

Vrij Universiteit Amsterdam

# Gerichte immunisatiestrategie met onnauwkeurige informatie in schaalvrije netwerken

Hoe kunnen we een nieuwe griepedemie voorkomen?

Michael Lie



BWI Werkstuk

2010



## Voorwoord

Dit BWI werkstuk is een onderdeel van de Master Business Mathematics and Informatics, als een van de afsluitende onderdelen van deze opleiding. Business Mathematics and Informatics is een multidisciplinaire opleiding met als doel bedrijfsprocessen te optimaliseren met behulp van de disciplines bedrijfseconomie, wiskunde en informatica. In dit werkstuk zal ik onderzoek doen naar een nieuwe immunisatiestrategie binnen schaalvrije netwerken, met behulp van deze drie disciplines.

Toen ik met mijn werkstuk wilde beginnen brak er net de het Influenza A (H1N1) virus uit in Nederland. Vanwege mijn interesse in small-world netwerken, leek het me leuk om hier onderzoek naar te doen voor mijn BWI werkstuk. Daarom heb ik besloten om onderzoek te doen naar een nieuwe immunisatiestrategie binnen schaalvrije netwerken.

Ik wil graag mijn begeleider dr. Sandjai Bhulai bedanken voor zijn ondersteuning bij het maken van dit BWI werkstuk.



## Samenvatting

Het uitbreken van een griepepidemie is erg slecht voor de economie van een land. Daarom zijn er diverse immunisatiestrategieën om zo een uitbraak te voorkomen. Van deze strategieën is een gerichte immunisatiestrategie op dit moment het meest effectief. Het nadeel van deze strategie is dat je volledige informatie nodig hebt over de topologie van een netwerk om deze strategie uit te voeren.

In dit werkstuk wordt een onnauwkeurig gerichte immunisatiestrategie onderzocht. Met behulp van wiskundige analyses en simulaties wordt er onderzocht wat voor invloed de nauwkeurigheid heeft op de effectiviteit van deze strategie.

Een hogere nauwkeurigheid van deze strategie resulteert in een hogere epidemische drempelwaarde. Deze epidemische waarde stijgt steeds sneller, naarmate de nauwkeurigheid hoger is. De invloed van de nauwkeurigheid op het aantal mensen is klein wanneer de verspreidingskans onder de epidemische drempelwaarde ligt. Wanneer de verspreidingskans boven de epidemische drempelwaarde ligt is deze invloed wel groot en zal de besmettingskans steeds sneller dalen wanneer de nauwkeurigheid stijgt.

Een hoge nauwkeurigheid is dus erg belangrijk voor een onnauwkeurig gerichte immunisatiestrategie, omdat de effectiviteit van de strategie steeds sneller afneemt naarmate de nauwkeurigheid daalt.



## Inhoudsopgave

Voorwoord .....	1
Samenvatting .....	3
1 Inleiding.....	7
2 Probleemstelling.....	9
2.1 Complexe netwerken .....	9
2.2 Epidemie verspreiding.....	9
2.3 Immunisatie .....	10
2.3.1 Willekeurige immunisatie .....	10
2.3.2 Gerichte immunisatie .....	10
2.3.3 Acquaintance immunisatie .....	10
2.3.4 Random walk immunisatie.....	10
2.3.5 Imperfect gerichte immunisatie.....	11
2.3.6 Onnauwkeurig gerichte immunisatie .....	11
2.4 Probleem .....	11
3 Wiskundige modellen .....	13
3.1 Barabási-Albert Model .....	13
3.2 SIR Model.....	13
3.3 Immunisatie .....	14
4 Wiskundige analyse .....	15
4.1 Epidemische drempelwaarde .....	15
4.2 Gemiddeld aantal geïnfecteerden .....	19
5 Simulatie .....	23
5.1 Invloed van de nauwkeurigheid.....	23
5.2 Verschillende strategieën vergeleken .....	24
6 Conclusie .....	27
7 Discussie.....	29
8 Literatuurlijst .....	31



## 1 Inleiding

Het afgelopen jaar brak het Influenza A (H1N1) virus uit in Nederland. Dit virus is makkelijk overdraagbaar op andere mensen en zorgde voor een epidemie op wereldwijde schaal. Een uitbraak van zo een virus is erg slecht voor de economie van het land. Er zijn veel strategieën om een uitbraak tegen te gaan. Een van deze strategieën is massavaccinatie, waarbij iedereen wordt gevaccineerd. Iedereen vaccineren kost erg veel geld, daarom is het efficiënter om slimmere strategieën te gebruiken die een uitbraak voorkomen, waarbij het niet nodig is om iedereen in te enten.

Pastor-Satorras en Vespignani hebben onderzoek gedaan naar immunisatiestrategieën in schaalvrije netwerken [1]. Schaalvrije netwerken zijn netwerken waarbij de verdeling van de graden een machtsfunctie volgt, ongeacht de grootte van het netwerk. Netwerken die deze eigenschap hebben zijn onder andere computernetwerken, biologische netwerken en sociale netwerken. Uit het onderzoek bleek dat in een oneindig groot netwerk met een willekeurige immunisatiestrategie het noodzakelijk is alle knopen te immuniseren om een epidemie te voorkomen. Ze stelden een strategie voor die eerst de knopen immuniseert die de hoogste graad hebben. Het aantal knopen dat geïmmuniseerd moest worden om een epidemie te voorkomen daalde significant.

Een strategie die zich volledig richt op de knopen met de hoogste graad, is in de praktijk vaak onmogelijk aangezien deze strategie volledige kennis vereist van de netwerktopologie. Wel is het in sommige gevallen mogelijk te schatten welke knopen de hoogste graad hebben, maar hier zullen dan altijd knopen bij zitten met een lagere graad. Hier kun je denken aan computernetwerken waarbij niet alle verbindingen bekend zijn.

Een schaalvrij netwerk kenmerkt zich door een relatief klein aantal knopen met een zeer hoge graad. Deze knopen met een hoge graad worden hubs genoemd. De robuustheid van zo een netwerk hangt sterk af van deze hubs, omdat deze veel van de knopen met elkaar verbinden. Hoe effectief is een gerichte strategie nog, wanneer er hubs worden gemist of wanneer knopen onterecht worden gezien als hubs?

In dit werkstuk gaan we onderzoeken hoe effectief een gerichte strategie is waarbij men niet beschikt over volledige kennis van de netwerktopologie. Er wordt bij deze strategie een schatting gemaakt van hoe nauwkeurig de beschikbare informatie is.

Het werkstuk is als volgt opgebouwd: In hoofdstuk 2 wordt uitgelegd wat precies de probleemstelling is. Vervolgens worden in hoofdstuk 3 de modellen beschreven die gebruikt worden tijdens het onderzoek. In hoofdstuk 4 zal het probleem wiskundig geanalyseerd worden. Hoofdstuk 5 bevat de resultaten van de simulatie. Vervolgens worden in hoofdstuk 6 de resultaten van de wiskundige analyse en de simulatie besproken en hoofdstuk 7 wordt afgesloten met een discussie.



## 2 Probleemstelling

### 2.1 Complexe netwerken

Veel netwerken die je tegenkomt in het echte leven zijn een netwerk met een zeer complexe structuur. Je komt ze bijvoorbeeld tegen in technologische netwerken, biologische netwerken, transport netwerken en sociale netwerken. De eigenschappen van deze netwerken zijn vaak complex, bijvoorbeeld dat de gradenverdeling van een netwerk een machtsfunctie volgt. Dit soort eigenschappen komen niet voor in simpele netwerken en daarom worden deze netwerken complexe netwerken genoemd.

Binnen de complexe netwerken wordt er veel onderzoek gedaan naar small-world netwerken. Een small-world netwerk is een netwerk waarbij de meeste knopen geen burens van elkaar zijn, maar het kortste pad tussen twee knopen altijd klein is. Een small-world netwerk kenmerkt zich door een hoge clusteringscoëfficiënt en een lage diameter. De clusteringscoëfficiënt geeft aan in welke mate de knopen trios vormen en de diameter is het langste pad van alle kortste paden tussen twee knopen.

In 1998 is het eerste wiskundig model gemaakt voor een small-world netwerk [2]. Dit model was gebaseerd op een willekeurig gegenereerd netwerk, en voegde dan een aantal extra verbindingen tussen de knopen toe. Een beperking van dit model is dat het een homogene gradenverdeling heeft, terwijl veel netwerken in het echt, bijvoorbeeld het internet, een gradenverdeling hebben die een machtsfunctie volgen.

Schaalvrije netwerken hebben deze eigenschappen wel, naast de eigenschappen van een small-world netwerk. Hierdoor heeft een schaalvrij netwerk een groot aantal knopen met een lage graad en een klein aantal knopen met een zeer hoge graad. De knopen met een zeer hoge graad worden hubs genoemd en zijn kenmerkend voor een schaalvrij netwerk.

### 2.2 Epidemie verspreiding

Er is veel onderzoek gedaan naar hoe een virus zich verspreidt binnen een schaalvrij netwerk. Uit onderzoek blijkt dat de snelheid waarop een virus zich verspreid in een schaalvrije netwerk niet hoog hoeft te zijn om te zorgen voor een epidemie [3].

In de samenleving spreekt men van een epidemie wanneer er een ziekte zich verspreidt onder de mensen, op een onverwacht hoge snelheid. Dit is verschillend voor elke ziekte. In Nederland spreekt men bijvoorbeeld van een griepedemie als er meer dan vijftig mensen op de honderdduizend mensen de griep hebben. Een ziekte als hondsdolheid zal bij een klein aantal gevallen al kunnen worden aangeduid als een epidemie, terwijl dat niet gebeurt als datzelfde aantal mensen verkouden is.

In dit werkstuk wordt met een epidemie bedoeld wanneer een virus zich met hoge snelheid verspreidt binnen een netwerk en het niet beperkt blijft tot een klein aantal gevallen. Het wel of niet uitbreken van zo een virus hangt af van verschillende factoren zoals de snelheid waarop het virus zich verspreidt, hoe lang het duurt voordat een persoon genezen is en ook het aantal mensen dat geïmmuniseerd is.

Dit hangt samen met de epidemische drempelwaarde. Wanneer het virus onder de epidemische drempelwaarde zit, zal het virus snel worden uitgeroeid, terwijl als het virus boven de epidemische drempelwaarde zit, er een epidemie zal uitbreken.

Afhankelijk van de ziekte kan het virus zich op verschillende manieren verspreiden en ook hebben ze verschillende effecten. Ziektes kunnen bijvoorbeeld genetisch overdraagbaar en niet meer te genezen zijn. Veel ziektes gedragen zich volgens het SIR-model. Dit model gaat er vanuit dat zieken de ziekte kunnen overdragen, dat alle zieken weer genezen en vervolgens immuun zijn voor deze ziekte.

## 2.3 Immunisatie

Er zijn verschillende onderzoeken gedaan naar het voorkomen van een uitbraak van een epidemie door een gedeelte van de populatie te immuniseren. Hiervoor zijn diverse strategieën ontwikkeld. Een aantal bekende immunisatiestrategieën zijn beschreven in de volgende paragrafen.

### 2.3.1 Willekeurige immunisatie

Bij een willekeurige immunisatiestrategie worden de individuen geheel willekeurig geïmmuniseerd. In een schaalvrije netwerk is deze strategie niet heel effectief, omdat er een groot aantal individuen met weinig contacten zijn en slechts een klein aantal individuen met veel contacten. Juist deze laatste groep heeft veel invloed op het wel of niet uitbreken van een epidemie, omdat deze zoveel individuen met elkaar verbindt. Met een willekeurige immunisatiestrategie is de kans klein dat deze groep wordt geïmmuniseerd. Uit onderzoek bleek dan ook dat in een oneindig groot netwerk met deze strategie de hele populatie geïmmuniseerd moet worden om een epidemie te voorkomen [1].

### 2.3.2 Gerichte immunisatie

Met een gerichte immunisatiestrategie worden de individuen met het hoogst aantal contacten geïmmuniseerd. Door de gradenverdeling van een schaalvrije netwerk is deze strategie juist zeer effectief, waardoor het mogelijk is om een uitbraak van een epidemie te voorkomen door een klein deel van de populatie te immuniseren [1]. De gerichte strategie is op dit moment de beste strategie om een uitbraak van de epidemie te voorkomen. Het nadeel van deze strategie is dat volledige kennis van de netwerktopologie is vereist om deze strategie uit te voeren, wat vaak in de praktijk niet bekend is.

### 2.3.3 Acquaintance immunisatie

Bij acquaintance immunisatie wordt alleen gebruik gemaakt van lokale informatie. Bij deze strategie worden eerst willekeurig individuen geselecteerd en van die individuen worden willekeurig contacten geselecteerd die vervolgens geïmmuniseerd worden. De gedachte achter deze strategie is dat de kans dat een individu met veel contacten geselecteerd wordt groter is dan een individu met weinig contacten. Deze strategie blijkt effectiever te zijn dan de willekeurige immunisatiestrategie [7].

### 2.3.4 Random walk immunisatie

Ook bij de random walk immunisatie wordt alleen gebruik gemaakt van lokale informatie. Er wordt bij deze strategie eerst willekeurig een individu gekozen. Dan wordt er 'gelopen' naar een van zijn contacten. Wanneer dit individu nog niet is geïmmuniseerd, wordt deze geïmmuniseerd en wordt er verder naar een willekeurig contact van dit individu gelopen. Mocht het contact al geïmmuniseerd zijn,

wordt er teruggelopen. Dit wordt herhaald totdat de gewenste fractie van de populatie is geïmmuniseerd. Ook deze strategie is effectiever dan de willekeurige immunisatiestrategie [8].

### **2.3.5 Imperfect gerichte immunisatie**

Er is onderzoek gedaan naar hoe effectief een gerichte immunisatiestrategie is wanneer er hubs worden gemist door gebrek aan informatie. Ondanks het gebrek aan deze kennis bleek de strategie het aantal besmette individuen significant te verminderen zelfs wanneer een klein gedeelte van de populatie geïmmuniseerd is [9].

### **2.3.6 Onnauwkeurig gerichte immunisatie**

In dit werkstuk gaan we de effectiviteit van een gerichte immunisatiestrategie onderzoeken, wanneer de informatie die er is over de netwerktopologie niet geheel nauwkeurig. Hierbij zijn er dus knopen onterecht als hubs geclassificeerd, en worden er hubs gemist.

## **2.4 Probleem**

Wanneer de gerichte immunisatiestrategie wordt toegepast is volledige kennis van de netwerktopologie vereist. In de praktijk is het echter niet altijd mogelijk om dit precies te weten. Bij het toepassen van deze strategie zal de groep die wordt ingeënt altijd wel mensen bevatten die niet zoveel in aanraking komen met andere mensen en worden er mensen niet ingeënt die juist wel met veel mensen in aanraking komen.

Hoe effectief is de gerichte immunisatiestrategie nog in een schaalvrije netwerk als je er vanuit gaat dat informatie over de netwerktopologie niet geheel nauwkeurig is? De effectiviteit van zo een strategie gaan we onderzoeken in dit werkstuk met de volgende probleemstellingen:

- Wat is de invloed van de nauwkeurigheid van een gerichte immunisatiestrategie op de epidemische drempelwaarde?
- Wat is de invloed van de nauwkeurigheid van een gerichte immunisatiestrategie op het aantal mensen dat besmet raakt?



### 3 Wiskundige modellen

#### 3.1 Barabási-Albert Model

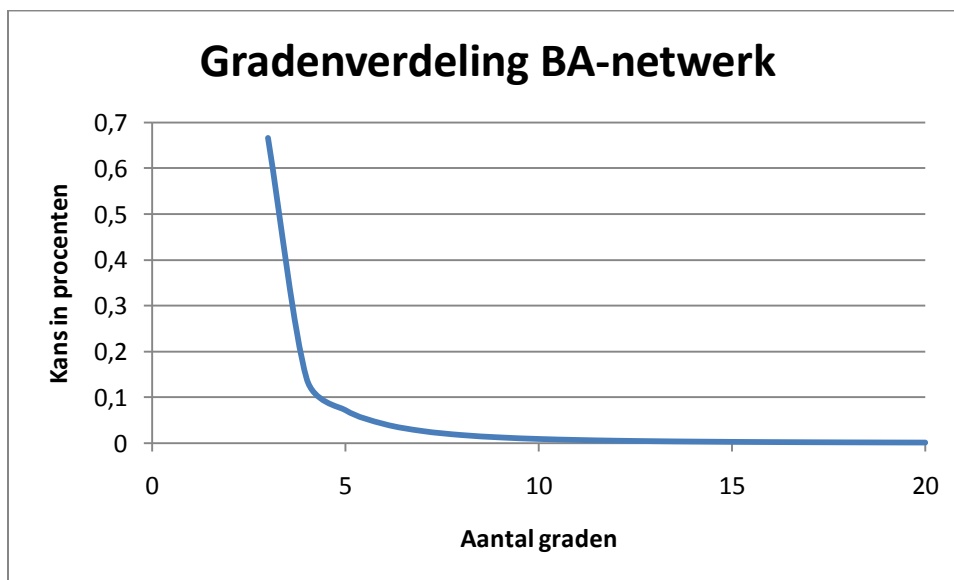
Het netwerk dat gebruikt wordt voor dit onderzoek, zal gebaseerd zijn op het Barabási-Albert model. Dit model genereert een schaalvrije netwerk met behulp van 'preferential attachment'. Dit betekent dat de knopen met de meeste verbindingen, de grootste kans hebben om een verbinding te maken met een nieuwe knoop.

Het netwerk begint met  $m$  knopen, waarbij  $m$  groter dan of gelijk aan 2 moet zijn. Elke keer wordt er een knoop aan het netwerk toegevoegd met  $m$  verbindingen. Deze verbindingen worden gemaakt met de knopen met een kans die afhankelijk is van het aantal verbindingen dat elke knoop heeft. De kans  $p_i$  dat knoop  $i$  een verbinding krijgt met de nieuwe knoop is als volgt,

$$p_i = \frac{k_i}{\sum_j k_j}, \quad (1)$$

waarbij  $k_i$  de graad is van een knoop.

Het netwerk heeft de gradenverdeling  $P(k) = 2m^2/k^3$ , zoals te zien in afbeelding 1. Verreweg de meeste knopen hebben een klein aantal verbindingen, en slechts een klein gedeelte van de knopen hebben veel verbindingen.



Afbeelding 1: De gradenverdeling van een Barabási-Albert netwerk, met  $m=3$ .

De knopen in dit netwerk staan symbool voor individuen, de verbindingen die een knoop heeft, staan symbool voor de contacten die een individu heeft.

#### 3.2 SIR Model

Het SIR model is een ziektecompartimentenmodel dat bestaat uit drie compartimenten.

- Susceptible: In deze fase is een individu vatbaar voor de ziekte,
- Infectious: In deze fase is een individu geïnfecteerd met de ziekte en is besmettelijk voor anderen,
- Removed: In deze fase is een individu genezen van de ziekte. Het individu is niet vatbaar meer voor de ziekte en kan deze ook niet meer overdragen.

Dit is een eenvoudig model om een ziekte zoals de griep te analyseren, waarbij de individuen die genezen, immuun zijn geworden tegen de ziekte.

Elke tijdseenheid zal een individu die vatbaar is voor een ziekte, besmet worden met de ziekte met verspreidingskans  $\lambda$ , mits het individu tenminste één verbinding heeft met een geïnfecteerde individu. Een geïnfecteerde individu, zal met genezingskans  $\delta$  weer genezen. Zonder verlies van algemeenheid stellen we in dit werkstuk  $\delta = 1$  [4].

In dit model wordt aangenomen dat de populatie constant blijft, dus dat er geen geboorte en sterfte plaatsvindt.

### 3.3 Immunisatie

De immunisatiestrategie die wordt toegepast is een onnauwkeurig gerichte immunisatiestrategie. Deze strategie betreft een combinatie van een gerichte en een willekeurige immunisatiestrategie, welke afhankelijk is van nauwkeurigheid  $c$ . Een fractie  $a$  van de populatie zal worden geïmmuniseerd.

Bij een gerichte immunisatiestrategie zouden alle individuen die meer dan  $k_\alpha$  contacten hebben, worden geïmmuniseerd. Hierbij is  $k_\alpha$  de grens van het aantal contacten dat een individu heeft, dat niet behoort tot de bovenste fractie  $a$  individuen betreffende het aantal contacten.

Bij deze onnauwkeurig gerichte immunisatiestrategie zal van de individuen die meer dan  $k_\alpha$  contacten hebben tenminste met nauwkeurigheid  $c$  worden geïmmuniseerd. De overgebleven fractie  $(1 - c) \cdot a$  van de populatie wordt dan willekeurig geïmmuniseerd. Dit resulteert in de volgende kansen om geïmmuniseerd te worden:

$$k \leq k_\alpha: (1 - c) \cdot a, \quad (2)$$

$$k > k_\alpha: c + (1 - c) \cdot a. \quad (3)$$

Een  $c$  van 1 resulteert in een normale gerichte immunisatiestrategie, en een  $c$  van 0 resulteert in een willekeurige immunisatiestrategie.

## 4 Wiskundige analyse

### 4.1 Epidemische drempelwaarde

Individueen in het model vallen onder één van de volgende categorieën: susceptible, infected, removed en immunized. De dichtheden op tijdstip  $t$  van deze categorieën worden respectievelijk aangeduid met  $S_k(t)$ ,  $\rho_k(t)$ ,  $R_k(t)$  en  $I_k(t)$ . De som van deze dichtheden is dus 1.

Deze dichtheden voldoen aan de volgende differentiaal vergelijkingen [4]:

$$\frac{dS_k(t)}{dt} = -\lambda k S_k(t) \theta(t), \quad (4)$$

$$\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \lambda k S_k(t) \theta(t), \quad (5)$$

$$\frac{dR_k(t)}{dt} = \rho_k(t), \quad (6)$$

$$\frac{dI_k(t)}{dt} = 0. \quad (7)$$

Hierbij is  $\theta(t)$  de kans dat een willekeurige link verbinding maakt met een geïnfecteerde individu.  $kS_k(t)\theta(t)$  geeft dan de kans dat een vatbaar individu met  $k$  contacten contact heeft met een geïnfecteerde individu. Wanneer je deze term vermenigvuldigt met de overdragskans  $\lambda$ , krijg je de kans dat een vatbaar individu met  $k$  contacten besmet raakt. Hieruit volgt vergelijking (4).

Omdat de genezingskans  $\delta = 1$ , volgt vergelijking (6). Het verschil in aantal geïnfecteerde individuen is het verschil in het aantal individuen dat vatbaar is voor de ziekte minus het verschil in aantal individuen dat genezen is. Het aantal geïmmuniseerden blijft altijd gelijk.

$\theta(t)$  is als volgt te berekenen:

$$\theta(t) = \frac{\sum_k k P(k) \rho_k(t)}{\sum_s s P(s)} = \frac{\sum_k k P(k) \rho_k(t)}{\langle k \rangle}. \quad (8)$$

$P(k)$  is de kans dat een individu  $k$  contacten heeft en  $\langle k \rangle$  is het gemiddeld aantal contacten per individu.  $\sum_k k P(k)$  geeft dus het gemiddeld aantal contacten per individu aan, en wanneer deze term vermenigvuldigd wordt met  $\rho_k(t)$  krijg je het gemiddeld aantal geïnfecteerde contacten per individu. Het gemiddeld aantal geïnfecteerde contacten per individu gedeeld door het gemiddeld aantal contacten per individu geeft vervolgens de kans dat een willekeurige link verbinding maakt met een geïnfecteerde individu.

De start condities zijn als volgt:

$$\begin{aligned}
 k \leq k_\alpha: \rho_k(0) = 0, \quad R_k(0) = 0, \quad S_k(0) = 1 - (1 - c)a, \quad I_k(0) = (1 - c)a, \\
 k > k_\alpha: \rho_k(0) = 0, \quad R_k(0) = 0, \quad S_k(0) = 1 - (c + (1 - c)a), \quad I_k(0) = c + (1 - c)a.
 \end{aligned}$$

We nemen hierbij aan dat in het begin alleen een hele kleine fractie van de populatie geïnfecteerd is. De fractie individuen dat genezen is, is aan het begin nul, de fractie vatbare individuen is  $1 -$  de fractie geïmmuniseerden. De populatie wordt geïmmuniseerd volgens vergelijkingen (2) en (3).

Met deze start condities, kunnen we door vergelijking (4) te integreren, de dichtheidsfunctie  $S_k(t)$  berekenen.

$$S_k(t) = \begin{cases} e^{-\lambda k \int_0^t \theta(t) dt + \ln(1 - (1 - c)a)}, & \text{als } k \leq k_\alpha, \\ e^{-\lambda k \int_0^t \theta(t) dt + \ln(1 - (c + (1 - c)a))}, & \text{als } k > k_\alpha. \end{cases} \quad (9)$$

We bekijken de integraal in deze functie apart en noemen deze functie  $\phi(t)$ .

$$\begin{aligned}
 \phi(t) &= \int_0^t \theta(t) dt = \int_0^t \frac{\sum_k kP(k) \rho_k(t)}{\langle k \rangle} dt \\
 &= \frac{\sum_k kP(k) \int_0^t \rho_k(t) dt}{\langle k \rangle} = \frac{1}{\langle k \rangle} \sum_k kP(k) R_k(t).
 \end{aligned} \quad (10)$$

We zien dat functie  $\phi(t)$  de kans geeft dat een willekeurige link, verbinding maakt met een verwijderde contact. Door de differentiaal te nemen van deze functie, krijgen we de volgende vergelijking:

$$\begin{aligned}
 \frac{d\phi(t)}{dt} &= \frac{1}{\langle k \rangle} \sum_k kP(k) \rho_k(t) \\
 &= \frac{1}{\langle k \rangle} \sum_k kP(k) [1 - R_k(t) - I_k(t) - S_k(t)] \\
 &= 1 - \phi(t) - \frac{1}{\langle k \rangle} \sum_k kP(k) I_k(t) - \frac{1}{\langle k \rangle} \sum_k kP(k) S_k(t).
 \end{aligned} \quad (11)$$

Wanneer we de limiet  $t \rightarrow \infty$  nemen, zal er een evenwicht ontstaan waarbij er geen geïnfecteerden meer zijn. De overige compartimenten zullen dan constant blijven, dus  $\lim_{t \rightarrow \infty} \frac{d\phi(t)}{dt} = 0$ . Dan volgt de volgende vergelijking:

$$\begin{aligned}
\phi(\infty) &= 1 - \frac{1}{\langle k \rangle} \sum_k kP(k)I_k(\infty) - \frac{1}{\langle k \rangle} \sum_k kP(k)S_k(\infty) \\
&= 1 - \frac{1}{\langle k \rangle} \sum_{k_l}^{k_\alpha} kP(k)(1-c)a - \frac{1}{\langle k \rangle} \sum_{k_{\alpha+1}}^{k_u} kP(k) \cdot \\
&\quad (c + (1-c)a) \\
&\quad - \frac{1}{\langle k \rangle} \sum_{k_l}^{k_\alpha} kP(k)e^{-\lambda k \phi(\infty) + \ln(1-(1-c)a)} \\
&\quad - \frac{1}{\langle k \rangle} \sum_{k_{\alpha+1}}^{k_u} kP(k)e^{-\lambda k \phi(\infty) + \ln(1-(c+(1-c)a))} \\
&= F(\phi(\infty)),
\end{aligned} \tag{12}$$

waar  $k_l$  en  $k_u$  de ondergrens en bovengrens zijn van het aantal contacten dat een individu kan hebben en  $k_{\alpha+1}$  het aantal contacten is van de individu die net boven de grens  $k_\alpha$  zit.

We stellen de rechterkant van de vergelijking (11) gelijk aan  $F(\phi(\infty))$ . We willen nu de vergelijking  $F(\phi(\infty)) = \phi(\infty)$  oplossen.  $\phi(\infty) = 0$  is een oplossing, in het geval dat er geen epidemie is uitgebroken. In het interval  $[0,1]$  moet er nog een oplossing zijn die niet nul is, in het geval dat er wel een epidemie uitbreekt, met de voorwaarde dat  $\phi(\infty) = 0$  een oplossing is. De differentiaal van  $F(\phi(\infty))$  moet dan groter of gelijk aan één zijn, anders zal deze functie nooit de lijn  $\phi(\infty)$  snijden [4].

$$\left. \frac{dF(\phi(\infty))}{d\phi(\infty)} \right|_{\phi(\infty)=0} \geq 1, \tag{13}$$

$$\begin{aligned}
\frac{dF(\phi(\infty))}{d\phi(\infty)} &= \frac{1}{\langle k \rangle} \left[ (1 - (1-c)a) \sum_{k_l}^{k_u} k^2 P(k) \lambda \right. \\
&\quad \left. - c \sum_{k_{\alpha+1}}^{k_u} k^2 P(k) \lambda \right].
\end{aligned} \tag{14}$$

In het geval dat vergelijking (13) gelijk aan één is, kunnen we de epidemische drempelwaarde  $\lambda_\alpha$  definiëren als:

$$\lambda_\alpha^{-1} = \frac{1}{\langle k \rangle} \left[ (1 - (1 - c)a) \sum_{k_l}^{k_u} k^2 P(k) - c \sum_{k_{\alpha+1}}^{k_u} k^2 P(k) \right]. \quad (15)$$

Voor ons netwerkmodel geldt:  $P(k) = 2m^2/k^3$ ,  $\langle k \rangle = 2m$ ,  $k_l = m$ ,  $k_u = mN^{1/2}$  [5,6]. Door dit in te vullen, en de continuüm theorie toe te passen, is vergelijking (14) om te schrijven in:

$$\lambda_\alpha^{-1} = (1 - (1 - c)a) \frac{1}{2} m \ln(N) - c m \ln \left( \frac{mN^{1/2}}{k_{\alpha+1}} \right). \quad (16)$$

De enige onbekende term in deze vergelijking is  $k_{\alpha+1}$ , welke is te berekenen door de volgende vergelijking op te lossen:

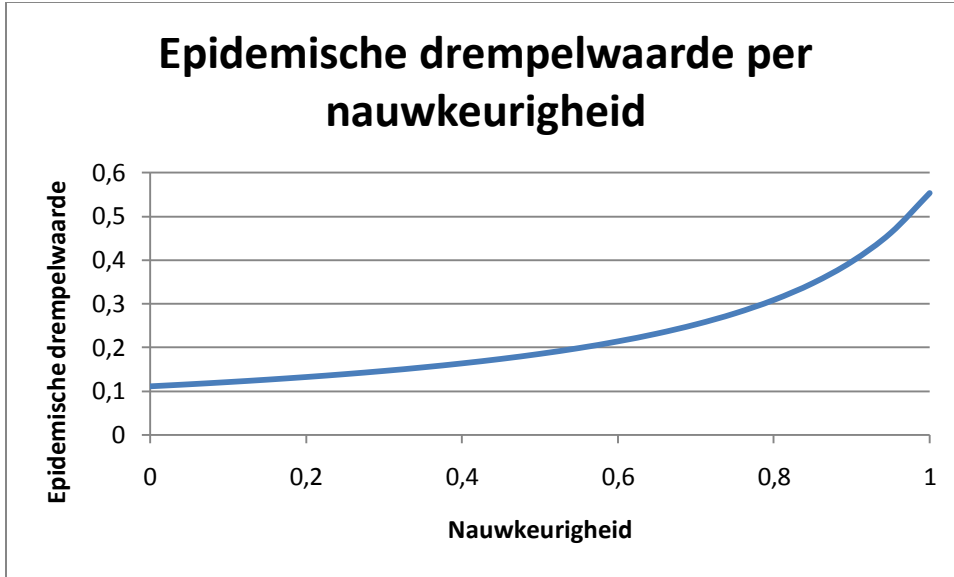
$$\frac{\sum_{k_l}^{k_{\alpha+1}} P(k)}{\sum_{k_l}^{k_u} P(k)} = 1 - a. \quad (17)$$

Met behulp van de continuüm theorie kunnen we hieruit afleiden dat

$$k_{\alpha+1} = \sqrt{\frac{m^2}{\frac{(1-a)}{N} + a}}. \quad (18)$$

We kunnen nu met vergelijking (16) de epidemische drempelwaarde bepalen, afhankelijk van de grootte van het netwerk, de minimale graad, de immunisatiefractie en de nauwkeurigheid. Uit deze vergelijking kunnen we afleiden dat de epidemische drempelwaarde zal stijgen wanneer de immunisatiefractie en de nauwkeurigheid toeneemt, en zal dalen wanneer de netwerkgrootte en de begingraad toeneemt.

Op afbeelding 2 is te zien dat hoe hoger de nauwkeurigheid is, hoe sneller de epidemische drempelwaarde toeneemt wanneer de nauwkeurigheid stijgt.



Afbeelding 2: De invloed van de nauwkeurigheid op de epidemische drempelwaarde.  $N=5000$ ,  $m=3$ ,  $a=0,3$ .

## 4.2 Gemiddeld aantal geïnfekteerden

We gaan nu het gemiddeld aantal geïnfekteerden  $R_\infty$  berekenen, wanneer de infectiesnelheid dicht bij de epidemische drempelwaarde ligt.

$$R_\infty = 1 - \sum_k kP(k)I_k(\infty) - \sum_k kP(k)S_k(\infty). \quad (19)$$

We nemen aan dat het netwerk oneindig groot is, dus  $k_u \rightarrow \infty$ . Door de continuüm theorie toe te passen in vergelijking (19), is  $R_\infty$  ook te schrijven als:

$$\begin{aligned} R_\infty = & 1 - (1 - c)a - \frac{cm^2}{k_{\alpha+1}} \\ & - (1 - (1 - c)a) \int_{k_t}^{\infty} \frac{2m^2}{k^3} e^{-\lambda k \phi(\infty)} dk \\ & - c \int_{k_{\alpha+1}}^{\infty} \frac{2m^2}{k^3} e^{-\lambda k \phi(\infty)} dk. \end{aligned} \quad (20)$$

Neem  $z_1 = \lambda k_t \phi(\infty)$  en  $z_2 = \lambda k_{\alpha+1} \phi(\infty)$ . Wanneer de infectiesnelheid  $\lambda$  dicht bij de epidemische drempelwaarde  $\lambda_\alpha$  ligt in een oneindig groot netwerk, krijgen we  $z_1 \rightarrow 0$  en  $z_2 \rightarrow 0$ . Door middel van substitutie krijgen we de volgende vergelijking:

$$R_\infty = 1 - (1 - c)a - \frac{cm^2}{k_{\alpha+1}} \quad (21)$$

$$\begin{aligned}
& -(1 - (1 - c)a) \frac{2m^2 z_1^2}{k_t^2} \int_{z_1}^{\infty} x^{-3} e^{-\lambda k \phi(\infty)} dx \\
& -c \frac{2m^2 z_2^2}{k_{\alpha+1}^2} \int_{z_2}^{\infty} x^{-3} e^{-x} dx.
\end{aligned}$$

De integralen in de vergelijking (21) hebben nu de vorm van een incomplete gamma functie  $\Gamma(a, z) = \int_z^{\infty} t^{a-1} e^{-t} dt$  met de volgende eigenschappen in het geval dat  $z \rightarrow 0$ :

$$\begin{aligned}
\Gamma(a + 1, z) &= a\Gamma(a, z) + z^a e^{-z}, \\
\Gamma(0, z) &= -r_E - \ln(z) + z + o(z^2).
\end{aligned} \tag{22}$$

Hierbij is  $r_E$  de Euler constante [9]. Door deze eigenschappen toe te passen, en de hogere orde termen te negeren, kunnen we vergelijking (21) omschrijven in:

$$\begin{aligned}
R_{\infty} &= 1 - (1 - c)a \\
&\quad - \frac{cm^2}{k_{\alpha+1}} \\
&\quad - 2m^2 \left[ (1 - (1 - c)a) \frac{1}{k_t^2} \left( \frac{1}{2} - z_1 \right) - c \frac{1}{k_{\alpha+1}^2} \left( \frac{1}{2} - z_2 \right) \right] \\
&= 2m\lambda\phi(\infty) \left( 1 - (1 - c)a - \frac{cm}{k_{\alpha+1}} \right),
\end{aligned} \tag{23}$$

waar  $k_{\alpha+1}$  de volgende waarde heeft in een oneindig groot netwerk, volgens vergelijking (18):

$$k_{\alpha+1} = \sqrt{\frac{m^2}{a}}. \tag{24}$$

$\phi(\infty)$  is nu de enige onbekende term. Als we vergelijking (12) uitwerken met behulp van de continuüm theorie krijgen we:

$$\begin{aligned}
\phi(\infty) &= 1 - (1 - c)a - \frac{cm}{k_{\alpha+1}} \\
&\quad - (1 - (1 - c)a)m \int_{k_t}^{\infty} \frac{1}{k^2} e^{-\lambda k \phi(\infty)} dk \\
&\quad - cm \int_{k_{\alpha+1}}^{\infty} \frac{1}{k^2} e^{-\lambda k \phi(\infty)} dk.
\end{aligned} \tag{25}$$

Door  $z_1 = \lambda k_t \phi(\infty)$  en  $z_2 = \lambda k_{\alpha+1} \phi(\infty)$  te nemen krijgen we door middel van substitutie weer integralen in de vorm van een incomplete gamma functie:

$$\begin{aligned} \phi(\infty) = & 1 - (1 - c)a - \frac{cm}{k_{\alpha+1}} \\ & - (1 - (1 - c)a) \frac{mz_1}{k_t} \int_{z_1}^{\infty} x^{-2} e^{-\lambda k \phi(\infty)} dx \\ & - c \frac{mz_2}{k_{\alpha+1}} \int_{z_2}^{\infty} x^{-2} e^{-x} dx. \end{aligned} \quad (26)$$

Door de eigenschappen uit vergelijking (19) toe te passen, krijgen we voor  $\phi(\infty)$  het volgende:

$$\begin{aligned} \phi(\infty) = & (1 - (1 - c)a)(z_1 - r_E z_1 - \ln(z_1) z_1) \\ & - \frac{cm}{k_{\alpha+1}} (z_2 - r_E z_2 - \ln(z_2) z_2). \end{aligned} \quad (27)$$

Dit is te schrijven als:

$$A\phi(\infty) + B\phi(\infty)\ln\phi(\infty) = 0, \quad (28)$$

$$\begin{aligned} A = & 1 - (1 - (1 - c)a - c)\lambda m + (1 - (1 - c)a - c)\lambda m R_E \\ & + (1 - (1 - c)a)\lambda m \ln(\lambda m) - c\lambda m \ln(\lambda k_{\alpha+1}), \end{aligned} \quad (29)$$

$$B = (1 - (1 - c)a - c)\lambda m, \quad (30)$$

$$\phi(\infty) = e^{-\frac{A}{B}}. \quad (31)$$

We kunnen nu  $\phi(\infty)$  invullen in vergelijking (23) en het gemiddeld aantal geïnfecteerden berekenen in een oneindig groot netwerk, wanneer de verspreidingskans dicht bij de epidemische drempelwaarde ligt. De uitkomst van deze analyse was niet te controleren met simulaties, vanwege de oneindige grootte van het netwerk.

Door verschillende waarden in te voeren in deze vergelijking, is te zien dat het gemiddeld aantal geïnfecteerden daalt wanneer de nauwkeurigheid stijgt. Hoe deze precies daalt is moeilijk te zeggen, omdat deze vergelijking alleen het gemiddeld aantal geïnfecteerden geeft wanneer de verspreidingskans rondom de epidemische drempelwaarde ligt.



## 5 Simulatie

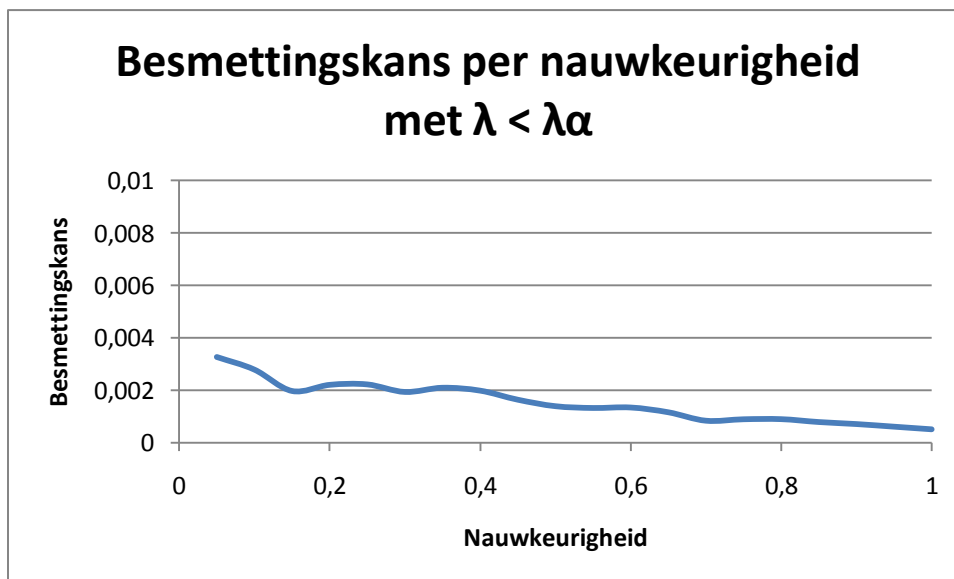
Om te bepalen hoeveel individuen gemiddeld geïnfecteerd raken na de uitbraak van een virus in een netwerk die niet oneindig groot is, maken we gebruik van simulatie. Deze simulatie wordt geprogrammeerd in R en wordt uitgevoerd op een schaalvrij netwerk dat gegenereerd worden volgens het Barabási-Albert model. Na het bouwen van het netwerk, worden volgens de immunisatiestrategie individuen immuun gemaakt. Vervolgens zal willekeurig 0,5% van de bevolking geïnfecteerd worden met een virus, waarna het virus zich verspreidt volgens het SIR-model.

Met deze simulatie onderzoeken we wat de invloeden zijn van de nauwkeurigheid op de kans dat een individu geïnfecteerd wordt door het virus. De resultaten zijn gebaseerd op een gemiddelde van vijftig waarnemingen.

### 5.1 Invloed van de nauwkeurigheid

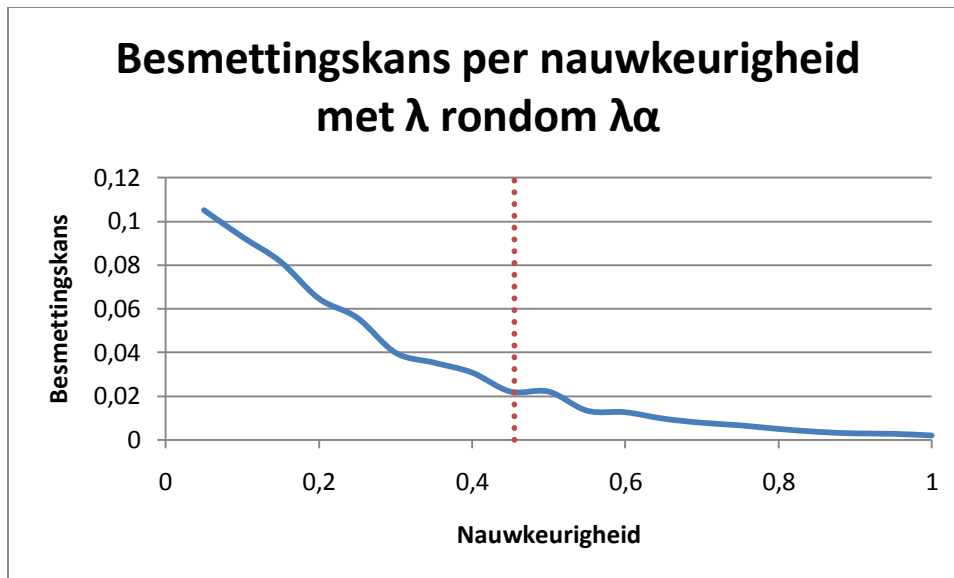
De invloed van de nauwkeurigheid van een gerichte immunisatiestrategie onderzoeken we in drie verschillende situaties. Namelijk wanneer de verspreidingskans onder, rondom en boven de epidemische drempelwaarde ligt. Zoals ook uit de wiskundige analyse blijkt, zorgt een toename van de nauwkeurigheid voor een afname van de besmettingskans.

Wanneer de verspreidingskans onder de epidemische drempelwaarde ligt, blijkt de nauwkeurigheid weinig invloed te hebben op de besmettingkans. Het virus breekt in alle gevallen niet uit en het aantal mensen dat gemiddeld besmet raakt blijft beperkt. Er is een lichte daling van de besmettingskans waar te nemen wanneer de nauwkeurigheid toeneemt (zie afbeelding 3 voor de resultaten).



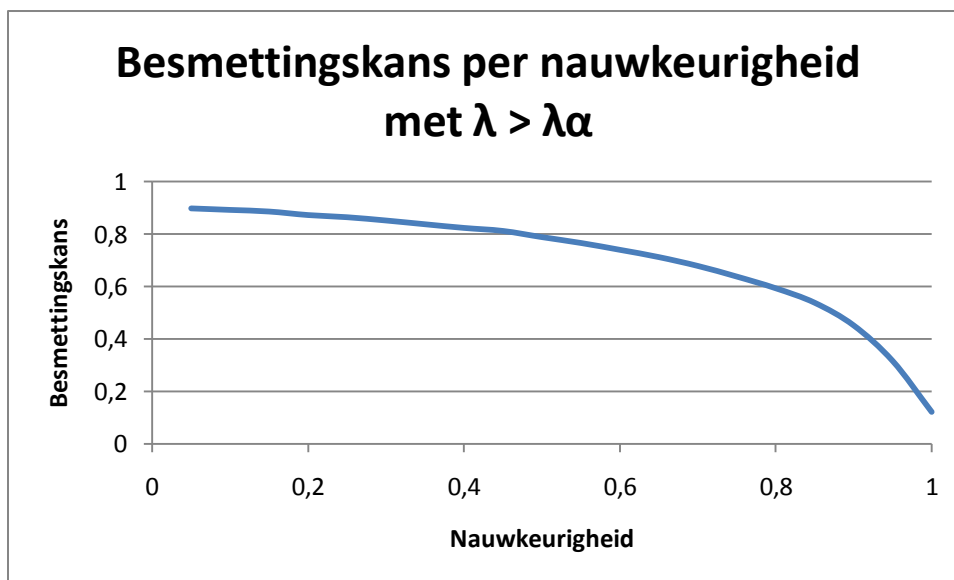
Afbeelding 3: De invloed van de nauwkeurigheid van een gerichte immunisatiestrategie op de verspreidingskans van een individu, wanneer de verspreidingskans onder de epidemische drempelwaarde ligt.  $N=5000$ ,  $m=3$ ,  $a=0.2$ ,  $\lambda=0.05$ .

Als de verspreidingskans rondom de epidemische drempelwaarde ligt, daalt de besmettingskans steeds langzamer. Voor de epidemische drempelwaarde blijkt de besmettingskans een stuk sneller te dalen dan na de epidemische drempelwaarde (zie afbeelding 4 voor de resultaten).



Afbeelding 4: De invloed van de nauwkeurigheid van een gerichte immunisatiestrategie op de besmettingskans van een individu, wanneer de verspreidingskans rond de epidemische drempelwaarde ligt.  $N=5000$ ,  $m=3$ ,  $a=0.15$ ,  $\lambda=0.2$ . De rode stippellijn geeft de epidemische drempelwaarde aan.

Wanneer de verspreidingskans boven de epidemische drempelwaarde ligt, is de snelheid waarmee de besmettingskans daalt in het begin juist het laagst. De snelheid waarmee de besmettingskans daalt blijkt juist toe te nemen wanneer de verspreidingskans boven de epidemische drempelwaarde ligt (zie afbeelding 5 voor de resultaten).

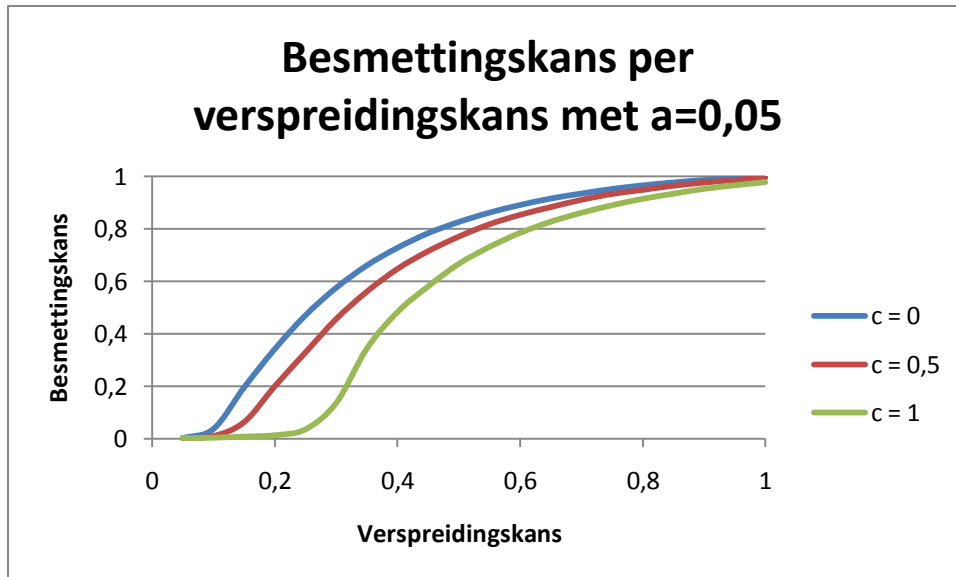


Afbeelding 5: De invloed van de nauwkeurigheid van een gerichte immunisatiestrategie op de besmettingskans van een individu, wanneer de verspreidingskans boven de epidemische drempelwaarde ligt.  $N=5000$ ,  $m=3$ ,  $a=0.2$ ,  $\lambda=0.7$ .

## 5.2 Verschillende strategieën vergeleken

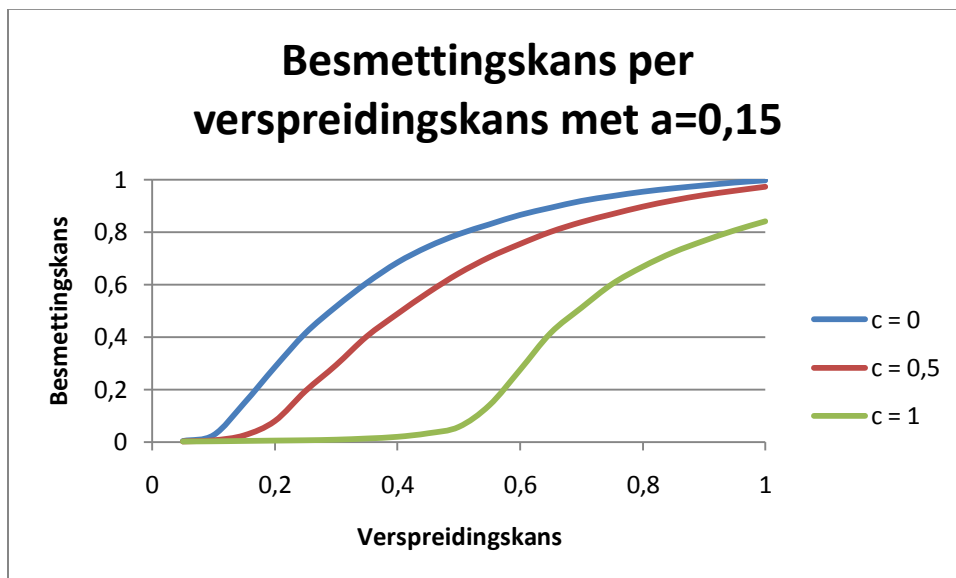
Bij een nauwkeurigheid van 0% is de onnauwkeurig gerichte immunisatiestrategie gelijk aan een willekeurige immunisatiestrategie en bij een nauwkeurigheid van 100% is deze strategie gelijk aan een

gerichte immunisatiestrategie. We gaan nu onderzoeken hoe een onnauwkeurig gerichte immunisatiestrategie presteert, vergeleken met deze twee immunisatiestrategieën.



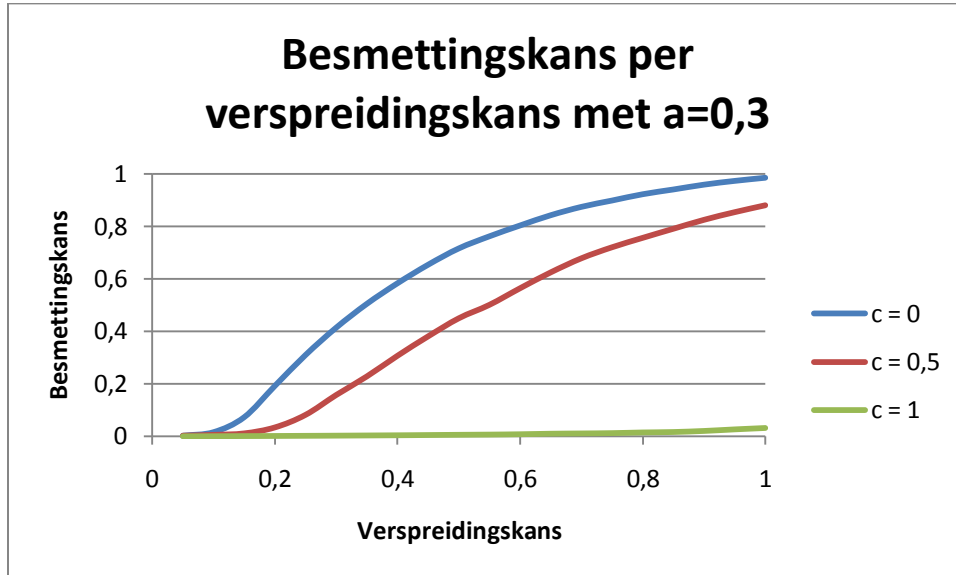
Afbeelding 6: De invloed van de verspreidingskans op de besmettingskans van een individu.  $N=5000$ ,  $m=3$ ,  $a=0,05$ .

Wanneer vijf procent van de populatie wordt geïmmuniseerd, blijven de verschillen relatief klein tussen de drie immunisatiestrategieën. Een uitbraak van epidemie wordt bij de gerichte immunisatiestrategie wel een stuk langer voorkomen. Ook zien we dat het verschil tussen de willekeurige immunisatiestrategie en de onnauwkeurig gerichte immunisatiestrategie kleiner is dan de gerichte immunisatiestrategie en de onnauwkeurig gerichte immunisatiestrategie (zie afbeelding 6 voor de resultaten).



Afbeelding 7: De invloed van de verspreidingskans op de besmettingskans van een individu.  $N=5000$ ,  $m=3$ ,  $a=0,15$ .

Wanneer vijftien procent van de populatie wordt geïmmuniseerd, worden de verschillen tussen de immunisatiestrategieën groter. Waar de willekeurige immunisatiestrategie en de onnauwkeurig gerichte immunisatiestrategie nog relatief weinig verschillen, is het verschil van de gerichte immunisatiestrategie en de andere twee immunisatiestrategieën erg groot geworden (zie afbeelding 7 voor de resultaten).



Afbeelding 8: De invloed van de verspreidingskans op de besmettingskans van een individu.  $N=5000$ ,  $m=3$ ,  $a=0.3$ .

Bij een immunisatiefractie van dertig procent zijn deze verschillen nog meer toegenomen (zie afbeelding 8). De onnauwkeurig gerichte immunisatiestrategie begint zich meer te onderscheiden van de willekeurige immunisatiestrategie, en de gerichte immunisatiestrategie voorkomt zelfs in alle gevallen een uitbraak van een epidemie. Verder zien we nog steeds dat het verschil tussen de willekeurige immunisatiestrategie en de onnauwkeurig gerichte immunisatiestrategie kleiner is dan het verschil tussen de gerichte immunisatiestrategie en de onnauwkeurig gerichte immunisatiestrategie.

## 6 Conclusie

Met behulp van wiskunde analyses and simulaties hebben we in dit werkstuk de volgende twee probleemstellingen onderzocht:

- Wat is de invloed van de nauwkeurigheid van een gerichte immunisatiestrategie op de epidemische drempelwaarde?
- Wat is de invloed van de nauwkeurigheid van een gerichte immunisatiestrategie op het aantal mensen dat besmet raakt?

Om de eerste probleemstelling te onderzoeken hebben we een wiskundige analyse uitgevoerd naar de epidemische drempelwaarde. Uit de analyse blijkt een hogere nauwkeurigheid te leiden tot een hogere drempelwaarde. Verder blijkt de epidemische drempelwaarde steeds sneller te stijgen, wanneer de nauwkeurigheid hoger wordt. De nauwkeurigheid van een gerichte immunisatiestrategie heeft dan ook meer invloed op de epidemische drempelwaarde wanneer de nauwkeurigheid hoger is.

Om te onderzoeken wat de invloed van de nauwkeurigheid van een gerichte immunisatiestrategie is op het aantal mensen dat besmet raakt hebben we simulaties uitgevoerd. Uit deze simulaties blijkt dat de invloed van de nauwkeurigheid verschilt per verspreidingskans.

Wanneer de verspreidingskans onder de epidemische drempelwaarde ligt, is de invloed van de nauwkeurigheid beperkt. Rondom de epidemische drempelwaarde blijkt de besmettingskans steeds langzamer te dalen, terwijl boven de epidemische drempelwaarde de besmettingskans steeds sneller afneemt. De invloed van de nauwkeurigheid is dus erg afhankelijk de epidemische drempelwaarde, waarbij de invloed het sterkst is als de verspreidingskans boven de epidemische drempelwaarde ligt.

Ook de immunisatiefractie blijkt een rol te spelen. Als een klein gedeelte van de populatie wordt geïmmuniseerd zijn de verschillen nog relatief klein tussen respectievelijk 0%, 50% en 100% nauwkeurigheid. Deze verschillen worden echter steeds groter wanneer de fractie van de populatie dat geïmmuniseerd wordt steeds groter wordt.

Het verschil van de besmettingskans tussen een nauwkeurigheid van 0% en 50% is een stuk kleiner dan bij een nauwkeurigheid van 50% en 100%. Ook bij de besmettingskans is de invloed van de nauwkeurigheid groter wanneer de nauwkeurigheid hoger is.

Concluderend is een hoge nauwkeurigheid dus erg belangrijk voor een onnauwkeurige gerichte immunisatiestrategie en blijkt de effectiviteit steeds sneller af te nemen naarmate de nauwkeurigheid daalt.



## 7 Discussie

De resultaten van dit werkstuk kunnen van belang zijn bij het voorkomen van een uitbraak van een epidemie. De onnauwkeurig gerichte immunisatiestrategie zou bijvoorbeeld toegepast kunnen worden om virussen te bestrijden binnen computernetwerken en sociale netwerken.

De uitbraak van seksueel overdraagbare ziektes zou voorkomen kunnen worden door zich te richten op de meest actieve groepen zoals prostituees. Aan de hand van anonieme metingen onder prostituees en niet-prostituees zou de nauwkeurigheid van een gerichte immunisatiestrategie berekend kunnen worden wanneer alleen de prostituees geïmmuniseerd worden.

Met nieuwe technologieën zoals GPS en de OV-chipkaart zou men in de toekomst schattingen kunnen maken hoeveel contact iemand heeft. Zodoende is het dan mogelijk door slechts een gedeelte van de populatie te immuniseren een uitbraak van griep epidemie te voorkomen.

Tijdens het onderzoek was het niet mogelijk om simulaties uit te voeren met groter dan 5000 knopen, vanwege beperkingen in R. Het zou interessant zijn om te onderzoeken of de resultaten hetzelfde blijven wanneer we de grootte van het netwerk een stuk groter wordt.



## 8 Literatuurlijst

- [1] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E 65, 036104 (2002).
- [2] D.J. Watts and S.H. Strogatz, Nature 393 (6684): 409-10 (1998).
- [3] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B26, 521 (2002).
- [4] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E65, 036104 (2002).
- [5] A.-L. Barabási and R. Albert, Science, 286: 509-512 (1999).
- [6] S.N. Dorgovtsev and J.F.F. Mendes, Adv. Phys. 51, 1079 (2002).
- [7] R. Cohen, S. Havlin and D. ben-Avraham, Phys. Rev. Lett. 91, 247901 (2003).
- [8] W. Pei, Z. Chen and Z. Yuan, JCTA 7, 2: 151-156 (2009).
- [9] Y. Wang, G. Xiao, J. Hu, T.H. Cheng and L. Wang, Phys. A 388: 2535-2546 (2009).