Graph Modelling on SEIR Dynamics

Oyeniyi Ridwan
200453228

MATH 802- Essay Writing
Department of Mathematics and Statistics
University of Regina, Canada.

1 Introduction

Planning and forecasting tasks are important for mitigating the sudden and potentially catastrophic impact of the infectious disease pandemic on society; however, it is not an easy task. During a pandemic, decisions are made with limited experience in a rapidly changing and uncertain situation which we observed during the novel coronavirus pandemic. The historically occurring pandemics have caused the death of 10 million people all over the world [31]. Although, currently, the existence of vaccines against some infectious diseases is reassuring, the cities and countries connected via air transportation facilitate the rapid transmission of viruses such as COVID-19. In particular, large number of individuals spend a relatively long period of incubation and hence, can transmit the disease to other people without knowing for 10-14 days when no symptoms are manifested [9]. In this regard, models are important tools for planning the pandemic and carrying out response measures (see, [7, 31]). Certainly, it is not possible to forecast the location or time of occurrence of next pandemic; however, models preserve large potential for increasing the effectiveness of response measures once pandemic occurs (see, [7, 13, 15]).

Modelling is one of the broadly used tools for the forecasting related to the epidemic situation in the society. At present, it is a topical problem to develop imitation models of complex systems such as the process of spread of an infection. The existence of a adequate mathematical model is paramount for obtaining an accurate forecast regarding the level of the spread of a disease and studying the process (see, [7, 23]). The modelling of pandemic processes can be viewed as an integral component of electronic medical demographics

system of the government and play a prominent role in prediction and effective decision making. Most epidemiological models are compartmental models, with the population divided into classes and with assumptions about the rate of transfer from one class to another. Example of such models is the $Susceptible - Exposed - Infectious - Removed/Recovery$ (SEIR) model describing the disease transmission and the rate of infected individuals. The model, and some of its applications is considered as a starting point for describing the spread of various diseases, e.g tuberculosis, measles, MERS and COVID-19. The basic idea of the SEIR model (see, [15, 24]), is to describe the number of infected and recovered individuals based on the number of contacts, probability of disease transmission, incubation period, recovery, and mortality rate. Since the model focuses on a very short time with respect to demographic dynamics, it postulates that births and natural (i.e, not connected with epidemics) while deaths balance each other.

In some cases, the behaviour of average quantities of longer period of time is sufficient to provide useful insight into the spread of the epidemics from the available data, the spatial component of many transmission systems has been recognised to be of pivotal importance. Due to this, spatially heterogeneous features must be included in the model to properly represent the transmission pattern. Then, a reasonable hypothesis about the phenomena may consider that the spatial aspects of transmission heavily influence the aggregation characteristic of the epidemic: we need hence to investigate data by using models that include such spatial connections. For example, the understanding of human mobility and the development of qualitative and quantitative theories is of key importance for the modelling and comprehension of human infectious disease dynamics, on geographical scales of different sizes. Also for the spread of infectious diseases in livestock comprehensive information on livestock movements, cattle movement, and contacts is required to devise appropriate disease control strategies. Understanding contact risk when herds mix extensively, and where different pathogens can be transmitted at different spatial and temporal scales, remains a major challenge [8]. For example, using data related to cattle movements and focusing on the geographical distribution of these movements is possible to improve the analysis of the spread of epizootic diseases [22]. To introduce spatial heterogeneity we consider metapopulation-based models, where the population is partitioned into large, spatially segregated sub-populations. A similar approach could be used in a more general way, irrespective of the biological interpretation: different ages, small interacting communities e.t.c (see, [21, 28, 29]). Understanding the dynamics of coupled systems on graphs mod-

elling connectivity in real-life systems can be quite challenging. On the other hand, one is often interested in the analysis of average quantities over large networks [18].

1.1 Basic Graph Theory

Graph theory was introduced and developed in recent years by Lov$\grave{a}$sz, Szegedy, Borgs, Chayes, S$\grave{o}$s, and Vesztergombi among others (see, [1, 2, 3, 4, 19]). A graph $G = (V, E)$ is a set of nodes (vertices) $V(G) = \{1, \ldots, n\}$, usually $n = |V(G)|$ where $|.|$ denotes the cardinality of a set and a set of edges $E(G) \subseteq V \times V$ between the vertices. The graphs will be simple, without loops or multiple edges, and finite unless otherwise specified. Weights (real numbers) will be given to the edges of a graph to make it an edge-weighted graph. Moreover, we assume that the graph is undirected, i.e., we identify the edges (i, j) and (j, i). Let $A = (a_{ij})$ be the adjacency matrix of a graph:

$$a_{ij} = \begin{cases} 1 & \text{if } (i, j) \text{ is an edge,} \\ 0 & \text{otherwise.} \end{cases}$$

Let G and H be graphs, a map ϕ from $V(H)$ to $V(G)$ is a homomorphism if it preserves edge adjacency, i.e if for every edge (i, j) in $E(H), (\phi(i), \phi(j))$ is an edge in $E(G)$. Denote by $hom(H, G)$ the number of homomorphisms from H to G. For example for H with one node and no edge $hom(H, G) = |V(G)|$, instead when H has two nodes and one edge $hom(H, G) = |E(G)|$, or when H has three vertices and $E(H) = \{(1; 2); (2; 3); (3; 1)\}, hom(H, G)$ is 6 times the number of triangles in G. Normalising by the total number of possible maps, we get the density of homomorphisms from H to G, $t(H, G) = \frac{hom((H,G)}{|V(G)|^{|V(H)|}}$. It is defined that a sequence of simple graphs $\{G\}_{n \in \mathbb{N}}$ is said to be convergent if G_n become more and more similar as n goes to infinity.

Definition1 : A sequence of graphs $\{G\}_{n \in \mathbb{N}}$ is said to be left convergent if the sequences $t(H, G_n)$ converges for $n \to \infty$ for every simple graph H.

Definition2 : A graphon is a bounded measurable functions $W : [0, 1]^2 \to \mathbb{R}$ that satisfy $W(x, y) = W(y, x)$ for all $x, y \in [0, 1]$.

2 Review of Related literature

The papaer [27] investigates various regression approaches for the modelling of the spread of COVID-19 and its impact on stock market. The forecast of the spread of coronavirus is studied with the application of logistic curve and

Bayesian regression. The impact of COVID-19 is studied with a regression approach and its impact is compared with the impact of other crises. The opportunity of the quantitative measurement of uncertainty in Bayesian regression can be useful information for experts while choosing models.

[20] proposes a conceptual model of COVID-19 epidemic in Wuhan by considering individual behavioral reactions and government actions, for example, the extension of leave, travel restriction, hospitalization and quarantine measures. The model has a simple structure, it can successfully predict the progress of COVID-19 epidemic and thus allows for understanding the spread trends. [17] focus on the analysis and prediction of the spread of COVID-19 pandemic. The analysis of existing data characterizing the epidemiological situation in Hubei shows that the error produced by the model is sufficiently small in comparison with official data. The study investigates the factors affecting the spread of COVID-19, for instance, the number of recovered individuals, the period of incubation and the average number of days of treatment.

[36] investigates the multiple routes transmitted epidemic process on multiplex networks. The authors propose detailed theoretical analysis to accurately calculate the epidemic threshold and outbreak size. The main outcomes of the study that the epidemic can spread across the multiplex network even if all the network layers are well below their respective epidemic thresholds. Noted that a strong positive degree correlation of nodes in a multiplex network could lead to a much lower epidemic threshold and a relatively smaller outbreak size.

In [5] proposed Bats-Hosts-Reservoir-People network model for the modelling of transmission potential from the source of infection (bats as assumed) to human. However, as the investigation of Bats-Hosts-Reservoir-People network is complex and public opinion is inclined towards the assumption that the virus is transmitted from seafood market (Huanan Seafood Wholesale Market) to human, the model is simplified and Reservoir-People (RP) transmission network is studied. Research results show that, compared to other severe respiratory syndromes, the transmission potential of COVID-19 is higher than Middle East Respiratory Syndrome (MERS), similar to severe acute respiratory syndrome (SARS), but lower than MERS syndrome observed in the Republic of Korea.

[37] the study is dedicated to estimating the unreported number of novel Coronavirus (2019-nCoV) cases in China in the first half of January 2020. On the based of the proposed approach modelled the epidemic curve of 2019-nCoV cases, in mainland China from 1 December 2019 to 24 January 2020 through the exponential growth. The number of unreported cases was determined by the maximum likelihood estimation. The author confirmed that the initial growth

4

phase followed an exponential growth pattern. As a result, noted that the reporting rate after 17 January 2020 was likely to have increased compared with the situation from 1 to 17 January 2020 on average, and it should be considered in future investigation.

[16] focused on projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. The study used estimates of seasonality, immunity, and cross-immunity for human coronavirus OC43 (HCoV-OC43) and HCoVHKU1 using time-series data from the United States to inform a model of SARS-CoV-2 transmission. The authors projected that recurrent wintertime outbreaks of SARS-CoV-2 will probably occur after the initial, most severe pandemic wave. The study used existing data to build a deterministic model of multiyear interactions between existing coronaviruses, with a focus on the United States, and used this to project the potential epidemic dynamics. The author reckons that the long-term dynamics of SARS-CoV-2 strongly depends on immune responses and immune cross-reactions between the coronaviruses, as well as the timing of introduction of the new virus into a population.

Different scenarios of the model will enable the assessment and forecast of detected and undetected cases of infection and death by considering the dynamics, time and location-based demographic characteristics (age, nationality, gender, profession, medical condition and etc.) starting from the moment of registration of the first infection case.

3 SEIR Models on Graphs

When investigating the transmission of infectious diseases, the analysis of the average behaviour of a large population is sufficient to provide useful insight and extract valuable information from the input data. However, the importance of the spatial component of many transmission systems is being increasingly recognised [28]. The main approaches for spatial models concern different scales: an individual-based simulation, a meta-population model, or a network model. Individual-based models explicitly represent every individual and usually assume a variable probability that any infectious host can infect any other susceptible host. Then the model should be able to account for the states of all N individuals in the population in an independent manner, and at the same time, it should allow for arbitrary interactions among them. The analysis of these models is a difficult task as the computational cost of numerical simulations is very onerous, and the extraction of the collective behaviours is very complex. Conversely, in the meta-population models the number of individu-

als at different space locations is in some states. These models often assume that each location is connected to others, with possible variable strengths of connection.

To describe a mathematical model for the spread of infectious disease, one has to make some assumptions about the disease transmission. We consider here, as a basic model, the SEIR compartmental model, where individuals are classified into different population groups based on the infection status. The model tracks the number of people in each of the following categories: *Susceptibles* (individuals that may become infected), *Exposed* (individuals that have been infected with a pathogen, but due to the pathogen incubation period, are not yet infectious), *Infectious* (individuals that are infected with a pathogen and may transmit it to others), and *Recovered* (individual that is either no longer infectious or has been "removed" from the population). For this report, we consider diseases with a latent phase during which the individual is infected but not yet infectious: a recent example of the application of this type of model is the description of the transmission of the COVID-19 disease [30].

The initial population N is subdivide into four classes, namely, $S(t)$ (susceptibles), $E(t)$ (exposed), $I(t)$ (infected-infectious), and $R(t)$ (recovered), where t is the time variable. The basic assumption for the scalar model is the homogeneously mixing population hypothesis which, roughly speaking, means that a given infectious individual may transmit the disease to any susceptible individual at the same rate. Also, one postulates that all the individuals in the population have the same chances of interacting with each other. People move from S to E based on the number of contacts with I individuals, multiplied by the probability of infection β, where $\beta I(t)/N$ is the average number of contacts with infection per unit time of one susceptible person. The other processes taking place at time t are: the exposed E become infectious I with a rate μ and the infectious recover R with a rate γ. Recovered individuals do not flow back into the S class, as lifelong immunity is postulated. The fractions $1/\mu$ and $1/\gamma$ are the average disease incubation and infectious periods, respectively. We assume that the total population remain constant, i.e., $S(t) + E(t) + I(t) + R(t) = N$. Next, we consider the rescaled variables, which for simplicity we do not relabel, $S(t)/N \to S(t), E(t)/N \to E(t), I(t)/N \to I(t), R(t)/N \to R(t)$. The system of

ordinary differential equations (ODEs) becomes:

$$\frac{dS}{dt} = -\beta S(t)I(t)$$
$$\frac{dE}{dt} = \beta S(t)I(t) - \mu E(t)$$
$$\frac{dI}{dt} = \mu E(t) - \gamma I(t)$$
$$\frac{dR}{dt} = \gamma I(t). \qquad (1)$$

A meta-population model and a network that is represented by a simple graph $G = (V, E)$ with n vertices (nodes, regions, patches) and m edges (connections) was considered. Each edge is described by a couple of nodes $(u, v), u, v \in V$. We order the nodes and label them with an integer index, and assume that the adjacency matrix of G is an irreducible matrix: there are no isolated and unreachable groups. In node j the corresponding sub-population possesses N_j individuals and $\sum_{j=1}^{n} N_j = N$. Also, individuals can move to a different node, interact with people in that node, and then return to the original one. The total amount of sub-population j that goes into node k and interacts with people in the node is denoted by a_{jk}. The matrix with entries a_{jk}, so that $\sum_{j=1}^{n} a_{jk} = N_j \quad j = 1, \ldots, n$. let P^{out} be the probability outing matrix with entries P_{jk}^{out} where the probability (percentage) that the sub-population j goes to node k is P_{jk}^{out}. In addition, we denote P^{in} the probability incoming matrix with entries P_{jk}^{in}, where P_{jk}^{in} is now the probability (percentage) of the sub-population in k that arrived from j. Let $M_j = \sum_{j=1}^{n} a_{jk}$ be the total amount of people arrived in node $j = 1, \ldots, n$, so that $= \sum_{j=1}^{n} M_j = N$. Then, for any $j = 1, \ldots, n$, $= \sum_{k=1}^{n} P_{jk}^{out} = \sum_{k=1}^{n} P_{jk}^{in} = 1$. Therefore, we have

$$A = Diag(N_1, N_2, \ldots, N_n)P^{out} = P^{in}Diag(M_1, M_2, \ldots, M_n),$$

where $Diag(x_1, x_2, \ldots, x_n)$ is the diagonal matrix with the vector $(x_1, x_2, \ldots, x_n)^T \in \mathbb{R}^n$ on the main diagonal.

Let $S_j(t), E_j(t), I_j(t)R_j(t)$ be the number of individuals in the node j at time t, $S_j(t) + E_j(t) + I_j(t) + R_j(t) = N_j$: we consider a time interval in which we can neglect demographics. Without any interaction with other nodes, within a deterministic approach of the compartmental models, with continuous time t, the epidemic dynamics can be described by the following system of differential

7

equations given below:

$$\frac{dS_j}{dt} = -\lambda S_j(t)$$
$$\frac{dE_j}{dt} = \lambda S_j(t) - \mu E_j(t)$$
$$\frac{dI_j}{dt} = \mu E_j(t) - \gamma I_j(t)$$
$$\frac{dR_j}{dt} = \gamma I_j(t), \tag{2}$$

where the parameter λ is the force of infection (i.e, the rate at which susceptible individuals become infected or exposed, and it is a function depending on the number of infectious individuals: it contains information about the interactions between individuals that concur with the infection transmission). Suppose that the population of N_j individuals mix at random, meaning that all pairs of individuals have the same probability of interacting, the force of infection may be computed as:

$$
\begin{aligned}
\lambda = \ & \text{transmission rate} \\
& \times \ \text{effective number of contacts per unit time} \\
& \times \ \text{proportion of contents infection} \\
& \approx \frac{\beta i_j}{N_j}, \tag{3}
\end{aligned}
$$

where β is the infectious rate. Then the system state,

$$\frac{dS_j}{dt} = -\beta \frac{i_j}{N_j} S_j(t)$$
$$\frac{dE_j}{dt} = \beta \frac{i_j}{N_j} - \mu E_j(t)$$
$$\frac{dI_j}{dt} = \mu E_j(t) - \gamma I_j(t)$$
$$\frac{dR_J}{dt} = \gamma I_j(t), \tag{4}$$

with the rescaled (percentage) quantities of susceptible, infected, removed at time t at the node j normalised to the number N_j of individuals associated to

the node j. Then, we obtain:

$$\frac{dS_j}{dt} = -\beta i_j(t) S_j(t)$$

$$\frac{dE_j}{dt} = \beta i_j(t) S_j(t) - \mu E_j(t)$$

$$\frac{dI_j}{dt} = \mu E_j(t) - \gamma I_j(t)$$

$$\frac{dR_j}{dt} = \gamma I_j(t), \tag{5}$$

where i_j denotes the derivative of the function i_j.

Now, we take a node j that is connected with the other nodes as encoded in matrix A. Then $S_j(t)$ can change due to the contribution of susceptible people that come from j reached an adjacent node k and met infectious people in that node. Then the contribution to $\frac{dS_j}{dt}$ due to the interactions in node k is given by the $P_{jk}^{out} S_j = a_{jk} S_j$ susceptible people that met a population in node k with a proportion of infectious people given by:

$$\frac{(\text{no. of infectious people in node k})}{(\text{no. of total people in node k})} = \frac{\sum_{l=1}^{n} P_{lk}^{out} I_i}{\sum_{l=1}^{n} a_{lk}} = \sum_{l=1}^{n} P_{lk}^{in} i_l.$$

Let the vector $X(t) = (x_1(t), x_2(t), \ldots, x_n(t))^T \in \mathbb{R}^n$, with $X = S, E, I, R$, the $SEIR$ model on the graph G is given below:

$$\frac{dS}{dt} = -\beta Diag(S(t)) BI(t)$$

$$\frac{dE}{dt} = \beta Diag(S(t)) BI(t) - \mu E(t)$$

$$\frac{dI}{dt} = \mu E(t) - \gamma I(t)$$

$$\frac{dR}{dt} = \gamma I(t), \tag{6}$$

where $B = P^{out} Diag(M_1, \cdots, M_n)^{-1} {P^{out}}^T$, and after rescaling (obtained by a premultiplication with $Diag(N_1, N_2, \cdots, N_n)^{-1}$ we have,

$$\frac{dS}{dt} = -\beta Diag(S(t)) Ai(t)$$

$$\frac{dE}{dt} = \beta Diag(S(t)) Ai(t) - \mu E(t)$$

$$\frac{dI}{dt} = \mu E(t) - \gamma I(t)$$

$$\frac{dR}{dt} = \gamma I(t), \tag{7}$$

where $A = P^{out}(P^{in})^T$.

Owing to the classical Cauchy Lipschitz Picard Lindelof Theorem, the Cauchy problem obtained by coupling system (7) with the initial data

$$S_j(0) = S_{j0}, I_j(0) = E_{j0}, I_j(0) = 1_{j0}, R_j(0) = R_{j0}, \text{ for every } j = 1, \ldots, n \quad (8)$$

has a local in-time solution. Since S_j, E_j, I_j, R_j represent percentages of individuals, the modelling range is then

$$0 \le S_j(t), E_j(t), I_j(t), R_j(t) \le 1, \quad S_j(t) + E_j(t) + I_j(t) + R_j(t) = 1 \quad (9)$$

for every $t \ge 0$ and $j = 1, \ldots, n$.

Lemma 1 : Assume that $0 \le S_j(t), E_j(t), I_j(t), R_j(t) \le 1$ and $S_j(t) + E_j(t) + I_j(t) + R_j(t) = 1$ for every $j = 1, \ldots, n$. Then, the solution of the Cauchy problem obtained by coupling (7) with (8) is global in time and satisfies [1].

Proof : To establish the second condition in (9), it suffices to observe that $\dot{S}_j(t) + \dot{E}_j(t) + \dot{I}_j(t) + \dot{R}_j(t) = 0$ for every $j = 1, \ldots, n$. We recall that, if the solution is bounded, then it can also be extended for every $t \ge 0$. Indeed, it suffices to show that $S_j(t), E_j(t), I_j(t), R_j(t) \ge 0$ for every t and every $j = 1, \ldots, n$. By the continuous dependence of solutions of ODEs on parameters, it suffices to establish the same property for the family of perturbed systems:

$$\frac{dS^e}{dt} = -\beta S_j^e(t) \sum_{k=1}^n a_{jk} i_k^e(t)$$

$$\frac{dE^e}{dt} = \beta S_j^e(t) \sum_{k=1}^n a_{jk} i_k^e(t) - \mu E_j^e(t) + \epsilon$$

$$\frac{dI}{dt} = \mu E_j^e(t) - \gamma I_j^e(t) + \epsilon$$

$$\frac{dR}{dt} = \gamma I_j^e(t), \quad (10)$$

where $\epsilon > 0$ is a parameter converging to 0^+. First, we observe that, if $S_j^e(t) = 0$ for some t, then $\dot{S}_j^e(t) = 0$. By the uniqueness part of the Cauchy Lipschitz Picard Lindelof Theorem, this implies that $S_j^e(t) \equiv 0$. If $S_j^e(0) = 0$ and hence that $S_j^e(t) > 0$ for every t if $S_{j0} > 0$. We set,

$$\bar{t}_e := min\{t \ge 0 : e_j^e(t) = 0 \text{ or } i_j^e(t) = 0 \text{ for some } j = 1, 2, \ldots, n\}$$

and we separately consider the following cases. If there is $j = 1, \ldots, n$ such that $e_j^e(\bar{t}_e) = 0$, then since $s_j^e(\bar{t}_e), i_j^e(\bar{t}_e) \ge 0$, we have $e_j^e(\bar{t}_e) \ge \epsilon$ and hence $\bar{t}_e = 0$ and $e_j^e > 0$ in a left neighborhood of 0. If there is $j = 1, \ldots, n$ such that $i^e(\bar{t}_e) = 0$, then $i^e(\bar{t}_e) \ge \epsilon$ and hence $\bar{t}_e = 0$ and $i_j^e > 0$ in a left neighborhood of 0. This

implies that $0 \leq s_j^e(t), e_j^e(t), i_j^e(t)$ for every $t \geq 0$. Since $\dot{r}_j^e \geq 0$, then $r_j^e(t) \geq 0$ for every $t \geq 0$.

Lemma 2 : Considering the SEIR system (9), the equilibrium points are $P_\infty = (S_\infty, 0, 0, R_\infty)$. Also, consider the Cauchy problem obtained by coupling (9) with (10). Under the same assumptions as in Lemma 1, very trajectory converges to an equilibrium point $P_\infty = (S_\infty, 0, 0, R_\infty)$ with $S_\infty^T 1 + R_\infty^T 1 = 1$.

Proof: Let (S, E, I, R) be an equilibrium point. From the fourth equation in (9), we conclude that $I_j = 0$ for every $j = 1, \ldots, n$. By plugging this equality in the third equation of (9), we conclude that $E_j =$ for every $j = 1, \ldots, n$. This implies that the equilibria are in the form $P_\infty = (S_\infty, 0, 0, R_\infty)$. We are left to prove the second part of the lemma. We recall Lemma 1 and by using the fourth equation in (9) we conclude that, for every $j = 1, \ldots, n$, the function R_j is monotone non-decreasing and hence has a limit as $t \to +\infty$. Also, the limit is confined between 0 and 1. By adding the third and the fourth equation in (9) we get that $I_j + R_j$ is monotone non-decreasing and also has a limit as $t \to +\infty$. We conclude that I_j has a limit as $t \to +\infty$. By using the first equation in (9) we infer that, for every $j = 1, \ldots, n$ S_j is monotone non-increasing and hence has a limit, which is confined between 0 and 1, as $t \to +\infty$. By adding the first and the second equation in (9) we get that, for every $j = 1, \ldots, n$ $S_j + E_j$ is monotone non-increasing and hence has a limit as $t \to +\infty$. We eventually conclude that E_j has a limit for $t \to +\infty$. Since the limit point is finite, it must be an equilibrium. The condition $S_\infty^T 1 + R_\infty^T 1 = 1$ follows from the equality $S_j(t) + E_j(t) + I_j(t) + R_j(t) = 1$ for every $j = 1, \ldots, n$ and $t \geq 0$.

3.1 Factor affecting the spread of Infectious diseases and main characteristics

Researchers utilize models in order to assess important epidemiological characteristics of the disease such as the impact of public health interventions including the period of incubation, ability to infect, asymptomaticity and severity as well as social distancing, medical examination at airports, travel restrictions and contact tracing [23]. It is to be noted that existing models can be considered as an important tool for understanding the disease, implementing response measures and political decisions; however, at the same time, this leads to disagreements as their approaches and proposed assumptions differ from each other significantly (see, [11, 12]). Most people are vulnerable to the infection at early stages of epidemic and the spread of this disease from human to human can be modelled as a stochastic "branching process". If an infected person on average infects

two individuals, the number of infected individuals doubles at each stage and this process grows exponentially.

It is evident that an infected person does not necessarily infect others. Several factors affect the probability of being infected. In case of a pandemic, the speed of infection depends on average number of people one person can infect and the time needed for these people to become contagious. Hence, it is important to determine the main dynamics of transmission for accurate prediction of COVID-19. According to WHO, following factors affect the spread of COVID-19 [35]:

- How many individuals does one person infect on average? (According to WHO information, "reproduction number" is considered to be 1.4e2.5 currently)

- The period of infection of a person in the same environment with the infected? (15 min or more and less than 2 m distance [9])

- What is the period after infection before the symptoms start to be manifested? ("incubation period" is assumed to be approximately 5.1 days)

- What is the ratio of spread before symptoms start (if present)?

The collection of preliminary data enables precise predictions regarding COVID-19 progress in order to take into consideration mentioned factors and known specific characteristics in the model.

Currently, several applications are being used in China, South Korea, Singapore, Israel and other countries to monitor social distancing among preventive measures. Considering that practically everyone owns smartphones, it is possible to detect the location of these devices using geolocation systems. Based on the data mining methods, a warning signal can be sent regarding the maintenance of social distance and undesirable approach distance when needed by monitoring the distance between individuals in densely populated areas. In particular, it is possible to determine risk groups and maintain strict monitoring between them. At present, contact tracing applications are used by governments in several countries ([26]; TraceTogether; Stopp Corona APP; QR health code; eRouska [25]; COVID Community Alert). Among those,Stopp Corona (Austria), Alipay Health Code (Chine), eRouska (Czechia), StopCovid (France), COVID Community Alert (Italy), TraceTogether (Singapore), COVID-19 Apple / Google App (US) and etc. Apple and Google Companies state that it is possible to apply contact tracing systems based on joint Bluetooth Low Energy technology and privacy-preserving cryptography [26]; Apple & Google partner

on COVID-19). For instance, TraceTogether app is used in Singapore in order to obtain information via smartphone regarding possible contact with patients infected with coronavirus (TraceTogether). According to criterion of the ministry of health of Singapore, a "close contact" implies a distance less than 2 m for 30 min. The ministry of health of Israel checks the presence of risk of infection with coronavirus in case of being in proximity of or in contact with a virus host with "HaMagen" app developed for mobile phones (HaMagen). Two different methods are applied in order to measure the distance between individuals via smartphones. One of those is the use of geolocation information utilizing GPS (Global Positioning System) or GNSS (Global Navigation Satellite System) systems - in this case, the positioning accuracy is 0.6 m and can reach 10 cm with the time [6]. But the loading of local network or unsustainability and unreliability of the system in indoor locations are considered to be main drawbacks (see, [6, 32]). The measurement of distance with Wi-Fi and Bluetooth signals is characterized with large measurement error in close distances. From this point of view, the use of ultra-wideband (UWB) radio signals can significantly increase the measurement accuracy in close distances (5e10 cm) [6]. At present, UWB technology is preferred for the development of devices determining the location in specific indoor locations. The utilization of this technology will play in important role in measuring social distance and determining the risk of infection.

4   Spectrum of the SEIR model on graphs

Despite the presence of various approached to the modelling of the spread of infectious diseases, it is known that an infectious agent is spread via interaction between individuals as in social network. Formally, social networks can be described with graph and in this case, the edges of the graph denote individuals and arrows denote the set of interactions. In this regard, graph models can be utilized while modelling and visualizing the pandemic processes. Graph models play an important role in forecasting the disease spread and supporting decision making. In particular, the use of colored graphs enables the visualization of a pandemic process and can be paramount for forecasting, assessment of the effectiveness of pandemic processes, implementation of proactive processes and decision making. Alongside the hypotheses proposed in existing approaches, uncertainty regarding epidemic characteristics is one of the drawbacks of these models.

In this section, we always assume that all the components of $S(t)$ are strictly

positive. Based on the proof of Lemma 1, this amounts to assuming that they are strictly positive at $t = 0$. The coefficient matrix $B(t) := Diag(S(t))A$ of the SEIR model on a general simple graph, can be interpreted as a weighted adjacency matrix, where the weight associated with each edge is $b_{ij}(t) := S_i(t)a_{ij}$; the weighted adjacency matrix is not symmetric and the weighted graph is directed. To study the spectrum of $B(t)$ and eventually work with a symmetric weighted adjacency matrix, we introduce its "symmetrisation"

$$B_{sym}(t) := Diag(S(t))^{\frac{1}{2}} A Diag(S(t))^{\frac{1}{2}},$$

where the weight associated with the edge (i, j) $b_{ij}^{sym}(t) : S_i^{\frac{1}{2}}(t)a_{ij}S_j^{\frac{1}{2}}(t)$. Furthermore, the matrices $B(t)$ and $B_{sym}(t)$ have the same eigenvalues; in fact

$$B(t) - \nu I = Diag(S(t))^{\frac{1}{2}} \left(Diag(S(t))^{\frac{1}{2}} A Diag(S(t))^{\frac{1}{2}} - \nu I \right) Diag(S(t))^{-\frac{1}{2}}$$

$$= Diag(S(t))^{\frac{1}{2}} (B_{sym}(t) - \nu I) Diag(S(t))^{-\frac{1}{2}}.$$

Let v be an eigenvector of $b(t)$ with eigenvalue ν. Then

$$(t)v = \nu v \Leftrightarrow Diag(S(t))^{-\frac{1}{2}} Bv = \nu Diag(S(t))^{-\frac{1}{2}} v = \Leftrightarrow y := Diag(S(t))^{-\frac{1}{2}} v,$$

$$Diag(S(t))^{-\frac{1}{2}} B Diag(S(t))^{\frac{1}{2}} y = \nu y \Leftrightarrow B_{sym}(t)v = \nu v$$

This shows that $y = Diag(S(t))^{-\frac{1}{2}} v$ is an eigenvector of $B_{sym}(t)$ with eigenvalue ν i.e, the the spectrum of $B(t)$ is real and it uniquely identifies the spectrum of $B_{sym(t)}$, and vice versa. The eigenvectors $y_i = Diag(S(t))^{-\frac{1}{2}} v_i, i = 1, \cdots, n$ are not orthonormal.

Based on these properties, SEIR and symmetric SEIR models on a general simple graph can be reduced (the vector r can be reconstructed from vectors s, v_i, e).

$$\begin{cases} \dot{S} & = -\beta Diag(S(t))A_i = \beta B(t)i \\ \dot{E} & = \beta B(t)i - \mu E \\ \dot{I} & = \mu E - \gamma i \end{cases} , \quad \begin{cases} \dot{S} & = -\beta B_{sym}(t)i \\ \dot{E} & = \beta B_{sym}(t)i - \mu E \\ \dot{I} & = \mu E - \gamma i. \end{cases}$$
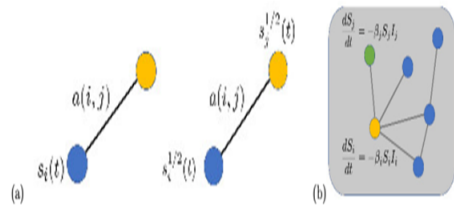
Figure 1

(a) Non-symmetric and Symmetric weights for the SEIR models on a graph.
(b) definition of the SEIR model on graphs starting from the governing differential equations of 1D SEIR model at each node of the input graph.
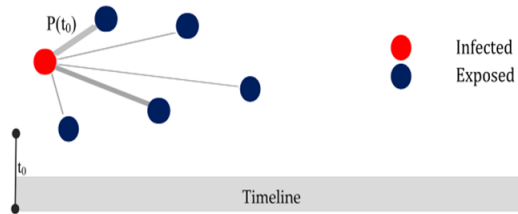

Figure 2

First case of infection is recorded at t_0 time point and the coloured graph of individuals in contact with infected person via flourish visualization tool.
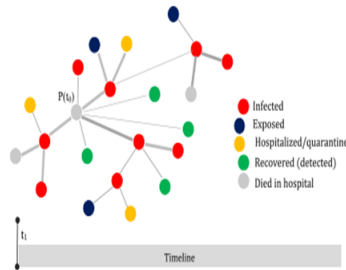

Figure 3

It is possible to visualize the graph model at t_1 time period by tracking the process of human-human transmission of the virus. Only confirmed case of infection are included in the graph model.

5 Conclusion

SEIR models defined on graphs were considered and their basic features such as the spectral properties of the weighted adjacency matrix were discussed. This analysis provides a useful step in understanding the spread of diseases in populations with complex social structures. The development of metapopulation models able to capture the spatial population structure, the development of computationally efficient methods for calculating key epidemiological model quantities, and the integration of within- and between-host dynamics in models. Despite the existence of large number of proposed models, the majority of them do not take into account the epidemiological characteristics of COVID-19 such as social distancing, duration of contact with a patient as well as important factors affecting the scale of infection. Currently, technologies for monitoring

the maintenance of social distance, contact tracing apps adopted by several governments are used as preventive measures among others.

The governments, non-governmental organizations, experts and epidemiologists attempt to utilize models in order to understand how to respond to, fight and treat the pandemic. The main challenges are the substantial differences between models and the use of different data sets and methodologies in each country. Here is a need for developing realistic models considering the human behavior for the purpose of fighting COVID-19 pandemic at global level and preventing future cases of pandemic. It must be noted that the application of technologies will play an important role in measuring the social distance and determining the risk of infection and also the analysis of the factors affecting the number of infections will allows for developing more realistic models in future research. Moreover, in some applications, it is necessary to estimate an underlying graphon to perform some network analysis. Some non parametric estimation methods have been proposed, and some are provably consistent. However, if certain useful features of the nodes (e.g age, social group, health information) are available, it may possible to incorporate this source of information to help with the estimation, using both the adjacency matrix and node features.

References

[1] C. Borgs, J.T. Chayes, H. Cohn, Y. Zhao, (2018). "An Lp theory of sparse graph convergence II: LD convergence, quotients, and right convergence." *Annals of Prob.* 45, 337-396.

[2] C. Borgs, J.T. Chayes, H. Cohn, Y. Zhao, (2019). "An Lp theory of sparse graph convergence I: Limits, sparse random graph models, and power law distributions. "*Trans. Am. Math. Soc.* 372, 3019-3062.

[3] C. Borgs, J.T. Chayes, L. Lovàsz, V.T. Sòs, K. Vesztergombi (2006). "Counting graph homomorphisms." *Top. Discr. Math. Ser.Algor. Combin* 26, 315-371.

[4] C. Borgs, J.T. Chayes, L. Lovàsz, V.T. Sòs, K. Vesztergombi (2012). "Convergent sequences of dense graphs II, Multiway cuts and statistical physics." *Ann. Math* 176, 151-219.

[5] Chen, T.-M., Rui, J., Wang, Q.-P., Zhao, Z.-Y., et al. (2020). "A mathematical model for simulating the phase-based transmissibility of a novel coronavirus." *Infectious Diseases of Poverty* 9. https://doi.org/10.1186 s40249-020-00640-3.

[6] Connell, C. (2015). "What's the difference between measuring location by UWB, Wi-Fi, and Bluetooth?" www.electronicdesign.com/technologies/communications/article/21800581/whats-the-difference-between-measuring-location-by-uwb-wifi-and bluetooth.

[7] Currie, C. S. M., Fowler, J. W., Kotiadis, K., Monks, T., et al. (2020). "How simulation modelling can help reduce the impact of COVID-19." *Journal of Simulation* https://doi.org/10.1080/17477778.2020.1751570.

[8] D. Ekwem, T.A. Morrison, R. Reeve, R. et al.,(2021). "Livestock movement informs the risk of disease spread in traditional production systems in East Africa." *Sci Rep* 11 16375 (2021), DOI:10.1038/s41598-021-95706-z.

[9] ECDC Technical Report. (2020). "Contact tracing: Public health management of persons, including healthcare workers, having had contact with COVID-19 cases in the European Union." *Technical Report* https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-public-health-management-contact-novel-coronavirus-cases-EU.pdf.

[10] F. Brauer,(2017). "Mathematical epidemiology: past, present, and future." *Infect Dis Model.* 2, 113-27.

[11] Gleick, P. H. (2020). "No COVID-19 models are perfect, but some are useful. " https://time.com/5838335/covid-19-prediction-models/.

[12] Holmdahl, I., & Buckee, C. (2020). "Wrong but useful-what Covid-19 epidemiologic models can and cannot tell us." https://www.nejm.org/doi/full/10.1056/ NEJMp2016822.

[13] Holmes, E. C., Rambaut, A., & Andersen, K. G. (2018). "Pandemics: Spend on surveillance, not prediction. Nature." *558(7709)* 180-182. https://doi.org/10.1038/ d41586-018-05373-w.

[14] H.W. Hethcote (2000). "The mathematics of infectious diseases" *SIAM Rev* 42, 599-653.

[15] Kerkhove, M. D. V., & Ferguson, N. M. (2012). "Epidemic and intervention modellingea scientific rationale for policy decisions? Lessons from the 2009 influenza pandemic." *Bulletin of the World Health Organization* 90(4), 306-310. https://doi.org/10.2471/BLT.11.097949.

[16] Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. G., & Lipsitch, M. (2020). "Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period." *Science* 368(6493), 860-868.

[17] Li, L., Yang, Z., Dang, Z., et al.(2020). "Propagation analysis and prediction of the COVID-19." *Infectious Disease Modelling* 5, 282-292.

[18] L. Lovàsz, B. Szegedy (2007). "Szemerèdys̀ lemma for analyst." *Geom. Funct. Anal* 17, 252-270.

[19] L. Lovàsz Large (2012). "Networks and graph limits." *American Mathematical Society.*

[20] Lin, Q., Zhao, S., Gao, D., Lou, Y., et al.(2020). "A conceptual model for the coronavirus disease 2019 (COVID-19) outbreak in Wuhan, China with individual reaction and governmental action." *International Journal of Infectious Diseases* 93, 211-216.

[21] M.E.J. Newman, (2002). "Spread of epidemic disease on networks." *Phys. Rev. E 66* 016128.

[22] M. Hässig, A.B. Meier, U. Braun, B. Urech Hässig, R. Schmidt, F. Lewis,(2015). "Cattle movement as a risk factor for epidemics." *Schweizer Archiv fur Tierheilkunde* 157(8), 441-448 DOI:10.17236/sat00029.

[23] Michaud, J., Kates, J., & Levitt, L. (2020). "COVID-19 models: Can they tell us what we want to know?" https://www.kff.org/coronavirus-policy-watch/covid-19- models.

[24] M.J. Keeling, P. Rohani, (2008). "Modeling Infectious Diseases in Humans and Animals." *Princeton University Press* Princeton, NJ.

[25] Morrow, A. (2020). "France to test controversial Covid-19 tracking app during lockdown exit." www.rfi.fr/en/science-and-technology.

[26] Panzarino, M. (2020). "Apple and Google are launching a joint COVID-19 tracing tool for iOS and Android. TechCrunch." https://techcrunch.com/2020/04/10/ apple-and-google-are-launching-a-joint-covid-19-tracing-tool/.

[27] Pavlyshenko, B. M. (2020). "Regression approach for modeling COVID-19 spread and its impact on stock market (Preprint)." https://arxiv.org/pdf/2004.01489.

[28] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani (2015). "Epidemic Processes in Complex Networks." *Rev. Mod. Phys* 87, 925.

[29] S. Eubank, H. Guclu, V.A. Kumar, M.V. Marathe, A. Srinivasan, Z. Toroczkai, N. Wang (2004). "Modelling disease outbreaks in realistic urban social networks." *Nature* 429(6988), 180-184.

[30] S. He, Y. Peng, K. Sun,(2020). "SEIR modelling of the COVID-19 and its dynamics." *Nonlinear Dyn* 101, 1667-1680.

[31] Shearer, F. M., Moss, R., McVernon, J., Ross, J. V., & McCaw, J. M. (2020). "Infectious disease pandemic planning and response: Incorporating decision analysis." *PLoS Medicine* 17(1). https://doi.org/10.1371/journal.pmed.1003018.

[32] Slamich, M. (2020). "Bluetooth vs ultra-wideband: Which indoor location system?" https://blog.pointr.tech/bluetooth-vs-ultra-wideband-which-technologyto-use-for-indoor-location.

[33] P. Blanchard, D. Volchenkov,(2011). "Random Walks and Diffusions on Graphs and Databases:An Introduction." *Springer-Verlag Berlin Heidelberg.*

[34] Panzarino, M. (2020). "Apple and Google are launching a joint COVID-19 tracing tool for iOS and Android. TechCrunch." https://techcrunch.com/2020/04/10 apple-and-google-are-launching-a-joint-covid-19-tracing-tool/.

[35] WHO IHR. (2020). "Statement on the meeting of the international health regulations (2005) emergency committee regarding the outbreak of novel coronavirus (2019-nCoV). " https://www.who.int/news-room/detail/23-01-2020-statement-on-the-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov).

[36] Zhao, D., Li, L., Peng, H., & Yang, Y. (2014). "Multiple routes transmitted epidemics on multiplex networks. "*Physics Letters A* 378(10), 770-776.

[37] Zhao, S., Musa, S. S., Lin, Q., Ran, J., Yan, G., et al. (2020). "Estimating the unreported number of novel coronavirus (2019-nCoV) cases in China in the first half of January 2020: A data-Driven modelling analysis of the early outbreak." *Journal of Clinical Medicine,* 9(2), 388. https://doi.org/10.3390/jcm902038810.3390/jcm9020388.