

De Shannon theorie

Citation for published version (APA):

Schalkwijk, J. P. M. (1974). *De Shannon theorie*. Technische Hogeschool Eindhoven.

Document status and date:

Gepubliceerd: 01/01/1974

Document Version:

Uitgevers PDF, ook bekend als Version of Record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

Take down policy

If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

DE SHANNON THEORIE

REDE

UITGESPROKEN BIJ DE AANVAARDING
VAN HET AMBT VAN GEWOON HOGLERAAR
IN DE TELECOMMUNICATIE AAN DE
AFDELING DER ELEKTROTECHNIEK
VAN DE TECHNISCHE HOGESCHOOL EINDHOVEN
OP VRIJDAG 8 NOVEMBER 1974

DOOR

J. P. M. SCHALKWIJK

'Hard core information theory' of 'Shannon theory' is een deel van het vakgebied informatie theorie. De Shannon theorie dateert van juli 1948, d.w.z. van de publicatie van Claude Shannon's artikel 'The Mathematical Theory of Communication' in het Bell System Technical Journal. De ontwikkeling van de Shannon theorie kan op de voet worden gevolgd in een nieuw boek van de IEEE (Institute of Electrical and Electronics Engineers) Press, getiteld 'Key Papers in the Development of Information Theory', samengesteld door David Slepian. In de volgende bladzijden wil ik proberen een aantal centrale aspecten van de Shannon theorie toe te lichten.

In Fig. 1 is een communicatie situatie (net) voorgesteld met vier deelnemers (punten) A, B, C en D. Informatie overdracht kan plaatsvinden in de richting van de pijltjes, d.w.z. voor het pijltje $A \rightarrow B$ staat de zender in A en de ontvanger in B. De maximale informatie overdracht [1] tussen twee willekeurige punten P en Q in zo'n net is gelijk aan het minimum

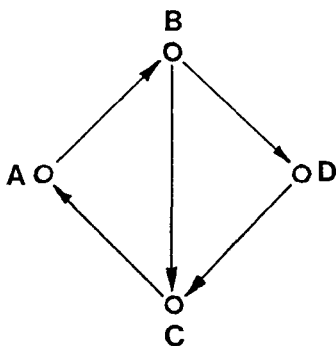


Fig. 1. Communicatie net.

van de capaciteit over alle mogelijke scheidingsvlakken tussen P en Q (Ford-Fulkerson theorema). In het algemeen is er weinig informatie theoretisch onderzoek aan communicatienetten gedaan, daar men zich vrijwel direct op de enkele verbinding, zoals de pijl $A \rightarrow B$ in Fig. 1, is gaan concentreren. Wij zullen dit nu ook doen, echter niet zonder een recente opbloei [2-10] van interesse in communicatie netten te noemen, waarop we later terugkomen.

De enkele verbinding, zoals $A \rightarrow B$ in Fig. 1, is in het blokschema van

Fig. 2 nader uitgewerkt. De informatie bron B genereert een signaal $u(t)$ met bandbreedte W_b Hz. Dit signaal $u(t)$ wordt door de modulator M omgezet in een signaal $s(t)$ met bandbreedte W_k Hz. Het quotient

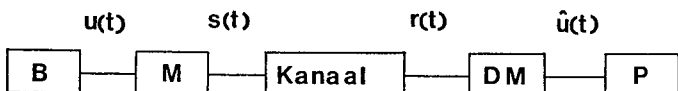


Fig. 2. Blokschema van de enkele verbinding.

$N = W_k/W_b$ wordt de 'bandwidth expansion' genoemd. De bedoeling van het modulatie proces is een signaal $s(t)$ te genereren, dat minder gevoelig voor de verstoring op het kanaal is dan de informatie $u(t)$ in zijn oorspronkelijke vorm. Hoe groter de 'bandwidth expansion' N des te beter kan aan de eis van storingsongevoeligheid van $s(t)$ worden voldaan. Het verband tussen N en de storingsongevoeligheid van het systeem zal nader worden toegelicht. De demolulator, DM in Fig. 2, heeft tenslotte tot taak uit het ontvangen signaal $r(t)$, een verstoorde versie van het verzonden signaal $s(t)$, een schatting $\hat{u}(t)$, van $u(t)$ te vormen. De enkele verbinding eindigt dan in de informatie put P, zie Fig. 2.

Volgens het bemonsteringstheorema [11, p. 599] kunnen we een signaal $x(t)$ met bandbreedte W Hz vervangen door een equivalente rij x_1, x_2, \dots van $2W$ monsters per seconde. Voor de volgende discussie denken we ons nu de signalen in Fig. 2 vervangen door equivalente rijen van monsters. In communicatie termen noemt men dit het overgaan van de golfvorm representatie op de vector representatie van het communicatie probleem. Het bron signaal $u(t)$ wordt dus vervangen door de rij u_1, u_2, \dots , waarbij de bron $2W_b$ monsters per seconde produceert. Het gemoduleerde signaal $s(t)$ wordt vervangen door de rij $s_{11}, s_{12}, \dots, s_{1N}, s_{21}, s_{22}, \dots, s_{2N}, s_{31}, \dots$, waarbij $2W_k$ monsters per seconde worden verzonden. Bij een 'bandwidth expansion' van $N = W_k/W_b$ kunnen dan N kanaal monsters $s_{k1}, s_{k2}, \dots, s_{kN}$ worden gereserveerd voor het zo goed mogelijk overbrengen van een enkel bron monster $u_k, k=1, 2, \dots$. In Fig. 3a is de signaal locus weergegeven voor het geval $N=1$. Als het bron monster $u_k, k=1, 2, \dots$ toeneemt van $u_k = u_{\min}$ tot $u_k = u_{\max}$ groeit het bijbehorende signaal monster van $-|s|_{\max}$ tot $+|s|_{\max}$. Door $|s|_{\max}$ te vergroten neemt de relatieve ruisverstoring af. De signaal energie per bron monster is echter evenredig met s_{k1}^2 , m.a.w. gedeelten van de signaal locus, die ver van de oorsprong 0 in Fig. 3 liggen, corresponderen met een hoge waarde van de verzonden energie per bron monster. De oplossing van dit dilemma, te weten een

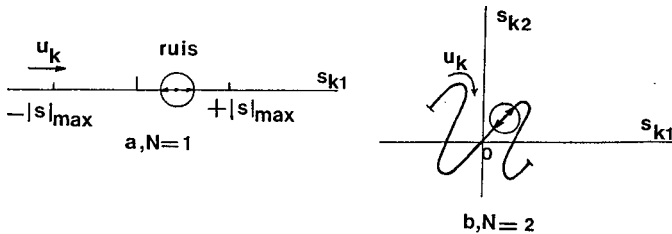


Fig. 3. De ruil tussen vermogen en bandbreedte.

lange signaal locus, m.a.w. relatieve ongevoeligheid voor ruis en toch weinig verzonden energie per bron monster werd in 1947 door Kotelnikov [12] (een recipient van de Lenin prijs) in zijn proefschrift gegeven. Voor $N \geq 2$, in Fig. 3b is het geval van $N=2$ kanaal monsters s_{k1}, s_{k2} per bron monster $u_k, k=1, 2, \dots$ weergegeven, kunnen we de signaal locus namelijk oprollen, zodat we een lange locus of een grote ruis ongevoeligheid krijgen zonder ver van de oorsprong 0 verwijderd te raken, m.a.w. voor weinig energie per bron monster. Voor het Gaussisch kanaal wordt de maximale verbetering in signaal-ruis verhouding t.g.v. 'bandwidth expansion' gegeven door de formule

$$(\text{SNR})_0 = \left[1 + \frac{(\text{SNR})_i}{N} \right]^N \quad (1)$$

Slepian [13] heeft uitgebreide berekeningen aan 'twisted modulation systems' gedaan. Een eenvoudige afleiding van deze fundamentele formule (1) werd door Schalkwijk en Bluestein [14] gegeven.

Net als Kotelnikov concentreren we ons nu verder op het overbrengen van een enkel bron monster $u_k, k=1, 2, \dots$. Het modulatie proces, waarbij u_k de grootte van de signaal monsters $s_{k1}, s_{k2}, \dots, s_{kN}$ bepaalt, kan worden beschreven m.b.v. een signaal locus in een N dimensionale ruimte, zie Fig. 3b. Voor grote N kan deze locus met analoge circuits niet voldoende nauwkeurig worden vastgelegd. M.a.w. geavanceerde modulatie en demodulatie technieken kunnen slechts met digitale hulpmiddelen (computers) worden gerealiseerd. Het over te brengen bron monster $u_k, k=1, 2, \dots$ is echter een continue grootheid. Op dit punt neemt Shannon de draad van Kotelnikov over! Aangezien de continue grootheid u_k toch niet feilloos, voor een Gaussisch kanaal geeft formule (1) b.v. $(\text{SNR})_0 < \infty$, is over te brengen, kunnen we u_k zonder verlies

meteen quantiseren. Bij de discretisering van u_k wordt dan een controleerbare fout of distortie ϵ geïntroduceerd. In het algemeen zal het minimum aantal binaire digits, H_ϵ , per bron monster toenemen naarmate een kleinere quantiseringsfout ϵ wordt geëist. Later zullen we zien dat deze H_ϵ binaire digits of bits per bron monster foutloos kunnen worden overgebracht zolang $H_\epsilon < C$, waarbij de capaciteit C in bits per bron sample een parameter van het transmissie kanaal is. De totale transmissie fout of distortie van het communicatie systeem is dan gelijk aan ϵ . Het verband tussen ϵ en H_ϵ wordt de 'rate-distortion function' [15] genoemd. Voor een Gaussisch bron monster u_k , $k=1, 2, \dots$, met variantie A en een distortie maat $\epsilon = E[(\hat{u}_k - u_k)^2]$ is de 'rate-distortion function',

$$H_\epsilon = \frac{1}{2} \log_2 \frac{A}{\epsilon}, \quad (2)$$

afgebeeld in Fig. 4. Nadat het bron monster u_k , $k=1, 2, \dots$ aldus is gequantiseerd kunnen nog slechts 2^{H_ϵ} punten op de signaal locus van Fig. 3b worden gerealiseerd. We hebben nu een aan ons oorspronkelijke modulatie systeem verwante kanaal code gekregen met $M=2^{H_\epsilon}$ code

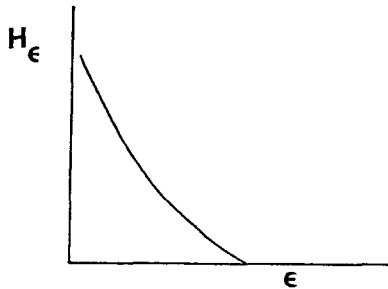


Fig. 4. 'Rate-distortion function'.

woorden. Alvorens dieper op deze kanaal codes in te gaan geven we in de volgende alinea nog een korte beschouwing over bron informatie die van nature reeds discreet is, zoals b.v. een serie uitkomsten van een berekening op een digitale computer.

Een discrete informatie bron produceert een rij letters u_1, u_2, \dots , van een eindig alfabet. Een binaire bron, met alfabet $\{a,b\}$, zonder geheugen kan worden gerealiseerd m.b.v. de roulette van Fig. 5. M.a.w. de roulette wordt gebruikt om te bepalen of u_k , $k=1, 2, \dots$, a dan wel b is. In het in Fig. 5 afgebeelde geval is $P[b]=p < P[a]=q=1-p$. De bron informatie

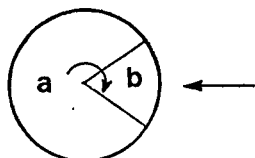


Fig. 5. Binaire informatie bron zonder geheugen.

kan nu worden gecodeerd volgens $a \rightarrow 0, b \rightarrow 1$, m.a.w. één binair digit per bron letter. Als p echter belangrijk van $\frac{1}{2}$ verschilt kan de bron informatie veel economischer worden gecodeerd. In een rij van B bron symbolen zal de letter b ongeveer pB maal voorkomen. Maar er zijn slechts $\binom{B}{pB}$ rijen ter lengte B met een fractie b 's gelijk aan p . M.a.w., we hebben slechts

$$\lceil \log_2 \binom{B}{pB} \rceil / B \text{ binaire digits per bron letter} \quad (3)$$

nodig, waar $\lceil x \rceil$ het kleinste gehele getal juist groter dan x voorstelt. M.b.v. Stirlings formule vinden we de volgende asymptotische uitdrukking voor (3) voor $B \rightarrow \infty$:

$$H(p) \triangleq -p \log_2 p - (1-p) \log_2 (1-p). \quad (4)$$

De entropie $H(p)$ van de discrete informatie bron is in Fig. 6 afgebeeld als een functie van p . Voor $p \downarrow 0$ en $p \uparrow 1$ kan de bron informatie, u_1, u_2, \dots , dus worden gecodeerd [16, 17, 18] m.b.v. zeer weinig, $H(p)$ binaire digits

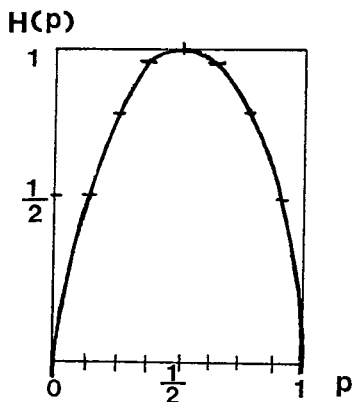


Fig. 6. Binaire entropie functie $H(p)$.

of bits per bron letter. De laatste twee alineas kunnen nu als volgt worden geresumeerd. Zowel continue- als discrete bron informatie kan worden gecodeerd in onafhankelijke, even waarschijnlijke, binaire digits en wel H_s en $H(p)$ bits per bron monster respectievelijk. Deze onafhankelijke, even waarschijnlijke, binaire digits of bits kunnen nu m.b.v. een kanaal code foutloos worden overgebracht zolang $H < C$, waar C de eerder genoemde kanaal capaciteit in bits per bron monster is en H of H_s of $H(p)$ voorstelt, afhankelijk van het al dan niet continue zijn van de informatie bron. Door de bron informatie aldus om te zetten in een rij van H bits per bron monster hebben we het met het oorspronkelijke modulatie systeem verwante coderingsprobleem gesplitst in een bron- en een kanaal coderingsprobleem. M.a.w. het stochastische proces van de bron en het stochastische proces van het transmissie medium zijn ontkoppeld! Deze ont koppeling (Shannon) geeft een belangrijke vereenvoudiging in de theoretische behandeling van geavanceerde modulatie systemen en heeft bovendien een belangrijk praktisch aspect. Een communicatie net kan nu namelijk worden geoptimaliseerd voor de standaard input (vergelijk de standaard impedantie van 50Ω voor coaxiale kabels), d.w.z. onafhankelijke, even waarschijnlijke, binaire digits, zonder dat daarvoor een prijs in de vorm van extra distortie betaald behoeft te worden. We gaan nu verder met de transmissie van informatie, d.w.z. met de diverse kanaal codes en met het begrip kanaal capaciteit!

Bij de volgende beschouwing over fouten corrigerende (kanaal) codes beperken we ons tot het 'binary symmetric channel' (BSC) van Fig. 7.

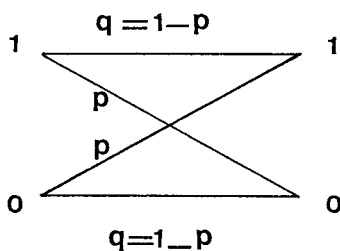


Fig. 7. 'Binary symmetric channel'.

De kans dat het ontvangen symbool, 0 of 1, van het verzonden symbool verschilt is $p < \frac{1}{2}$. Als input voor de codeer schakeling nemen we blokken van de standaard input van onafhankelijke, even waarschijnlijke binaire digits ter lengte K , d.w.z. dat onze blok code $M = 2^K$ code woorden moet

hebben. Als deze code woorden elk uit N binaire digits bestaan, dan wordt de efficiëntie van deze blok code gedefinieerd als

$$R \triangleq K/N \quad (5)$$

Een (R,N)-code heeft dus $M=2^{RN}$ code woorden ter lengte N. In de volgende tabel zijn drie blok codes weergegeven. Fig. 8 geeft een geo-

| M=2, N=3 | M=2, N=5 | M=4, N=3 |
|----------|-----------|----------|
| 0 0 0 | 0 0 0 0 0 | 0 0 0 |
| 1 1 1 | 1 1 1 1 1 | 0 1 1 |
| | | 1 0 1 |
| | | 1 1 0 |

metrische representatie van de derde code met M=4 code woorden ter lengte N=3. De optimale decodeerschakeling, d.w.z. de schakeling met

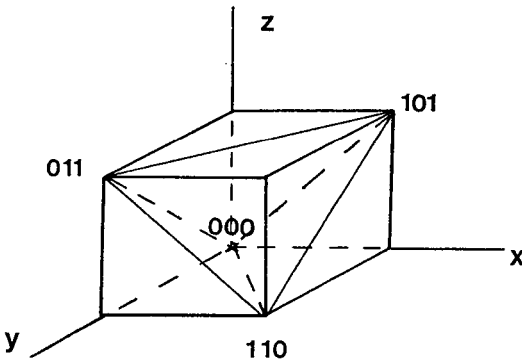


Fig. 8. Tetraheder code.

de kleinste kans $P_{e, opt}(M, N)$ op een decoderingsfout, vergelijkt ieder van de bij de bepaalde code behorende code woorden met de ontvangen digits en kiest het codewoord op minimale Hammingafstand, d.w.z. het code woord, dat in de meeste posities met de ontvangen digits overeenstemt. Als er meerdere code woorden voor keuze in aanmerking komen wordt een willekeurige keuze tussen de code woorden in kwestie gemaakt. Voor een kanaal met ruwe foutenkans $p=0,1$ bedragen de kansen op foute decoding voor de bovenstaande codes respectievelijk

$$\begin{aligned}
 P_{e,opt}(2,3) &= \binom{3}{2}p^2q + \binom{3}{3}p^3 = 0,028 \\
 P_{e,opt}(2,5) &= \binom{5}{3}p^3q^2 + \binom{5}{4}p^4q + \binom{5}{5}p^5 = 0,00856 \\
 P_{e,opt}(4,3) &= \binom{3}{1}pq^2 + \binom{3}{2}p^2q + \binom{3}{3}p^3 = 0,19.
 \end{aligned}$$

In het algemeen neemt $P_{e,opt}(M,N)$ af als N toeneemt en neemt $P_{e,opt}(M,N)$ toe met toenemende M . Als R constant wordt gehouden neemt $M=2^{RN}$ exponentieel toe met toenemende N . Zoals we net zagen heeft het toenemen van N een gunstig effect op de foutenkans $P_{e,opt}(M,N)$, terwijl het toenemen van $M=2^{RN}$ een nadelig effect heeft. De vraag is nu welk effect het wint. Shannon's antwoord luidt als volgt: voor vele kanalen kunnen we een constante C , de capaciteit, en een reeks codes met toenemende blok lengte N aangeven, zodat voor $R < C$ de foutenkans $P_{e,opt}(M,N)$ afneemt met toenemende blok lengte N , terwijl voor $R > C$ de foutenkans $P_{e,opt}(M,N)$ voor toenemende N tot 1 nadert. Voor het BSC van Fig. 7 met ruwe foutenkans $p=0,1$ bedraagt de kanaal capaciteit

$$C = 1 - H(0,1) = 0,47 \text{ bits per transmissie,} \quad (6)$$

waar de functie $H(\cdot)$ eerder werd gedefinieerd. Het bovenstaande wordt soms op hippe manier uitgedrukt door te zeggen dat C bits per transmissie foutloos kunnen worden overgedragen! In de volgende twee alinea's wordt op respectievelijk twee coderingstechnieken iets dieper ingegaan.

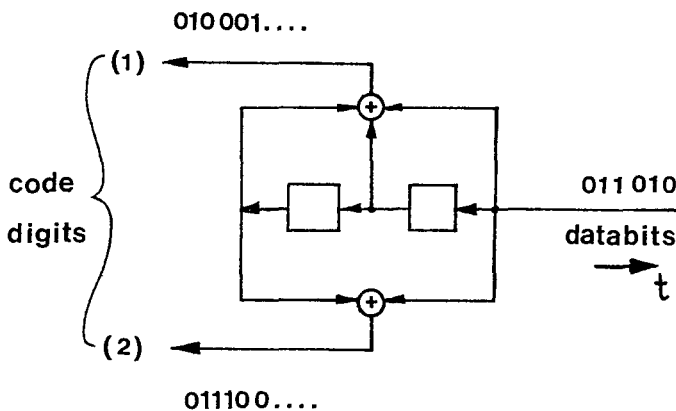


Fig. 9. Codeerschakeling voor een $R=1/2$ convolutie code.

Het in Fig. 9 afgebeelde circuit, met twee modulo-2 optellers en twee geheugen cellen, genereert een convolutie code. Voor iedere data bit worden twee code digits gegenereert, m.a.w. deze code heeft een efficiëntie van $R = \frac{1}{2}$ bit per digit. Met $v=2$ geheugen cellen heeft de in Fig. 9 afgebeelde codeerschakeling $2^v=4$ toestanden, t.w. $a=00$, $b=01$, $c=10$ en $d=11$. De code digits, bepaald door de input data bit en de toestand van de codeerschakeling, kunnen uit het 'trellis' diagram [19] van Fig. 10 worden afgelezen. In dit diagram correspondeert een getrokken lijn met een data 0 en een gestippelde lijn met een data 1. Een data input 11010... geeft b.v. een code output 1101010010..., corresponderend met het

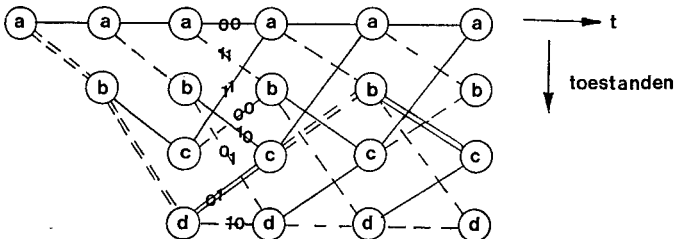


Fig. 10. De trellis van de codeerschakeling.

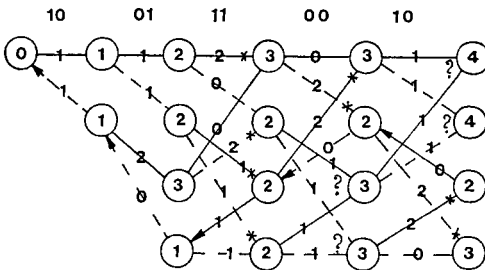


Fig. 11. De trellis van de decodeerschakeling.

dubbele pad in Fig. 10. Het decodeeralgoritme van een decodeerschakeling, die de Hammingafstand minimaliseert, kan eveneens m.b.v. een trellisdiagram, Fig. 11, worden beschreven. Neem aan dat de code digits 1101010010... worden ontvangen als 100110010.... Iedere overgang in de trellis van Fig. 11 kennen we nu een 'lengte' toe, die gelijk is aan de Hammingafstand tussen de bij die overgang behorende code digits, zie Fig. 10, en de corresponderende ontvangen digits. De 'metric' van een toe-

stand in Fig. 11 is nu de lengte van het kortste pad naar die betreffende toestand. De metric waarden op tijdstip $t+1$ kunnen recursief uit die op tijdstip t worden bepaald. Een $(t+1)$ -toestand kan uit twee t -toestanden worden bereikt. Tel de lengte van de overgang op bij de metric waarde van de bijbehorende t -toestand. Vergelijk het resultaat voor de twee overgangen naar de betreffende $(t+1)$ -toestand. De laagste waarde bepaalt de nieuwe metric van de $(t+1)$ -toestand. De overgang waarvoor de t -toestand metric plus de lengte van de overgang het kleinst is, wordt de 'survivor' genoemd. De andere overgang wordt doorgekruist. Decodeer van rechts naar links beginnend bij de toestand met de laagste metric, toestand c in Fig. 11, door steeds de niet doorgekruiste overgang (de survivor) te kiezen.

Het probleem met de decodeerschakeling volgens Viterbi (en in feite met iedere decodeerschakeling) is een probleem van 'computational complexity'. Het benodigde aantal componenten (IC's) is evenredig met het aantal, 2^v , trellis toestanden. Door deze exponentiële groei van het aantal componenten met v is de decodeerschakeling volgens Viterbi voor kleine foutenkans, d.w.z. voor grote v (≈ 10), praktisch niet langer realiseerbaar. De dialoog is in sommige asymmetrische situaties effectiever in het corrigeren van transmissie fouten dan Viterbi's monoloog. Gegeven een perfect terugmeldingskanaal, dat de zender informeert over eventuele transmissiefouten. Schalkwijk's [7] coderingsstrategie werkt dan als volgt: 1) precodeer de informatie zodanig dat de patronen 0111 en 1000 niet voorkomen, 2) codeer door iedere digit die fout overkomt drie (i.h.a. $k \geq 3$) maal te herhalen, en 3) corrigeer door $0111 \rightarrow 1$ en $1000 \rightarrow 0$ te vervangen.

Voorbeeld: 1) de te verzenden informatie is 10010..., 2) t.g.v. twee transmissiefouten (dakjes) verzenden we $1\hat{0}(0\hat{0}(000)0)010\dots$, 3) we ontvangen 11010000010..., en 4) decoderen $110(1000)0010 \rightarrow 1(1000)010 \rightarrow 10010$. Deze dialoog is optimaal in die zin, dat voor een bepaalde efficiëntie R de grootste fouten fractie f kan worden gecorrigeerd. Bovendien is het correctie algoritme zeer eenvoudig te implementeren.

Tenslotte wil ik mijn rede eindigen met een recente toepassing [8] van de Shannon theorie op het communicatie net van Fig. 12, waarbij een heilig huisje omver werd geworpen. Wanneer ik alleen met ontvanger (1) of (2) communiceer kan ik respectievelijk het efficiëntie-paar $(C_1, 0)$ of $0, C_2$) bereiken met

$$C_i = W \log \left(1 + \frac{P}{N_i W} \right); i=1,2 \quad (7)$$

Met tijd- of frequentie multiplexing kan ik ieder efficiëntie-paar op de lijn $(C_1,0) \rightarrow (0,C_2)$ bereiken. Volgens de traditionele leer zijn efficiëntie-

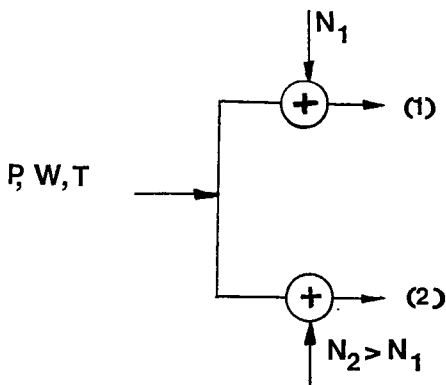


Fig. 12. Een zender (P,W,T) en twee ontvangers $(N_1$ en $N_2)$.

paren buiten het gearceerde gebied in Fig. 13 niet bereikbaar. Het blijkt nu echter dat code multiplexing beter is dan tijd- of frequentie multiplexing! De redenering gaat als volgt. Zend de informatie stromen naar

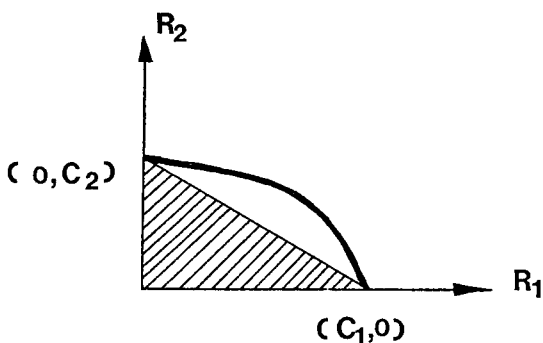


Fig. 13. Toelaatbare efficiëntie paren.

ontvanger (1) en (2) door elkaar in hetzelfde tijdsinterval $(0,T)$ en over dezelfde bandbreedte W . Als de slechte $(N_2 > N_1)$ ontvanger (2) de informatie stroom R_2 kan decoderen kan de goede $(N_1 < N_2)$ ontvanger (1) dat zeker ook. We kunnen nu de informatie voor ontvanger (2) bij ontvanger (1) aftrekken en een efficiëntie

$$R_1 = W \log \left(1 + \frac{P_1}{N_1 W} \right), \quad (8)$$

realiseren. Bij ontvanger (2) beschouwen we P_1 als extra ruis, m.a.w.

$$R_2 = W \log \left(1 + \frac{P_2}{P_1 + N_2 W} \right) \quad (9)$$

De efficiëntie-paren (R_1, R_2) volgens (8) en (9) liggen nu op de gebogen lijn in Fig. 13.

De moraal van de vorige alinea is de volgende. Volgens de traditionele leer had ik eerst de experts (N_1) moeten toespreken en vervolgens de leken (N_2) . Door mijn toespraak tot deze beide groepen tegelijk te richten heb ik in theorie meer informatie kunnen overdragen. Ik hoop dat ik hier ook in geslaagd ben!

Referenties

1. P. Elias, A. Feinstein and C. E. Shannon, 'A note on the maximum flow through a network', IRE Trans. Information Theory, vol. IT-2, pp. 117-119, December 1956.
2. J. P. M. Schalkwijk and T. Kailath, 'A coding scheme for additive noise channels with feedback - Part I: no bandwidth constraint', (1967 IEEE Information Theory Award) IEEE Trans. Inform. Theory, vol. IT-12, April 1966, pp. 172-182.
3. J. P. M. Schalkwijk, 'A coding scheme for additive noise channels with feedback - Part II: bandlimited signals', IEEE Trans. Inform. Theory, vol. IT-12, April 1966, pp. 183-189.
4. J. P. M. Schalkwijk, 'Centre-of-gravity information feedback', IEEE Trans. Inform. Theory, vol. IT-14, March 1968, pp. 324-331.
5. J. P. M. Schalkwijk, 'Recent developments in feedback communication', (Invited Paper) Proc. IEEE, vol. 57, July 1969, pp. 1242-1249.
6. J. P. M. Schalkwijk and M. E. Barron, 'Sequential signaling under a peak power constraint', IEEE Trans. Inform. Theory, vol. IT-17, May 1971, pp. 278-282.
7. J. P. M. Schalkwijk, 'A class of simple and optimal strategies for block coding on the binary symmetric channel with noiseless feedback', IEEE Trans. Inform. Theory, vol. IT-17, May 1971, pp. 283-287.
8. T. M. Cover, 'Broadcast channels', IEEE Trans. Inform. Theory, vol. IT-18, January 1972, pp. 2-14.
9. J. P. M. Schalkwijk and K. A. Post, 'On the error probability for a class of binary recursive feedback strategies', IEEE Trans. Inform. Theory, vol. IT-19, July 1973, pp. 498-511.
10. J. P. M. Schalkwijk, 'A coding scheme for duplex channels', to be published in the IEEE Trans. on Comm. Technology.
11. J. M. Wozencraft and I. M. Jacobs, Principles of communication engineering. New York: Wiley, 1965.
12. V. A. Kotelnikov, The theory of optimum noise immunity. New York: McGraw-Hill, 1959. (Doctoral dissertation presented in January 1947 before the academic council of the Molotov Energy Institute in Moscow).
13. D. Slepian, 'Bounds on communication', Bell Sys. Tech. J., vol. 42, May 1963, pp. 681-707.
14. J. P. M. Schalkwijk, L. I. Bluestein, 'Transmission of analog waveforms through channels with feedback', IEEE Trans. Inform. Theory, vol. IT-13, October 1967, pp. 617-618.
15. R. G. Gallager, Information theory and reliable communication. New York: Wiley, 1968.

16. J. P. M. Schalkwijk, 'An algorithm for source coding', IEEE Trans. Inform. Theory, vol. IT-18, May 1972, pp. 395-399.
17. J. P. M. Schalkwijk, F. Antonio and R. Petry, 'An efficient algorithm for data reduction', Proceedings of the Fifth Hawaii International Conference on System Sciences, 1972.
18. T. M. Cover, 'Enumerative source coding', IEEE Trans. Inform. Theory, vol. IT-19, January 1973, pp. 73-76.
19. A. J. Viterbi, 'Convolutional codes and their performance in communication systems', IEEE Trans. Communication Technology, vol. COM-19, October 1971, pp. 751-772.