De statespace van Small World Networks

Emiel Suilen, Daan van den Berg, Frank van Harmelen
epsuilen@few.vu.nl, daanvandenberg1976@gmail.com,
Frank.van.Harmelen@cs.vu.nl
VRIJE UNIVERSITEIT AMSTERDAM

2 juli 2010

Inhoudsopgave

1 Inleiding

Dit artikel is een vervolg op het eindverslag van het vak Heuristieken, waar een small world network werd onderzocht van 100 knopen en 300 verbindingen. Deze vervolgopdracht heeft als doel het verder verkennen van de statespace van deze configuratie, en het identificeren van de theoretische maxima en minima, met name de punten die een lokaal minimum of maximum vormen.

Deze statespace kan niet volledig worden doorgerekend. Een volledig verbonden graaf met K knopen heeft namelijk $\frac{K\times(K-1)}{2}$ mogelijke verbindingen. Dit zijn in totaal $\frac{100\times(100-1)}{2} = 4950$ verbindingen. Van deze 4950 verbindingen worden er echter maar 300 gebruikt. Dit levert een combinatie $\binom{4950}{300}$ op. Dit is gelijk aan $7,58 \times 10^{479}$.

De cluster coëfficiënt van een graaf C_g is het gemiddelde van de cluster coëfficiënten van de individuele knopen in graaf C_g. De cluster coëfficiënt van een knoop C_i is het aantal verbindingen dat zijn buren met elkaar hebben gedeeld door het aantal mogelijke verbindingen dat zijn buren zouden kunnen hebben.

De karakteristieke padlengte L_g van een graaf G is de gemiddelde kortste padlengte tussen twee knopen in die graaf.

Het vermoeden bestaat dat er een strikte relatie is tussen cluster coëfficiënt, en karakteristieke padlengte. Deze relatie wordt in dit artikel weerlegd.

2 Grenzen van Cluster coëfficiënt en Karakteristieke Padlengte

In de volgende paragrafen worden de minima en maxima voor zowel de cluster coëfficiënt en de karakteristieke padlengte besproken. Beide waardes hebben theoretische minima en maxima die onafhankelijk zijn van de daadwerkelijke grootte van de graaf.

2.1 Karakteristieke Padlengte

Voor de padlengte is er zowel een bovengrens als een ondergrens.

2.1.1 Minimale Padlengte: Pluis

De laagst mogelijke karakteristieke padlengte voor een graaf met K knopen en V verbindingen is 1. In dat geval kan je vanuit elke knoop in één stap naar elke andere knoop. Dit is alleen het geval voor een volledig verbonden graaf.

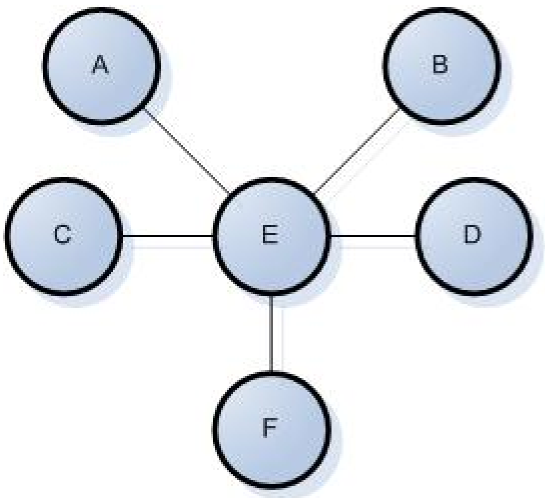
Voor grafen die minder dan $K - 1$ verbindingen hebben, is de padlengte ongedefineerd omdat er dan paren knopen zijn waartussen geen pad bestaat.

Voor verbonden grafen geldt wel een maximum. Omdat een pad tussen knoop k_i en k_j alleen 1 is dan en slechts dan als er een verbinding $v_{ij} \in V$ bestaat. Als er geen verbinding $v_{ij} \in V$ bestaat, dan is de kortst mogelijk padlengte tussen de knopen k_i en k_j 2. Dus is de kleinste karakteristieke padlengte van een verbonden graaf met K knopen en V verbindingen met $V \geq k$ gelijk aan

$$2 - \frac{2V}{K(K-1)}$$

Voor 100 knopen en 300 verbindingen is dit:

$$2 - \frac{2 \times 300}{100 \times 99} = 1,93939394$$



Figuur 1: Een graaf met zes knopen en vijf verbindingen met een kleinste padlengte. De padlengte is: $1,66667$.

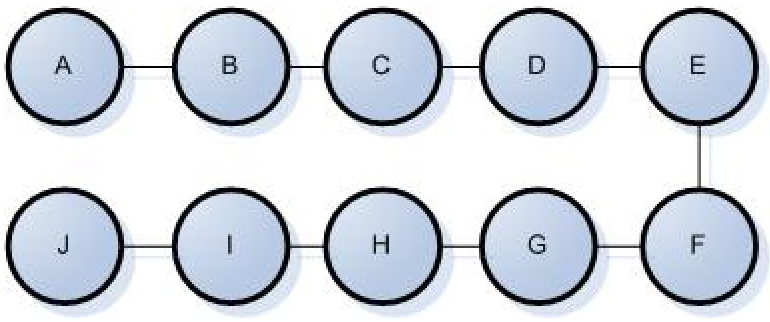
Merk op dat dit minimum triviaal is om te bereiken. Neem als voorbeeld figuur 1. Hier zijn slechts zes knopen en vijf verbindingen aanwezig. Omdat knoop E met alle andere knopen is verbonden zijn alle padlengtes vanaf deze knoop 1. Vanaf A zijn er vier paden met lengte twee, namelijk $(A;B),(A;C),(A;D),(A;F)$ en één pad met lengte 1, namelijk $(A;E)$. De karakteristieke padlengte van A is dus $\frac{9}{5} = 1,8$. Ditzelfde geldt natuurlijk ook voor B, C, D en F. De gemiddelde padlengte van deze graaf is dus $\frac{5 \times 1,8 + 1}{6} = 1,66667$.

Grafen met een minimale karakteristieke padlengte worden Pluis-grafen genoemd.

De theoretische kleinste padlengte van een graaf met 6 knopen en 5 verbindingen, is $2 - \frac{2 \times 5}{6 \times 5} = 1,66667$. Dit is niet heel bijzonder omdat de graaf in figuur 1 precies aan de optimale beschrijving voldoet. Wanneer twee knopen geen directe verbinding hebben, en dus geen padlengte van 1, dan hebben ze via knoop E altijd een indirecte verbinding met lengte 2.

2.1.2 Maximale padlengte: Staart

De hoogst mogelijke padlengte voor een graaf met K knopen is een lange sliert aan knopen, waarmee knoop k_i alleen verbonden is met de knopen $k_{(i-1)}$ en $k_{(i+1)}$. Deze graaf kan alleen gemaakt worden als er precies $K - 1$ verbindingen zijn. Figuur 2 geeft zo'n graaf met 10 knopen weer.



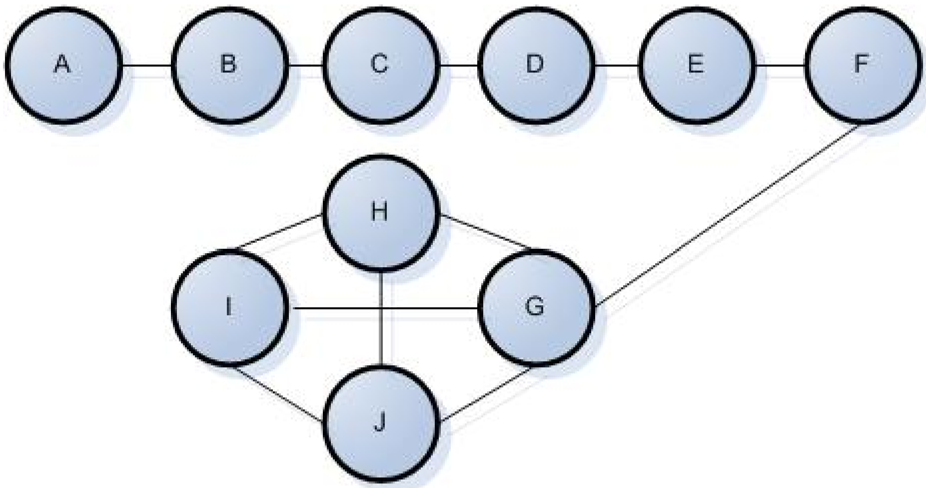
Figuur 2: Een graaf met de maximale karakteristieke padlengte voor 10 knopen. De padlengte is: $3,66667$.

Indien er meer dan $K - 1$ verbindingen zijn, is zo'n structuur onmogelijk: er blijven verbindingen over. Deze verbindingen moeten in zo'n configuratie worden gelegd, zodanig dat er veel verbindingen op een klein punt geconcentreerd zijn, die allen een pad naar één punt moeten vormen dat zo ver mogelijk van deze configuratie af ligt. Lovejoy en Loch[1] defineren zo'n figuur als volgt voor een graaf S met K knopen en V verbindingen:

1. Een volledig verbonden sub-graaf S_c van K_c knopen en V_c verbindingen.

2. Een staart S_s van $K - K_c$ knopen en $K - K_c - 1$, of V_s verbindingen.

Figuur 3 geeft zo'n figuur weer voor 10 knopen en 12 verbindingen. De clustering van de knopen G, H, I en J introduceert weliswaar paden met een padlengte van 1, maar deze vier knopen hebben ook ieder een pad met een hoge padlengte naar knoop A. Zo'n type graaf wordt een Staart-graaf genoemd.



Figuur 3: Een graaf met de maximale karakteristieke padlengte voor 10 knopen en 12 verbindingen. De padlengte is: $3,17778$.

In het algemeen geeft dit de volgende karakteristieke padlengte:

$$\frac{\gamma}{K(K-1)} \tag{1}$$

Met γ is als volgt gedefinieerd:

$$\gamma = K_c(K_c-1) + \frac{K-K_c}{3}(9K_c - 7 + (K-K_c)(K+2K_c)) \tag{2}$$

Voor een graaf met 10 knopen en 12 verbindingen geeft dit een karakteristieke padlengte van $3,17778$.

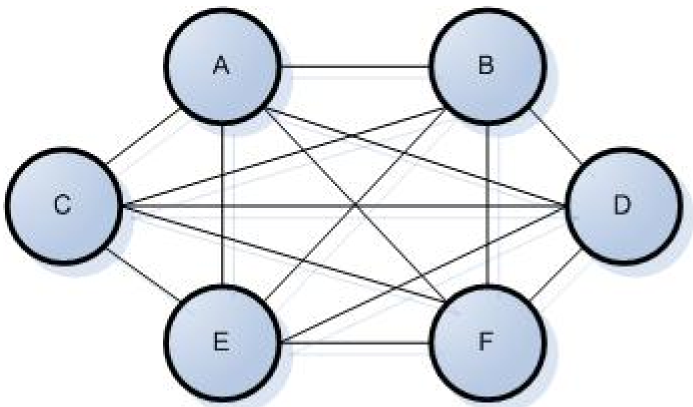
Merk in figuur 3 op dat knoop G zowel onderdeel is van de staart als van de sub-graaf. Deze knoop wordt de stuit-knoop genoemd.

Echter, ook zo'n graaf-configuratie is beperkt: hij is alleen mogelijk als er, bij K knopen, een subgraaf met K_c knopen en V_c verbindingen en een staart van V_s verbindingen, er precies $V_c + V_s$ verbindingen zijn. De oplossing voor dit probleem is echter simpel. Als er te veel verbindingen zijn, moet er een subgraaf met $K_c + 1$ knopen gemaakt worden, met een staart van $K_s - 1$ knopen. Dit levert een graaf op met te weinig verbindingen. Door verbindingen bij de stuit-knoop weg te halen, wordt het gewenste aantal verbindingen bereikt.

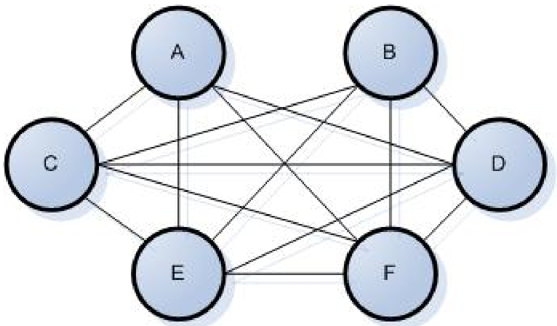
In het geval van een graaf van 100 knopen en 300 verbindingen zijn er 9 verbindingen te weinig om een sub-graaf van 22 knopen en een staart van 78 knopen te maken. Deze worden van de stuit-knoop verwijderd, zodat er in totaal 300 verbindingen nodig zijn. Dit levert een karakteristieke padlengte van $30,19010$ op.

2.2 Cluster coëfficiënten

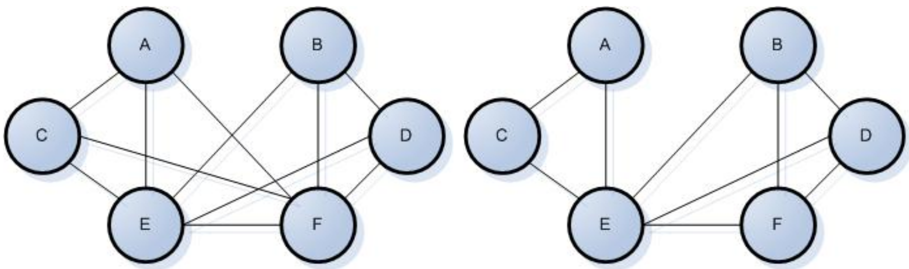
Het hoogst haalbare cluster coëfficiënt voor een ongerichte graaf is 1. Voor een verbonden graaf betekent dit dat ze volledig verbonden is. Wanneer er, in een volledig verbonden graaf, een verbinding tussen twee knopen wordt verbroken dan zullen alle andere knopen in de graaf een lagere cluster coëfficiënt krijgen. In figuur 5 is de verbinding tussen A en B verbroken. De knopen C tot en met F zijn nog wel verbonden met zowel A als B. Omdat A en B elkaar niet meer kennen daalt de cluster coëfficiënt van C tot en met F van 1 naar $0,9$. De cluster coëfficiënt van de graaf daalt hierdoor van 1 naar $0,933$.



Figuur 4: Een graaf met zes knopen. De cluster coëfficiënt is 1.



Figuur 5: Dezelfde graaf als in figuur 4 maar dan met de verbinding tussen A en B verbroken. De cluster coëfficiënt daalt naar 0,933.



(a) Graaf met 6 knopen en 11 verbindingen. De cluster coëfficiënt is 0,866.

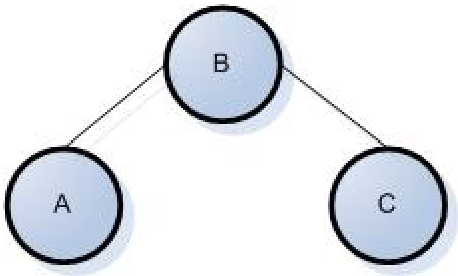
(b) Graaf met 6 knopen en 9 verbindingen. De cluster coëfficiënt is 0,9.

Figuur 6: Twee grafen met 6 knopen. Hoewel figuur 6b minder verbindingen heeft dan figuur 6a heeft deze toch een hoger cluster coëfficiënt.

Het is dus niet mogelijk om een verbonden graaf te maken die niet volledig verbonden is, maar wel een cluster coëfficiënt van 1 heeft. Hierdoor is het, met minder dan het maximaal aantal verbindingen, niet mogelijk om een verbonden graaf te maken met een cluster coëfficiënt van 1. Het is echter niet zo dat het verbreken van een verbinding altijd leidt tot een lagere cluster coëfficiënt. Figuur 6a toont een graaf met 11 verbindingen en een cluster coëfficiënt van $0,8666$. Merk op dat de verbindingen zo gelegd zijn dat er zoveel mogelijk knopen zijn met een cluster coëfficiënt van 1. Figuur 6b laat een graaf zien met 9 verbindingen en cluster coëfficiënt van $0,9$. Deze graaf heeft twee verbindingen minder maar wel een hogere cluster coëfficiënt. Het verwijderen van verbindingen in figuur 6b zal de cluster coëfficiënt van deze graaf weer verkleinen.

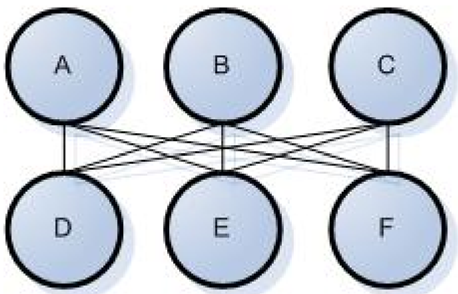
2.2.1 Minimale Cluster Coëfficiënt

De minimale cluster coëfficiënt voor een ongerichte graaf is afhankelijk van de verbindingsdichtheid. Als voor ieder drietal knopen (A, B, C) in de graaf geldt dat als $(A; B)$ én $(B; C)$ en niet $(A; C)$ dan is de cluster coëfficiënt van de graaf nul. Figuur 7 geeft zo'n graaf weer.



Figuur 7: Een graaf met 3 knopen en een cluster coëfficiënt van nul.

Het aantal verbindingen waarvoor dit mogelijk is, is in ieder geval $K-1 \leq V \leq \frac{1}{2}K^2$. In deze ongelijkheid is $\frac{1}{2}K^2$ het aantal verbindingen waarmee de knopen in twee evengrote groepen worden verdeeld en iedere knoop in één groep wordt verbonden met alle knopen in de andere groep; $K - 1$ is de ondergrens voor verbondenheid van de graaf. Omdat $99 \leq 300 \leq 1250$ is de minimale cluster coëfficiënt van een graaf van 100 knopen en 300 verbindingen nul.



Figuur 8: Graaf met 6 knopen en 9 verbindingen. De cluster coëfficiënt is hier 0.

2.2.2 Maximale Cluster Coëfficiënt: JEB-graaf

De maximale cluster coëfficiënt voor een ongerichte graaf is 1. Voor een niet-verbonden graaf met K knopen en V verbindingen betekent dit een verzameling losse modules $m_1, m_2 \ldots m_k$ waarvoor geldt:

- $m_1 + m_2 + \ldots + m_k = K$
- $\frac{1}{2}(m_1(m_1 - 1)) + \frac{1}{2}(m_2(m_2 - 1)) + \frac{1}{2}(m_k(m_k - 1)) = V$

Een verbonden graaf met de maximale cluster coëfficiënt is volledig verbonden, zodat $V = \frac{1}{2}K(K-1)$. Voor een niet volledig verbonden graaf van K knopen en V verbindingen geldt, dat als $v \geq 1\frac{1}{2}(K-1)$, er een maximale cluster coëfficiënt van $\frac{K-1}{K}$ bestaat. Voor een graaf van 100 knopen betekent dit, dat mits het aantal verbindingen groter of gelijk is aan 144, er een graaf bestaat met een cluster coëfficiënt van $\frac{99}{100}$.

Ondanks dat de verwachte cluster coëfficiënt voor random grafen van K knopen lineair proportioneel is aan V, laat bovenstaande zien dat zelfs voor grafen met extreem lage connectiedichtheden zeer hoge cluster coëfficiënten mogelijk zijn.

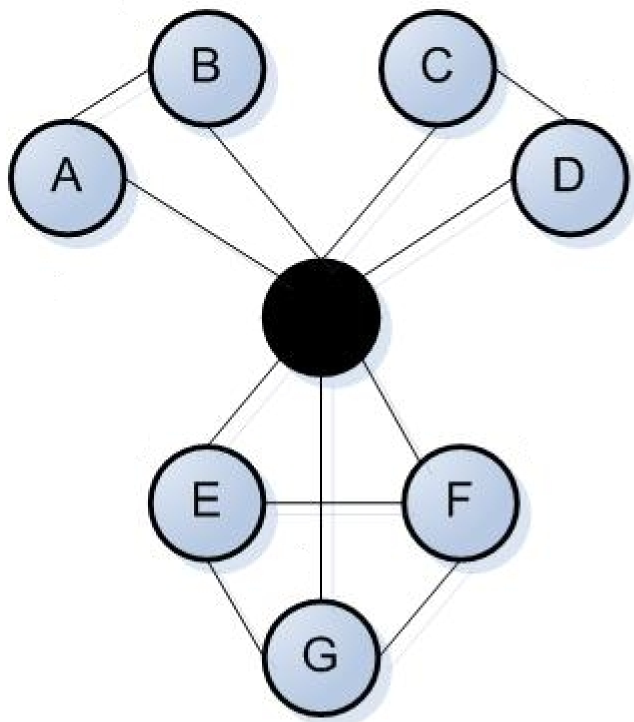
Dit soort grafen, met een maximale cluster coëfficiënt, hebben tevens een minimale padlengte, en zijn dus een onderklasse van de Pluis-grafen. Een graaf uit deze onderklasse heet de JEB-graaf, en heeft de volgende eigenschappen:

1. Één centrale knoop, die met alle andere knopen verbonden is.
2. Alle overige knopen zijn verbonden in clusters, die allemaal onderling verbonden zijn. Elke knoop in zo'n cluster heeft hierdoor een cluster coëfficiënt van 1.

Dit zorgt er voor, dat een graaf met K knopen een cluster coëfficiënt bereikt van $\frac{K-1}{K}$. De cluster coëfficiënt zal zelfs iets hoger zijn, want ook de centrale knoop heeft een cluster coëfficiënt waarde, ongelijk aan nul. Dit levert in een graaf met K knopen en V verbindingen de volgende cluster coëfficiënt op:

$$\frac{\frac{2|V-K+1|}{(K-1)(K-2)} + K - 1}{K} \tag{3}$$

Voor een graaf met 100 knopen en 300 verbindingen is de cluster coëfficiënt dan $0,99041$.



Figuur 9: Graaf met 8 knopen en 12 verbindingen. De cluster coëfficiënt is hier $0,90476$. De cluster coëfficiënten van de knopen A, B, C, D, E, F en G zijn allemaal 1. Tevens heeft deze graaf de minimale karakteristieke padlengte: $1,5714$

3 Empirische resultaten

In het vorige hoofdstuk zijn de boven- en ondergrenzen van de statespace voor een graaf van 100 knopen en 300 verbindingen beschreven. In dit hoofdstuk zal een algoritme besproken worden dat grafen binnen dit gebied vindt, en zullen de resultaten van dit algoritme weergegeven worden.

3.1 Algoritme

Om de statespace door te rekenen is er een algoritme gemaakt dat werkt als een serie van non-deterministische hill-climbers. Het algoritme bestaat uit drie onderdelen:

1. Punt-algoritme: zoek naar een punt in de statespace.

2. Lijn-algoritme: zoek naar een lijn van punten in de statespace.

3. State-algoritme: zoek naar een serie aan lijnen in de statespace.

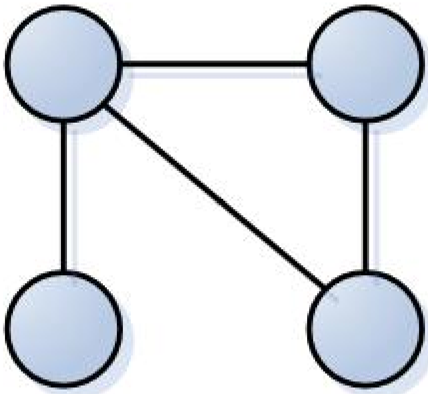
Het algoritme is in eerste instantie opgezet om op cluster coëfficiënt te zoeken. Het kan echter ook gebruikt worden om op karakteristieke padlengte te zoeken.

3.1.1 Punt-algoritme

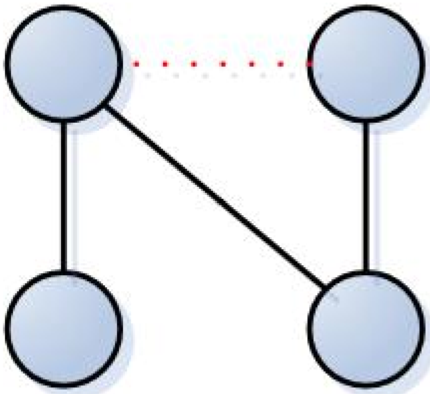
Dit onderdeel van het algoritme zoekt een nieuw punt in de statespace vanuit een gegeven graaf. Hierbij wordt gezocht naar een maximale cluster coëfficiënt. Vanuit de oorspronkelijke graaf worden 100 nieuwe grafen gemaakt, door telkens random één verbinding te verleggen. Alle 100 grafen die hieruit ontstaan en volledig verbonden zijn worden hierna beoordeeld op de cluster coëfficiënt. De graaf met de hoogste cluster coëfficiënt, is dan de nieuwe graaf én een nieuw punt in de statespace.

In plaats van de maximale cluster coëfficiënt, kan er ook naar een minimale cluster coëfficiënt, of een gegeven waarde gezocht worden. Dit wordt gebruikt in het Lijn-algoritme.

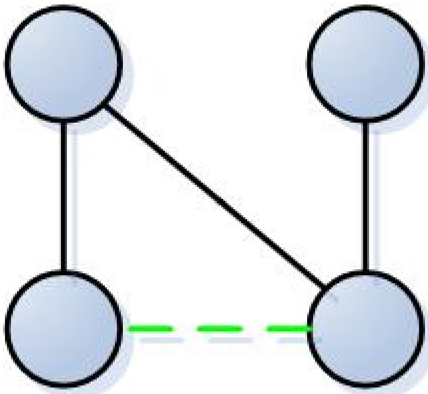
Hier volgt een voorbeeld van het Punt-algoritme, op een graaf van 4 knopen en 4 verbindingen:



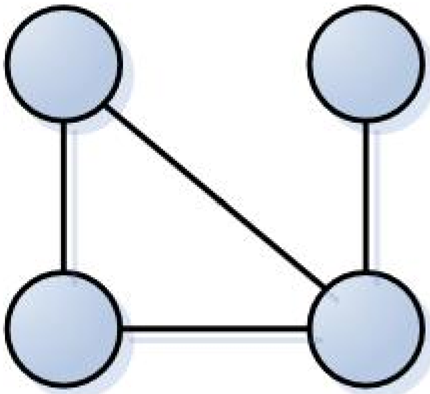
(a) De gegeven graaf



(b) De selectie van een willekeurige verbinding om te verwijderen



(c) De selectie van twee willekeurige knopen om een verbinding tussen te leggen



(d) De resulterende graaf

Figuur 10: Het Punt-algoritme

3.1.2 Lijn-algoritme

Met dit algoritme wordt een lijn aan punten binnen de statespace gezocht. Dit algoritme begint weer met een gegeven graaf. Vervolgens wordt met behulp van het Punt-algoritme naar een gegeven cluster coëfficiënt waarde gezocht, binnen een vaste afwijking. Als een graaf is gevonden met de gegeven cluster coëfficiënt, wordt de zoek-eigenschap omgekeerd: er wordt nu met behulp van het Punt-algoritme naar de maximale karakteristieke padlengte gezocht. Echter, elke graaf die nu wordt gevonden moet ook nog de gegeven cluster coëfficiënt hebben, binnen een vaste afwijking. Dit vormt een lijn aan punten met een gegeven cluster coëfficiënt.

Het Lijn-algoritme maakt 1000 keer gebruik van het Punt-algoritme, om zowel de gegeven cluster coëfficiënt waarde te vinden, en ook om vervolgens de lijn te vinden. Er wordt dus in totaal naar maximaal 1000 punten gezocht. Dit is om te voorkomen dat het algoritme in een lokaal maximum verzeild raakt, en niet meer naar de gegeven waarde kan rekenen, zonder meerdere stappen terug te doen, die afzonderlijk een slechter resultaat op zouden leveren.

Het algoritme levert ook een beste resultaat op. Dit is, als de gegeven cluster coëfficiënt is gevonden, de graaf met de hoogste karakteristieke padlengte, en anders de graaf die het dichtst bij de gegeven cluster coëfficiënt in de buurt ligt.

Hier volgt een voorbeeld van het Lijn-algoritme. Er wordt gezocht naar een lijn aan punten met een cluster coëfficiënt van 0,800, en er wordt gezocht vanuit de JEB-graaf.

1. De JEB-graaf wordt als beginpunt meegegeven aan het Lijn-algoritme.

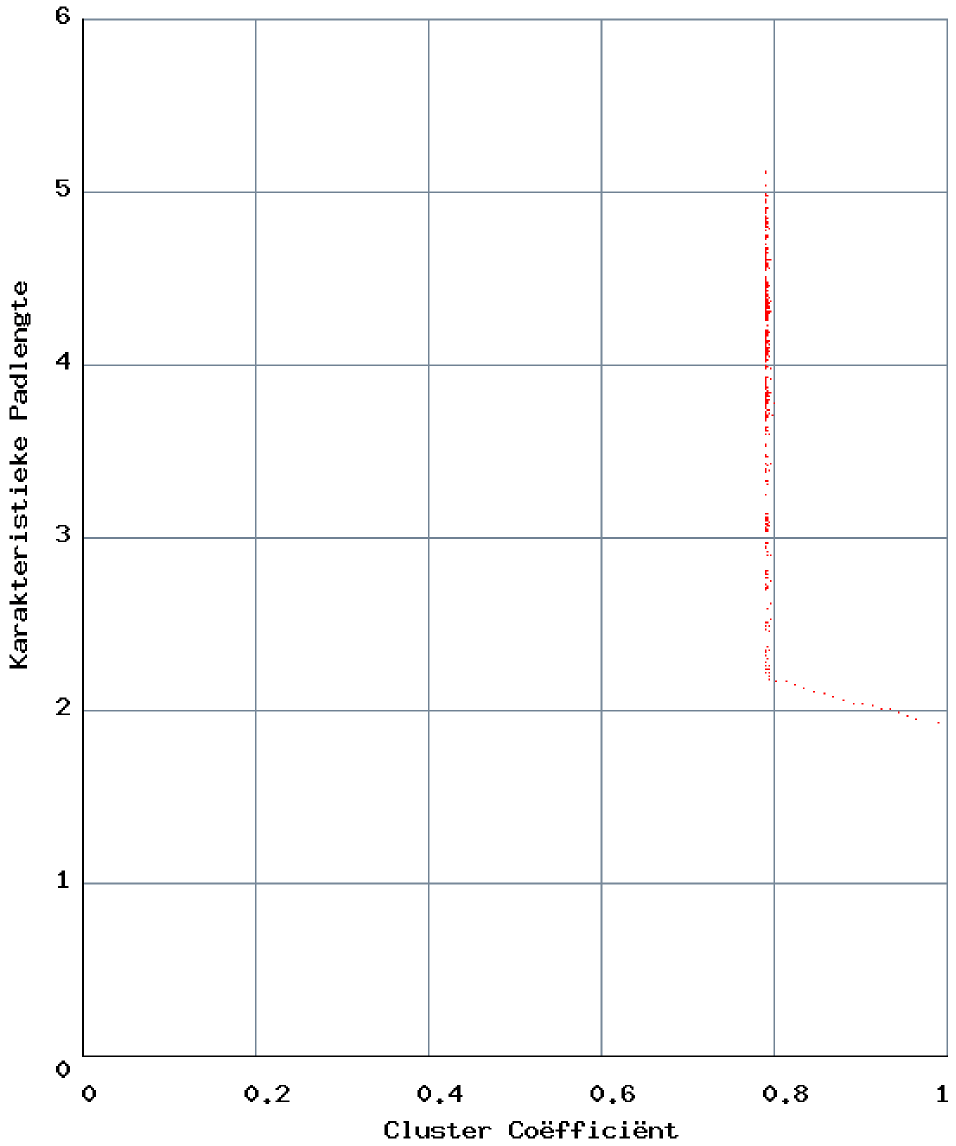
2. Met behulp van het Punt-algoritme wordt er een nieuwe graaf berekend, met een cluster coëfficiënt die zo dicht mogelijk bij 0,800 ligt. Deze nieuwe graaf heeft een cluster coëfficiënt van 0,9650.

3. Het Punt-algoritme wordt in totaal 15 keer herhaald, totdat de 16^e graaf een cluster coëfficiënt van 0,8039 heeft.

4. Nu wordt er naar de maximale karakteristieke padlengte gezocht. De graaf moet wel een cluster coëfficiënt van $0,8000 \pm 0,01$ hebben.

5. Na 984 keer het Punt-algoritme gebruikt te hebben, is er uiteindelijk een graaf gevonden met een cluster coëfficiënt van 0,7905, en een karakteristieke padlengte van 5,1226.

Net zoals het Punt-algoritme, kan er in plaats van eerst naar de cluster coëfficiënt te zoeken, eerst naar de karakteristieke padlengte gezocht worden. Tevens kan er, in plaats van naar de maximale padlengte te zoeken, er naar de minimale padlengte gezocht worden.



Figuur 11: Resultaat van Lijn-algoritme, rekenend naar een cluster coëfficiënt van 0,800

3.1.3 State-algoritme

Met dit laatste deel van het algoritme wordt binnen de complete statespace gezocht, door herhaaldelijk het Lijn-algoritme toe te passen. Er wordt een begin-graaf gedefineerd, vanuit waar de lijnen worden gezocht. Vervolgens wordt vanaf de ondergrens van de cluster coëfficiënt door de statespace gezocht, totdat de bovengrens is bereikt. Dit gebeurt in intervallen van 0,001 groot.

Elk interval is als volgt gedefineerd: de begingraaf wordt aan het Lijn-algoritme meegegeven, plus de waarde van het interval. Als deze waarde is bereikt, wordt de zoek-eigenschap omgedraaid, en wordt naar de maximale karakteristieke padlengte gezocht, totdat het Lijn-algoritme 1000 punten heeft gevonden.

Zodra het Lijn-algoritme 1000 punten heeft gevonden, wordt het beste resultaat meegegeven aan het Lijn-algoritme, om nog een keer 1000 punten uit te rekenen. Er wordt opnieuw naar de maximale karakteristieke padlengte gezocht. Deze stap is om te voorkomen dat, als het algoritme na een goed resultaat gevonden te hebben verzeild raakt in een lokaal maximum, het algoritme in dit lokale maximum blijft steken.

Deze tweede uitvoering van het Lijn-algoritme levert opnieuw een beste resultaat op. Dit beste resultaat wordt opnieuw als basis meegegeven aan het Lijn-algoritme, maar nu wordt er naar een minimale karakteristieke padlengte gezocht. Dit levert weer een beste resultaat op, wat een vierde keer wordt meegegeven aan het Lijn-algoritme, en er wordt voor de tweede keer naar een graaf met een minimale karakteristieke padlengte gezocht.

In totaal wordt het Lijn-algoritme vier keer uitgevoerd, wat 4000 punten oplevert. Elk interval bestaat dus uit 4000 punten. Vervolgens wordt de waarde van het interval met de intervalgrootte verhoogt, en wordt er opnieuw 4 keer het Lijn-algoritme toegepast, net zolang tot de bovengrens van de cluster coëfficiënt, naar boven afgerond, bereikt is.

Hier volgt een voorbeeld van het State-algoritme dat op cluster coëfficiënt 0,8 en 0,801 zoekt, met als begingraaf de JEB-graaf.

1. Het State-algoritme begint met zoeken naar een graaf met cluster coëfficiënt rond de $0,8000 \pm 0,01$.

2. De JEB-graaf wordt als beginpunt meegegeven aan het Lijn-algoritme.

3. De 16^e graaf heeft een cluster coëfficiënt van $0,8039$. Er wordt nu naar de maximale karakteristieke padlengte gezocht.

4. Na 1000 punten berekend te hebben, heeft het Lijn-algoritme een graaf

gevonden met een cluster coëfficiënt van $0,7905$, en een karakteristieke padlengte van $5,1226$. Figuur 12 geeft dit weer.

5. Deze graaf wordt als nieuw beginpunt meegegeven aan het Lijn-algoritme, die opnieuw naar een graaf gaat zoeken die een cluster coëfficiënt heeft rond de $0,8000 \pm 0,01$, en die een maximale cluster coëfficiënt heeft.

6. Na 1000 punten uitgerekend te hebben, heeft het Lijn-algoritme een graaf gevonden met een cluster coëfficiënt van $0,7905$, en een karakteristieke padlengte van $5,8378$. Figuur 13 geeft dit weer.

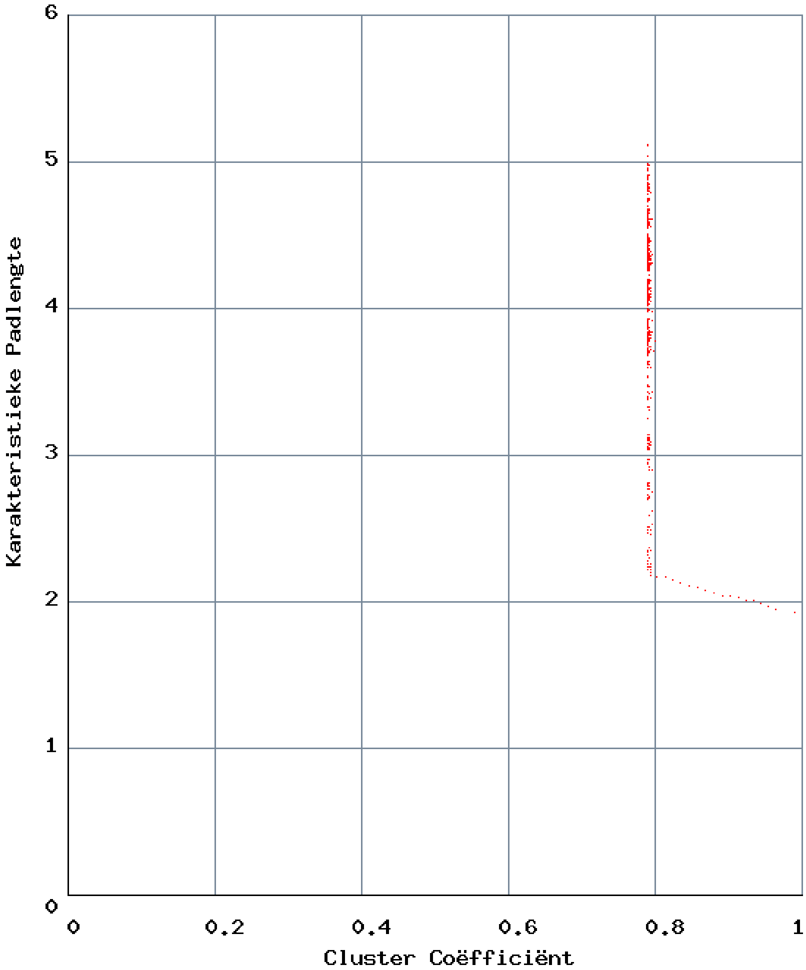
7. Met deze nieuwe graaf wordt een derde keer het Lijn-algoritme gestart. Deze gaat wederom op zoek naar een graaf met een cluster coëfficiënt rond de $0,8000 \pm 0,01$, maar dit keer zoekt hij naar een minimale karakteristieke padlengte.

8. Na 1000 punten uitgerekend te hebben, heeft het Lijn-algoritme een graaf gevonden met een cluster coëfficiënt van $0,7902$ en een karakteristieke padlengte van $1,9394$. Figuur 14 geeft dit weer.

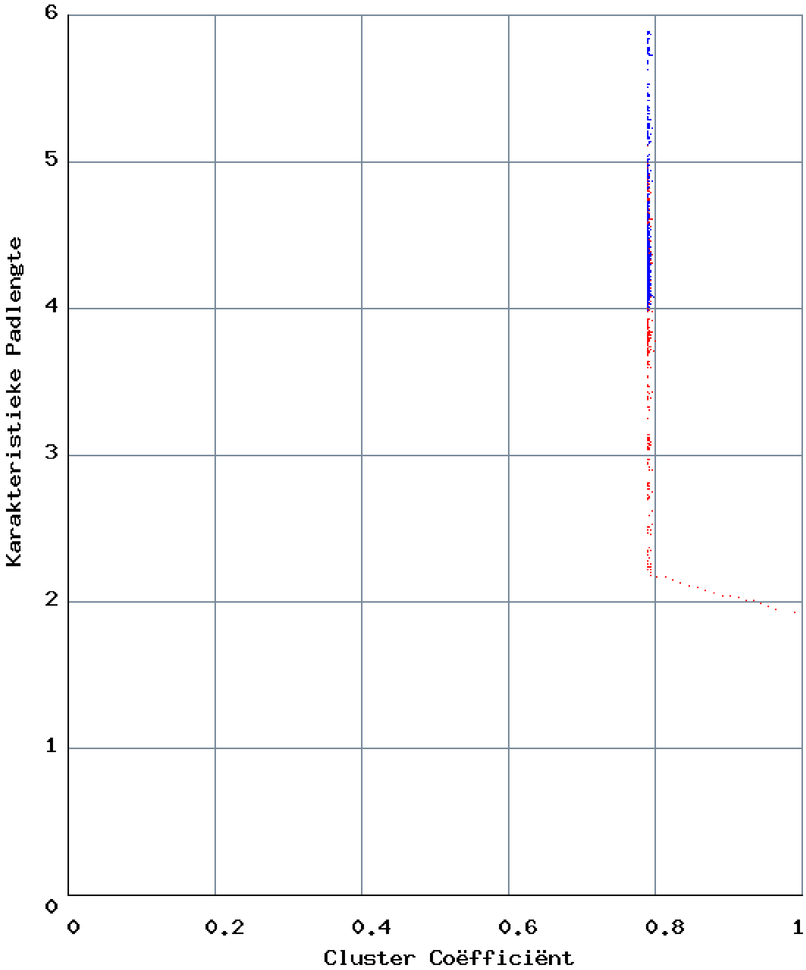
9. Met deze graaf wordt een laatste maal het Lijn-algoritme gestart, om opnieuw 1000 punten met een cluster coëfficiënt rond de $0,8000 \pm 0,01$ te zoeken, en met een minimale karakteristieke padlengte. Figuur 15 geeft dit weer.

10. Na vier keer het Lijn-algoritme gestart te hebben, is het State-algoritme klaar met het berekenen van de Statespace met een cluster coëfficiënt van $0,8$. Het State-algoritme gaat nu zoeken naar de volgende interval, namelijk een cluster coëfficiënt van $0,8010 \pm 0,01$. Figuur 16 geeft dit weer.

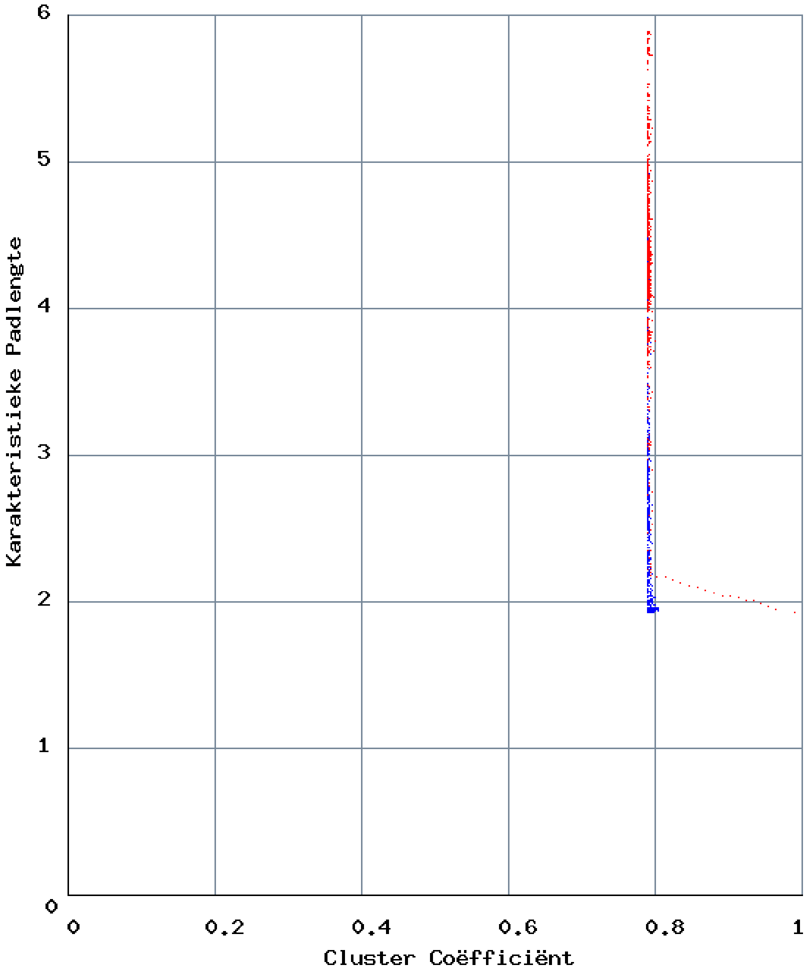
De volgende 5 pagina's bevatten illustraties over de werking van het State-algoritme.



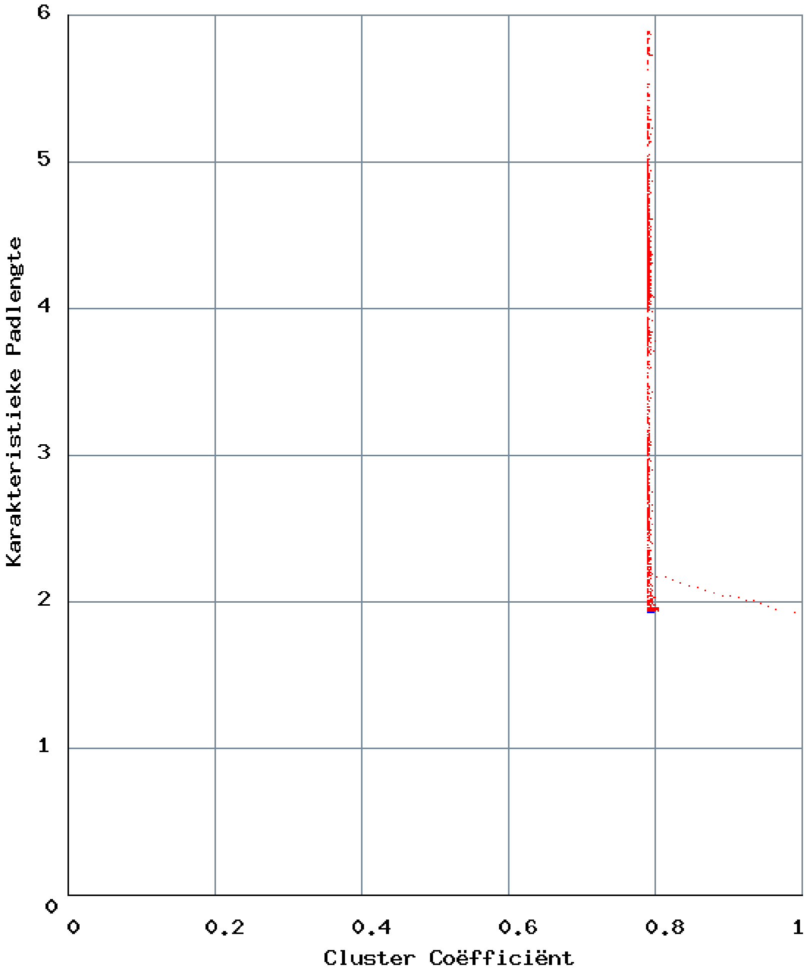
Figuur 12: De eerste stap van het State-algoritme, gezocht naar een cluster coëfficiënt van 0,800. 1000 punten



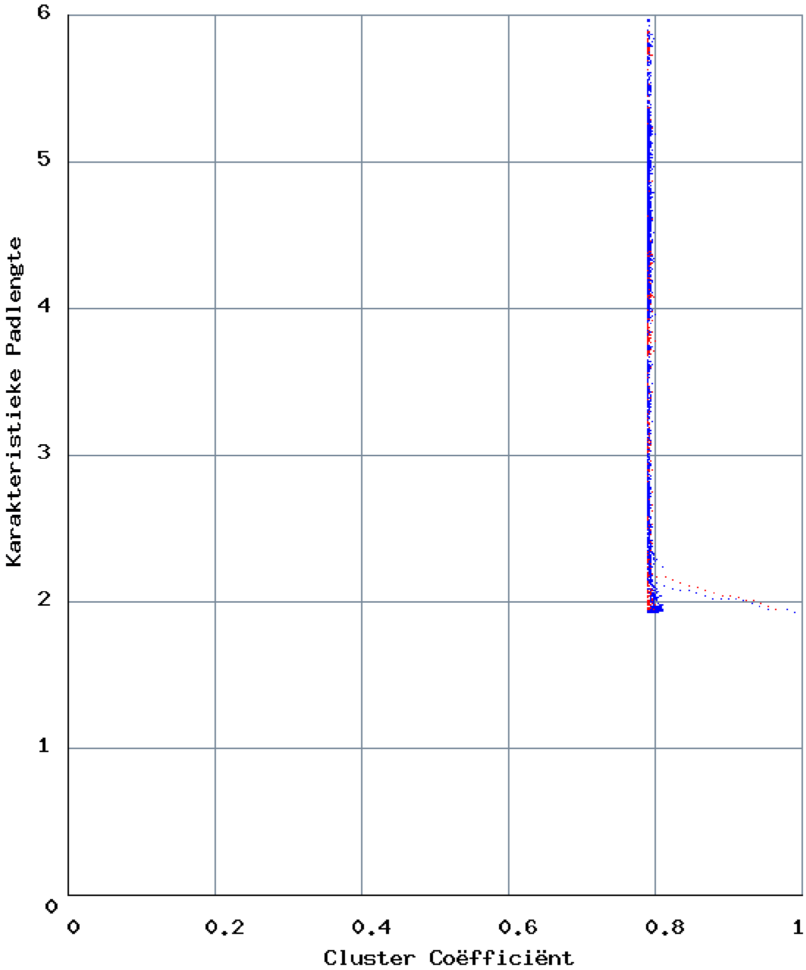
Figuur 13: De tweede stap van het State-algoritme, gezocht naar een cluster coëfficiënt van 0,800. De nieuwe punten zijn blauw. 2000 punten



Figuur 14: De derde stap van het State-algoritme, gezocht naar een cluster coëfficiënt van 0,800. De nieuwe punten zijn blauw. 3000 punten



Figuur 15: De vierde stap van het State-algoritme, gezocht naar een cluster coëfficiënt van 0,800. De nieuwe punten zijn blauw. 4000 punten

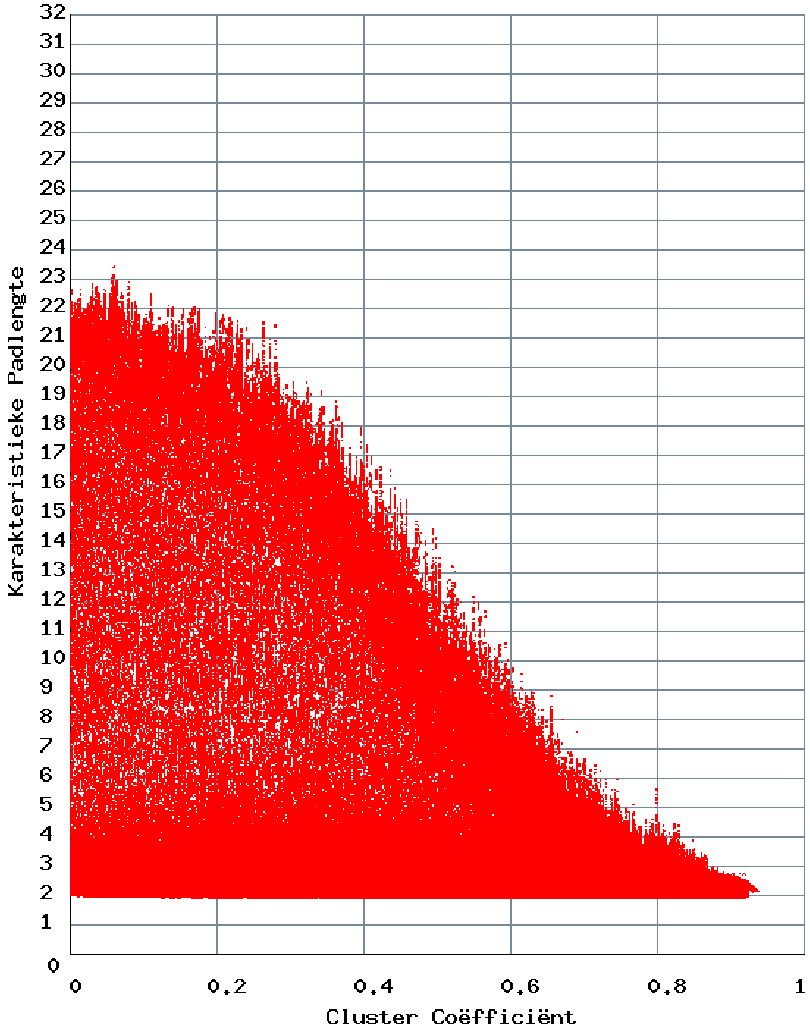


Figuur 16: Het volgende interval van het State-algoritme, gezocht naar een cluster coëfficiënt van 0,801. De nieuwe punten zijn blauw. 8000 punten

3.2 Statespace

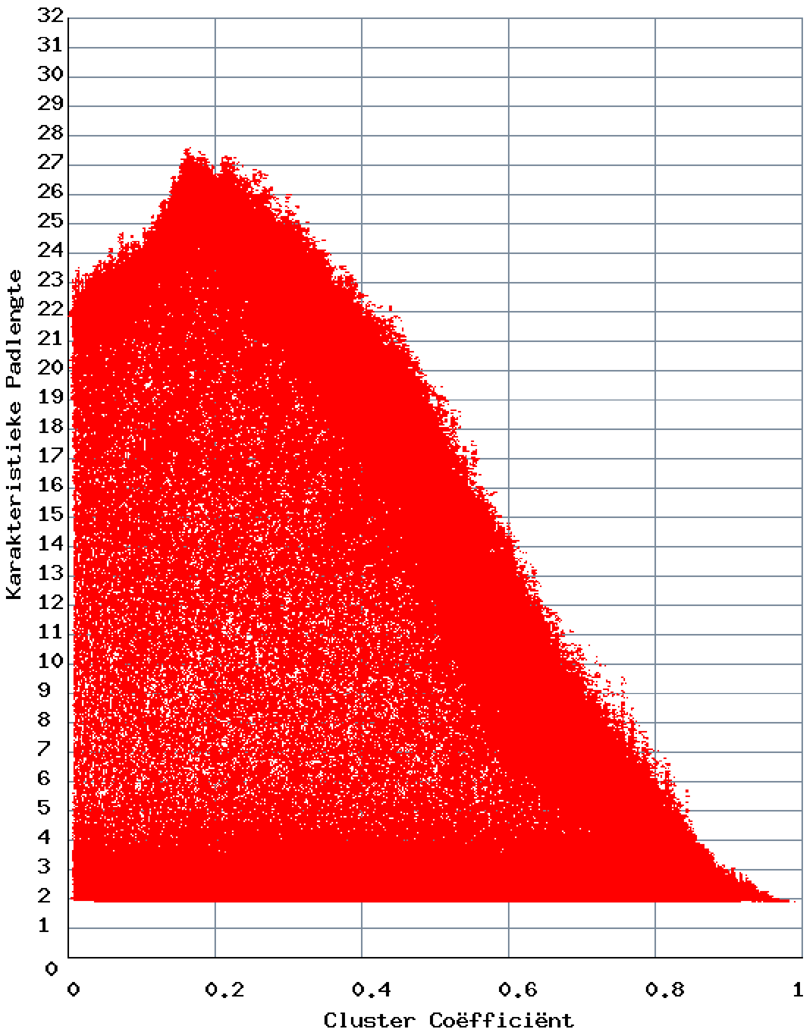
Het State-algoritme is in totaal vier keer uitgevoerd. Drie keer werd er naar de cluster coëfficiënt gezocht, en één keer naar de karakteristieke padlengte. De cluster coëfficiënt werd vanaf 3 begingrafen doorzocht: een telkens willekeurig gekozen graaf, een JEB-graaf en een Staart-graaf. De karakteristieke padlengte werd vanaf een willekeurig gekozen graaf doorzocht. De volgende vijf pagina's laten de resultaten hiervan zien.

3.2.1 Cluster Coëfficiënt vanuit random graaf



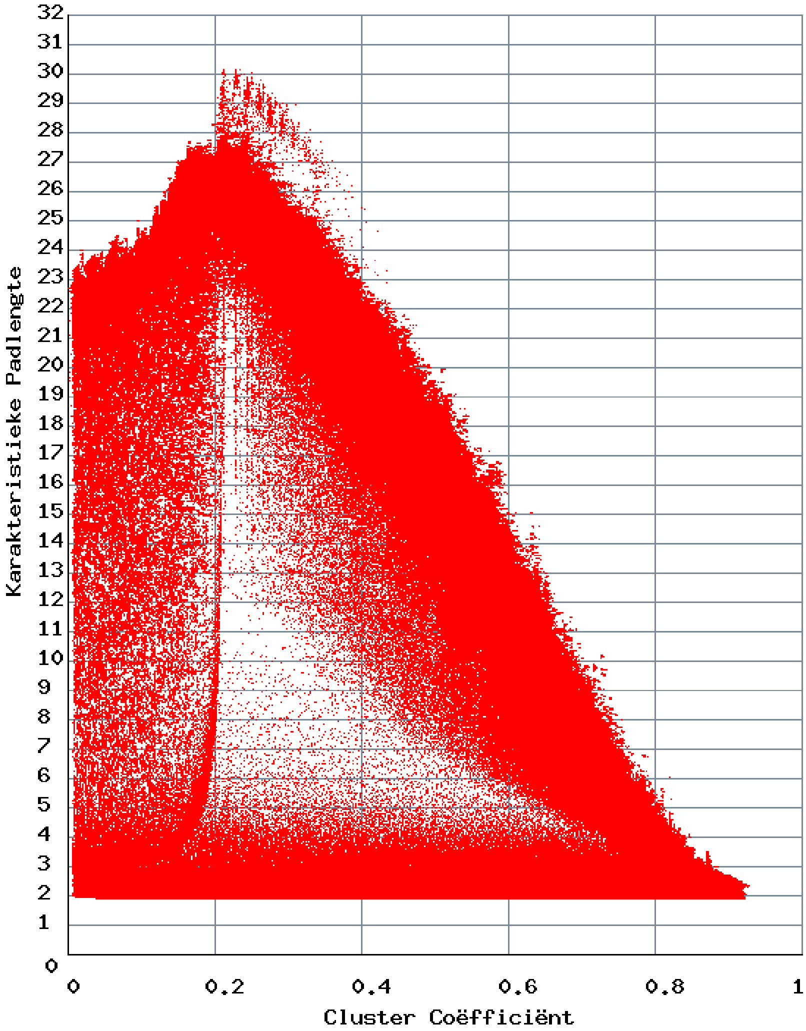
Figuur 17: Statespace doorzocht op Cluster Coëfficiënt vanuit een random graaf. 4.004.000 punten, met een toegestane afwijking van 0,001.

3.2.2 Cluster Coëfficiënt vanuit JEB graaf



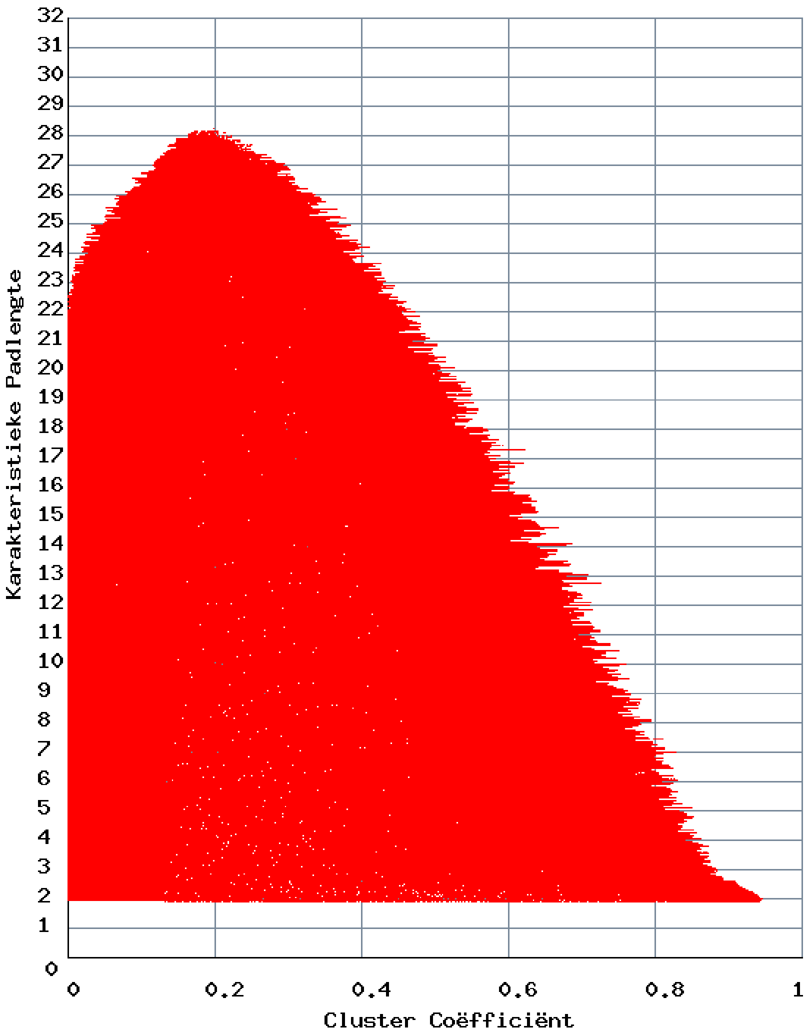
Figuur 18: Statespace doorzocht op Cluster Coëfficiënt vanuit de JEB-graaf. 4.004.000 punten, met een toegestane afwijking van 0,01.

3.2.3 Cluster Coëfficiënt vanuit Staart graaf



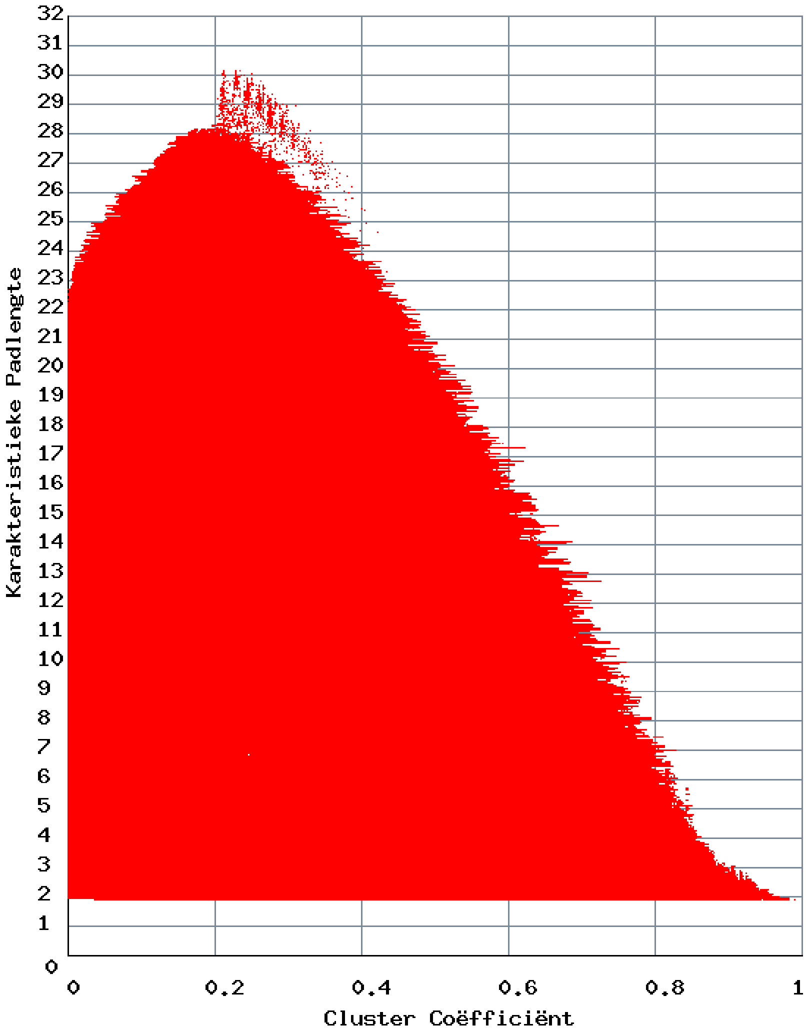
Figuur 19: Statespace doorzocht op Cluster Coëfficiënt vanuit de Staart-graaf. 4.004.000 punten, met een toegestane afwijking van 0,01.

3.2.4 Karakteristieke Padlengte vanuit random graaf



Figuur 20: Statespace doorzocht op Karakteristieke Padlengte vanuit een random graaf. 11.312.000 punten, met een toegestane afwijking van 0,01.

3.2.5 Combinatie van statespaces



Figuur 21: Combinatie van alle vier de voorgaande Statespaces. 23.324.000 punten.

4 Conclusie

Voor dit onderzoek heerste de opvatting dat er een strikte relatie aanwezig is tussen de cluster coëfficiënt en de karakteristieke padlengte. Een hoge cluster coëfficiënt betekent een hoge karakteristieke padlengte, en omgekeerd: een lage cluster coëfficiënt betekent een lage karakteristieke padlengte.

Deze opvatting is nu op twee manieren weerlegd, namelijk door de statespaces zelf, en door de JEB- en Staart-grafen.

De statespaces laten een ruime verdeling van grafen met 100 knopen en 300 verbindingen zien, waarbij de cluster coëfficiënt onafhankelijk is van de karakteristieke padlengte: er zijn grafen met een hoge cluster coëfficiënt die een lage karakteristieke padlengte hebben, grafen met een lage cluster coëfficiënt die een hoge karakteristieke padlengte hebben, en vele grafen daartussen in. Het enige soort grafen dat niet wordt gevonden is met een hoge cluster coëfficiënt en een hoge karakteristieke padlengte.

De JEB- en de Staart-graaf zijn een uitzondering: deze hebben wel een strikte relatie tussen de cluster coëfficiënt, en de karakteristieke padlengte. Om de maximale cluster coëfficiënt te bereiken, moet de padlengte een vaste waarde hebben: minimaal. Om een maximale karakteristieke padlengte te bereiken, moet de cluster coëfficiënt een vaste waarde hebben, namelijk het aantal knopen in de sub-graaf.

Referenties

[1] William S. Lovejoy, Christoph H. Loch, *Minimal and maximal characteristic path lenghts in connected sociamatrices.* www.sciencedirect.com, 2003.