

Simulation Paper

Benjamin Vandersmissen¹, Lars Van Roy²,
Evelien Daems³, and Frank Jan Fekkes⁴

¹ `benjamin.vandersmissen@student.uantwerpen.be`

² `lars.vanroy@student.uantwerpen.be`

³ `evelien.daems@student.uantwerpen.be`

⁴ `franciscus.fekkes@student.uantwerpen.be`

Abstract. This paper will give an in depth view of the performed adaptations to the Stride project, as well as an overview of the findings that were obtained by comparing these additions to the original stride project.

1 Introduction

The stride project is designed to simulate and evaluate the lifetime of various infectious diseases. In order to properly analyze which factors have an effect on various diseases, the stride project allows multiple factors to be varied so that we can get an idea of which factors are the main influence for the evolution of diseases. This expansion to the original stride project will include 2 major corrections, to make the simulation more realistic, being an improved workplace size distribution and an improved household composition distribution, a major addition to the pools in which diseases can spread, being daycare and preschool pools, a minor addition to make the input more variable by adding two addition input types being HDF5 and JSON and finally a visualizer that will allow users to get a graphical overview of where the diseases are most active, along with a list of parameters to give further information of the evolution of the disease for a given location.

Other than these additions, there will also be a section about the simulation over entire Belgium, compared to isolated within Flanders. other than just the simulated area, there will also be a comparison between the data derived from the Belgian population and the data derived from the Flemish population.

2 Daycare & Preschool

3 Data Formats

3.1 JSON

3.2 HDF5

4 Data Visualization

5 Workplace Size Distribution

5.1 Introduction

The original stride project made use of a set workplace size distribution, being an average size of 20, that would be randomly populated. So it might be that there would be smaller and larger workplaces, but the average would be around 20, where this is not accurate. The majority of the workplaces is smaller than 20 and there are workplaces that are way larger than 20. It would therefore be more accurate to add an input file that would give a distribution of which workplace size occurs with which chance.

5.2 Implementation

The input is given in a CSV file. The data itself consists of the chance for a workplace size being within the given range, along with the bounds for that range.

The generation will consist of a major loop that will iterate over the different locations. For each location we will calculate the maximal number of workers that is presumed to be working at this location, this will not be the exact value however, as many of the people who could be working, might not be, due to studies, or just general unemployment. After the number of workers is calculated there will be another loop, where workplaces will be randomly generated (using the provided distribution when available) until every possible worker is capable of working.

The population of the generated workplaces will start with a loop, iterating over the possible locations. For each location we will loop over the population of that location in a secondary loop. For each person we will check if the person is of age, and that the person is not a student. If this is ok, we will perform a coin flip, to determine whether or not the person is employed. If the person turns out to be employed, we will do one final coin flip, to determine whether or not the person commutes. When this is all done, we will select a random workplace that still has free space (if there are any workplaces left with free space), or a random workplace in case there were no spaces left.

5.3 Expected Impact

6 Household Composition Distribution

7 Belgian Simulations