Modelling COVID-19 Spread in a Supermarket

CS7DS1 Data Analytics, Trinity College Dublin, Ireland

Abstract

The COVID-19 pandemic has claimed countless lives and continues to pose a danger to human beings in the near future. The spread of this disease raised alarms all over the world resulting in socio-economic changes in day to day life. Research has been conducted to explore contributing elements that are relevant to COVID-19 transmission hazards in order to control the virus' spread. Geometric modelling of the pandemic spread is one of the methods used to simulate the virus spread and examine the methods to restrict it. In this paper, we formulate an agent based customer mobility and virus transmission model to simulate virus spread in a supermarket. The simulation is created using python libraries piglet and networkx. We generate customer mobility paths inside the store based on the customer historic data using a Machine Learning based recommender system. Waiting time is assigned to every customer which is directly proportional to the probability of visiting an aisle. Virus transmission is simulated based on the distance between infected and non-infected customers. The results from the simulation are presented in the paper to understand the impact of human movement on spatial networks such as store path which can further help restrict the virus spread.

Index Terms: Geometric Modelling, Customer Behavior Prediction, Agent Based Customer Mobility Model, Virus Transmission Model

1 Introduction

The COVID-19 (Coronavirus Disease 2019) pandemic led to worldwide lockdowns in order to minimise the spread of the virus. However, regardless of how strict a country's lockdown was, citizens still needed to go to supermarkets to buy food. It is a necessity and, as such, supermarkets remained one of the few places still open where many people congregated indoors. It is important from both a logistical and a safety perspective to make sure that the transmission of COVID-19 is kept to a minimum in these situations. This can be achieved in a variety of ways, such as limiting waiting times, reducing the maximum number of customers in the store, creating specific pathways through the store, and so on. Simulations are a good way to find the best possible conditions for reducing the spread of this virus in a safe way, without using real people and supermarkets as experiments. Our research aims to simulate a supermarket in the hopes of gaining information on how to make supermarkets safer.

COVID-19 itself is a novel virus known as Severe Acute Respiratory Syndrome Coronavirus (SARS-COV-2) and affects the respiratory system in humans [21]. The first case was reported in December 2019 and the World Health Organization (WHO) later declared COVID-19 a global pandemic on March 11, 2020 [17]. The observed incubation period of COVID-19 is estimated to have a mean of 5 days and a median of 3 days with a range of 0-24 days [21]. Only after the incubation may an infected individual display symptoms. Infected individuals are quarantined for a period of time depending on varying policies, such as 14 days and so on. However, infected individuals may be asymptomatic or display mild symptoms which go undetected and pass on the virus to other people.

COVID-19 is a contagious disease spread through physical contact between individuals [17]. Respiratory droplets produced from coughing, sneezing and talking are the primary mode of COVID-19 transmission. Studies demonstrate that these droplets can remain in the environment further infecting other individuals [7]. The estimated number of death due to COVID-19 is 18.2 million, although only 5.94 million deaths were reported between 1st January 2020 to 31st December 2021 [23]. There were no vaccines or antiviral treatments available for medical interventions in curing the disease in the early stages [12]. Instead, a set of non-pharmaceutical interventions (NPIs) were found to be effective in curtailing the spread of COVID-19, such as social distancing, clean hygienic practices, wearing face masks, voluntary quarantine, etc [17] [7] [12]. Eventually, countries throughout the world enforced these NPI practices, affecting many lives throughout the world.

In this paper, we aim to use simulations to model the spread of COVID-19. Section 2 provides an overview of previous research in the area, as well as some of the theoretical background that underlies our approach. In Sections 3 and 4, the methods by which the simulation is implemented are outlined. Section 5 details the results of our simulations, with some concluding remarks being given in Section 6. Finally, Section 7 contains some ideas as to future improvements and additions that could be made to the simulation.

2 Related Work

There has been a lot of work done in the field of modelling COVID-19. Many different models have been applied at different scales, ranging from supermarkets to countries. This section discusses some of these previous studies.

Models differ in terms of the type of the model, the acquisition method, the distribution of key input parameters, the hypothesis of the model, and variation in terms of the distribution used to describe the key time period of the COVID-19 infection [25]. Prediction differences under varying public health strategies were observed and it was found that travel restrictions showed the most significant effect. It was noted that the use of such models to decide on public health strategies needs to be done with extreme care and caution as deviation in the input epidemiological parameters can greatly affect the value of R0. As such, we tried to take great care to make informed decisions when choosing parameters for our study.

COVID-19 infection spread can be modelled using a graph based model [3]. They consider the social distance, the duration of contact, and the location based demographic characteristics among factors affecting spread. It is noted that there are economic, social, demographic, and population density factors which affect the rates leading to complex models [25] [3].

There are various complexities related to modelling and the different factors which need to be analysed. The broad categories of models are statistical, mathematical-mechanistic state space, and machine-learning based models. [19]

Affect of lockdowns on infection spread can be assessed using equation based SIR(D) models [15]. A quantum computational approach was used in order to deal with the space and time complexities of the huge amount of data related to epidemics and modelling. They find that the curve flattens once lockdowns are imposed. Generally epidemic models are created using Agent Based Models or Equation Based Models which have their advantages and disadvantages over one another [8].

Agent Based Models allow for the creation of heterogeneous agents with different characteristics which can affect their behaviours and hence their infection likelihood. Since in these types of models, the agents can make their own decisions and factor in more sorts of realistic scenarios. Agent based models are better than Equation Based Models because equation based models can get difficult to analyse as they are designed for homogeneous populations and analysing a large number of factors increases the number of equations to be solved in the model. Because of the homogenity, this implies that agents have equal probability of coming into contact with one another which is not the case in real life [8]. Hence by making use of Agent Based Models and reccommender systems to predict the customer buying behaviours to make it as realistic as possible.

Python can be used to create a simulation of a supermarket and its customers [20]. Although the objective of the current and previous studies' are very similar, they simulated a number of shops in a larger shopping centre with less complex inner structures. They had customers move around pseudo-randomly based on a correlation matrix between shops. This left a gap for us to improve upon this by by using Machine Learning models and shortest path algorithms to define where and how the customers move.

Modelling Covid-19 spread in a supermarket can be split into two parts: modeling the customer mobility in the supermarket and modelling the transmission of virus among the customers [26]. Agent-based model can be used to model the customer movements in a supermarket with a simple virus transmission model based on the virus exposure time. However, most of these papers do not consider the shopping behaviour of the customer while modelling the customer mobility. Different supermarkets have a varying base of customers with very different buying behaviours. In this paper, we have tried to fill this gap by introducing a Machine Learning based recommender system to predict customer buying behaviours for the next item order. Customer behavior modelling can be performed using two approaches: Individual-level and segment-level [16]. The individual-level approach predicts the customer behavior for an individual customer, whereas the segment-level approach predicts it for a group of customer sharing common properties such as buying habits, age group, etc.

Markov Chains and Matrix Factorization matrix methods can be used to model sequential behavior of users to predict the shopping basket of the customer [18]. To accommodate both a customer's sequential behavior and personal choices, a hierarchical representation model was used to predict customer behavior based on transactions and user profile data [24]. A novel Neural Network model is suggested to predict what a user is going to buy during their next purchase [22]. A dynamic Recurrent Basket Model is proposed based on a Recurrent Neural Network (RNN) to accommodate for dynamic customer representation and to predict customer behavior [27]. Both the Neural Network and the Recurrent Neural Network approaches use historical transaction data of customers. In an extension of previous work, another study used five different machine learning approaches: Logistic Regression, Support Vector Machines (SVM), Decision Trees, Random Forests, and Neural Networks [16].

Since the focus of this work is to model the spread of the virus in a supermarket, we predict the probability of customer visiting an aisle based on historical data to create customer paths in the supermarket. Extending the existing work, in this paper we use Recurrent Neural Networks and Gradient Boosting Algorithm (GBM) methods for customer behaviour prediction. [1]

3 Methodology

In this section we discuss about the customer shopping behaviour prediction done by a recommender system, the customer mobility model (which decides how a customer gets in to the store and moves around until they leave) and the virus transmission model (which forms the basis of the virus transmission amongst the visiting customers).

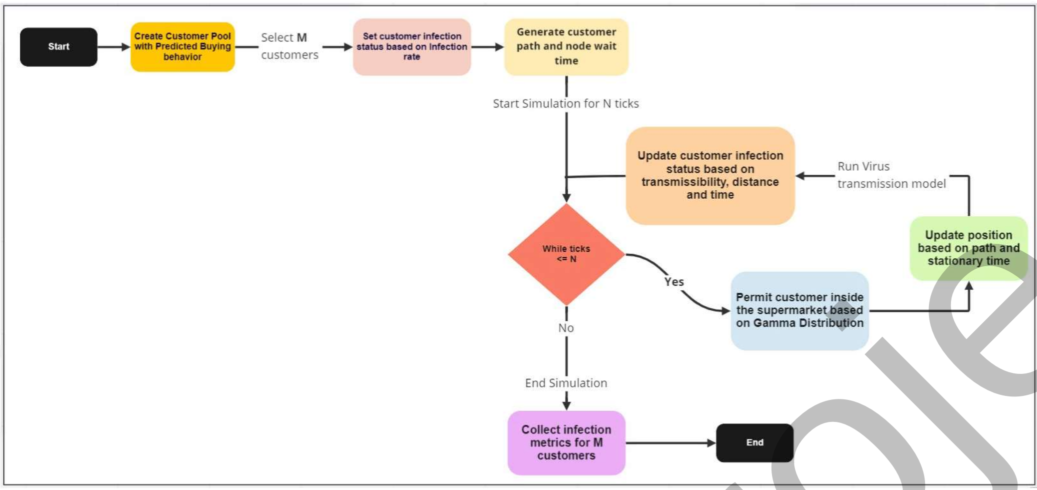

Figure 1: Flowchart representing the simulation process flow

The flow starts with the creation of a customer pool, with some customers initially set as infected based on the infection rate for the county with peak transmission rate [2] along with the infection duration representing the number of days the customer has been infected. Customer arrival into the supermarket at every tick is decided based on a Gamma distribution. A Machine Learning model is used to understand the buying patterns of customers using the Instacart dataset [9]. This recommender system creates a vector for each customer containing the probabilities of them visiting different aisles in the store. The store path and waiting time (in ticks) for each customer are then generated based on the aisle vectors. Simulation is then run for n ticks, with the position of the customers updated every tick once they get into the store based on the waiting time values assigned to them at each node. During the simulation, our virus transmission model decides the transmissibility based on a customer's infection status, contacts, and exposure time. The set of infected customers are then used for further analysis.

3.1 Customer Shopping Behaviour Prediction

The mobility of customers inside a supermarket depends on the products they are going purchase. Predicting the customers' shopping behaviour can help make the customer mobility model more realistic, which would help to extend this model to multiple supermarkets by using their own customer data. Being able to model customer mobility as per their own customer data would help supermarkets make COVID-19 policies specific to their customer shopping patterns. In this paper, we have used the Instacart dataset [9] to develop a customer shopping behaviour prediction model.

The instacart dataset is a public dataset released by *instacart.com* containing 3 million online grocery shopping transaction records from more than 200,000 users. A customer's purchase sequence is captured in the `order_products` table along with a time and date. The dataset contains almost all of the products found in most supermarkets in the `Products` table, grouped by `Aisles` and `Departments`. The dataset contains between 4 and 100 orders for each user, together with the product sequence for each order. It also tells you what week and hour the order was placed, as well as how long of a gap there was between orders. This dataset is used to train the recommendation system to output a customer's buying pattern for their next order.

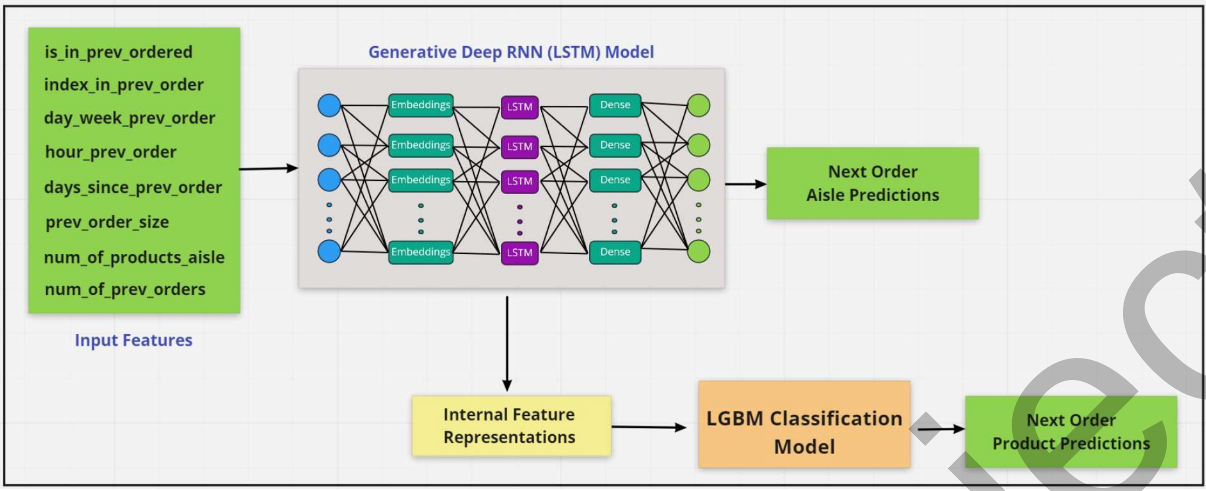

Figure 2: Recommender System Design : LGBM stacked on top of Deep RNN (LSTM)

The recommender system must forecast whether or not a certain product will be re-ordered in the user's next order given a user, a product, and the user's prior purchase history. Figure 2 shows the design of recommender system used in this paper. The recommender system was built with a stacking of models. First, a deep recurrent neural network (DRNN) was used as a generative model to fit part of the data. The internal representation of the DRNN model was used as feature for the second level model. The second level model was a Light Gradient Boosting (LGB) model which used the features from the first model and predicted whether a user buys a particular product in their next order. We mapped the product predictions to `Aisle` levels to get the recommendation at an aisle level, i.e. predicting whether a customer is going to visit an aisle in their next order.

The model was run on test data containing 5819 customers. Later, this aisle level prediction data of the customers was used to model the customer mobility inside the supermarket.

3.2 Customer Mobility Model

In this section, the design of customer mobility across a custom supermarket floor plan is discussed. The output from the previous customer behaviour prediction model was fed into to customer mobility model. Here we designed a custom supermarket and based on the customer buying behaviour sampled independent paths for each customer. In order for customers to move around the store, they need to follow a path to get from an initial point A to their destination B. As part of the simulation, a customer will be assigned a number of places, or coordinates, that they will go to as part of their trip. The path generator's job is to find the best path for them to achieve this, i.e. to solve the travelling salesman problem. We implemented two different methods of achieving this. The initial one involved a brute-force approach that iterates through all permutations of the places a customer needs to go to. The shortest path among these is then chosen.

The second method, which is the final chosen method for generating paths, is a form of the nearest neighbour approach. This method was initially chosen as it seemed like a more natural approach for the customer to take. In real life, a person would usually try to find the nearest next point that they want to go to, rather than optimising their entire overall path. Furthermore, it is more efficient than a complete brute-force search. The idea is to start at the customer's initial coordinate, or 'node', and find out which of the nodes they want to go to is currently the closest to them. This node is chosen as the next one in the path and we continue to iterate in a similar fashion until all of the planned nodes are visited. 3 shows a sample path in our custom supermarket design.

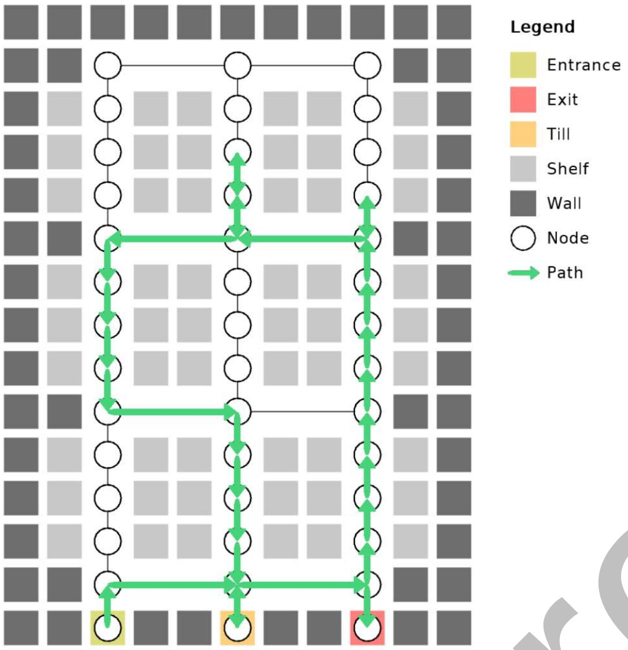

Figure 3: Customer Mobility Model: Network representation of sample store with example shopping path in green.

3.3 Virus Transmission Model

We model the Virus Transmission Probability by means of the dimensionless quantity R_0, also called the basic reproduction number [10]. R_0 is defined in Equation (1). It signifies the number of secondary cases that a single infected agent can produce in a completely susceptible population.

$$R_0 \propto \tau \times \bar{c} \times d \tag{1}$$

Where τ is the transmissibility, which is the probability of infection given contact between a susceptible and infected individual, $\bar{c}$ is the average rate of contact between susceptible and individuals infected, and d is the duration of infectiousness. Re-writing Equation (1) in terms of transmissibility we get:

$$\tau = \frac{R_0}{\bar{c} \times d} \tag{2}$$

We chose the value of R_0 as 2.5 since the pathogen that causes COVID-19 has an R0 range of 2-2.5 [6]. Considering the 14 day recommended quarantine period and the mean incubation period of COVID-19 is 7 days (around a week) [4]. Since an infected individual is capable of spreading the infection after the incubation period, the value of d was chosen as a number between 1-7. For a lower value of d, the transmissibility is high according to Equation (2). For example, on the 8th day of infection, which is 1 day after the incubation period (d=1), a person is more likely to infect others compared to the 14th day of infection, which is 7 days after the incubation period (d=7). We chose $\bar{c}$ as a reasonable estimate of 3 [13] for the purpose of the simulation.

Initial infection status of both the customers are checked first and the node (aisle) of the customers are checked. If either of the customers are infected then the longer the two customers stay in same node, the higher is the chance of the susceptible customer being infected.

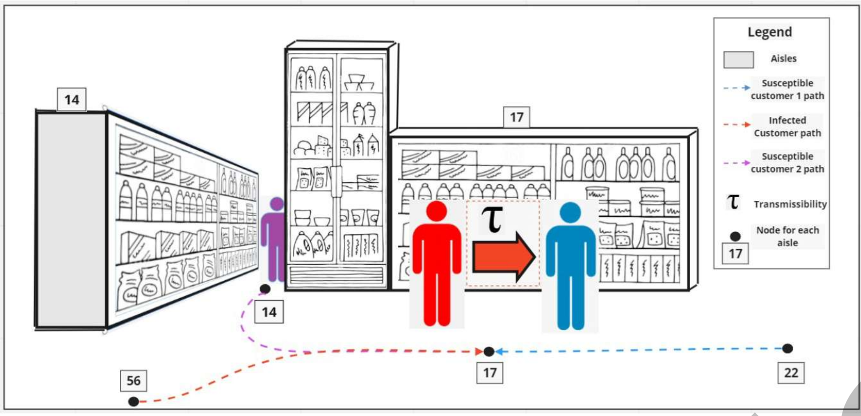

Figure 4: Susceptible customer (blue, purple) is infected with a transmissibility of τ by infected customer (red)

4 Implementation

4.1 Simulation

The simulation used for this report is written in Python and attempts to simulate the behaviour of numerous customers in a single store on a particular day. The visualisation and simulation of the experiments are created from the `pyglet` [1] and `networkx`[2] libraries. The `pyglet` library is commonly used in simulation environments and contains a rich Application Programming Interface (API) for cross platform windowing and multimedia. The `networkx` library is used for network analysis supporting non linear data structures such as trees, digraphs and so on with standard graph algorithms. It is used as the path of customers towards the aisle, cash register, and exit. Prior to starting the simulation, a number of configuration parameters are loaded from a file. The list of potential customers, as well as the graph-based layout of the store, are generated based on these parameters.

The design of the supermarket's floor plan as shown in figure (3) follows a standard layout from a study on designs of supermarket store layouts, involving the arrangement of entrances, exits, and aisles [14]. The design implements 134 aisles from the *Instacart dataset*, with a node from the `networkx` used to represent an aisle.

The progress of the simulation is carried out via a *tick* counter. Each tick corresponds to a certain amount of real time, e.g. five seconds. The simulation, starting at the beginning of the simulated day, repeatedly calls the `tick()` function which increments the tick counter and carries out a number of computational tasks such as determining if a new customer should enter the store or if an existing customer should leave, updating the positions of each customer and determining if two customers are within range of infecting each other. Once the simulation has been completed, i.e. the tick counter has reached a certain maximum value corresponding to the time the store closes, the raw results of the simulation are collated and presented.

4.2 Customer Arrival

New customers are added to the queue using a bimodal gamma distribution that models the probability of a customer's arrival. The gamma distribution, or probability density function (PDF), of customer arrival is hence modelled to mimic the entire length of the simulation. The probability that a customer will join the queue at a particular tick is hence the corresponding value of the PDF for that tick/ point in time, scaled

[1] http://pyglet.org/
[2] https://networkx.org/

appropriately based on the total simulation time and `xlim` of the distribution. The choice of whether a customer will be added to the queue is generated randomly as True/False using this value of probability. The PDF for a Gamma Distribution is as follows:

$$f_{k,\theta}(x) = \frac{x^{k-1}e^{\frac{-x}{\theta}}}{\theta^k\Gamma(k)}, \; x > 0 \tag{3}$$

Where k is the shape parameter that controls the overall characteristic of the distribution and θ is the scale parameter that controls the horizontal and vertical stretch of the distribution. For $k \leq 1$, the gamma PDF is strictly decreasing, and for $k > 1$ it modifies into a left skewed bell curve, which gets increasingly more symmetrical and less skewed as the value of k increases. As the value of θ increases, the distribution becomes wider and more shallow, but maintains its inherent shape and skew determined by k.

Figure 5: Bimodal Distributions of Footfall in different stores [5]

Figure 6: Modelled Bimodal Distribution for Customer Arrival

We initially considered the use of a Poisson distribution to model customer arrival but decided against it as this kind of distribution is mostly used to model the number of events that occur in a fixed interval of time, whereas we seek to find the probability that a single customer will enter the store at a particular tick. We also obtained footfall data for stores like Lidl, IKEA, Tesco, etc. from Google [5] in order to gain an understanding of what the busy times of these stores are and what kind of a distribution that would correspond to. Examples of sample data we came across are shown in Figure [5]. We hence did not use a normal distribution as this does not accurately reflect the bimodality of the footfall data, i.e. there can be different peaks of customer arrival within a single day. We picked the shape and scale parameters for the bimodal gamma distribution with care to effectively model the observed footfall data from Google. Figure [6] shows the final modelled distribution.

4.3 Customer Stationary Time in Each Aisle

Once the customers are in the store, their paths will be generated in the simulation by the model based on their past purchases. As in real life, during a particular visit each customer spends some time in each of the aisles. We aim to replicate this behaviour in the simulation based on two factors. First, since the probabilities of each customer visiting each node will be obtained from the model, the higher the probability of visiting an aisle, the higher the stationary/waiting time at that aisle will be for that customer. Secondly, as explained in [11], the stationary time was modelled based on a Poisson distribution and it gave deep insights into the relationship between customer in-aisle stationary time and the probability of visiting that particular aisle. A Poisson distribution will model stationary time within a fixed time interval and each stationary time will be an iid and will solely be related to the probability of a customer visiting that aisle. The PDF of a Poisson Distribution is as follows:

$$f(k,\lambda) = \frac{\lambda^k e^{-\lambda}}{k!} \tag{4}$$

Where k is the number of events or, in this case, the number of ticks a customer will spend on a specific aisle. λ will define the probability of these events, or the number of ticks on each aisle for the customer.

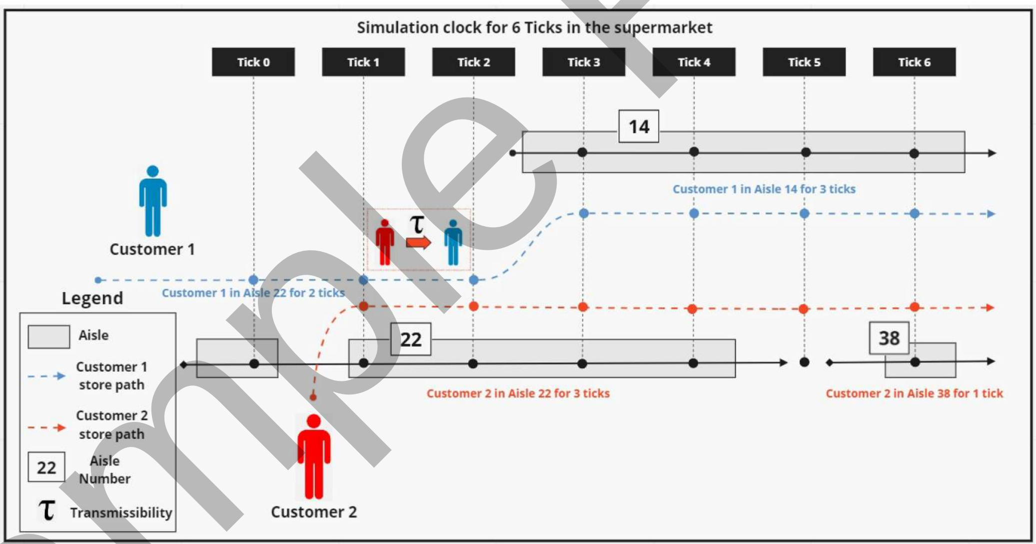

Figure 7: Customers Stationary time in simulation clock for 6 ticks in the supermarket

It can be seen in Figure [7] how 2 customers visit and spend their modelled stationary time in each of the aisles. Figure [7] is based on a simulation clock time for just 6 ticks in the supermarket. It can be seen that customer 1 is the first customer in the store and they spend a duration of 2 ticks in aisle 22. Meanwhile, another customer comes in and spends nearly 4 ticks in the same aisle. These stationary times were modelled from a Poisson distribution, with stationary time assigned based on the probability of the customer visiting the aisle to maximise sales profit, as mentioned in [11].

Once each customer has completed their stationary time in their respective aisles, they will move to a different aisle. For example, it can be seen in Figure [7] that customer 1 moved from aisle 22 to aisle 14 once their stationary time on aisle 22 was completed. This was done for all the customers to mimic the same routines of customers spending time in each aisle based on their past preferences. This also brings into account how this habit can have a negative affect during a pandemic as spending more time in each aisle can aggravate the spread of the epidemic among customers.

5 RESULT

In this section we demonstrate the results observed from 100 simulation performed. Each simulation represents one working day of the supermarket. In each simulation the number of customers in store, number of newly infected, amount of exposure time is recorded at the end of each tick (5 seconds of time).

Table 1 holds the default parameters used for simulation of the spread.

Parameter	Default value / Assumption	Reference
Basic Reproduction Number, R0	2.5	[6]
Average contact between susceptible customer and infected customer, $\bar{c}$	3	[13]
Infectiousness duration, d	1-7 days	[4]
Customer arrival probability modelled using bimodal gamma distribution	arrival_probability = (gamma.pdf(x, a=3.5, scale=4.5) * 3) + (gamma.pdf(x, a=18, scale=2) * 4)	[5]
Tick duration	5 seconds	
Initial infection rate	1.2362%	[2]

Table 1: Parameter values in our customer mobility and virus transmission models

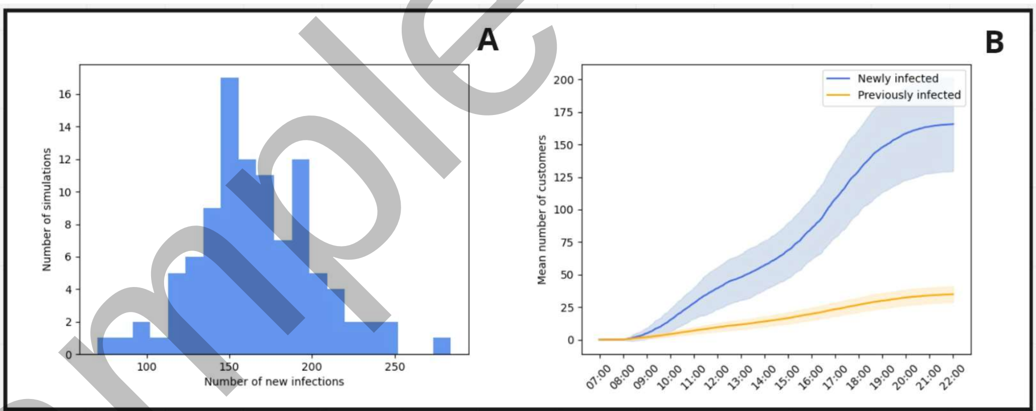

Figure 8: **(A)** - Distribution of "Number of new infections" over 100 simulations. **(B)** - Trend of new infections over the opening hours of the supermarket.

Over the 100 simulations, for every simulation the number of newly infected customers and previously infected customers were recorded. Furthermore, the exposure time of susceptible customers were also recorded. This helped us understand with the given transmissibility τ discussed in section 3.3 how the virus is transmitted in the discussed customer mobility model in section 3.2.

From Figure 8 we can observe that the number of new infections over 100 simulations follows a normal distribution with mean of nearly 175 new infections every working day of super market. Surprisingly, even the lowest number of new infections stays close 100 out of average 3000 customers per day. As the day proceeds, we can observe in Figure 8 plot (B) that number of new infections increases significantly over time.

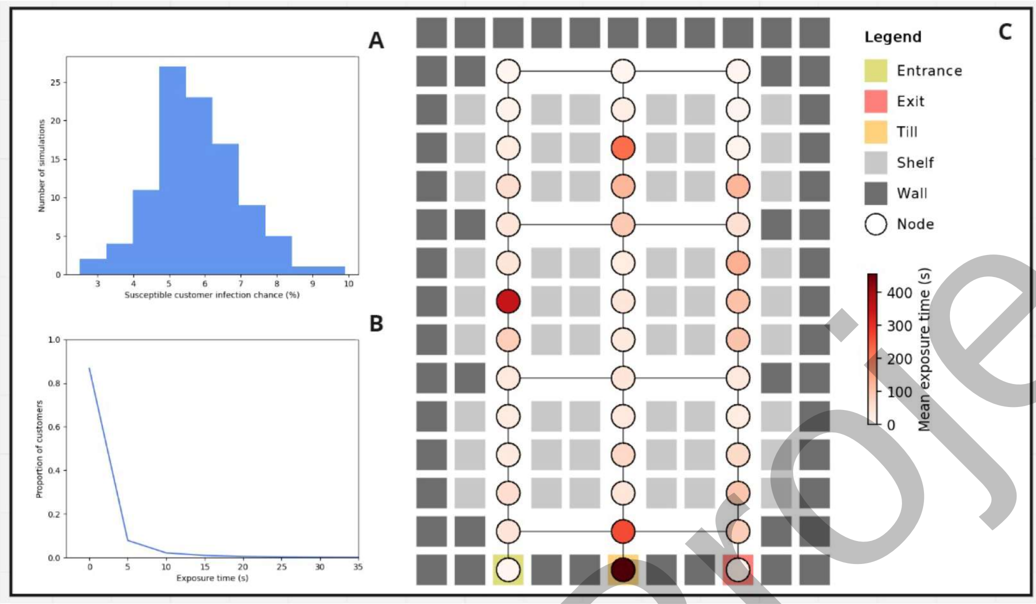

Figure 9: Plot for exposure time and chance of infection across 100 simulations. **(A)**- Distribution of chance of infection per susceptible customer across 100 simulations. **(B)**- Distribution of total customer exposure time for all exposed customers (i.e., susceptible customers with positive exposure time) across 1000 simulations.. **(C)**- Total exposure time per node across 100 simulations. Some nodes of the store show significantly higher amount of exposure time than others.

The chance of infection and exposure time for each susceptible customer is recorded in every simulation. In Figure 9 plot (A) shows that the chance of infection of susceptible customer follows a normal distribution with mean of nearly 6% customers being susceptible out of average 3000 customers per day. The plot (B) shows that the exposure time follows exponential distribution and mean exposure time of susceptible customer is around 1.08 seconds. The plot (C) shows the high exposure nodes in the supermarket. It is obvious that the exposure near till is significantly high. Interestingly, some other nodes with high exposure time could be because of more popular items being present in those aisles. More popular items in an aisle could lead to more customers coming to that area and increasing the average exposure time there. These areas are potential hotspots of transmission.

Table 2 shows the important metrics observed from the simulation. The default setting of simulation can be changed, along with change of store layout, change of infection rate (due to facemasks), introducing self-checkout etc. Multiple simulations can be done with varying settings which can be further studied to introduce policies to reduce the spread of Covid-19.

Metric	Mean	Standard Deviation
Number of Daily Customers	2918.75	44.08
Number of infected customers	34.95	6.1
Number of susceptible customers	2883.8	43.97
Mean shopping time (sec)	264.69	1.28
Total exposure time (sec)	6209.8	1155.65
Mean exposure time (sec) per susceptible customer	1.08	0.2
Total exposure time (sec) per infected customer	89.29	9.42
Proportion of susceptible customer with any exposure	0.1225	0.0216
Number of new infections	165.71	36.07
Proportion of infections per susceptible customer	0.05746	0.0125

Table 2: **Simulation Resuls** : The mean and standard deviation of each metric across 100 simulations.

6 Conclusion

In this paper, we presented a model for Covid-19 virus transmission in supermarkets based on an agent-based model of customers traversing from aisle to aisle and being exposed to potential virus infection when in the same aisle as an infected customer. We calculated the risk of virus transmission from an infected customer to a susceptible client based on a number of parameters, including the likelihood of infection, the average exposure with sick customers, and how long the infected customers have been infected.

We used machine learning based customer behaviour predictions methods to simulate realistic movement of customers in the supermarket and ran over 100 simulations to conclusively draw the results.

Based on our Covid-19 spread model, supermarket owners can identify potential hotspots in their supermarkets and introduce new policies like limiting the number of customers, rearraging items to avoid hotspots and introducing one aisle policy. Our simulation can be used by retailers to reduce the risks of spread of epidemics like Covid-19 and consequently reduce the economical and societal impact of such epidemics.

7 Further Work

The project can be further extended through the customer shopping behaviour prediction, customer mobility model and virus transmission model. Different floor plan design can be used in the simulation with varying sizes and layout of the supermarket. The customer's entry of supermarket store can be determined by their transaction history involving their shopping behaviour prediction instead of a fix value imposed by existing regulations. Further more different non-pharmaceutical intervention settings such as face masks, social distancing rules and etc can result in different probability of spreading the virus in store. The customer's health history information such as vaccine status and prior health history can affect the virus transmission model. The output simulation and results based on the layout of store can be used to generate different recommendations for customer in different supermarket.

References

[1] Instacart Market Basket Analysis. https://github.com/sjvasquez/instacart-basket-prediction.

[2] Profile of county & lea statistics. https://covid-19.geohive.ie/pages/detailed-county-statistics.

[3] R. Alguliyev, R. Aliguliyev, and F. Yusifov. Graph modelling for tracking the covid-19 pandemic spread. *Infectious Disease Modelling*, 6:112–122, 2021. doi: 10.1016/j.idm.2020.12.002

[4] C. Cheng, D. Zhang, D. Dang, J. Geng, P. Zhu1, M. Yuan, R. Liang, H. Yang, Y. Jin, J. Xie, S. Chen, and G. Duan1. The incubation period of covid-19: a global meta-analysis of 53 studies and a chinese observation study of 11 545 patients. 2021.

[5] M. D'Zmura. Behind the scenes: Popular times and live busyness information, 2020.

[6] M. Fisher. R0, the messy metric that may soon shape our lives, explained. 2020.

[7] T. Galbadage, B. M. Peterson, and R. S. Gunasekera. Does covid-19 spread through droplets alone? *Frontiers in public health*, 8:163, 2020.

[8] E. Hunter, B. Mac Namee, and J. D. Kelleher. A comparison of agent-based models and equation based models for infectious disease epidemiology. In *AICS*, pp. 33–44, 2018.

[9] Instacart. The Instacart Online Grocery Shopping Dataset. https://www.instacart.com/datasets/grocery-shopping-2017, 2017.

[10] J. H. Jones. Notes on r0, 2007.

[11] Y. Kaneko and K. Yada. Bayesian hidden markov model for evaluating the influence of in-store stationary time of customers on their purchase behavior. In *2018 5th Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE)*, pp. 142–148, 2018. doi: 10.1109/APWConCSE.2018.00031

[12] P. Keskinocak, B. E. Oruc, A. Baxter, J. Asplund, and N. Serban. The impact of social distancing on covid19 spread: State of georgia case study. *Plos one*, 15(10):e0239798, 2020.

[13] C. G. McAllen, F. B. Patrick Wall, M. Codd, E. Gormley, C. Walsh, J. Duggan, T. B. Murphy, P. Nolan, B. Smyth, K. O'Brien, C. Teller, M. J. Green, L. O'Grady, K. Culhane, C. Buckley, C. Carroll, S. Doyle, J. Martin, and S. J. More. Numbers of close contacts of individuals infected with sars-cov-2 and their association with government intervention strategies. 2021.

[14] M. Ohta and Y. Higuchi. Study on the design of supermarket store layouts: The principle of "sales magnet ". *International Journal of Economics and Management Engineering*, 7(1):209–212, 2013.

[15] A. Padhi, S. Pradhan, P. P. Sahu, S. Kalyani, B. K. Behera, and P. K. Panigrahi. Studying the effect of lockdown using epidemiological modelling of covid-19 and a quantum computational approach using the ising spin interaction. *medRxiv*, 2020. doi: 10.1101/2020.06.06.20124024

[16] S. Peker, A. Kocyigit, and P. E. Eren. An empirical comparison of customer behavior modeling approaches for shopping list prediction. In *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, pp. 1220–1225, 2018. doi: 10.23919/MIPRO.2018.8400221

[17] N. Prakash, B. Srivastava, S. Singh, S. Sharma, and S. Jain. Effectiveness of social distancing interventions in containing covid-19 incidence: International evidence using kalman filter. *Economics & Human Biology*, 44:101091, 2022.

[18] S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme. Factorizing personalized markov chains for next-basket recommendation. In *Proceedings of the 19th International Conference on World Wide Web*, WWW '10, p. 811–820. Association for Computing Machinery, New York, NY, USA, 2010. doi: 10.1145/1772690.1772773

[19] C. I. Siettos and L. Russo. Mathematical modeling of infectious disease dynamics. *Virulence*, 4(4):295–306, 2013.

[20] A. Teehan, C. Shrivastava, P. Sajeekumar, P. Madaan, S. V. Kulkarni, S. Tomar, S. Jacob, and S. Ramenahalli. Geometric/epidemic modelling of covid-19 spread. 2021.

[21] T. P. Velavan and C. G. Meyer. The covid-19 epidemic. *Tropical Medicine & International Health*, 25(3):278–280, 2020. doi: 10.1111/tmi.13383

[22] S. Wan, Y. Lan, P. Wang, J. Guo, J. Xu, and X. Cheng. Next basket recommendation with neural networks. In *RecSys Posters*, 2015.

[23] H. Wang, K. R. Paulson, S. A. Pease, S. Watson, H. Comfort, P. Zheng, A. Y. Aravkin, C. Bisignano, R. M. Barber, T. Alam, et al. Estimating excess mortality due to the covid-19 pandemic: a systematic analysis of covid-19-related mortality, 2020–21. *The Lancet*, 2022.

[24] P. Wang, J. Guo, Y. Lan, J. Xu, S. Wan, and X. Cheng. Learning hierarchical representation model for nextbasket recommendation. SIGIR '15, p. 403–412. Association for Computing Machinery, New York, NY, USA, 2015. doi: 10.1145/2766462.2767694

[25] Y. Xiang, Y. Jia, L. Chen, L. Guo, B. Shu, and E. Long. Covid-19 epidemic prediction and the impact of public health interventions: A review of covid-19 epidemic models. *Infectious Disease Modelling*, 6:324–342, 2021.

[26] F. Ying and N. O'Clery. Modelling covid-19 transmission in supermarkets using an agent-based model. *PLOS ONE*, 16(4):e0249821, 2021. doi: 10.1371/journal.pone.0249821

[27] F. Yu, Q. Liu, S. Wu, L. Wang, and T. Tan. A dynamic recurrent model for next basket recommendation. SIGIR '16, p. 729–732. Association for Computing Machinery, New York, NY, USA, 2016. doi: 10.1145/2911451.2914683