First Annual Progression Report

Ethan Kelly

Wednesday 9th June, 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

Contents

1 Introduction

This Annual Progress Report, the first of our PhD programme, serves to detail the work we have done in the first year of our research. We first outline the research title we have worked under and outline the structure of the discussion to follow.

1.1 Thesis Statement

Our working research title is *game-theoretic and probabilistic methods applied to spatial network models of contagion.* Over the past year, we have spent a large portion of time reading and developing ideas related to graph models of contagion, such as FIREFIGHTER and similar games on graphs that are already used to model contagion (or could fit in such a context easily). We have also explored game theory to understand optimal strategies and probabilistic methods in order to adapt the games we have studied to stochastic contexts. Such a broad approach has meant that we have a wide range of tools available to us in understanding and developing models of disease and other contagions.

1.2 Report Roadmap

Our research over the past year has focused into two main strands: a formal, mathematical approach and an experimental approach. Our motivation for this two-pronged *modus operandi* is to ultimately develop a model for disease spread formally, using dynamical system and graph theoretic approaches, which we then use an experimental implementation to algorithmically explore different approaches to the model, for instance in examining candidates for theoretically optimal public health strategies in the context of a disease. To explain this work in detail, we will begin this report by discussing the literature examined in the past year, with particular focus on any papers that have proved foundational to our own work. This literature review section will provide background to our research so far, which is discussed in the section that follows it. We will then outline the direction we hope to take this initial work moving forward and detail milestones we hope to reach in the coming year.

2 Literature Review

2.1 Context

To better understand the relevance of the following literature review, we first provide a brief overview of where the concepts to follow fit into our work so far, beginning with the Firefighter Problem and how it can be adapted to a rudimentary model of disease. A huge amount of our research (particularly the experimental strand) has started with this problem as its foundation; in Section 3, we will explain the ways we have extended the classic problem in stochastic and game-theoretic directions.

We then move on to discuss compartmental graph models of disease, examining the differential equations required to describe such models exactly. This has formed the basis of our more formal work over the past year, so again a discussion of the results and knowledge in this field that pre-dates our work provides important background for Section 3. While this work appears fairly distinct in character to the more experimental prong of our approach, the two aspects are related in principle: both aim to take existing approaches to disease modelling and introduce a new way of representing protection from disease, both as an external defence strategy and as an internal inclination in individuals.

We end this section with a discussion of Percolation Theory and its use in a graph-theoretic context. We have identified this as an important tool for disease modelling using graphs, but we have not yet employed it to a significant extent. We will, nonetheless, discuss the areas that we believe it will prove useful in Section 3.

2.2 The Firefighter Problem

The Firefighter Problem, which we refer to as simply FIREFIGHTER, was first introduced by Hartnell [5] and models on a given graph an outbreak of fire with a firefighter strategically blocking its path.

We formalise FIREFIGHTER as follows: at $t = 0$, a fire breaks out at some vertex v_0 of graph G. The firefighter then 'protects' another vertex of G. A protected vertex is protected for the remainder of the game. Similarly, a vertex that is on fire is 'burning' for the rest of the game and cannot ignite again. The fire spreads to any immediate neighbouring vertices that are neither protected nor burnt. Then, the firefighter may protect another vertex, the fire spreads again and so on. The following is a decision formulation given by Finbow and MacGillivray for FIREFIGHTER [2]:

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?

Common problems in classic FIREFIGHTER often involve how to minimise the number of vertices that will be burnt and, in a given class of trees, determining the average number of burnt vertices. There are several results available to us in classic FIREFIGHTER - for one, the decision problem is NP-complete, even if restricted to trees with maximum degree three. Further, the problem is solvable in polynomial time for graphs of maximum degree three, so long as the fire starts at a vertex of degree two [2]. We also have some results regarding game strategies, such as that the greedy algorithm is a 1/2-approximation for the problem on trees [2].

If we let each vertex be an individual and edges between them represent social contact, we reach a simple model for disease infection, which is where our interest in the game enters. We note that are many natural contextualisations of FIREFIGHTER beyond modelling disease spread, which we may choose to apply our work to in future - for instance, we may think of edges representing virtual contact between individuals on social media, yielding a model for the spread of viral internet memes [10].

2.3 Markovian SIR epidemics on networks

Compartmental models are mathematical ways of simulating how sections of a population interact. This involves systems of equations describing the behaviour of each compartment and then seeing how each compartment interacts with the others. These compartments can describe many states an individual (typically a person or an animal) can be in - for instance, we could use predator/prey compartments to simulate a hunting situation or susceptible/infected/recovered compartments to model disease spread.

In this discussion, we will begin by discussing the standard SIR model, allowing us to later explain how it can be extended to a graph-theoretic context and then explain the current work on generating a system of equations to exactly describe such a model. The work detailed here on Markovian SIR epidemic graph models is hugely foundational for our formal approach to research so far.

2.3.1 Standard SIR model with fixed population

The SIR Model is a compartmental epidemiological model with three compartments (generally referred to as 'states') related to an epidemic: *susceptible, infected* and *recovered.* We define $S(t)$ as the number of susceptible people (i.e. those who do not currently have but are able to contract the infection) at time t, $I(t)$ as the number of infected people (i.e. individuals who have the disease and are infectious) at time t and finally $R(t)$ as the number of recovered people (i.e. those who have had the disease and subsequently recovered, granting them at least some level of immunity in certain contexts). For a fixed population N, we have that $S(t) + I(t) + R(t) = N$ - that is, we assume a fixed population where we do not wish to consider vital dynamics, which is a reasonable assumption to make when an epidemic is short-lived. Then, for β and γ the rates of infection and recovery respectively, the standard SIR model without vital dynamics (birth and death rates) is given as follows [6]:

$$\frac{dS}{dt} = -\beta\frac{SI}{N}, \tag{1}$$

$$\frac{dI}{dt} = \beta\frac{SI}{N}\gamma I, \tag{2}$$

$$\frac{dR}{dt} = \gamma I - \mu R. \tag{3}$$

2.3.2 Extending the SIR model to graphs

In order to extend our SIR model to network graphs, we first examine the probability of an agent being in a given class: let $\langle A_i \rangle$ represent the time-independent probability of vertex i being in state A, meaning that the expression of the form $\langle A_i B_j \rangle$ represents the probability of vertices i and j being in states A and B respectively [8]. We begin with a contact network, where vertices represent individuals and the edges between them represent social contact which may serve as an infection pathway. Then, the adjacency matrix G of this network is constructed by assigning $G_{ij} = 1$ when i and j share an edge and $G_{ij} = 0$ otherwise.[1] Then, we extend this contact network to a transmission network: let β_i represent the per-link infection rate for individual i and γ_i represent the recovery rate for i. For the transmission matrix T, we assign $T_{ij} = \beta_i$ if there is a single-edge route of infection between i and j and $T_{ij} = 0$ otherwise. Often, we will consider unweighted and undirected graphs, but in general T_{ij} may not equal T_{ji}.

We now note that we can replace β_i with a term involving such a transmission matrix of a network in order to begin extending the usual SIR model into a network realm. Using the substitution $\beta_i \frac{SI}{N} = \sum_{j=1}^{N} T_{ij}\langle S_i I_j \rangle$, the equations become

$$\langle \dot{S_i} \rangle = -\sum_{j=1}^{N} T_{ij}\langle S_i I_j \rangle$$

$$\langle \dot{I_i} \rangle = \sum_{j=1}^{N} T_{ij}\langle S_i I_j \rangle - \gamma_i \langle I \rangle$$

$$\langle \dot{R_i} \rangle = \gamma_i \langle I \rangle,$$

which are the evolution equations given in [8].

In the past year, we have also spent significant time understanding, applying and expanding results around equations describing Markovian SIR graph models of disease exactly [8] so we will now provide a review of the key work in this area and its utility. We will go on to detail the work done so far in extending the results from this paper in the following section.

In [8], 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 second of these three describes the approach we have taken to our work so far, as will be seen in Section 3. The latter of these three approaches is the one used in the work on Markovian SIR graph epidemics being discussed currently: 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.

[1]Because we are infrequently interested in self-transmission, we often set $G_{ii} = 0$.

The work done in [8] has two goals. Firstly, the authors aim to provide an exact, deterministic representations of Markovian SIR epidemics on graphs with and without loops. Secondly, they seek 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.

The significance of cut-vertices is that they permit us to create closures in the systems of equations that describe a compartmental model. To better explain closures, consider the equations below for calculating singles and pairs [8]:

$$\begin{aligned}
\langle \dot{S_i} \rangle &= -\sum_{j=1}^{N} T_{ij} \langle S_i I_j \rangle, \\
\langle \dot{I_i} \rangle &= \sum_{j=1}^{N} T_{ij} \langle S_i I_j \rangle - \gamma_i \langle I_i \rangle, \\
\langle \dot{S_i I_j} \rangle &= \sum_{k=1, k\neq i}^{N} T_{jk} \langle S_i S_j I_k \rangle - \sum_{k=1, k\neq j}^{N} T_i k \langle I_k S_i I_j \rangle \\
&\quad - T_{ij} \langle S_i I_j \rangle - \gamma_i \langle S_i I_j \rangle, \\
\langle \dot{S_i S_j} \rangle &= -\sum_{k=1, k\neq j}^{N} T_{ik} \langle I_k S_i S_j \rangle - \sum_{k=1, k\neq i}^{N} T_{jk} \langle S_i S_j I_k \rangle,
\end{aligned}$$

This is *not* a closed system - we require equations for triples, although we have only sought to express singles and doubles. Obtaining closures here would involve substituting triple terms for lower-order terms (singles and doubles).

The main result of the work reveals an important relation between the structure of the graph used in the epidemic model and the types of closures that are feasible. Let $G = \{V, E\}$ be a graph on N vertices with a set of edges, E. Consider a connected subset of vertices $F = \{v_1, v_2, \ldots, v_k\} \subset V$ and assume that $\exists v_{i^*} \in F$, a cut-vertex in G such that $F \setminus \{v_{i^*}\}$ is partitioned into at least two disjointed components with vertices $F_1 = \{v_1, v_2, \ldots, v_{i-1}\}$ and $F_2 = \{v_{i+1}, v_{i+2}, \ldots, v_k\}$ belonging to any such two, distinct and disjointed components or subnetworks. Then the following equation holds:

$$\langle Z_{v_1} Z_{v_2} \ldots Z_{v_{i-1}} S_{v_{i^*}} Z_{v_{i+2}} \ldots Z_{v_k} \rangle (t) = \frac{\langle Z_{v_1} Z_{v_2} \ldots Z_{v_{i-1}} S_{v_{i^*}} \rangle (t) \langle S_{v_{i^*}} Z_{v_{i+2}} \ldots Z_{v_k} \rangle (t)}{\langle S_{v_{i^*}} \rangle (t)},$$

where in each Z_{v_i} term, Z may stand for any state symbol in our model (such as S for susceptible, P for protected and so on) for all vertices v_i that are not cut-vertices.

The authors further 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 [8]:

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

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 [8]:

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

where P is the number of distinct sub-graphs produced when the original graph is spliced into 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 [8].

In their concluding remarks, the authors recommend the implementation of an algorithmic approach to this equations generation with closure detection as the examples provided in their work were done by hand [8]. This is something we have been working on for several months, which will be discussed in Section 3.2.3.

2.4 Percolation Theory

Widely known and used in physics, statistics and mathematics, Percolation theory involves modelling scenarios as n-dimensional graphs, so application to FIREFIGHTER is not entirely unexpected. We now examine percolation theory and later, in Section 3.3, explain the utility we have found it may yield in extending and expanding modelling work from FIREFIGHTER.

In percolation, the edges between vertices in the graph can be either 'open' or 'closed' with probability p and $1-p$ respectively. We can think of percolation problems as liquid being poured onto a porous material and whether there is a path from hole to hole along open paths through the material. Note that removing more and more edges moves us towards a critical point at which removing further edges would cause the graph to fall apart into smaller clusters of vertices and edges that have no access to each other [3]. This is known as 'bond' percolation, as edges correspond to bonds in many of its applications.

Several authors have suggested percolation as a possible approach to FIREFIGHTER [2]. In this context, we could determine the critical point to see how we might contain the fire to a smaller cluster that cannot spread to the wider graph. For FIREFIGHTER, site percolation is more applicable: rather than considering open or closed *edges* ('bonds') between vertices as in bond percolation, we consider each *vertex* ('site') as being 'occupied' or 'unoccupied' with probability p and $1-p$ respectively.

Formally, we consider a point lattice $\mathbb{L}$ and denote the open cluster as $C(x)$, where $x \in \mathbb{L}$ is the local origin of the cluster. This cluster $C(x)$ is defined as the set of all vertices that can be reached from open paths beginning at the nucleation site, x. Then, we are particularly interested in the *percolation probability*:

$$\theta(p) = \mathbb{P}_p(\,|C(0)| = \infty\,),$$

and the *critical probability* (or *percolation threshold*):

$$p_c = \sup\{\, p \mid \theta(p) = 0 \,\}.$$

Here, $\mathbb{P}_p$ is the product measure given by:

$$\mathbb{P}_p = \prod_{v \in \mathbb{L}^d} \mu_v$$

where μ_v is the *Bernoulli measure*, which returns p when v is open and $1-p$ when v is closed [9, p. 28]. Analytically, others have shown that in the case of a two-dimensional regular point lattice, the critical probability is $p_c = 1/2$ [7].

3 Work done so far

3.1 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. By agency, we mean the ability of an individual to make decisions and choose to pursue them. In our context of disease modelling, agency might manifest in a modelled individual's capacity to make decisions about mitigating their own risk of contracting the infection (given the extent to which they are able to do so). To address this, we have designed and implemented a compartmental graph model of infectious disease, which notably includes associating with each individual a 'defence rating,' which is a probability corresponding to their internal inclination towards self-protection (which also accounts for external constraints, for instance whether their circumstances mandate increased social contact). In this section, we first describe the implementation carried out so far before discussing the initial results obtained. We will then discuss the conference at which we presented some of these results and the significance of the approach employed.

3.1.1 Implementation of agency-based models

In our implemented model, we begin with a graph on a particular number of vertices and edges. This can be either given by the user or generated using code we have written for several random graph types.

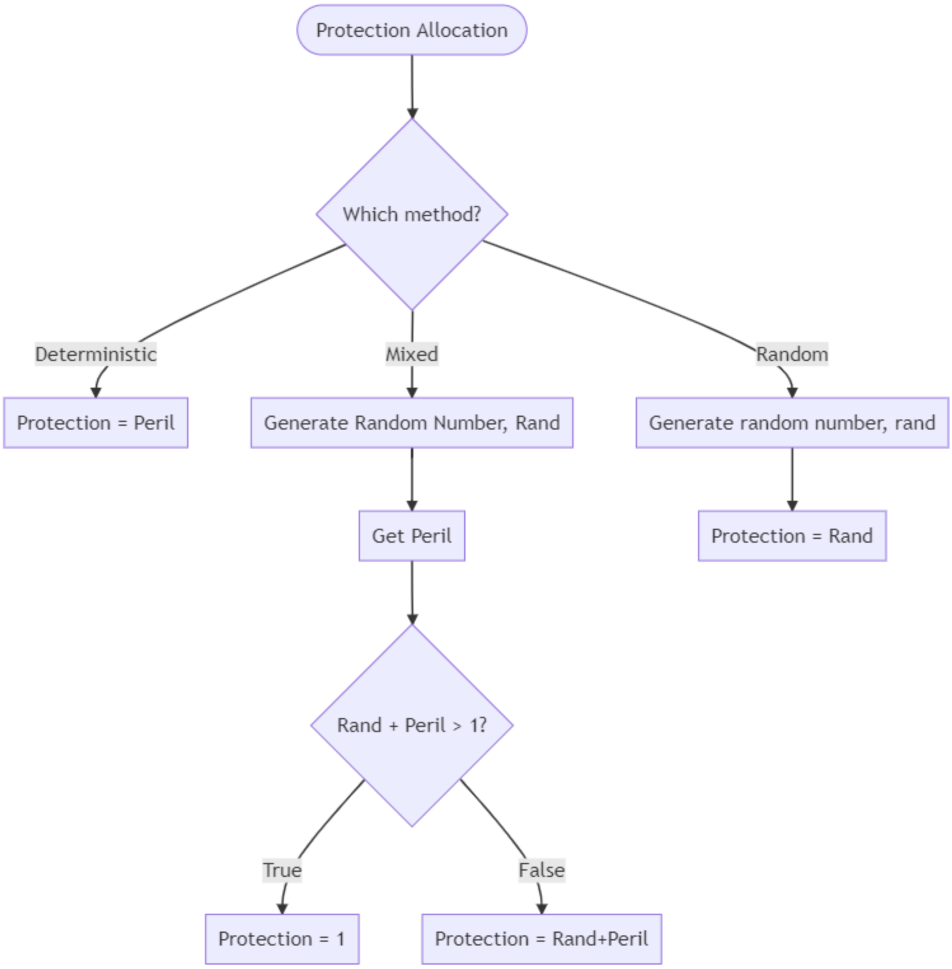


Figure 1: Visualisation of protection allocation - this is allocated either based on proximity to infection, randomly or by combination of both.

We then assign an agent to each vertex. Agents can be in a number of states, such as 'susceptible' (could contract), 'infected' (currently has the infection and is infectious), 'recovered' (previously had the infection) and 'protected' (cannot contract the infection). The latter of these state is where much of our work has been focused: we have been studying the dynamics of the system with this state using the 'protection ratings' of vertices. We currently assign this rating in one of three ways: purely randomly, based upon proximity to infection or a combination of the two approaches. Figure **??** shows how each agent is initialised and Figure 1 shows in particular how such agents have their protection ratings assigned.

Once these agents are initialised, a defensive move is made, which is adapted to account for the inherent individual protection ratings. The current existing defence strategies are detailed in Figure 2. Currently, there are three main defence strategies that are deployed.

The defence strategies are represented visually in Figure 2. These strategies are:

- Defend based on highest degree, breaking ties on greatest proximity to infection;
- Defend based on greatest proximity to infection, break ties on highest degree; and

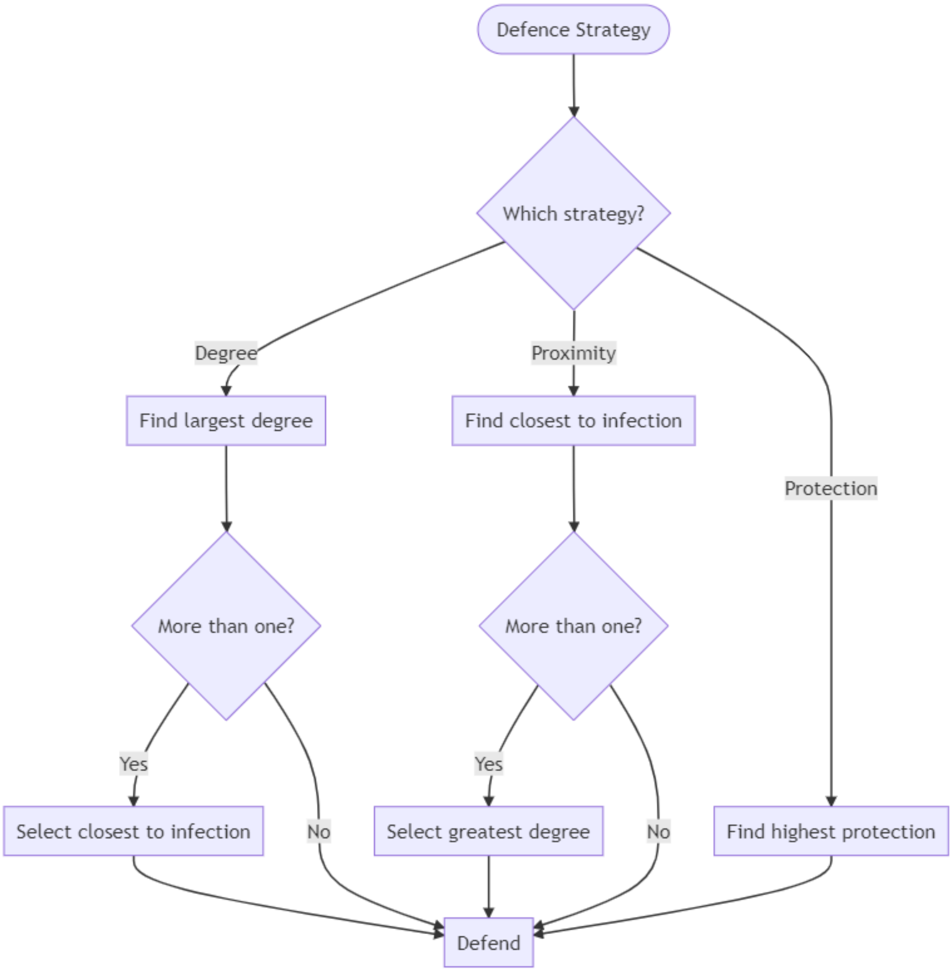


Figure 2: Representation of three implemented defence strategies.

- Defend the agents who currently have the highest protection rating.

The first two of these strategies are common in classic Firefighter - defending based on degree is effective in fairly dense graphs and defending based on proximity to fire is effective in sparse and tree-like graphs [2].

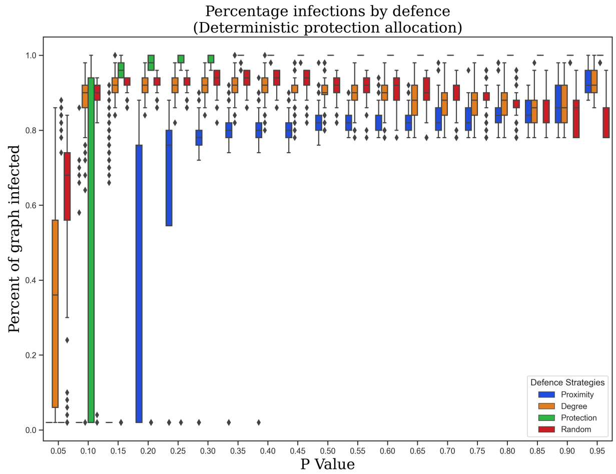
3.1.2 Computational Biology Conference

On 27$^{\text{th}}$ May 2021, we presented a talk on agency-oriented modelling to the First University of Glasgow Computational Biology Conference. The title of the talk was *introducing features of agency into computational models of infectious disease.* The key points of discussion were:

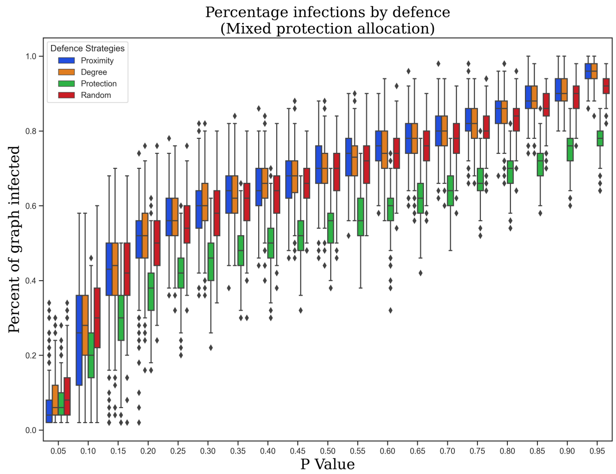
- Current computational approaches to graph models of disease
- Extending existing graph models to better account for individual agency
- Introducing a compartmental modelling approach

3.2 Graph-based compartmental models

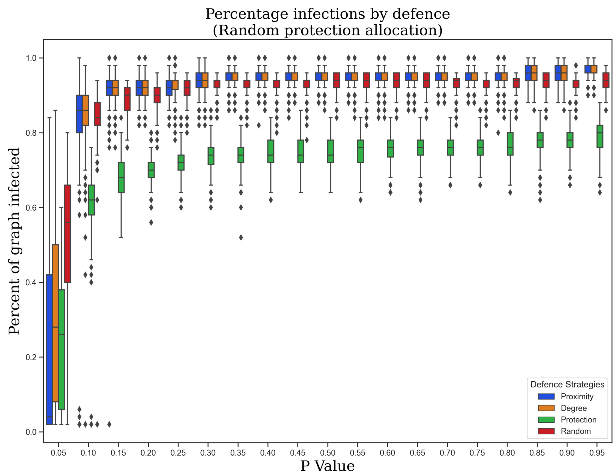
In this section, we will detail and explain the work we have done in extending the work outlined in Section 2.3. We begin by exploring the addition of a new compartmental state to a graph-based SIR model and then we determine the impact this has on the total system of equations describing these models. The main focus of our work so far in this area has been to develop an algorithmic approach for the work detailed in [8] (discussed in Section 2.3), which the authors of that work left as an open problem.



(a) Protection ratings allocated based on proximity to closest infection.



(b) Protection rating allocated using a baseline random number which is increased based on proximity to closest infection.



(c) Protection rating allocated randomly.

Figure 3: Charts showing percentage of Erdös Rényi graphs generated with varying probability p infected with agent protection ratings allocated in three different ways

3.2.1 Adding a new state to the model

Let ζ_i be the probability that we defend individual i. Similarly, let α_i represent the efficacy of the protection measure for individual i, which may decay over time and vary from person to person. Using these rates of protection and effectiveness, for fixed population size the differential equations become:

$$\langle \dot{S_i} \rangle = \alpha_i \langle P_i \rangle - \sum_{j=1}^{N} T_{ij} \langle S_i I_j \rangle - \zeta_i \langle S_i \rangle \tag{4}$$

$$\langle \dot{I} \rangle = \sum_{j=1}^{N} T_{ij} \langle S_i I_j \rangle - \gamma_i \langle I \rangle \tag{5}$$

$$\langle \dot{R_i} \rangle = \gamma_i \langle I \rangle \tag{6}$$

$$\langle \dot{P_i} \rangle = \zeta_i \langle S_i \rangle - \alpha_i \langle P_i \rangle. \tag{7}$$

3.2.2 Results for total system of equations

As an example, we consider the 'triangle network' - a loop of three nodes. The equations required to precisely express the system SIR dynamics of this network are as follows [8]:

$$\text{6 singles: } \langle \dot{S_1} \rangle, \langle \dot{S_2} \rangle, \langle \dot{S_3} \rangle, \langle \dot{I_1} \rangle, \langle \dot{I_2} \rangle, \langle \dot{I_3} \rangle. \tag{8}$$

$$\text{6 doubles: } \langle \dot{S_1 I_2} \rangle, \langle \dot{I_1 S_2} \rangle, \langle \dot{S_1 I_3} \rangle, \langle \dot{I_1 S_3} \rangle, \langle \dot{S_2 I_3} \rangle, \langle \dot{I_2 S_3} \rangle. \tag{9}$$

$$\text{6 triples: } \langle S_1 \dot{I_2} I_3 \rangle, \langle S_1 \dot{I_2} S_3 \rangle, \langle S_1 \dot{S_2} I_3 \rangle, \langle I_1 \dot{S_2} S_3 \rangle, \langle I_1 \dot{I_2} S_3 \rangle, \langle I_1 \dot{S_2} I_3 \rangle. \tag{10}$$

Now, using the equations for the $SIRP$ model, we have the following equation requirements:

$$\begin{aligned}
\text{9 singles: } & \text{(8) and } \langle \dot{P_1} \rangle, \langle \dot{P_2} \rangle, \langle \dot{P_3} \rangle. \\
\text{18 doubles: } & \text{(9) and } \langle \dot{S_1 P_2} \rangle, \langle \dot{P_1 S_2} \rangle, \langle \dot{I_1 P_2} \rangle, \langle \dot{P_1 I_2} \rangle, \langle \dot{S_1 P_3} \rangle, \langle \dot{P_1 S_3} \rangle, \\
& \langle \dot{I_1 P_3} \rangle, \langle \dot{P_1 I_3} \rangle, \langle \dot{S_2 P_3} \rangle, \langle \dot{P_2 S_3} \rangle, \langle \dot{I_2 P_3} \rangle, \langle \dot{P_2 I_3} \rangle. \\
\text{24 triples: } & \text{(10) and } \langle S_1 \dot{S_2} P_3 \rangle, \langle S_1 \dot{P_2} S_3 \rangle, \langle S_1 \dot{I_2} P_3 \rangle, \langle S_1 \dot{P_2} I_3 \rangle, \langle S_1 \dot{P_2} P_3 \rangle, \langle I_1 \dot{S_2} P_3 \rangle, \\
& \langle I_1 \dot{P_2} S_3 \rangle, \langle I_1 \dot{I_2} P_3 \rangle, \langle I_1 \dot{P_2} I_3 \rangle, \langle I_1 \dot{P_2} P_2 \rangle, \langle P_1 \dot{S_2} S_3 \rangle, \langle P_1 \dot{S_2} I_3 \rangle, \\
& \langle P_1 \dot{I_2} S_3 \rangle, \langle P_1 \dot{I_2} I_3 \rangle, \langle P_1 \dot{P_2} I_3 \rangle, \langle P_1 \dot{P_2} S_3 \rangle, \langle P_1 \dot{S_2} P_3 \rangle, \langle P_1 \dot{I_2} P_3 \rangle.
\end{aligned}$$

Note that the reason we dispense with the cases of all three vertices being in the same state is that this would not result in any dynamics - no vertices would ever change state in this case.

3.2.3 Implementation

We have been working on an implementation of the above work in equations generation. Our goal is to produce code that can accept a particular graph, for instance as a CSV file, and determine the number of equations that could be required to fully describe a compartmental model (for instance, SIR or $SIRP$) on that graph. This could be by providing upper and lower bounds if the exact graph structure is unknown, or by providing the exact number of equations if the graph is known. With this information, the user can request the software returns the full list of differential equations that exactly describe the epidemic dynamics using an algorithmic approach to generation.

3.3 Uses of Percolation in The Firefighter Problem

We have identified three main avenues that we would like to explore further regarding the use of Percolation Theory in Firefighter:

- the firefighter may use percolation in order to defend the graph,
- the fire might spread with percolation probability p or

- we might use percolation on the graph to form a more useful model (i.e. one that can more accurately represent an irregular population density).

The first of these options might be used when the firefighter can save more than one vertex at each turn, or (as will be seen further on in this section) defence is a stochastic process; the third of the three options may be more useful when modelling disease spread.

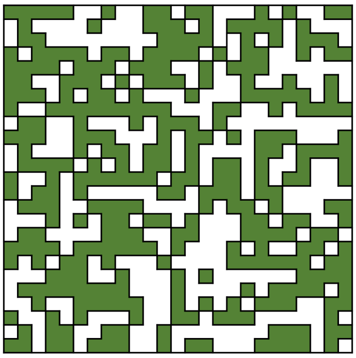


Figure 4: A 4-regular (25×25) graph, percolated with probability $p = 2/3$.

3.3.1 Better than random

One potential use of percolation is a baseline test. In most scenarios, a method for obtaining defence strategies should be at least as effective as a random defence sequence. We could find such a random sequence by allowing the firefighter to defend in a percolation-like strategy for comparitive purposes. Consider a sequence of vertices in graph G, written as $d_1, d_2, \ldots, d_t$. An optimal defence sequence could be found using integer programming as provided by Finbow and MacGillivray [2]:

$$\begin{array}{lll}
\text{Maximise} & \displaystyle\sum_{v \in V(G)} d_v w(v) & \text{for each level } i \\
\text{subject to} & \displaystyle d_v + \sum_{\text{level}(v)=i} d_v \leq 1 & \text{for each level } i \\
& \displaystyle d_v + \sum_{u \succ v} d_u \leq 1 & \text{for every outer vertex } v \text{ of } T, \\
& d_v \in \{0, 1\}. &
\end{array}$$

where $u \succ v$ indicates that u is an ancestor of v. The optimal strategies provided for different classes and densities of graphs here will provide an upper bound (which may indeed be impossible to attain in some cases) for success of a given strategy. We can find a lower bound using percolation, and so we have a range of success values as a starting point: if some strategy is better than random percolation, then it is worth considering, but below the particular expected optimal solution from integer programming and we can, in principle, improve or find a better strategy.

We conjecture that, at the lowest graph densities, the random strategy will be close to the optimal strategy and thus finding an improvement is at once difficult and lacking in great utility. At the very highest graph densities, random strategies will have a very low expected best-case scenario but so will most strategies, since the constraint on the firefighter that they have only one vertex to save per turn does not go as far towards fire containment per turn compared to sparser graphs.

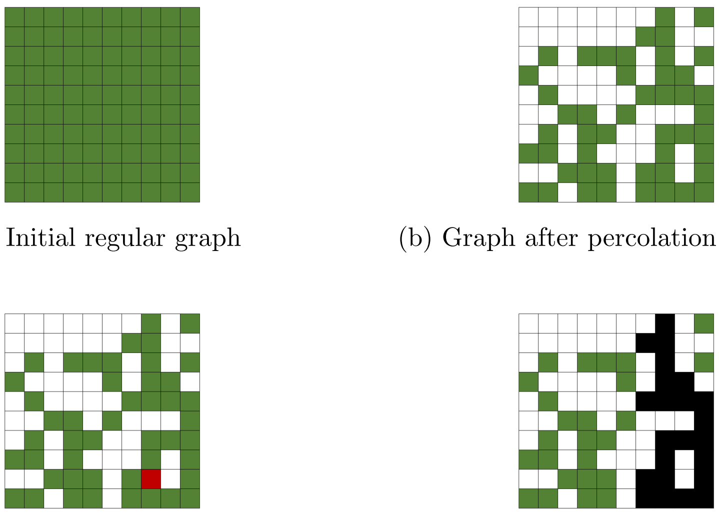
3.3.2 Reproduction rate

We now focus our attention on the fire spread being determined by percolation (rather than the firefighter's defence sequence). Diseases, when there is a large enough sample size, have a basic reproduction rate associated with them, denoted R_0: for instance, measles has a basic reproduction rate $12 \leq R_0 \leq 18$ [4] and the Influenza strain responsible for the 1918 pandemic has a basic reproduction rate of $1.4 \leq R_0 \leq 2.8$ [1]. These baseline, theoretical values can be implemented as an internal probability to a propagating fire: to formulate a stochastic version of Firedfighter, we let the fire propagate with some probability (which could be determined by reproduction rate of a real infectious disease) in a percolation-like process and examine the change to optimal defence.

Where we wish to consider vertices as individuals and edges as the connections between them, percolation may give us a more useful model for disease spread when we do not assume the population is well mixed and instead introduce probability functions to correspond to the likelihood one vertex is connected to another.

3.3.3 Irregular population density

A great deal of the literature surrounding Firefighter assumes a regular graph - that is, in the context of disease we assume a well-mixed population where everyone has equal probability of coming into contact with their neighbours. Of course, this is a simplification of reality: some individuals are very well connected and have lots of contact with others, and others have significantly less contact. In the context of a forest fire, the density of a forest is naturally irregular and there is a probability in the unit interval that fire can spread between two trees depending on their proximity (among other factors): an example of this can be seen in figure 4. Thus, percolation on regular grids to more closely resemble populations or forest density could lead to far more useful and realistic modelling results. If we start models on graphs which have been percolated with some given probability, we can ensure the model better resembles the density and irregularity of a natural population.



(a) Initial regular graph

(b) Graph after percolation ($p = 1/2$)

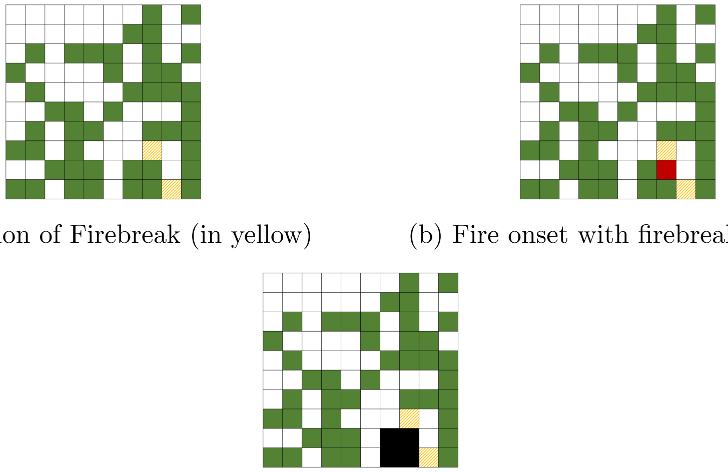
(c) Initial outbreak ($t = 0$)

(d) End of outbreak ($t = 12$)

Figure 5: Outbreak of fire on a percolated graph

To illustrate this idea, we will consider a similar problem to Firefighter, called The Firebreak Problem or simply Firebreak. Rather than have a firefighter reactively combatting the fire on each turn, we begin with an allocated amount of funding with which to mitigate the expected damage of the fire, and spend all of that before the fire begins. Generally, this means we have a number of edges to remove (trees to cut down or place a barrier in between) before the fire begins. Figure 5 depicts the use of percolation on a graph to model a forest fire. We can see that the model would have little utility if we began the outbreak on the graph in 5a (all of the graph would be burnt), but percolation has made this a more instructive and realistic model in figure 5b. We can see where the fire has spread in 5d, and we can see where - if we had a finite amount of resources - we must focus our efforts and funding in introducing a firebreak.

If, for instance, we could only remove two vertices as a firebreak, we ought to make those two the ones shown in yellow in figure 6. This is an example of how percolating a graph for use in Firefighter (and indeed Firebreak) can assist in generating a more realistic and useful model for contagion. There is much interest in the regular graph examples, but our modelling contexts



(a) Position of Firebreak (in yellow)

(b) Fire onset with firebreak ($t = 0$)

(c) End of outbreak ($t = 3$)

Figure 6: Outbreak of fire on a defended percolated graph

of interest call for a more sporadic graph density and so we, in future research, will try and prove conjectures regarding the minimum number of vertex defences required in FIREBREAK for percolated graphs of given percolation threshold and dimension (or infinite graphs) and analogous questions in FIREFIGHTER.

These random graphs are very reminiscent of Erdös Rényi random graphs and this is for good reason. In fact, the Erdös Rényi process of graph generation - adding edges into a graph with some given probability p - can be thought of as bond percolation on the complete graph with probability $1 - p$.

4 Research Plan

4.1 Goals

Over the next year, we hope to extend the two strands of work we have been pursuing with a number of goals and milestones in mind. In the more formal, mathematical side of our work, we aim to achieve the following:

- Formalise a probabilistic, agency-based graph model of disease that can be implemented in the PRISM model checking language.
- Complete and verify the equations generation work that we have started over the last year and write down the algorithmic approach used.
- Explore further formal approaches to agency-based modelling that could be used to formulate models or understand better our existing models (for instance, integer programming approaches or model-checking using PRISM).

Moreover, in the experimental side of our work, we aim to achieve the following:

- Further generalise the current implementation to accept further parameters and output requested results.
- Implement, test and explore further defence strategies in the implementation, using this to reason about optimal defence candidates in different contexts.
- Write down a series of (provable) statements regarding defence strategies in this implementation on various graph classes.
- Use real-world data to model different epidemic scenarios and contextualise defence strategies.

Some significant milestones we aim to reach are as follows:

- TODO Conference?
- TODO Paper?
- Request, write and present a talk for the FATA research section in the department in a post-graduate slot in the Autumn of 2021.

4.2 Timeline

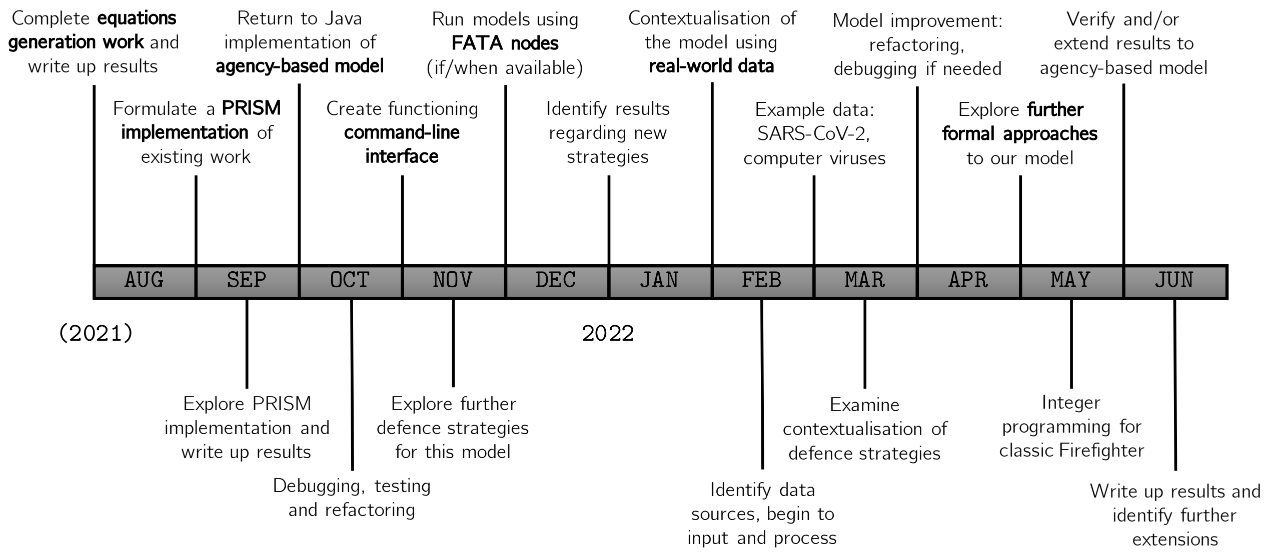


Figure 7: Timeline of work for the next year

References

[1] N. M. Ferguson, D. A. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke, *Strategies for mitigating an influenza pandemic*, Nature, 442 (2006), pp. 448–452.

[2] S. Finbow and G. MacGillivray, *The firefighter problem: A survey of results, directions and questions*, The Australasian Journal of Combinatorics, 43 (2009).

[3] G. Grimmett, *Percolation*, Springer-Verlag, Berlin, 2 ed., 1999.

[4] F. M. Guerra, S. Bolotin, G. Lim, J. Heffernan, S. L. Deeks, Y. Li, and N. S. Crowcroft, *The basic reproduction number (r_0) of measles: a systematic review*, The Lancet Infectious Diseases, 17 (2017), pp. 420–428.

[5] B. L. Hartnell, *Firefighter! an application of domination*, in 25th Manitoba Conference on Combinatorial Mathematics and Computing, University of Manitoba in Winnipeg, Canada, 1995.

[6] H. Hethcote, *The mathematics of infectious diseases*, SIAM Review, 42 (2000), pp. 599–653.

[7] H. Kersten, *The critical probability of bond percoation on the square lattice equals 1/2*, Communications in mathematical physics, 74 (1980), pp. 41–59.

[8] I. Z. Kiss, C. G. Morris, F. Sélley, P. L. Simon, and R. R. Wilkinson, *Exact deterministic representation of markovian SIR epidemics on networks with and without loops*, Journal of Mathematical Biology, 70 (2014), pp. 437–464.

[9] A. Klenke, *Probability Theory*, Springer, Berlin Heidelberg, 2 ed., 2014.

[10] J. D. O'Brien, I. K. Dassios, and J. P. Gleeson, *Spreading of memes on multiplex networks*, New Journal of Physics, 21 (2019).

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