Bridging Compartmental Models and Network Analysis in Epidemiological Modelling

Jorge M Mendes1,2,*; Helena Baptista1; Ying C MacNab2
1NOVA Information Management School (NOVA IMS), New University of Lisbon, Campolide Campus, 1070-312 Lisbon, Portugal.
2Epidemiology and Biostatistics, School of Population and Public Health, University of British Columbia Vancouver, Canada.
3NOVA Cairo, The Knowledge Hub Universities, New Administrative Capital, Cairo, Egypt.

Abstract

In this study, the authors undertake a comparative analysis of compartmental models and network analysis as means of simulating the propagation of infectious diseases. Compartmental models operate under the assumption of homogeneous mixing, which is often a flawed assumption as individuals tend to engage in varied contact patterns based on their surrounding environments. On the other hand, social network analysis accounts for the intricate web of interpersonal connections between individuals, thereby offering a more realistic portrayal of social behaviour. However, network analysis can be computationally demanding, rendering its application in real-time epidemic modelling challenging. By conducting a series of simulations utilising the SIR model and network analysis, the authors accentuate the merits of a hybrid modelling approach that integrates the strengths of both compartmental models and network analysis while mitigating their respective limitations. Additionally, the authors suggest plausible avenues for future research.

Background

Epidemiological modelling has played a crucial role in comprehending and managing the transmission of infectious diseases (Fitzpatrick et al. (2019), Gopal et al. (2021)). One of the primary strategies for tracking the course of an epidemic is compartmental modelling, which has been enhanced (e.g. Hethcote (2000), Abou-Ismail (2020), Mendes & Coelho (2021)) since the pioneering work of Kermack and McKendrick in 1927 (Kermack & McKendrick (1927), Kermack & McKendrick (1932), Kermack & McKendrick (1933)). Another approach is network analysis (Kiss et al. (2017)), which has been used to highlight specific disease spread mechanisms (e.g. Eames & Keeling (2002)). Both methods provide unique perspectives and limitations for simulating disease spread, and this study explores these methodologies in brief, highlighting their differences, limitations, and potential integrations for more accurate and computationally efficient simulations.

Compartmental models, such as the SIR model, are widely used to study the spread of infectious diseases. These models divide populations into distinct compartments, such as susceptible, infectious, and removed/recovered individuals, and rely on differential equations to capture the movement of individuals between these compartments. However, recent research has been using stochastic approaches (e.g. Ward et al. (2022) and Mendes & Coelho (2023)) for modelling the spread of diseases in real-world settings. The main assumption of compartmental models is homogeneous mixing, which assumes that individuals have equal contact probabilities with one another, regardless of their location or social context. This assumption is often unrealistic, as individuals have different contact patterns according to the environments they move around. For example, individuals in larger areas are less likely to interact with those who are far away (Anderson & May (1992)). In contrast, social network analysis takes into account the complexity of interpersonal connections between individuals (Figure 1). Each node in a social network can be in different states, and transition probabilities are influenced by the number and nature of contacts with infectious individuals. The homogeneous network model assumes equal connections for each node, while the hetero-
Consider a time discrete version of the SIR model given by the following equations:

\[ S(t) = S(t - 1) - dI, \]
\[ I(t) = I(t - 1) + dI - dR, \]
\[ R(t) = R(t - 1) + dR, \]
\[ dI \sim \text{Bin}(S(t - 1)p_{\text{in}}, t = 1,\ldots,T), \]
\[ N = S(t) + I(t) + R(t), \]
\[ S(0) = N - i_0, I(0) = i_0, R(0) = 0, p_{\text{in}} = 1 - \exp(-\beta(t - 1)), t = 1,\ldots,T, p_{\text{dr}} = 1 - \exp(-\gamma(t - 1)), t = 1,\ldots,T. \]

We used a population size of \( N = 1,000 \) to compare the SIR baseline model with network models, which was sufficient to convey our points and simultaneously computationally feasible. The epidemic began with 10 infectious individuals, \( \beta = 0.0003, \gamma = 1/10 \) (10-day infectious period), and \( t = 0.110 \). To account for uncertainty, we ran the process 1,000 times, generating 1,000 replicates. The evolution of compartment \( I(t) \) is shown in Figure 1 (Baseline SIR), with the blue line representing the mean of the replicates and the light-blue shaded area indicating the 95% credible interval. We then transformed the SIR model into a network with 1,000 nodes and 50 connections each, ensuring homogeneity and reflecting the comprehensive nature of the network of potential links in the SIR model. The infection dynamics with 1,000 nodes and 50 links were identical to those in the SIR model without a network structure (SIR Network) after calibration of \( \beta \) to obtain a similar prevalence curve. The red line in Figure 1 approaches the blue line due to the small population size, falling within the 95% credible interval.

The second scenario (shown in Figures 2 and 3) as network model I and II assumes there are two clusters. One, the Urbans, has 800 nodes highly connected (85 connection each), and the second, the Rurals, has 200 nodes with 8 connections each, keeping the same average number of nodes, that is 50. Some of the Rurals are connected to nodes in the Urban cluster. We tried two distinct scenarios only differing on the location of the initial infections. Then we assume that the infection starts with five infectious Rurals and five infectious Urbans (Network model II). They are not very different. Indeed, the epidemic spreads at a slightly slower rate in this scenario compared to the previous one, which is expected due to the fewer number of connections among the Rurals. However, the rapid spread among the Urbans more than compensates for this, leading to a faster spread overall and a similar spike of the outbreak, although it ends earlier than the SIR network. In the long term, the infected population does not reach the levels of the first scenario due to the slower evolution of the outbreaks caused by the limited connections among the Rurals and the fast reduction of number of susceptible nodes in the Urbans cluster.
Integrating the two approaches: Frequently, when we have two infeasible extremes, the ideal solution lies in between. Here, using a hybrid approach to combine the strengths of both compartmental models and network analysis and mitigate their limitations can be achieved by:

1. Downscaling compartmental models to smaller geographical areas, for multi-compartment models, with homogeneous mixing within but heterogeneous mixing among them. Spatial neighbourhood and similarity matrices (e.g. Besag et al. (1991), MacNab (2022), MacNab (2023), Baptista et al. (2016), Baptista et al. (2020)), as used in disease mapping, can be incorporated to capture the regional dynamics of disease spread, considering factors such as population density and mobility patterns (mobile-generated data may be used in establishing dynamic mobility patterns).

2. Incorporating stochastic, time-varying heterogeneous contact matrices (Mendes & Coelho (2023)) based on localisation, age, and sex may help on assess the impact of public health measures (e.g. gathering bans, travel restrictions, lockdowns, etc.) across different demographics and geographical areas.

3. Incorporating the impact of various pathogenic variants on the infection mechanism model. Additionally, the infection mechanism should account for spatial variation in infection probability through spatial processes, such as spatial random fields or Markov random fields (e.g. Banerjee et al. (2015), MacNab (2023), Baptista et al. (2020)).

4. Incorporating, the impact and effectiveness of public health measures targeting either the reduction of probability of being infected, given a contact with an infectious individual (e.g. facial coverings) or the reduction of contacts among individuals (e.g. gathering bans) (e.g. Mendes & Coelho (2023)).

5. Incorporating additional compartments (e.g. Mendes & Coelho (2023)) to account for differential recovery rates, disease severity, and the consequent effects on healthcare resource utilisation, including hospitalisation and fatality rates.

Conclusion and future directions

As we illustrated herein, simulations can provide valuable insights into the dynamics of disease spread under different scenarios. These simulations will not only reinforce the theoretical aspects discussed but also offer practical examples of how epidemiological models can be adapted and applied to real-world scenarios to gain a deeper understanding of the disease dynamics.

The fusion of compartmental models and social network analysis, as well as incorporation of spatial neighbourhoood and/or similarity-matrices commonly seen in disease mapping, offers a promising pathway for more accurate and efficient epidemiological modelling. This integrative approach accommodates, as much as possible, the complexity of real-world social interactions and geographical variations as described, enhancing our ability to predict and manage infectious disease outbreaks.

Our future research shall focus on refining these hybrid models, exploring the impact of various social behaviours and public health interventions. Additionally, leveraging advancements in computational power and algorithmic efficiency will be crucial in realising the full potential of these sophisticated epidemiological models.

References


