Mathematical Statistics
Stockholm University

Epidemics on networks and early stage vaccination

Shaban Mbare

Research Report 2007:12
Licentiate thesis

ISSN 1650-0377

Postal address:

Mathematical Statistics
Dept. of Mathematics
Stockholm University
SE-106 91 Stockholm
Sweden

Internet:

http://www.math.su.se/matstat

Epidemics on networks and early stage vaccination

Shaban Mbare [†], [*]

May 2007

Abstract

This thesis consists of two articles.

1. Mbare S., Andersson M., Svensson Å, Britton T. (2007): Social networks, epidemics and vaccination through contact tracing, (Manuscript).
2. Mbare S., Andersson M., Svensson Å, Britton T. (2007): Network epidemics and early stage vaccination: the effect of latent and infectious periods and their randomness, (Manuscript)

Both articles deal with epidemics on social networks and early stage vaccination, and in particular different vaccination strategies in order to prevent a major outbreak. The social structure is described by a random graph having pre-specified degree distribution, from which friends of an infectious individual are traced and a given vaccination strategy is implemented after delay. We compare the effectiveness of the strategies and investigate the effects of the randomness of the latent period, infectious period and detection (delay) time in the dynamics of the disease. Branching process approximations of the early stage of the epidemic make it possible to compute the probability of outbreaks.

KEY WORDS: contact tracing, degree distribution, detection time, epidemic model, social networks, vaccination strategies.

[*]Present address: Mathematical Statistics, Stockholm University, SE-106 91, Stockholm, Sweden. [†]Permanent address: Mathematics Department, University of Dar es Salaam, Box 35062, Dar es Salaam, Tanzania. *E-mail*: shaban@math.su.se, shabanmbare@yahoo.com

Acknowledgements

I wish to express my gratitude to the International Science Programme (ISP) for its continuous financial support in my studies at Stockholm University, without which this work would not have been possible.

My sincere thanks goes to my supervisor Tom Britton for his tireless encouragement that enabled me to accomplish this thesis. He was always ready to meet and talk about my ideas and to ask me good questions to help me think through the problems (whether philosophical, analytical or computational). A special thank you goes to my co-supervisors Åke Svensson and Mikael Andersson. They taught me how to write academic papers, have confidence in me when I doubted myself and brought out the good ideas in me. It is their encouragement and constant guidance that enabled me to feel at home while working with this thesis.

Let me also say 'thank you' to all members of the Division of Mathematical statistics to have confidence in me. You dedicated your precious time to encourage me at this hard moment as I am away from my family.

Last, but not least, I thank my family: my wife Lela I. Kikoti for patience and encouragement she was giving me whenever I called home. My daughters, Nagma and Hamida for always giving me a call wanting to know when I will join them.

Introduction

The social structure is important in understanding the spread of infectious diseases in the community, and it is often modelled by a random network or graph (see e.g. Newman, 2003). In such a graph, nodes (vertices) represent individuals and the links (edges) correspond to some social relations. It is through these links that the disease can spread. The dynamics of the disease in the network depends on the distribution of the number of friends (degree) of each individual in the community. It has been shown that most empirical networks display heavy-tailed degree distributions, implying that in a large population there is a small proportion of individuals who have many links (friends), thus imposing a threat of infecting more individuals.

An important quantity in epidemics is the basic reproduction number, denoted by R_0, since if $R_0 \leq 1$ only minor outbreaks can occur, while if $R_0 > 1$ a major outbreak is possible (Anderson and May 1991). Any form of vaccination aims at reducing the reproduction number below one in order to prevent the disease from taking off. Vaccination strategies can be implemented before the epidemic starts (see e.g. Cohen *et al.* 2003, Madar *et al.* 2004), or as soon as the disease starts spreading, the latter being the focus of this thesis. The aim of the thesis is to propose two vaccination strategies during the early stages of an epidemic after the detection of an infectious individual. The epidemic model used is a Susceptible-Infectious-Removed (SIR) model, which confers lifelong immunity after recovery. In the first strategy, once an infectious individual is detected, his friends are vaccinated independently with probability ρ, and the second strategy sets a bound on the maximum number of possible infections from a given infectious individual before he is detected and vaccinates instantly all susceptible friends.

The first article assumes that a susceptible individual becomes infectious immediately if he/she has close contact with an infectious individual. The infectious period and the time to detection are modelled by independent exponential distributions. The main question is how the two vaccination strategies influence the reproduction number and hence the spread of the disease. For each vaccination strategy, the reproduction number is derived and simulations are performed for a Poisson and a heavy-tail degree distribution, and results show that the second strategy is more effective in both situations. The second article is an extension of the first article. It assumes that once an individual gets infected he/she is first latent (infected but not infectious) for some time, and then becomes infectious before he/she recovers and becomes immune. The latent period, infectious period and time to detection are modelled by independent gamma distributions. The main question is how the random properties of the latent period, infectious period and detection time affect the dynamics of the disease. For each vaccination strategy we approximate the initial stages of the epidemic by a branching process, thus allowing the computation of the probability of a major outbreak.

References

[1] Anderson, R. M. and May, R. M: Infections of diseases in humans - Dynamics and Control, Oxford University press, 1991.

[2] Cohen, R., Havlin, S and Ben-Avraham, D.: Efficient immunization strategies for computer networks and populations. *Phys. Rev. Lett.* **91** (2003), 247901.

[3] Madar N., Kalisky T., Cohen R., Ben-Avraham D., and Havlin S.: Immunization and epidemic dynamics in complex networks, *The European Physical Journal* **B** (2004).

[4] Newman, M.E.J: The structure and function of complex networks, *SIAM Rev.* **45(2)** (2003) 167-256.

Paper I

Social networks, epidemics and vaccination through contact tracing

Shaban Mbare[†] *, Mikael Andersson [†], Åke Svensson[†] , Tom Britton[†]

May 2007

Abstract

We consider a social network whose structure can be represented by a simple random graph having a pre-specified degree distribution F. A simple Markovian SIR network epidemic model is defined on such a social graph. We propose two real time vaccination strategies through contact tracing during the early stages of an epidemic outbreak. The first strategy considers vaccination of each friend of an identified infectious individual independently with probability ρ. The second strategy sets a bound m on the maximum number an infectious individual can infect before being identified. Expressions are derived for how these strategies influence the reproduction number. We give some numerical examples and simulation results based on the Poisson and heavy-tail degree distributions. We show that both vaccination strategies are effective in the Poisson degree distribution. In the heavy-tail degree distribution, the second vaccination strategy is much more effective.

Key words: social networks, degree distribution, reproduction number, contact tracing, vaccination strategies, delay time.

1 Introduction

Social networks are often described by simple undirected random graphs in order to capture social relationships among different individuals (Scott, 2000). Usually the vertices of the graph correspond to individuals and the edges to some social relations (Newman, *et al.*, 2001). On such a social graph an epidemic model may be defined, where initially individuals are free from the disease. An infectious individual can infect its susceptible friends (those who do not have the disease yet, but can catch it), before it recovers and becomes immune. The identification of individuals that have been in contact with an infectious individual (contact tracing) has attracted attention as a disease control measure that seeks to uncover newly infected cases preferably before they become infectious (e.g.

*Department of Mathematics, University of Dar es Salaam, Postal address: Box 35062, Dar es Salaam, Tanzania. *E-mail*: shabanmbare@yahoo.com

[†]Department of Mathematics, Stockholm University, SE-106 91 Stockholm, Sweden.

Kiss, *et al.,* 2005, Eames and Keeling, 2003, Eames 2006). The so traced individuals who are still susceptible can be vaccinated or immunized in order to prevent a major outbreak. How to contain the disease before it takes off is the question that can be addressed by the choice of a vaccination strategy.

In the present paper we study issues arising from such modelling. In particular we consider a simple social graph of a fixed population where the vertex degree (number of friends) follows a pre-specified distribution F. The social graph is assumed to be otherwise completely random. The simplest epidemic model is the Markovian susceptible-infectious-removed (SIR) for the spread of the disease in the social network (Anderson and May, 1992, Andersson and Britton, 2000). Initially one randomly selected individual is externally infected. Any individual who gets infected infects each of his/her friends independently of each other at the rate λ during the infectious period, and then recovers and become immune (removed). The infectious periods of different individuals are independent and identically exponentially distributed with mean $1/\gamma$. For this social graph and epidemic model we study two vaccination strategies: vaccination of located friends after delay (the first strategy) and vaccination of friends after delay also bounding on the maximum number of possible infections (the second strategy). Both strategies are implemented during the early stages of the epidemic after tracing the contacts or friends of the infectious individual in order to uncover possible chains of infection.

In a fixed large population of size n we derive the expected number of secondary cases generated by a typical infectious individual, in a susceptible population, often called basic reproduction number and denoted by R_0, for an epidemic without intervention. The quantity R_0 is of fundamental importance to the dynamics of infections, since a major outbreak is possible if and only if $R_0 > 1$ (Andersson and Britton, 2000). Using Poisson and the heavy-tail (scale-free) degree distributions for social networks, we show that both strategies can be effective for the Poisson case. The first strategy is less effective in the heavy-tailed social network because infected individuals with high degree still have the chance to infect many individuals.

Much work has been done on vaccination strategies prior to arrival of the disease. For example, targeted vaccination (Cohen *et al.,* 2003), uniform vaccination (Madar *et al.* 2004, Pastor-Satorras and Vespignani 2001, 2002) and acquaintance vaccination (Cohen *et al.,* 2003). In the uniform and acquaintance vaccination strategies individuals are chosen randomly, and targeted vaccination strategy requires the knowledge of individuals with high degrees. However, during the early stages of an epidemic, contact tracing can be used as a control measure of epidemic in a social network (Huerta and Tsimring, 2003), where susceptible individuals that have been in contact with an identified infectious person are found and get vaccinated (or some other type of immunization). The present paper contributes to this work by proposing two real time vaccination strategies that can be performed during the early stages of an epidemic outbreak, and requiring no knowledge of individual's degree prior to the detection of an infectious person. Another

aim of the paper is to derive the expressions of the reproduction number and compare the effectiveness of the vaccination strategies in preventing major outbreaks. The rest of the paper is organized as follows. In Section 2 we describe a model to generate a social network with a given degree distribution. We also define a simple epidemic model for a disease spreading on the social network. Section 3 treats vaccination as a measure to prevent major outbreak, and discusses two vaccination strategies. Section 4 provides numerical examples and simulation results of the model. Finally, Section 5 discusses the results and provides concluding remarks.

2 The Model

2.1 Social structure

Given that the degree distribution F is the only information we have, the model describes a way to construct an undirected graph of the social structure and we define it as follows. Take a set of n vertices and for each vertex assign a number of stubs $D = k$ independently from a random variable D with distribution $F = \{p_k\}$, where $p_k = P(D = k)$ is the probability that a randomly selected vertex has degree k. Then choose those stubs completely at random pairwise and join them up to form edges between the vertices. That is, first pick two stubs randomly among all stubs in the graph and join them. Then pick two stubs at random from the remaining stubs and join them, and so on. The vertices i and j are neighbours if there is an edge between them. For details on how to generate simple undirected graph we refer the reader to Britton, *et al*, (2006) and references therein. This procedure produces a graph with the desired degree distribution as $n \to \infty$, but which in all other respects is random. That is, we generate a graph which is drawn uniformly at random from the set of graphs with the given distribution.

2.2 A Simple epidemic model on the social graph

We now define an epidemic process taking place on the social network described above. As mentioned before, we consider three states Susceptible, Infectious and Recovered (and immune) that an individual can experience during an epidemic process. The Markovian SIR network epidemic model is used to describe the dynamic process of infections in the population through individual contacts (which are edges in the graph), and the epidemic process is defined as follows. Assume initially that all individuals are susceptible except one randomly selected individual who is externally infected. The infected individual remains infectious for a time period according to a random variable I which follows an arbitrary distribution G, and after the period I the individual recovers and becomes immune. During this infectious period I an infectious individual makes contacts with each of his/her friends according to independent Poisson processes with intensity λ. This implies that the first contact between the infectious individual and a particular neighbour takes place time T after the infection, where T is exponentially distributed with mean $\frac{1}{\lambda}$. For this situation, we assume that there is no latent period and any susceptible friend contacted by an infectious individual becomes immediately infectious. Those who

become infected behave similarly, and the epidemic process goes on until there are no more infectious individuals, when the epidemic stops. It is worth noting that the time T_i to contact a specific friend i is different from time T_j to contact another friend j, where $i \neq j$, and the sequence of times $\{T_i\}_{i=1}^{k}$, k being the number of friends, are independent and exponentially distributed, and only contacts occuring in the interval $[0, I]$ lead to infections.

Given an infectious individual, transmission of infection to a susceptible friend occurs if and only if the first contact takes place during the infectious period. That is, the time T of the infectious individual to contact a specified susceptible friend is less than the infectious period I. We compute the probability p that transmission of infection can occur by conditioning on time T as follows,

$$p = P(T < I) = \int_0^\infty P(T < I | T = t)\lambda e^{-\lambda t} dt = \int_0^\infty (1 - G(t))\lambda e^{-\lambda} dt. \tag{2.1}$$

To determine the basic reproduction number of the epidemic process, we consider an infectious individual in the second generation since the initial infectious individual is atypical with all friends being susceptible and having degree distribution F. In the second generation an infectious individual with k neighbours is selected with probability proportional to kp_k. If the selected person has degree k, he/she generates on average $(k-1)p$ new cases since the individual he/she was infected by is not susceptible. Hence, it follows that the basic reproduction number then becomes;

$$R_0 = p \sum_k (k-1) \frac{kp_k}{\sum_j jp_j},$$

which can be represented as (see Andersson (1999) for a similar argument)

$$R_0 = p \left(E(D) + \frac{Var(D) - E(D)}{E(D)} \right). \tag{2.2}$$

Suppose the infectious period follows an exponential distribution with parameter γ, so that $G \equiv \text{Exp}(\gamma)$. The probability p (defined in Equation (2.1)) that transmission of the infection takes place then becomes

$$p = P(T < I) = \int_0^\infty P(I > T | T = t)\lambda e^{-\lambda t} dt = \int_0^\infty e^{-\gamma t}\lambda e^{-\lambda t} dt = \frac{\lambda}{\lambda + \gamma}. \tag{2.3}$$

The corresponding basic reproduction number in Equation (2.2) then becomes

$$R_0 = \frac{\lambda}{\lambda + \gamma} \left(E(D) + \frac{Var(D) - E(D)}{E(D)} \right). \tag{2.4}$$

3 Vaccination Strategies

Vaccination is a major tool which is used to protect individuals in a population against infectious diseases. Similar effects to vaccination may be obtained through isolation and quarantine. In this section we discuss two vaccination strategies that can be used during the early stages of an epidemic upon detection of the disease case. If a traced individual is still susceptible, he/she is vaccinated and immediately makes the transition from susceptible to removed (meaning immune). We assume that a perfect vaccine is available, and that vaccination has no effect on an individual who has already been infected. To understand the dynamics of the disease and the performance of the control polices we derive expressions for the reproduction number of the vaccination strategies and compare their effectiveness.

3.1 Vaccination of located friends after delay

We assume that it takes some time between the point an individual gets infected to the point at which it is recognized by the health authorities and vaccinations are performed, because in reality it is not easy to detect the individual immediately at the point of infection. The time that elapses before the disease case is found and friends located, is herein referred to as delay (or detection) time. This delay time is a random quantity which may include, time before symptoms appear, time for authorities to detect the disease case, and time to find friends in order to vaccinate them. We assume for simplicity that all located friends are vaccinated at the same time.

Let S be the delay time whose distribution is H with parameter θ. The strategy relies on the infectious individual for information of his friends (contacts) in the event that he is detected, and assumes that only a proportion ρ of the contacts are found and become vaccinated. In this case we introduce an indicator random variable X such that $X = 1$ (which happens with probability ρ) if a given friend of an infectious individual is found and $X = 0$ otherwise. Taking into account the infectious period I and the time T to contact a given friend, transmission of infection occurs always if $T < \min(S, I)$, or if $S < T < I$ and $\{X = 0\}$. Conditioning on $T = t$ the probability q that a contact results into the transmission of infection is given by

$$\begin{aligned} q &= P(T < \min(S, I)) + P(S < T < I \cap \{X = 0\}) \\ &= \int_0^\infty \Big(P\big(T < \min(S, I) | T = t\big) + (1 - \rho) P\big(S < T < I | T = t\big) \Big) \lambda e^{-\lambda t} dt. \end{aligned}$$

When we consider the arbitrary distributions of I and S as before, that is, G and H respectively, we get the following general relation for the probability of infection,

$$q = \int_0^\infty \Big((1 - G(t))(1 - H(t)) \Big) \lambda e^{-\lambda t} dt + (1 - \rho) \int_0^\infty \Big(H(t)(1 - G(t)) \Big) \lambda e^{-\lambda t} dt. \quad (3.1)$$

As before, during the early stages an infectious individual has degree k with probability $\frac{k p_k}{E(D)}$, and will then on average infect $(k - 1)q$ individuals. The corresponding general

reproduction number R hence equals

$$R = q\sum_k (k-1)\frac{kp_k}{E(D)}. \tag{3.2}$$

From now on, we assume that the infectious period I and detection time S are independent exponentially distributed with parameters γ and θ respectively, meaning that $G \equiv \text{Exp}(\gamma)$ and $H \equiv \text{Exp}(\theta)$. The probability of infection in Equation (3.1) then becomes

$$\begin{aligned} q &= \int_0^\infty \lambda e^{-(\lambda+\gamma+\theta)t}dt + (1-\rho)\int_0^\infty (1-e^{-\theta t})\lambda e^{-(\lambda+\gamma)t}dt \\ &= \frac{\rho\lambda}{\lambda+\gamma+\theta} + \frac{(1-\rho)\lambda}{\lambda+\gamma}. \end{aligned} \tag{3.3}$$

The corresponding reproduction number $R_{\theta,\rho}$ in Equation (3.2) equals

$$R_{\theta,\rho} = \sum_k \Big(\frac{\rho\lambda}{\lambda+\gamma+\theta} + \frac{(1-\rho)\lambda}{\lambda+\gamma}\Big)(k-1)\frac{kp_k}{E(D)} \tag{3.4}$$

Using Equation (2.4) and some algebra, Equation (3.4) can be simplified to yield the representation of $R_{\theta,\rho}$ in terms of basic reproduction number R_0 as

$$R_{\theta,\rho} = \Big(1 - \frac{\rho\ \theta}{\lambda+\gamma+\theta}\Big)R_0. \tag{3.5}$$

It is obvious from Equation (3.5) that if the basic reproduction number is very large then so is $R_{\theta,\rho}$. In order to surely prevent an epidemic outbreak when this vaccination strategy is applied the corresponding reproduction number should be less than or equal to one, that is $R_{\theta,\rho} \leq 1$ implying that

$$\Big(1 - \frac{\rho\ \theta}{\lambda+\gamma+\theta}\Big)R_0 \leq 1.$$

To be more precise, the ρ and θ must hence satisfy

$$\frac{\rho\theta}{\lambda+\gamma+\theta} \geq 1 - \frac{1}{R_0},$$

and an approximate conservative bound is thus given by

$$\rho\theta \geq \Big(1 - \frac{1}{R_0}\Big)\big(\lambda+\gamma\big).$$

In this strategy there are two special cases. First, if all friends of an infectious person are located and become vaccinated instantly, then $\rho = 1$. Equation (3.3) reduces to

$$q = \frac{\lambda}{\lambda+\gamma+\theta},$$

and the corresponding reproduction number in (3.5) becomes

$$R_{\theta,\rho=1} = \frac{\lambda+\gamma}{\lambda+\gamma+\theta} R_0. \tag{3.6}$$

In this special case, the reproduction number $R_{\theta,\rho=1}$ is assured to be below unity if the choice of parameters satisfy the inequality,

$$\theta \geq (\lambda+\gamma)(R_0-1),$$

or equivalently if the expected time to detection $\frac{1}{\theta}$ satisfies

$$\frac{1}{\theta} \leq \frac{1}{(\lambda+\gamma)(R_0-1)}$$

implying that the detection intensity θ must be large enough in order to avoid the spread of infection in the social network. A second special case is when the detection intensity θ is so large that the first quantity in the second equality of Equation (3.3) approaches zero, and thus probability of transmission of infection becomes,

$$q = \frac{(1-\rho)\lambda}{\lambda+\gamma}.$$

This is equivalent to detecting an infectious individual immediately when he/she becomes infected and vaccinate a proportion ρ of all friends. Using the same argument as before, the corresponding reproduction number in Equation (3.5) reduces to

$$R_\rho = (1-\rho)R_0, \tag{3.7}$$

which is linear in ρ for a known value of R_0, and surely there will be only a minor outbreak if $\rho \geq 1 - \frac{1}{R_0}$. The interpretation of this is that, if R_0 is very large, for instance a heavy tailed network, then nearly all friends must be vaccinated. The criterion that $\rho \geq 1 - \frac{1}{R_0}$ is the same as in a general vaccination programme in order to obtain herd immunity. The difference with the approach suggested here is that it is only necessary to vaccinate around those who get infected. If that is possible, the spread can be controlled with considerable fewer vaccinations than if the vaccination took place before an outbreak.

3.2 Vaccination after delay with bounds on the maximum number of infections.

In the first strategy, the reproduction number $R_{\theta,\rho}$ becomes large if the basic reproduction number R_0 is large, which of course, is often the case with heavy tail degree distributions. We are then motivated to introduce a new vaccination strategy which aims at reducing further the reproduction number by controlling the individuals who have many friends (super-spreaders) by setting a bound on the number of possible infections from a given infectious individual. Let S and I be the detection time and infectious period as before, and we assume that they are exponentially distributed with parameters θ and γ

respectively. The strategy is defined as follows. An infectious individual contacts his/her friends independently at different times $T_i^{(k)}$, $i = 1, \ldots, k-1$ with rate λ. This implies that the ordered times $T_{(1)}^{(k)} \leq T_{(2)}^{(k)} \leq \ldots \leq T_{(k)}^{(k)}$ satisfy $T_{(1)}^{(k)} \sim \text{Exp}((k-1)\lambda)$, $T_{(2)}^{(k)} - T_{(1)}^{(k)} \sim \text{Exp}((k-2)\lambda)$, and so on. The individual infects friend i if $T_{(i)}^{(k)} < \min(S, I)$. We set an arbitrary maximum bound m say, of the number of friends that can be infected before an infectious individual is detected. This means that such an infectious person can at most infect m friends after which all remaining susceptible friends are vaccinated at the same time.

The number of infected friends by a given infectious individual is a random variable Y, given by

$$Y = \min\left\{ \max\left(i : T_{(i)}^{(k)} < \min(S, I)\right),\ m \right\}.$$

In order to compute the expected number of infected friends caused by a typical infectious individual during the early stages, we first determine the distribution of Y. During infectious period, infection occurs before detection with probability $\lambda/(\lambda + \gamma + \theta)$, hence, conditioning on the degree $D = k$ of an infectious individual, the probability that at least one friend is infected is

$$P(Y \geq 1 | D = k) = \frac{(k-1)\lambda}{(k-1)\lambda + \gamma + \theta}.$$

Similarly the probability that at least two friends are infected is

$$P(Y \geq 2 | D = k) = \frac{(k-1)\lambda}{(k-1)\lambda + \gamma + \theta} \cdot \frac{(k-2)\lambda}{(k-2)\lambda + \gamma + \theta}.$$

In general, for $Y \geq i$, $i = 0, \ldots, m$, the probability that at least i friends get infected is

$$\begin{aligned} P(Y \geq i | D = k) &= \frac{(k-1)\lambda}{(k-1)\lambda + \gamma + \theta} \cdot \frac{(k-2)\lambda}{(k-2)\lambda + \gamma + \theta} \cdots \frac{(k-i)\lambda}{(k-i)\lambda + \gamma + \theta} \\ &= \prod_{j=1}^{i} \left(\frac{(k-j)\lambda}{(k-j)\lambda + \gamma + \theta} \right) \end{aligned} \tag{3.8}$$

Since an infectious individual cannot infect more than m friends, then $P(Y > m) = 0$. From the law of total probability we obtain the expression for the probability that at least i friends get infected as

$$P(Y \geq i) = \sum_{k=i+1}^{\infty} P(Y \geq i | D = k) \frac{k p_k}{E(D)},\ i = 0, \ldots, m. \tag{3.9}$$

The sum in (3.9) starts at $k = i + 1$ since in order to infect i friends an infectious person needs at least $i + 1$ friends, one being the infector. The expected number of individuals

infected by the infectious person, which is also the reproduction number, is

$$\begin{aligned} R_{\theta,m} = E(Y) &= \sum_{i=1}^{m} P(Y \geq i) = \sum_{i=1}^{m} \sum_{k=i+1}^{\infty} P(Y \geq i | D = k) \frac{k p_k}{E(D)} \\ &= \sum_{i=1}^{m} \sum_{k=i+1}^{\infty} \prod_{j=1}^{i} \left(\frac{(k-j)\lambda}{(k-j)\lambda + \gamma + \theta} \right) \frac{k p_k}{E(D)}. \end{aligned} \tag{3.10}$$

The epidemic outbreak is surely avoided if $R_{\theta,m} \leq 1$, implying that for the strategy to have a strong impact as an attempt to avoid outbreaks, the expected number of infected people $E(Y)$ should be less than or equal to one. We should note here that since $R_{\theta,m}$ depends on the parameters λ and γ, then it also depends on R_0, but not as explicit as $R_{\theta,\rho}$ in the first strategy.

4 Examples and simulations

We have compared our two vaccination strategies by simulations for different parameter values always using the same population of size $n = 1000$ individuals. We used two different degree distributions, the Poisson degree distribution with mean 4 and a heavy-tail (scale-free) degree distribution, to compare the effectiveness of the strategies. The heavy-tail degree distribution follows the power law of the form of $p_k = ck^{-\alpha}$, where $c \approx (1.34)^{-1}$ is a normalizing constant and $\alpha = 2.5$ in our case. In order to have a reasonable comparison, we want a heavy tail distribution with the same mean as in the Poisson distribution.

We modify the power law to the form $p_k = c(k+1)^{-\alpha}$ $k = 0, 1, \ldots$, allowing that some individuals may not have contacts with others (that is $k = 0$) in the social network. The mean degree of the heavy-tail distribution is approximately 0.9. We modified this by defining a new random variable Z as the sum of two independent random variables D_p from the Poisson distribution with mean 3.1 and D_h from the heavy-tail distribution with mean 0.9. The distribution of Z is the convolution of the distributions of D_p and D_h. Since the Poisson degree distribution has a shorter tail and its variance is much less than the variance of the heavy-tail distribution, then we assume that Z is approximately heavy-tail distributed with mean 4, the same as the Poisson distribution.

We have chosen both the contact rate λ and the recovery rate γ to be one in all simulations. For the Poisson degree distribution with mean 4, the basic reproduction number is exactly 2 (from Equation (2.4)). In the heavy-tail degree distribution, the theoretical basic reproduction number is infinite ($R_0 = +\infty$), but in a finite population the basic reproduction number is of course finite. For our population size, the basic reproduction number of our heavy-tail distribution is approximately 7.2, and it is computed (from Equation (2.4)) with mean degree equals to 4 and the degree variance is 45.6 which is the difference between the second moment and the square of the mean of the degrees of a heavy tail distribution.

In order to compare the performance of the two vaccination strategies in preventing an outbreak, we have chosen the detection rate θ, the proportion of located friends ρ and the bound of the maximum number m one person can infect to assume the values $\theta = 1, 5, 20$, $\rho = 0.2, 0.5, 1$ and $m = 2, 5, 10$. These parameter values are just a representation of many other values which can be chosen for the same purpose. Based on Equations (3.5) and (3.10) the interest is to observe the behaviour of the reproduction numbers in the two vaccination strategies as the values of θ, ρ and m vary. Hence, these parameter values are designated as small ($\theta = 1$, $\rho = 0.2$, $m = 2$), intermediate ($\theta = 5$, $\rho = 0.5$, $m = 5$) and large ($\theta = 20$, $\rho = 1$, $m = 10$) with reference to the simulation results.

We have performed 500 simulations for an epidemic without vaccination and each of the two vaccination strategies. The social graph is generated once in order to have a common social structure, and in each simulation a different initial infectious individual is chosen randomly from the graph. We then obtain the proportions of minor outbreaks and the average sizes among major outbreaks. We define $\pi^{(1)}$ and $\pi^{(2)}$ as the proportions of 500 simulations having ten or less infected individuals in the first and the second vaccination strategies respectively. It implies that more than ten infected individuals is interpreted as a major outbreak. We also let $\mu^{(1)}$ and $\mu^{(2)}$ respectively be the average sizes among major outbreaks in the first and the second vaccination strategies. The summary of the results are shown in Tables 1 and 2.

In Table 1 numerical results from the Poisson degree distribution are presented. There are nine combinations of parameter values comprised of the pairs (θ, ρ) from the first strategy, and nine combinations of the pairs (θ, m) from the second strategy. Each pair of the parameter values is used in the simulation to obtain the proportions $\pi^{(i)}$ of minor outbreaks and the average sizes among major outbreaks $\mu^{(i)}$, for $i = 1, 2$. Similarly, using Equations (3.5) and (3.10) respectively, the reproduction numbers corresponding to the pairs (θ, ρ) and (θ, m) were computed. Results indicate that both strategies are effective in reducing the reproduction numbers below one when the degree distribution is Poisson. The first strategy is effective when the detection rate of an infectious individual is intermediate or high (that is $\theta = 5$ or $\theta = 20$) and the proportion of located friends is large ($\rho \approx 1$). We also note that the second strategy performs well for all values of our choice for the maximum bound ($m = 2, 5, 10$) when the detection rate is high ($\theta = 20$).

Table 2 shows the corresponding results, but for the heavy-tail degree distribution. When the basic reproduction number $R_0 = 7.2$ (in our case) the first strategy performs fairly well when the detection rate is large ($\theta = 20$) and requires the vaccination of all friends of the infectious individuals ($\rho = 1$). From Equation (3.5), the first strategy seems to be less efficient since in reality R_0 is large in heavy-tail social networks as the population size n becomes large. The second strategy works well for the intermediate or high detection rate $\theta = 5$ or $\theta = 20$ and the maximum bound of possible infections is small ($m = 2$), but it can guarantee prevention of an outbreak when we have high detection rate and small bound of maximum number of individuals one infectious person can infect.

Table 1: Numerical values from a Poisson degree distribution with mean 4 and $\lambda = \gamma = 1$ and $R_0 = 2$. $R_{\theta,\rho}$ and $R_{\theta,m}$ are computed numerically whereas π and μ are obtained from the simulations.

Strategy 1	$R_{\theta,\rho}$	$\pi^{(1)}$	$\mu^{(1)}$	Strategy 2	$R_{\theta,m}$	$\pi^{(2)}$	$\mu^{(2)}$
$\theta = 1,\ \rho = 0.2$	1.87	0.372	651	$\theta = 1,\ m = 10$	1.33	0.604	201
$\theta = 1,\ \rho = 0.5$	1.67	0.460	470	$\theta = 1,\ m = 5$	1.30	0.622	178
$\theta = 1,\ \rho = 1$	1.33	0.598	208	$\theta = 1,\ m = 2$	0.98	0.762	42
$\theta = 5,\ \rho = 0.2$	1.71	0.426	494	$\theta = 5,\ m = 10$	0.57	0.966	17
$\theta = 5,\ \rho = 0.5$	1.28	0.616	157	$\theta = 5,\ m = 5$	0.56	0.946	15
$\theta = 5,\ \rho = 1$	0.57	0.968	14	$\theta = 5,\ m = 2$	0.51	0.982	13
$\theta = 20,\ \rho = 0.2$	1.64	0.438	418	$\theta = 20,\ m = 10$	0.18	1	–
$\theta = 20,\ \rho = 0.5$	1.09	0.704	67	$\theta = 20,\ m = 5$	0.18	1	–
$\theta = 20,\ \rho = 1$	0.18	1	–	$\theta = 20,\ m = 2$	0.17	1	–

Table 2: Numerical values from a heavy-tail degree distribution with mean 4 and $\lambda = \gamma = 1$ and $R_0 = 7.2$ $R_0 = 2$. $R_{\theta,\rho}$ and $R_{\theta,m}$ are computed numerically whereas π and μ are obtained from the simulations.

Strategy 1	$R_{\theta,\rho}$	$\pi^{(1)}$	$\mu^{(1)}$	Strategy 2	$R_{\theta,m}$	$\pi^{(2)}$	$\mu^{(2)}$
$\theta = 1,\ \rho = 0.2$	6.7	0.418	608	$\theta = 1,\ m = 10$	3.5	0.662	168
$\theta = 1,\ \rho = 0.5$	6.0	0.492	431	$\theta = 1,\ m = 5$	2.5	0.610	114
$\theta = 1,\ \rho = 1$	4.8	0.612	227	$\theta = 1,\ m = 2$	1.1	0.758	32
$\theta = 5,\ \rho = 0.2$	6.2	0.462	462	$\theta = 5,\ m = 10$	2.3	0.940	21
$\theta = 5,\ \rho = 0.5$	4.6	0.652	187	$\theta = 5,\ m = 5$	1.6	0.948	16
$\theta = 5,\ \rho = 1.0$	2.1	0.918	40	$\theta = 5,\ m = 2$	0.85	0.972	15
$\theta = 20,\ \rho = 0.2$	5.9	0.496	388	$\theta = 20,\ m = 10$	1.2	0.994	16
$\theta = 20,\ \rho = 0.5$	3.9	0.706	92	$\theta = 20,\ m = 5$	0.89	0.998	11
$\theta = 20,\ \rho = 1$	0.7	0.998	19	$\theta = 20,\ m = 2$	0.5	1	–

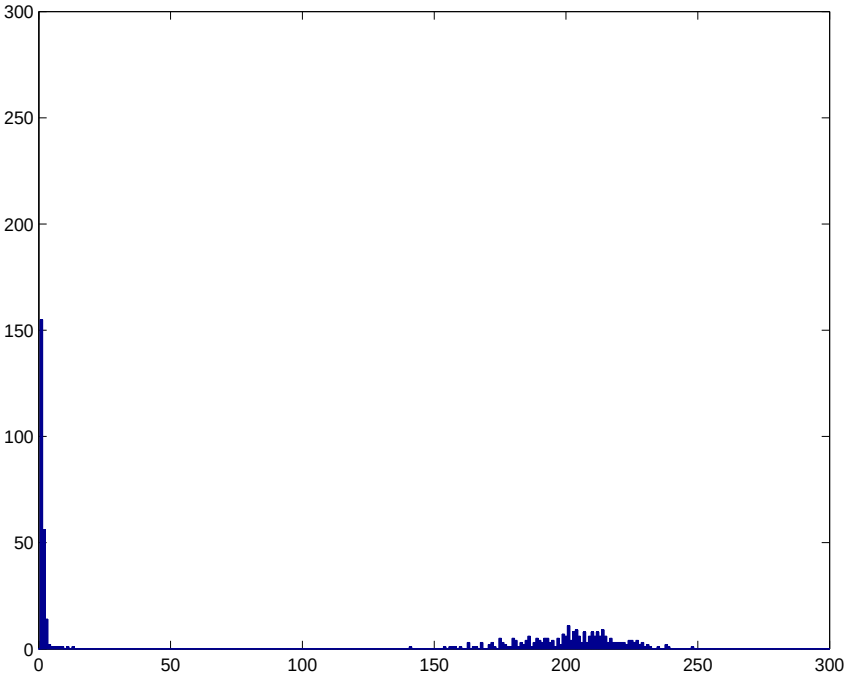

Figure 1: Histogram of final sizes for 500 simulations of an epidemic without vaccination in a heavy-tail degree distribution with $n = 1000$ individuals and $R_0 = 7.2 > 1$, i.e above threshold, indicating that there is a major outbreak.

Figures 1 and 2 $(A - B)$ report the outbreak sizes of the epidemic for the 500 simulations in the heavy-tail degree distribution. Figure 1 presents the outbreak size from 500 simulations of the epidemic without vaccination and the basic reproduction number R_0 equals $7.2 > 1$. Figure $2A$ shows the outbreak size after the implementation of the first vaccination strategy with detection rate $\theta = 20$ and the proportion of located friends $\rho = 1$. The proportion of minor outbreaks $\pi^{(1)} = 0.998$, showing that the first vaccination strategy is effective. This is in agreement with the computed reproduction number $R_{\theta,\rho} = 0.7$, which is below threshold. Figure $2B$ shows corresponding results but for the second vaccination strategy with the detection rate $\theta = 20$ and the bound of the maximum number of individuals one person can infect $m = 2$. The proportion of minor outbreak is $\pi^{(2)} = 1$ and null average size of major outbreak, an indication that the performance of the second vaccination strategy is satisfactory. This too agrees with the computed reproduction number $R_{\theta,m} = 0.5$, below threshold. Hence, we note from Figure 2 that both vaccination strategies are effective when compared with Figure 1 for epidemics without intervention, which clearly shows that there is an outbreak.

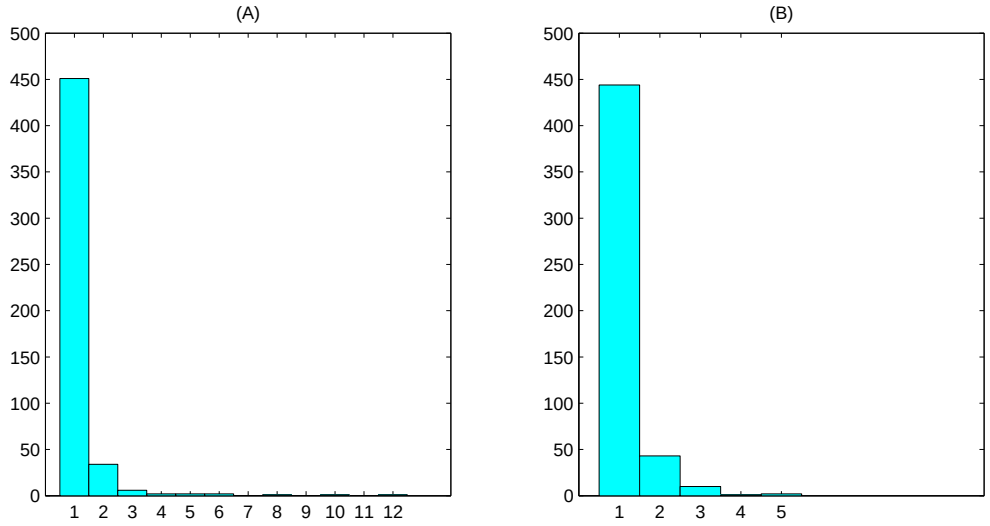

Figure 2: Histograms of outbreak sizes for 500 simulations with $n = 1000$ individuals in the heavy-tail degree distribution. In (A) we implemented the first vaccination strategy with $\theta = 20$ and $\rho = 1$, whereas in (B) is the second vaccination strategy with $\theta = 20$ and $m = 2$. Both strategies are effective since only few individuals are infected.

5 Discussion

In the present paper we studied two vaccination strategies during an epidemic outbreak. We used contact tracing as part of epidemic control in the social networks. The aim was to incorporate and determine the role of delay time, from the point an individual becomes infectious until his friends are vaccinated. Models of the reproduction numbers for the vaccination strategies were derived and compared through simulation and some numerical examples are obtained for Poisson and a heavy-tail degree distribution.

We have shown that both strategies are efficient when the number of friends, follows Poisson degree distribution. This result was expected, since Poisson distribution has a short tail and small degree variance. In the heavy-tail degree distribution the variance is always large (though finite in finite populations), which induces a large reproduction number. Our numerical example shows that the first strategy has less effect on the reproduction number than the second strategy. The second vaccination strategy is better since the effect of individuals with many friends (super-spreaders) is reduced.

However, it is worth to note that a large reproduction number does not necessarily imply a large outbreak, or high probability for a major outbreak. For instance, in the first strategy of Table 1, when $\theta = 20$ and $\rho = 0.5$, the reproduction number is above threshold ($R_{\theta,\rho} = 1.09$), but the proportion of minor outbreaks is high ($\pi^{(1)} = 0.704$) and the average of major outbreak is small ($\mu^{(1)} = 67$). Similarly, the first strategy in

Table 2, for $\theta = 5$ and $\rho = 1$ the reproduction number is above threshold ($R_{\theta,\rho} = 2.1$), but the proportion of minor outbreaks is high ($\pi^{(1)} = 0.918$) and the average of major outbreak is small ($\mu^{(1)} = 40$). Also the second strategy in Table 2, when $\theta = 5$ and $m = 5$, the reproduction number $R_{\theta,m} = 1.6$ (above threshold), giving high proportion of minor outbreak ($\pi^{(2)} = 0.948$) and small average of major outbreak ($\mu^{(2)} = 16$).

The model can be made more realistic in several ways. One underlying assumption is that an individual chooses his/her friends independently of each other. In real life there might be some assortative mixing (see Newman, 2003), meaning that individuals with many (few) friends are connected to individuals with many (few) friends. Also many networks show strong clustering, implying that there is positive probability that two individuals with a common friend are also friends. However, to include these in the model of social networks would make the analysis and comparison of the vaccination strategies harder.

Another assumption is that the social network is considered fixed over time. This is appropriate for diseases with short infectious periods but for the diseases with long infectious period, a dynamic social network would be preferred. The advantage we have in the fixed networks is that the expression of the reproduction number can easily be derived and the performance of the vaccination strategies can easily be analyzed and compared in more detail.

An important question not addressed in this paper is to study the effects of different distributions for the infectious period and delay time and to introduce a latency period. To derive expressions for the reproduction number in such distributions after the implementation of our vaccination strategies are important problems. Other interesting problems could be to determine the probability of outbreaks and the final size when these vaccination strategies are implemented and to compare the theoretical results with the corresponding simulation results in Table 1 and Table 2. More work is thus needed for the realistic epidemic modelling in social networks, and hence of the vaccination strategies. We believe the findings of the present paper will give some insight into possible effects of different vaccination strategies that are valid, also in more complex models.

Acknowledgements

Shaban Mbare gratefully acknowleges financial support from the International Science Programme (ISP) based at Uppsala University. He also thanks the Division of Mathematical Statistics at Stockholm University for hosting him for this research.

References

[1] Anderson, R. M. and May, R. M: Infections of diseases in humans - Dynamics and Control, Oxford University press, 1991.

[2] Andersson, H. and Britton, T.: Stochastic Epidemic Models and Their Statistical Analysis. Lecture Notes in Statistics, 151, Springer-Verlag, New York, 2000.

[3] Andersson, H.: Epidemic Models and Social Networks, *Math. Scientist* **24** (1999) 128-149

[4] Britton, T., Deijfen M., Martin-Löf A.: Generating random graphs with prescribed degree distribution. *J. Stat. Phys.*, **124** (2006) 1377-1397.

[5] Cohen, R., Havlin, S and Ben-Avraham, D.: Efficient immunization strategies for computer networks and populations. *Phys. Rev. Lett.* **91** (2003), 247901.

[6] Eames, K.T.D.: Contact tracing strategies in heterogeneous populations, *Epidemiol. Infect* (2006) 1-12.

[7] Eames, K.T.D., and Keeling, M.J.: Contact tracing and disease control, *Proc. R. Soc. Lond.* **B 270** (2003) 2565-2571.

[8] Huerta, R. and Tsimring, L.S: Contact tracing and epidemics control in social networks, *Phys. Rev. E* **66** (2002), 056115.

[9] Kiss, I.Z., Green, D.M. and Kao, R.R.: Disease contact tracing in random and clustered networks, *Proc. R. Soc.* **B 272** (2005) 1407-1414.

[10] Madar N., Kalisky T., Cohen R., Ben-Avraham D., and Havlin S.: Immunization and epidemic dynamics in complex networks, *The European Physical Journal* **B** (2004).

[11] Newman, M.E.J: The structure and function of complex networks, *SIAM Rev.* **45(2)** (2003) 167-256.

[12] Newman, M.E.J., Strogatz, S.H. and Watts, D.J.: Random graphs with arbitrary distributions and their applications. *Phys. Rev. E* **64** (2001), 026118.

[13] Pastor-Satorras, R. and Vespignani, A.: Immunization of complex networks, *Phys. Rev. E* **65** (2002), 036104.

[14] Pastor-Satorras, R. and Vespignani, A.: Epidemic and immunization in scale free networks. In *"Handbook of Graphs and Networks from the Genome to the Internet"*, eds. S. Bornholdt and H. G. Schuster, Wiley-VCH, Berlin, 2002, pp. 113-132.

[15] Scott, J.: Social Networks Analysis, *A Handbook*. 2nd ed., Sage, London, 2000.

[16] Tsimring, L.S. and Huerta, R.: Modelling of contact tracing in social networks *Physica A* **325** (2003) 33-39.

Paper II

Network epidemics and early stage vaccination: the effect of latent and infectious periods and their randomness

Shaban Mbare[†] *, Mikael Andersson [†], Åke Svensson[†] , Tom Britton[†]

May 2007

Abstract

Empirical evidence shows that the distributions of the latent and infectious periods affect the dynamics of the spread of an infectious disease. This paper treats the SEIR epidemic model describing the spread of an infectious disease giving life-long immunity, in a community whose social structure can be represented by a simple random graph having a pre-specified degree distribution. Two real time vaccination strategies, based on tracing the friends of infectious individuals during the early stages of an epidemic are proposed. The first strategy considers vaccination of each friend of a detected infectious individual independently with probability. The second strategy sets an upper bound on the number of friends an individual can infect before being detected. We approximate the initial phase of the epidemic by a branching process. We give two numerical examples: the Poisson and a heavy tail degree distribution, and show how the random properties of the latent period, infectious period and detection time affect the reproduction number and the probability of an outbreak.

Key words: degree distribution, social networks, epidemic models, vaccination strategies, coefficient of variation, branching approximation.

1 Introduction

Infectious diseases spread through contacts within the population of susceptible and infectious individuals. In network theory, contact structure can be represented by simple undirected random graphs (see e.g. Newman 2003, Andersson and Britton 2000), where vertices correspond to individuals and edges to some type of social relations, here referred to as friendships. On such a graph a model for the spread of a disease can be defined,

*Department of Mathematics, University of Dar es Salaam, Postal address: Box 35062, Dar es Salaam, Tanzania. *E-mail*: shabanmbare@yahoo.com

[†]Department of Mathematics, Stockholm University, SE-106 91 Stockholm, Sweden.

where individuals are at first susceptible. The disease spread is started by introducing a disease agent to one individual in the population, and after the latent period such an individual has the potential to spread the disease to her not yet infected friends. It is argued (see e.g. Wearing *et al.* 2005, Lloyd 2001) that the dynamics of the disease is affected by the distributions of the latent and infectious periods, which is the focus of the present paper. The period from the point an individual gets infected to the beginning of the state of becoming infectious is what we term as the latent period.

In this paper we model the latent period, infectious period and detection (delay) time by independent gamma distributions. In particular we consider a simple random social graph in a closed population, where the degree distribution (number of friends) follows some pre-specified distribution F. A simple epidemic model that can mimic the dynamics of the disease is the susceptible-exposed-infectious-removed (SEIR) model (see e.g. Anderson and May 1991, Hethcote 2000). Any individual who becomes infected, infects each of her friends independently at rate λ, before she recovers and becomes immune (a state known as removed). For this social graph and epidemic model we study two real time vaccination strategies: vaccination of located friends after delay (the first strategy) and vaccination of friends after delay with a bound on the maximum number of possible infections (the second strategy). By real time we mean that both strategies are implemented during the early stages of the epidemic after tracing the friends of infectious individuals in order to uncover possible chains of infection. We investigate the effects of random properties of latent and infectious periods on the disease outbreaks. The same vaccination strategies were treated in Mbare *et al.,* (2007), but then assuming no latent period and exponentially distributed infectious periods and detection times.

The SEIR epidemic model has been extensively studied (e.g. Bailey 1975, Diekmann and Heesterbeek 2000, Hethcote 2000) with the assumption that the latent (when considered) and infectious periods are exponentially distributed. The exponential distribution is mathematically convenient, but it corresponds to the assumption that the chance of recovery in a given time interval is independent of the time since infection. It is argued (e.g. Gough 1977, Lloyd 2001), that in practice the chance of recovery (in a given time interval) is initially small but increases overtime. This indicates that the infectious period distribution is less dispersed and more closely centred around its mean than the exponential distribution. This can be interpreted that the exponential distribution overestimates the number of individuals whose duration of infection is much shorter or much longer than the mean.

A key quantity in planning vaccination strategies is the basic reproduction number which is denoted R_0. The basic reproduction number determines whether a major outbreak can occur or not. If $R_0 \leq 1$, only minor outbreaks can occur whereas if $R_0 > 1$ then there is a positive probability for a major outbreak (Andersson and Britton 2000). Given a closed large population of size n we derive R_0 for the epidemic without vaccination and show that the latent period has no effect on R_0 and the final size distribution. We also

derive the reproduction numbers for the two vaccination strategies, and use branching processes to approximate the probability of having a major outbreak. All this is done when assuming independent gamma distributions for the latent period, infectious period and detection times. We investigate how the random behaviour of the infectious period and detection times affect the reproduction number and the probability of major outbreaks.

The rest of the paper is structured as follows. The social graph and epidemic models are defined in Section 2, where the basic reproduction number is derived. Section 3 contains our two vaccination strategies and the corresponding branching process approximations. Examples and numerical illustrations are given in Section 4, and in Section 5 we discuss the results and provide concluding remarks.

2 The Model

2.1 Social structure

Let n be the population size which is assumed to be large and fixed. We define the model of the social network having a prespecified degree distribution $F = \{p_k\}_{k=0}^{n-1}$. That is, we are given the probabilities p_k that a randomly chosen vertex in the network has degree $D = k$. The model is defined as follows (see, eg. Molloy and Reed, 1995). Assign independent degrees $D_1, \ldots, D_n$ from F to the vertices and give a vertex with degree k, k stubs. Then join the stubs randomly pairwise to form edges between them. That is, first pick two stubs randomly among all stubs in the graph and join them. Then pick two stubs at random from the remaining stubs and join them, and so on. Vertices i and j are neighbours (friends) if there is an edge between them. For details on how to generate simple undirected random graphs we refer the reader to Britton *et al.* (2006) and references therein. This procedure produces a graph with the desired degree distribution as $n \to \infty$, but which in all other respects is completely random. That is, we generate a graph which is drawn uniformly at random from the set of graphs with the given distribution. This procedure underestimates the degree for finite n, but the difference is negligible for larger n.

2.2 Epidemic model on the social graph

In this fixed network, we now describe a model of the spread of an infectious disease giving life-long immunity after recovery. Once an individual gets infected she is first latent for a $\Gamma(n_L, n_L\delta)$-distributed time, after which she becomes infectious and remains so for a period I being $\Gamma(n_I, n_I\gamma)$-distributed. This corresponds to the subdivision of latent and infectious periods into n_L and n_I stages respectively. The time spent in each stage is independent exponentially distributed with average lengths of $1/(n_L\delta)$ in the latent period and $1/(n_I\gamma)$ in the infectious period. The rates of movements between stages in the latent and infectious periods are $n_L\delta$ and $n_I\gamma$ in order to ensure that the average time spent in the latent and infectious periods are $1/\delta$ and $1/\gamma$ respectively. When the infectious period is over the individual recovers and becomes immune. This means that

the average infectious period is $1/\gamma$ with standard deviation $1/(\gamma\sqrt{n_I})$ and the squared coefficient of variation $1/n_I$. Similarly, the expected length of the latent period is $1/\delta$, the standard deviation is $1/(\delta\sqrt{n_L})$ and the squared coefficient of variation is $1/n_L$. During the infectious period an individual has close contact with a given friend according to a Poisson process with rate λ. By close contact we mean that a contact results in infection if the friend is susceptible. All Poisson processes describing infectious contacts, as well as latent and infectious periods, are defined to be mutually independent. The epidemic continues until there are no more latent or infectious individuals in the population, because then all individuals are either still susceptible or immune.

To determine the probability of transmission of infection we argue as follows. For this to happen an infectious individual must make the first contact before the end of the infectious period. If T denotes the time of the first contact (after the beginning of the infectious period) the probability of transmission of infection is hence

$$P(T < I) = \int_0^\infty P(T < I | T = t)\lambda e^{-\lambda t} dt = \int_0^\infty (1 - P(I \leq t))\lambda e^{-\lambda t} dt. \tag{2.1}$$

Since $I \sim \Gamma(n_I, n_I\gamma)$ with n_I integer, the distribution function of the infectious period I is

$$F_I(t) = 1 - \sum_{j=0}^{n_I-1} \frac{(n_I\gamma t)^j}{j!}\, e^{-n_I\gamma t}. \tag{2.2}$$

A related expression to (2.2)is obtained in Bain and Engelhardt (1991). As a consequence

$$1 - P(I \leq t) = \sum_{j=0}^{n_I-1} \frac{(n_I\gamma t)^j}{j!}\, e^{-n_I\gamma t}.$$

Hence (2.1) becomes

$$\begin{aligned} P(T < I) &= \lambda \sum_{j=0}^{n_I-1} \frac{(n_I\gamma)^j}{j!} \int_0^\infty t^j e^{-(n_I\gamma+\lambda)t} dt \\ &= 1 - \left(\frac{n_I\gamma}{\lambda + n_I\gamma}\right)^{n_I}, \end{aligned} \tag{2.3}$$

since

$$\int_0^\infty t^j e^{-(n_I\gamma+\lambda)t} dt = \frac{j!}{(\lambda + n_I\gamma)^{j+1}}.$$

During the early stages of the epidemic, an infectious individual has degree k with probability $\tilde{p_k} = \frac{kp_k}{E(D)}$, where $E(D) = \sum_j jp_j$. Given the degree k, the expected number of new cases she will infect is $(k-1)P(T < I)$, since all individuals except the one she was infected by are susceptible during the early stages. As a consequence, the basic reproduction number R_0 becomes

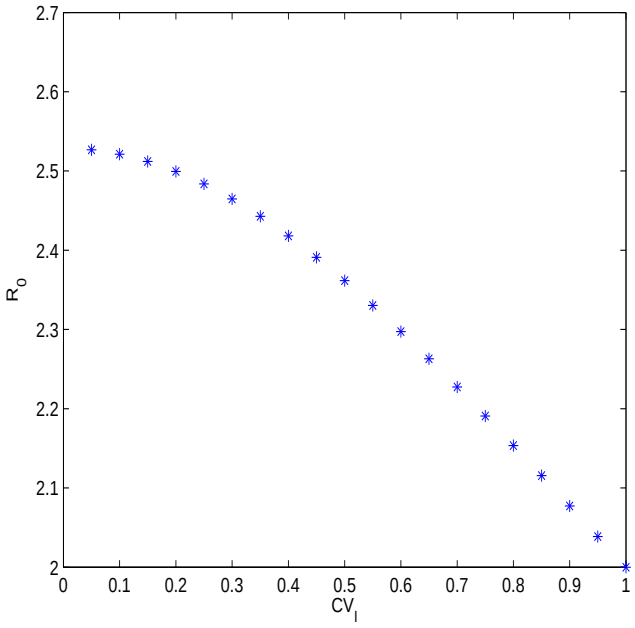

Figure 1: Basic reproduction number as a function of the coefficient of variation of the infectious period when $\lambda = \gamma = 1$. The degree distribution is Poisson with mean 4.

$$\begin{aligned} R_0 =& \Big[1 - \Big(\frac{n_I\gamma}{\lambda + n_I\gamma}\Big)^{n_I}\Big] \sum_k (k-1)\frac{kp_k}{E(D)} \\ =& \Big[1 - \Big(\frac{n_I\gamma}{\lambda + n_I\gamma}\Big)^{n_I}\Big] \Big(E(D) + \frac{Var(D) - E(D)}{E(D)}\Big). \end{aligned} \tag{2.4}$$

We note from Equation (2.4) that the latent period has no effect on the basic reproduction number R_0. For a fixed $E(D)$, λ and γ and finite n_I, R_0 is increasing in var(D), so the more variance in the degree distribution , the higher R_0, and if the degree distribution has infinite variance then $R_0 = +\infty$, a case which we do not pursue in this paper. Also, R_0 increases in λ when all other quantities in equation (2.4) are fixed, corresponding to more contacts leads to higher expected number of infected individuals. Similarly, R_0 increases when γ is decreasing, implying that the longer the average infectious period, the higher the basic reproduction number.

Using equation (2.4), we fix $\lambda = \gamma = 1$ and choose the degree distribution to be a Poisson with mean 4. We show in Figure 1, the basic reproduction number R_0, as a function of CV_I, the coefficient of variation of the infectious period I, that R_0 is decreasing with increasing coefficient of variation of the infectious period.

3 Vaccination strategies

Suppose that a vaccine is available which prevents individuals from becoming infected. For simplicity we only consider vaccines that give complete and life-long immunity. The goal is to contain the disease such that when the epidemic stops only a few or a small proportion of individuals have been infected. We propose two vaccination strategies based on tracing the friends following the detection of an infectious individual, and we investigate how effective the two strategies are in preventing the disease from taking off. We assume that vaccination has no effect on those individuals who have already been infected.

3.1 Vaccination of located friends after delay

It is always the case that some random time period elapses from the point an individual gets infected to the point at which she is recognized by the health authorities and vaccinations are performed, because in reality it is not easy to detect the individual immediately at the point of infection. The time that elapses, here referred to as detection (delay) time, is a random quantity which may include; time before symptoms appear, time for authorities to detect the disease case, and time to find friends in order to vaccinate them. We assume for simplicity that all located friends are vaccinated at the same time.

We assume that the latent and incubation periods are equal. The incubation period is defined as the time interval from the point of infection to the appearance of symptoms (see Anderson and May, 1991). We assume that detection time will always be larger than the latent or incubation period, which implies that an infectious individual can only be detected after showing disease symptoms, that is after the end of the incubation period. The strategy is then defined as follows. An infectious individual is detected after a time period S having $\Gamma(n_S, n_S\theta)$-distribution, where θ is the detection rate and n_S is a positive integer. Once an individual is detected each friend is vaccinated independently with probability ρ. The mean length of the detection time is $1/\theta$, the standard deviation is $1/(\theta\sqrt{n_S})$ and the squared coefficient of variation is $1/n_S$.

To determine how the strategy works, the reproduction number is an important quantity. Let X be an indicator variable such that $X_i = 1$ if a given friend i of an infectious individual is found (which happens with probability ρ), and $X_i = 0$ otherwise. Suppose T_i is the an infectious requires in order to contact friend i for the first time, then transmission of infection occurs if $T_i < \min(S, I)$, or if $S < T_i < I$ and $\{X_i = 0\}$. Conditioning on $T = t$, this event has probability p given by

$$\begin{aligned} p &= P(T < \min(S, I)) + P(S < T < I \cap \{X = 0\}) \\ &= \int_0^\infty \Big(P\big(T < \min(S, I) | T = t\big) + (1-\rho) P\big(S < T < I | T = t\big) \Big) \lambda e^{-\lambda t} dt \\ &= \int_0^\infty \Big(1 - P(I \le t)\Big)\Big(1 - P(S \le t)\Big) \lambda e^{-\lambda t} dt + (1-\rho) \int_0^\infty P(S < t) \Big(1 - P(I \le t)\Big) \lambda e^{-\lambda t} dt. \end{aligned} \tag{3.1}$$

Using independent gamma distributions for S and I in (3.1) we get that

$$\begin{aligned} p = &\int_0^\infty \sum_{j=0}^{n_I-1} \frac{(n_I\gamma t)^j}{j!}\ e^{-n_I\gamma t}\ \sum_{i=0}^{n_S-1} \frac{(n_S\theta t)^i}{i!}\ e^{-n_S\theta t}\ \lambda e^{-\lambda t} dt \\ &+ (1-\rho)\int_0^\infty \sum_{j=0}^{n_I-1} \frac{(n_I\gamma t)^j}{j!}\ e^{-n_I\gamma t}\ \left(1 - \sum_{i=0}^{n_S-1} \frac{(n_S\theta t)^i}{i!}\ e^{-n_S\theta t}\right)\ \lambda e^{-\lambda t} dt. \end{aligned} \tag{3.2}$$

Interchanging integration and sum and performing some algebra, (3.2) simplifies to the following expression

$$p = (1-\rho)\left[1 - \Big(\frac{n_I\gamma}{\lambda + n_I\gamma}\Big)^{n_I}\right] + \frac{\rho\lambda}{\lambda + n_I\gamma + n_S\theta}\ \xi, \tag{3.3}$$

where

$$\xi = \sum_{i=0}^{n_S-1}\sum_{j=0}^{n_I-1} \binom{i+j}{i} \Big(\frac{n_S\theta}{\lambda + n_I\gamma + n_S\theta}\Big)^i \Big(\frac{n_I\gamma}{\lambda + n_I\gamma + n_S\theta}\Big)^j.$$

For fixed λ, γ and finite n_I and n_S, p is decreasing with increasing θ, hence high detection rate minimizes the probability of infection.

As before, during the early stages of an outbreak, an infectious individual has degree k with probability $\tilde{p_k} = \frac{kp_k}{E(D)}$ and of these are $(k-1)$ susceptible. The reproduction number hence becomes (using equation (2.4))

$$\begin{aligned} R_{\theta,\rho} &= p\sum_k (k-1)\frac{kp_k}{E(D)} \\ &= (1-\rho)R_0 + \frac{\rho\lambda}{\lambda + n_I\gamma + n_S\theta}\ \xi\ \sum_k (k-1)\frac{kp_k}{E(D)} \\ &= \left[(1-\rho) + \frac{\rho\lambda}{\lambda + n_I\gamma + n_S\theta}\left(1 - \Big(\frac{n_I\gamma}{\lambda + n_I\gamma}\Big)^{n_I}\right)^{-1}\xi\right] R_0. \end{aligned} \tag{3.4}$$

Note that $R_{\theta,\rho}$ is increasing in R_0, so if the basic reproduction number is very large during the early stage of epidemic so is $R_{\theta,\rho}$. Also, for fixed R_0, and finite n_I and n_S, $R_{\theta,\rho}$ is decreasing as θ increases, meaning that the higher the detection rate the smaller the mean number of infected individuals.

It is of interest to gain some insight in how detection rate affect the reproduction number at different values of ρ, the proportion of detected friends. We show this in Figure 2 using equation (3.4), for the exponential case (when $n_I = n_S = 1$) and $\lambda = \gamma = 1$ and we choose $R_0 = 2$. Figure 2 shows $R_{\theta,\rho}$, the reproduction number as a function of the detection rate θ, for $\rho = 0.2$ and $\rho = 0.8$. We note that the reproduction number is greatly reduced as θ increases, implying that a short average detection time is important. We also see that the reproduction number is further reduced if a high proportion of individuals are vaccinated.

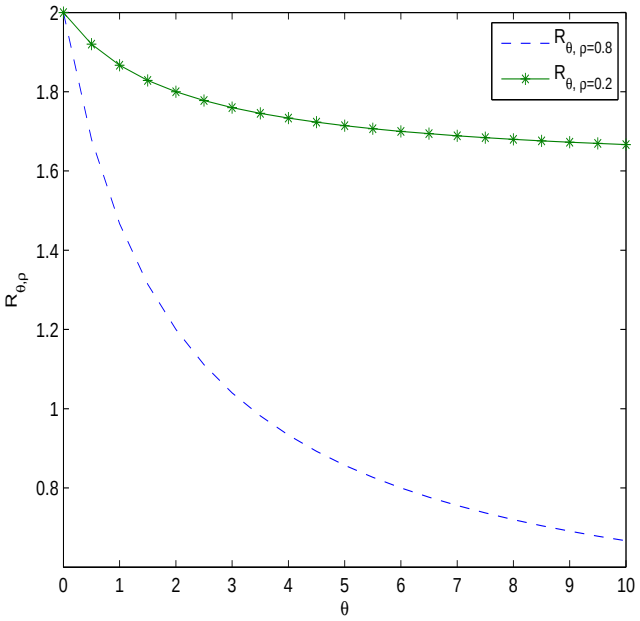

Figure 2: Reproduction number $R_{\theta,\rho}$, as a function of θ, the detection rate when $R_0 = 2$, and $\lambda = \gamma = 1$, for $\rho = 0.2$ and $\rho = 0.8$.

3.1.1 Branching process approximation

During the initial stages of an epidemic in a large population, the chance is high that contacted individuals are susceptible, implying that the number of infectious individuals may be approximated by a branching process (Ball, 1996). By "initial" we mean that during the onset of the disease only few or a small proportion of the population have been infected such that the probability of contact between two infectious individuals is negligible. In a branching process every infectious individual gives birth to (infects) a random number of offspring (infected) independently of each other, but with the same distribution. The process is assumed to start with Z_0 individuals and each individual capable of giving birth has a random lifespan equivalent to her infectious period I. In the following we assume that $Z_0 = 1$, meaning that initially there is only one infectious individual capable of infecting her friends. The approximation of the process can be made precise by coupling arguments as in Ball (1996), but that treatment is beyond the scope of the present paper.

A quantity of interest is the distribution of the number of offspring Z from one infectious individual, since given such a distribution the extinction probability can easily be derived. Given the infectious time period $I = i$ and detection time $S = s$, an individual contacts a particular friend with probability $\eta(s,i) = \rho(1-e^{-\lambda \min(s,i)}) + (1-\rho)(1-e^{-\lambda i})$. This means that $\eta(s,i) = \rho(1-e^{-\lambda s}) + (1-\rho)(1-e^{-\lambda i})$ if $s < i$ or $\eta(s,i) = 1-e^{-\lambda i}$ if $s > i$. So, given

that $D = k$, $S = s$ and $I = i$, the number of individuals infected in the first generation, is binomial$\Big(k-1,\ \eta(s,i)\Big)$. The parameter $k-1$ follows since the initially infectious individual was infected by one of her friends (who cannot be re-infected). Conditioning on $D = k$ alone, an infectious individual infects $Z = z$ friends with probability given by

$$\begin{aligned} P(Z = z|D = k) = &\int_0^\infty \int_0^i \binom{k-1}{z} \eta(s,i)^z \Big(1-\eta(s,i)\Big)^{k-1-z} f_S(s) f_I(i)\, \mathrm{d}s\ \mathrm{d}i \\ &+ \int_0^\infty \int_i^\infty \binom{k-1}{z} \eta(s,i)^z \Big(1-\eta(s,i)\Big)^{k-1-z} f_S(s) f_I(i)\, \mathrm{d}s\ \mathrm{d}i \end{aligned} \tag{3.5}$$

Unconditionally, the number of individuals infected in the next generation is hence given by

$$P(Z = z) = \sum_{k=0}^{\infty} P(Z = z|D = k) \frac{k p_k}{E(D)}. \tag{3.6}$$

All succeeding offspring give birth independently according to the same distribution $P(Z = z)$. An important tool in the analysis of the process is the probability generating function of the offspring distribution, defined by

$$\varphi(s) = \sum_{z=0}^{\infty} s^z P(Z = z), \text{ for some } |s| \leq 1. \tag{3.7}$$

The probability of extinction q can be found by conditioning on the number of offspring of the initial infectious individual. Ultimate extinction occurs if and only if all "daughter" families started by these offspring become extinct. Since each family is assumed to act independently, and since the probability that any particular branch dies out is just q, then it follows that $P(\text{extinction}|Z = z) = q^z$ and that

$$q = P(\text{extinction}) = \sum_{z=0}^{\infty} P(\text{extinction}|Z = z) P(Z = z),$$

implying that

$$q = \sum_{z=0}^{\infty} q^z P(Z = z) = \varphi(q). \tag{3.8}$$

The moments of the process, when they exist, can be expressed in terms of the derivatives of $\varphi(s)$ with respect to s and evaluated at $s = 1$. From branching process theory (e.g. Jagers, 1975), it is known that the extinction probability q is the smallest nonnegative solution to $\varphi(s) = s$, and that if the mean number of offspring $\mu = \varphi'(1) < 1$ the only solution is $q = 1$, and if $\mu > 1$ then there is a second solution $q < 1$ which is equal to extinction probability.

3.2 Vaccination after delay with bounds on the maximum number of infections

The reproduction number $R_{\theta,\rho}$ in the first strategy is a linear function of the basic reproduction number R_0. As a consequence, if R_0 is very large, so will $R_{\theta,\rho}$ be. The second strategy aims at reducing the reproduction number more than the first strategy. To achieve this we assume that it is possible to put an upper bound, m say, of the number of infections caused by one individual. By this we mean that an individual can at most infect m friends. We define the strategy by first stating the contact pattern of an individual. Let $T^{(k)}_{(i)}$, $i = 1, \ldots, k-1$, be the ordered times at which an infectious individual contacts the i^{th} friend among her $k-1$ susceptible friends. An individual contacts each friend independently at the rate λ. This means that $T^{(k)}_{(1)} \sim \text{Exp}((k-1)\lambda)$, $T^{(k)}_{(2)} - T^{(k)}_{(1)} \sim \text{Exp}((k-2)\lambda)$, ..., $T^{(k)}_{(i+1)} - T^{(k)}_{(i)} \sim \text{Exp}((k-i-1)\lambda)$, and all susceptible friends are vaccinated at the same time once an individual is detected.

We approximate the initial phase of the epidemic by a branching process. As before we have detection time S for an individual whose infectious time period is I. Conditioning on the degree $D = k$ and using Equation (3.3) with $\rho = 1$, (since all susceptible friends are vaccinated following detection in the present strategy), transmission of infection occurs with probability π given by

$$\pi = \begin{cases} p, & \text{if } k-1 < m \\ p\ \frac{m}{k-1}, & \text{if } k-1 \geq m \end{cases} \tag{3.9}$$

where

$$p = \frac{\lambda}{\lambda + n_I\gamma + n_S\theta} \sum_{i=0}^{n_S-1} \sum_{j=0}^{n_I-1} \binom{i+j}{i} \Big(\frac{n_S\theta}{\lambda + n_I\gamma + n_S\theta}\Big)^i \Big(\frac{n_I\gamma}{\lambda + n_I\gamma + n_S\theta}\Big)^j.$$

Equation(3.9) follows, since the bound m has a effect if it is less than $k-1$, the number of suceptible friends, otherwise the bound m has no role if it is larger than $k-1$. We proceed to determine the expected number of new cases say Z, generated by an individual during the early stage of the epidemic, and we get that

$$E(Z|D=k) = \pi(k-1) = \begin{cases} p(k-1), & \text{if } k-1 < m \\ p\ m, & \text{if } k-1 \geq m. \end{cases} \tag{3.10}$$

The reproduction number $R_{\theta,m}$, which in this case is the unconditional number of infected individuals, hence is given by

$$\begin{aligned} R_{\theta,m} = E(Z) &= \sum_k E(Z|D=k)\frac{kp_k}{E(D)} \\ &= \sum_{k=1}^{m} E(Z|D=k)\frac{kp_k}{E(D)} + \sum_{k=m+1}^{\infty} E(Z|D=k)\frac{kp_k}{E(D)} \\ &= \sum_{k=1}^{m} p(k-1)\frac{kp_k}{E(D)} + \sum_{k=m+1}^{\infty} p\ m\frac{kp_k}{E(D)}, \end{aligned} \tag{3.11}$$

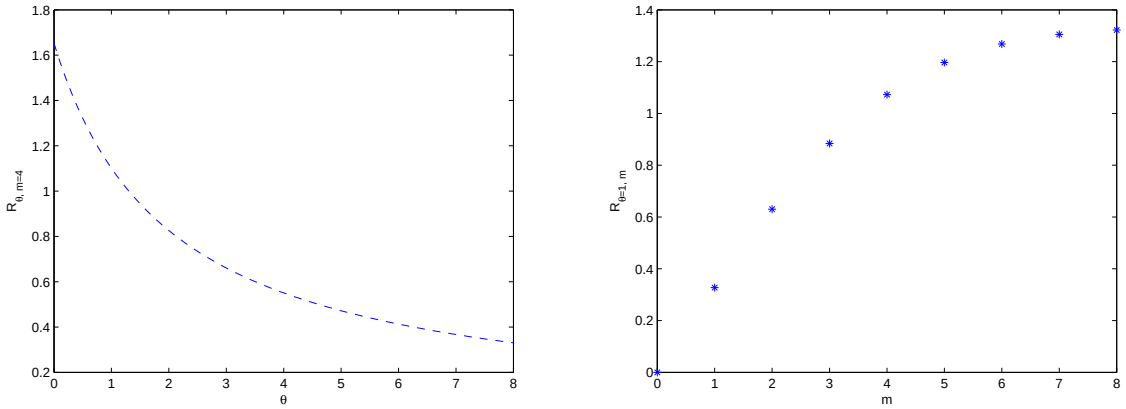

Figure 3: The reproduction number $R_{\theta,m}$, as a function of θ, the detection rate (left), and the bound of the maximum possible infections m (right). In both cases we treat the exponential case when $n_I = n_S = 1$, and fix $\lambda = \gamma = 1$.

where $P(D = k) = \frac{kp_k}{E(D)}$. If $R_{\theta,m} \leq 1$, only minor outbreaks can occur whereas if $R_{\theta,m} > 1$ a major outbreak can occur with a positive probability. In Figure 3, we show that $R_{\theta,m}$ is decreasing with increase in θ for a fixed $m = 4$. Similarly, $R_{\theta,m}$ is increasing with increasing m, for a fixed $\theta = 1$.

We now treat the special case where the infectious period and the detection time have independent exponential distributions with mean $1/\gamma$ and $1/\theta$ respectively, implying that $n_I = n_S = 1$. The probability that infection occurs before detection and the end of infectious period is $\lambda/(\lambda + \gamma + \theta)$. So given $D = k$, an individual will infect at least one friend in the first generation with probability given by (recall that $T^{(k)}_{(1)} \sim \text{Exp}((k-1)\lambda)$, $T^{(k)}_{(2)} - T^{(k)}_{(1)} \sim \text{Exp}((k-2)\lambda)$, and so on)

$$P(Z \geq 1) = \frac{(k-1)\lambda}{(k-1)\lambda + \gamma + \theta}.$$

Similarly, the probability that she infects at least two friends is equal to

$$P(Z \geq 2|D = k) = \frac{(k-1)\lambda}{(k-1)\lambda + \gamma + \theta} \cdot \frac{(k-2)\lambda}{(k-2)\lambda + \gamma + \theta}.$$

Hence, for $Z \geq l,\ l = 0,\ 1,\ \ldots, m$, an infectious individual will infect at least l individuals with probability given by

$$P(Z \geq l|D = k) = \prod_{j=1}^{l} \left(\frac{(k-j)\lambda}{(k-j)\lambda + \gamma + \theta} \right). \tag{3.12}$$

From (3.12) the unconditional probability that at least l individuals become infected,

hence becomes

$$P(Z \geq l) = \sum_{k=l+1}^{\infty} P(Z \geq l | D = k) \frac{kp_k}{E(D)}, \quad l = 0,\ 1, \ldots, m. \tag{3.13}$$

Consequently, the distribution of the number of individuals infected during the early stages of the epidemic is

$$\begin{aligned} P(Z = l) &= P(Z \geq l) - P(Z > l) \\ &= \sum_{k=l+1}^{\infty} P(Z \geq l | D = k) \frac{kp_k}{E(D)} - \sum_{k=l+2}^{\infty} P(Z \geq l | D = k) \frac{kp_k}{E(D)}, \end{aligned} \tag{3.14}$$

The first sum in equation (3.14) starts at $k = l + 1$, since in order to infect l friends (who must be susceptible) an infectious individual needs at least $l + 1$ friends. We get the probability generating function $\varphi(s)$ using equation (3.14) as

$$\varphi(s) = \sum_{l=0}^{\infty} s^l P(Z = l), \quad |s| \leq 1 \tag{3.15}$$

For the branching process to die out, all branching processes initiated by the individual during the early stages of the epidemic must die out, where the extinction probability q is the smallest nonnegative solution of the equation $q = \varphi(q)$ (see Jagers 1975).

4 Discussions

The aim in this paper was two-fold. First to study the effects of the random properties of the latent period, infectious period and detection time to the disease spreading on the networks, and secondly to approximate the initial phase of the epidemic by a branching process. All these are done after the implementation of the proposed vaccination strategies. The models which have been developed are not fully realistic, but we still believe that they may capture some of the relevant behaviour which would appear in more complex models. For example, a more realistic social structure would consider some network properties such as clustering, associativity and preferential attachment (see e.g Newman 2003). Also, from a more general point, it would be of interest to relax the assumption that an infectious individual contacts friends at the same rate. Another interesting extension could be to consider the varying susceptibility to the disease among individuals (see e.g. O'neill 2001).

Acknowledgements

S. Mbare thanks ISP at Uppsala University for funding him during this research. He would also like to express gratitude to the Department of Mathematics at Stockholm University for hosting and expose him to the World of Mathematics.

References

[1] Anderson, R. M. and May, R. M. (1991): Infections of diseases in humans - Dynamics and Control. Oxford University press.

[2] Andersson, H. and Britton, T. (2000): Stochastic Epidemic Models and Their Statistical Analysis. Lecture Notes in Statistics, 151, Springer-Verlag, New York.

[3] Bailey, N.T.J. (1975): The Mathematical Theory of Infectious Diseases, *2nd Ed.*, Griffin, London.

[4] Bain, J.L. and Engelhardt, M. (1991): Introduction to Probability and Mathematical Statistics, 2nd ed. *The Duxbury advanced series in statistics and decision sciences*, USA.

[5] Ball, F.G. (1996): Threshold behaviour in stochastic epidemics among households. In *Athens Conf. Applied Probability and Time Series, Vol. I: Applied Probability*

[6] Diekmann, O. and Heesterbeek, J.A.P (2000): Mathematical epidemiology of infectious diseases: model building, analysis and interpretation, John Wiley and sons, New York.

[7] Gough, K.J. (1977): The estimation of latent and infectious periods, *Biometrika* **64**, 559-565.

[8] Hethcote, W.H. (2000): The Mathematics of Infectious Diseases, *SIAM REVIEW* **42(4)**, 599-653.

[9] Jagers, P. (1975): Branching Processes with Biological Applications, A Wiley-Interscience Publications, London.

[10] Lloyd, A.L. (2001): Realistic Distributions of Infectious Periods in Epidemic Models: Changing Patterns of Persistence and Dynamics, *Theoretical Population Biology* **60**, 59-71.

[11] Mbare, S., Britton, T., Andersson, M. and Svensson, Å. (2007): Social networks, epidemics and vaccination through contact tracing, (Manuscript).

[12] Molloy, M. and Reed, B. (1995): A critical point for random graphs with a given degree sequence, *Random Structures and Algorithms* **6**, 161-179.

[13] Newman, M.E.J (2003): The structure and function of complex networks, *SIAM Rev.* **45(2)** 167-256.

[14] O'Neill, P.D. (2001): Inference for an epidemic when susceptibility varies, *Biostatistics* **2(1)**, pp 99-108.

[15] Wearing, H.J., Rohani, P. and Keeling, M.J. (2005): Appropriate Methods for the Management of infectious Diseases, *PLoS Med* **2(7)**: e174. 621-627