UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia "Galileo Galilei"

Master Degree in Physics (LM-17)

Final Dissertation

Modelling COVID-19 spreading in a network

Candidate

Riccardo Milocco

Thesis supervisor

Prof./Dr. Marco Baiesi

Academic Year 2020/2021

Contents

Abstract

The usual simplified description of epidemic dynamics predicts an exponential growth. This is due to the mean field character of the dynamical equations. However, a recent paper (Thurner S, Klimek P and Hanel R 2020 Proc. Nat. Acad. Sci. 117, 22684) [4] showed that in a network with fixed connectivity, the nodes become infected at a rate that increases linearly rather than exponentially. Experimental data for COVID-19 seem to validate this approach. In this thesis we plan to study this model by tuning its parameters. In particular, we monitor the effect induced by a significant presence of hubs in the network.

Thesis Summary Riccardo Milocco

Last Update: 25-06-2021

1.1 Introduction

1.2 Goals Outline

1. Assuming that a typical social network has an average number of contact (degree) $D \simeq 5$; while lockdown measures reduce it to the household size ($D \simeq 2.5$ [references??]), our goal is to reproduce the dynamics of an epidemic spreading on a sparser network due to the Non-Pharmaceutical Interventions (NPIs) generally known as lockdown[1]. A direct comparison of the obtained curves with the states of Austria and the USA [4] is satisfactory even at a plain-eye inspection. However, the project aims at the study of how the topology of a network could characterize a paradigmatic model as the SIR one. So, no statistical fit has been performed in this work, since a relevant work on the available data, e.g. *GitHub Protezione Civile* , has to be performed. Specifically, the number of infected at the early stages has to be refined due to the heterogeneous way of defining an "Infected" person. In fact, In particular, these states express the (quasi-)linearity in the cumulative number of cases similarly to the simulated ones. Moreover, it has to be pointed out that the previous quoted states are representative of two opposite political policies facing the Covid-19 pandemic and undertaken by many others: a immediate "severe lockdown" (Austria) and "light restriction" on contacts (USA).

2. As claimed in [4], we want to find that for any fixed transmission rate, recovery rate and "long-range" probability (β, μ, p) there exists a critical number of social contacts (D_c) which describes a first-order phase transition of the disease growth obtained with the SIR model:

   - for $D < D_c$, linear growth, i.e. low infection prevalence, occurs;

   - for $D > D_c$, "classical" exponential growth as predicted by the SIR model.

   The rationale behind is that, since $R_0 := \beta D/\mu$, by augmenting the average degree of the udnerlying network also the R_0 is enanched, resulting in a more powerful spreading. On the way to obtain D_c, the simulated plots show a strongly dependence on p, i.e. the probability of a node to have a distant neighbor ("long-range" attachment), which is as crucial as β, μ and D in the characterisation of a epidemic spreading. Therefore, a new definition of $R_0 := R_0$ is proposed, which integrate the possibility that same effective parameter but with different long-range attachment (e.g. one local and the other random attached) could drive to different scenarios. Finally, by setting our parameters to the empirical estimates of β and μ, we are going to estimate D_c.

---

[1]For brevity, the word "lockdown" is an alias for all the "lockdown-like" measures undertaken to prevent the diffusion of Covid-19 without requiring Pharmaceutical drug treatment, e.g. confinment (the direct translation of "lockdown") but also contact tracing, higher hygiene, face masks,

At 8 May 2020, none of the affected Covid-19 states have reached "herd immunity" [2], but still they have reached the "epidemic peak" due to the containment restrictions on social contacts. The most striking observation is that the epidemic curves, especially when NPIs are introduced, exhibit a linear growth for an extended time interval in contrast with the "S-shaped" logistic curve as predicted by the standard compartmental models. The extended of the linear regime depends on the onset of the measures; while for the early stages, as it was the case for many countries [4] (8 May 2020), an exponential growth dominates the spreading of the disease. By taking care of the modification of the underlying social network structure, it's possible to recover the spreading trends for the states that have or have not applied the NPIs measures.

Infection may occur for two reasons:

1. interaction between an infected and a susceptible person;

2. contact is "intense" (e.g. long, close,...) enough to lead to a disease transmission.

So, the rationale behind social distancing is that it takes to a reduction of both of these factors. Moreover, the standard SIR model assumes that there's the same probability that an individual encounters an infected person ("well-mixed population") and all the nodes have the same number of neighbours ($N - 1$ or $\langle k \rangle$). However, since their first introduction, considering the complex network structure as provided remarkable results. e.g. how the vaccination threshold depends on the network topology [3]. No focus was still be put on the spreading of the disease below the epidemic threshold, since few works where published trying to analyze the lockdown effects, e.g. [2]. In this thesis work, we try to grasp the relevant features of a complex social network. In fact, they present richer structures with respect to the well-known model ("Erdoes-Rényi","Scale-Free", etc.), including multilevel organization; weak ties between communities; and temporal aspects that suggest a degree of fluidity, however, with stable social cores. To this scope, we proceeded by following steps.

At first, we pruned a fully connected graph of $N = 1000$ nodes, by keeping fixed N while having the nearest neighbors k. In parallel, since $R_0 = \frac{\beta \langle k \rangle}{\mu}$, we doubled the β parameter to keep the power of the epidemic R_0 fixed (1.1).

Since the aim is to understand at which extent the underlying network structure force the spreading to differ with respect to the "well-mixed" population model, i.e. the mean-field theory, we put in the same plot the number of total infected cases with the epidemics that has spread either on the network (in green) or in a "mean-field" fashion (in orange)[3]. More precisely, the "mean-field" is a "quenched-mean-field" where every node has D nearest neighbors picked at random on the whole set of nodes. In this way, the hypothesis of the "well-mixed" population is recovered, i.e. every-one has the same probability to get in touch with an infected, and the spreading could be modeled with the exact same parameters (β and μ) as the network epidemics.

Then, we introduced the "rewiring probability" which enables the distant nodes to connect faster, thus, obtaining a standard Erdoes-Rényi with non-zero p. In this case, we could recover the "linearity" of the SIR infection as is shown in 1.2.

To further test the recovered linearity with respect to the mean-field benchmark, we used a Poissonian "small-world" network which enabled us to taking care of degree heterogeneity (i.e. different social contacts), family-clusters and their overlaps, small-world feature, given the fact that "leisure" activities may connect nodes that are diametrically opposite. Moreover, by changing the number of the average degree D we may pass from light to severe NPIs measures, i.e. strict lockdown. The Poissonian "small-world" network could be reproduced by imposing, at first, a Poissonian degree sequence on all nodes ("Configuration Model"); then, re-link the nodes to their closest neighboring nodes on the circle. Thus, obtaining a network with poissonian degree distribution and the small-world property. The linearity obtain with this graph is reported in the 1.3

---

[2]The total infected cases at the first peak were 0.3% of the total population, remarkably low with respect to the SARS reported level that are $0.5\% - 0.8\%$ of the entire population.

[3]Moreover, the daily infected of network (in blue) and of the mean-field (in magenta)

The further works are focused in analyzing the network spreading on a fully-connected cavemen graph and on a network with the presence of hubs. Moreover, the critical average degree D_c (cfr. the initial part of this review) is going to be recovered for the quoted network topologies.
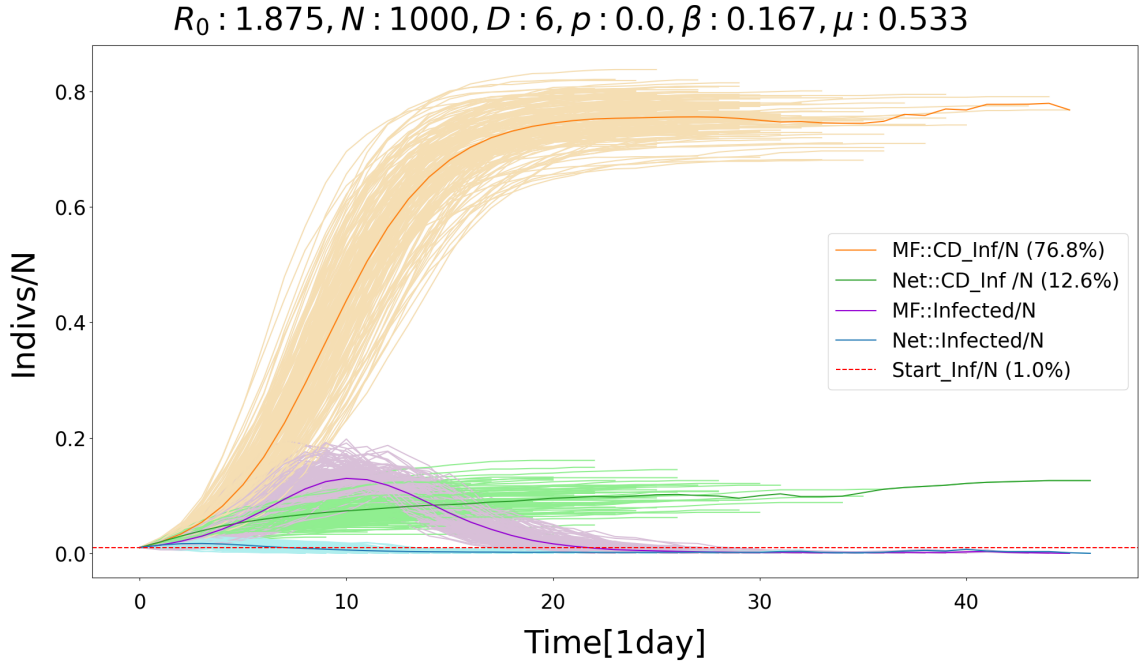
Figure 1.1: Linearity for "pruning" method with $p = 0$

Figures
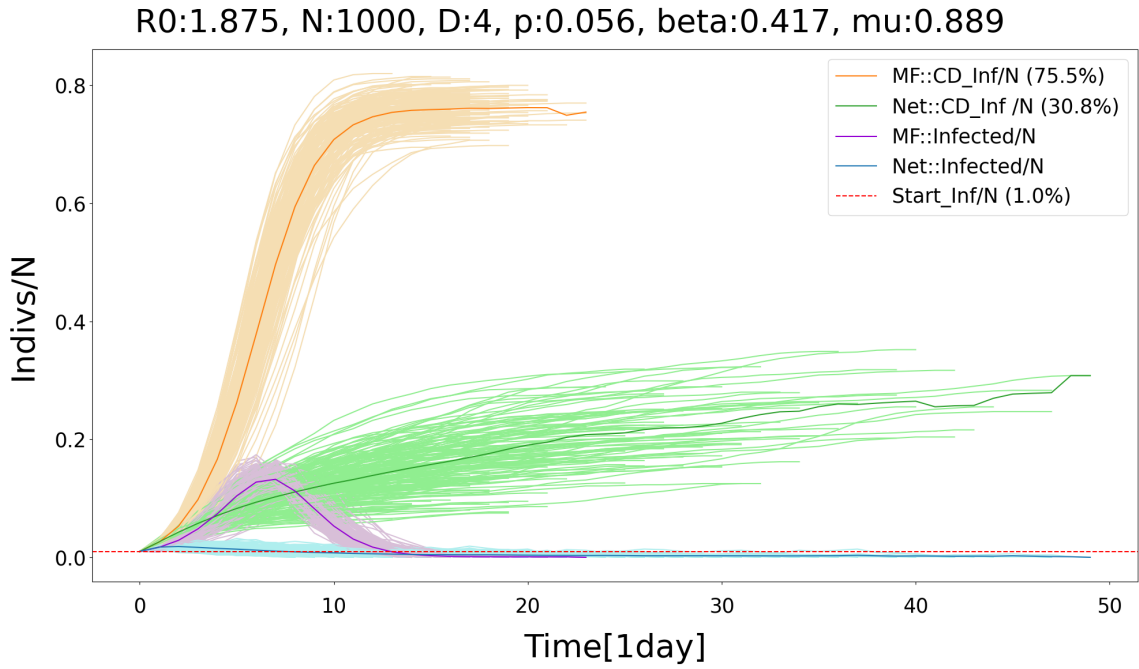
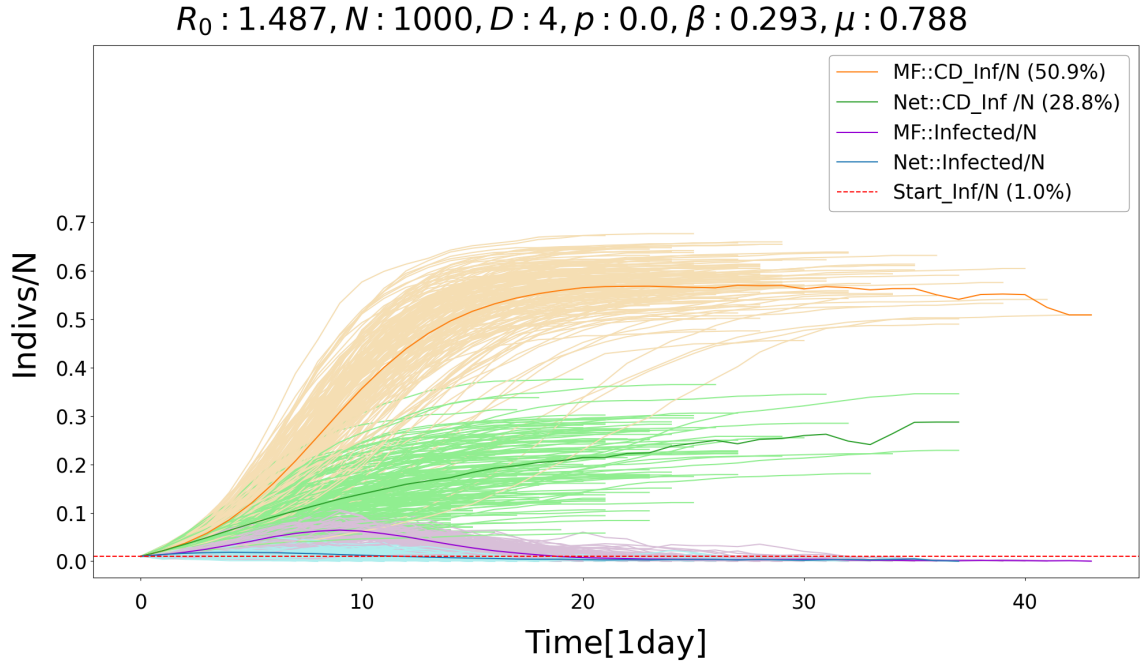Figure 1.2: Linearity of the total number of cases with $p = 0.056$

Figure 1.3: Linearity for the nearest neighbors rewiring with $p = 0$

Figure

Model

The main assumptions of the model are:

- fixed number of total individuals N connected by a social link;

- "undirected" adiaciency matrix A, where $A_{i,j} = 1$ if there's a link between the ith- and jth-node or $A_{i,j} = 0$ otherwise;

- transmission per day probability r and d days for an infected individual to be recovered, i.e. immune or naturally dead;

- average degree D and shortcut (or rewiring) probability ϵ

As one may expect, the r and d parameters are in a one-to-one map with to the SIR ones, respectively, for the recovery and transmission rates:

$$\gamma = 1/d, \quad \beta = rD/N. \tag{2.1}$$

At each time step (day), the infection curve of positive cases, $P(t)$, is the cumulative sum of the number of new cases $C(t)$. So, if $D > D_c$ the "well-mixed" (or homogeneous/mean-field) approximation may be applied, with the result that $C(t)$ is associated to $R(t)$ up to a time-shift of d, since all the infected are predicted to become recovered after d days.

Results

3.1 Infection Dynamics

To gain some intuition about the importance of modelling the underlying network, we've made the disease spread for two different values of D but fixing $N_{total} = 1000$, $D = 8$, $\epsilon = 0.1$, $d = 6$ days, $r = 0.1$ and 10 initially infected nodes as seeds ($N_{seeds} = 10$).
The results are the following:

- In the large D and ϵ limit, the mean-field condition is fulfilled and the model resembles the compartmental SIR model. By mapping the parameters as done in 2.1, we may compare $P(t)$ from the network model with the compartmental $R(t)$ by shifting it of $-d$ days. As a relevant reference for the comparison to the "small D" evolution, we've plotted the histogram of the daily cases which shows the typical peak in the early-exponential stages and, then, a decrease towards herd immunity, which for [4] has been reached at the 98% of the overall population.

- By reducing the average degree to $D = 3$, we've find a changing in the spreading behaviour, since $P(t)$ increases almost linearly for a remarkable timespan. At the end of the infection, nearly $1/5$ of the entire population has been infected which is noticeably smaller than the SIR predicted herd immunity nearly at 80%.

With this rough perception of the main results by considering the underlying network structure, we've further analyzed the parameter dependence.

3.2 Parameter Dependence and Phase Transition

The aim of this section is to obtain the value of the "critical average degree" D_c, such that, for $D < D_c$, the infection curve is (roughly) linear. In order to take care of the two growth trends, we've defined an order parameter

$$\mathbb{O} := SD(C(t)) \tag{3.1}$$

as the standard deviation of the daily new cases, having removed the days without new infected. In this way, $\mathbb{O}! = 0$ signals the presence of a nonlinear increase of the daily new cases and, therefore, of its cumulative sum $P(t)$. Other way around, $\mathbb{O} = 0$ corresponds to a linear growth of $P(t)$.

Appendix

4.1 Notes on Network book of A.Barabasi

4.1.1 Compartimental Models 10.2

SI model

At (10.1) β is the probability of 1 spread only. So, $\frac{\beta \langle k \rangle S(t)}{N}$ expresses the probability of 1 spread over $\langle k \rangle S(t)/N$ neighbors.

$C := \ln \frac{i_0}{1-i_0}$

SIS model

In (10.7), define $v_{SIS} := 1 - \frac{\mu}{\beta * \langle k \rangle} = 1 - R_0^{-1}$ the "characteristic velocity". If $v > 0$ spread; otherwise it dies out. So, $\tau^{SIS} = \frac{1}{\beta \langle k \rangle v_{SIS}} = \frac{\tau^{SI}}{v_{SIS}}$ In, SI model $v_{SIS} = 1$ [1]

The endemic fraction of infected is given by $\frac{\beta - \mu}{\beta}$ slides 16 of the LDE course.

4.1.2 Epidemics on Nets 10.3

SI Model

10.16 :

$$\tau_n^{SI} = \frac{\tau_c^{SI}}{\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle^2}}$$

10.17 : is obtained by imposing the initial condition $i(t = 0) = i_0$. Anyway, $i_k = i_0 + i_0 f(t)$

SIS

$$\tau_N^{SIS} = \frac{\tilde{k}^2}{\beta \langle k \rangle - \mu \tilde{k}^2}$$

where $\tilde{k}^2 := \frac{\langle k \rangle^2}{\langle k^2 \rangle}$ and ???? if the network is homogeneous $\langle k \rangle^2 = \langle k^2 \rangle$. Check better what's the behaviour of $\langle k^2 \rangle$ wrt the considered network.

4.1.3 10.B

If a nets lacks degree of correlation, i.e. $e_{ij} = q_i q_j$, the probability $p_{kk'}$ is independent on k. Thus,

$$p_{kk'} = \frac{k' p'_k}{\langle k \rangle}.$$

Typo under (10.50) multiply by $k p_k$ not $k - 1) p_k$

Tight Binding model

5.1   Second Quantization formulatio of the SIR Model

Bibliography

[1] Albert-László Barabási and Márton Pósfai. *Network science*. Cambridge University Press, 2016.

[2] Neil Ferguson. Capturing human behaviour. *Nature*, 446:733, 05 2007.

[3] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic Spreading in Scale-Free Networks. *Physical Review Letters*, 86(14):3200–3203, April 2001.

[4] Stefan Thurner, Peter Klimek, and Rudolf Hanel. A network-based explanation of why most covid-19 infection curves are linear. *Proceedings of the National Academy of Sciences*, 117(37):22684–22689, 2020.