volgende wijze vinden. Ken aan B en C 1 stem toe. Stel het quotum op q en geef A x stemmen. Aangezien de coalitie BC niet winnend is moet gelden

$$q > 2 \tag{1}$$

De coalities AB en AC zijn winnend en derhalve geldt

$$x + 1 \geq q \tag{2}$$

Tenslotte heeft A een van de partijen B of C nodig voor een winnende coalitie waaruit volgt

$$x < q \tag{3}$$

We zoeken nu gehele waarden van x en q die voldoen aan de bovestaande ongelijkheden. Een van de mogelijke oplossingen wordt gegeven door $x = 3$ en $q = 4$.

Op deze wijze kunnen we elk systeem met vetorecht vertalen naar een zogenaamd gewogen stemreglement. In de veiligheidsraad van de Verenigde Naties hanteert men een reglement met vetorecht. De Veiligheidsraad van de VN bestaat uit 15 leden. China, Engeland, Frankrijk, Rusland en de Verenigde Staten zijn de zogenaamde permanente leden van de raad. De andere 10 leden hebben

slechts gedurende een beperkte periode zitting in de raad. Voor de anname van een voorstel zijn 9 van de 15 stemmen nodig. De vijf permanente leden hebben echter een vetorecht. Een tegenstem van één van deze leden blokkeert de aanname van een voorstel. Om dit systeem te vertalen naar een gewogenstemreglement gaan we weer als volgt te werk. Ken aan de tien niet permanente leden een gewicht van 1 toe. Stel het gewicht van de permante leden op x en het quotum op q. De vijf permante leden tezamen met vier niet permante leden vormen een winnende coalitie. Er geldt dus

$$5x + 4 \geq q \tag{4}$$

Anderzijds is een coalitie met minder dan 4 niet permanente leden niet winnend dus geldt:

$$5x + 3 < q \tag{5}$$

Als één van de permante geen deel uitmaakt van de coalitie is deze niet winnend, zelfs niet als alle niet permante leden tot de coalitie behoren. Daaruit volgt dat

$$4x + 10 < q \tag{6}$$

Uit de twee ongelijkheden volgt

$$4x + 10 < q \leq 5x + 4 \tag{7}$$

We zoeken weer gehele waarden voor x en q die aan de ongelijkheden voldoen. Uit de laatste ongelijkheid leiden we af dat $x > 6$. Laten we $x = 7$ proberen. Uit $4x + 10 < q$ en $q \leq 5x + 4$ volgt dan dat $38 < q \leq 39$. We vinden voor de permanente leden een gewicht van 7 en een quotum van 39. Het is nu niet moeilijk meer om aan te tonen dat het reglement met vetorecht gelijkwaardig is met dit gewogen stemreglement. Immers een winnende coalitie omvat altijd de vijf permante leden en minstens vier niet permantente leden. Het gewicht van zo'n coalitie is minstens $5 \times 7 + 4 = 39$. Dus iedere winnende coalitie heeft een gewicht dat minstens gelijk is aan het quotum. In een verliezende coalitie ontbreekt minstens één permanent lid in welk geval het totale gewicht hoogstens $4 \times 7 + 10 = 38$ is, of er zijn tezamen met de vijf permanente leden hoogstens drie niet permanente leden aanwezig; in dit geval is het gewicht hoogstens $5 \times 7 + 3 = 38$. Iedere verliezende coalitie bereikt het quotum dus niet. Voor de machtsverdeling binnen de veiligheidsraad mogen we dus uitgaan van een stemmenverdeling $(7, 7, \ldots 7, 1, 1 \ldots 1)$ en een quotum van 39.

Iedere gewogen systeem waarbij bovendien een aantal leden een vetorecht hebben, kan worden vertaald naar een simpel gewogen systeem. Zo kunnen we een vergadering van n personen waarvan k personen een vetorecht hebben en waarbij p stemmen nodig zijn voor de aanname van een voorstel, opvatten als een gewogen systeem. Ken aan de leden met vetorecht het gewicht $n + 1$ en aan de andere leden een gewicht van 1 toe en stel het aantal stemmen nodig voor een geldige meerderheid op $kn + p$.

Een systeem met vetorecht geeft zoals we hebben gezien geen andere machtsstructuren dan gewogen stemreglementen.

9. De Shapley-Shubik-index

Bij een Unie van drie landen geven de diverse stemmenverdelingen aanleiding tot 4 verschillende machtsstructuren, waarvan slechts 2 zonder dummy's. Als we het aantal landen in de Unie uitbreiden, neemt ook het aantal mogelijke machtstructuren toe. De volgende voorbeelden geven een aantal mogelijke machtsstructuren.

quotum=9

	A	B	C	D	totaal
aantal stemmen	9	4	3	1	17
machtsindex	1	0	0	0	1

quotum=13

	A	B	C	D	totaal
aantal stemmen	7	6	3	1	17
machtsindex	1/2	1/2	0	0	1

quotum=10

	A	B	C	D	totaal
aantal stemmen	5	4	4	2	15
machtsindex	1/4	1/4	1/4	1/4	1

De bovenstaande voorbeelden leveren weinig problemen op, maar hoe bepalen we de machtsstructuur in het volgende voorbeeld?

quotum=11

	A	B	C	D	totaal
aantal stemmen	7	4	3	1	15
machtsindex	?	?	?	?	1

Om hier de machtsstructuur te vinden, introduceren we het begrip van de *spil* of *pivot*. Stel dat de coalitie ACB wordt gevormd in de aangegeven volgorde, d.w.z. C sluit zich aan bij A waarna B toetreedt tot de inmiddels gevormde coalitie AC. De aansluiting van C bij A geeft de coalitie AC welliswaar een sterkere positie maar winnend is ze nog niet, d.w.z. de coalitie heeft nog geen gekwalificeerde meerderheid. Pas na het toetreden van B wordt de coalitie winnend. In die zin is B de belangrijkste partij. Zo'n partij wordt de spil genoemd. Als we de coalitie vormen in de volgorde BCA dan wordt de coalitie pas winnend na de toetreding van partij A. In dit geval is partij A de spil. De macht van een partij wordt nu bepaald door de mogelijkheid om als belangrijke spilpartij op te treden. Een dummy is in dit verband een partij die nooit een spilfunctie heeft. Voor het berekenen van de machtsindex schrijven we alle permutaties van de partijen of landen op. Vervolgens gaan we na hoe vaak een partij als spil optreedt. De index wordt dan berekend door het aantal malen dat een partij de spil is te delen door het totaal aantal permutaties van de partijen. Dat we de berekening uitvoeren over alle mogelijke permutaties reflecteert het feit dat we alle mogelijke coalities even waarschijnlijk achten. De op de bovenstaande

wijze berekende index wordt de *Shapley-Shubik-index*[1] genoemd. De berekening van de index voor ons voorbeeld gaat als volgt.

A**B**CD	B**A**CD	CA**B**D	DA**B**C
A**B**DC	B**A**DC	CA**D**B	DA**C**B
AC**B**D	BC**A**D	CB**A**D	DB**A**C
AC**D**B	BC**D**A	CB**D**A	DB**C**A
AD**B**C	BD**A**C	CD**A**B	DC**A**B
AD**C**B	BD**C**A	CD**B**A	DC**B**A

Er zijn in het totaal 4! permutaties van de 4 landen. In de bovenstaande lijst hebben we ze alle 24 opgeschreven. Het vetgedrukte land is in de volgorde de spil. Zoals uit de bovenstaande tabel is af te lezen zijn A en B beide 8 keer de spil en C en D beide 4 keer. De Shapley-Shubik-indices zijn dus $m(A) = m(B) = 8/24$ en $m(C) = m(D) = 4/24$.

De lezer kan gemakkelijk nagaan dat de machtsindices die we tot nu toe tegengekomen zjn in overeenstemming zijn met de Shapley-Shubik-index.

10. De Europeese Gemeenschap

Bij de oprichting van de Europeese Unie in 1958 werd in het verdrag van Rome de stemmenverdeling binnen de unie vastgelegd. De drie grote lidstaten Frankrijk, Italië en Duitsland kregen alle 4 stemmen. België en Nederland kregen 2 stemmen en Luxemburg kreeg 1 stem. Het quotum werd vastgesteld op 12 van de 17 stemmen.

Europeese Unie 1958

Italië	4	Nederland	2
Frankrijk	4	België	2
Duitsland	4	Luxemburg	1

Hoe groot is de Shapley-Shubik-index van Nederland bij deze verdeling? In de volgordes waarin Nederland de spilfunctie vervult, moeten precies twee grote landen Nederland vooraf gaan. Dit geeft de volgende 2 permutaties: (442$\underline{2}$41) en 4421$\underline{2}$4. De eerste permutatie geeft $3!.\binom{3}{2}2! = 36$ verschillende volgordes. De tweede permutatie levert $4!\binom{3}{2} = 72$ verschillende volgordes op. Nederland is dus de spil in 108 van de 6!=720 volgordes. De Shapley-Shubik-index van Nederland is derhalve $108/720 = 3/20$. Uiteraard geldt hetzelfde voor België. Luxemburg kan in een volgorde alleen dan een spilfunctie vervullen als de aan Luxemburg voorafgaande landen tezamen precies 11 stemmen hebben. Daar alle landen behalve Luxemburg een even aantal stemmen hebben is dit onmogelijk. Met andere woorden, Luxemburg vervult nooit een spilfunctie en de Shapley-Shubik-index van Luxemburg is derhalve gelijk aan 0. De index voor de grote landen kan nu eenvoudig berekend worden als $(1-2\cdot 6/20)/3 = 14/60$. In overzicht van de machtstructuur vindt men in de volgende tabel.

[1] Martin Shubik en Lloyd Shapley zijn beide wiskundigen en economen met een grote verdienste op het gebied van de speltheorie.

Europeese Unie 1958

Land	Stemmen	Percentage van de stemmen	Index	Percentage van de macht
Italië	4	23.5	14/60	23.3
Frankrijk	4	23.5	14/60	23.3
Duitsland	4	23.5	14/60	23.3
Nederland	2	11.8	9/60	15.0
België	2	11.8	9/60	15.0
Luxemburg	1	5.9	0	0

Zoals uit be bovenstaande tabel blijkt, geeft de stemmenverdeling een redelijk beeld van de machtstructuur.

Bij de eerste uitbreiding van de Europeese Unie in 1973 met drie landen besloot men tot de volgende stemmenverdeling.

Europeese Unie 1973

Frankrijk	10	Nederland	5	Engeland	10
Italië	10	België	5	Denenmarken	3
Duitsland	10	Luxemburg	2	Ierland	3

Het quotum werd vastgesteld op 41 van de 58 stemmen. We merken op dat het aantal stemmen van de oorspronkelijke leden van de Unie met $2\frac{1}{2}$ vermenigvuldigd zijn behalve Luxenburg dat slechts 2 keer zoveel stemmen kreeg toebedeeld. De onderlinge verhoudingen bleven dus gelijk met uitzondering van Luxemburg dat er wat op achteruitgaat. In 1973 werd het quotum gesteld op 70.7% van de stemmen hetgeen ongeveer gelijk is aan het qoutum van 1958 dat uitkomt op 70.6% van de stemmen. Men mag verwachten dat door de toetreding van drie landen de macht van de oorspronkelijke leden enigszins verwatert. De berekening van de diverse indices is hier een tijdrovende bezigheid vandaar dat de volgende tabel met behulp van een computer tot stand gekomen is[2].

Europeese Unie 1973

Land	Stemmen	Percentage van de stemmen	Index
Frankrijk	10	17.9	0.178571
Duitsland	10	17.9	0.178571
Italië	10	17.9	0.178571
Engeland	10	17.2	0.178571
Nederland	5	8.6	0.080952
België	5	8.6	0.080952
Denenmarken	3	5.2	0.057143
Ierland	3	5.2	0.057143
Luxemburg	2	3.4	0.009524

[2] Een programma voor het berekenen van diverse powerindices kan gevonden worden op `www.uni-konstanz.de/FuF/Verwiss/koenig/`

In de tabel valt op dat de macht van Luxemburg, zoals gemeten door de Shapley-Shubik-index, is toegenomen. Dit komt omdat er in de nieuwe situatie minstens één volgorde van de landen is waarvoor Luxemburg de spil is. De macht van Luxemburg is toegenomen ondanks het feit dat Luxemburg er qua stemmenaantal relatief op achteruitgegaan is. Dit fenomeen staat in de literatuur bekend als de *"nieuwe-ledenparadox"*.

11. Machtsblokken

Met de uitbreidingen van de EU verwatert de macht van de individuele lidstaten. Hierdoor wordt het met name voor de kleine lidstaten aantrekkelijk om als één blok te opereren. Een dergelijke blokvorming heeft uiteraard gevolgen voor de machtsposities binnen de EU. We illustreren dit aan de hand van de situatie in 1973. Stel dat de Beneluxlanden een onderlinge afspraak maken om in de vergaderingen van de Unie als één blok te stemmen. Hoe verandert hierdoor de machtspositie van de drie Beneluxlanden? De situatie zou in dit geval als volgt worden

Europeese Unie 1973

Land	Stemmen	Percentage van de stemmen	Index	Percentage van de macht
Benelux	12	20.7	88/420	21.0
Italié	10	17.2	81/420	19.3
Frankrijk	10	17.2	81/420	19.3
Duitsland	10	17.2	81/420	19.3
Engeland	10	17.2	81/420	19.0
Denemarken	3	5.2	4/420	1.0
Ierland	3	5.2	4/420	1.0

De bundeling van de krachten van de Beneluxlanden leidt tot een gezamenlijke macht van 21% terwijl de landen tezamen een macht van 17% hadden. Zoals uit de tabel is af te leiden, gaat de samenwerking tussen de Beneluxlanden vooral ten koste van Denemarken en Ierland.

12. De Veiligheidsraad

Bij de oprichting van de Verenigde Naties en de instelling van de Veiligheidsraad in 1945 kwam men de volgende verdeling overeen. De raad zou bestaan uit vijf permanente leden met een vetorecht en zes niet permanente leden. Voor de aanname van een motie waren de vijf permanente leden en minstens twee niet permanente leden nodig. Een niet permanent lid kan derhalve alleen dan als spil optreden als het wordt vooraf gegaan door de vijf permanente leden en één niet permanent lid. Een niet permanent is dus de spil in $\binom{5}{1} \cdot 6! \cdot 4! = 86400$ permutaties. Daar er 11! permutaties van de landen zijn, is de machtsindex van een niet permanent lid gelijk aan $(5 \cdot 6! \cdot 4!)/(11!) = 0.0022$. De totale macht van de niet permanente leden was in 1945 dus gelijk aan $6 \cdot 0.0022 = 0.0132$ terwijl de gezamenlijke macht van de permanente leden gelijk was aan 0.9868.

Veiligheidsraad 1945		
Leden	Permanent lid	Tijdelijke leden
index	0.1974	0.0022

Om aan deze wel erg scheve verhouding een einde te maken, besloot de raad in 1965 op voorstel van een hervormingscommissie meer invloed toe te kennen aan de tijdelijke leden. Hiertoe werd het aantal niet permante leden uitgebreid met vier leden. Voor de aanname van een motie waren nu behalve de vijf permanente leden minstens vier niet permanente leden nodig. De nieuwe samenstelling levert de volgende machtsverhoudingen op.

Veiligheidsraad 1965		
Leden	Permanent lid	Tijdelijke leden
index	0.196	0.0018

De totale macht van de niet permanente leden is van 0.0132 gestegen tot $10 \cdot 0.0018 = 0.018$. De leden van de Algemene Vergadering waren zeer ingenomen met het voorstel van de hervormingscommissie!

13. Slotopmerkingen

In dit artikel hebben we een vaak gebruikte machtsindex besproken. Deze index kwantificeert op een formele wijze de macht die een partij kan uitoefenen. In de literatuur komen behalve de Shapley-Shubik-index nog enkele andere indices voor. Bekende indices zijn onder andere de Banzhaf-index, de Johnston-index en de Deegan-Packel-index. Welke van deze indices het beste de werkelijke macht van een partij refecteert is niet zo eenvoudig te zeggen. Bij het onderzoek met betrekking tot machtsindices stelt men vaak vooraf een aantal eisen of axioma's op waaraan een goede index moet voldoen. Vervolgens gaan we na welke index aan de axioma's voldoet. Het blijkt echter dat zelfs een gering aantal voor de hand liggende axioma's tot een onmogelijkheid leidt. Een situatie die vergelijkbaar is met de onmogelijkheidsstelling van Arrow uit 1951.

Literatuur

1. R.Bosch, *Gewogen stemsystemen*, Euclides 73-6, 1997/98.
2. P.C. Ordeshook, *Game Theory and Political Theory*, Cambridge University Press, 1986.
3. A. Rapoport, *N-person Game Theory*, Dover Publications Inc, 2001.
4. D.G. Saari, *Chaotic Elections*, American Mathematical Society, 2001.

CWI Vakantiecursus 2003

pp. 41 – 57

Bilevel optimization: anticipatory dynamic traffic management

Henk van Zuylen
Henk Taale
Delft University of Technology
Transportation Planning and Traffic Engineering section
Netherlands Ministry of Transport
Transport Research Center (AVV)
e-mail: H.J. van Zuylen@tudelft.nl

1. Introduction

If two persons play a game with the intention to win, they choose actions that maximize their chances to win. It is not simple to do so, because if the players take an action in turn and if the action of player 1 has an influence on the chances of player 2 to win, and vice versa, the players have to anticipate the possible actions of their partner. If we assume that player 1 and 2 both try to maximize their chances to win, and one player (the leader) knows how the other player (the follower) will respond to any decision he may make, we have a game that is known as a *Stackelberg game*. If, on the contrary, we assume that the two players do not know each others' strategy, they will each try to optimize their own objective function. The equilibrium that can be achieved int hat situation is that no player can improve his objective function by changing his own decision without co-operation of the other player. This equilibrium is called the Nash equilibrium. The term 'Nash equilibrium' is used for systems with many non-co-operating players. If there are only two players, the term Cournot equilibrium is used.

Mathematically the situation is that players 1 and 2 try to **minimize** the objective function $P_i(x_1, x_2)$, where x_i is the vector of decision that can be taken by player i. In the Stackelberg game, the reaction of player 2 is determined by the action of player 1:

$$x_2 = T(x_1). \tag{1}$$

Then, at the optimum choice made by player 2 to optimize his objective function, player 1 chooses x_1 such that

$$P_1(x_1, T(x_1)) \geq P_1(x_1^*, T(x_1^*)), \forall x_1 \in A_1, \tag{2}$$

where x_1^* is the optimum solution for player 1, A_1 is the set of possible actions for player 1 and the action $T(x_1)(T(x_1 \in A_2))$ is given by

$$P_2(x_1, T(x_1)) \leq P_2(x_1, x_2), \forall x_2 \in A_2, \tag{3}$$

i.e. $T(x_1)$ is the best possible action for player 2 within the set of possible actions A_2. In the Nash or Cournot equilibrium, the situation (x_1^*, x_2^*) is defined by

$$P_1(x_1, x_2^*) \geq P_1(x_1^*, x_2^*), \forall x_1 \in A_1 \text{ and} \tag{4}$$
$$P_2(x_1^*, x_2) \geq P_2(x_1^*, x_2^*), \forall x_2 \in A_2.$$

As an example we take the following objective functions:

$$P_1 = x_1^2 - x_1 x_2 + 2x_2^2 + x_1, \tag{5}$$

$$P_2 = 2x_1^2 - x_1 x_2 + x_2^2. \tag{6}$$

The players will optimize their own profit, i.e. player 1 will choose

$$\frac{\partial P_1}{\partial x_1} = 0 = 2x_1 - x_2 + 1 \tag{7}$$

and player 2 will choose

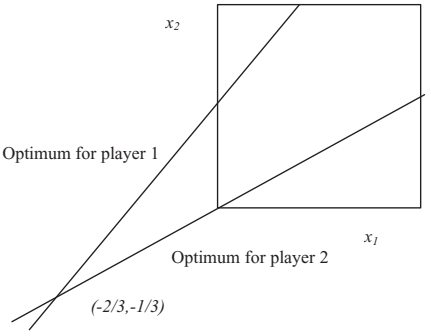
$$\frac{\partial P_2}{\partial x_2} = 0 = -x_1 + 2x_2. \tag{8}$$

The situation in which both players realized their own optimum – without knowledge of the objectives and strategy of the other player – is

$$x_1^* = -2/3 \text{ and } x_2^* = -1/3, \tag{9}$$

with objective functions

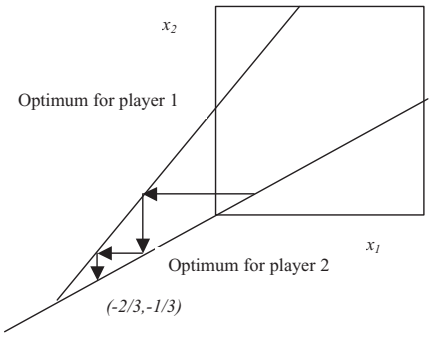
$$P_1 = -2/9 \text{ and } P_2 = 7/9. \tag{10}$$



Under certain circumstances, the Nash equilibrium can be obtained by iteration:

$$\text{Min } P_1(x_1, x_2^{k-1}) \text{ with respect to } x_1 \text{ and}$$
$$\text{Min } P_2(x_1^{k-1}, x_2) \text{ with respect to } x_2.$$

The iteration process is given in the following figure



1.1. Non converging iterations

It is clear that there are situations that iterations do not converge to a Nash equilibrium. For instance, suppose that the set of optimum solutions are interchanged:

$$\frac{\partial P_2}{\partial x_2} = 0 = 2x_1 - x_2 + 1 \text{ and } \frac{\partial P_1}{\partial x_1} = 0 = -x_1 + 2x_2$$

The Nash equilibrium still is the same, but iterations will not converge.

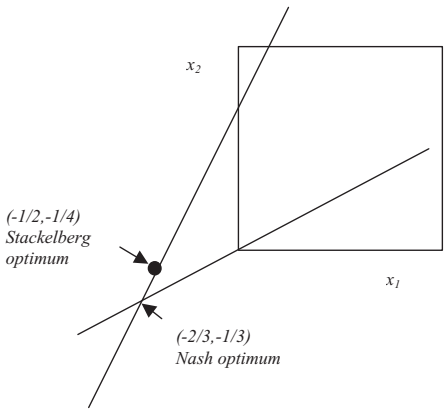
The *Stackelberg equilibrium* is different. No iteration is used, but player 2 anticipates the reaction x_2 on his action x_1 and with this knowledge he chooses an optimum action:

$$x_2 = T(x_1) = \frac{1}{2}x_1. \tag{11}$$

With this knowledge, the objective function for player 1 becomes

$$\begin{aligned} P_1 = x_1^2 \quad & -x_1x_2 + 2x_2^2 + x_1 \\ & = x_1^2 - \tfrac{1}{2}x_1^2 + 2.(\tfrac{1}{2}xx_1)^2 + x_1 = x_1^2 + x_1 \end{aligned} \tag{12}$$

with a minimum for $x_1 = -\frac{1}{2}$ and $P_1 = -\frac{1}{4}$. The optimum is shown in the figure below.



First of all, it is clear that the Stackelberg optimum gives a better result for player 1 ($P_{1,Stackelberg} = -\frac{1}{4}, P_{1,Nash} = -2/9$) and also a better result for player 2 ($P_2 = 7/16$ instead of $7/9$). Second, if player 1 tries to improve his objective function, it is possible to get a further reduction, but the reaction of player 2 eliminates this reduction again. Thus a further iteration after the realization of the Stackelberg equilibrium does not improve the outcomes, even though this may look advantageous for player 1. Third, the game is not symmetrical. If we assume that player 2 can predict the response of player 1, the optimisation problem becomes

$$\text{Min } P_2 = 2x_1^2 - x_1x_2 + x_2^2$$
$$\text{Under the constraint } x_1 = \tfrac{1}{2}(x_2 - 1)$$

This gives the objective function

$$P_2 = x_2^2 - \frac{1}{2}x_2 + \frac{1}{2}$$

With the solution

$$x_2 = \frac{1}{4} \text{ and } x_1 = -3/8$$

and

$$P_2 = 7/16 \text{ and } P_1 = -13/64$$

which gives a worse result for player 1, as could be expected, than in the case that player 1 has the leading role.

2. The one-level optimization: traffic control

2.1. The delay as objective function

In Dynamic Traffic Management it is not common to optimize the measures with respect to a certain goal. If one wants to do so, mathematical procedure is needed. It is necessary to have a goal that is expressed in quantitative terms, e.g. total delay, total travel time, queue length, average speed etc. This qualitative entity is called the *objective function*. The second requirement is, that a unique relation exists between certain control parameters and the value of the objective function. It should be possible to determine the effect of a change of a parameter on the value of the objective function.. The third requirement is, that there is a methodology to search in a systematic way in the space of feasible control parameter settings for an optimum value of the objective function.

For example, if we want to optimize traffic control on an intersection, the control parameters are – for fixed time control – the cycle time, the sequence of the signal groups, the combination of the signal groups in green combinations and the greentimes. The boundary conditions are, e.g. the minimum green time, the conflicts, the clearance times between conflicting signal groups etc. As objective function a weighted sum of stops and delays summed for all road users in a certain time interval. The relationship between the total delay W

for a signal group per cycle and the relevant parameters, $t_{r,eff}$, the effective red time and C, the cycle time is given by Webster as

$$W = \frac{1}{2}Vt_{r,eff}^2/(1-y) + \frac{1}{2}x^2C/(1-x) - 0.65 * V^{1/3}C^{4/3}x^{2+5u}, \qquad (13)$$

with

- V = flow (pcu per second)
- $y = V/s$, the load ratio with s = the saturation flow in pcu/s
- x = degree of saturation, the ratio between the number of vehicles arriving per cycle ($V.C$) and the number that can depart ($\{C - t_{r,eff}\}.s$)
- u = ratio between green time and cycle time ($(1 - t_{r,eff})/C$)

Often, formula (13) is used in its simplified form

$$W = 0.9\{\frac{1}{2}Vt_{r,eff}^2/(1-y) + \frac{1}{2}x^2C/(1-x)\}. \qquad (14)$$

A systematic methodology to search for the best cycle time, composition of green phases and duration of red or green phases, does not exist. Several methods are used, but none of them can be proven to find the real optimum. If the problem is simplified by choosing a-priori the green combinations, it is possible to search in the space of cycle times and green splits for an optimal solution.

This optimization problem is already rather complicated, due to the degrees of freedom (a mixture of discrete and continuous parameters) and a rather complicated function that describes the relationship between the control parameters and W. Furthermore, the expression (1) is only valid for uniform arrivals. In networks with rather short distances between the different controlled intersections, the traffic arrives in platoons giving a different expression for the delay. The best approach in that case is followed in the TRANSYT program. TRANSYT calculates the delay from the arrival and departure pattern at the intersection. The cycle time is subdivided in intervals (e.g. intervals of 1 or 2 seconds). For each time interval the number of arrivals are calculated from the departures from other intersections in the network. From the arrival pattern and the greentimes the departure pattern is calculated and the difference between arrivals and departures gives the queue length per time interval, which is the basis of the delay calculation (Figure 1).

It is obvious that there is no longer a closed mathematical form that gives the relationship between the control parameters and the delay. The optimization procedure applied in TRANSYT is a hill-climbing method, where the parameters are changed and after each change the delay is recalculated.
If we want to optimize traffic actuated control continuously, it can be done by simulation. In most cases micro-simulation is needed: individual vehicles are simulated as they depart from an intersection, pass detectors, arrive at the back of the queue etc. Also the traffic control has to be simulated, given the simulated detector signals, the control algorithm and certain control parameters. If we want to optimize the parameters of vehicle actuated controllers, we

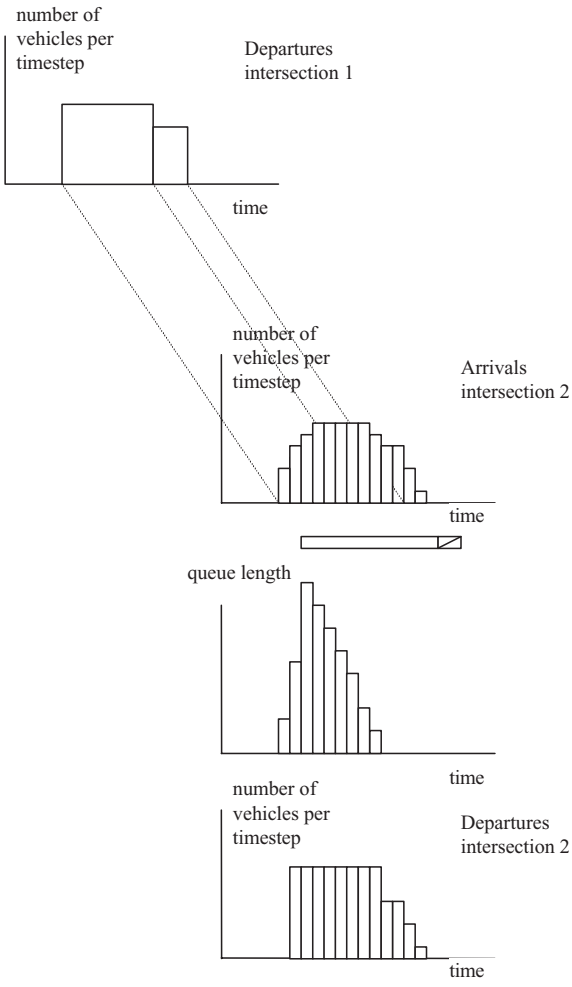


Figure 1. *The calculation of delays in TRANSYT*

have to do it similar to the method applied in TRANSYT: after each change of a parameter we have to simulate the traffic and control process again to calculate the delay. For the search process that is applied for such optimizations sometimes genetic algorithms are used.

2.2. The bilevel optimization

The problem to optimize an objective function by choosing optimal system parameters becomes even more complicated if we see that the system itself is not simply a function of the system parameters, but that the system itself is the result of an optimization process by other actors. In terms of a game we have two players with different objectives. Each optimizes his own objective

function, but they should do so anticipating the reaction of their opponent. In general bot optimization processes are non-cooperative, which means that it is possible that the optimization of player 1 may limit the possibilities for player 2, but it is also possible that player 2 gets better opportunities for optimization thanks to an action of player 1.

In the example of the previous section we assumed that the optimization of the traffic lights could be done for given system conditions: the saturation flows and flows are fixed. In reality it is quite common that people change their routes after an improvement of traffic control. Traffic control has an influence on travel times and drivers try to minimize their travel time by choosing the quickest route (let's assume that this is true; in reality the route choice behavior is much more complicated). In fact we have a kind of Stackelberg game, where the road administrator tries to minimize his objective function, the total delay in the network, while the users minimize their individual travel times. Mathematically this can be represented as

$$\mathrm{Min}_{\boldsymbol{S}}\ D(\boldsymbol{V},\boldsymbol{S}) \tag{15}$$

and

$$\mathrm{Min}_{\boldsymbol{V}}\mathrm{Z}(\boldsymbol{V},\boldsymbol{S}) \tag{16}$$

$$Z(\boldsymbol{V},\boldsymbol{S}) = \sum_a \int_0^{V_a} T(\boldsymbol{S},x)dx, \tag{17}$$

where D is the total delay on all links in the network, $\boldsymbol{V}$ is the vector of all link flows, $\boldsymbol{S}$ is the vector of the signal settings, $T(\boldsymbol{S},x)$ is the travel time as a function of the volume and signal settings and the summation is over all links in the network. It is clear that the vector of decision variables for the road administrator is $\boldsymbol{S}$ and for all road users $\boldsymbol{V}$. However, although the volumes are determined by the minimization of individual travel times, there is no closed form which determines the volumes on the different links, given the control scheme.

We could reduce the bilevel optimization problem to a single level optimization by assuming that it is a Stackelberg game, where one player can predict the decision of the second player and optimize its decision width he assumption that the reaction of the other player is the known output of a predictable decision process. E.g., we can assume that the optimum traffic signal setting chosen by the road administrator are a function of the volumes:

$$\boldsymbol{S} = f(\boldsymbol{V}), \tag{18}$$

where

$$D(\boldsymbol{V},\boldsymbol{S}) \geq D(\boldsymbol{V},f(\boldsymbol{V})), \forall \boldsymbol{S} \in P \tag{19}$$

and P is the set of all feasible signal settings.
Thanks to the simple formulas derived by Webster a long time ago, we know that the minimum delay settings of an isolated pretimed intersections is obtained for the following setting of the control parameters:

$$C = \frac{1.5t_l + 5}{1 - \sum_i V_i/S_i} \tag{20}$$

with t_l the lost time of the control scheme, V_i the flow on stream i and S_i the saturation flow of stream i, while the summation is done over all streams of a conflict group. The optimum distribution of the greentimes is done proportionally to the ratio V_i/S_i.

As an alternative to this approach we might also assume that we optimize traffic control under the assumption that the flows follow directly from an assignment. There are, however, in general no closed formulas as equation (9) to give the flows in a network, given the travel times.

So, we can reduce the bilevel optimization problem of delay minimization for traffic control and route optimization for drivers by a single optimization of route choice with the assumption that the intersection delays are determined by formula (13) and (20). For simple networks, the optimization can easily be done. For instance take the following example.

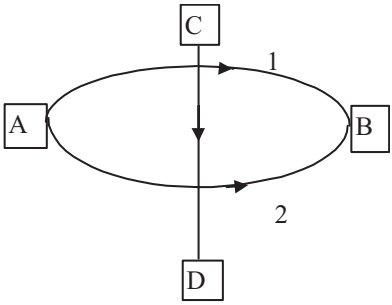


Figure 2. *Example of a network with symmetric choice possibilities*

Between Origin A and destination B two routes exist, each of them have one controlled intersection with the road between C and D. We have now the optimization problem to find the solution for (16). However, we know that this solution is the state where no driver can choose another route with a faster travel time. In this case we can look for those solutions for the route choice that make the travel time over both alternative routes equal. We can calculate the travel time difference between route 1 and 2 as a function of the split over the two routes (Figure 3).

The situation was calculated for 1000 veh/h flow from A to B, 500 veh/h from C to D, saturation flows of 1800 veh/h and internal lost times of 9 seconds, with a minimum green time of 6 seconds.

The equilibrium optimal solution is the symmetrical distribution of 50% over route 1 and 50% over route 2.

However, if we change the flow from A to B to 550 veh/h, the picture changes a lot (Figure 4). Two a-symmetric stable equilibrium states exist (20–80%) and the 50–50% distribution is an equilibrium, but if changes in the distribution occur, the change is enhanced by the subsequent adaptation of the traffic control scheme: the control scheme for the route with the largest flow

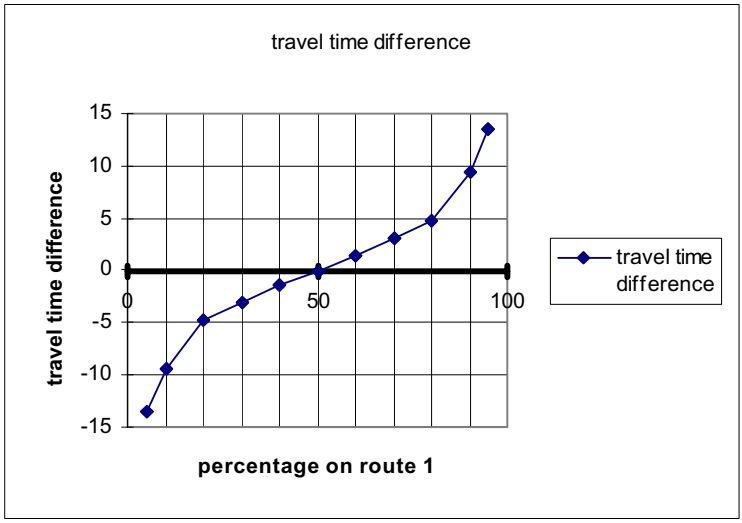


Figure 3. *Travel time difference between route 1 and 2, giving one single symmetric equilibrium. The travel time difference is given in seconds*

gives shorter delays. The total travel time is in both cases at a minimum for the 50–50% distribution.

The conclusion from this little exercise is that the optimum solution of the assignment problem is not always simple to find. Apparently we can find a solution $\boldsymbol{S}(\boldsymbol{V})$ for the signal settings, but if we introduce this solution in the optimization problem (16)–(17), we may find solutions which satisfy the conditions of equation (15)–(16), but which are not stable. In order to identify the character of this optimization problem, we shall solve the interaction between traffic control and route choice in more detail. We can also look at this example using the formulation of equation (16)–(17). The objective function is given by the following graph:

It is obvious that under certain conditions the shape of the objective function changes considerably which makes completely different solutions possible. A closer examination of the objective function Z shows that this peculiar behavior only occurs if the complete delay function is used and the minimum green time constraints are applied to the cycle time. If the minimum green time is assumed to be $0s$, the equilibrium state appears to be the 0–100% distribution between routes 1 and 2.

2.3. The route choice and traffic optimization in two dimensions

In this subsection we shall look at a simple network in which we can study the bilevel optimization process in detail using analytical formulas. In order

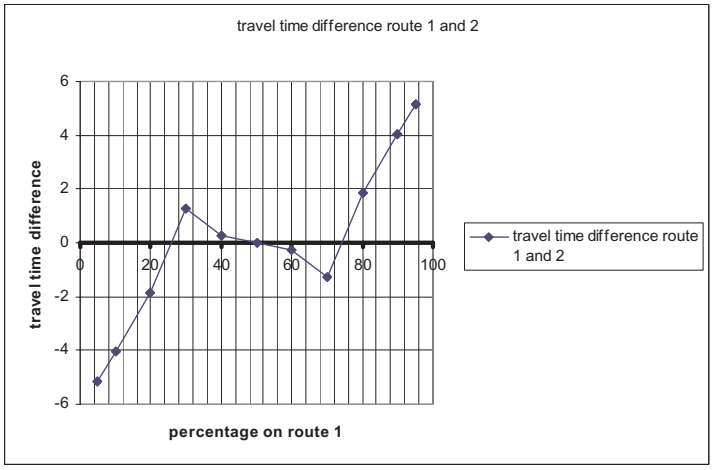


Figure 4. *Travel times differences (in seconds) for route 1 and 2 with asymmetric equilibrium*

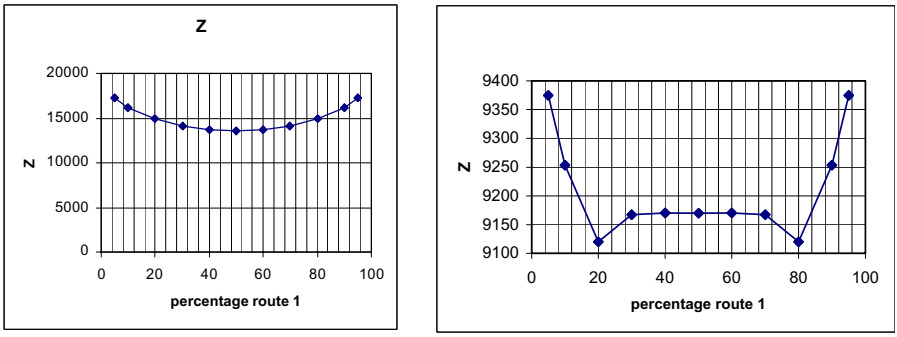


Figure 5. *The objective function for the route choice for the conditions of fig 3 (left) and 4 (right). The function Z was obtained by numerical integration*

to get some more insight in the characteristics of the problem, we shall reduce the traffic control problem to one single dimension. The combined assignment and optimization problem can be visualized in a two dimensional space which makes further analysis easier. We assume that the cycle time remains fixed, the only parameter left is the green split.

The delay D for a single controlled flow, given by equation (13) (difference with (13) is that (21) is delay in sec/veh and (13) gives $sec/cycle$) can be approximated by

$$D(V, C, t_g) \approx 0.9[\frac{1}{2}(C - t_g)^2(1 - V/s)^{-1}C^{-1} + \frac{1}{2}x_2/V(1 - x)] \qquad (21)$$

with

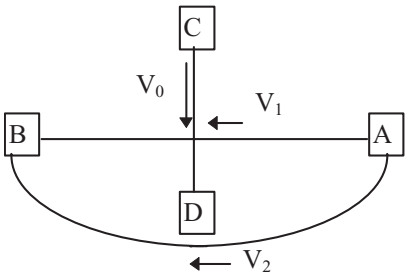


Figure 6. *Simple network for the bilevel optimization of route choice and traffic control*

- V = volume (veh/sec)
- $x = (V/s)(C/t_g)$
- s = saturation flow
- C = cycle time calculated with Websters' method (formula (20)
- t_g = green time.

For both approaches of the intersection together the total delay is given by

$$\begin{aligned}\sum D_i \approx \quad & 0.9[V_1\{\tfrac{1}{2}(C-t_{g1})2(1-V_1/s_1)^{-1}C^{-1} \\ & +\tfrac{1}{2}(V_1C/t_{g1}s_1)^2/V_1(1-V_1C/t_{g1}s_1)\} \\ & +V_0\{\tfrac{1}{2}(C-t_{g0})^2(1-V_0/s_0)^{-1}C^{-1} \\ & +\tfrac{1}{2}(V_0C/t_{g0}s_0)^2/V_0(1-V_0C/t_{g0}s_0)\}],\end{aligned} \tag{22}$$

with

$$\begin{aligned}0 \le V_1 \le V,\\ t_{g1}+t_{g0} = C - t_l,\end{aligned}$$

t_l is the internal lost time of the control scheme. See also Figures 7 and 8.

The route choice problem can be formulated in the following way

$$\min_{\{V\}} Z\{V_1, V_2\} \text{ with } V_1 + V_2 = V \tag{23}$$

where Z can be elaborated to

$$\begin{aligned}Z &= V_1L_1/v + V_2L_2/v + 0.45[(C-t_{g1})^2/C\textstyle\int(1-z/s)dz \\ &+(C/s.t_{g1})^2\textstyle\int z(1-zC/s.t_{g1})dz] \\ &= V_1L_1/v + V_2L_2/v + 0.45[s(C-t_{g1})^2/C\ln(1-V_1/s)^{-1} \\ &-V_1C/s.t_{g1} - \ln(1-V_1C/s.t_{g1})],\end{aligned} \tag{24}$$

where L_i is the length of link i. For equation (18) the boundary condition $V_1 + V_2 = V$ applies, the function Z becomes also a function of two variables,

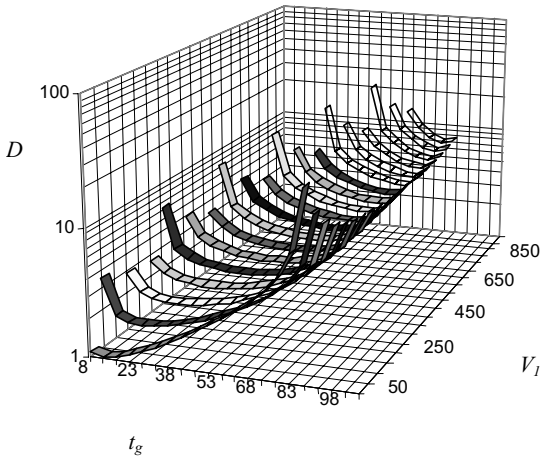


Figure 7. *Total delay as a function of the green time t_{g1} and volume V_1*

V_1 and t_{g1}. Figure 9 gives a graphical representation of $Z(V_1, t_g)$. The equilibrium solutions are on the line drawn in Figure 10 where $\delta Z/\delta V_1 = 0$. If we combine the lines which give the optimum green split (from Figure 8) and the equilibrium assignment (Figure 10) we get Figure 11.

We see, in this example that 3 situations exist were the traffic control is optimized with respect to the traffic volumes and the traffic volumes are consistent with the travel times. If we assume that the process of adjustment of traffic control and route choice are iterative, we find patterns given by the arrows in Figure 11: an adjustment in traffic control will give a change in travel time, with the consequence that some drivers choose another route. The changed traffic volumes make it necessary to adjust the control scheme etc. The process stops if a situation has been reached where the drawn and dotted curves intersect (i.e. points 1, 2 and 3).

If in situation 2 the route choice would change slightly and the traffic control is adapted to the changed flows, we see that a positive feedback mechanism exists: a small variation in route choice is reinforced by the mechanism in which more traffic leads to more green time which reduces delay and attracts more traffic etc. Only at the extremes, where all traffic chooses the same routes or where congestion prevents further growth, does the positive feed disappear.

So in this example two stable equilibrium situations exist: 1 and 3. The total travel time is minimum (for this calculation) in situation 3.

Also in this case the form of the two curves of Figure 11 depends critically on control parameters and (saturation) flows. The optimal green split depend on minimum green time and the internal lost time, which makes that the shape of the curves in Figure 11, is determined for a great deal by the boundaries of the space of feasible solutions. Changes in the boundaries will change the shape of the curves, which can have the consequence that the curves intersect

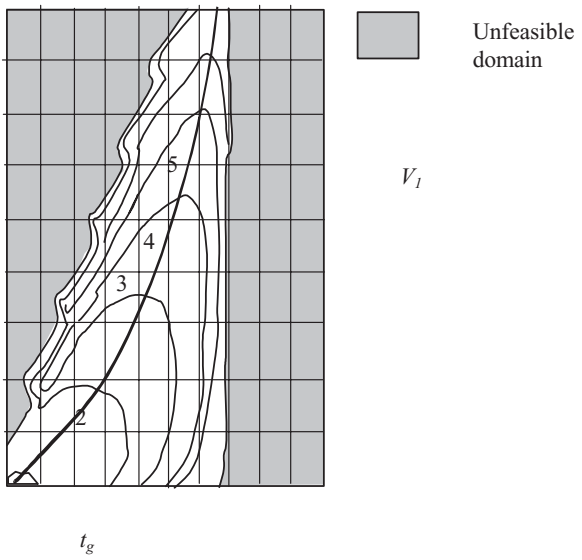


Figure 8. *Iso-curves with equal total delay in the $t_g - V_1$ plane. The thick line gives the green time that minimizes the total delay for a given V_1*

on one, two or more points and that the intersection point can move irregularly after small changes in the system parameters or boundary conditions.

3. Solving the assignment / control problem using the Stackelberg approach

For the simple network with the delay function given by equation (21) and the Z function given by equation (24), it is possible to solve the problem as a Stackelberg game analytically. In the previous section the iteration gives a Nash equilibrium, but if we take equation (24) and solve it for V_1, the solution has to satisfy:

$$\partial Z/\partial V_1 = 0 \text{ or } V_1 C/s.t_{g1} = 1, \tag{25}$$

the first for values of V_1 in the undersaturated condition, the last if the degree of saturation is 1.

The solution can be written as:

$$(L_1 - L_2)/v + 0.45 \left\{ st_r^2/C. \frac{1}{s - V_1} - \frac{C}{s.t_{g1}} + \frac{C}{s.t_{g1}} \frac{1}{1 - V_1 C/st_{g1}} \right\} = 0 \tag{26}$$

or

$$V_1 C/s.t_{g1} = 1 \tag{27}$$

or

$$V_1 = 0. \tag{28}$$

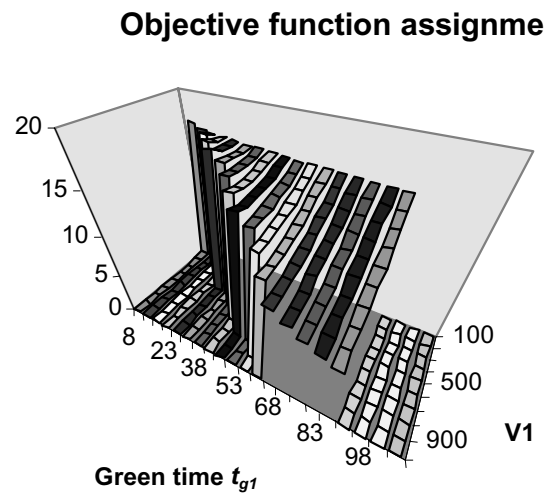


Figure 9. *Objective function Z as a function of t_{g1} and V_1*

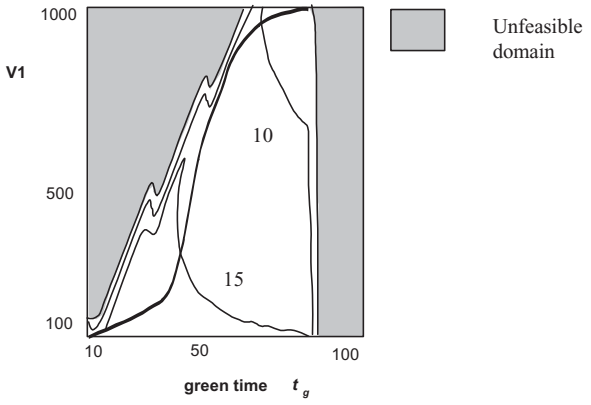


Figure 10. *Iso-lines with equal values of the objective function Z and the (thick) line giving the equilibrium assignment for a given green time*

where equation (27) and (28) correspond to the boundary conditions and $t_r = C - t_{g1}$. If we substitute $s' = s.t_{g1}/C$ and $(L_1 - L_2)/v = \Delta t$, the equation (26) can be rewritten as

$$V_1^2(\Delta t - 0.45/s') - V_1[\Delta t(s + s') + 0.45st_r^2/C - 0.45s/s'] \\ +\Delta tss' + 0.45ss't_r^2/C. \tag{29}$$

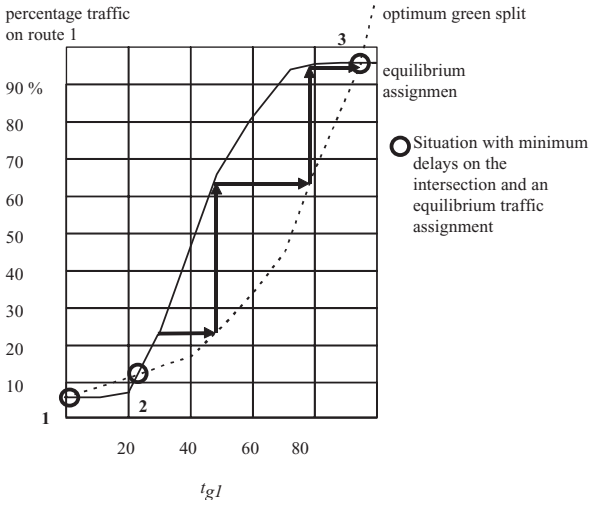


Figure 11. *Optimum green time (dotted line) and equilibrium assignment (drawn line) with three equilibrium situations*

This equation can be solved directly and the solution for V_1 can be substituted in equation (22) giving a solution

$$\begin{aligned} V_1 = {} & 0.5(\Delta t - 0.45/s')^{-1}\{[\Delta t(s+s') + 0.45st_r^2/C - 0.45s/s'] \\ & \pm\{[\Delta t(s+s') + 0.45st_r^2/C - 0.45s/s']^2 - \\ & 4(\Delta t - 0.45/s')(\Delta tss' + 0.45ss't_r^2/C)\}^{1/2}\}. \end{aligned} \tag{30}$$

With this solution for the equilibrium network flows, it is relatively simple to find the solution for the minimum total travel time, substituting the flow V_1 from equation (30) in the total travel given by the delay of equation (22) plus the travel times over both routes. In Figure 11 this means that we search along the dotted line (given by equation 30) for the point with the minimum total travel time. For the example in Figure 11 it happens to be point 3.

4. More bilevel problems

In the previous section we have studied in some detail the bilevel optimization problem for traffic control and route choice. Traffic control can be much more than the control of an intersection. Also ramp-metering can be considered, tidal lanes, variable message signs and travel information are possible tools control traffic. Also pricing (road pricing, parking, and tariffs of public transport) can be seen as instruments at the supply side. Next to route choice one may also look at departure time choice, mode choice, destination choice etc. All these bilevel optimization problems can be categorized as the mixture of demand (travel behavior) and supply (utilization of the traffic infrastructure).

The number of possible applications of bilevel optimization in demand/supply problems is nearly unlimited. For instance, the consumer behavior in response to certain optimization in retailing is an example that may have little to do with traffic, but has problems with a similar structure: the retailer optimizes her business process, which has an impact on the customer who might change her shopping behavior. Another example of bilevel problems is the estimation of an OD-matrix from traffic counts. The counts of traffic on links in a network can be assigned to certain origin-destination pair if it is known which routes are used and how many trips use this particular link:

$$\sum_{ij} T_{ij} p_{ij}^a = V_a \tag{31}$$

where T_{ij} are the number of trips between origin i and destination j, V_a is the volume on link a and p_{ij}^a is the fraction of the trips between i and j that use link a. The matrix estimation problem can be formulated as a minimization problem:

$$\mathrm{Min}_{\boldsymbol{t}} \boldsymbol{Z}_{ME}(\boldsymbol{T}, \boldsymbol{V}) = \boldsymbol{D}(\boldsymbol{T}, \boldsymbol{t})\boldsymbol{U}^{-1}\boldsymbol{D}(\boldsymbol{T}, \boldsymbol{t}) + \boldsymbol{d}(\boldsymbol{V},\ \boldsymbol{v})\boldsymbol{W}^{-1}\boldsymbol{d}(\boldsymbol{V} - \boldsymbol{v}), \tag{32}$$

where $\boldsymbol{Z}_{ME}$ is the objective function for the matrix estimation, $\boldsymbol{T}$ is the OD-matrix to be estimated, $\boldsymbol{t}$ is an a-priori matrix, $\boldsymbol{V}$ is the matrix with observed flows and $\boldsymbol{v}$ is the matrix of flows as calculated by equation (31). The (matrix) operator $\boldsymbol{W}$ is a weight function for the observed flows and $\boldsymbol{U}$ is the weight function for the difference between the a-priori and the estimated matrix. The distance function $\boldsymbol{D(T, t)}$ depends on the model used for the OD-matrix. The difference function $\boldsymbol{d}$ gives a numerical measure for the difference between the observed flows and the flows as calculated by the assignment. (Maher et al. 2001). The optimization on the second level is simply the one given by equation (17).

The quantity p_{ij}^a can be determined with a route assignment, but before we can do so, we have to know the OD-matrix T_{ij}. The od-matrix is calculated by the minimization of the difference between the matrix T_{ij} and some a-priori matrix t_{ij}. The assignment afterwards gives p_{ij}^a and can be done by the technique as described before. So the solution of the OD-matrix estimation problem is a bilevel problem where on one level some distance between the unknown T_{ij} and the apriori matrix t_{ij} is minimized, the other level is the individual optimization of travel times. A very interesting example of bilevel optimization is described by M. Bell (2001). He studied the robustness of infrastructure by playing a game of two parties. One party tries to cause as much harm as possible by removing capacity from a network, the other party tries to restructure the network such that the impact of the destructive action of the other party has the least impact as possible. The final state of the network has to be as robust as possible. The most interesting subject of the different bilevel problems is the way the solution is searched. Since the optimization function are often rather simple, but the interaction between both optimization processes can be very complicated, the solution algorithm has to satisfy the requirement that if finds all relevant solutions in a satisfactory way.

Literature

1. Allsop, R.E., 1974. Some possibilities for using traffic control to influence trip distribution and route choice, *Proceedings of the Sixth International Symposium on Transport and Traffic Theory*. New York: Elsevier.
2. Bell, M.G.H., S. Grosso, 1998. The Path Flow Estimator as a network observer. *Traff. Engng. & Control* Oct. 1998, 540–549.
3. Bell, Michael G.H., 2000. Game theory approach to measuring the performance reliability of transport networks. *Transportation Research Part B* **34** (6), 533–545.
4. Cascetta, E., M. Gallo, B. Montella, 1998. An asymmetric SUE model for the combined assignment-control problem. WCTR congers stream C3, Antwerp.
5. Charlesworth, J.A., 1977. The calculation of mutually consistent signal settings and traffic assignments for a signal controlled network. *Proceedings of the Seventh International Symposium on Transport and Traffic Theory*. Kyoto: The Institute of Systems Science Research 545–569.
6. Chen H.-K., C.-Y. Wang, 1999. Dynamic Capacitated User-Optimal Route Choice Problem. TRB Conference 1999, Washington DC.
7. Fisk. C.S., 1984. Game theory and transportation system modeling. *Transpn. Res.* **18** B 310–310.
8. Maher, M.J., Z. Zhang, D. van Vliet, 2001. A bi-level programming approach for trip matrix estimation and traffic control problems with stochastic user equilibrium link flows. *Transpn Res.* B **35** (2001) 23–40.
9. Mogridge, M.J.H., 1997. The self defeating nature of urban road capacity policy; a review of theories, disputes and available evidence, *Transport Policy Vol. 4* Number 1, pp. 5–24.
10. Mokhtarian, P.L., E.A. Raney, 1997. Behavioral response to congestion: identifying patterns and socio-economic differences in adaption. *Transport Policy* **4** (3), 147–160.
11. Nakayama, S., R. Kitamura, S. Fujii, 1999. Drivers' Learning and Network Behaviour: A dynamic Analysis of Driver-Network System as a Complex System. TRB Conference presentation 990808, Washington DC.
12. H. Pendyala, R.M.R. Kitamura, E.I. Pas, 1997. An Activity Based microsimulation analysis of transportation control measures. *Transport Policy* **4** (3) pp. 183–192.

CWI Vakantiecursus 2003

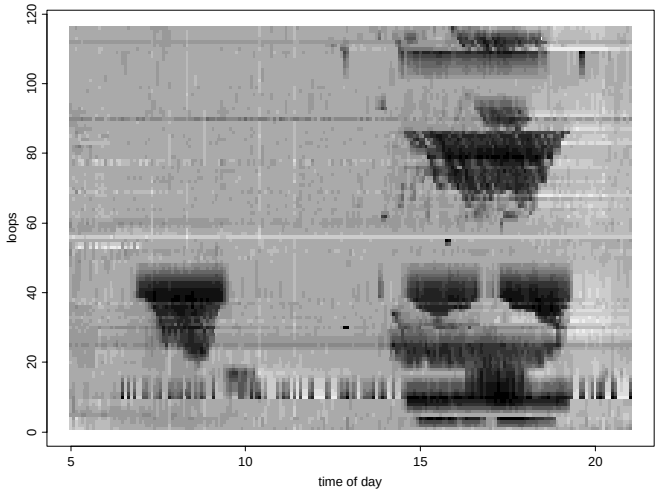
pp. 59 – 66

Reistijden voorspellen op snelwegen

Erik van Zwet
Universiteit Leiden
e-mail: vanzwet@math.leidenuniv.nl

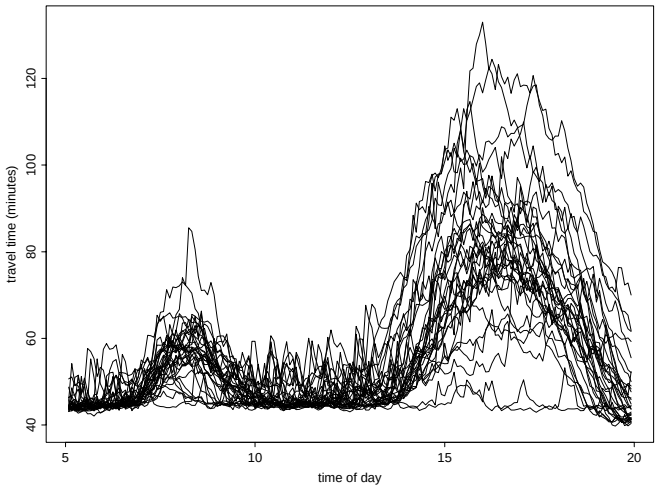
1. Inleiding

Op bijna alle snelwegen in Nederland liggen zogeheten lusdetectoren. Deze detectoren tellen het aantal passerende voertuigen, en meten hun snelheid. Hetzelfde snelheidsmetingen worden ook gedaan in Californië, waar ik de afgelopen drie jaar onderzoek heb gedaan. In Figuur 1 zien we snelheidsmetingen uit Californië op 16 Juni 2000 van 116 lussen langs ongeveer 80 km van Interstate 10 (East) door Los Angeles.



Figuur 1. *Snelheidsveld v op 16 Juni 2000 langs Interstate 10 East in Los Angeles. De donkere, driehoekige vormen zijn files die ontstaan en weer oplossen. De horizontale strepen zijn slecht werkende detectoren*

Met de lusgegevens kunnen we een ruwe berekening maken van de reistijd tussen een tweetal punten A en B. In Figuur 2 zien we een grafiek van de reistijden op ons traject langs I10 East op weekdagen tussen 16 Juni en 8 September 2002. De ochtend– en middagspits zijn duidelijk te herkennen. Op 3 en 4 Juli, feestdagen in Amerika, zijn de reistijden veel korter dan normaal.



Figuur 2. *Reistijden op 34 weekdagen op een traject van 80 km door Los Angeles. Opvallend zijn de grote verschillen in reistijd, vooral in de middagspits*

Op basis van reistijden uit het verleden kunnen we voorspellingen hoe lang het op een gegeven moment in de toekomst zal duren om van A naar B te komen. Dat is het onderwerp van dit artikel.

Als U of ik ergens naar toe gaan, proberen we natuurlijk te voorspellen wanneer we aan zullen komen. We gebruiken daarbij onze ervaring uit het verleden en luisteren naar de verkeersinformatie op de radio. Onze voorspeller doet in zekere zin hetzelfde. De "ervaring" van de voorspeller is een databestand met de lusdetector metingen langs de route op elke minuut van de dag in de afgelopen maanden. De voorspeller maakt een combinatie van deze historische gegevens en de meest recente lusmetingen van, zeg, een paar minuten geleden. Om de optimale combinatie te bepalen maken we gebruik van een statistische methode die (lineaire) regressie heet. Deze term komt van de uitdrukking "regressie naar het gemiddelde". In de praktijk betekent het dat als de verkeers-situatie uitzonderlijk slecht is, we verwachten dat het beter zal worden en vice versa. Hoe snel de verbetering of verslechtering in de richting van het gemiddelde zal plaatsvinden hangt af van verschillende factoren, zoals de locatie en de tijd van de dag.

Bepaalde routes in Nederland (en Amerika) zijn nu eenmaal erg druk, en daardoor is de gemiddelde reistijd langer dan gewenst. Dat is op zich vervelend en kostbaar genoeg, maar wellicht nog schadelijker is de grote variabiliteit van de reistijd op zulke routes. Als het soms een kwartier en soms een uur duurt om naar het werk te gaan, is men gedwongen om ruim op tijd te vertrekken en veelal te vroeg aan te komen. En als men in de file terecht komt, vinden de meeste mensen de onzekerheid over het wel of niet halen van een belangrijke

afspraak erger dan het eigenlijke tijdverlies. Dat de reistijd van dag tot dag flink kan verschillen is duidelijk zichtbaar in Figuur 2. Tijdens de middagspits varieren de reistijden van 40 minuten tot maar liefst twee uur.

Een reistijd voorspeller vermindert niet de gemiddelde reistijd, maar kan wel een groot deel van de onzekerheid elimineren. Er blijft natuurlijk nog wel enige onzekerheid over, want het is niet mogelijk om tot op de minuut precies te voorspellen hoe lang het zal duren om van Amsterdam naar Breda te rijden. Hoeveel data, rekenkracht of slimme trucs we ook hebben, we kunnen niet daadwerkelijk in de toekomst kijken. Het blijkt echter, dat als de reistijd op een zekere route een spreiding (standaard deviatie) van, zeg, 20 minuten heeft, onze voorspeller die kan terugbrengen tot 8 minuten. De winst van 12 minuten is zeker zo nuttig als een daadwerkelijke vermindering van de reistijd.

2. Het Probleem

Zij $v(i,d,t)$ de snelheid gemeten bij lus i op dag d en tijd t. We willen de reistijd voorspellen van een reis langs lussen $i = 1,\ldots,I$ als we vertrekken op een zeker tijdstip in de toekomst. In Figuur 1 zien we een voorbeeld van het snelheidsveld v voor één dag. Het is zeker niet duidelijk hoe we alle informatie die we hebben verzameld tot op het huidige moment het beste kunnen gebruiken om een reistijd voorspeller te definiëren. We hebben echter een compressie van deze informatie gevonden, die voor ons doeleinde bijzonder effectief is (Rice en van Zwet, 2001, Zhang en Rice, 2001).

Merk op dat we met behulp van het snelheidsveld v de reistijd $T(d,t)$ kunnen berekenen bij vertrek op dag d, tijd t. Deze reistijd kunnen we ons voorstellen als behorend bij een pad door het veld v. We merken op dat we informatie van *na* het tijdstip t nodig hebben om de berekening uit te voeren. Met behulp van de informatie die we ter beschikking hebben *op* tijd t kunnen we wel de zogeheten instantane reistijd $T^*(d,t)$ uitrekenen.

$$T^*(d,t) = \sum_{i=1}^{I-1} \frac{2d_i}{v(i,d,t) + v(i+1,d,t)}, \tag{1}$$

waarbij d_i de afstand tussen de lussen i en $i+1$ is. De instantane reistijd zou daadwerkelijk worden gerealiseerd als de snelheid na tijd t ongewijzigd zou blijven, totdat de reis voltooid is.

Als we de reistijd $T(d,t)$ hebben berekend voor een aantal dagen $d = 1,2,\ldots,n$ in het verleden, dan kunnen we ook het historisch gemiddelde berekenen

$$\bar{T}(t) = \frac{1}{n}\sum_{d=1}^{n} T(d,t). \tag{2}$$

Ons doel is de reistijd $T(d,t+\delta)$, $\delta \geq 0$, te voorspellen met behulp van alle gegevens die we hebben op dag d en tijd t. Hier is δ de tijd tot vertrek, en zoals we al opmerkten is ons probleem ook voor $\delta = 0$ niet triviaal. Twee naïeve voorspellers liggen voor de hand, en worden ook in de praktijk vaak gebruikt. Het historisch gemiddelde $\bar{T}(t+\delta)$ en de instantane reistijd op tijd t, $T^*(d,t)$.

We verwachten—en dat blijkt ook zo te zijn—dat $\bar{T}(t+\delta)$ het het beste zal doen voor grote δ, en $T^*(d,t)$ voor kleine δ. Onze nieuwe voorspeller is een gewogen gemiddelde van deze twee naïeve voorspellers, en doet het beter dan beide voor alle δ.

3. Lineaire Regressie

Bij de bestudering van grote hoeveelheden snelwegdata, is ons een empirisch feit opgevallen: dat er een lineaire relatie bestaat tussen $T^*(d,t)$ en $T(d,t+\delta)$. In Figuren 3 en 4 zetten we $T^*(d,t)$ uit tegen $T(d,t+\delta)$ voor onze data van Interstate 10 East. Merk op dat de expliciete relatie varieert met de keuze van t en δ, maar dat de lineariteit gehandhaafd blijft. Met deze observatie in gedachten, stellen we het volgende model op

$$T(d,t+\delta) = \alpha(t,\delta) + \beta(t,\delta)T^*(d,t) + \varepsilon. \tag{3}$$

waarbij ε een stochastische grootheid is met verwachting nul die toevallige veranderingen en meetfouten modelleert. Merk op dat de parameters α en β mogen varieren met t en δ. Lineaire modellen met veranderende parameters worden besproken door Hastie en Tibshirani (1983).

We kunnen $\alpha(t,\delta)$ en $\beta(t,\delta)$ bepalen door de methode van de kleinste kwadraten te gebruiken. Dat wil zeggen, we kiezen $\alpha(t,\delta)$ en $\beta(t,\delta)$ zó dat de kwadratische fout genomen over de historische data zo klein mogelijk is. Met andere woorden, we minimaliseren

$$\sum_{d=1}^{n}(T(d,t+\delta) - \alpha(t,\delta) - \beta(t,\delta)T^*(d,t))^2 \tag{4}$$

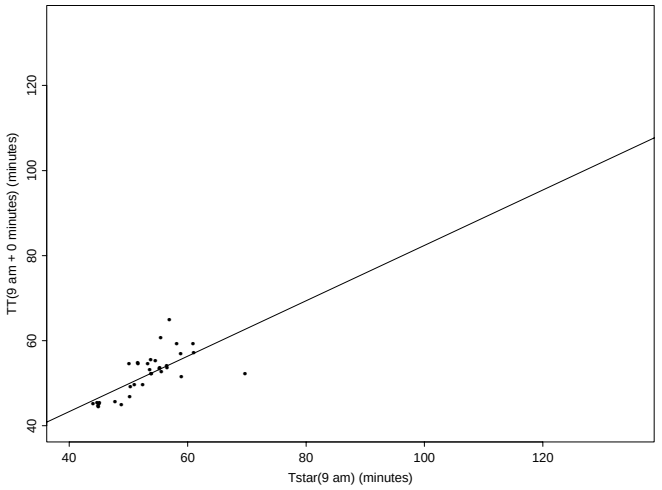
over α en β. We duiden de optimale parameter waarden aan met $\hat{\alpha}(t,\delta)$ en $\hat{\beta}(t,\delta)$ en definiëren onze voorspeller als

$$\hat{T}(d,t+\delta) = \hat{\alpha}(t,\delta) + \hat{\beta}(t,\delta)T^*(d,t). \tag{5}$$

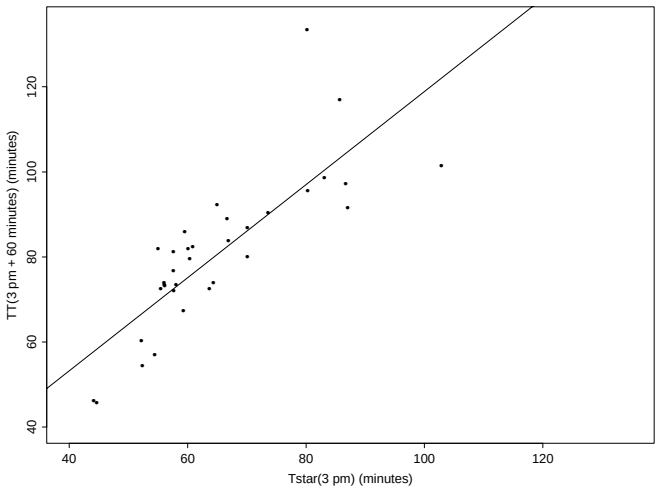
Als we $\alpha(t,\delta)$ herschrijven als $\alpha'(t,\delta)\bar{T}(t+\delta)$, dan zien we dat (3) en (5) de toekomstige reistijd uitdrukken als een lineaire combinatie van de twee naïeve voorspellers $\bar{T}(t+\delta)$ en $T^*(d,t)$. We kunnen onze nieuwe voorspeller dus interpreteren als de beste lineaire combinatie van de twee naïeve voorspellers. We mogen verwachten dat onze voorspeller het beter doet dan beide en dat blijkt in de praktijk ook zo te zijn.

4. Voorbeeld

We hebben onze voorspeller uitgeprobeerd op onze data van 116 lusdetectoren op een traject van 80 km in Los Angeles. Lusmetingen werden gedaan van 5 uur 's ochtends tot 9 uur 's avonds gedurende 34 weekdagen tussen 16 Juni en 8 September 2000. De metingen werden geaggregeerd over intervallen van 5 minuten. Het snelheidsveld v voor 16 Juni zien we in Figuur 1. De opvallende horizontale lijnen zijn slecht werkende lussen. Het automatisch identificeren en



Figuur 3. *T^*(9 uur) versus T(9 uur). De regressie lijn snijdt de y-as in het punt α=17.3 en heeft helling β=0.65*



Figuur 4. *T^*(15 uur) versus T(16 uur). De regressie lijn snijdt de y-as in het punt α=9.5 en heeft helling β=1.1*

corrigeren van slechte data is een belangrijk probleem, maar we laten dat hier voor wat het is. We hebben de slechte metingen "met de hand" verwijderd en vervangen door middel van lineare interpolatie. Vervolgens hebben we de reistijden $T(d,t)$ voor alle dagen berekend, en deze zien we in Figuur 2.

We vergelijken de nauwkeurigheid van onze nieuwe voorspeller $\hat{T}(d,t+\delta)$ met de twee naïeve voorspellers $\bar{T}(t+\delta)$ en $T^*(d,t)$ voor verschillende keuzes van t (5 uur, 6 uur,..., 20 uur) en δ (0 minuten, 60 minuten). We hebben de wortel van de kwadratische fout geschat door steeds één dag weg te laten, de voorspeller voor die dag te bepalen op basis van de andere dagen, en de kwadratische fouten te middelen. De wortel van de kwadratische fout (root mean squared error, ofwel RMSE) is

$$RMSE(t,\delta) = \left(\frac{1}{34} \sum_{d=1}^{34} (T(d,t+\delta) - \hat{T}(d,t+\delta))^2 \right)^{1/2} \tag{6}$$

De resultaten voor $\delta = 0$ en $\delta = 60$ minuten zijn weergegeven in Figuren 5 and 6. Allereerst merken we op dat de instantane reistijd T^* een redelijk goede voorspeller is voor kleine δ, maar zeer slecht voor grote δ. Het historisch gemiddelde is slechter dan T^* voor kleine δ, maar beter voor grote δ. Het belangrijkste resultaat is dat onze nieuwe voorspeller het beter doet dan beide. De RMSE van onze voorspeller blijft onder de 10 minuten, zelfs als we een uur van tevoren voorspellen (δ =60 minuten).

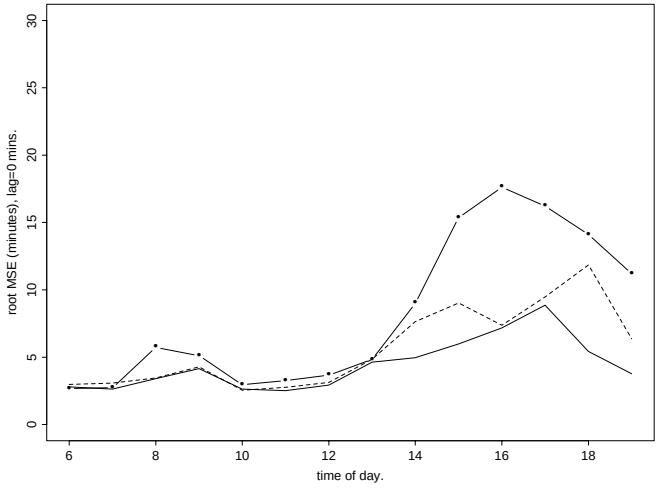
De RMSE van het historisch gemiddelde is een schatting voor de standaard deviatie van de reistijd. We zien in Figuur 5 hoe onze voorspeller deze in de middagspits terugbrengt van 20 minuten naar 8.

5. Opmerkingen

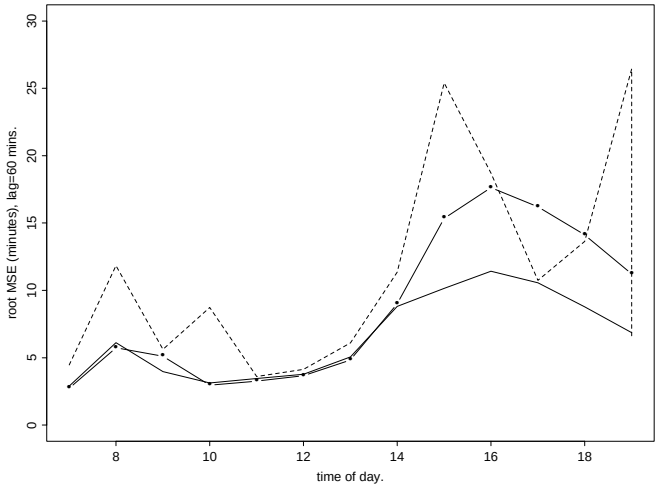
Reistijdvoorspellen is een levendig onderzoeksgebied, en onze voorspeller is zeker niet de enige. Allerlei methoden, bijvoorbeeld gebaseerd op tijdreeks analyse, of op het principe van nabije buren (nearest neighbours), zijn eerder al geprobeerd. Een interessante aanpak die onlangs in Delft is ontwikkeld, maakt gebruik van een neuraal netwerk (van Lint et al, 2002).

Hèt grote voordeel van onze methode—volgens ons—is zijn eenvoud. De berekening van de optimale parameter waarden $\hat{\alpha}(t,\delta)$ en $\hat{\beta}(t,\delta)$, voor alle t en δ, kost enig gereken, maar dit kan "off-line" gedaan worden. Om nu op dag d, tijd t een reistijd te voorspellen hoeven we alleen de instantane reistijd $T^*(d,t)$ te berekenen en uit te vermenigvuldigen met de parameters $\hat{\alpha}(t,\delta)$ en $\hat{\beta}(t,\delta)$ voor de juiste δ.

We hebben onze voorspeller geïmplementeerd voor het netwerk van snelwegen dat Los Angeles doorkruist. We hebben zo'n 40 vertrekpunten en bestemmingen gekozen en voor elk paar de 10 kortste routes berekend. Een gebruiker kan via het Internet een reistijd voorspelling en bijbehorende beste route opvragen voor iedere vertrektijd in de toekomst. Als δ (de tijd tot vertrek) groot wordt, convergeert onze vooorspeller vanzelf naar het historisch gemiddelde, en telt de huidige verkeerssituatie dus niet meer mee.



Figuur 5. *Geschatte wortel van de kwadratische fout voor δ= 0 minuten. Historisch gemiddelde $\bar{T}$ (– · –), instantane reistijd T^* (- - -) en onze voorspeller $\hat{T}$ (—)*



Figuur 6. *Geschatte wortel van de kwadratische fout voor δ= 60 minuten. Historisch gemiddelde $\bar{T}$ (– · –), instantane reistijd T^* (- - -) en onze voorspeller $\hat{T}$ (—)*

Literatuur

1. T. Hastie and R. Tibshirani (1993). Varying coefficient models. *Journal of the Royal Statistical Society Series B*, **55(4)** pp. 757–796.
2. H. van Lint, S. P. Hoogendoorn, H. J. van Zuylen (2002). State space neural networks for freeway travel time prediction. ICANN pp. 1043–1048
3. J. Rice and E.W. van Zwet (2001). A simple and effective method for predicting travel times on freeways. Aangeboden aan: *IEEE Transactions on Intelligent Transportation Systems.*
4. X. Zhang and J. Rice (2001). Short-term travel time prediction using a time-varying coefficient linear model. Te verschijnen in: *Transportation Research C.*

CWI Vakantiecursus 2003

pp. 67 – 92

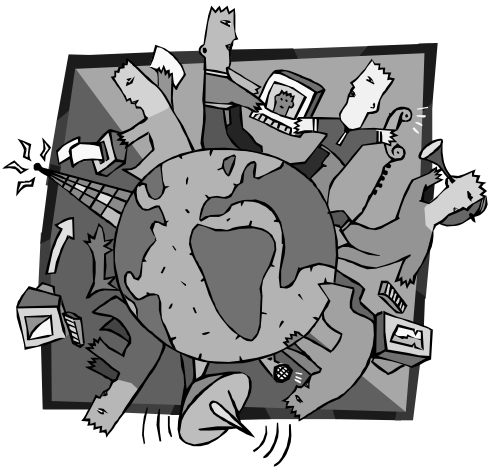
Nieuwe generatie telecommunicatietechnieken en -diensten

George Huitema
TNO Telecom
e-mail: g.b.huitema@telecom.tno.nl

1. De wereld van Telecommunicatie

1.1. Wiskunde laat de mobiele telefoon piepen

Het is meer dan 100 jaar geleden dat Bell de telefoon uitvond. Sinds die tijd heeft telecommunicatie de wereld veroverd. De communicatiemogelijkheden in het dagelijkse leven zijn door de invoering van digitale transmissietechnieken en de opkomst van internet enorm toegenomen. Dankzij de nieuwste technieken is het mogelijk om wereldwijd vanaf elke locatie met elke andere locatie te communiceren.



Figuur 1. *Wereldwijde communicatie*

Een groot aantal informatie- en datadiensten staat de huidige gebruikers ter beschikking. In de nabije toekomst zullen de nieuwe generatie vaste en mobiele breedbandige diensten de gebruikers een nog groter scala aan nieuwe communicatiediensten verschaffen. Consumenten krijgen beschikking over een grote diversiteit aan randapparaten en communicatiemiddelen. Door de technologische ontwikkelingen vloeien communicatie, informatie, commercie en entertainment samen in een nieuwe industrie.

In dit verhaal wordt eerst een globaal, schetsmatig overzicht gegeven van de steeds maar voortgaande ontwikkelingen in de telecommunicatiewereld. Hierbij zal telkens kort worden aangeduid welke specifieke wiskundige disciplines een rol spelen. De nieuwe generaties telecommunicatienetwerken worden in hoofdstuk 2 beschreven waarna in hoofdstuk 3 een overzicht volgt van de nieuwe communicatie- en informatiediensten die door deze nieuwe netwerken mogelijk worden. In de daaropvolgende hoofdstukken 4 en 5 wordt nader ingegaan op het proces van aanbieden en verrekenen van telematicadiensten waarbij enige wiskundige cases worden uitgewerkt. Om daadwerkelijk telecommunicatiediensten te gebruiken zullen klanten zich eerst aan het netwerk kenbaar moeten maken. Hierbij wordt geverifieerd of de persoon is wie hij zegt te zijn, en of hij in kwestie gerechtigd is om van de dienst gebruik te mogen maken (*authenticatie* en *autorisatie*). Ten slotte -slechts de zon gaat voor niets op- zal ook het telecommunicatiegebruik van klanten moeten worden geregistreerd en verrekend (*accounting* en *billing*).

Met dank aan Jan van Maanen en Rob van der Mei voor het kritisch doorlezen van de tekst van deze bijdrage.

1.2. Telecommunicatie

Technisch gezien is telecommunicatie de overdracht van informatie over afstand, van een zendende partij (*transmitter*) naar een ontvangende partij (*receiver*). Deze overdacht gebeurt in een aantal achtereenvolgende stappen. Bij elk van deze stappen spelen wiskundige methoden en technieken een belangrijke rol. Op het eerste gezicht zou men niet denken dat bijvoorbeeld een mobiel telefoontoestel zo veel wiskunde in zich draagt. De voortgaande ontwikkelingen van mobiele telecommunicatie maken het straks nog zo dat een mobiel meer rekenkracht in zich draagt dan de meest geavanceerde zakrekenmachine van tegenwoordig!

Om een idee van de wiskundige toepassingen bij telecommunicatie te krijgen lopen we de verschillende stappen bij het transport van informatie nu bij langs. Eerst wordt de informatiestroom omgezet in een gecomprimeerde, digitale gegevensstroom (*digitalisering* en *codering*). Daarna worden aan de datastroom allerlei redundante gegevens toegevoegd zodat foutdetectie en foutcorrectie op de te transporteren gegevens mogelijk is (*foutdetectie* en *foutcorrectie*). Vervolgens wordt de datastroom versleuteld (encryptie) ter voorkoming van afluisteren en ongewenste veranderingen. En ten slotte wordt voordat het signaal daadwerkelijk verzonden wordt, dit aan een geschikte zogeheten draaggolf toegevoegd (*modulatie*). Na versturing van het signaal op basis van een *communicatieprotocol*, vindt bij de ontvangende partij het omgekeerde proces plaats: het signaal wordt van de draaggolf gescheiden (demodulatie), de gegevens worden ontcijferd (*encryptie*) en ten slotte in de voor de ontvangende partij weer in begrijpelijk vorm teruggebracht (*encodering*).

Naast technische oplossingen voor het transporteren van informatie in netwerken, wordt in de telecommunicatiewereld ook veel techniek ingezet voor het verdelen van schaarste. De exploitatie van een communicatienetwerk wordt namelijk gekenmerkt door het zoeken naar de meest economische manier van

het beheren van schaarste aan bandbreedte en transport- en dienstencapaciteit voor gebruikers. Immers gebruikers willen vanaf elke locatie, op elk tijdstip, in elke gewenste kwaliteit en met voldoende beveiliging en privacy, met andere gebruikers communiceren. Het opzetten en onderhouden van een goed telecommunicatienetwerk kost veel geld. De kunst voor een telecommunicatieaanbieder is dan ook om tegen lage kosten een zo hoog mogelijke kwaliteit aan gebruikers te leveren. Dat hier veel wiskunde bij komt kijken, moge duidelijk zijn

1.3. Voortdurende groei

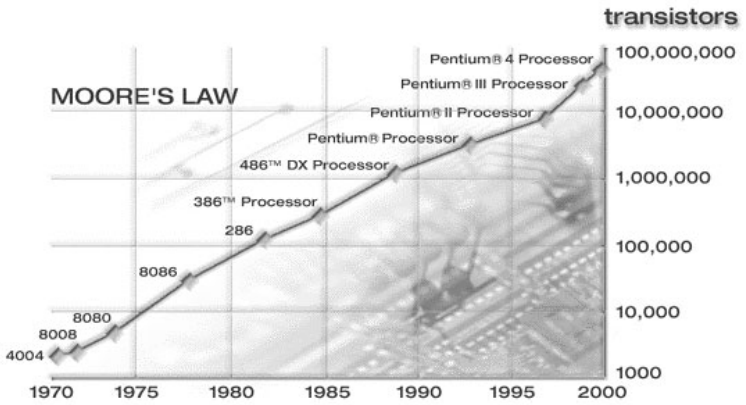
De voortdurende groei aan communicatiemogelijkheden kent een tweetal belangrijke oorzaken. Een daarvan is de continue miniaturisatie van de microelektronica. De ander is het feit dat het gebruik van een bepaald medium aantrekkelijker wordt indien meer personen gebruik maken van dezelfde communicatiemiddelen. Deze twee bekende waarnemingen gaan door het leven als de wetten van Moore en Metcalfe.

In 1965 constateerde Moore, medeoprichter van de chipfabrikant Intel, dat ieder jaar het aantal transistoren per cm^2 op een chip verdubbelde. Moore voorspelde dat dit in de toekomst gewoon door zou gaan. In de daarop volgende jaren is het tempo wel een beetje afgenomen maar de dichtheid verdubbelt nog steeds elke 18 maanden (vergelijk Figuur 2). Een soortgelijke verdubbelingsnelheid zien we bij opslag- en geheugencapaciteit van computers, en bij telecommunicatiebandbreedte. Dit is nu wat de wet van Moore ons zegt.

Wet van Moore

Voor verschilende elementen uit de ICT-wereld geldt een exponentiële groei. Verdubbelingstijden lopen uiteen van 0.7 jaar tot 3 jaar.

Experts verwachten dat deze 'wet' nog enkele decennia zal blijven gelden.



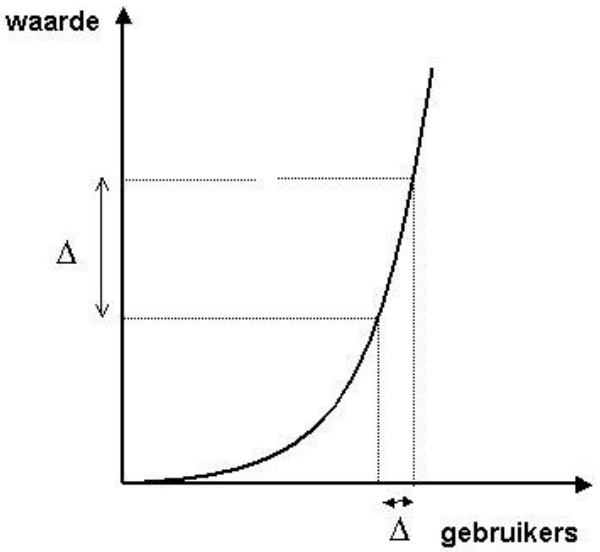
***Figuur 2**. Illustratie van de Wet van Moore aan de groei van de dichtheid van transistoren*

Door de Wet van Moore wordt het steeds aantrekkelijker voor zowel consumenten als voor bedrijven om de juiste elektronische middelen te hebben voor communicatie en voor handel te bedrijven.

Metcalfe, ontwerper van het veel gebruikte Ethernet-protocol voor computernetwerken en oprichter van 3COM, leverancier van netwerksystemen, ontdekte (vergelijk Figuur 3):

Wet van Metcalfe
De waarde van een communicatienetwerk is gelijk aan het kwadraat van het aantal gebruikers.

Het versturen van mobiele tekstberichten (SMS) is hier een goed voorbeeld van. Een SMS-je versturen is van beperkt nut als er weinig mensen in je vrienden- of relatienetwerk zijn die er gebruik van maken. Maar als het merendeel van je relaties een mobiele telefoon heeft met deze mogelijkheid, wordt het ineens veel interessanter. Dit verklaart de snelle doorbraak van SMS de afgelopen jaren.



Figuur 3. *Wet van Metcalfe*

Bij de introductie van een nieuwe communicatietechnologie zal in het begin niet veel meer te merken zijn dan dat er een toenemend aantal gebruikers is. Maar als de technologie eenmaal gemeengoed geworden is, zullen er ook economische en sociale veranderingen volgen. Dit fenomeen wordt disruption genoemd. Bekende voorbeelden uit het verleden zijn de invoering van radio en TV. Moore's wet voorspelt dat in onze huidige digitale maatschappij het bereiken van de kritische massa steeds sneller gaat. Heden ten dage kunnen we dit met onze eigen ogen en oren constateren bij het gebruik van de PC en de mobiele telefoon.

2. Nieuwe generatie telecommunicatienetwerken

In het algemeen bestaat een telecommunicatienetwerk uit een verzameling van systemen die met elkaar kunnen communiceren. Voorbeelden van telecommunicatienetwerken zijn de traditionele telefoonnetwerken en de moderne mobiele netwerken. In dit laatste geval zijn de communicerende systemen mobiele telefoons (*handhelds*). Op het gebied van netwerken zijn een aantal belangrijke ontwikkelingen te noemen.

2.1. Pakketgeschakelde netwerken

Telefoonnetwerken werken op basis van *circuitschakeling*. Wanneer een abonnee een andere abonnee belt, wordt tussen de twee telefoons een vaste geschakelde verbinding tot stand gebracht. Tijdens het bellen blijft de verbinding open staan en geen enkele andere beller kan gebruik maken van deze verbinding voordat een van beide bellers het lopende gesprek beëindigt. Tegenover deze technologie staat *pakketschakeling*. Hierbij kan elke verbinding in het netwerk tegelijkertijd gebruikt worden door meerdere gebruikers. Er is geen specifieke, open verbinding tussen twee gebruikers, het netwerk is *connectionless*. De stroom informatiegegevens tussen gebruikers wordt opgeknipt in losse pakketjes, die elk een eigen weg door het netwerk volgen. Bij aankomst bij de ontvangende partij worden de pakketjes weer in de goede volgorde gezet tot de oorspronkelijke gegevensstroom.

Het voordeel van pakketschakeling is dat het de bandbreedte van de verbindingen beter benut (efficiëntie) en dat het goedkoper en eenvoudiger valt te implementeren. Circuitschakeling biedt in het algmeen betere beveiliging daar verbindingen niet gedeeld worden met anderen. Pakketschakeling wordt in bijna alle nieuwe generatie netwerken toegepast. Hét voorbeeld hier is het *Internet* dat bestaat uit een netwerk van pakketgeschakelde netwerken.

2.2. Lokale en persoonlijke netwerken

Wireless Local Area Network (WLAN)

WLAN's zijn lokale netwerken waarbij draadloze dataoverdracht tussen gebruikers, plaatsvindt door middel van gestandaardiseerde netwerktechnologie. Momenteel wordt veelal de IEEE 802.11b communicatiestandaard (WiFi) gebruikt die theoretisch een overdrachtssnelheid van 11 Mbps mogelijk maakt. Via kleine PC-insteekkaartjes kunnen zo laptops en kleine handcomputertjes (PDA's, Personal Digital Assistant) draadloze verbindingen met internet opbouwen. Deze tekst van ongeveer 400kB zou in een fractie van een seconde van internet gehaald kunnen worden. WLAN's, ook wel *hotspots* genoemd, zijn momenteel al te vinden op vliegvelden, zakenhotels en bij sommige ketens van coffeeshops.

Personal Area Network (PAN)

De vrij nieuwe naam *Personal Area Network* wordt gebruikt voor het netwerk van persoonlijke communicatie- en informatiesystemen dat in een straal van

ongeveer 10 m rondom een persoon valt te onderscheiden. Hieronder vallen bijvoorbeeld de verbindingen (al of niet draadloos) tussen printers, mobiele telefoons, digitale camera's en PDA's. De communicatie bestaat doorgaans uit het overbrengen van files (o.a. plaatjes) en de synchronisatie van gegevens. De communicatietechniek *Bluetooth* is hierbij sterk in opkomst doordat de bijbehorende communicatiechips relatief goedkoop zijn en weinig energie verbruiken. De netwerken in deze categorie hebben meestal een ad hoc karakter, dat wil zeggen dat per gelegenheid er een netwerkconfiguratie ontstaat die het mogelijk maakt dat applicaties op de verschillende apparaten met elkaar communiceren. Hiervoor is wel een vereiste dat apparaten van verschillende leveranciers goed op elkaar aansluiten. Veel lezers zullen ondertussen wel eens in hun eigen omgeving hebben meegemaakt dat laptops, mobiele telefoons of camera's niet altijd goed samenwerken. Ondanks dat dezelfde standaarden zijn gebruikt, blijken er helaas soms verschillende implementaties door producenten gebruikt te worden.

2.3. Mobiele netwerken

Sinds de eerste mobiele toepassingen waarbij de communicatie tussen apparaten niet via een vaste draad verliep maar via draadloze transmissie, zijn er ondertussen al drie verschillende generaties mobiele netwerken te onderscheiden.

1G, de eerste generatie, is de naam voor de analoge mobiele netwerken uit het midden van de jaren 1980. Deze netwerken waren nationaal van karakter en ondersteunden slechts spraakdiensten of daaraan gerelateerde diensten zoals doorschakelen. In Nederland kenden we deze mobiele netwerken onder de naam *ATF* (Auto TeleFoon). Door het nationale karakter waren deze netwerken incompatibel met elkaar en mobiele communicatie werd meer beschouwd als een bijzonderheid, zie Figuur 4.

Door de steeds groeiende behoefte aan mobiele communicatie werd echter ook de behoefte aan een globaal mobiel netwerk groter. Door inspanningen van internationale standaardisatieorganisaties is vanuit Europa het huidige GSM netwerk (Global System for Mobile communications) ontstaan. Deze 2e generatie mobiele digitale communicatietechniek is een groot technisch en commercieel succes geworden. Vergeleken met de lappendekens aan mobiele netwerken in Amerika die niet goed op elkaar aansluiten, is GSM een grote Europese verworvenheid.

De op handen zijnde 3^{e} generatie mobiele netwerktechnologie, 3G, is ontworpen als een volledig overkoepelend mobiel netwerk voor globale mobiele communicatie. In Europa is UMTS (Universal Mobile Telecommunication System) de 3G-standaard geworden. In 3G worden zowel spraak als multimedia ondersteund, zie Figuur 5. Bovendien omvat 3G internetdiensten.

Hoewel 3G op dit moment vooral nog in een experimenteel stadium is, zijn recentelijk met name in de Aziatische wereld de ontwikkelingen voor 4G al opgestart, zie [1]. Deze volgende generatie mobiele netwerken moet 3G weer verder brengen en alle vormen van spraak en datadiensten in mobiele context omvatten.

Doordat mobiele technologie steeds kleiner en goedkoper wordt, vinden we

Figuur 4. *ATF-Telefoon (Carvox 2453, 1985, Museum voor Communicatie, Den Haag)*

Figuur 5. *Nieuwe 3G-telefoon*

in allerlei apparaten en gebruiksvoorwerpen steeds vaker mobiele devices geplaatst die onderling communiceren. De verwachting is dat deze specifieke communicatie tussen apparaten (we spreken hier van Machine-2-Machine communicatie) op den duur ruimschoots de omvang van mobiele communicatie tussen personen overtreft. Voor een visie op deze onzichtbare mobiele ontwikkelingen zie [16].

Wiskunde voor mobiele netwerken

Vergeleken met de traditionele vaste netwerken kennen mobiele netwerken geheel eigen specifieke wiskundige problemen. Zo zijn bijvoorbeeld voor de realisatie van 1^e en 2^e generatie mobiel netwerken uitgebreide netwerkplanningstools ontwikkeld, die aangeven waar mobiele masten moeten worden geplaatst opdat mobiele gebruikers volledige en uniforme dekking hebben van hun mobiele telefoon. Doordat het gebruik van 3G-netwerken een geheel andere verdeling van dataverkeer over het netwerk laat zien dan bij de huidige 2G-netwerken, worden momenteel geheel nieuwe 3G-specifieke planningtools ontwikkeld [10]. Verder zijn er ook nieuwe toegangstechnieken ontwikkeld, onder de naam *WCDMA* (Wideband Code Division Multiple Access). Op basis van lineaire algebra wordt er voor gezorgd dat gebruikers in een 3G-netwerk elkaar zo weinig mogelijk storen.

Een andere nieuwe ontwikkeling is dat 3G-Netwerken ook verschillende kwaliteitsklassen voor diensten mogelijk maken, zoals het bekijken van videobeelden bij verschillende bandbreedten. Het toekennen en garanderen van bepaalde kwaliteitsniveaus (*QoS*, Quality of Service) berust op geavanceerde wiskundige technieken, zie [14, 22].

Voor de nieuwe generatie mobiele netwerken valt in het algemeen op te merken dat steeds méér verschillende wiskundedisciplines worden toegepast. Om enkele gebieden te noemen:

- Security (complexe cryptografische algoritmen),
- Op Internet gebaseerde communicatie (verkeersmodellering), zie [15],
- Bestellen van goederen via mobiele telefoon (digitale handtekeningen en sleutelsystemen),
- Op locatie gebaseerde diensten (snelle geometrische algoritmen),
- Systeemspecificaties en prototypen (wiskundige modelleringstechnieken),
- Beheer van mobiele toestellen en van de dienstenplatformen in het mobiele netwerk (management van complexe softwaresystemen).

3. Een scala aan nieuwe communicatie- en informatiediensten

De wereld van nieuwe telecommunicatiediensten kan op verschillende manieren worden ingedeeld. Hieronder volgt een eenvoudige, brede indeling (zie [3]), die een duidelijk beeld geeft van de verschillen tussen de soorten diensten, zoals wie communiceert met wie, welke transacties vinden er plaats, etc.

– Interpersoonlijke Communicatie
– Informatie en Entertainment (Infotainment)
– Zakelijke Diensten
– Persoonlijk Ondernemen

Bij Interpersoonlijke Communicatie vindt het communiceren plaats tussen personen onderling, tegenwoordig ook vaak Person-2-Person (P-2-P) genoemd. De communicatie gebeurt via telefoongesprekken, e-mail en het naar elkaar sturen van mobiele tekstberichten (SMS). Opkomende diensten in deze categorie zijn onder andere het sturen van mobiele multimedia berichten (MMS), zoals een fotootje dat gemaakt is met de mobiele telefoon zelf.

Diensten die informatie verschaffen of de gebruiker enig vermaak geven, vallen onder de categorie Infotainment. Tot dit gebied behoren onder andere nieuwsdiensten, spelletjes (games), moppen, muziek en informatie over sport. Een groot marktaandeel in deze categorie vormen de sexinformatiediensten (adult content).

De belangrijkste Zakelijke Diensten zijn e-mail en toegang tot bedrijfsnetwerken. Andere diensten in deze categorie hebben betrekking op het raadplegen van lijsten met informatie zoals adressen en plaatjes en verder het beheer van elektronische agenda's.

Tot de categorie Persoonlijk Ondernemen rekenen we diensten waarmee individuele personen financiële transacties uitvoeren. Voorbeelden hiervan zijn het overboeken van geld tussen rekeningen, bijvoorbeeld het gebruik van girofoon en het doen van betalingen vanuit een soort elektronische portemonnee (*wallet*) die op een PC of mobiel toestel is geïnstalleerd. Per betaaldienst zal de mate van security verschillend zijn. Bij betalingen van grote bedragen zal het hogere incasso-risico natuurlijk afgedekt moeten worden door een strengere vorm van security (lees: complexere cryptografische technieken). Verder behoren tot deze categorie ook financiële groepsdiensten zoals het opwaarderen van de prepaidtegoeden van de kinderen vanuit het vaste telefoonabonnement van één van de ouders. Veel diensten uit de financiële categorie zijn gekoppeld aan diensten uit de andere categorieën. Zo zijn bijvoorbeeld veel infotainmentdiensten gekoppeld aan betaaldiensten, zie [2].

4. Aanbieden van diensten aan klanten

Telecommunicatiediensten kunnen niet zomaar door klanten gebruikt worden. Voordat de klant toegang krijgt zal hij eerst zijn identiteit kenbaar moeten maken. Dan kan de dienstaanbieder eerst verifiëren of de klant is wie hij zegt te zijn en of hij geautoriseerd is om de dienst te gebruiken. Deze verificatiestappen heten respectievelijk *authenticatie* en *autorisatie*. Een andere belangrijke stap bij het aanbieden van diensten is het registreren van hoe klanten precies de diensten gebruiken. Dit zogenaamde *accountingproces* is een belangrijke bron van gegevens voor de verrekening met de klant van het dienstgebruik. In totaal spreken we van een *AAA-proces* (Authenticatie, Autorisatie, Accounting).

De voorbeelden hieronder zijn gekozen uit de mobiele wereld. Immers, juist bij mobiele netwerken spelen AAA-processen een belangrijke rol vanwege de mobiliteit van de gebruikers, het gebruik van de radioweg voor transmissie en het niet nagelvast zijn van mobiele toestellen. Er is zo des meer kans op het aannemen van een andere identiteit door gebruikers. Bij elk van de AAA-processen spelen wiskundige methoden en technieken een rol.

4.1. Case: authenticatie en autorisatie

Het authenticatie- en autorisatieproces kent twee fasen, namelijk toegang krijgen tot het mobiele toestel en daarna toegang krijgen tot het mobiele netwerk.

Toegang tot het mobiele toestel

Dit is een eenvoudige stap. Voordat een gebruiker zijn mobiel toestel kan gebruiken dient hij eerst een Persoon Identificatie Nummer (PIN) in te voeren. Dit nummer wordt vergeleken met de code die de gebruiker ooit in zijn toestel op de zogenaamde SIM-kaart (klein chipkaartje dat een klant krijgt van zijn mobiel belbedrijf) heeft opgeslagen. De PIN dient alleen maar om lokaal de toegang tot het mobiele toestel te autoriseren en wordt niet over de radioweg gezonden.

Toegang tot het mobiele netwerk

Na PIN-controle zal verder voordat een klant diensten kan gebruiken, de identiteit van de klant door middel van zijn SIM-kaart gecheckt worden. Basis van dit authenticatieproces is het rekenalgoritme A3 dat zowel in de SIM-kaart als in een beheersysteem (het Authenticatie Centrum, AUC) van het mobiele netwerk bevindt. Het A3-algoritme is een zogenaamd *one-way* algoritme.

Definitie

Een one-way algoritme

$$H(m) = h$$

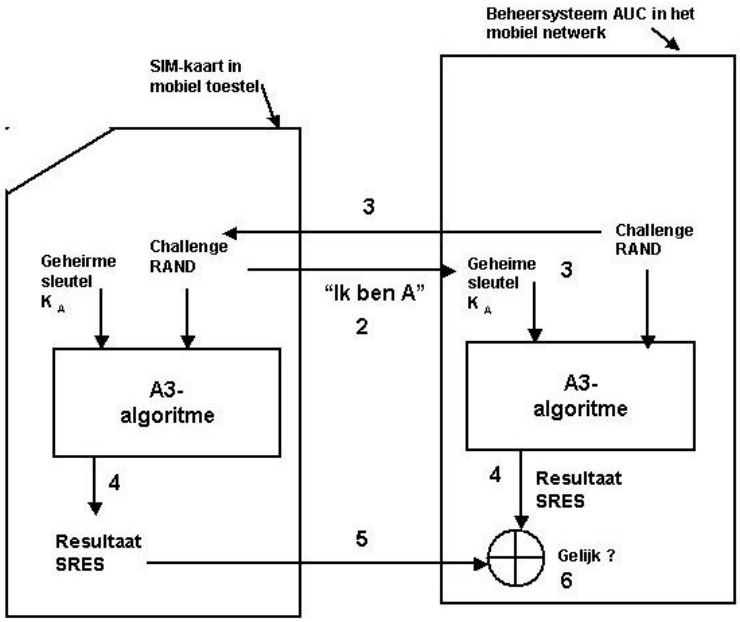
berekent aan de hand van een input m op eenvoudige en snelle wijze het resultaat h. De weg terug, het berekenen van de input m bij een gegeven resultaat is praktisch onmogelijk. Bovendien geldt dat het praktisch onmogelijk is om bij een input m een andere input m' , $m \neq m'$ te vinden met $H(m) = H(m')$.

Het authenticatieproces verloopt in een aantal stappen:

1. Na de PIN-controle zoekt het mobiele toestel automatisch contact met het mobiele netwerk.
2. De identiteit A die in de SIM-kaart zit, wordt door het toestel automatisch naar het mobiele netwerk gestuurd. In de SIM-kaart bevindt zich ook een geheime sleutel K_A.
3. Het beheersysteem AUC zoekt vervolgens in een klanteninformatiesysteem de bij A behorende geheime sleutel K_A. In het AUC wordt bij deze gebruikerssessie random een 128-bitsgetal RAND gegenereerd die het mobiele netwerk op haar beurt weer naar het toestel stuurt.
4. Zowel het AUC als de SIM-kaart berekenen met het A3-algoritme aan de hand van de input RAND en K_A het resultaat SRES (Signed RESult).
5. Het mobiel toestel stuurt het berekende getal SRES naar het mobiele netwerk.
6. In het AUC wordt het dáár berekende getal SRES vergeleken met het ontvangen getal SRES. Indien beide getallen overeenkomen krijgt het mobiele

toestel (gebruiker A) toegang tot het netwerk om de beschikbare mobiele diensten te gebruiken. Bij elk gebruik van een dienst kan deze authenticatieprocedure opnieuw doorlopen worden.

Door het one-way karakter van het A3-algoritme is het gebruik van de radioweg voor het versturen van de identiteit A en de getallen RAND en SRES zonder veel risico. Omdat het mobiele toestel door het netwerk uitgedaagd wordt om SRES te berekenen heet het getal RAND hier een '*challenge*'. Voor een zeer lezenswaardige inleiding over het beveiligd toegang krijgen tot netwerken en systemen, zie [8].



Figuur 6. *Authenticatie van mobiele gebruiker*

Nadat de gebruiker in het netwerk bekend is geworden vindt er nog een nadere autorisatie plaats. Zo wordt onder andere voor prepaid gebruikers gecheckt of er voldoende beltegoed is.

4.2. Accounting

Bij elk telefoongesprek, of meer in het algemeen, bij elke telecommunicatiedienst die een klant gebruikt, worden in het netwerk gebruiksgegevens vastgelegd over, onder andere, wie de dienst gebruikt, op welk tijdstip, hoe lang de sessie duurt en specifieke technische netwerkgegevens. We spreken hier van *accounting* of *metering*. (Opmerking: de term accounting wordt in andere contexten ook wel gebruikt om aan te duiden dat er tussen twee partijen een verrekening plaats vindt, zoals bij internationaal belverkeer.)

Een belangrijk aspect van accounting is het vastleggen van informatie over wat gebruikers doen in het telecommunicatienetwerk om de afgenomen diensten te kunnen verrekenen. De geregistreerde gebruiksgegevens zijn echter bovendien uitgangspunt voor een groot aantal andere bedrijfsinterne beheerprocessen, zoals het dynamisch configureren van netwerken naar actuele behoeften, het tijdig signaleren van noodzakelijk onderhoud, het beheren van service-afspraken met klanten en het maken van data-analyses voor onder andere fraudebestrijding, prijsstelling en marketing.

De accountinggegevens worden op allerlei plaatsen in het netwerk gegenereerd. Dat kan zijn in specifieke netwerkelementen zoals een centrale (switch), router, of gateway maar ook in specifieke servers die deel uitmaken van dienstenplatformen voor het leveren van informatiediensten. Bij een gewoon telefoongesprek wordt na afloop van het gesprek softwarematig een blok van gegevens aangemaakt, een Call Detail Record (CDR). Dit gespreksrecord is ongeveer 150 bytes groot en bevat, onder andere, het nummer van degene die belt (A-nummer), het gebelde nummer (B-nummer), het begintijdstip en gespreksduur. Bij accounting van telefoniediensten gaat het om gigantische hoeveelheden informatie die uiteindelijk verwerkt moeten worden tot nota's voor klanten. Naar schatting worden er in Nederland dagelijks alleen al meer dan 100 miljoen gewone, mobiele telefoongesprekken en SMS-jes verrekend.

Geldt de accounting van telefoniediensten al als een behoorlijk omvangrijk dataverwerkingsproces, het registratieproces bij nieuwe generatie telecommunicatiediensten is nog complexer. Ten eerste worden bij het gebruik van zo'n dienst vaak tegelijkertijd op meerdere plaatsen in het netwerk gegevens aangemaakt. Immers bij deze, meestal internetachtige diensten zijn vaak simultaan meerdere dienstaanbieders betrokken die allemaal een bepaalde dienstcomponent leveren, zie [13].

Verder vraagt de aard van de dienst doorgaans om registratie van geheel nieuwe record-parameters. Zo is de registratie van verstreken tijd niet altijd meer aan de orde maar is juist de hoeveelheid getransporteerde bits of de waarde van de opgehaalde beltoon relevant.

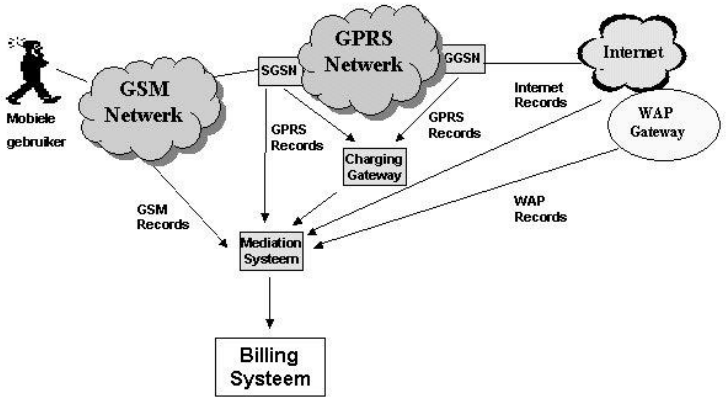
Ook is de registratie complexer omdat bij deze nieuwe generatie van diensten niet slechts één gegeven wordt geregistreerd maar een continue stroom van internetgegevens (events en loggings) wordt gegenereerd. Deze stroom aan gebruiksgegevens kan oplopen tot tientallen gegevens per seconde.

Verder geldt in het bijzonder voor mobiele diensten dat de klant deze overal wil gebruiken en hij dus ook terechtkan komen in de netwerken van andere (buitenlandse) telecommunicatieaanbieders. Dit kan leiden tot ingewikkelde accountingsituaties. Indien bijvoorbeeld een mobiele abonnee in het buitenland belt (we spreken van *roaming*), wordt in dat bewuste buitenlandse netwerk een call record aangemaakt dat vervolgens naar het thuisnetwerk van de mobiele beller wordt gestuurd.

Nadat de accountinggegevens zijn gegeneerd worden deze vervolgens centraal in het netwerk verzameld om daarna uiteindelijk in het *facturering-* of *billingsysteem* tot nota's verwerkt te worden.

Hieronder, in Figuur 7, staat ter illustratie een schematisch overzicht van

de diverse plaatsen in het huidige mobiele GPRS-netwerk (General Packet Radio Service) waar accountinggegevens worden gegenereerd. Het GPRS-netwerk bestaat uit het GSM-netwerk waaraan een pakketgeschakeld datanetwerk is toegevoegd. Het GPRS-netwerk vormt daarmee een generatie mobiel netwerk (2.5G) tussen GSM en 3G in.



***Figuur 7**. Accounting van GPRS-diensten*

Bij GPRS-diensten worden er op minstens vijf verschillende netwerklocaties gebruiksgegevens in de vorm van call records aangemaakt. Te weten:

1. De GGSN (*Gateway GPRS Support Node*); informatie over de verbindingen vanuit het GPRS-netwerk met het internet.
2. De SGSN (*Serving GPRS Support Node*); registratie van het gebruik van het GPRS-netwerk.
3. De *Charging Gateway*; gecombineerde gegevens van de informatie uit de SCSN en GGSN
4. Internet; registratie van gebruik van specifieke internet-componenten
5. *WAP* gateway (Wireless Access Protocol); registratie van het gebruik van specifieke informatie voor mobiele gebruikers.

4.3. Verzamelen van accountinggegevens

De accountinggegevens die op diverse plaatsen in een telecommunicatienetwerk zijn geregistreerd worden door een centraal verzamelsysteem bij elkaar gebracht om ten slotte in een billingsysteem tot rekeningen voor klanten verwerkt te worden. Het verzamelsysteem, ook wel *Mediation* geheten, verzamelt niet alleen de gegevens maar checkt deze ook nog op formaatfouten, corrigeert ze eventueel en brengt alle gegevens op een uniform formaat zodat de verdere verwerking vergemakkelijkt wordt. Het verzamelen kan op twee manieren gebeuren. Ten eerste kunnen de netwerksystemen die gegevens hebben gegeneerd deze zelf opsturen naar de Mediation, of de Mediation zelf vraagt om de gegevens, het zogenaamde pollen. Afhankelijk van de soort verrekening die door klanten of

aanbieder gekozen wordt zal een verzamelwijze worden gekozen. Hier wordt verder ingegaan op het pollen van de accountinggegevens.

Pollen van accountinggegevens
Er wordt uitgegaan van het verzamelsysteem Mediation, dat de verschillende netwerklocaties om de beurt vraagt hun gegenereerde accounting- of gebruiksgegevens op te sturen. Dit *pollen* van de verschillende netwerklocaties kost echter tijd. Deze tijd is, onder andere, afhankelijk van de hoeveelheid gegevens die in een bepaalde periode op deze locatie worden gegenereerd. We zoeken nu naar een efficiënt polling-schema waardoor in het billingsysteem zo actueel mogelijk het gebruik (lees: saldo) per klant bijgehouden kan worden. Daar de gegevens van de Mediation direct doorgestuurd worden naar het billingsysteem, definiëren we de volgende doelstelling:

Doelstelling
Vind een polling-schema zodanig dat de accountinggegevens zo weinig mogelijk verouderd in de Mediation terechtkomen.

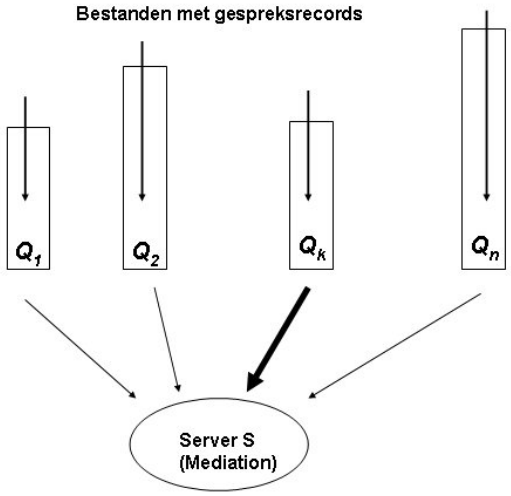
Veel kwantitatieve aspecten van telecommunicatienetwerken- en systemen worden bestudeerd als toepassingsgebied van de theorie van wachtrijen en prestatieanalyse, zie [18]. De algemene theorie van pollingmodellen is hierbinnen weer een geheel eigen wiskundige discipline, zie bijvoorbeeld [19, 20]. Het klassieke pollingmodel gaat uit van een aantal wachtrijen waar klanten arriveren die door een gemeenschappelijke bediende om de beurt bezocht worden. Van belang is nu de volgorde van de wachtrijen die de bediende afloopt, *het pollingschema*. Daarnaast is de *service discipline* van belang hoeveel klanten bediend worden tijdens een bezoek en is van belang de tijd die het kost om te schakelen tussen de verschillende wachtrijen.

Er zijn diverse manieren om een polling-schema voor accountinggegevens op te stellen. Een eenvoudige manier is om cyclisch elk netwerklocatie aan te doen en alle daar gegenereerde gegevens op te halen. De volgorde van te bezoeken locaties kan vast liggen (*statische polling*) of voortdurend aangepast worden. Zo kunnen we bijvoorbeeld na elke pollingsessie van een locatie de volgende te pollen locatie bepalen (*dynamische polling*) of pas een nieuwe tabel met volgorden opstellen na het doorlopen van de huidige pollingtabel (*semi-dynamische polling*). In het algemeen geldt dat de gemiddelde wachttijd (*veroudering*) van files bij statische polling-schema's groter is dan bij semi-dynamische en dynamische schema's.

Hier bespreken we echter specifieke statische polling-schema's waarbij het polling-schema bestaat uit *vaste* tijden waarop de locaties met gebruiksgegevens worden aangedaan. Dit zijn de zogenaamde *Fixed Time Polling-schema's*. Deze schema's zijn van belang omdat mediationsoftware soms alleen zulke schema's ondersteunen. Fixed Time Polling-schema's worden ook in bredere context gebruikt op het gebied van communicatieprotocollen, computerarchitecturen, roostering van personeel en taken en dienstregelingen in het openbaar vervoer gebruikt.

4.4. Case: Fixed Time Polling-schema's

We volgen de notaties en concepten uit [5, 6]. Uitgangspunt is een server S die de wachtrijen $Q_1, \ldots, Q_n$ bedient. In de context van het verzamelen van accountinggegevens is S de Mediation en zijn de Q_i de netwerklocaties waar de gebruiksgegevens opgehaald moeten worden, zie Figuur 8. De gegevens worden niet per stuk opgehaald maar per bestand (*file*). Als de gespreksrecords binnenkomen worden deze eerst in bestanden opgeslagen. Als het bestand een vaste lengte heeft bereikt, wordt het bestand afgesloten en is vervolgens beschikbaar om gepold te worden. De klanten in de wachtrijen Q_i zijn zo deze afgesloten databestanden.



Figuur 8. *Netwerklocaties Q_k wordt gepold*

We veronderstellen dat de wachtrijen Q_i oneindige opslagcapaciteit hebben. Immers in de praktijk hebben de netwerklocaties, meestal centrales, een grote opslagcapaciteit welke voldoende is voor meerdere dagen.

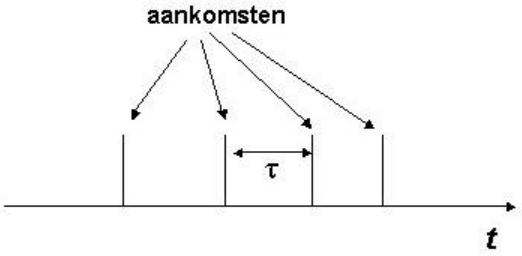
Verder veronderstellen we dat gesprekken op een centrale aankomen volgens een *Poisson-proces*, zie Figuur 9.

Definitie

Een stochastisch proces van gebeurtenissen (aankomsten) heet een Poisson-proces met intensiteit λ als geldt dat voor elke $t > 0$ de stochastische variabele X die het aantal aankomsten telt in T seconden, een Poisson-verdeling heeft met parameter $\alpha = \lambda T$.

Dat wil zeggen, voor de kansen p_i geldt

$$p_i = P\{X = i\} = \frac{\alpha^i e^{-\alpha}}{i!} \text{ voor i} = 0, 1, 2, \ldots$$



Figuur 9. *Een Poisson-aankomstproces*

Simpel is na te gaan dat het verwachte aantal aankomsten in T seconden gelijk is aan

$$EX = \sum_{i=0}^{\infty} ip_i = \alpha = \lambda T.$$

Equivalent met de eis dat X een Poisson-verdeling heeft, is de eis dat de tijdsintervallen Y tussen de aankomsten van het stochastische proces (zie Figuur 9) onafhankelijke stochastische variabelen zijn met een negatief-exponentiële verdeling gegeven door de kansdichtheid

$$f_Y(\tau) = \lambda e^{-\lambda\tau} \text{ voor } \tau \geq 0.$$

Het is eenvoudig na te gaan dat de verwachte waarde EY gelijk is aan

$$EY = \int_0^{\infty} \tau f_Y(\tau) d\tau = \frac{1}{\lambda}.$$

De aankomsttijden op de netwerklocaties van deze accountingbestanden worden bepaald door de tijd die het kost om een bestand te vullen met de geregistreerde gespreksrecords. De bedieningstijd van een wachtrij hangt af van de lengte van deze bestanden en van de transportsnelheid tussen de netwerklocatie en de Mediation.

We nemen aan dat de server S de wachtrijen volgens de *gated* procedure bedient: alle bestanden die tijdens het begin van een pollingsessie afgesloten waren op deze netwerklocatie, worden opgehaald. Merk op dat voor een pollingtabel met vaste tijden hierdoor wel eens een conflict kan ontstaan indien de server nog niet klaar is bij een wachtrij. In de praktijk gaan we ervan uit dat de Mediation altijd meerdere communicatiekanalen heeft en dat dit probleem zich dus niet voordoet.

Definitie
Nummer de te pollen netwerklocaties $Q_1, \ldots, Q_n$ van $1, \ldots, n$. Een Fixed Time Polling-schema is een paar (P, T) met de vector $P = (P_1, \ldots, P_m)$, de polling-

Index pollings-tabel (i)	Netwerk-locatie (P_i)	Tussen-bezoektijd (T_i)
1	3	T_1
2	1	T_2
3	4	T_3
4	3	T_4
5	2	T_5
6	3	T_6
7	4	T_7
8	1	T_8
9	2	T_9
10	4	T_{10}

Figuur 10. *Voorbeeld van Fixed Time Polling-tabel* ($m = 10, n = 4$)

tabel van te pollen netwerklocaties, en de vector $T = (T_1, \ldots, T_m)$ de tussenbezoektijden van het polling-schema. Hier is $P_k \in \{1, \ldots, n\}$ en T_k is de tijd tussen de start van het k^{e} bezoek en de start van het $(k+1)^{\text{e}}$ bezoek voor $k = 1, \ldots, m$. Na m keer is de gehele pollingtabel doorlopen en wordt de tabel weer van voren af aan opnieuw doorlopen. Op basis van de tussentijden T_k kunnen zo vanaf een gegeven begintijdstip de exacte tijdstippen van het polling-schema worden vastgesteld.

Voor een voorbeeld, zie Figuur 10. Er zijn 4 netwerklocaties die volgens de tabel gepold worden. Zo wordt bijvoorbeeld de locatie Q_3 per cyclus 3 keer gepold en wel op de 1e, 4e en 6e keer.

Stel λ_i $(i = 1, \ldots, n)$ is de aankomstintensiteit bij wachtrij i. EW_i is de gemiddelde wachttijd (exclusief bedieningsduur) van een willekeurige klant bij wachtrij i. Er geldt dat voor de bezettingsgraad ρ_i van wachtrij $i, \rho_i = \lambda\beta_i$, waarbij β_i de verwachte bedieningsduur is van een klant in i. Definieer λ de totale aankomstintensiteit, dan is

$$\lambda = \sum_{i=1}^{n} \lambda_i.$$

Voor EW, de gemiddelde totale wachttijd voor een willekeurige klant, geldt dat

$$EW = \sum_{i=1}^{n} \frac{\lambda_i}{\lambda} EW_i$$

Zij C de totale cyclustijd die de server nodig heeft om één keer de gehele pollingtabel te doorlopen. Verder definiëren we de deelcyclustijden C_k $(k = 1, \ldots, m)$ als de k^{e} deelcyclustijd, de tijd tussen de start van het k^{e} bezoek aan de netwerklocatie P_k en de start van het volgende bezoek aan dezelfde locatie. Hier geldt dat $m + 1$ gelijk is aan 1, etc.

Er geldt dan dat de cyclustijd C gelijk is aan het totaal van de bezoektijden

$$C = \sum_{l=1}^{m} T_l.$$

Eenvoudig is in te zien dat de cyclustijd gelijk is aan het totaal van alle deelcyclustijden corresponderend bij elke keuze van een netwerklocatie i

$$C = \sum_{\{k:P_k=i\}} C_k \text{ voor i} = 1, \ldots, \text{n}.$$

Ten slotte, definieer de bezoekfrequenties m_i aan locatie i

$$m_i = |\{k : P_k = i\}| \text{ voor i} = 1, \ldots, \text{n}$$

Er geldt dan voor de lengte m van de pollingtabel

$$m = \sum_{i=1}^{n} m_i$$

Gemakkelijk valt na te gaan dat voor de pollingtabel uit Figuur 10 geldt $m_1 = 2, m_2 = 2, m_3 = 3$ en $m_4 = 3$.

Probleemstelling
Construeer een Fixed Time Polling-schema zodanig dat EW, de gemiddelde totale wachttijd voor een willekeurige klant minimaal is.

Oplossing
We zoeken een geschikt paar (P, T). Dat wil zeggen we bepalen de lengte m van de tabel, de bezoekfrequenties voor elke wachtrij, de bezoekvolgorde en de bezoektussentijden.

Dit optimaliseringprobleem is uiterst complex, echter door middel van een heuristische aanpak, uitgaande van een aantal benaderingen, blijkt in drie opeenvolgen de stappen een bijna optimale oplossing gevonden te kunnen worden.

1. Bepaling van de bezoekfrequenties $m_1, \ldots m_n$ voor elke wachtrij $1, \ldots, n$.

Op basis van de volgende benadering van de gemiddelde totale wachttijd EW

$$EW = \approx C \sum_{i=1}^{n} \frac{\lambda_i(1+\rho_i)}{2m_i}$$

vinden we de bezoekfrequenties.

2. Bepaling van de bezoekvolgorde van de netwerklocatie $P_1, \ldots, P_m$.

Om de bezoekvolgorde vast te stellen, wordt gebruik gemaakt van een methode die op basis van de Gulden Snede zo goed mogelijk de bezoeken aan de verschillende wachtrijen gelijk spreidt. Dat wil zeggen, zij X_k het aantal bezoeken tussen de start van het k^{e} bezoek aan netwerklocatie P_k en het volgende bezoek aan deze netwerklocatie, dan verdeelt deze methode zo goed mogelijk de bezoeken opdat de getallen X_k gelijk zijn. Voor meer details zie [7].

3. Bepaling van de tussenbezoektijden $T_1, \ldots, T_m$.

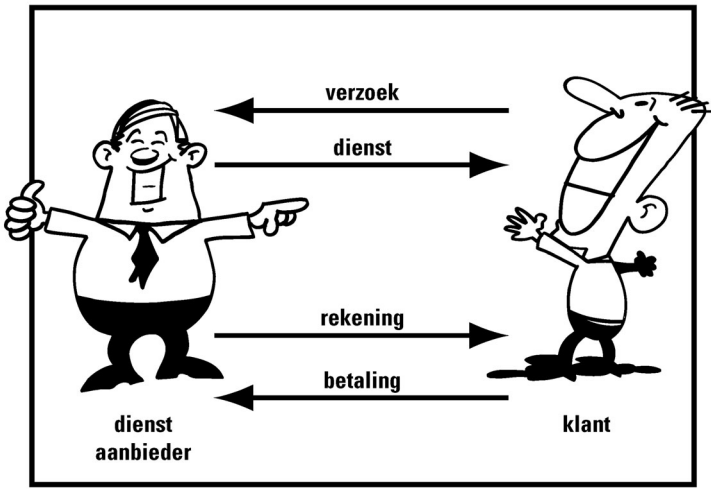
Ten slotte, leiden we de tussenbezoektijden af uit een stel lineaire vergelijkingen waarvan de coëfficiënten bepaald zijn door de aankomstintensiteiten λ_i, de transmissietijden en de omschakeltijden tussen de wachtrijen i.

Uit numerieke simulaties met echte data van een aantal telefooncentrales is gebleken dat ten opzichte van de standaard mediationsoftware, het volgens bovenstaande stappen geconstrueerde Fixed Time Polling-schema, een reductie van 80% van de gemiddelde totale wachttijd opleverde.

5. Verrekenen van telecommunicatiediensten (billing)

5.1. Het verrekeningsproces

Verrekening of *billing* is de financiële afhandeling van het leveren van diensten aan klanten, zie Figuur 11.



Figuur 11. *Het verrekenen van diensten*

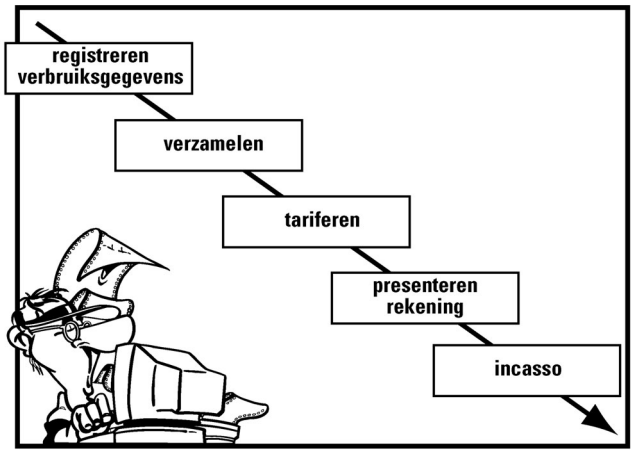
De afrekening met klanten is voor bedrijven hét belangrijkste gereedschap om de behoeften van klanten te begrijpen, en om in te schatten hoe goed daaraan voldaan wordt, zie [11].

Typerend voor de billing van telecommunicatiediensten is dat van veel gebruikers, verspreid over een uitgestrekt netwerk, gebruik makend van veel ver-

schillende randapparaten, de gebruiksgegevens worden geregistreerd. Dit is precies de situatie die we ook aantreffen bij veel andere bedrijfstakken. Denk daarbij bijvoorbeeld aan de afrekening van utiliteitsdiensten zoals gas-, water- of energiebedrijven, de betaling van reizen en ondersteunende diensten bij openbaarvervoerbedrijven en de verrekening door de overheid van het gebruik van tolwegen.

In het totale billingproces zijn een vijftal voor de hand liggende stappen te onderscheiden, zie Figuur 12. De eerste twee zijn boven al besproken.

- de registratie van het gebruik (accounting)
- het verzamelen van alle geregistreerde verbruiksgegevens,
- het tariferen van de gegevens,
- het op de nota plaatsen van de getarifeerde gegevens,
- het presenteren van de nota aan de klanten en tot slot
- het incasseren van de verschuldigde bedragen.

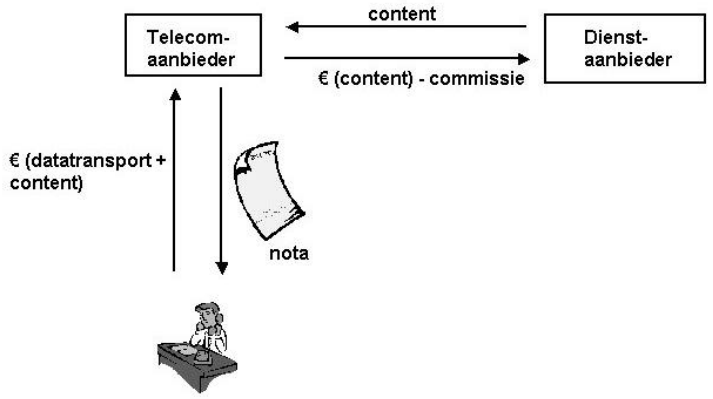


Figuur 12. *De verschillende stappen van het billing- of verrekeningsproces*

Het verrekeningsproces is in totaal een complex bedrijfsproces doordat er een hele verzameling van technische en organisatorische processen mee gemoeid is, die gecoördineerd uitgevoerd moeten worden. Deze deelprocessen strekken zich uit door alle beheerlagen van het bedrijf en hebben directe interactie met soortgelijke processen in het domein van klanten en businesspartners.

Juist bij de nieuwe generatie telecommunicatiediensten zullen business partners nauwgezet op billinggebied moeten samenwerken. Dit heeft tot gevolg dat allerlei billingprotocollen en - interfaces gestandaardiseerd moeten worden en op elkaar afgestemd. Een voorbeeld van zulke samenwerking is de huidige veel voorkomende situatie waarbij dienstaanbieders (De Telegraaf, ANWB, KNMI, etc.) meeliften op de nota van de grote telecommunicatiebedrijven. De telecombedrijven hebben reeds een verrekeningsrelatie met klanten. Op deze nota

wordt naast het aandeel van het telecombedrijf (netwerkaccess en datatransport) ook het informatie- of contentdeel van de dienst de klant in rekening gebracht en geïncasseerd. Na aftrek van commissie, immers het telecombedrijf draagt het incassorisico, worden de opbrengsten van de informatie aan de dienstaanbieder gegeven, zie Figuur 13.



Figuur 13. *Verrekening van telecommunicatiediensten*

Als grote billinguitdaging geldt het kunnen correleren van de gebruiksgegevens van het datatransportdeel van een dienst en het informatiedeel. Anders gezegd, precies te weten welke bitten bij welke dienst hoort. Met deze functionaliteit zouden aanbieders aantrekkelijke prijsplannen aan klanten kunnen aanbieden, zoals happy hours waarbij klanten slechts een stukprijs betalen voor een opgehaald plaatje maar niet het benodigde datatransport. Het correlatieprobleem is momenteel een open probleem op het gebied van gegevens- en procesmodellering. Dit probleemgebied vereist nader onderzoek naar simultane internet accountingprocessen bij aanbieders van diensten en telecomoperators. In het bijzonder wordt gezocht naar snelle matchingsalgoritmen voor grote hoeveelheden data.

5.2. Soorten van verrekeningsmethoden

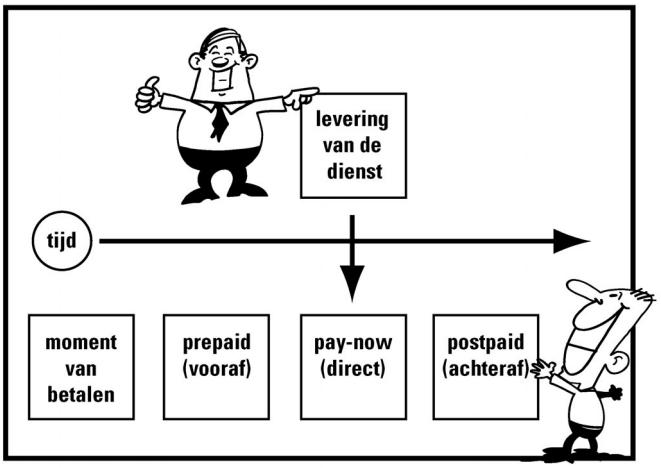
Moment van betalen

Er zijn diverse soorten van verrekening mogelijk afhankelijk van de betalingswijze door klanten, vooraf aan het daadwerkelijke gebruik van de dienst (*prepaid*), direct (*pay-now*) of achteraf (*postpaid*), zie Figuur 14.

In gang gezet door Telecom Italia deed "prepaid-bellen" vanaf 1995 zijn intrede in de telecomwereld. Mobiele operators boorden in rap tempo nieuwe marktsegmenten aan door naast een gewoon standaardabonnement de mogelijkheid te bieden om te bellen op basis van een beltegoed. Prepaid-bellen is erg cultureel bepaald, zo wordt in Scandinavische landen en in Japan nauwelijks prepaid gebruikt. Doordat belbedrijven aantrekkelijke abonnementen en toestellen bieden bij de invoering van nieuwe telecommunicatiediensten lijkt het

aandeel prepaid-bellers in het algemeen af te nemen.

Voor de volledigheid van dit overzicht geldt dat er ook telecommunicatiediensten zijn waar klanten niet voor betalen. Voorbeelden van deze gratis diensten zijn de gratis 0800-telefoonnummers en veel gratis internetdiensten. De kosten van het gebruik worden hierbij door derden betaald. Voor het 0800-nummer is dat de eigenaar van het nummer en voor veel gratis internetdiensten zijn dat de adverteerders.



Figuur 14. *Soort verrekening is afhankelijk van het moment van betalen*

Mobiele- en internetbetalingen

Onderdeel van het billingproces is het incasso-proces waarbij de klant gevraagd wordt om te betalen. Het gebied van mobiele- en internetbetalingen wordt gekenmerkt door telkens opduikende nieuwe initiatieven, proeven en systemen, zie voor de aardigheid [9]. Er wordt driftig gezocht naar eenvoudige, veilige en goedkope betalingssystemen. Tot nu toe is er nog steeds niet een algemeen geaccepteerd elektronisch betaalmiddel op de markt verschenen. Er is ondertussen een kip-ei-situatie ontstaan waarbij on line aanbieders geen grote omzetten hebben door het ontbreken van een grootschalig betaalsysteem. En er is geen alom beschikbaar betaalsysteem omdat niemand wil investeren in de bijbehorende introductie. Bij billing via internet, ook wel *e-billing* genaamd, kunnen klanten hun gebruiksgegevens en nota's via een vaste of mobiele terminal direct opvragen en eventueel betalen. Klanten hebben hierdoor sneller, altijd en overal inzage in hun telecommunicatiekosten.

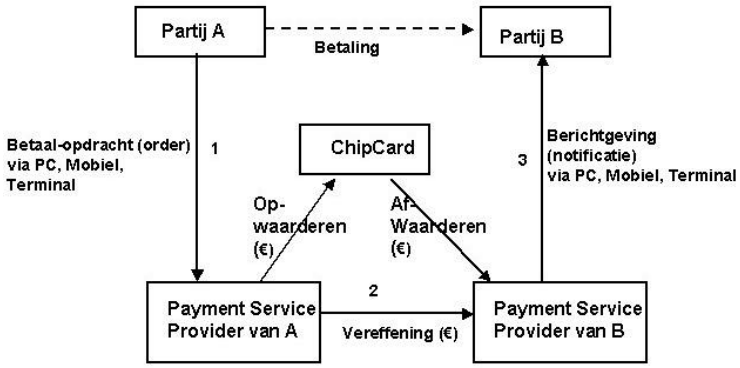
Het betalingsproces omvat een aantal stappen, zie Figuur 15. Voor meer details en een goede inleiding over beveiliging van netwerken en diensten, zie [21].

Stel dat Partij A geld wil overmaken naar Partij B op elektronische wijze via een mobiel toestel, een PC of een geldterminal in een winkel of op straat.

1. Partij A geeft een betaalopdracht door middel van een van de genoemde terminals aan zijn *Payment Service Provider*. Veelal is dit zijn eigen bank maar het kan ook een andere gespecialiseerde financiële partij zijn.
2. De Payment Provider van A zoekt nu contact met de Payment Provider van de begunstigde Partij B. De Payment Provider van B verhoogt nu het banksaldo van B met het bedrag dat A wil overmaken.
3. Ten slotte ontvangt Partij B een bevestiging van zijn Payment Provider dat Partij A een bepaalde hoeveelheid geld naar hem heeft overgemaakt.

Een andere manier om elektronisch geld over te maken is het gebruik maken van elektronisch geld (*digital money*) dat gebruikers op bijvoorbeeld een chipcard hebben staan.

Het gebruik van elektronisch geld en mobiel- internetbetalingen vereist dat deze financiële methoden beschermd zijn tegen fraudeurs en hackers. Het moet bijvoorbeeld onmogelijk zijn om nieuw geld aan te maken, geld dubbel uit te geven of geld van andermans rekeningen te plukken. Hiervoor zijn uitgebreide beveiligingstechnieken ontwikkeld die sterk gebaseerd zijn op wiskundige cryptografische technieken, zie onder andere [8, 21].



Figuur 15. *Afhandeling van elektronische betalingen via Payment Service Providers*

Hieronder wordt nader ingegaan op het gebruik van openbare-sleutelsystemen (*public-key cryptosystem*) om bij de betalingsstappen 1, 2 en 3 te verifiëren dat het verzonden betalingsbericht niet veranderd is en van de genoemde partij afkomstig is. Dit wordt toegelicht aan de hand van het bekende RSA-algoritme. Dit algoritme is genoemd naar de ontdekkers Rivest, Shamir en Adleman (1978) van het bekende onderzoeksinstituut MIT in Boston.

Case: Openbare-sleutel algoritme RSA
Het algoritme is een voorbeeld van de eerder genoemde one-way methoden. De gebruikte grote priemgetallen maken het praktisch onmogelijk om het algoritme te kraken. Het algoritme bestaat uit de volgende stappen.

1. Kies twee grote priemgetallen p en q.
2. Op basis hiervan bereken de getallen $n = p \times q$ en $z = (p-1) \times (q-1)$.
3. Kies een getal d dat relatief priem is met het getal z.
4. Ten slotte, bereken het getal e zodanig dat $e \times d = 1 \mod z$.

Stel nu dat de Payment Service Provider van A, PSP_A een vertrouwelijk betalingsbericht m wil sturen naar de Payment Service Provider PSP_B van B. PSP_A wil dat alleen PSP_B het bericht m kan lezen.

Neem aan dat de lengte van het bericht m kleiner is dan n. Deel anders het bericht m eerst op in kleinere stukken. PSP_A versleutelt nu het bericht m met de openbare coderingssleutel EB van PSP_B door te berekenen

$$c = EB(m) = me(\text{mod } n)$$

Het bericht c wordt nu verzonden door PSP_A naar PSP_B en daar door PSP_B ontcijferd door het toepassen van zijn eigen geheime decoderingssleutel DB

$$DB(c) = cd(\text{mod } n)$$

Door de getaltheoretische eigenschappen van het RSA -algoritme geldt nu

$$DB(c) = DB(EB(m)) = m$$

De Payment Service Provider PSP_B heeft hiermee het oorspronkelijke bericht van PSP_A kunnen ontcijferen.

In de praktijk is het gebruik van RSA behoorlijk langzaam (7,4 kbit/s bij 1024-bits priemgetallen p en q). Het gebruik van een ander algoritme, DES (*Data Encryption Standaard*), levert bij een geschikte implementatie duizenden factoren snelheidswinst op, zie [17].

5.3. Rekenen aan verrekening

Bedrijven zoeken naar kosteneffectieve billingmethoden
Vergeleken met veel andere bedrijfsprocessen is billing een omvangrijk, complex en daardoor duur proces. De meeste kosten voor het leveren van diensten zitten in de kosten van billing. Voor sommige dienstaanbieders is het onmogelijk om bepaalde diensten aan te bieden omdat de kosten voor billing meer zouden bedragen dan het bedrag van de dienst zelf.

Billing vereist dat op een kosteneffectieve manier gebruiksgegevens worden gegenereerd en verzameld. Voor informatiediensten waarbij gebruikers steeds maar kleine hoeveelheden informatie (*content*) opvragen is onderzoek naar geschikte, betaalbare billingmechanismen, de zogenaamde micro payments zeer

gewenst. De huidige situatie leidt tot een dilemma. Enerzijds moeten contentitems juist zeer laag geprijsd zijn willen gebruikers het afnemen, terwijl anderzijds de waarde van het item niet te laag moet zijn wil de registratie en de verwerking van betalingsgegevens nog rendabel zijn.

Mogelijke oplossingsrichtingen zijn het aan gebruikers beschikbaar stellen van elektronische portemonnees en het registreren van gebruik op handhelds in plaats van in het netwerk. Een geheel andere oplossingsrichting is *statistische* billing. Hierbij worden niet precies alle giga-hoeveelheden gebruiksgegevens in netwerken bijgehouden, dat is juist te kostbaar, maar worden, gegeven een zeker betrouwbaarheidsniveau, slechts schattingen gemaakt over het verbruik. Ook kan, in plaats van kostbare real-time transacties, kosten worden bespaard door het opsparen op netwerklocaties van accountinggegevens. Een kenmerkend voorbeeld vormen de voor de detailhandel goedkopere chipbetalingen dan pinbetalingen.

Klanten zoeken naar lage tarieven

Naast dat billing geld kost levert billing juist ook geld op. Marketeers in bedrijven bepalen de hoogte van de tarieven en de verschillende soorten prijsplannen en tariefstructuren. Ook hier zal er veel gerekend moeten worden om de meest geschikte (lees: lucratieve) tarieven, ook in vergelijking met andere aanbieders te vinden. In de praktijk worden veel simulaties met de verschillende voorstellen gedaan op verzamelingen met echt netwerkverkeer (gebruiksgegevens van klanten).

Voor klanten geldt het omgekeerde. Zij zijn ook geïnteresseerd in de tarieven maar willen juist de meest voordelige tarieven bepalen, in het bijzonder tussen de verschillende aanbieders. Om klanten een goed advies te geven, verzamelt de succesvolle internetondernemer Ben Woldering (op 14-jarige leeftijd al begonnen!), alle tariefgegevens van telecommunicatieleveranciers (mobiel, gewone telefonie, internet). Bezoekers van zijn sites (o.a. www.bellen.nl) krijgen zo gemakkelijk antwoorden op vragen welke aanbieder de goedkoopste is.

Literatuur

[1] 4G. Zie website `http://users.ece.gatech.edu/~jxie/4G/#NEW`

[2] T.T. Ahonen (2002). *m-Profits, Making Money from 3G Services*, Wiley.

[3] T.T. Ahonen, J. Barrett (2002). *Services for UMTS, Creating Killer applications in 3G*, Wiley.

[4] R. Bekkers (1999). *GSM in detail*. Kluwer.

[5] S.C. Borst (1994). *Polling Systems*, Proefschrift Katholieke Universiteit Brabant.

[6] S.C. Borst, O.J. Boxma, J.H.A. Harink, G.B. Huitema (1994). Optimization of fixed time polling schemes, *Telecommunication Systems* **3**, 31–59.

[7] O.J. Boxma, H.Levy, J.A. Westrate (1991). Optimization of polling systems. In: *Proc. Performance '90*, (eds.) P.J.B. King, I. Mitrani, R.J. Pooley, North-Holland, Amsterdam, 349–361.

[8] J. van de Craats (1991). *Pasjes en pincodes*, Aramith Uitgevers Bloemendaal.
[9] Electronic Payment Schemes. Zie `http://www.w3.org/ECommerce/roadmap.html`
[10] E. Fledderus (2000). Wiskundig modelleren in Mobiele Telecommuncatie. *ITW Nieuws* **10** (1).
[11] G.B. Huitema (2002). *Van de nota een deugd maken*. Oratie Rijksuniversiteit Groningen. Voor een pdf-file, zie `http://www.research.kpn.com/headline_bekijk.asp?hid=1758`.
[12] R.E. Kooij, R.D. van der Mei, J.G. Beerends (2001). Internet Telefonie, *Natuur & Techniek*, juli, 66–72.
[13] M. van Le, B.J.F van Beijnum, G.B. Huitema (2003). *Flexible Realtime Service Accounting Architecture*, TU Twente.
[14] R.D. van der Mei (2000). Verkeersmodellering van netwerken. *Nieuw Archief voor Wiskunde* **5**, 390–396.
[15] R.D. van der Mei (2001). Wiskunde in Telecommunicatienetwerken, Vacantiecursus 2001, Experimentele Wiskunde. *CWI Syllabus* **49**, 111–123.
[16] R. Meijer (2002). *De Telerevolutie*. Oratie Universiteit van Amsterdam.
[17] RSA Fast (1999). `http://www.rsasecurity.com`, RSA Laboratories.
[18] F.C. Schoute (1989), *Prestatie-analyse van telecommunicatiesystemen*, Kluwer. Voor een pdf-file, zie `http://mmc.et.tudelft.nl/presan/pab.html`.
[19] H. Takagi (1986). *Analysis of Polling Systems*, The MIT Press, Cambridge.
[20] H. Takagi (1990). Queueing analysis of polling models: an update. *Stochastic Analysis of Computer and Communication Systems*, (ed.) H. Takagi, North-Holland, Amsterdam, 267–318.
[21] A.S. Tanenbaum, M. van Steen (2002). *Distributed Systems*. In het bijzonder Hoofdstuk 8 over Security en sectie 8.7 over Electronic Payment Systems. Prentice Hall.
[22] R. Vranken, R.D. van der Mei, R.E. Kooij, J.L. van den Berg (2002). Performance of TCP with Multiple Priority Classes. *Proceedings International Seminar Telecommunication Networks and Teletraffic Theory*, 78–87

CWI Vakantiecursus 2003

pp. 93 – 102

Mannen in snelle pakken – de weerstand van schaatsers

Nando Timmer
DUWIND – Delfts Universitair Windenergie Instituut
Technische Universiteit Delft
e-mail: w.a.timmer@citg.tudelft.nl

1. Inleiding

De International Herald Tribune schreef: "speed skaters make history", USA Today sprak van "the new edge in speedskating". In eigen land kopte de Volskrant "Romme leidt stayers naar nieuw tijdperk" en het NRC Handelsblad jubelde "Nederland zet de schaatswereld op z'n kop met strips". Het is maandag 9 februari 1998, een dag na de 5000 m schaatsen voor heren op de Olympische Winterspelen van Nagano, Japan. De eerste 4 plaatsen van het eindklassement van deze 5 km werden bezet door Nederlanders. Gianni Romme breekt zijn oude wereldrecord met 8,43 seconden. De nummer 2, Rintje Ritsma, haalt 6 seconden van zijn eigen toptijd af en nummer 3, Bart Veldkamp, uitkomend voor België, haalt de eerste olympische schaatsmedaille voor onze Zuiderburen binnen door zijn eigen persoonlijk record met maar liefst 11 seconden te verbeteren. Hij eindigde eveneens onder het oude record van Romme. Eindelijk weer eens goud voor Nederland op de 5 km, de tweede gouden medaille ooit op deze afstand, na die van Ard Schenk in Sapporo 1972. De Nederlanders droegen in Nagano als enigen in het internationale gezelschap zigzag strips op het hoofd en de onderbenen. De twee onderzoekers van de TU Delft, die de strips ontwikkeld hadden, waren even het middelpunt van het universum en de TU Delft zette die week paginagrote advertenties in de grote landelijke dagbladen om zichzelf te promoten met dit prachtige resultaat van Delfts onderzoek. Euforie en scepsis, protesten en triomf, het was er allemaal en een paar stukjes foam op de schaatskleding van de heren en dames coryfeeën van de Nederlandse schaatsploeg speelden daarbij een belangrijke rol (Figuur 1). Kwam het door de strips?

Niemand zal het ooit weten. Er is nu eenmaal geen vergelijkingsmatriaal voorhanden; tijden gereden door dezelfde schaatsers op dezelfde dag onder dezelfde omstandigheden, maar zonder de zigzagstrips. Het is mogelijk dat de strips alleen een placebo-werking hadden, maar het is niet waarschijnlijk. Echter, de helaas te voeg overleden uitvinder van de moderne klapschaats, Gerrit-Jan van Ingen Schenau, heeft eens gezegd: "wetenschappers maken geen kampioenen". Dat is de enige zekerheid; de Nederlandse medaillewinnaars mogen vooral zichzelf op de borst slaan met het succes van Nagano. Toptijden rijdt men niet door "ribbelstrookjes", zoals de strips ook wel genoemd werden, op het pak te plakken. Daarvoor zijn allereerst talent, training, techniek en doorzettingsvermogen noodzakelijke vereisten. Maar de juiste materiaalkeuze speelt daarbij

Figuur 1. *Gianni Romme tijdens zijn recordrace op de 5000m in Nagano 1998*

een niet geheel ondergeschikte rol. In het volgende zullen de achtergronden van de schaatsstrips belicht worden. Vervolgens zal de lijn doorgetrokken worden naar de huidige snelle schaatspakken.

2. De weerstandskracht

Ruwweg 80% van de weerstand die een schaatser ondervindt is luchtweerstand. Enkele procenten winst in deze aërodynamische weerstand vertaalt zich al snel in een betere eindtijd. De luchtweerstandskracht D, uitgedrukt in Newton wordt gegeven door:

$$D = C_d.\frac{1}{2}\rho V^2.S \tag{1}$$

Hierin is ρ de luchtdichtheid in kg/m^3, V de snelheid in m/s, S het frontaal oppervlak in m^2 en C_d de weerstandscoëfficiënt. Een ieder die de Japanse sprinter Shimizu heeft zien schaatsen met zijn neus bijna op het ijs, beseft dat het verkleinen van het frontaal oppervlak directe gevolgen moet hebben voor de snelheid.

De luchtdichtheid speelt eveneens een grote rol bij de eindtijden. Helaas voor de Heerenveense baan Thialf zullen banen op grote hoogte de wereldrecords blijven leveren. De verschillen met laaglandbanen zijn dermate groot, dat dit niet goed gemaakt kan worden door bijvoorbeeld “beter” ijs te maken. In tabel 1 is de luchtdichtheid gegeven voor de banen in Heerenveen en Salt Lake City, met daarbij de baanrecords voor heren. Die van Salt Lake City zijn tevens de huidige wereldrecords. Voor beide locaties is de haltemperatuur 15^oC.

Locatie	hoogte (m)	ρ (kg/m^3)	500 m	1000 m	1500 m	5 km	10 km
Heerenveen	0	1.225	34.83	1.10.41	1.46.63	6.29.31	13.03.40
Salt Lake City	1400	1.035	34.22	1.07.18	1.43.95	6.14.66	12.58.92

Tabel 1. *De relatie tussen de baanhoogte en de baanrecords voor een laag- en een hooglandbaan*

De derde belangrijke parameter die de snelheid bepaald is de weerstandscoëfficiënt C_d. De hoogte van deze parameter toont in welke mate het object dat in de stroming staat "gestroomlijnd" is. Stroomlijn is belangrijk, weten we uit de praktijk. Een vliegtuig (of een auto) moet stroomlijn hebben, anders is de weerstand te hoog en wordt te veel brandstof verbruikt.

3. Drukweerstand

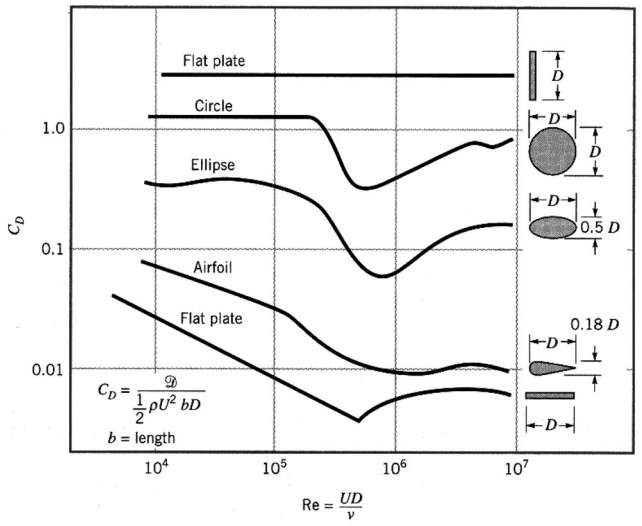
Bij de weerstandscoëfficiënt speelt de vorm van het object een belangrijke rol. Een stomp voorwerp heeft een hogere C_d dan een druppelvormig object. De stroming om lichaamsdelen zoals benen en hoofd heeft wel iets weg van die om bollen en cilinders. Daarom wordt hierna de werking van de strips uitgelegd aan de hand van deze vormen. Van voorwerpen met ronde vormen, zoals bollen en cilinders, varieert de C_d met het Reynoldsgetal. Dit schalingsgetal geeft de verhouding weer tussen de wrijvingskrachten en de mechanische krachten in de stroming. Het Reynoldsgetal wordt gegeven door:

$$Re = V.D/\nu \tag{2}$$

Hierin is V de snelheid, D een karakteristieke lengte, (voor een cilinder of bol is dat de diameter) en ν (spreek uit "nu") de kinemaethische viscositeit van het medium. Deze "nu" is afhankelijk van de temperatuur en de druk in het medium. Voor lucht bij 15 graden Celsius en 1013 mBar is "nu" ongeveer 15×10^{-6}. Als van twee objecten met dezelfde vorm, maar andere afmetingen, in een ander medium, met een andere snelheid het Reynoldsgetal overeenkomt, is het stromingspatroon eveneens hetzelfde. Ze hebben dan dezelfde C_d. In Figuur 2 is het verloop van de weerstandscoëfficiënt met het Reynoldsgetal weergegeven.

4. Het kritieke Reynoldsgetal

Voor niet-gestroomlijnde objecten met ronde vormen blijkt er een significante vermindering van de C_d plaats te vinden bij een bepaald Reynoldsgetal. Dit noemen we het kritieke Reynoldsgetal. De verlaging van C_d, voor cilinders zelfs tot eenderde van de oorspronkelijke waarde, hangt samen met een verandering in het karakter van de stroming rond de cilinder. Het luchtlaagje dat over het oppervlak stroomt noemen we de "grenslaag". Bij lage Re is de grenslaag laminair, dwz netjes gelaagd en rustig. Het nadeel van deze laminaire grenslaag is dat hij -door de wrijving- snel zijn energie kwijtraakt en niet de ronding van de bol helemaal kan blijven volgen. Ergens in de buurt van het dikste punt van de cilinder laat de grenslaag los van het oppervlak en duwt de buitenstroming weg, zodat er een breed "zog" ontstaat. In dit zog zijn de snelheid en de druk



Figuur 2. *Het verloop van de weerstandscoëfficiënt met het Reynoldsgetal voor diverse objecten*

laag. Een en ander is te vergelijken met de "slipstream" achter een vrachtauto. Men wordt als het ware meegezogen.

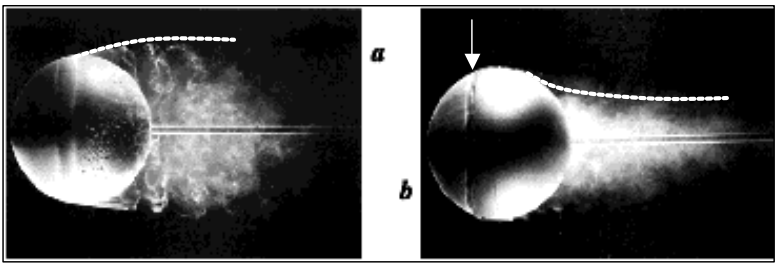
Het resultaat is dat er een groot drukverschil is tussen de voorkant (waar de lucht tegen op botst) en achterkant van de cilinder (waar de lage druk heerst). De cilinder heeft dan een grote "drukweerstand". Deze drukweerstand is veel groter dan de wrijvingsweerstand van de luchtdeeltjes langs het oppervlak

5. Glad is niet altijd beter

Als we het Reynoldsgetal groter maken (hetzij door de snelheid te vergroten bij vaste diameter en omgevingscondities, hetzij door de diameter te vergroten bij vaste snelheid, hetzij door het medium te variëren) tot aan het kritieke Reynoldsgetal, dan verandert het karakter van de grenslaag gaandeweg van laminair naar "turbulent", d.w.z. warrelige en chaotische bewegingen van de luchtdeeltjes van de grenslaag. Hierdoor vindt er energie-uitwisseling plaats met de stroming die eerst nog buiten de grenslaag zat, en niet te maken had met de wandwrijving, zodat de turbulente grenslaag meer energie heeft en veel langer de ronding van de cilinder kan blijven volgen. Het resultaat is dan dat het zog kleiner is en de druk in het zog hoger. De drukweerstand is dan drastisch afgenomen en de totale weerstand van de cilinder ook.

De snelheid waarbij deze overgang van laminair naar turbulent op de juiste manier plaatsvindt hangt af van de grootte van ν en de diameter van de cilinder. Voor cilinders met een diameter van ca. 15 cm ligt deze snelheid op minstens 30 m/s. Er is echter een mogelijkheid deze snelheid te verlagen. Op kunstmatige wijze kan de grenslaag turbulent gemaakt worden door het oppervlak ruw te

maken of door een ander soort verstoring aan te brengen. Dit is in beeld gebracht in Figuur 3.



Figuur 3. *Visualisatie van de stroming om een bol met behulp van rook. Bij (a) laat de laminaire grenslaag los. Bij (b) wordt via een groef langs kunstmatige weg de stroming turbulent gemaakt, zodat het zog kleiner wordt*

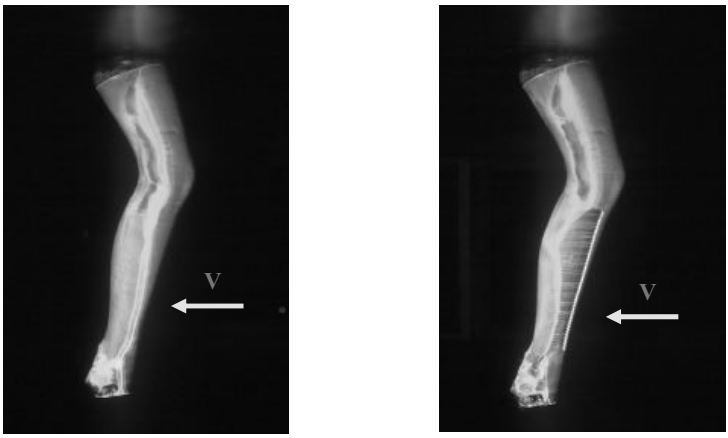
Door een juiste keuze kan zo het kritieke getal van Reynolds verlaagd worden tot dat wat bijvoorbeeld in het schaatsen voorkomt. Ook de putjes in een golfballetje zijn aangebracht met hetzelfde doel: verlaging van de drukweerstand, zodat het balletje verder komt.

6. Zigzagtape

Dat voor de benen en het hoofd van de schaatsers zigzagtape gebruikt werd, verbaast insiders niet. Bij windtunnelproeven aan profielen voor zweefvliegtuigen werd dit tape al langer gebruikt. Ook bij deze profielen komt loslating van de laminaire grenslaag voor. Het zigzagtape bleek door de werveltjes die van de puntjes afkomen een zeer efficiënte manier om de grenslaag turbulent te krijgen.

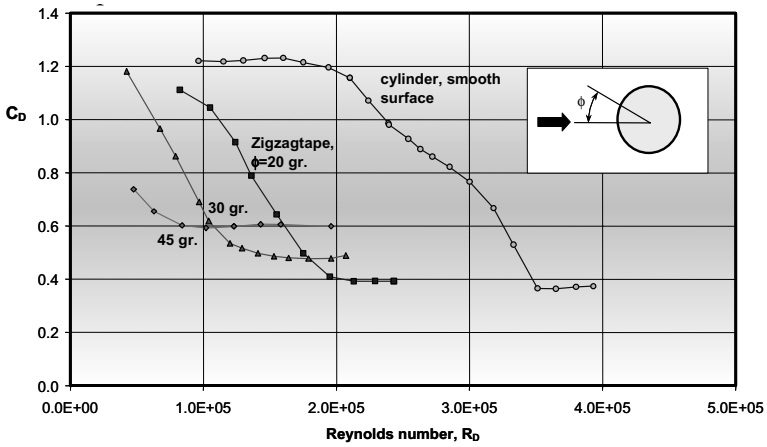
Hoe dat in de praktijk werkt laat Figuur 4 zien. Met behulp van fluorescerende olie is de stroming op het been zichtbaar gemaakt. Links het kale been, rechts het been met zigzagtape. De donkergele lijn op het been links laat de loslaatlijn zien. De stroming laat daar los en van achteren wordt olie aangevoerd door terugstroming in het zog achter het been. Rechts is te zien dat de stroming het been beter omvat. De loslaatlijn ligt een stuk meer naar achteren. Goed zijn ook de werveltjes te zien die van de puntjes van het tape afkomen.

De vraag is nu, waar het tape precies moet worden aangebracht en hoe dik moet het zijn. Daar is met behulp van cilindermetingen informatie over verkregen. Figuur 5 laat de relatie zien tussen de positie van het tape en de resulterende weerstandscoëfficiënt van de cilinder als functie van het Reynoldsgetal. De onderbenen van een schaatser hebben een gemiddelde Reynoldsgetal in de orde van 10^5. Een hoek van 30 tot 45 graden vanaf het voorste punt van het been lijkt daarvoor het meest geschikt. Het blijkt, dat het kunstmatig storen van de grenslaag ook een prijs heeft. Het kritieke Reynoldsgetal komt wel bij een lagere waarde te liggen, maar het minimum van de C_d komt niet



Figuur 4. *Het effect van zigzag tape op de loslaatlijn. Het been links vertoont vroege laminare loslating (de dikke lichte lijn. Het been rechts heeft turbulente loslating ten gevolge van de zigzag strip*

meer zo laag. Variatie van de dikte van het tape geeft een vergelijkbaar verloop te zien.

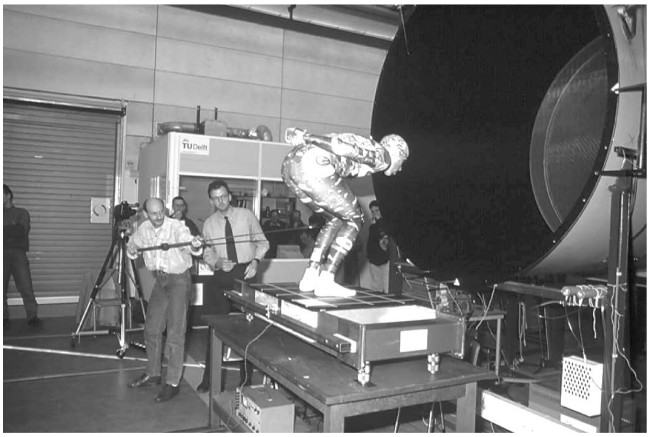


Figuur 5. *Het effect van de plaats van het zigzag tape (gemeten vanuit het voorste punt) op de weerstandscoëfficiënt van een cilinder. Het tape werd aan weerszijden van de neus aangebracht*

7. De openstraal windtunnel

Op echte benen is een en ander geverifieerd met topschaatsers als proefkonijn. In Figuur 6 ondergaat Bart Veldkamp een test in de openstraal windtunnel van

de TUDelft. Hij staat op een plateau, waarmee de weerstand wordt gemeten met behulp van een rekstrookbalans.



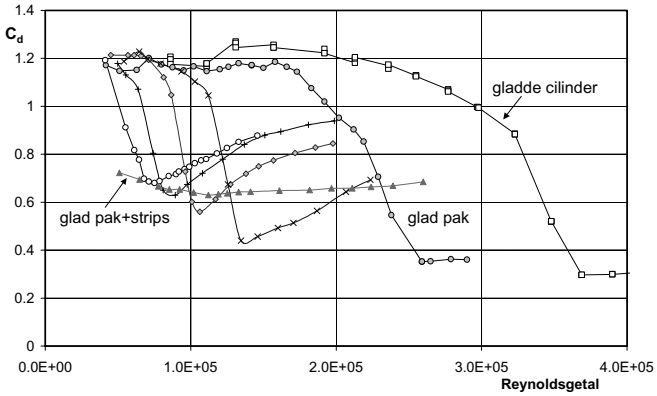
Figuur 6. *Bart Veldkamp in de openstraal windtunnel van de TUDelft. Hij staat op een rekstrookbalans om de weerstandskracht te meten. Met een woldraadje worden loslaatgebieden getraceerd*

Gemiddeld werd bij de topschaatsers in de wintunnel een verlaging in weerstand ten gevolge van de strips gevonden van ca 5%, hetgeen overeenkomt met ca. 0,5 seconde per rondje. In de bochten bevindt de schaatser zich uiteraard niet in de standaardhouding, zodat de totale winst wat lager uitvalt. Men heeft nogal getwijfeld aan de waarde van de metingen voor de praktijk, omdat men vond dat de statische houding van de schaatser in de windtunnel niet overeenkomt met de schaatspraktijk. Dat blijkt reuze mee te vallen. De dynamiek bij het schaatsen op de wat langere afstanden zou zich moeten vertalen in variatie van de invalshoek van het been. Omdat de lucht stilstaat is deze bij het grootste gedeelte van de schaatsbeweging -zelfs in de bochten- betrekkelijk gering. Metingen aan het been van Figuur 4 lieten zien, dat bij invalshoek variaties van ca. ± 15 graden gebruik van het tape nog steeds een weerstandsverlaging oplevert.

8. Nieuwe pakken

Het nadeel van de zigzagstrips is het feit dat ze bij iedere wedstrijd weer opnieuw moeten worden aangebracht. Bovendien is het moeilijk om voor de armen en de bovenbenen, die veel meer dynamiek vertonen dan de rest van het lichaam, een eenduidige locatie te vinden waarop de strips kunnen worden aangebracht. In de huidige generatie wedstrijdpakken is daarom het verstorende effect van de strips verwerkt via het gebruik van verschillende ruwe stoffen. Afhankelijk van het Reynoldsgetal (c.q. de lokale dikte van arm of been en de snelheid) moet dan de stof worden gekozen die de laagste weerstand geeft. In samenwerking met het Rotterdamse adviesbureau Flow*Tec* Aerodynamics

is onderzoek verricht naar stoffen die in aanmerking komen voor toepassing in sportkleding. Figuur 7 laat de experimenteel bepaalde weerstandscoëfficiënten zien van een cilinder bekleed met verschillende stoffen, ieder met zijn eigen specifieke ruwheid.



Figuur 7. *De weerstandscoëfficiënt van een cilinder bekleed met stoffen met verschillende ruwheid als functie van het Reynoldsgetal*

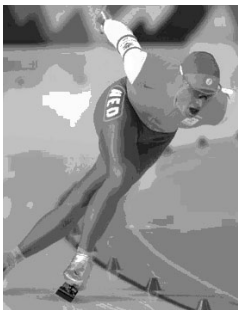
Gaan we ervan uit dat de onderbenen van een schaatser een gemiddeld Reynoldsgetal van ca. 100.000 hebben, dan zijn er volgens Figuur 7 stoffen die beslist niet in aanmerking komen, omdat ze wel effect hebben, maar bij een te hoge snelheid. Duidelijk is het afwijkende gedrag te zien van een gladde stof met zigzagstrips. In tegenstelling tot de ruwe stoffen, die een duidelijke “dip” in de weerstand laten zien, blijft de weerstandscoëfficiënt van een cilinder met gladde stof en zigzagtape constant over een breed gebied van snelheden. Als de strips niet gebruikt worden, moet de ontwerper dus gerichter keuzes maken, meer weten van het materiaal waarmee hij werkt, om met de juiste configuratie te komen.

De ruwheid van de stof kan gemaakt worden door verschillende breisels toe te passen, maar ook stoffen met applicaties komen in aanmerking. In Figuur 8 is de Poolse schaatsers Zygmunt te zien, die een windtunneltest ondergaat gehuld in een pak van Nederlands fabrikaat, waarop een driehoekige applicatie is aangebracht op bovenbenen en hoofd. Op de onderbenen en armen is een stof met verschillend breisel toegepast. De rest van het lichaam heeft een gladde stof.

De nieuwe generatie pakken zonder strips werd voor het eerst gebruikt op de Olympische Winterspelen van Salt Lake City, 2002. De Nederlandse schaatsers kwamen uit in een pak van Amerikaanse makelij. Figuur 9 toont Jochem Uytdehage in zo’n pak tijdens zijn rit op de 5000 m.

Maar om de betrekkelijkheid van deze materie aan te geven wil ik besluiten met Figuur 10. Het materiaal kan nog zo goed in orde zijn, op de been blijven is veel belangrijker.

Figuur 8. *De Pool Pavel Zygmunt in de windtunnel tijdens proeven aan een pak met driehoekige applicaties en verschillende stoffen op armen en benen*



Figuur 9. *Jochem Uytdehage in een nieuw schaatspak van een Amerikaanse fabrikant*

Figuur 10. *De val van favoriet Jeremy Wotherspoon vlak na de start van de 500 meter op de Winterspelen van Salt Lake City*

CWI Vakantiecursus 2003

pp. 103 – 112

De wiskunde achter de eurodiffusie

Misja Nuyens
Universiteit van Amsterdam
e-mail: mnuyens@science.uva.nl

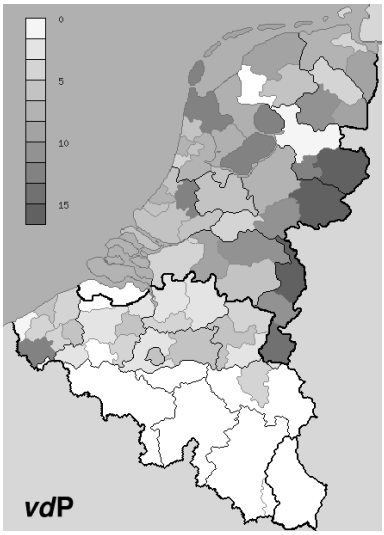
Bob Planqué
CWI, Amsterdam
e-mail: r.planque@cwi.nl

1. Inleiding

Op 1 januari 2002 deed zich een uitzonderlijke situatie voor: in twaalf verschillende landen werd tegelijkertijd de euro ingevoerd. Elk land had euro's geslagen met een eigen nationale zijde. Dit stelde ons voor het eerst in staat om het betalingsverkeer met munten te analyseren. De begintoestand was immers volledig bekend en het is duidelijk waar elke munt vandaan komt. In totaal werden 65 miljard munten in omloop gebracht, waarvan 3,3 miljard Nederlandse.

Direct begonnen de munten zich te verspreiden over Europa; het proces van de 'eurodiffusie' was van start gegaan. Hoe lang zou het duren voordat de eerste Finse euro in Amsterdam op zou duiken? Gezien het relatief kleine percentage Nederlandse euromunten, hoe lang zou het duren voordat bijvoorbeeld de helft van de munten in Nederland van buitenlandse komaf zou zijn? Zouden alle soorten euromunten zich even snel verspreiden, of zouden we een verschil kunnen opmerken tussen bijvoorbeeld twee-euromunten en één-centstukken?

Een zestal wiskundigen uit Amsterdam (het Eurodiffusieteam) greep deze kans aan om een indruk te krijgen hoe de uitwisseling van munten in Nederland en daarbuiten geschiedt.
Een website (`http://www.wiskgenoot.nl/eurodiffusie`) werd geopend waar zogenaamde *EuroMeters* konden opgeven welke munten ze in hun portemonnee hadden. Ruim tweehonderdvijftig schoolklassen voerden zo ook maandelijks een gezamenlijke meting in. Daarnaast werd tijdens de Studiegroep Wiskunde met de Industrie 2002 een poging gedaan om een wiskundige beschrijving te geven van de verspreiding van de euro's. In dit stuk wordt een indruk gegeven van drie wiskundige modellen waarmee de eurodiffusie beschreven zou kunnen worden. Hierbij wordt gebruik gemaakt van twee verschillende wiskundige technieken, namelijk partiële differentiaalvergelijkingen en Markovketens. Het doel van deze exercitie is niet zozeer om uitgebreid in te gaan op één specifiek wiskundig model, maar meer om een indruk te krijgen hoe we een keuze kunnen maken tussen de mogelijke invalshoeken bij het modelleren. De realistische toepassing van een model is vaak zeer afhankelijk van de beschikbare meetgegevens. Het wiskundig modelleren van de verspreiding van munten is

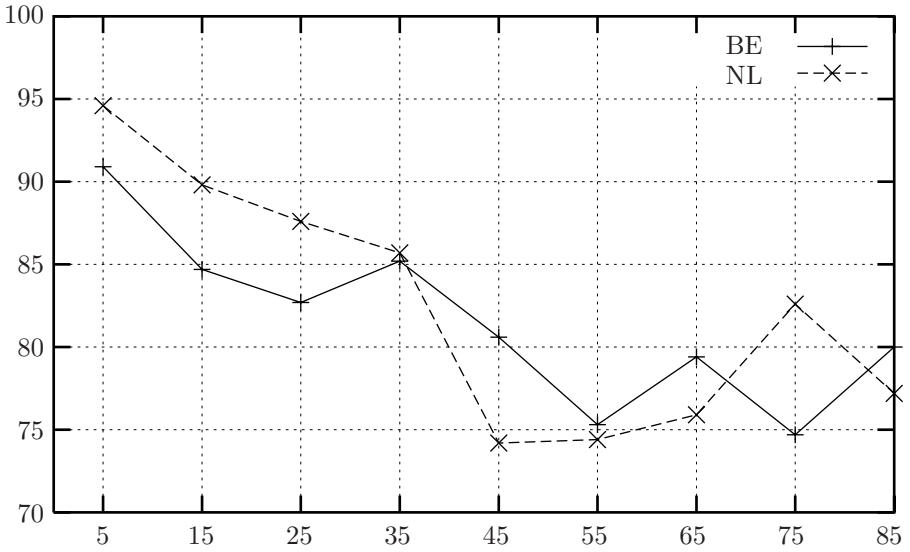


Figuur 1. *Percentage Duitse munten op 1 februari 2003*

hierop geen uitzondering. Zodoende geven we nu eerst een kort overzicht van de beschikbare data die de EuroMeters ons via de website hebben verschaft.

2. De meetgegevens

De vele EuroMeters die zich vrijwillig hadden opgegeven konden de inhoud van hun portemonnee op ieder moment invoeren. Men werd verzocht om dit in ieder geval ook aan het begin van elke maand te doen. Deze *Grote Eurometingen* vormen het meest betrouwbare overzicht van de toestand van de muntenpopulatie in Nederland en België. Bij zo'n Grote Eurometing werden soms wel 80.000 munten ingevoerd. Niettemin zien we grote fluctuaties in het verloop van de diffusie. Figuur 2 laat de percentages vaderlandse euromunten zien in Nederland en België tijdens de eerste 85 dagen. Hoewel een gestage afname waar te nemen is – zoals verwacht – zijn ook de nodige pieken en dalen te zien. Een mogelijke verklaring voor deze schommelingen is dat de EuroMeters geen aselecte groep mensen vormen. Bovendien is het niet uit te sluiten dat sommige Meters juist munten hebben opgegeven op het moment dat ze een buitenlandse euro in hun portemonee hebben gevonden. Het is immers veel spannender om een Finse munt in te voeren dan een handvol Nederlandse. Hoe sterk deze storende invloeden zijn is niet duidelijk, maar het geeft wel aan dat we de data niet altijd volkomen serieus kunnen nemen. Dit heeft zijn weerslag in de toepasbaarheid van de modellen.



***Figuur 2**. Percentage Nederlandse munten in Nederland en België in de eerste 85 dagen*

3. Continue modellen

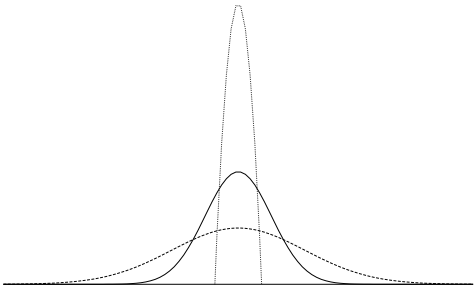
De verspreiding van euro's over Europa is het gevolg van vele verschillende processen. Mensen gaan op vakantie en geven de munten in hun portemonnee uit in het land van bestemming. Centrale banken geven nieuw geld uit. Mensen potten geld op of verliezen het in afvoerputjes. Maar misschien wel het belangrijkste proces is het dagelijks uitwisselen van muntjes bij de buurtsuper of op de markt.

Er zijn zoals we al aangaven verschillende soorten modellen die we zouden kunnen gebruiken om de eurodiffusie beter te begrijpen en voorspellen. Eén scheiding die we kunnen aanbrengen tussen deze modellen is hun nadruk op korte- of lange-afstandsprocessen. In deze paragraaf geven we een korte indruk van het soort modellen dat juist het lokale muntenverkeer als uitgangspunt neemt. Ze worden gemaakt d.m.v. zogenaamde diffusievergelijkingen, een belangrijke klasse van (partiële) differentiaalvergelijkingen (PDV's).

Het klassieke voorbeeld van een diffusieproces is dat van het mengen van twee soorten gas. Elk molecuul beweegt zich volgens een stochastische wandeling, onafhankelijk van de omringende moleculen. Kenmerkend voor diffusie is dat de gassen zich mengen totdat de concentraties van beide gassen overal hetzelfde zijn.

Een (mogelijk) nadeel van deze continue processen is het volgende: als we beginnen met een druppel rode vloeistof in een bad met water, dan voorspelt de theorie dat op elk willekeurig tijdstip na aanvang er *overal* een klein beetje

rode vloeistof zal zijn gekomen (zie Figuur 3). Deze minieme hoeveelheden zijn echter voor veel toepassingen verwaarloosbaar.



Figuur 3. *Een begrensde piek verliest zijn scherpe randen direct na aanvang van de diffusie en verspreidt zich uiteindelijk homogeen over heel* $\mathbb{R}$

Doordat er zoveel verschillende processen aan de gang zijn bij het verspreiden van de euro's is er een heel scala aan mogelijke modellen die met behulp van PDV's kunnen worden gemaakt. Bijvoorbeeld:

1. We kunnen aannemen dat er een duidelijke scheiding bestaat tussen verspreiding van munten op korte afstand en over langere afstand (de bakker vs. het vliegveld). In werkelijkheid zullen munten natuurlijk over allerlei afstanden worden uitgewisseld, maar een tweedeling is wellicht een aardige benadering.

2. We zouden plaatsafhankelijke effecten kunnen opnamen. Zo zal er bijvoorbeeld in steden meer diffusie plaatsvinden dan op platteland, simpelweg omdat er meer mensen zijn.

3. Voor gewone diffusieprocessen bestaat er geen goedgedefinieerde rand tot waar de diffusie op een bepaald tijdstip is gekomen. Dit is wellicht onwenselijk. We zouden dan kunnen kiezen voor een vergelijking die de verspreiding van euro's beschrijft op een manier die overeenkomt met het zich uitspreiden van een olievlek. Zo'n vergelijking houdt de randen wel intact.

4. We kunnen er tot slot voor kiezen om aan te nemen dat de invoering van nieuwe munten (door de centrale banken) en het verlies van munten verwaarloosbaar is.

Helaas is het niet goed mogelijk om een wijze keuze te maken uit bovenstaand rijtje: de hoeveelheid beschikbare data is te klein en te onbetrouwbaar. We beperken ons nu dan ook tot een relatief eenvoudige keuze.

Ons **Voorbeeld-model** wordt gegeven door de volgende vergelijking:

$$\frac{\partial}{\partial t}m(x,t) = D\nabla^2 m(x,t) + \int_\Omega K(x,z,t)m(z,t)dz.$$

Hier is $m(x,t)$ de hoeveelheid munten op plaats x en tijd t, D de zogenaamde diffusie-constante, Ω het gebied waar het proces plaatsvindt (de eurozone) en K een zogenaamde *intregraalkern* voor het modelleren van lange-afstandsverspreiding van munten.
We zullen nu de integraalkern K toelichten. In K zouden we bijvoorbeeld het volgende scenario kunnen opnemen. Reizigers verplaatsen zich van plaats x naar z met een zekere kans, en nemen dan een aantal munten mee van huis. Afhankelijk van hoe de munten op plaats z verdeeld zijn, worden meer of minder munten uitgewisseld. We nemen aan dat er diffusie overal even snel plaatsvindt. Het verschil tussen stad en platteland is gemakshalve onder het tapijt geschoven. Verder nemen we aan dat er geen munten kwijtraken of opgepot worden en dat er nieuwe munten worden ingevoerd. K heeft onder deze aannamen dan de volgende vorm:

$$K(x,z,t)m(z,t) = f(x,z,t) + f(z,x,t),$$

waar

$$f(x,z,t) = \epsilon p(x,z,t)\rho[m(z,t) - m(x,t)].$$

Hier is ϵ een maat voor het aantal munten dat men in bezit heeft en uitwisselt, is $p(x,z,t)$ de kans dat een persoon uit plaats x zich in z bevindt op tijd t en is ρ de bevolkingsdichtheid.

Aangezien we zeer weinig gegevens hebben die ons helpen enkele van de parameters te schatten, laat Figuur 4 de grafieken zien voor enkele keuzes van de parameters.

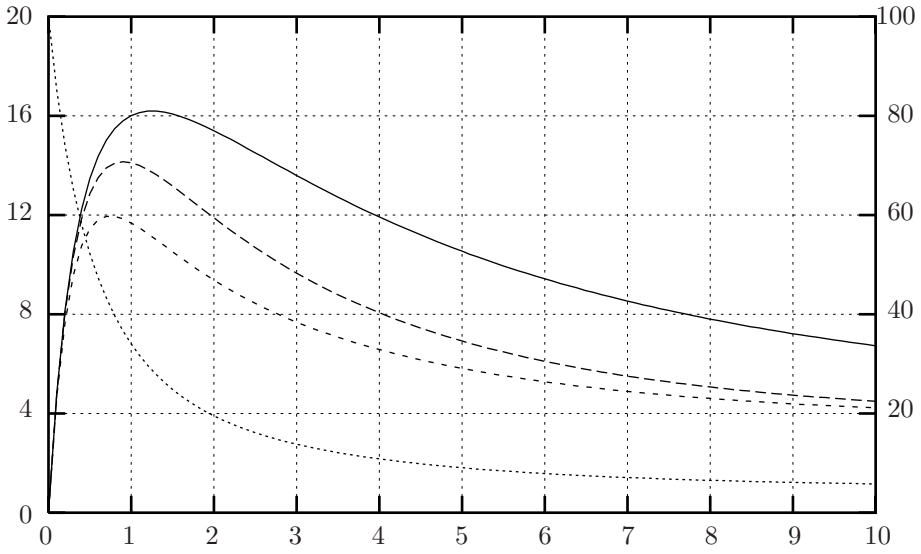
In deze grafiek is te zien hoe het aantal Nederlandse munten in ons land monotoon daalt in de tijd. Het aantal Belgische munten stijgt eerst doordat er in het begin nog veel Belgische munten in de buurt van Nederland zijn. Na een tijdje zijn de munten echter verder uitgespreid en daalt ook het aandeel Belgische munten tot een paar procent.

We merken nog op dat het lange-termijngedrag voor alle modellen die we hebben voorgesteld hetzelfde is: aangezien geen van de modellen een mechanisme bevat dat de munten weer scheidt is voor alle modellen de voorspelling dat in Nederland op de lange duur nog maar 5% van de munten van Nederlandse oorsprong is.

Continue modellen zijn een intuïtief voor de hand liggende optie om de eurodiffusie te modelleren. We moeten echter concluderen dat ze door de complexiteit van de werkelijkheid en het gebrek aan goede meetgegevens maar beperkt inzetbaar zijn.

4. Discrete modellen

Een andere manier om de diffusie van de euromunten te modelleren is de volgende. We maken een kansmodel dat het gedrag van één munt beschrijft. Het



Figuur 4. *Percentage Nederlandse (fijnst gestippeld, schaal rechts) en Belgische munten (schaal links) in Nederland in de eerste 10 maanden voor verschillende parameterwaarden voor* $\epsilon\rho/D$

gedrag van deze ene munt wordt door het gebruik van de zogenaamde *Wet van de grote aantallen* als volgt geëxtrapoleerd naar het gedrag van de totale populatie van munten: als de munt inn 200 dagen met kans 1/8 van Nederland van Spanje verhuist, zal ruwweg 1/8 deel van alle Nederlandse munten na 200 dagen zich in Spanje bevinden.
Voor de ene munt bouwen we een Markov-keten. Als mogelijke toestanden nemen we de twaalf landen. We kunnen dan een overgangsmatrix M opstellen die aangeeft hoe waarschijnlijk het is dat de munt zich in een tijdsstap – bijvoorbeeld een maand – naar land j verhuist, als hij zich in land i bevond. Als we de overgangsmatrix M n keer met zichzelf vermenigvuldigen, krijgen we de n-staps-overgangskansen. Als we de matrix maar met zichzelf blijven vermenigvuldigen, ontstaan er langzaam maar zeker overal dezelfde kolommen en komen we dichter en dichter bij de evenwichtstoestand. In het geval van de euromunten is de evenwichtstoestand de totale vermenging van alle munten uit alle landen. Dit betekent dat in elk van de eurolanden de concentraties Duitse, Spaans en Italiaanse munten hetzelfde zijn.

Uitbreidingen van het model zijn mogelijk. We kunnen een extra toestand invoeren voor munten die kwijtraken doordat ze verdwijnen in afvoerputjes of door een Peruaanse mee worden genomen naar Zuid-Amerika. Tegenover deze put kunnen we ook een bron plaatsen: de DNB brengt elk jaar een aantal nieuwe munten in omloop. Hiervoor kunnen we een extra toestand in het model kunnen creëren.

De getallen in de hierboven beschreven overgangsmatrix zijn echter onbekend. Het enige gegeven dat de ingevoerde metingen op de webpagina opleve-

ren is het percentage buitenlandse munten in de populatie van munten die in Nederland aanwezig zijn. We besluiten daarom het mooie, uitgebreide model hierboven te laten voor wat het is en een simpelere Markov-keten te bekijken.

5. De praktijk: een simpel(er) model

De meetgegevens geven ons alleen informatie over het percentage buitenlandse munten in Nederland. Dit zijn de enige gegevens die we mogen gebruiken. De simpele(re) Markov-keten die we gaan gebruiken om voorspellingen te doen voor het verloop van de eurodiffusie kent daarom slechts twee toestanden: Nederland en Buitenland. De overgangsmatrix bevat vier getallen. Noem p de kans dat een munt die nu in Nederland is een tijdsstap later in het buitenland is.

Als we aannemen dat evenveel euromunten ons land binnenkomen als verlaten, is het percentage buitenlandse munten in Nederland even groot als het percentage Nederlandse euromunten dat inmiddels haar heil heeft gezocht in het buitenland. Dit percentage na een maand hebben we hierboven p genoemd. In onze matrix M staat p linksonders, en $1-p$ linksboven. Volgens de gegevens zijn er twintig keer zoveel buitenlandse als Nederlandse euromunten. Als van alle munten in Nederland een gedeelte van grootte p in een tijdsstap de grens oversteekt, moet van alle munten in het buitenland een gedeelte $p/20$ naar Nederland komen om het evenwicht te bewaren. Een gedeelte van $1 - p/20$ zal (nog) in het buitenland blijven. We hebben nu de volgende vorm van de overgangsmatrix M afgeleid:

$$M = \begin{pmatrix} 1-p & p/20 \\ p & 1-p/20 \end{pmatrix}. \tag{1}$$

Rest ons nog om een goede waarde voor p te vinden!

Om uit de meetgegevens een waarde van p af te leiden, hebben we informatie nodig over de invloed van p op het gedrag van het proces. In het bijzonder willen we weten hoe p de n-staps-overgangskansen beïnvloedt. De volgende stelling geeft een antwoord op deze vraag.

Stelling *Laat de matrix M zijn als in (1). Dan geldt dat de n-staps-overgangskansen p_n van Nederland naar het buitenland monotoon stijgen in p als $0 < p < 20/21$, maar altijd kleiner zijn dan $20/21$.*

Bewijs We bewijzen deze stelling met inductie: we laten eerst zien dat de stelling waar is voor $n = 1$. Vervolgens tonen we aan dat als de stelling waar is voor een bepaalde $m \in \mathbb{N}$, deze ook waar is voor $m + 1$. Op deze manier hebben we laten zien dat de stelling geldt voor alle $n \in \mathbb{N}$.

Het geval $n = 1$ is eenvoudig te zien: p_1 is gelijk aan p en stijgt daarom in p. Bovendien geldt $p_1 < 20/21$ vanwege de aanname.

Neem nu aan dat p_m monotoon stijgend is in p en kleiner is dan $20/21$. Dit noemen we de inductiehypothese. Dan moeten we laten zien dat ook p_{m+1} stijgend is p en kleiner is dan $20/21$. De $(m + 1)$-staps-overgangskans wordt gegeven door

$$p_{m+1} = p(1 - p_m) + p_m(1 - p/20) = p + p_m(1 - \frac{21}{20}p). \tag{2}$$

Om te laten zien dat p_{m+1} stijgt in p, berekenen we de afgeleide naar p en laten zien dat deze (onder de voorwaarde!) positief is:

$$\frac{dp_{m+1}}{dp} = \frac{d}{dp}\Big(p + p_m(1 - \frac{21}{20}p)\Big) = 1 + (1 - \frac{21}{20}p)\frac{dp_m}{dp} - \frac{21}{20}p_m.$$

Vanwege de aanname $p < 20/21$ en de inductiehypothese geldt dat

$$(1 - \frac{21}{20}p)\frac{dp_m}{dp} > 0.$$

Dus

$$\frac{dp_{m+1}}{dp} > 1 - \frac{21}{20}p_m > 0$$

vanwege het tweede deel van de inductiehypothese. Daar p_{m+1} monotoon stijgt in p, kunnen we $p = 20/21$ invullen in (2) en krijgen we

$$p_{m+1} = p + p_m(1 - \frac{21}{20}p) < 20/21.$$

Hiermee is het bewijs voltooid. QED

6. Voorspellingen

Voor het doen van voorspellingen is de waarde van p nodig. Om p te vinden zouden we simpelweg kunnen kijken naar het percentage buitenlandse munten in Nederland na één maand. Echter, als we gegevens van de rest van het jaar gebruiken, zullen we waarschijnlijk een nauwkeurigere schatting krijgen. Hoe we dat doen zullen we nu uitleggen.

Eerst merken we op dat het percentage buitenlandse munten na bijvoorbeeld acht maanden de 8-staps-overgangskans p_8 is. Deze kans kunnen we in Figuur 5 tekenen.

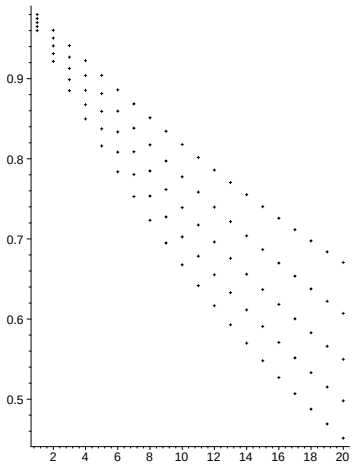
We vinden dan twee andere 8-staps-overgangskansen waar hij tussen ligt, met twee bijbehorende waarden voor p. De stelling hierboven zegt dan dat alle andere 8-staps-overgangskansen groter of kleiner zijn. De twee gevonden waarden voor p leveren zodoende een model op dat de gegevens het beste benadert (voor $n = 8$).

Om het percentage buitenlandse euromunten op een bepaald tijdstip in de toekomst te voorspellen, hoeven we nu alleen maar de curve die hoort bij de gevonden waarde(n) van p 'af te lopen' tot we bij dat tijdstip zijn.

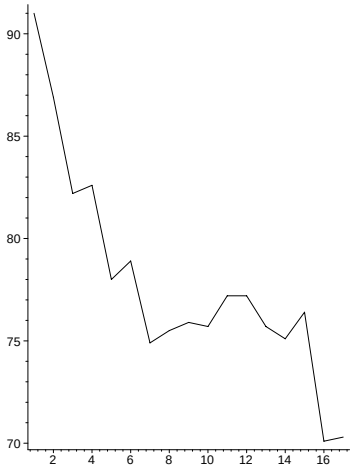
7. Resultaten en conclusie

De eerste maanden schatte het Eurodiffusieteam de waarde van de parameter p in het model hierboven op 0.04. Zoals uit Figuur 5 valt af te leiden, zou dat betekenen dat rond juni 2003 *het omslagpunt* zou zijn bereikt: evenveel buitenlandse als Nederlandse euro's in ons land.

Deze voorspelling bleek echter aan de vroege kant, zoals te zien is in Figuur 6.



Figuur 5. *Percentage buitenlandse munten in Nederland na* $n = 1, \ldots, 20$ *maanden voor* $p = 0.04 - 0.035 - 0.03 - 0.025 - 0.02$



Figuur 6. *Percentage Nederlandse euromunten in omloop in Nederland als functie van het aantal maanden sinds 1 januari 2002*

Na de zomervakantie werd de waarde van p door het Eurodiffusieteam *verlaagd* tot 0.03, maar ook deze schatting lijkt onjuist. Op dit moment schatten we p op 'slechts' 0.02.
Hierbij zij opgemerkt dat elk muntstuk eigenlijk zijn eigen p-waarde heeft: munten van twee euro blijken veel sneller te reizen dan die van één cent. (Kunt u verklaren waarom?) Het omslagpunt voor twee-eurostukken zal dan ook veel eerder komen dan dat voor één-centstukken.

De voornaamste les die we kunnen leren uit dit experiment is dat het wiskundig modelleren van een fenomeen zoals de eurodiffusie vaak een munt met twee kanten is: aan de ene kant is het wenselijk om zoveel mogelijk processen realistisch te beschrijven, maar aan de andere kant wordt je beperkt door de beschikbare meetgegevens.
Zo is het in de huidige opzet nog niet mogelijk gebleken om met een (eenvoudig) model de vreemde stagnatie van de verspreiding, zie Figuur 6, te verklaren. Ook het toevoegen van een bron (DNB) en een put (spaarpotten en niet-EU toeristen) aan het model biedt geen uitkomst.
Of het simpele model van paragraaf 5 op de lange termijn – wanneer alle mappen van muntenspaarders vol zijn – beter werkt, zullen we over een paar jaar weten...

8. Referenties

1. Geertje Hek, Misja Nuyens, Harmen van der Ploeg, Bob Planqué, Erick Vermeulen, *Het grote internationale eurodiffusie-experiment*, Natuur & Techniek **11** (2002) p. 56–62.
2. `http://www.wiskgenoot.nl/eurodiffusie`: site van het Eurodiffusieproject in 2002
3. Erick Vermeulen en de Studiegroep Wiskunde met de Industrie, *Het grote internationale eurodiffusie-experiment*, Natuur & Techniek **01** (2002) p. 11 and 22–25.
4. `http://www.eurodiffusie.nl`: site van het Eurodiffusieproject in 2003 (en verder?)
5. K. van Harn en P.J. Holewijn, *Markov-ketens in diskrete tijd,* Epsilon Uitgaven, 1991

CWI SYLLABI

1 Vakantiecursus 1984: *Hewet - plus wiskunde.* 1984.

2 E.M. de Jager, H.G.J. Pijls (eds.). *Proceedings Seminar 1981–1982. Mathematical structures in field theories.* 1984.

3 W.C.M. Kallenberg, et al. *Testing statistical hypotheses: worked solutions.* 1984.

4 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 1.* 1984.

5 J.G. Verwer (ed.). *Colloquium topics in applied numerical analysis, volume 2.* 1984.

6 P.J.M. Bongaarts, J.N. Buur, E.A. de Kerf, R. Martini, H.G.J. Pijls, J.W. de Roever. *Proceedings Seminar 1982–1983. Mathematical structures in field theories.* 1985.

7 Vacantiecurus 1985: *Variatierekening.* 1985.

8 G.M. Tuynman. *Proceedings Seminar 1983–1985. Mathematical structures in field theories, Vol.1 Geometric quantization.* 1985.

9 J. van Leeuwen, J.K. Lenstra (eds.). *Parallel computers and computations.* 1985.

10 Vakantiecursus 1986: *Matrices.* 1986.

11 P.W.H. Lemmens. *Discrete wiskunde: tellen, grafen, spelen en codes.* 1986.

12 J. van de Lune. *An introduction to Tauberian theory: from Tauber to Wiener.* 1986.

13 G.M. Tuynman, M.J. Bergvelt, A.P.E. ten Kroode. *Proceedings Seminar 1983–1985. Mathematical structures in field theories, Vol.2.* 1987.

14 Vakantiecursus 1987: *De personal computer en de wiskunde op school.* 1987.

15 Vakantiecursus 1983: *Complexe getallen.* 1987.

16 P.J.M. Bongaarts, E.A. de Kerf, P.H.M. Kersten. *Proceedings Seminar 1984–1986. Mathematical structures in field theories, Vol.1.* 1988.

17 F. den Hollander, H. Maassen (eds.). *Mark Kac seminar on probability and physics. Syllabus 1985–1987.* 1988.

18 Vakantiecursus 1988. *Differentierekening.* 1988.

19 R. de Bruin, C.G. van der Laan, J. Luyten, H.F. Vogt. *Publiceren met LATEX.* 1988.

20 R. van der Horst, R.D. Gill (eds.). *STATAL: statistical procedures in Algol 60, part 1.* 1988.

21 R. van der Horst, R.D. Gill (eds.). *STATAL: statistical procedures in Algol 60, part 2.* 1988.

22 R. van der Horst, R.D. Gill (eds.). *STATAL: statistical procedures in Algol 60, part 3.* 1988.

23 J. van Mill, G.Y. Nieuwland (eds.). *Proceedings van het symposium wiskunde en de computer.* 1989.

24 P.W.H. Lemmens (red.). *Bewijzen in de wiskunde.* 1989.

25 Vakantiecursus 1989: *Wiskunde in de Gouden Eeuw.* 1989.

26 G.G.A. Bäuerle et al. *Proceedings Seminar 1986–1987. Mathematical structures in field theories.* 1990.

27 Vakantiecursus 1990: *Getallentheorie en haar toepassingen.* 1990.

28 Vakantiecursus 1991: *Meetkundige structuren.* 1991.

29 A.G. van Asch, F. van der Blij. *Hoeken en hun Maat.* 1992.

30 M.J. Bergvelt, A.P.E. ten Kroode. *Proceedings seminar 1986–1987. Lectures on Kac-Moody algebras.* 1992.

31 Vakantiecursus 1992: *Systeemtheorie.* 1992.

32 F. den Hollander, H. Maassen (eds.). *Mark Kac seminar on probability and physics. Syllabus 1987–1992.* 1992.

33 P.W.H. Lemmens (ed.). *Meetkunde van kunst tot kunde, vroeger en nu.* 1993.

34 J.H. Kruizinga. *Toegepaste wiskunde op een PC.* 1992.

35 Vakantiecursus 1993: *Het reële getal.* 1993.

36 Vakantiecursus 1994: *Computeralgebra.* 1994.

37 G. Alberts. *Wiskunde en praktijk in historisch perspectief. Syllabus.* 1994.

38 G. Alberts, J. Schut (eds.). *Wiskunde en praktijk in historisch perspectief. Reader.* 1994.

39 E.A. de Kerf, H.G.J. Pijls (eds.). *Proceedings Seminar 1989–1990. Mathematical structures in field theory.* 1996.

40 Vakantiecursus 1995: *Kegelsneden en kwadratische vormen.* 1995.

41 Vakantiecursus 1996: *Chaos.* 1996.

42 H.C. Doets. *Wijzer in Wiskunde.* 1996.

43 Vakantiecursus 1997: *Rekenen op het Toeval.* 1997.

44 Vakantiecursus 1998: *Meetkunde, Oud en Nieuw.* 1998.

45 Vakantiecursus 1999: *Onbewezen Vermoedens.* 1999.

46 P.W. Hemker, B.W. van de Fliert (eds.). *Proceedings of the 33^{rd} European Study Group with Industry.* 1999.

47 K.O. Dzhaparidze. *Introduction to Option Pricing in a Securities Market.* 2000.

48 Vakantiecursus 2000: *Is wiskunde nog wel mensenwerk?* 2000.

49 Vakantiecursus 2001: *Experimentele wiskunde.* 2001.

50 Vakantiecursus 2002: *Wiskunde en gezondheid.* 2002.

51 G.M. Hek (ed.). *Proceedings of the 42^{nd} European Study Group with Industry.* 2002.

52 Vakantiecursus 2003: *Wiskunde in het dagelijks leven* 2003.

MC SYLLABI

1.1 F. Göbel, J. van de Lune. *Leergang besliskunde, deel 1: wiskundige basiskennis.* 1965.

1.2 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 2: kansberekening.* 1965.

1.3 J. Hemelrijk, J. Kriens. *Leergang besliskunde, deel 3: statistiek.* 1966.

1.4 G. de Leve, W. Molenaar. *Leergang besliskunde, deel 4: Markovketens en wachttijden.* 1966.

1.5 J. Kriens, G. de Leve. *Leergang besliskunde, deel 5: inleiding tot de mathematische besliskunde.* 1966.

1.6a B. Dorhout, J. Kriens. *Leergang besliskunde, deel 6a: wiskundige programmering 1.* 1968.

1.6b B. Dorhout, J. Kriens, J.Th. van Lieshout. *Leergang besliskunde, deel 6b: wiskundige programmering 2.* 1977.

1.7a G. de Leve. *Leergang besliskunde, deel 7a: dynamische programmering 1.* 1968.

1.7b G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7b: dynamische programmering 2.* 1970.

1.7c G. de Leve, H.C. Tijms. *Leergang besliskunde, deel 7c: dynamische programmering 3.* 1971.

1.8 J. Kriens, F. Göbel, W. Molenaar. *Leergang besliskunde, deel 8: minimaxmethode, netwerkplanning, simulatie.* 1968.

2.1 G.J.R. Förch, P.J. van der Houwen, R.P. van de Riet. *Colloquium stabiliteit van differentieschema's, deel 1.* 1967.

2.2 L. Dekker, T.J. Dekker, P.J. van der Houwen, M.N. Spijker. *Colloquium stabiliteit van differentieschema's, deel 2.* 1968.

3.1 H.A. Lauwerier. *Randwaardeproblemen, deel 1.* 1967.

3.2 H.A. Lauwerier. *Randwaardeproblemen, deel 2.* 1968.

3.3 H.A. Lauwerier. *Randwaardeproblemen, deel 3.* 1968.

4 H.A. Lauwerier. *Representaties van groepen.* 1968.

5 J.H. van Lint, J.J. Seidel, P.C. Baayen. *Colloquium discrete wiskunde.* 1968.

6 K.K. Koksma. *Cursus ALGOL 60.* 1969.

7.1 *Colloquium moderne rekenmachines, deel 1.* 1969.

7.2 *Colloquium moderne rekenmachines, deel 2.* 1969.

8 H. Bavinck, J. Grasman. *Relaxatietrillingen.* 1969.

9.1 T.M.T. Coolen, G.J.R. Förch, E.M. de Jager, H.G.J. Pijls. *Colloquium elliptische differentiaalvergelijkingen, deel 1.* 1970.

9.2 W.P. van den Brink, T.M.T. Coolen, B. Dijkhuis, P.P.N. de Groen, P.J. van der Houwen, E.M. de Jager, N.M. Temme, R.J. de Vogelaere. *Colloquium elliptische differentiaalvergelijkingen, deel 2.* 1970.

10 J. Fabius, W.R. van Zwet. *Grondbegrippen van de waarschijnlijkheidsrekening.* 1970.

11 H. Bart, M.A. Kaashoek. H.G.J. Pijls, W.J. de Schipper, J. de Vries. *Colloquium halfalgebra's en positieve operatoren.* 1971.

12 T.J. Dekker. *Numerieke algebra.* 1971.

13 F.E.J. Kruseman Aretz. *Programmeren voor rekenautomaten; de MC ALGOL 60 vertaler voor de EL X8.* 1971.

14 H. Bavinck, W. Gautschi, G.M. Willems. *Colloquium approximatietheorie.* 1971.

15.1 T.J. Dekker, P.W. Hemker, P.J. van der Houwen. *Colloquium stijve differentiaalvergelijkingen, deel 1.* 1972.

15.2 P.A. Beentjes, K. Dekker, H.C. Hemker, S.P.N. van Kampen, G.M. Willems. *Colloquium stijve differentiaalvergelijkingen, deel 2.* 1973.

15.3 P.A. Beentjes, K. Dekker, P.W. Hemker, M. van Veldhuizen. *Colloquium stijve differentiaalvergelijkingen, deel 3.* 1975.

16.1 L. Geurts. *Cursus programmeren, deel 1: de elementen van het programmeren.* 1973.

16.2 L. Geurts. *Cursus programmeren, deel 2: de programmeertaal ALGOL 60.* 1973.

17.1 P.S. Stobbe. *Lineaire algebra, deel 1.* 1973.

17.2 P.S. Stobbe. *Lineaire algebra, deel 2.* 1973.

17.3 N.M. Temme. *Lineaire algebra, deel 3.* 1976.

18 F. van der Blij, H. Freudenthal, J.J. de Iongh, J.J. Seidel, A. van Wijngaarden. *Een kwart eeuw wiskunde 1946-1971, syllabus van de vakantiecursus 1971.* 1973.

19 A. Hordijk, R. Potharst, J.Th. Runnenburg. *Optimaal stoppen van Markovketens.* 1973.

20 T.M.T. Coolen, P.W. Hemker, P.J. van der Houwen, E. Slagt. *ALGOL 60 procedures voor begin- en randwaardeproblemen.* 1976.

21 J.W. de Bakker (red.). *Colloquium programmacorrectheid.* 1975.

22 R. Helmers, J. Oosterhoff, F.H. Ruymgaart, M.C.A. van Zuylen. *Asymptotische methoden in de toetsingstheorie; toepassingen van naburigheid.* 1976.

23.1 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 1.* 1976.

23.2 J.W. de Roever (red.). *Colloquium onderwerpen uit de biomathematica, deel 2.* 1977.

24.1 P.J. van der Houwen. *Numerieke integratie van differentiaalvergelijkingen, deel 1: eenstapsmethoden.* 1974.

25 *Colloquium structuur van programmeertalen.* 1976.

26.1 N.M. Temme (ed.). *Nonlinear analysis, volume 1.* 1976.

26.2 N.M. Temme (ed.). *Nonlinear analysis, volume 2.* 1976.

27 M. Bakker, P.W. Hemker, P.J. van der Houwen, S.J. Polak, M. van Veldhuizen. *Colloquium discretiseringsmethoden.* 1976.

28 O. Diekmann, N.M. Temme (eds.). *Nonlinear diffusion problems.* 1976.

29.1 J.C.P. Bus (red.). *Colloquium numerieke programmatuur, deel 1A, deel 1B.* 1976.

29.2 H.J.J. te Riele (red.). *Colloquium numerieke programmatuur, deel 2.* 1977.

30 J. Heering, P. Klint (red.). *Colloquium programmeeromgevingen.* 1983.

31 J.H. van Lint (red.). *Inleiding in de coderingstheorie.* 1976.

32 L. Geurts (red.). *Colloquium bedrijfsystemen.* 1976.

33 P.J. van der Houwen. *Berekening van waterstanden in zeeën en rivieren.* 1977.

34 J. Hemelrijk. *Oriënterende cursus mathematische statistiek.* 1977.

35 P.J.W. ten Hagen (red.). *Colloquium computer graphics.* 1978.

36 J.M. Aarts, J. de Vries. *Colloquium topologische dynamische systemen.* 1977.

37 J.C. van Vliet (red.). *Colloquium capita datastructuren.* 1978.

38.1 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part I.* 1979.

38.2 T.H. Koornwinder (ed.). *Representations of locally compact groups with applications, part II.* 1979.

39 O.J. Vrieze, G.L. Wanrooy. *Colloquium stochastische spelen.* 1978.

40 J. van Tiel. *Convexe analyse.* 1979.

41 H.J.J. te Riele (ed.). *Colloquium numerical treatment of integral equations.* 1979.

42 J.C. van Vliet (red.). *Colloquium capita implementatie van programmeertalen.* 1980.

43 A.M. Cohen, H.A. Wilbrink. *Eindige groepen (een inleidende cursus).* 1980.

44 J.G. Verwer (ed.). *Colloquium numerical solution of partial differential equations.* 1980.

45 P. Klint (red.). *Colloquium hogere programmeertalen en computerarchitectuur.* 1980.

46.1 P.M.G. Apers (red.). *Colloquium databankorganisatie, deel 1.* 1981.

46.2 P.G.M. Apers (red.). *Colloquium databankorganisatie, deel 2.* 1981.

47.1 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60: general information and indices.* 1981.

47.2 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 1: elementary procedures; vol. 2: algebraic evaluations.* 1981.

47.3 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3A: linear algebra, part I.* 1981.

47.4 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 3B: linear algebra, part II.* 1981.

47.5 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 4: analytical evaluations; vol. 5A: analytical problems, part I.* 1981.

47.6 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 5B: analytical problems, part II.* 1981.

47.7 P.W. Hemker (ed.). *NUMAL, numerical procedures in ALGOL 60, vol. 6: special functions and constants; vol. 7: interpolation and approximation.* 1981.

48.1 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 1.* 1982.

48.2 P.M.B. Vitányi, J. van Leeuwen, P. van Emde Boas (red.). *Colloquium complexiteit en algoritmen, deel 2.* 1982.

49 T.H. Koornwinder (ed.). *The structure of real semisimple Lie groups.* 1982.

50 H. Nijmeijer. *Inleiding systeemtheorie.* 1982.

51 P.J. Hoogendoorn (red.). *Cursus cryptografie.* 1983.