07-103.

NETWERKEN ZIJN VERZAMELINGEN VAN ONDERLING VERBONDEN ELEMENTEN. ZE ZIJN OP ALLE NIVEAUS AANWEZIG, ZOWEL FYSISCH ALS ABSTRACT. MEER DAN OOIT WORDEN WE GECONFRONTEERD MET HET BESEF ALS INDIVIDU DEEL UIT TE MAKEN VAN NETWERKEN. IN WELKE ZIN GELIJKEN NETWERKEN VAN CELLEN, PROTEÏNEN, MENSEN EN COMPUTERS OP ELKAAR? BESTAAN ER ALGEMENE ONDERLIGGENDE PRINCIPES IN HET GEDRAG VAN COMPLEXE NETWERKEN? EN HOE GEVOELIG OF ROBUUST IS DE WERKING VAN NETWERKEN TEN OPZICHTE VAN EXTERNE AANVALLEN?

Gedrag in complexe netwerken

Johan A.K. Suykens



© Jan de Meue

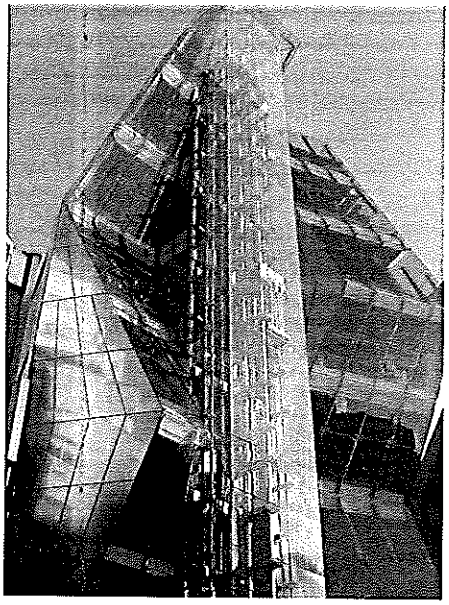
De opkomst van het internet heeft de mens meer dan ooit geconfronteerd met het besef dat we als individu deel uitmaken van netwerken. Het begrip netwerk kunnen we op verschillende niveaus opvatten, zowel fysisch als abstract. We maken deel uit van gemeenschappen of netwerken van vrienden, familie en professionele organisaties. In onze dagelijkse bezigheden zijn we sterk afhankelijk van het gebruik van elektriciteit en transportsystemen. Ook hier spelen netwerken een belangrijke rol, denken we maar aan het gevaar op het uitvallen van het elektriciteitsnetwerk of het ontstaan van eindeloze files in het verkeer bij een ongeval. Ook onze gezondheid wordt in hoge mate bepaald door de coördinatie in netwerken van genen, proteïnen en cellen. Een slechte werking ervan kan leiden tot levensbedreigende ziekten zoals kanker.

De structuur van een netwerk wordt wiskundig meestal voorgesteld in termen van een graaf.

De structuur van een netwerk wordt wiskundig meestal voorgesteld in termen van een *graaf*. Een graaf is een verzameling van knopen waartussen op bepaalde plaatsen verbindingen (*links*) bestaan. Voor het modelleren van vriendschapsbanden in een sociaal netwerk bijvoorbeeld stellen de knopen personen voor. Een vriendschapsband tussen twee personen wordt aangegeven met een verbinding tussen twee knopen. Als er geen vriendschapsband is, dan is er ook geen verbinding. In de wiskunde heeft men aanvankelijk vooral grafen bestudeerd waarbij verbindingen tussen knopen optreden op een willekeurige manier (de zogenaamde willekeurige netwerken). In willekeurige netwerken vertrekken uit de meeste knopen ongeveer een gelijk aantal verbindingen. In dit type netwerken vindt men zeer zelden knopen met opvallend meer of minder verbindingen dan het gemiddelde aantal. Die willekeurige netwerken werden meer dan veertig jaar geleden bestudeerd door de Hongaarse wiskundigen Paul Erdös en Alfréd Rényi.

In het laatste decennium is bij de studie van verschillende grootschalige netwerken, zoals het internet, wetenschappelijke publicatienetwerken en proteïne-interactienetwerken, echter aangetoond dat deze klassieke willekeurige netwerken geen geschikt

Karakter, nr. 20, 2007, pg. 26-27



model vormen. De bestudeerde netwerken vertonen immers een totaal verschillend karakter. Typisch hier is dat een beperkt aantal knopen het netwerk domineert (*hubs*). Men spreekt in dit geval van zogenaamde schaalvrije netwerken. In tegenstelling tot vroeger zijn er tegenwoordig veel data beschikbaar om de netwerkmodellen empirisch uit te testen en om de eigenschappen ervan te toetsen. Bij een studie van het *world wide web* door Albert-László Barabási (van de universiteit van Notre Dame) en zijn medewerkers bleek dat tachtig procent van de webpagina's minder dan vier links bezitten en slechts een kleine minderheid (0,01 procent van het aantal knopen) meer dan duizend links. Een statistische eigenschap van dit grootschalige netwerk is dat als men hoeveel webpagina's een bepaald aantal links bezitten, men vaststelt dat de (kans)verdeling ervan een machtswet volgt. De kans dat een knoop verbonden is met k andere knopen is evenredig met $1/k^{\gamma}$ waarbij γ in dit geval ongeveer gelijk is aan twee.

Opvallend is dat de studies van netwerken in totaal verschillende domeinen tot een gelijkaardige machtswet leiden. Het cellulaire metaboolnetwerk van verschillende levende organismen vertoont eveneens een schaalvrij karakter. In dit geval stellen de knopen moleculen voor waarbij een link verwijst naar een biochemische reactie. Volgens de bevindingen van Barabási zijn de meeste moleculen betrokken bij maximaal twee reacties, terwijl slechts een zeer beperkt aantal moleculen een rol speelt in bijna alle reacties. Er bestaat met andere woorden een klein aantal moleculen dat het metaboolnetwerk qua betrokkenheid domineert. Eenzelfde fenomeen doet zich voor in proteïne-interactienetwerken waarbij schaalvrij gedrag werd vastgesteld in de interactie bij duizenden proteïnen in gist. Een grote meerderheid van die proteïnen interageert met één of twee andere proteïnen, terwijl slechts een klein aantal betrokken is bij interacties met een zeer groot aantal proteïnen. Voor de bestudeerde netwerken in uiteenlopende domeinen is het niet alleen opmerkelijk dat ze alle de schaalvrije machtswet volgen, maar dat bovendien de waarde γ van de machtswet zich telkens typisch bevindt tussen de waarde twee en drie.

Vele netwerken zijn niet constant doorheen de tijd, maar zijn voortdurend in evolutie. Het schaalvrije gedrag kan ook worden begrepen vanuit die optiek. Bij het opstellen van groeimodellen voor de netwerken blijkt dat er een mechanisme van preferentiële aanhechting plaatsvindt. Dit betekent dat de knopen die in het begin vrij populair zijn en vele links bezitten in populariteit blijven toenemen. Een typisch voorbeeld van dit 'rijken worden rijker'-principe is het citeergedrag in de wetenschappelijke literatuur. Een veelgeciteerd artikel zet andere onderzoekers aan om het te lezen en eventueel ook te citeren. Ook in marketing is dit een veelgebruikte techniek voor het aanprijzen van nieuwe producten. Verrassend is dat vele biologische systemen aan dezelfde principes onderhevig zijn en dat men gelijkaardige bevindingen kan vaststellen voor verschillende netwerken onafhankelijk van het domein.

Een typisch voorbeeld van dit 'rijken worden rijker'-principe is het citeergedrag in de wetenschappelijke literatuur

Robuustheid en gevoeligheid zijn belangrijke aspecten voor netwerken in het algemeen. Hiermee wordt bedoeld in welke mate de werking van het globale systeem kan worden beïnvloed door externe verstoringen of door storingen aangebracht in de links of de structuur van het netwerk. Denken we bijvoorbeeld aan aanvallen door virussen in informatienetwerken, de verspreiding van epidemieën, fouten die optreden in cellen of proteïnen die verkeerd kunnen opvouwen. Een sterk punt is dat schaalvrije netwerken erg resistent zijn ten opzichte van toevallige fouten. In een proteïne-interactienetwerk van een cel bijvoorbeeld worden willekeurige mutaties goed doorstaan, de onaangetaste proteïnen blijven nog steeds in staat om samen verder te functioneren. De reden hiervoor is dat er in die netwerken toch voldoende knopen zijn die de taken kunnen overnemen. De netwerkeigenschappen worden er weinig door aangetast. De zwakke plek van schaalvrije netwerken is echter dat ze erg kwetsbaar kunnen zijn voor meer gerichte aanvallen die zich specifiek toespitsen op de *hubs*. Studies geven aan dat een gelijktijdige uitschakeling van een klein percentage (vijf tot vijftien procent) van het aantal *hubs* reeds kan leiden tot een globale netwerkcrash.

Het beschermen van de *hubs* lijkt daarom belangrijk voor het veiligstellen van het globale netwerk. Die algemene inzichten kunnen leerzaam zijn voor het ontwerp en bescherming van verschillende systemen. Bijvoorbeeld om de verspreiding van epidemieën tegen te gaan, zou het efficiënter kunnen zijn om meer gericht te vaccineren en zich toe te spitsen op de *hubs* van het netwerk. Ook bij de aanmaak van nieuwe geneesmiddelen zou men zich meer kunnen richten op de moleculen die de rol van *hubs* vervullen in het netwerk. Cyberaanvallen in informatienetwerken kunnen het best worden opgevangen door de *hubs* beter te beschermen.

Naast de besproken kwalitatieve eigenschappen zijn er ook andere karakteristieken die een belangrijke rol kunnen spelen, zoals bijvoorbeeld de diameter van het netwerk. De diameter is de maximale padlengte in het netwerk die moet worden afgelegd tussen een knoop en om het even welke andere knoop die men via het netwerk wenst te bereiken. Dit moet gebeuren volgens de kortst mogelijke route die beschikbaar is in het netwerk. Een netwerk kan tegelijk schaalvrij zijn en sterke clustervorming vertonen met clusters op verschillende hiërarchische niveaus. Voorbeelden hiervan zijn verschillende groepen van webpagina's met betrekking tot een specifiek onderwerp en groepen van moleculen die samen bepaalde functies verrichten binnen een cel. Ondanks de veelvuldige aanwezigheid van schaalvrij gedrag in netwerken, bestaan er netwerken en fysische systemen die andere eigenschappen vertonen, bijvoorbeeld atomen in een kristalrooster die alle een gelijk aantal links bezitten met de naburige atomen.

Naast de statistische eigenschappen van het netwerk, waarbij dikwijls abstractie wordt gemaakt van de onderliggende dynamica van het systeem, kan de niet-lineaire dynamica ook een grote rol spelen. Vanuit de optiek van dynamische systemen kan men synchronisatie-eigenschappen in complexe systemen bestuderen en begrijpen onder welke voorwaarden zich clustergroepen binnen netwerken kunnen vormen. Een misvatting is dat complex gedrag alleen maar in grootschalige systemen zou kunnen optreden. Vanuit de chaostheorie weten we dat zelfs zeer eenvoudige niet-lineaire systemen complex gedrag kunnen vertonen. *The Structure and Dynamics of Networks* van Newman *et al.* (2006) bevat een bundeling van een groot aantal reprints van tijdschriftpublicaties die de auteurs essentieel achten voor het nieuwe interdisciplinaire onderzoeksdomein van netwerken. Die publicaties zijn verdeeld in een aantal groepen die telkens beknopt begeleid worden door een inleidende tekst. Het boek biedt een interessant vertrekpunt voor de gespecialiseerde onderzoeker die zich een overzicht wenst te vormen van de huidige *state-of-the-art* in dit snelgroeiende domein. ●

M. Newman, A.-L. Barabási, D.J. Watts, *The Structure and Dynamics of Networks* (Princeton: Princeton Studies in Complexity 2006).
A.-L. Barabási, 'The Architecture of Complexity', in: *IEEE Control Systems Magazine*, 2007, 27 (4), 33-42.