First Annual Progress Report

Ethan Kelly

May 12, 2021

Ethan Kelly
e.kelly.1@research.gla.ac.uk
Principal supervisor: Jessica Enright
Jessica.Enright@glasgow.ac.uk
Second supervisor: Alice Miller
Alice.Miller@glasgow.ac.uk

1 Introduction

This Annual Progress Report, the first of my PhD programme, serves to detail the work I have done in the first year of my research. I will begin by discussing the literature I have examined in that year, with particular focus on some key papers from which I have spent a lot of time understanding and extending results. This literature review will provide background to the research I am carrying out, with specific attention to any areas I have found to be lacking. In particular, I aim to show how the disciplinary-specific approaches to disease modelling have created a significant gap in an important aspect of the field.

Then, I will move on to discuss the work I have done so far.

2 Literature Review

2.1 Overview of the literature

2.1.1 Firefighter

The following is a decision formulation given by Finbow and MacGillivray for Firefighter on a tree [?]:

Firefighter
Instance: A rooted graph (G, r) and an integer $k \geq 1$.
Question: Is there a finite sequence $d_1, d_2, \ldots d_t$ of vertices of the graph G such that:

i d_i is neither burned nor defended at time i,

ii At time t, no undefended vertex is adjacent to a burning vertex, and

iii At least k vertices are saved at the end of time t?

2.2 Markovian SIR epidemics on networks

In the past year, I have spent significant time understanding, applying and expanding the results from the paper "Exact deterministic representation of Markovian SIR epidemics on networks with and without loops," [?] so I will spend some time providing a review of this work and its utility. I will go on to detail the work I have done in extending the results from this paper in the following section.

In [?], the authors note that there are generally three approaches to compartmental models of disease. We can:

1. Take averages at population level,
2. Maintain a probabilistic view by considering the full state space, or
3. Begin modelling at the level of vertices and build up to larger structures from there.

The latter of these three approaches is the one used by the authors: begin by considering equations for single vertices, then consider dependencies on pairs, then triples and so on until we reach the full system size. Such equations are well defined and consistent, which is not difficult to see.

The work presented has two aims. Firstly, to provide an exact, deterministic representations of Markovian SIR epidemics on graphs with and without loops. Secondly, to identify a link between the structural properties of the graphs and the viability of closures that can be used to write down exact systems of equations that can be numerically evaluated. In particular, the authors show this structural link is founded on cut-vertices and bridges. Cut-vertices are vertices that, if removed from a connected graph, result in the formation of two (or more) disconnected sub-graphs. Bridges are edges that lie between two cut-vertices.

The authors spend significant time expanding intuitions on the identification of closures, which allow us to approximate or even exactly specify higher-order moments in terms of lower-order moments. They claim this is well known to be feasible for tree-like graphs and for graphs with loops starting from some specific initial conditions. They present some examples to develop the intuition that "loops cannot be closed by breaking them down to their component parts."

The main result of the work reveals an important relation between structure of the graph used in the epidemic model and types of closures that are feasible using cut-vertices and bridges.

They also prove an impressively general result: if a graph with N vertices and E edges has T triangles and no larger loops than size 3 (meaning also that triangles cannot have overlapping edges), an upper bound on the size of the system of equations describing the system dynamics can be calculated:

$$2N + 3E + 7T \leq 10N$$

The authors also provide a "recipe-like" approach to establish the feasibility of writing down an exact representation for a given graph even more generally. They use this to provide an upper bound for the number of equations required to describe epidemic dynamics exactly:

$$N_{EQ}(G) = \sum_{i=1}^{P} m_i f_i - 2 \sum_{j=1}^{L} (\mathrm{Ind}(v_{i_j}) - 1).$$

where P is the number of distinct sub-graphs produced when the original graph is spliced into independent sub-graphs through cut-vertices, m_i represents the number of equations required to describe the corresponding sub-graph i, f_i is the frequency or count of the sub-graph G_i and $\mathrm{Ind}(v_{i_j})$ is the number of sub-graphs to which the cut-vertex v_{i_j} belongs.

This takes a sum across the number of equations for all sub-graphs and adjusts to account for unnecessary multiplications caused by cut-vertices being part of multiple sub-graphs, which is a move made in illustrative examples throughout the work.

3 Work done so far

3.1 Compartmental Models

3.1.1 Agency-oriented modelling

One of the key deficiencies in existing approaches to compartmental models of disease, as described in Section 2, is the lack of agency in individuals. To address this, I have designed and implemented (in Java) a compartmental graph model of infectious disease.

In my implemented model, we begin with a graph. This graph has a particular number of vertices and edges and can be generated using code I have written for the following random graph types:

- Simple (unweighted, undirected, no loops or multiple edges)
- Erdős Rényi (edge between two vertices exists with given probability)
- Complete (every pair of vertices connected by single unique edge)
- Bipartite (vertices can be divided into two, disjoint, independent sets)
- Complete bipartite (every pair in each set connected by single unique edge)
- Path (graph containing a single path through each vertex)
- Binary tree (each vertex has at most three children)
- Cycle (graph containing only a single cycle through all vertices)
- Eulerian Path (graph containing only a single Eulerian path through all vertices)
- Eulerian Cycle (graph containing only a single Eulerian cycle through all vertices)
- Wheel (a single vertex connected to every other vertex in a cycle)
- Star (a tree with a single internal vertex and a given number of leaves)
- Regular (regularly uniform k-regular graph)
- Tree (uniformly random tree, generated using Prüfer sequence)
- A Barabási–Albert-generated preferential attachment graph (can be non-linear)
- A Watts-Strogatz model generated small-world graph.

3.1.2 Equations describing graph-based compartmental models

4 Thesis Statement

5 Research Plan

5.1 Goals

5.2 Resources required

5.3 Measures of success

5.4 Timeline

Ethan Kelly
e.kelly.1@research.gla.ac.uk
Principal supervisor: Jessica Enright
Jessica.Enright@glasgow.ac.uk
Second supervisor: Alice Miller
Alice.Miller@glasgow.ac.uk

Department of Computing Science
Sir Alwyn Williams Building
University of Glasgow
G12 8QN