The Effect of Disaggregation on Infection Spreading in a social network: *More* may not be *Merrier*

Varsha S. Kulkarni

School of Informatics, Indiana University, Bloomington, IN 47408, USA
and
Indian Institute of Technology, Kanpur 208016, U.P, India
(Email: vskulkar@umail.iu.edu)

In the past, researchers have applied epidemiological models in the framework of complex social networks to explain dynamics of contagious diseases therein. This paper analyzes the same in the framework of a disaggregated network *viz.* the world as a 'network' disaggregated into diverse countries as 'clusters'. First, the random matrix analysis carried out on WHO data on 'Hepatitis/other viral' infections of 13 countries for period 1985-2008, rules out systematic correlations in prevalence between countries or a 'common influence'. Next the effect of disaggregation is discussed in terms of how the connectivity and population diversities affect epidemic threshold (λ_c). Connectivity, k is interpreted as a cultural determinant and estimated from 'ethnic fractionalization' data. Statistical analysis is used to estimate the model parameters- spreading rate, curing rate -and to examine the effect of population on prevalence. Agent-based modeling (abm) is employed to investigate the roles of k and population N. The analyses reveal that (i) the patterns of infection spreading from the model match well with those observed (ii) the differences in disease patterns of countries revealed by the disaggregated framework are actually a reflection of the heterogeneity across the countries. A less perfect power law fit is obtained for distribution of k. (iii) λ_c depends on not just k (as established in previous studies that higher k adversely affects epidemics) but actually on its interplay with population of the region/country. It turns out that isolation may not be as effective in lowering infection spread in crowded communities as in sparsely populated regions (iv) a small but positive significant effect of population on prevalence got from statistical analysis confirms the above (v) the result of abm's for 3 configurations – [N, k], [N, $k/4$] and [3N, $k/4$] is consistent with above findings. Thus disaggregated modeling framework explains disease dynamics in spatially separated regions in terms of their cultural and demographic aspects and the variations revealed are interesting.

1. Introduction

Spreading dynamics of infectious diseases has long captured the attention of researchers across various disciplines including physics and mathematics [1-2]. Both stochastic and deterministic epidemic models have been proposed to explain the outbreak and spread of contagious diseases. Some studies have focused on epidemic models based on applications of statistical mechanics in the framework of a complex social network [3]. These models enable us to understand the disease dynamics and make predictions by assuming a suitable network structure with certain connectivity and other properties which play a crucial role in prediction and understanding. It has been found at large that higher connectivity in any network makes it easier to spread infection from infected to susceptible individuals (nodes) due to their 'interactions'. However the role of 'activity' or 'interaction' in disease transmission studied does not always address issues of disease aetiology [4], which may depend on numerous factors such as spatial, cultural, demographic aspects. These factors are often omitted from mathematical models and more stress is laid on a suitable network structure to know when and how people get infected therein, rather than what happens when the network becomes disaggregated based on spatial, cultural and demographic differences. Geographic models may help us understand disease dynamics given the local structure of barriers and etc [5]. A common strategy since the time of Black Plague epidemics (14^{th} century) to control geographic spread of diseases has been to inhibit movement, to isolate or "quarantine". This way the probability of geographical disease may be negligible. As cited in [5] Gould (1989) said " Ignoring spatial dimensions of an epidemic is like predicting the time of eclipse but being unable to tell people where they can see it"

Further WHO policies on disease control are based on the understanding of dynamics of disease transmission in the world as whole network. This is likely to work effectively by considering a disaggregated framework of the world that will lead to country specific knowledge. The differences in culture and populations of the spatially separated countries/regions lead to differences in patterns of infection spread over time. It is in terms of these factors that differences in patterns of disease reflect heterogeneity across the groups. One of the aims here is to see the extent of heterogeneity across a subset of countries. The paper mainly investigates the effect of disaggregation of the 'world' as a network into different (some) countries on disease dynamics therein, in terms of their cultural (ethno-linguistic and religious) and demographic (populations) aspects.

This paper is organized as follows. Section 2 describes the data analyzed. The basic modeling framework and its implications are discussed in section 3. Random matrix analysis is shown section 4. Sections 5,6 describe the cultural and demographic aspects and estimate the model parameters statistically. The effects of these aspects are shown by agent based modeling in section 7. The paper concludes with a discussion of key findings in section 8.

2. Data

The paper uses data provided by WHO on yearly incidence, prevalence of common infections 'Hepatitis, other viral' for the time period 1985-2008 for 13 countries namely- Azerbaijan, Belarus, Georgia, Slovenia, Bosnia-Herzegovinia, Norway, Ireland, Denmark, Netherlands, Czechoslovakia, Iceland, Lithuania, Slovakia. They have been selected mainly according to the availability of data. The populations and ethnic fractionalization index data for these countries is used. The population growth rate for the countries considered is very small for the period of analysis.

3. Model Dynamics

3.1 Modeling Framework

The dynamics of infection spreading through a network has been studied using the standard susceptible – infected –susceptible (SIS) model [3]. This 'mean field' model described below applies to a wide class of networks that exhibit bounded connectivity fluctuations. It is got by computing the change in number of infected individuals (or nodes) $n_i(t)$ at a time t given as:

$\Delta n_i(t) = n_i(t+\Delta t) - n_i(t) =$ - # of nodes Recovered in Δt + # of nodes Newly Infected in Δt

$$= - h\, n_i(t)\, \Delta t + \lambda k\, p(t)\, [N - n_i(t)]\, \Delta t$$
$$\text{(i)} \qquad\qquad \text{(ii)}$$

$$\frac{\Delta p}{\Delta t} = -hp(t) + \lambda k\, p(t)\, [1 - p(t)] \qquad (1a)$$

when $\Delta t \to 0$ $\quad \frac{dp}{dt} = -hp(t) + \lambda k\, p(t)\, [1 - p(t)] \qquad (1b)$

where N = total number of nodes in the network;
h = probability per unit time an infected node is cured (Curing rate);
λ = probability per unit time per link that a healthy node is infected (Spreading rate);
$p(t)$ = fraction of infected nodes in the network at time $t = n_i(t)/ N$
k = average connectivity

Equation (1) can be divided into 2 parts because change in prevalence depends on 2 factors – (i) Recovery (R) and (ii) Incidence (I). These two can be thought of as outflux and influx respectively and prevalence is the net stock in the reservoir. Here of course we assume the recovered individuals are never fully immune. This SIS framework is apt for the kind of infection being studied here.
At the steady state of infection transmission (and taking h=1), dp/dt =0 and from (1) $p = \frac{\lambda - \lambda_c}{\lambda}$

implies $p = 0$ if $\lambda < \lambda_c$
and $p \sim \lambda - \lambda_c$ if $\lambda \geq \lambda_c$

where $\lambda_c = 1/k$ is the epidemic threshold which specifies the phase transition of infection spread (from non-endemic to endemic state).

3.2 Departure

Heterogeneity is introduced in the social network by dividing into almost disjoint clusters. The nodes within a cluster are uniformly connected with some average connectivity but connectivity between the clusters is sparse. In other words a community structure is seen in the social network which is disaggregated into homogenous groups. Heterogeneity may be observed across the groups based on differences in average connectivity, size etc. It may be noted that these clusters differ in terms of cultural and demographic aspects and may be considered spatially separated such that the probability of infection crossing barriers from one region/cluster to another is almost zero. In this set up equation (1) for dynamics of j^{th} group changes to:

$$\frac{\Delta p_j}{\Delta t} = -h_j p_j(t) + \lambda_j k_j p_j(t) [s_j - p_j(t)] \quad (2a)$$

$$\frac{dp_j}{dt} = -h_j p_j(t) + \lambda_j k_j p_j(t) [s_j - p_j(t)] \quad (2b)$$

(i) (ii)

where h_j, λ_j, k_j, p_j= curing rate, spreading rate, average connectivity, probability of infection in j^{th} group respectively. And s_j is N_j/N , the relative population of j^{th} group.

In this case steady state infection condition changes as $dp_j/dt = 0$ and from (2) $p_j = \frac{\lambda_j - \lambda_{cj}}{\lambda_j}$

implies $p_j = 0$ if $\lambda_j < \lambda_{cj}$
and $p_j \geq 0$ if $\lambda_j \geq \lambda_{cj}$

where $\lambda_{cj} = \frac{h_j}{k_j s_j} = \frac{1}{k_j s_j}$ taking $h_j = 1$

Some studies like [6] have focused on preventing epidemics by using community structure models to find critical immunization coverage for different communities. However this disaggregation analysis is intended to focus on specific groups by assuming that disease transmission across groups is negligible. The effect of disaggregation can be seen by comparing the epidemic thresholds for phase transitions in two cases - λ_c, λ_{cj}. The threshold for the disaggregated case depends on not just the average connectivity but the relative population of the cluster/region/country being considered; in fact it is the interplay of these two factors which determines the threshold.

There could be 2 interpretations of the above mentioned dependence. One is that $k_j s_j$ may be interpreted as k_{eff} –the effective average connectivity of the j^{th} group/network meaning if the population is higher, the

connectivity is bound to be large. However it must be distinguished from the case when average connectivity may not be affected by the population of the group/network. This is where cultural and other socio-epidemiological aspects play a key role in understanding the social structure of the network being considered and hence the epidemics therein. For instance 'ethnic fractionalization' in some social networks tends to be much higher than in others and also people may sometimes tend to isolate themselves especially in times of infection scare. These are some factors that impact *k* irrespective of the population of the cluster. Apart from that there are some latent socio-epidemiological factors which may impact the epidemics even when there are not many direct/immediate links. These factors become more crucial in some cases- for instance when diseases spread because of the individuals' propensity to imitate the lifestyles of their group/social network. Ethnic fractionalization data may be a sort of indicator of all such kinds of direct and indirect linkages that matter.
This means that the 'demographic' or population factor may adversely affect the epidemics of the region/country considered. Some of the very plausible reasons for this are that crowded communities offer low level of hygiene/ resources etc, airborne diseases are easily transmitted in crowded communities despite having lesser number of direct contacts. Thus it is this interplay of connectivity and population of the group that is interesting to compare given the data on different countries in the world. The role of above mentioned cultural and demographic factors is discussed and analyzed in detail in section 5.

4. Correlation Analysis

The framework described above is a network- the world, disaggregated into different clusters- countries. Before moving on to data analysis this section investigates the presence of any explicit correlation or dependence between the prevalence rates of different countries. The model framework employed in this paper considers prevalence dynamics in each country/region as unaffected by any other region/country. With the tools of Random Matrix theory (RMT) and statistics, we try to justify this. RMT approach here is on exactly similar lines as that outlined in studies [7-8] for analyzing financial cross-correlations.

We first construct the correlation matrix of 'proportionate' changes in yearly prevalence for 13 countries over a time period 1985-2008.

$$G_i(t) = \frac{p_i(t+1) - p_i(t)}{p_i(t)} \qquad i = 1,2\ldots..13$$

for i^{th} country at time *t*. *G* is MxT matrix where M=number of countries and T=length of time considered. The correlation matrix C is constructed from entries of *G* for each of the M countries using the standard procedure and $C_{ij} = 1$ implies complete correlation between countries *i,j*. $C_{ij} = 0$ implies no correlation between countries *i,j*; $C_{ij} = -1$ implies complete anti-correlation between countries *i,j*

The properties (eigenvalues and eigenvectors) of C are compared with those of a 'random matrix' R constructed from purely random entries with mean 0 and variance 1 that are mutually uncorrelated. The probability density of eigen-values of R [9] is given by

$$P_{rm}(\lambda) = Q\frac{\sqrt{(\Lambda_+ - \Lambda)(\Lambda - \Lambda_-)}}{2\pi\Lambda},$$

where Q=T/M. RMT applies well in the limit of large M and T and Q>1 (fixed). Λ_+ and Λ_- are maximum and minimum eigenvalues given as $\Lambda_\pm = 1 + 1/Q \pm 2\sqrt{1/Q}$ known as RMT bounds. (However RMT is known to apply well even when these restrictions about size and limiting distribution are relaxed).
The largest e-value, Λ^{13} of C reflects the nature of correlations. If it deviates from RMT limit ($>\Lambda_+$), it contains some systematic information and reflects the non-random character of the interactions. The closer it is to the trace of the C, more information it contains about actual correlations. Here Q =24/13 and

Λ_+ and Λ_- are 3 and 0.07 respectively. The figure 1 below shows that the max(min)imum e-value of C is 2.97 (0.07) just within or at the RMT bounds.

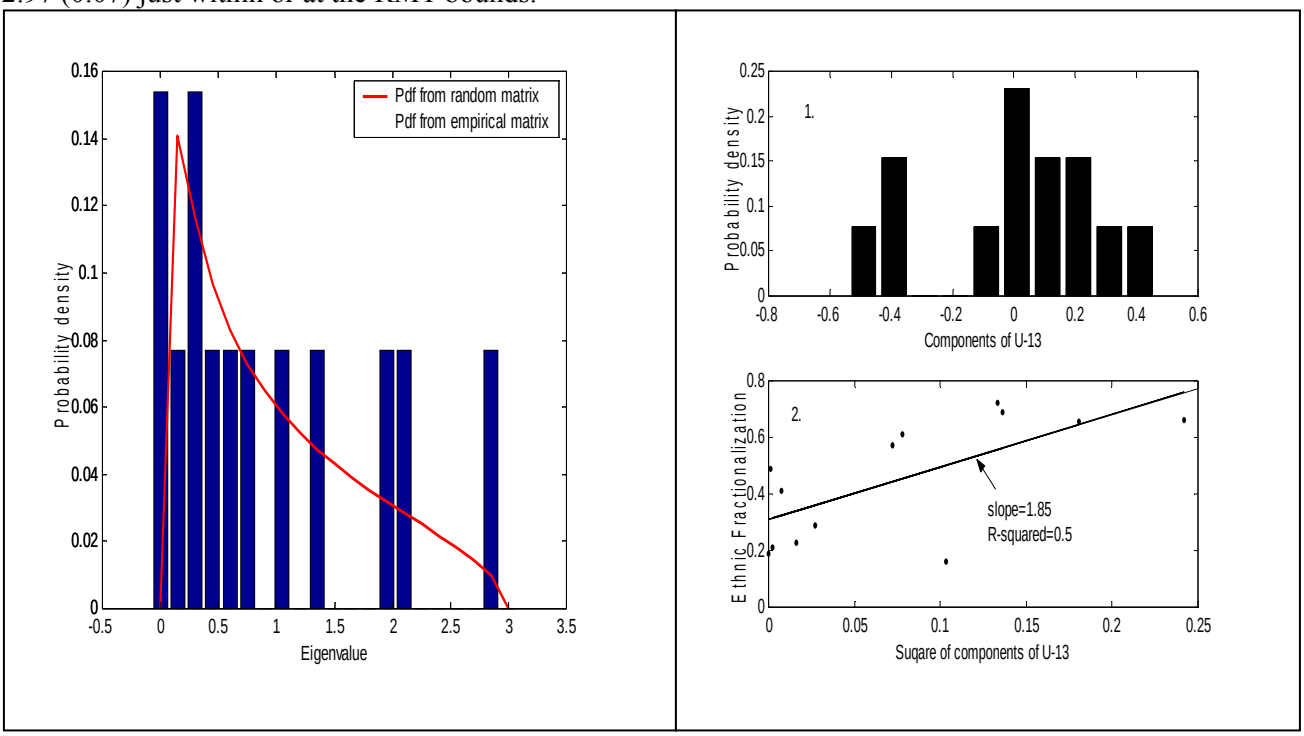

Figure 1 (Left) Comparison of probability density of empirical correlation matrix C (given by bars) and that of random matrix R (shown by solid line)
Figure 2 (Right) (i) (Top) Probability density of components of U^{13} ; (ii) (Bottom) Ethnic fractionalization of all 13 countries versus corresponding contributions to U^{13}

This suggests that there is no evidence of any systematic correlations between countries in this ensemble as far as prevalence is concerned. Earlier work [10] has enunciated the relevance of Λ^{13} along with its eigenvector U^{13} in reflecting the presence or absence of a common influence (shock) on the whole ensemble depending on whether and how much it exceeds the RMT limits. Social scientists have modeled to see if 'exogenous shocks' [11] (e.g. improvements in health technology) affect the health in all countries alike or not. In figure 2(i), the pdf of U^{13} components clearly shows that the countries behave quite independently of each other's or any common influence (shock) as both positive and negative components occur with finite probability. The contribution of every country in the ensemble to the magnitude of correlations (and anti-correlations) reflected by the square of corresponding components of U^{13}.

5. Ethnic Fractionalization, Population and Heterogeneity

The prevalence patterns are different in every country. In the disaggregated modeling framework described above, the epidemics in each country/region, is seen to be dependent on the interplay of 2 factors- connectivity and population. Hence the effect of disaggregation on epidemics is actually a reflection of the heterogeneity across the countries in terms of these. As argued previously, these factors may act independently and are determinants of cultural and demographic aspects respectively. Connectivity in any country/region is largely dependent on propensity of 'herding' to seek social capital [12] which often arises from ethno-linguistic, religious homogeneities.

The term 'ethnic fractionalization' is used by social scientists to refer to the extent of fragmentation occurring due diversity in language, ethnicity and religion. This is a cultural aspect and affects disease dynamics. The ethnic fractionalization index is commonly [13] defined as the probability that 2 randomly selected individuals do not belong to same group. For the j^{th} country of size N_j divided into M_j groups and m_g being population of g^{th} group, the index is given as :

$$EF_j = 1 - \sum_{g=1}^{M_j} (m_g / N_j)^2$$

Assuming $m_g = N_j/M_j$ for all g, $EF_j = 1 - 1/M_j$ and so $M_j = 1/(1-EF_j)$. Of course this assumption may not hold strictly but on a large scale it does not make a big difference. EF is considered as a measure of heterogeneity both within and across the countries in the world and it may serve as a measure of average connectivity of the country. Say k_g is a fraction f of the population of g^{th} group in j^{th} country *i.e.* $k_g = fm_g = k_j$. The epidemic threshold λ_{cj} is now given as $M_j / f s_j N_j$.
We find different values of k for different countries indicating heterogeneity. Figure 4 shows a less perfect fit to power law of cumulative probability of average connectivity P(k>x) computed from data on 30 countries taking $f=10^{-4}$. If data on country-wise average connectivity were available, this estimation could be validated. The heterogeneity within countries is being ignored but on a large scale, bounded connectivity fluctuations may be assumed. It is interesting to see a positive relationship between EF of countries and corresponding components of U^{13} (Figure 2(ii)). Since EF is an indicator of average connectivity, it implies lower the connectivity of a country; more (not necessarily significant) is its contribution to correlations. EF has been found to be negatively correlated with economic growth; latter being associated with healthy countries [11,13].

Another aspect in terms of which heterogeneity across countries/groups is described is their populations. No significant relationship between populations and magnitudes of U-13 components was found. The epidemic threshold depends inversely on population. As argued previously, isolation may bring down connectivity and hence the probability of infection but it may not help if the network population is very large. The following regression model would predict the effect of population on prevalence. (Note that the countries considered have a very small population growth rate per year)

$$P_{it} = \alpha + \beta_1 pop_i + \beta_2 time_t + \varepsilon_{it}$$

where P_{it} is the prevalence rate in country i at time t, pop_i is the population of i^{th} country relative to total population of all 13 countries, $time_t$ is time in years (1,2…) and ε_{it} is the error term. Table 1 below gives the results

	Estimate	**P-value**
α	0.00026 (0.0001)	0.0106
pop	8.067e-03 (0.00063)	0.1
time	-0.000036 (0.0000046)	9.41e-14

R squared = 0.3, p-value<2e-16

Table 1

The result confirms that population has a small but positive and quite significant effect on prevalence. However R squared is low and the residuals are high.

6. Estimation

This section estimates the parameters $(\lambda,h) \in [0,1]^2$. Two parts of equation (2) are analyzed separately as follows:
In time $\Delta t = 1$ year for j^{th} group, $\Delta p_j = - R_j + I_j$ where $R_j(t) = h_j\, pj(t)$ (i)
and $I_j(t) = a_j\, p_j(t)\, (s_j - p_j(t))$ (ii) where $a_j = \lambda_j k_j$

We use data on I_j , R_j and p_j and relative population s_j to estimate a_j , h_j for all j (13 countries). OLS is performed for (i) and (ii). Serial order correlation effects were not significant. Transformations did not increase significance. We report here the estimates from best models (least residuals) taking $p_j(t)$ as the predictor in (i) and $p_j(s_j - p_j)$ as the predictor in (ii).

$h = 1$ was obtained for all cases (consistent with [3]) and λ_j was determined from estimates of a_j by fixing k_j (computed using EF_j). The table below summarizes the results on λ, k for all countries for different values of f and compare with threshold:

Country	Relative Population	Ethnic Fractionalization Index	Number of groups, M	f	k approx range	λ	λ_c	Epidemic Observed
Azerbaijan	0.1044	0.49	2	10^{-5} 10^{-4} 10^{-3}	40 400 4000	0.21 0.021 0.0021	0.24 0.024 0.0024	Prevails, trails off to 0 toward the end
Georgia	0.0652	0.6543	3	10^{-5} 10^{-4} 10^{-3}	17 170 1700	0.74 0.074 0.0074	0.9 0.09 0.009	No No No
Belarus	0.1318	0.612	3	10^{-5}- 10^{-3}	34-3400	0.19-0.0019	0.22-0.0022	Prevails, trails off to 0 toward the end
Slovenia	0.0257	0.29	2	10^{-4}-10^{-3}	100-1000	0.21-0.021	0.38-0.038	No
Bosnia Herzegovinia	0.0457	0.69	3	10^{-4}-10^{-3}	117-1170	0.15-0.015	0.18-0.018	No
Norway	0.0574	0.21	1	10^{-5}-10^{-3}	44-4400	0.17-0.0017	0.39-0.0039	No
Ireland	0.0496	0.16	1	10^{-4}-10^{-3}	380-3800	0.12-0.012	0.05-0.005	Yes
Denmark	0.0692	0.31	2	10^{-5}-10^{-3}	26-2600	0.32-0.0032	0.56-0.0056	No
Netherlands	0.2023	0.722	4	10^{-5}-10^{-3}	39-3900	0.09-0.0009	0.13-0.0013	Yes
Czech Republic	0.1331	0.66	3	10^{-5}-10^{-3}	34-3400	0.15-0.0015	0.22-0.0022	No
Iceland	0.0037	0.19	1	10^{-3} and lesser	284 and more	0.9 and lesser	0.95 and lesser	No
Lithuania	0.0444	0.41	2	25 and lesser	Out of range			
Slovakia	0.0692	0.57	2	10^{-5}-10^{-3}	27-2700	0.4	0.54	No

Table 2

The results are based on approximate orders considered for *f* such that $\lambda \in [0,1]$, the feasible range. The values of λ for different regions show the effect of disaggregation on disease spreading. The figures 3 and 5 below confirm some of the findings.

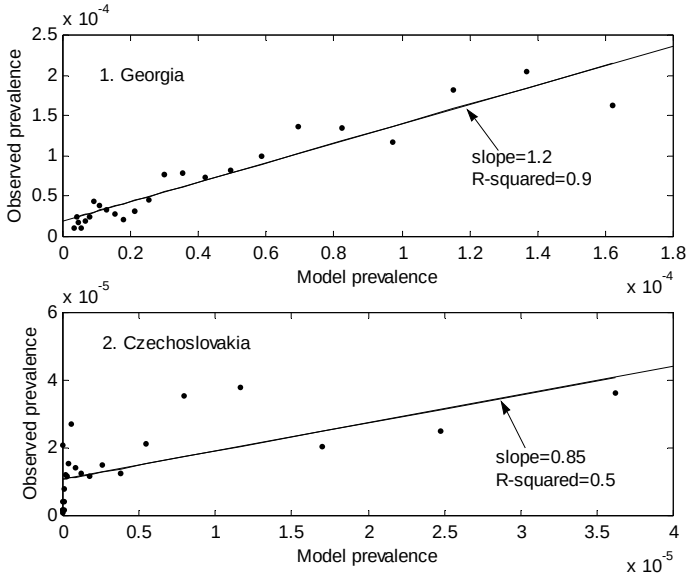

Figure 3 Observed prevalence versus the prevalence obtained by model simulation shown for 2 countries-(1) Georgia(Top), (2)Czech(Bottom).

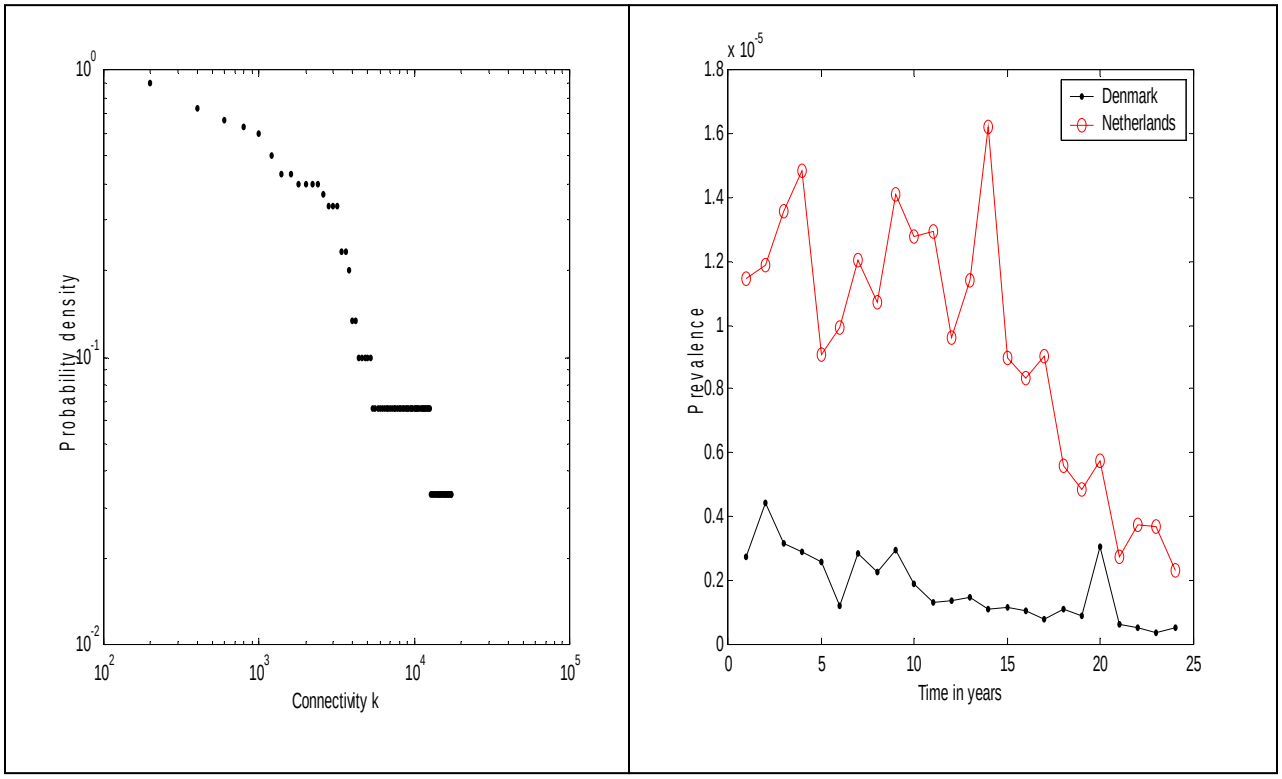

Figure 4 (Left) Log-log plot of Cumulative probability distribution of connectivity P(k>x) vs. k
Figure 5 (Right) Temporal evolution of Observed prevalence for Denmark and Netherlands

7. Agent Based Model

So far the effects of disaggregation of a network in terms of average connectivity and population have been analyzed. These factors may vary independently given the cultural and demographic diversity across the countries or clusters into which the network (world) is divided. It was found that epidemics in a disaggregated network depend mainly on the interplay of these two factors. And while isolation (lowering average connectivity) may work effectively in sparsely populated networks, it may not be the answer to epidemics in crowded networks. This section validates the above by explicitly comparing the disease transmission for 3 network configurations with parameters *N*, number of nodes and *k*, average connectivity. The schematic diagram below shows the sequence. To be seen first is the effect of lowering connectivity by fragmenting the network while keeping *N* fixed and second, the effect of increasing *N* at fixed *k*. An agent-based model (ABM) is simulated for each configuration (individuals as agents connected by links).

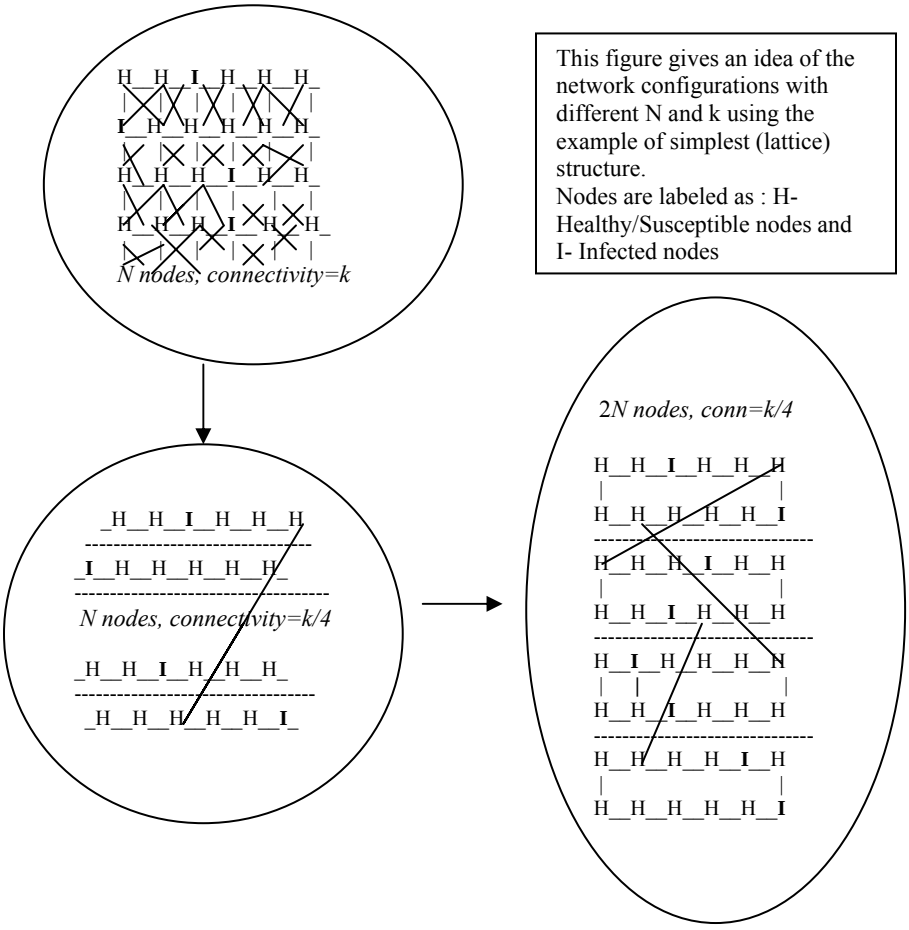

Figure 6 Sequence of network configurations used in simulation

ABM's have been found particularly useful as compared to analytical methods when the social network is more complex (e.g. scale free networks). The ABM's built using data on the social structure, connections of a network have also been validated [14]. The purpose of doing this experiment here is to incorporate some randomness in the process. Studies have shown community structure arising from evolving network properties such as degree distribution, clustering etc [15]. However here, the homogeneity of the network is preserved while changing either *k* or *N*. Although broadly it is the SIR approach, here only the intermediate state is considered *i.e.* what happens before nodes are immunized

or removed permanently. This is valid for common infections for which humans remain susceptible even after being cured. The rules of simulation are as follows.

(i) Start with a network structure of size N and p_i proportion of infected nodes, average connectivity k.

(ii) Specify an adjacency matrix A such that A_{ij} is 1 if nodes i and j are linked and 0 otherwise.

(iii) 'Infected' nodes are identified as 0 and 'healthy/susceptible' as 1.

(iv) At each time step, every healthy node is checked. If it is linked to an infected node then with some probability δ, it will be infected due to that node. Thus more the links a node has to an infected node, more will be the chance of it being infected.

(v) If a healthy node becomes infected, it's status is changed to 0.

(vi) Randomly pick an infected node and make it healthy.

(vii) At the end of every time step review how the proportion of infected has changed.

Next, the network of size N is divided into 2 clusters of size $N/2$ (very sparsely connected between them) so that average connectivity of each and of the network is $k/2$. The steps 2-7 are repeated to see the growth of infection.

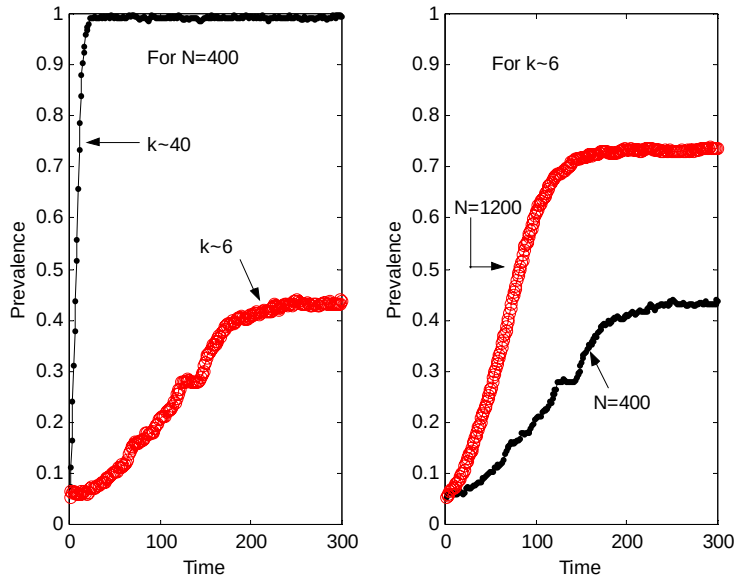

Figure 7 Prevalence patterns for 3 different combinations of N and k.

The above procedure is repeated for the network of size N divided into 4 groups, average connectivity $k/4$. From figure 7 (left), it is evident that not only the rate of spreading but also the steady state probability of infection has reduced in this case. Now for this case the network size is increased 3 times keeping connectivity = $k/4$ and the same simulation is carried out. Figure 7(right) shows the difference caused by larger network size. It confirms that a crowded network has not only higher growth rate of infection but also higher steady state infection level. This is happening because despite fewer direct links between healthy and infected, higher population increases the number of paths for transmission. This means that isolating oneself does not guarantee safety in crowded communities. This is different f-rom what has been generally believed that higher population implies higher connectivity,spreading rate.

8. Conclusion

This paper analyzes the impact that disaggregation of a network (the 'world') into groups (countries) with different cultural and demographic features, has on disease dynamics. Application of RMT methods reveals (confirms) no systematic correlation in the prevalence between countries of the

ensemble. Nor is there any apparent common influence or shock (e.g. new health technology). The basic modeling framework applied in some previous studies is extended to yield country specific patterns. In particular, it is seen that epidemic threshold in the new framework depends on not just the connectivity but also the population of the region/country. It is argued that these constitute the cultural (ethno-linguistic, religious), demographic aspects and are independent. Heterogeneity across countries is examined in terms of these aspects. A less perfect power law fit is obtained for connectivity distribution. Interestingly, from RMT analysis, it is seen that countries with higher EF seem to contribute more (but not significantly) to correlations and it has been established earlier that EF is negatively correlated with economic growth and hence national health[11, 13]. Even though it has been constructed as a rough estimate of connectivity, the results seem consistent with observed patterns. It would be interesting to see if/how the results would change if the probability of infection spreading across clusters/countries is considered finite (>>0).

Infection spreading and curing rates are estimated using statistical analysis of data on prevalence, incidence. The analytical results seem consistent with those observed. Population factor is found to have a small but positive and quite significant impact on prevalence. This is consistent with the results obtained by agent based simulation of 3 network configurations. It is confirmed that while lowering k brings down the prevalence rate and steady state probability of infection, both of these increase when population of the network increases at same level of k. While the disaggregated modeling framework brings out differences in disease patterns based on cultural, demographic diversities, it reveals unapparent yet interesting variations. For instance, during the same period, epidemic exists in Netherlands but disappears in Denmark. Hence it is important to consider the demographics of the region to make sure if isolation is the answer to epidemics. Crowded/populated countries run greater risk of infection/epidemic. *'More'* may not be *'Merrier'*

Acknowledgements

I thank Abhay G. Bhat, Raghav Gaiha and Suresh Kulkarni for useful clarifications and discussions. I am grateful to Santa Fe Institute summer school for providing this opportunity.

References

1. B. Pourbohloul and J. Miller, Network Theory and Spread of Communicable Diseases, http://www.math.ualberta.ca/~irl/summer_school/lecture_notes/network
2. N. Grassley and C. Fraser, Mathematical models of infectious disease transmission, *Nature Reviews Microbiology* V 6, 477-487, 2008.
3. R. Satorras and A. Vespignani, Immunization of Complex Networks, *Physical Review E*, V 65, No. 3, 036104, 2002.
4. K. Park, Textbook of Preventive and Social Medicine, 17th Edition, Banarasidas pub. India, 2002
5. L. Sattenspiel, The Geographical Spread of Infectious Diseases: Models and Applications, *Princeton University Press*, July 2009.
6. N. Beckner and S. Utev, The Effect of Community Structure on the Immunity Coverage Required to Prevent Epidemics, *Mathematical biosciences*, V 147, 23-39, 1998
7. V. Plerou, P. Gopikrishnan, B. Rosenow, L.A. Amaral, H.E. Stanley, A Random matrix theory approach to financial cross-correlations, *Physica A*, V 287, 374-382, 2000
8. V. Plerou, P. Gopikrishnan, B. Rosenow, L.A. Amaral, T. Guhr, H.E. Stanley, Random matrix approach to cross correlations in financial data, *Physical Review E*, V 65, 066126, 2002
9. M. Mehta, Random Matrices, *Academic Press*, 1991
10. V. Kulkarni and N. Deo, Correlation and volatility in an Indian stock market: A random matrix approach, *European Physical J. B*, V 60, 101-109, 2007

11. D. Bloom, D. Canning, G. Fink, Disease and Development Revisited, *working paper, National Bureau of Economic Research*, 2009
12. N. Lin, K. Cook, R. Burt, Social Capital (Theory & Research), *Aldine Transaction*, New Brunswick & New Jersey, 2001
13. A. Alesina, A. Devleeschauwer, W. Easterly, S. Kurlat, R. Wacziarg, Fractionalization, *Journal of Economic Growth*, V 8, 155-194, 2003
14. A. Skvortsov, R. Connell, P. Dawson, R. Gailis, Epidemic Modelling: Validation of Agent-based simulation using simple mathematical models, http://www.mssanz.org.au/modsim07/papers/13_s20/EpidemicModeling_s20_Skvortsov_.pdf
15. C. Li and G. Chen, Modeling of weighted evolving networks with community structures, *Physica A*, V 370, 869-876, 2006