A POWER-LAW MODEL FOR THE SPREAD OF COVID-19

Yuhong Chen, Fangzhou Liu*, Ni Dang, and Martin Buss

Chair of Automatic Control Engineering
Technical University of Munich, 80333 Munich

April 21, 2020

ABSTRACT

The current pandemic outbreak caused by the corona-virus COVID-19 has led to many infected cases and deaths globally. Mathematical modeling of the spread of the epidemic and predictions of its further course are critical to deploy effective control measures. In this note, we introduce a power-law model with an exponential cutoff to study the daily new confirmed case. The proposed model indicates that the epidemic spreads on a networked population but not a well-mixed one. The network structure discriminates our model from the macroscopic compartmental epidemic models, where the individuals are assumed to interact with each other with the same chance. Simulation results, especially the simulations about China and South Korea, illustrate the capability of the proposed model to fit and predict the outbreak trend of the COVID-19. Further comparison shows that the proposed model outperforms the macroscopic susceptible-infected-recovered (SIR) model.

1 Introduction

Since the first notification of COVID-19 by the World Health Organization on January 5, 2020, the epidemic has rapidly spread world-wide. As a fundamental step to predict and control the spread of this pandemic, an effective model is necessary to characterize the propagation process.

Much work has been done to model the diffusion of COVID-19 using macroscopic compartmental epidemic models, e.g., Susceptible-Infected-Recovered (SIR) and Susceptible-Exposed-Infectious-Removed (SEIR) [1]. These compartmental models mostly assume certain contagion and curing rates that need to be identified and can decide the trend. It worth noticing that the model works under the assumption that every susceptible person had the same possibility to contact infected people, which obviously omit the network structure in population and is not suitable for our highly grouped society [2]. Moreover, as the estimated curing rate in the very early stage can hardly match the real one, the prediction given by this model always falls into endemic equilibrium regardless of the epidemic genres, which is not accurate.

Spreading through respiratory tract droplets and contact, the virus easily infects people in close contact, such as families and colleagues. It can also spread in limited contact scenarios such as market and communities. Generally speaking, infection is highly relevant to physical distance, so quarantine at home, transportation restriction between cities and states blockade all greatly affect the spread of virus. In this situation, the diffusion process can hardly be seen as the classic one in a well mixed population with universal contact possibility. The communication networks between individuals clearly play an important role in the process. Inspired by [3], we use power-law model which is laid under free scale network structure to fit the COVID-19 diffusion data. The good fitness indicates an underlying small-world network of connections between susceptible and infected individuals.

Some endeavors has been made on building power-law models for COVID-19. Anna L. Ziff and Robert M. Ziff's [4] introduced the power-law model with an exponential cutoff to approximate and predict the trend in China, The death data in China has been well fitted with fractal exponent $\alpha \approx 2.25$. Maier *et al.* [5] studied the confirmed infection cases for different regions of China and found that the number can be effectively fitted by the pow-law with exponents 2.1 ± 0.3. Li *et al.* [6] also found power-law behavior in the number of infections and the number of recoveries with

*Email address: fangzhou.liu@tum.de

different exponents. However, the confirmed infection cases and death cases are not actually caused by diffusion. Moreover, the performance difference of SIR model and power-law model is not illustrated. Therefore, we carry approximation on daily infected cases in 7 countries and give comparisons on the performance of SIR and power-law model for China and Germany, as representatives for ending and ongoing phases. The details are given as follows.

2 Main Results

In this section, a power law model with an exponential cutoff is utilized to characterize the daily new infected cases of COVID-19. A power law with an exponential cutoff is simply a power law multiplied by an exponential function:

$$n(t) = Kt^x e^{-\frac{t}{t_0}}, \tag{1}$$

where K, x and t_0 are positive constant parameters. t^x represents the power-law term and $e^{-\frac{t}{t_0}}$ is the exponential part. In this evolution, in the early stage, the power law behavior dominates the growth, and the exponential decay term $e^{-\frac{t}{t_0}}$ eventually overwhelms the power-law section for very large values of t.

We use least squares method to identify the nonlinear function (1). Specifically, the parameters are obtained by solving the following optimization problem,

$$\min_{K,x,t_0} \sum_{t=1}^{l} \gamma^{l-t} \Big(n^*(t) - Kt^x e^{-\frac{t}{t_0}} \Big)^2, \tag{2}$$

where $n^*(t)$ is the infected cases at the t-th day and we consider the data up to l days. As the situation of the diffusion is quite dynamic, we introduce a forgetting factor $\gamma \in (0, 1]$ into the variance such that the older error samples are assigned with exponentially less weight. In the following simulation, we set $\gamma = 0.98$.

We acquire the daily new infected data from the website Worldometer [2] and using the fmincon function in MATLAB platform to realize the numerical algorithms. The fitting results of Korea, China, France, Germany, Italian, Spain, America and world wide are given as follows in Fig. 1. The period of data used and the output fitted parameters are given in Table 2. We make long-term prediction of the diffusion and provide the prediction of turning point date T_{peak} and total infected numbers over time I_{total}. The estimated end date T_{end} in this paper is given by the first time when the estimated daily new infection number goes below 0.5.

Country	K	x	t_0	Estimated I_{total}	Starting Date	Estimate T_{peak}	Estimated T_{end}
South Korea	3.68e-02	6.51	1.88	8.09e+03	18-Feb-2020	01-Mar-2020	31-Mar-2020
China	6.11e-01	5.08	3.07	7.65e+04	23-Jan-2020	07-Feb-2020	30-Mar-2020
France	1.20e-06	8.24	5.03	2.51e+05	26-Feb-2020	07-Apr-2020	16-Jul-2020
Germany	1.80e-06	8.70	3.86	1.63e+05	26-Feb-2020	30-Mar-2020	16-Jun-2020
Spain	1.33e-06	8.97	3.70	2.09e+05	26-Feb-2020	30-Mar-2020	14-Jun-2020
Italy	1.84e-02	5.11	6.27	1.96e+05	23-Feb-2020	26-Mar-2020	10-Jul-2020
USA	1.18e-06	8.54	5.33	1.33e+06	26-Feb-2020	11-Apr-2020	08-Aug-2020
Worldwide	3.63e-06	6.75	15.06	1.47e+07	01-Feb-2020	12-May-2020	18-Apr-2021

Table 1: Simulation Results

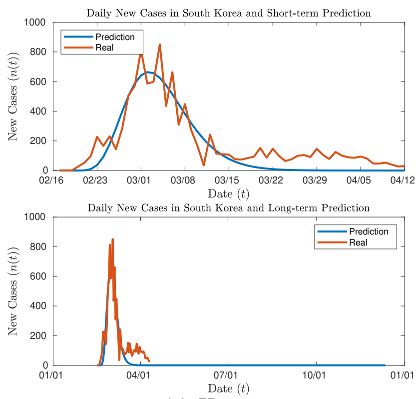
Table 2 presents rough estimation of the epidemic diffusion in every country. As the power-law part plays the dominate role in the initial increasing stage, the parameter x measures the increasing rate in certain degree, and so does t_0 to the decreasing rate. From the fitting results, close x are attributed to France, Germany and Spain, which is consistent with the fact that they went though similar diffusion process. In addition, those countries known for strict lock down measures, say China and Italy, show notably low x. Note that Korea results in the fastest decay with the smallest t_0 which may come from the wide testing and the accurate tracking.

From the perspective of prediction, we introduce the peak date $T_{peak} = T_{start} + xt_0$ by differentiating formula(1), which implies that lower x and t_0 both contribute to accelerate the arrival of the turning point. Unfortunately, America may be the last in the 7 countries to meet the peak and ending, due to the fast growth and slow decay. As the geographical isolation and national lock down, outbreaks of the pandemic in the world are not synchronized, but in regions and

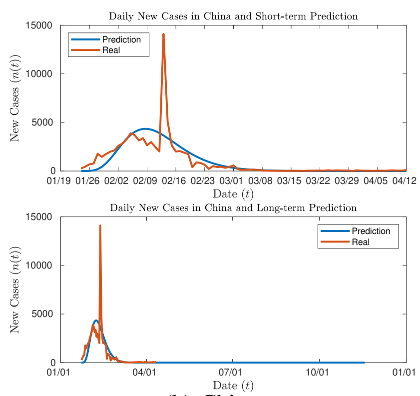
[2] https://www.worldometers.info/coronavirus/

batches like a relay. We obtain high t_0 for the spread of COVID-19 world-wide, which indicates rather low decaying rate and the end date is estimated to come in the April of next year.

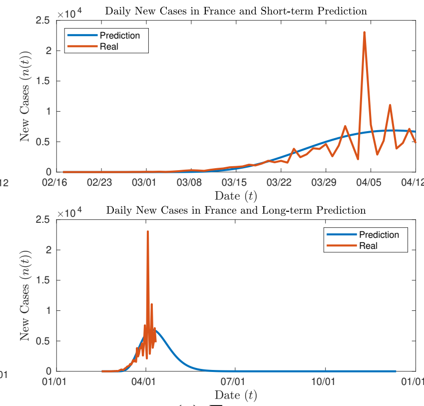
More intuitive results are given in Fig. 1. It shows that the power-law model fit the numbers well.



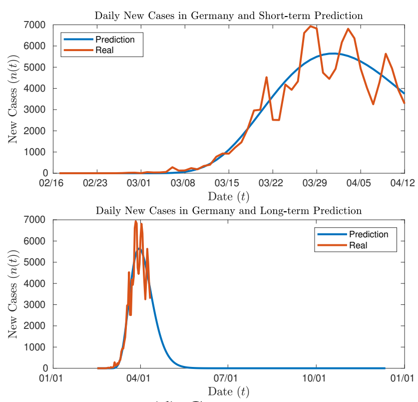
(a) Korea



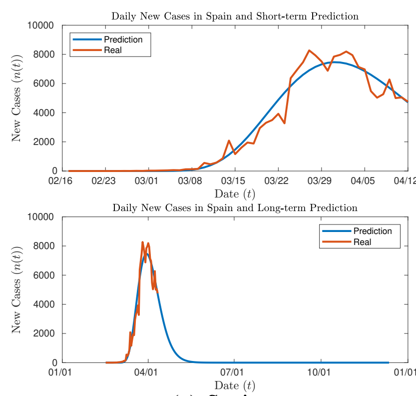
(b) China



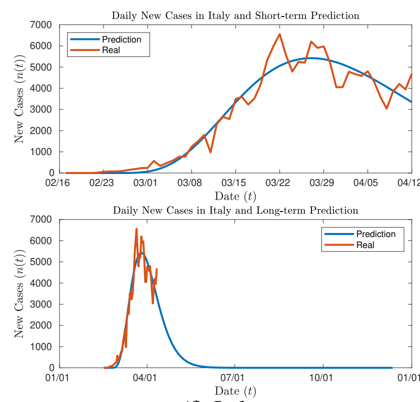
(c) France



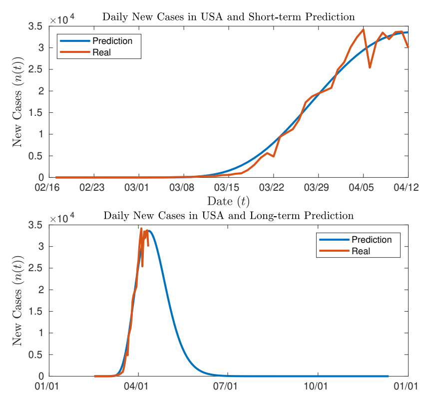
(d) Germany



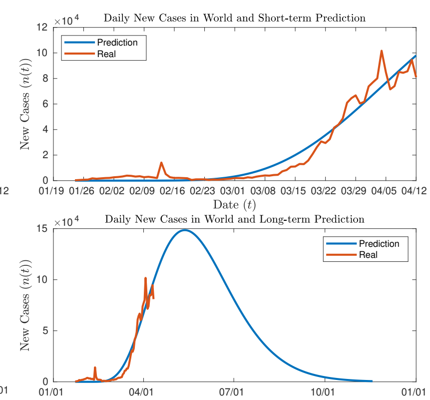
(e) Spain



(f) Italy



(g) USA



(h) world Wide

Figure 1: Simulation Results

3 Comparison with the SIR Model

In this section, we compare the power-law model with the macroscopic SIR model. The SIR model is given by the following difference equations

$$\begin{array}{l} I(t+1) = I(t) + \frac{\beta}{N}(N - I(t) - R(t))I(t) - \delta I(t) \\ R(t+1) = R(t) + \delta I(t), \end{array} \quad (3)$$

where $I(t)$ and $R(t)$ represent the active infected cases and recovered cases on the t-th day, respectively; β and δ are the infection rate and curing rate, respectively; N represents the population of the country. In the SIR model (3), it is assumed to hold for all $t \geq 0$ that $S(t) + I(t) + R(t) \equiv N$, where $S(t)$ is the number of the susceptible. Thus, the model (3) is valid for the epidemic spread in a fixed population. Note that the number of the susceptible can hardly obtained in real spreading processes. In this regard, we take into consideration the dynamics of $I(t)$ and $R(t)$.

The cost function in the regression is given by

$$\min_{\beta,\delta} \sum_{t=1}^{l} \lambda_1 (I^*(t) - I(t))^2 + \lambda_2 (R^*(t) - R(t))^2, \quad (4)$$

where $I^*(t)$ and $R^*(t)$ are the real number of active cases and recovered cases, respectively. λ_1, λ_2 are the weighting coefficients, both of which are set as 0.5 in the following simulation. It worth noticing that here we use the data of closed cases to fit the model instead of only recovered cases which omit the death cases.

The comparison between the power-law model (1) and the SIR model (3) is conducted by using the data of China and Germany. By solving the optimization problem (4), we obtain that the parameters for China are $\beta = 1.42, \delta = 1.41$, while those for Germany are $\beta = 2.85, \delta = 0.10$.

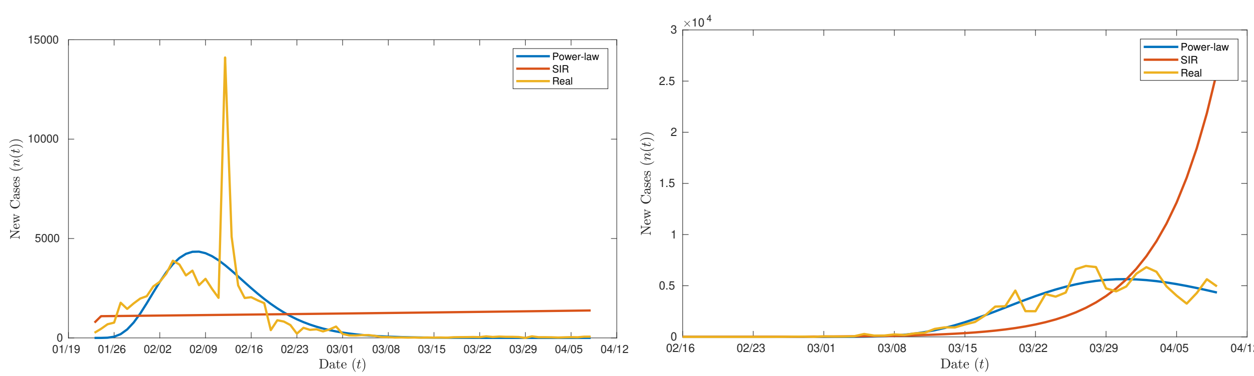
Note that the SIR model fits active cases while the power-law model fits daily new cases. Thus, we conduct the comparison from these two perspectives separately. In specific, based on the SIR model, the daily new cases at the t-th day, $n_{SIR}(t)$, can be obtained by

$$n_{SIR}(t) = I(t) + R(t) - I(t-1) - R(t-1).$$

As is shown in Fig.2, it is clear that $n(t)$ outperforms n_{SIR} in fitting $n^*(t)$. Apart from the daily new cases, we obtain the total active cases (I_{pl}) by using the power-law model as follows

$$I_{pl}(t) = \sum_{i=1}^{t} n(i).$$

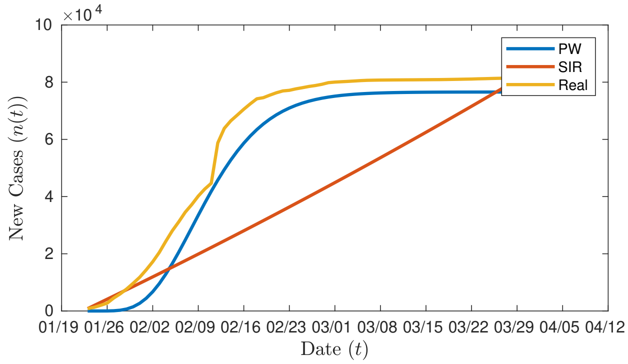
The comparison between $I(t)$ and $I_{pl}(t)$ as well as $I^*(t)$ are presented in Fig. 3 It can be observed that, in the analysis of China, the power-law model better simulates the trend of first rising rapidly and then tending to stay flat while the SIR model gives a longer rising forecast and more total infections expected. In the data of Germany, the power-law model also gives better prediction of the total infected cases than the SIR model.



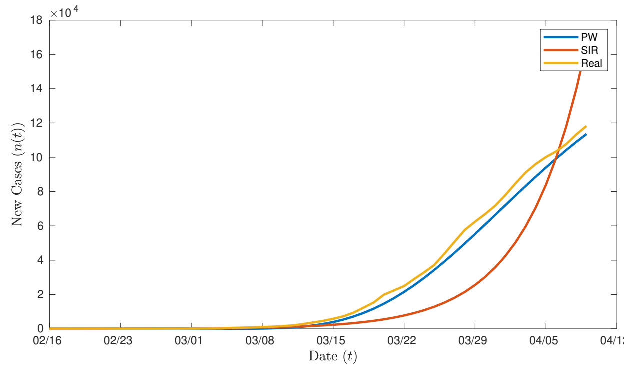
(a) Daily New Cases in China (b) Daily New Cases in Germany

Figure 2: The comparison between SIR Model and Power-law Model Using Daily New Cases

The estimated peak dates by SIR are 16-May-2020 and 01-Jun-2022 for Germany and China with end dates 09-Nov-2020 and 11-Nov-2030, respectively. Estimated total numbers of infection are 7.75e+07 for Germany which mount to



(a) Total Infected Cases in China



(b) Total Infected Cases in Germany

Figure 3: The comparison between SIR Model and Power-law Model Using Total Infected Cases

93.6% of the population and 3.36e+06 for China which takes 0.24% of the population. Apparently, the SIR model used can hardly predict the trend of daily new infection and total infection, even given the data of an ending diffusion. The reason possibly lies in the inflexible setting of fixed closing rate and infecting rate, as the network structured individuals do not spread the virus in a uniform way.

4 Conclusion

In this note, we introduce the power-law model with an exponential cutoff to approximate the daily new infection data of COVID-19. We fit the data from 7 different countries, and gives the estimation of total infection, peak date and end date for each country and would wide. We expect the epidemic to end in April next year with a total number of 1.47e+07 infections. By comparing with the SIR model, we show that the power-line model has better performance in estimating the number of daily new infections and the total number of infections, which further illustrates the necessity of modeling the infection process from a network-based perspective.

References

[1] Zhou T, Liu Q, Yang Z, et al. Preliminary prediction of the basic reproduction number of the Wuhan novel coronavirus 2019-nCoV[J]. Journal of Evidence-Based Medicine, 2020.

[2] Kadushin C. Power, influence and social circles: A new methodology for studying opinion makers[J]. American Sociological Review, 1968: 685-699.

[3] Vazquez A. Polynomial growth in branching processes with diverging reproductive number[J]. Physical review letters, 2006, 96(3): 038702.

[4] Ziff A L, Ziff R M. Fractal kinetics of COVID-19 pandemic[J]. medRxiv, 2020.

[5] Maier B F, Brockmann D. Effective containment explains sub-exponential growth in confirmed cases of recent COVID-19 outbreak in Mainland China[J]. arXiv preprint arXiv:2002.07572, 2020.

[6] Li M, Chen J, Deng Y. Scaling features in the spreading of COVID-19[J]. arXiv preprint arXiv:2002.09199, 2020.