

# Effective vaccination strategies for realistic social networks<sup>☆</sup>

Joel C. Miller<sup>a,b,\*</sup>, James M. Hyman<sup>a</sup>

<sup>a</sup>Mathematical Modeling & Analysis Group MS B284, Los Alamos National Laboratory, Los Alamos, NM 87545, USA

<sup>b</sup>Center for Nonlinear Studies, Los Alamos National Laboratory, USA

Available online 14 September 2007

---

## Abstract

We consider the effectiveness of targeted vaccination at preventing the spread of infectious disease in a realistic social network. We compare vaccination strategies based on no information (random vaccination) to complete information (PageRank) about the network. The most effective strategy we find is to vaccinate those people with the most unvaccinated contacts. However, this strategy requires considerable information and computational effort which may not be practical. The next best strategies vaccinate people with many contacts who in turn have few contacts.

Published by Elsevier B.V.

*Keywords:* Epidemic; Network; Vaccination strategy; PageRank

---

## 1. Introduction

The spread of infectious disease continues to have significant impact on society. In the first half of the twentieth century, Reed and Frost (as described by Ref. [1]) and Kermack and McKendrick [2] developed simple compartmental models of epidemic spread. Over time, these models were modified to more accurately represent social structures or other details influencing disease spread [3]. Increasingly sophisticated recent models consider the population at an individual rather than continuum scale. Some such models are analytically tractable [4–6], while others are agent-based simulations with complex interactions [7–11].

We consider a susceptible-infected-recovered (SIR) model [3] of disease spread on a network. The outbreak begins with a single infection, the *index case*. The remaining nodes begin susceptible and may become infected if in contact with an infected node. An infected node infects each of its susceptible neighbors independently with probability  $T$ , the *transmissibility*, and then recovers (with immunity). Outbreaks are classified as epidemics or non-epidemics depending on whether the outbreak dies out at early time. The probability  $\mathcal{P}$  of an epidemic is the probability that a random index case sparks an epidemic, while the attack rate  $\mathcal{A}$  is the fraction infected during the course of the epidemic. With constant  $T$  the problem can be mapped to bond percolation so the outbreak size is the component size of the index case. It follows that  $\mathcal{P} = \mathcal{A}$  [4–6].

---

<sup>☆</sup> Los Alamos Report LA-UR-06-8539.

\*Corresponding author. Mathematical Modeling & Analysis Group MS B284, Los Alamos National Laboratory, Los Alamos, NM 87545, USA.

E-mail address: [jomiller@lanl.gov](mailto:jomiller@lanl.gov) (J.C. Miller).

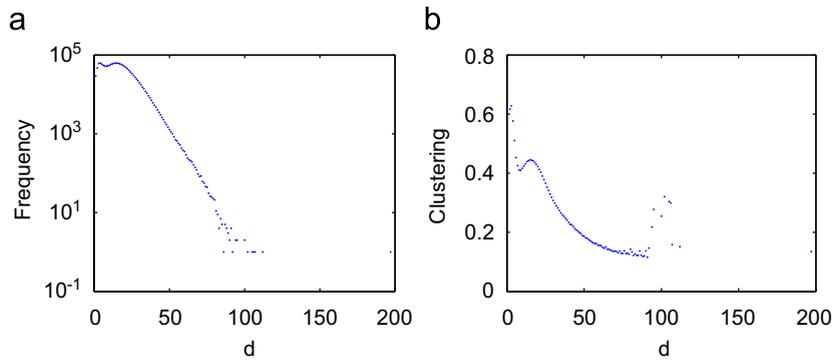


Fig. 1. Properties of the agent-based EpiSimS network modeling the contacts of people in Portland, OR [7–9]. The degree distribution has an exponential tail, except for one node with degree close to 200. The clustering coefficient is, on average, much higher for lower degree nodes than higher degree nodes. (a) Distribution of number of contacts between people in a day as measured by the degree  $d$  of the nodes. (b) Average clustering of the EpiSimS network as a function of degree.

Our network comes from an EpiSimS [7–9] simulation of Portland, Oregon where a virtual population travels through the city during a day. We represent each person by a node. Nodes are connected if the people visit a location simultaneously. Behavior does not change, so contacts remain fixed. We disregard duration or closeness of contact. The resulting static, unweighted network has approximately 1.6 million nodes, with average degree of about 16. The clustering coefficient (defined by  $c_i = n_{\Delta,i}/\binom{d_i}{2}$ , the number of triangles  $n_{\Delta,i}$  including  $i$  divided by the possible number given the degree  $d_i$  [12]) is significant. Some network properties appear in Fig. 1.

We measure a strategy’s effectiveness by its ability to prevent disease spread (in contrast, the current strategy for influenza vaccination targets those most at risk of complications if infected). Our goal is to identify effective strategies for controlling epidemics on networks from which guiding principles may be distilled to apply to real-world outbreaks.

## 2. The vaccination strategies

We assume that vaccination occurs before the outbreak and is completely effective. Vaccination makes an individual invisible to the disease, and has the effect of removing nodes from the network.

A *Random Vaccination* strategy vaccinates a fraction of the population randomly, using no knowledge of the network. *Acquaintance Vaccination* [13,14] selects a node randomly and vaccinates one of its neighbors. This requires minimal knowledge of the network, but preferentially selects higher degree nodes. If we have information about all degrees, we may rank nodes by degree, and use *Degree Vaccination* [15] to vaccinate nodes by descending degree. When a high-degree node is vaccinated, the effective degree of its neighbors drops. *Degree Vaccination with Dynamic Reranking* recalculates degrees and vaccinates the node with the most unvaccinated neighbors. This requires substantially more information. These four strategies have been considered before, but not on a network with the realistic structure of the EpiSimS network.

We introduce three new strategies. The first uses *PageRank* [16], the algorithm behind the Google web rankings. In this algorithm a random walker travels on the network, moving along a link from the current node with probability  $\delta$  and jumping to a random node with probability  $1 - \delta$ . For web rankings,  $\delta$  is typically close to 1, but we consider arbitrary values  $0 < \delta < 1$ . This algorithm depends on information about the entire network. The weight of a node  $w_i$  may be interpreted as the frequency with which it is visited at large time. The weights are calculated through an eigenvalue problem. We expect a random walker to most often visit nodes which can be infected or create infections along many paths, and so it should identify nodes important to the spread of the disease.

Based on observations from PageRank, we design a new *Neighbor-degree Based Vaccination* strategy (described in Section 3), which uses only information about the degrees of a node’s neighbors. This approach achieves a result comparable to dynamic reranking, but with significantly less computational effort.

The above strategies consider properties that are difficult to measure in real populations and change as the vaccination fraction increases. Our final strategy is easier to measure and is not affected by vaccinating. *Travel Based Vaccination* vaccinates those people who visit the most locations.

### 3. Results

We compare strategies quantitatively using the *mean attack rate*,  $M = \int_0^1 \int_0^1 \mathcal{A} dV dT$ , where  $V$  is the vaccination fraction and  $T$  is the transmissibility of the disease. Smaller  $M$  indicates a more effective average strategy, but does not ensure that the strategy is better for all  $V$  and  $T$ .

The results for the random, acquaintance, degree, and degree with dynamic reranking strategies (Fig. 2) show that additional information about the population can be used to improve the vaccination strategy.  $M$  drops from 0.38 (random) to 0.33 (acquaintance) to 0.29 (degree) and finally 0.25 (degree with dynamic reranking). For larger transmissibility, there is little difference in the strategies until  $V$  is close to the epidemic threshold.

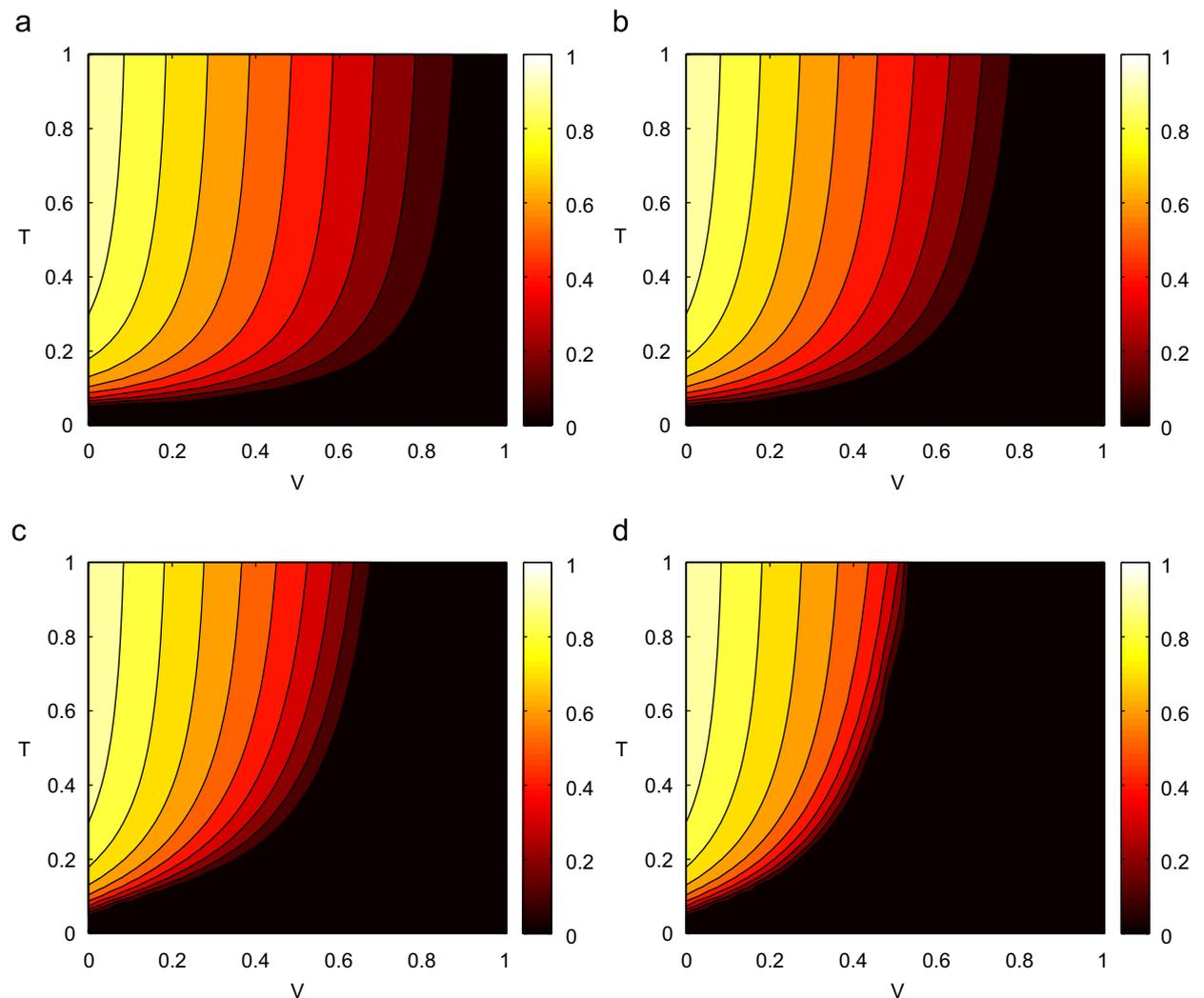


Fig. 2. The attack rate  $\mathcal{A}$  for varying transmissibility  $T$  and vaccination levels  $V$  under different vaccination strategies. Additional information about the contact network can be used to increase the effectiveness of the vaccination strategy. (a) Random:  $M = 0.38$ . (b) Acquaintance:  $M = 0.33$ . (c) Degree:  $M = 0.29$ . (d) Degree with dynamic reranking:  $M = 0.25$ .

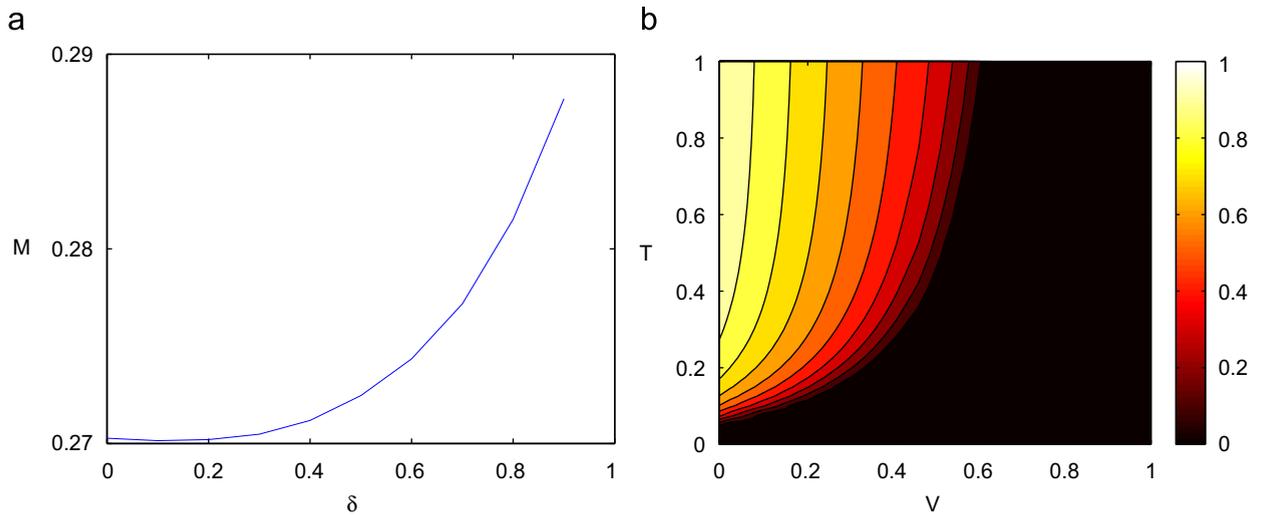


Fig. 3. Results from the vaccination strategy based on PageRank. The results are remarkably insensitive to changes in  $\delta$  at small values. (a) The PageRank algorithm is most effective at small  $\delta$ . (b) Results from vaccinating with  $\delta \rightarrow 0$ .  $M = 0.27$ .

In general PageRank (Fig. 3) performs as well or better than vaccination by degree. We find that the best performance occurs at  $\delta \approx 0.1$ , but the benefits change little for  $\delta < 0.2$ . We see that in the limit  $\delta \rightarrow 0^+$  PageRank outperforms even degree with dynamic reranking for large values of  $T$  and small values of  $V$ , although  $M$  is slightly larger.

Near-optimal results for PageRank are found in the  $\delta \rightarrow 0^+$  limit where we can make analytical progress

$$\begin{aligned}
 w_i &= \frac{1 - \delta}{N} + \delta \sum_{j \in \mathcal{N}(i)} \frac{w_j}{d_j} \\
 &= \frac{1}{N} \left[ 1 - \delta + \delta \left( \sum_{j \in \mathcal{N}(i)} d_j^{-1} \right) + \mathcal{O}(\delta^2) \right],
 \end{aligned}$$

where  $\mathcal{N}(i)$  denotes the neighbors of  $i$ ,  $N$  is the population size, and  $d_j$  is the degree of  $j$ . The difference in weightings occurs at  $\mathcal{O}(\delta)$ . The nodes are ranked based on  $\sum_{j \in \mathcal{N}(i)} 1/d_j$ , so only local information is needed. Nodes with more neighbors have higher weight, but having low degree neighbors increases their weight the most.

We now consider using neighbors' degrees to rank nodes. This requires less information than PageRank, more information than vaccination by degree, and takes less computational effort than vaccination by degree with dynamic reranking.

A simple way to incorporate neighbors' degrees is to rank nodes by the sum of the degrees of the neighboring nodes. However, the PageRank results suggest this will not perform well, and in fact it performs worse than vaccination by degree ( $M = 0.31$ ). This occurs because vaccinating a node with many high degree neighbors gives little protection to those neighbors. Either they will be vaccinated as well, or they are likely to get infected from another source.

This motivates a different strategy: consider an arbitrary edge from  $i$  to  $j$ . Let  $r$  denote the probability that this edge would transmit infection to  $j$ , without knowledge of whether  $i$  is infected. Consequently,  $r$  is  $T$  times the probability that  $i$  becomes infected in a network with  $j$  removed. The resulting value of  $r$  for an edge is difficult to calculate and may be correlated with other edges. For simplicity, we assume that  $r$  is edge-independent and that transmitting infection along one edge is independent of all others.

If  $j$  has  $d_j$  neighbors, these assumptions mean  $j$  becomes infected with probability  $1 - (1 - r)^{d_j}$ . If we vaccinate one neighbor  $i$ , then the risk for  $j$  is reduced by  $r(1 - r)^{d_j - 1}$ . As a crude measure of the amount of

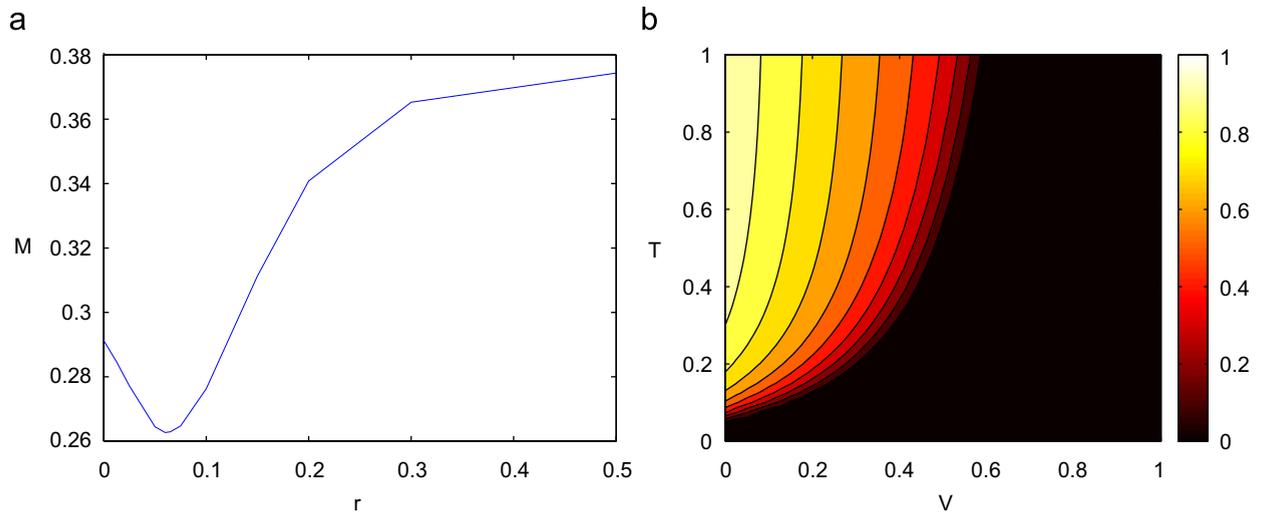


Fig. 4. Neighbor-degree based vaccination: results of vaccinating based on  $w_i = \sum_j (1 - r)^{d_j}$ . The results are sensitive to the choice of  $r$ , and the optimal value is likely to be network-dependent. (a) Neighbor-degree:  $M(r)$ . The optimal value of  $r$  is about 0.06. (b) Neighbor-degree with  $r = 0.06$ :  $M = 0.26$ .

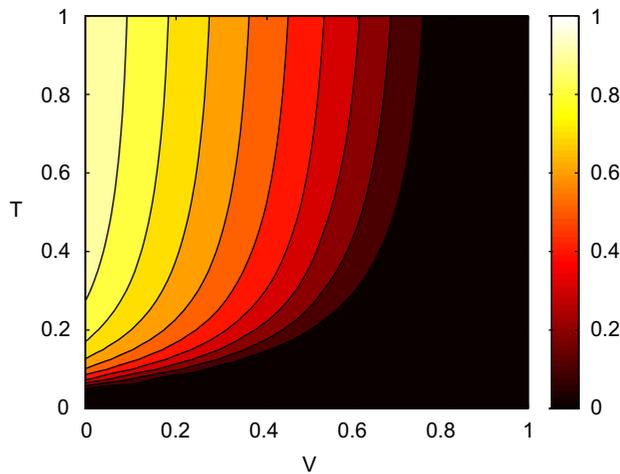


Fig. 5. Travel based vaccination.  $M = 0.33$ , comparable to acquaintance based vaccination.

protection vaccinating  $i$  provides its neighbors, we sum  $(1 - r)^{d_j - 1}$  over  $j \in \mathcal{N}(i)$ :

$$w_i = \sum_{j \in \mathcal{N}(i)} (1 - r)^{d_j - 1}.$$

We then vaccinate nodes based on their ranking from  $w_i$ .

Fig. 4 shows results. For  $r \ll 1$ ,  $w_i = d_i + \mathcal{O}(r)$ , so the  $r \rightarrow 0$  limit gives the same result as vaccination by degree. As  $r$  increases,  $M$  initially decreases, and at  $r = 0.06$  achieves its minimum<sup>2</sup>  $M = 0.26$ . At larger  $r$ , the result is comparable to random vaccination.

We finally consider vaccinating those people who visit the most locations (Fig. 5). This performs comparably to acquaintance based vaccination, with  $M = 0.33$ . It is an improvement over random vaccination, and requires very little effort to rank in a real-world situation.

<sup>2</sup>Although  $T = 0.06$  is also approximately the epidemic threshold in the unvaccinated case, this appears to be coincidence.

#### 4. Conclusions

We find that information about the contact networks structure can be used to significantly improve vaccination strategies. Unfortunately this information is rarely available for a real population. This suggests a need for guiding principles which may be used to design effective strategies with minimal information.

We have shown that a strategy based on the number of locations an individual visits is better than random vaccination. This ranking is relatively easy to measure, and (unlike degree) the ranking does not change as nodes are vaccinated. However, we find that in general our other strategies outperform travel based vaccination.

An obvious guiding principle which we find to perform well is to vaccinate those people (such as bus drivers or school teachers) who have the most contacts. However, a secondary principle we have identified is that the benefit gained by vaccinating someone with many contacts is reduced if his or her contacts also have many contacts. Therefore, consideration should be given to vaccinating those people who have many contacts with largely solitary populations, such as people who visit many residences during their work.

#### Acknowledgments

DOE at LANL supported this work under Contract DE-AC52-06NA25396 and the DOE Office of ASCR program in Applied Mathematical Sciences.

#### References

- [1] H. Abbey, An examination of the Reed-Frost theory of epidemics, *Hum. Biol.* 24 (3) (1952) 201–233.
- [2] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, *R. Soc. London Proc. Ser. A* 115 (1927) 700–721.
- [3] R.M. Anderson, R.M. May, *Infectious Diseases of Humans*, Oxford Science Publications, 1991.
- [4] M.E.J. Newman, Spread of epidemic disease on networks, *Phys. Rev. E* 66 (1) (2002) 16128.
- [5] L.A. Meyers, Contact network epidemiology: Bond percolation applied to infectious disease prediction and control, *Bull. Am. Math. Soc.* 44 (1) (2007) 63–86.
- [6] J.C. Miller, Epidemic size and probability in populations with heterogeneous infectivity and susceptibility, *Phys. Rev. E* 76 (1) (2007) 010101.
- [7] C.L. Barrett, S.G. Eubank, J.P. Smith, If smallpox strikes Portland., *Sci. Am.* 292 (3) (2005) 42–49.
- [8] S. Del Valle, P.D. Stroud, J.P. Smith, S.M. Mniszewski, J.M. Riese, S.J. Sydorik, D.A. Kubicek, EpiSimS: epidemic simulation system, Technical Report LAUR-06-6714, Los Alamos National Laboratory, 2006.
- [9] S. Eubank, H. Guclu, V.S.A. Kumar, M.V. Marathe, A. Srinivasan, Z. Toroczkai, N. Wang, Modelling disease outbreaks in realistic urban social networks, *Nature* 429 (6988) (2004) 180–184.
- [10] T.C. Germann, K. Kadau, I.M. Longini Jr., C.A. Macken, Mitigation strategies for pandemic influenza in the United States, *Proc. Nat. Acad. Sci. USA* 103 (15) (2006) 5935–5940.
- [11] N.M. Ferguson, D.A.T. Cummings, S. Cauchemez, C. Fraser, S. Riley, A. Meeyai, S. Iamsirithaworn, D.S. Burke, Strategies for containing an emerging influenza pandemic in Southeast Asia, *Nature* 437 (7056) (2005) 209–214.
- [12] D. Watts, S. Strogatz, Collective dynamics of ‘small world’ networks, *Nature* 393 (6684) (1998) 409–410.
- [13] N. Madar, T. Kalisky, R. Cohen, D. ben Avraham, S. Havlin, Immunization and epidemic dynamics in complex networks, *Eur. Phys. J. B* 38 (2004) 269–276.
- [14] R. Cohen, S. Havlin, D. ben Avraham, Efficient immunization strategies for computer networks and populations, *Phys. Rev. Lett.* 91 (24) (2003) 247901.
- [15] R. Pastor-Satorras, A. Vespignani, Immunization of complex networks, *Phys. Rev. E* 65 (3) (2002) 036104.
- [16] L. Page, S. Brin, R. Motwani, T. Winograd, The Pagerank Citation Ranking: Bringing Order to the Web, 1998.