

# The influence of age-age correlations on epidemic spreading in social network

Andrzej Grabowski<sup>a</sup>

Central Institute for Labour Protection, National Research Institute, 00-701 Warsaw, Poland

Received 11 February 2014 / Received in final form 8 May 2014

Published online 1 July 2014

© The Author(s) 2014. This article is published with open access at [Springerlink.com](http://Springerlink.com)

**Abstract.** On the basis of the experimental data concerning interactions between humans the process of epidemic spreading in a social network was investigated. It was found that number of contact and average age of nearest neighbors are highly correlated with age of an individual. The influence of those correlations on the process of epidemic spreading and effectiveness of control measures like mass immunizations campaigns was investigated. It occurs that the magnitude of epidemic is decreased and the effectiveness of target vaccination is increased.

## 1 Introduction

Strong heterogeneity and mixing patterns are one of the most important properties of social network [1–3], because it strongly influences dynamic phenomena in social networks. Some examples of such phenomena are epidemic spreading, information flow, rumour propagation or opinion formation [4–7]. Real-world, as well as on-line, social networks have attracted the attention of the physicists community. Because the available data sets are vast, it is possible to use techniques of statistical physics [2], e.g. large on-line networks like e-mail networks [8] and web-based social networks [9] or artificial communities (e.g. players who interact in the large virtual world of a massive multiplayer online role playing game [10]) were investigated. Recent questionnaire studies [11] have provided large data sets concerning real-world contacts.

Heterogeneity of contact patterns is recognized as an important feature for realistic modeling of many epidemics, e.g. contact heterogeneity has been shown to have a large impact on epidemic thresholds and the final size of epidemics [12,13]. One of the important factors influencing structure of the network (and increasing heterogeneity) is the age of the node [3,14]. This is so because many of the real networks evolve in time, hence the age influences connection creation probability (as well as the probability that connection will be removed) [15,16]. In real life networks, a time factor may also modulate the attachment probability, therefore in many models of evolving networks [17], especially growing networks, it is assumed that the connection probability is described by a function of connectivity and age of existing nodes. This is a more general approach and allows getting deeper insight in critical behavior of networks, e.g. phase transition from

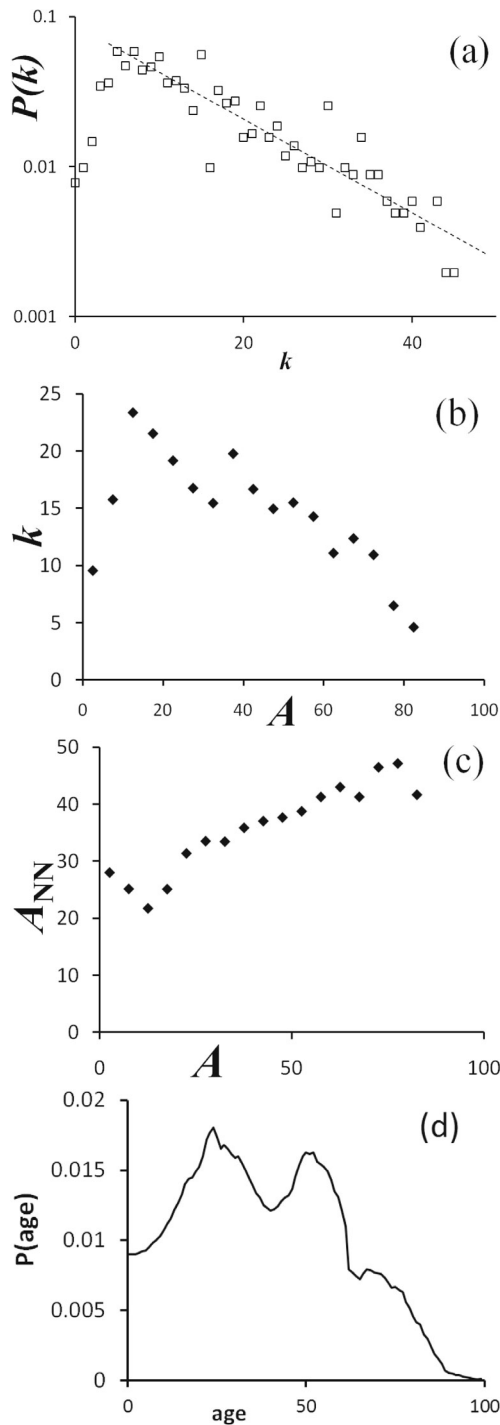
small-world network to scale-free network was investigated and phase diagram was obtained [18].

The aim of this study is to investigate droplet-/airborne infections, which require physical closeness between the infected and susceptible individual, therefore the social network model is based on observed one day contact patterns, including both types of physical contacts reported in the study – conversations and touch. To construct the social network model we incorporated data from contact survey conducted in Poland in the framework of POLYMOD [11,13]. Quota sampling was employed taking into consideration population distribution by age, sex, region and type of residence (rural, city < 100 000 and city > 100 000). The participants were recruited by trained interviewers visiting random households and each of participants was randomly assigned one day when they had to record all contacts. A contact was defined as a conversation in physical presence or a touch. Participants were asked to fill the total duration of all contact episodes with each contact person during the assigned day, as well as location of contacts and usual frequency of contacts with that person. Detailed description of the study is provided elsewhere [11].

The degree distribution of the social network is presented in Figure 1a. Initially total degree distribution increases and has a maximum for  $k = 5$ . However for large enough  $k$  ( $k > 5$ ) the degree distribution has an exponential form  $P(k) \sim e^{-\alpha k}$ . The value of the exponent  $\alpha$  equals 0.07.

We have found that an important factor significantly influencing the structure of a social network is an age  $A$  of each individual. The connectivity of an individual depends on its age (see Fig. 1b). Maximum connectivity is observed for teenagers and decreases (approximately linearly) with age increasing. The average age of nearest neighbors  $A_{NN}$  is highly correlated with the age of an

<sup>a</sup> e-mail: [angra@ciop.pl](mailto:angra@ciop.pl)



**Fig. 1.** The degree distribution (a), the relationship between age and degree (b) and the relationship between age and average age of nearest neighbors (c). Demographic structure in Poland (d), data source – Central Statistical Office.

individual (see Fig. 1c). For an age greater than 20 years the  $A_{NN}$  increases approximately linearly with age of an individual increasing. It should be noted that similar results we have found in on-line social networks [19]. The demographic structure in Poland is shown in Figure 1d.

## 2 The model

The second aim of our work is to investigate numerically the influence of age-age and age-degree correlations on the process of epidemic spreading in a social network. Therefore the following procedure of network creation was used. For a chosen number of individuals  $N$ , at first the age  $a$  of each an individual is chosen according to the demographic structure (individuals are binned in 5 years age groups). The number of connections  $k$  and age of neighbors of an individual is drawn from experimental distribution, i.e.  $i$ th individual initially has  $k_i$  free connections and each connection has assigned desired age of a neighbor.

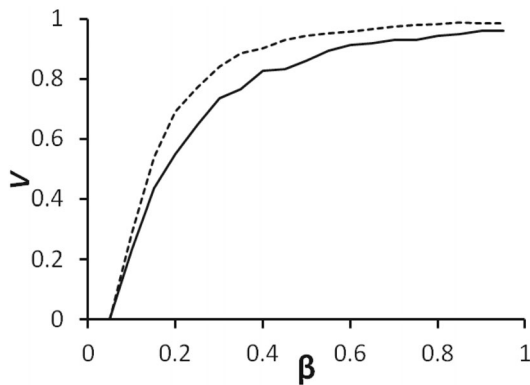
Initially all individuals are not connected. Next, connections between individuals are created. However, each pair of individuals can be connected only once, and a new connection is added to the  $i$ th individual only when its actual number of connections is smaller than the value  $k_i$ . Because we take into account age of individuals, a pair of nodes  $(i, j)$  can be connected only if  $i$ th individual has free connection to the individual with age  $a_j$ , and  $j$ th individual has free connection to the individual with age  $a_i$ . After creating a connection between individuals  $i$  and  $j$  additionally “clustering” procedure is applied. This means that, if possible a new connection is created between the  $i$ th individual and an individual randomly selected from available (i.e. having at least one free connection to which proper age is assigned) neighbors of the  $j$ th individual. Iteration of this procedure allows to obtain the desired distribution of connectivity (actual number of connections of a node equals the value which was drawn from desired distribution). Such a procedure allows us to create the network which has properties typical for social networks, e.g. small average shortest path, large clustering (0.2), assortative mixing [3,20]. Moreover we have taken into account age-age and degree-age correlations.

In the literature there are many models of epidemic spreading with different mechanisms of contagion [21–23]. However, to understand better the influence of human dynamics on the process of spreading, we used a simple SIR model [24,25]. In our model, each individual is in one of three permitted states: healthy and susceptible ( $S$ ), ill ( $I$ ), healthy and unsusceptible or isolated from the rest of the population ( $R$ ).

In SIR models based on differential equations it is often assumed that an increase in the number of ill individuals  $N_I$  is proportional to  $N_I$  [25,26]. We assume that the probability of an infection of an ill individual equals  $\beta$  (in other words  $\beta$  is probability of infection per day of contact). The probability of transition between states  $I$  and  $R$  equals 1. In our model  $\beta$  does not depend on age, however in the case of some diseases such dependency may occur.

## 3 Results

Computations were performed for the initial conditions with one ill ( $I$ ) and randomly located individual and the rest of the population healthy and susceptible ( $S$ ). Synchronous dynamics were used with the assumption that



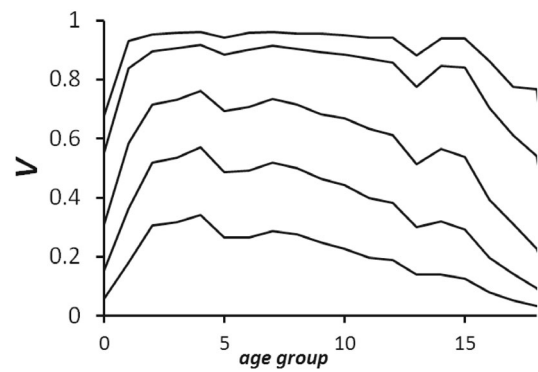
**Fig. 2.** Influence of the parameter  $\beta$  on the average magnitude of the epidemic,  $V$  for different types of networks (network with age structure – solid line, network without age structure – dashed line). Results were averaged over  $10^3$  independent simulations. The size of the network:  $10^5$  individuals.

individuals can change their state only once in each time step. To investigate the dynamics of the spreading process and the magnitude of an epidemic, we introduced two observables: the time  $t_{max}$  when the maximal number of ill individuals is reached and the magnitude of the epidemic,  $V$ , defined as the relative number of individuals who had the disease during the epidemic.

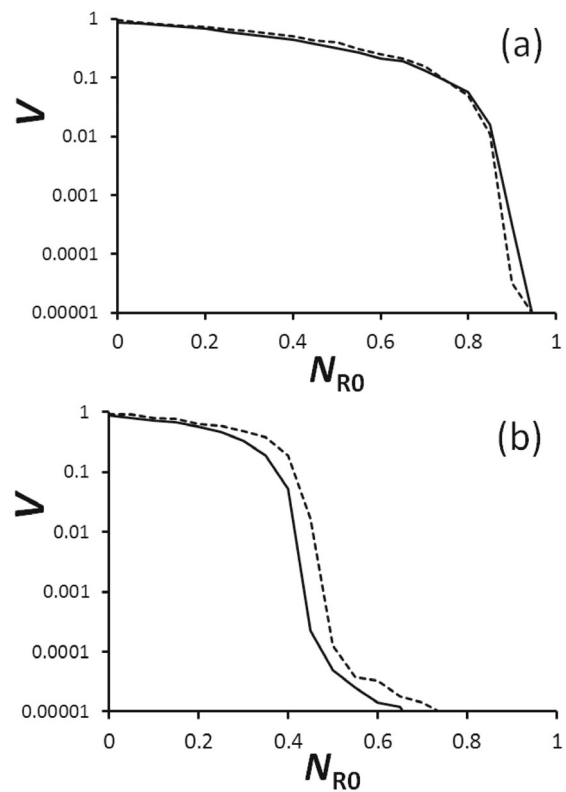
Computations were performed for initial conditions with one ill ( $I$ ) randomly located individual and the rest of the population healthy and susceptible ( $S$ ). To investigate the dynamics of the spreading process and the magnitude of an epidemic, we introduced one observable: the magnitude of the epidemic (i.e. final attack rate),  $V$ , defined as the proportion of individuals who had the disease during the epidemic.

The relation between the control parameter describing a disease and  $V$  is shown in Figure 2. To investigate the influence age-age and degree-age correlations computations for network with the same degree distributions but without age structure were performed. It occurs that the age structure decreases the magnitude of epidemic. This is so because the epidemic spreads mainly among individuals with large number of connections and number of connections is highly correlated with age (cf. Fig. 1b). Hence, the relative number of infected individuals depends on the age (Fig. 3). Note that in the network age-age correlations are observed also – most of connections are between individuals with similar age (cf. Fig. 1b). Therefore older individuals (and small children), who has small number of connections, are very rarely connected to individuals with large number of connections. The probability that they will be infected is relatively small, especially for small values of  $\beta$ . With an increase in  $\beta$ , there is an increase in probability that infection will spread from age groups with high number of connections to the age groups with small number of connections. As a result the curve on Figure 3 becomes more flat – the influence of age structure on the process of epidemic spreading decreases.

It should be noted that such properties of a network influence effectiveness of control measures like mass



**Fig. 3.** Relative number of infected individuals as a function of age group for different values of  $\beta$  (0.05, 0.1, 0.15, 0.25, 0.5 and 0.75; respectively from bottom to top). Individuals are binned in five years age groups. Results were averaged over  $10^3$  independent simulations. The size of the network:  $10^5$  individuals.



**Fig. 4.** The relationship between  $V$  and the number of preventively vaccinated individuals  $N_{RO}$  for  $\beta = 0.5$  and different types of networks (with age structure – solid line, without age structure – dashed line), (a) random vaccinations, (b) target vaccination. Results were averaged over  $10^3$  independent simulations. The size of the network:  $10^5$  individuals.

immunizations campaigns [27]. Figure 4 shows the relationship between magnitude of epidemic  $V$  and relative number of preventively vaccinated individuals  $N_{RO}$  (i.e. the relative number of individuals who are in the state  $R$  in time  $t = 0$ ) for  $\beta = 0.5$ . With an increase in the number of preventively vaccinated individuals  $N_{RO}$ , there is a decrease in the rate of spreading of infection.

This is so because an epidemic cannot spread freely in the presence of vaccinated individuals. However, for critical value  $N_{R0} = N_{RC}$  there is an abrupt decrease in the magnitude of epidemic  $V$ : the epidemic is suppressed. This phenomenon was described in vaccinated populations for many diseases and is known as herd immunity [28]. It can be regarded as a phase transition. Such phase transitions are observed in percolating systems [26,29–31]. It is visible that in the case of both types of networks (with or without age structure) the magnitude of epidemic is similar. It means that in the case of random immunization the herd immunity level ( $N_{RC}$ ) is not affected by age structure of social network.

Simulating effects of immunization targeted to individuals with the largest contact networks the herd immunity level ( $N_{RC}$ ) is generally lower than for random vaccination (see Fig. 4b), implying greater efficiency of this targeted approach. It is also visible that in the case of target vaccinations the critical value  $N_{RC}$  is lower if age structure of social network is taken into account. This is so because most of individuals with large number of connections are only in a few age groups and most of connections of these individuals are pointing to individuals from these age groups. Immunizations of these individuals give much better results because the probability that they will be infected is larger than in other age groups (cf. Fig. 3). For large values of  $\beta > 0.5$  the differences between results for different networks (with and without age structure) vanishes, because the epidemic can spread freely also among individuals in age groups with relative low number of connections (see Figs. 1b and 3).

The SIR model (where the probability of infection with  $k^I$  ill neighbors equals  $\beta k^I$ ;  $\beta$  is the microscopic spreading (infection) rate, and infected individuals decay into the removed class at the rate  $\gamma = 1$ ) shows that the expression for the critical threshold is a function of the moments of the degree distribution  $\beta_C = \frac{\langle k \rangle}{\langle k^2 \rangle}$ . If the value of  $\beta$  is above  $\beta_C$ , the disease spreads and infects a finite fraction of the population. On the other hand, when  $\beta$  is below the threshold, the total number of infected individuals is infinitesimally small in the limit of very large populations. In networks with a strongly fluctuating degree distribution, the epidemic threshold approaches zero for increasing sizes of networks. Such behaviour of percolating systems is especially clearly visible in the case of two dimensional systems. There are many mathematical models describing real world scenarios, e.g. models of epidemics of plant diseases or forest fire models. In colonized (saprotrophic) as well as parasitic pathogen fungal systems in soil activities occur practically in two-dimensional spaces (soil layers). For a review see e.g. [32]. It should be noted that in the case of finite two dimensional systems it is relatively easily to distinguish boundaries. In the case of complex networks it is much more difficult (however some kind of boundary is formed by nodes with only one connection, because they cannot pass the disease further). Therefore it is in some sense more difficult to get rid of finite size effect, e.g. by assuming periodic boundary conditions. It is interesting that the process of

a disease spreading is similar to random walk of carrier in a network (however this is constrained walk, only nodes in the state  $S$  are accessible). The problem of wandering at random in an lattice (or network) finds applications in virtually all sciences [32].

## 4 Conclusions

It was shown that age structure, i.e. correlations between age and degree and correlations between age of node and age of its neighbors, observed in real social networks influences the process of epidemic spreading. The results of numerical calculations indicate that age structure decreases the magnitude of epidemic and increases the effectiveness of target vaccination. Therefore age structure should be taken into account in order to build more plausible models of epidemic spreading.

Scientific work financed from the budget for science by the Ministry of Science and Higher Education within the scope of the Iuventus Plus program.

## References

1. R. Albert, A.-L. Barabási, *Rev. Mod. Phys.* **74**, 47 (2002)
2. S.N. Dorogovtsev, J.F.F. Mendes, *Evolution of Networks* (Oxford University Press, 2004)
3. M.E.J. Newman, *Phys. Rev. E* **67**, 026126 (2003)
4. J.-J. Cheng, Y. Liu, B. Shen, W.-G. Yuan, *Eur. Phys. J. B* **86**, 29 (2013)
5. A. Czaplicka, J. Hołyst, P. Słot, *Eur. Phys. J. Special Topics* **222**, 1335 (2013)
6. G. Abramson, M.F. Laguna, J.R. Iglesias, *Eur. Phys. J. B* **86**, 202 (2013)
7. M.T.Z. Ruan, Z. Liu, *Eur. Phys. J. B* **86**, 149 (2013)
8. H. Ebel, L.I. Mielsch, S. Bornholdt, *Phys. Rev. E* **66**, 035103(R) (2002)
9. A. Grabowski, N. Kruszewska, R.A. Kosiński, *Eur. Phys. J. B* **66**, 107 (2008)
10. A. Grabowski, N. Kruszewska, R.A. Kosiński, *Phys. Rev. E* **78**, 066110 (2008)
11. J. Mossong, N. Hens, M. Jit, P. Beutels, K. Auranen, R. Mikolajczyk, M. Massari, S. Salmaso, G.S. Tomba, J. Wallinga, J. Heijne, M. Sadkowska-Todys, M. Rosinska, W. John Edmunds, *PLoS Med.* **5**, e74 (2008)
12. E. Volza, *Eur. Phys. J. B* **63**, 381 (2008)
13. A. Grabowski, M. Rosińska, *Eur. Phys. J. B* **85**, 248 (2012)
14. H. Zhu, X. Wang, J.-Y. Zhu, *Phys. Rev. E* **68**, 056121 (2003)
15. S.N. Dorogovtsev, J.F.F. Mendes, *Phys. Rev. E* **62**, 1842 (2000)
16. J.A. Hołyst, A. Fronczak, P. Fronczak, *Phys. Rev. E* **70**, 046119 (2004)
17. D. Makowiec, *Eur. Phys. J. B* **48**, 547 (2005)
18. K.B. Hajra, P. Sen, *Phys. Rev. E* **70**, 056103 (2004)
19. A. Grabowski, R.A. Kosiński, *Acta Physica Polonica B* **39**, 1291 (2008)
20. S.N. Soffer, A. Vazquez, *Phys. Rev. E* **71**, 057101 (2005)

21. M. Boguna, R. Pastor-Satorras, Phys. Rev. E **66**, 047104 (2002)
22. P.S. Dodds, D.J. Watts, Phys. Rev. Lett. **92**, 218701 (2004)
23. D. Centola, V.M. Eguiluz, M.W. Macy, Physica A **374**, 449 (2007)
24. W.O. Kermack, A.G. McKendrick, Proc. Roy. Soc. Lond. A **115**, 700 (1927)
25. N.T.J. Bailey, *The Mathematical Theory of Infectious Diseases* (Springer, 1993)
26. Y. Moreno, R. Pastor-Satorras, A. Vespignani, Eur. Phys. J. B **26**, 521 (2002)
27. L.B. Shaw, I.B. Schwartz, Phys. Rev. E **81**, 046120 (2010)
28. P. Fine, Epidemiol. Rev. **15**, 265 (1993)
29. A. Allard, P.-A. Noel, L.J. Dubé, Phys. Rev. E **79**, 036113 (2009)
30. S.N. Dorogovtsev, A.V. Goltsev, Rev. Mod. Phys. **80**, 1275 (2008)
31. L.M. Sander, C.P. Warren, I.M. Sokolov, C. Simon, J. Koopman, Math. Biosci. **180**, 293 (2002)
32. A. Gadomski, J. Siódmak, J.J. Uher, Acta Physica Polonica B **32**, 1541 (2001)

**Open Access** This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.