OEC 2021 - Group R

Joseph Grosso, Andrew Fryer, Kyle Singer, Andrew Farley

January 23rd 2021

Contents

1 Introduction

The goal of this project is to identify and model the spread of ZBY1 virus through a given high school in one day. In order to achieve this, the team decided to create a simulation consisting of students, teachers, and teaching assistants (TAs). The goal of the simulation is to show each person's probability of getting infected, based on the courses they take, the extra-curriculars they are in, who they have lunch with, and who they unintentionally interact with. The simulation being created takes into account all of the different probabilities of contracting the virus each of their interactions bring. The output will then show how their risk increases or decreases after each event (i.e. attending a class). In order to do this, the team needs a good understanding of the interactions between individuals, a working algorithm to run the simulation, as well as a detailed visualization of the results.

1.1 Structure

Figure 1 shows our flowchart of the algorithm that we settled upon to model our probability of infection at each state. Our given list of associations and people in the school was fed into our algorithm, which was then used to calculate our risk for infection. This was to be calculated using proximity to infected persons and infectivity risk based on preexisting conditions and age. We would then use the updated state of risk for each person and their new locations over time as our next input to the algorithm, and then iteratively repeat this process.

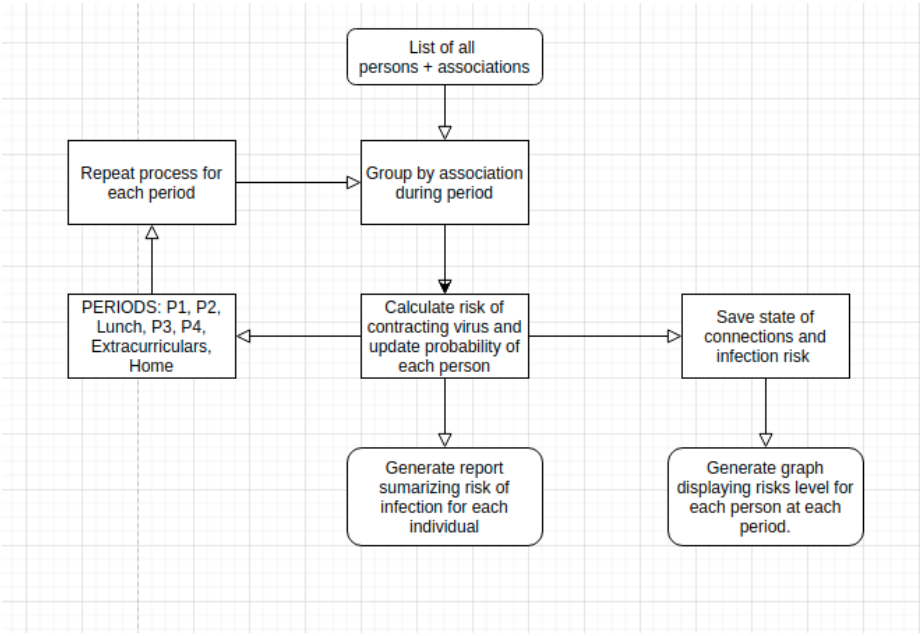

Figure 1: A flowchart of the high-level process our algorithm uses to simulate infections.

1.2 Interactions

The first thing considered when thinking about the different aspects of interactions between students, teachers, and teaching assistants was if there was a known number of people that could be infected in one interaction. To ensure that the probabilities are reasonable, we computed each probability using a baseline value, $r0$, of 3.

1.2.1 Types

Before starting to code, the team discussed the different possible types of interactions that could lead to a spread of infection. The ones identified were: in class (student-student, student-teacher, teacher-TA, TA-student), switching classes, lunch (teachers, TAs, students (grade dependent)), and extra curricular activities (student-student).

1.2.2 Probabilities

When considering the probability of someone getting infected, the team decided to create functions that would output probabilities given that one person is transmitting to another, assuming one of the two is infected. This way the situation they are in (i.e. in class), is considered and each probability can then

be adjusted to fit each student differently depending on their differences. This also helps to avoid any normalization issues. The different figures chosen in the calculation of each probability would consider the carefulness of students, teachers, and TAs, as well as the type of situation/period they are currently (see section 2.1.1).

1.2.3 Safety Measures

Noting that the situation laid out in this competition is similar to the one the world is currently living in (COVID-19), the team decided to consider the health and safety measures that are currently in place today. More specifically, the team chose to include the possibility to show how the spread of the virus through the school would be if students were wearing masks and social distancing (depending on the situation). This could then be accounted for while determining the probability of infection.

1.3 Algorithm

As soon as the problem was introduced, the team started brainstorming potential solutions. A common theme in the ideas of brainstorm was taking a simulated approach. This was decided to be optimal as it would allow us to create an iterative algorithm, a powerful approach. Besides this, there were two main lines of thought from the ideas generated. The first was to calculate the probabilities of an individual being infected using their previous probability as well as information about their current location and habits throughout each school period. The second was to determine if an individual was infected at each period using a Boolean value while still using information about their current location and habits each period. The advantage to the first method is it is more general, calculating the risk each individual had of contracting the virus. The advantage to the second method is it is easier to understand but will change will each run of the simulation as random numbers would be needed to compute the result.

It was decided that the first method would be implemented due to its generality and robustness. Another advantage is that if the team had enough time they would be able to use the first method to generate and instance of the second. Unfortunately, the team ran out of time but would have liked to implement both methods. The algorithm works by collecting all data from the provided excel spreadsheet and creating class instances with the data. Each person (i.e. student, teacher, and TA) has their own class instance. A query class was built to store the information for each person and to allow for queries to be executed, obtaining lists of people in different situations. A main loop was constructed which iterated over all possible times of the day, Period 1, Period 2, Lunch, Period 3, Period 4, and Extracurriculars. On each iteration, the query class was used to obtain the necessary data for each event, and probability functions were used to propagate the infection probabilities to the next iteration.