Geometric Modeling of COVID-19 epidemic Spread

Zhouhao Rui, BSc

A Dissertation

Present to the University of Dublin, Trinity College

in partial fulfilment of the requirements for the degree of

Master of Science in Computer Science (Data Science)

Supervisor: Dr.Bahman Honari

August 2022

Declaration

I, the undersigned, declare that this work has not previously been submitted as an exercise for a degree at this, or any other University, and that unless otherwise stated, is my own work.

Zhouhao Rui

August 19, 2022

Permission to Lend and/or Copy

I, the undersigned, agree that Trinity College Library may lend or copy this thesis upon request.

Zhouhao Rui

August 19, 2022

Geometric Modeling of COVID-19 Epidemic Spread

Zhouhao Rui, Master of Science in Computer Science University of Dublin, Trinity College,

2022

Supervisor: Dr.Bahman Honari

Abstract

Following the early COVID-19 pandemic in 2020, supermarkets in Europe implemented a policy of limiting the maximum number of people to control the spread of the epidemic. However, this did not prevent the clustering of customers within supermarkets, especially in the shopping area. In order to find ways to reduce the infection rate and risk of exposure to virus, this thesis constructs a supermarket model based on Delay tolerant network and Graph theory to simulate customer mobilities. The Susceptible-Exposed-Infected (SEI) model was also applied to implement the process of virus spread, where susceptible people are considered to be infected after more than 5 minutes of exposure to a virus. It was found that by controlling the purchasing time in shopping area, the infection rate could be reduced from 5.4% to 3%, and by accelerating the checking-out speed, the exposure time in the checking area could be reduced from 52% to 46% of the total exposure time. The experiment showed that the risk of virus transmission could be reduced by controlling the shopping time of customers and increasing the number of checkouts per unit time.
Air circulation is the predominant mode of transmission in enclosed areas, and the Susceptible-Exposed-Infected-Recovered model with indoor airborne transmission was used to study the long-term outbreak development. Experiments have shown that outbreaks last longer when the quanta exchange rate of infected individuals reaches 800 ACH or more because the R0 value exceeds 2.

Keywords: DTN, SEIR, Double Queuing System, Graph Theory, Epidemic Modeling

Acknowledgments

I would like to express my gratitude to my supervisor, Dr. Bahman Honari. He gave me a lot of guidance about the mathematical model and helped me understand the topic.
I also very much appreciate my friends and family who support and encourage during the dissertation period.

Zhouhao Rui

University of Dublin, Trinity College
August 2021

Contents

List of Figures

List of Tables

1 Introduction & Motivation

1.1 Introduction

Covid-19 has been a highly serious disease since early 2020. It is easy to spread, and the R0 value of the original variants is between 2-3, which is higher than the regular flu. The virus can transmit in several ways, including spreading from the mouth or nose in small liquid particles, long-range airborne transmission in aerosols and touching skin surface (WHO, 2022). According to (EPA, 2021), particles from an infected person can move throughout an entire room or indoor space. The particles can also linger in the air after a person has left the room. Therefore, it’s very easy for the virus to spread in public rooms. During last year's lockdown, the government made a policy of limiting the population to control the spread of the outbreak in some enclosed areas, such as supermarkets. For example, Lidl had an alarm system where a red light indicated that the number of people had reached the maximum and customers had to wait outside the door. This did help to control the epidemic spread, but it ignored the issue of customer mobilities. Customers may gather in a certain area and this cannot be prevented by only limiting the number of people. It is important to control the number of people and to ensure that the distance between customers is more than 6 feet, as studies have shown that within this distance the aerosols containing the virus can be transmitted through the respiratory tract (HSE, 2022).

In this paper, the supermarket model is used to limit the number of people and to control the shopping time of the customers and the speed of the checkout area in order to reduce the contagion rate and the possibility of exposure.

1.2 Motivation

The main motivation for research into the spread of the epidemic in supermarkets is the more significant potential for virus transmission in enclosed areas. In poorly ventilated rooms, as virus-carrying aerosols are exhaled, so is the CO_2. It's easy for the CO_2 to accumulate with the virus (Lewis, 2021). Therefore, controlling the exposure time of people in supermarkets is essential and is an improvement to the current policy. At the same time, there is no previous research focusing on modelling airborne infectious disease models in the supermarket scenario, which can be attempted in this project. The objectives of this paper are as follows.

1. to analyse and identify factors that can reduce the risk of exposure and transmission rates.
2. to build an epidemic model with indoor airborne factors and to assess the long-term risk of an outbreak.

1.3 Overview

The paper structure is explained as follows. To start with, Chapter one introduces the background of Covid-19 and current precautions in the supermarket. The motivations of this dissertation are listed after the background part. In Chapter 2, theoretical knowledge and documentation of epidemic spread in supermarkets will be reviewed, including the Susceptible-Exposed-Infected-Recovered model, Delay tolerant model, Double Queuing System and Customer Mobility Model. Chapter 3 introduces the mathematical knowledge and the methodologies applied. The simulation process and experiment environment are covered in Chapter 4, together with the source data included in this project. Chapter 5 focuses on introducing the result analysis and finding out the factors influencing epidemic spread in the supermarket and the long-term epidemic tendency. Final chapter makes a brief conclusion and explains the future work.

2 Literature Review

In the past two years, several studies have been produced on the spread of covid-19, including the well-known Susceptible-Exposed-Infectious-Recovered (SEIR) and Susceptible-Infectious-Recovered (SIR) models, which use a large amount of data to predict the probable spread size and the peak time of an outbreak (Xiang Zhou, 2020) and the spread of the epidemic in different communities. (Ian Cooper 1, 2020). SIR and SEIR models work well for the spread of viruses in public areas and for predicting outbreaks over a long period, as long as the influencing parameters in the models are determined accurately, i.e., the progression rate and recovery rate of exposure to the infected, the growth of infection can be simulated in a more optimal state. While in a closed area, the SEIR model would be flawed, as there would be no post-infection recovery in a short aggregation time, and the difference between susceptible and non-susceptible individuals would be minimal in a confined space. One new idea is to simplify the SEIR model and to classify customers into two categories, Infected and Susceptible, using exposure time as a condition for infection and the average exposure time of the group as an indicator of the severity of the epidemic transmission (Fabian Ying, 2021). These studies are a reasonable quantification of the efficiency and process of epidemic propagation in different scenarios.

Another problem differs from the macroscopic prediction of epidemic trends; people constantly move in supermarkets and need to be treated as moving nodes in the simulation. DTN (Delay Tolerant Network) is a model often used to simulate epidemic spread between mobile objects. DTN uses a store-carry-forward mechanism for message forwarding, which can effectively solve the uncertainty and intermittent interruption of customer movement. The combination with the SIR model allows for a comprehensive presentation of the characteristics of the transmission of epidemic disease in a supermarket. This chapter will then introduce and discuss some detailed recent works that are helpful for the topic of epidemic simulation in the supermarket. To begin with, the SEIR and relevant family models for the epidemic spread will be exploited. The second sub-section focuses on the findings about DTN and its advantages in carrying out the supermarket epidemic simulation. The following section will analyse the supermarket queuing system and customer mobilities. Final Conclusions contains the summary of the findings and explore some new perspective that are not studied in the previous documents.

2.1 SEIR Model

Research findings on epidemic spread date back to the early 20th century when Kermack and McKendrick first proposed the SIR (Susceptible, Infected, Recovered) model in 1927, which attempted to use partial differential equations to construct a classification and transformation of susceptible, infected, and recovered individuals by age. Three groups of healthy people who are susceptible (S), infected individuals (I), and removed individuals either by them being recovered and immunized or by their death (R) (Saina Abolmaali, 2021). The SIR model is defined in equation (2.1):

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

$$\frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

(2.1)

Where β represents the infection rate, which indicates the probability that one susceptible person will adapt to the infected one, N is the total number of the population in the model, which is the same with $N = S + I + R$. γ is the parameter that controls the recovery rate, which is the probability that an infected person will recover from the disease. Here the derivative represents the conversion rate of different individual populations per unit of time. $\frac{dS}{dt}$ equals the negative number of infected people from the susceptible group. $\frac{dR}{dt}$ equals the number of recovered people from the infected group. $\frac{dI}{dt}$ is the difference between the above two indicators.

The SIR model above only simulates the spread of the epidemic in an ideal short-period scenario. To take the vital dynamics of population into account, in 1932 and 1933, (McKendrick, 1927) added birth rates, mortality rates, and locations of spread into the model, which catered to the survival and migration of the population characteristics. Assume that the death rate μ, birth rate Λ (Wikipedia, 2022) , the SIR model should evolutes to equation (2.2):

$$\frac{dS}{dt} = \Lambda - \mu S - \frac{\beta IS}{N}$$

$$\frac{dI}{dt} = \frac{\beta IS}{N} - \gamma I - \mu I$$

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

(2.2)

In the above equations, the $\frac{dS}{dt}$ equals the difference between birth and death rates minus infection rate in the susceptible group, and infected group deaths were considered on the basis of the SIR model when exploring the new equation of $\frac{dI}{dt}$. In the same way, for the directive $\frac{dR}{dt}$, the influence of the death rate in the recovered group is also taken into account. To summarize all the equations and findings above, a new indicator called reproduction number R_0 was defined, which is equal to $\frac{\beta}{\mu+\gamma}$ (Ross Beckley, 2013). When R_0 is smaller than 1, the epidemic archives DFE (disease free equilibrium), which means the epidemic should be under control.

The vital dynamics SIR model has solved some problems and made the epidemic model more suitable for natural populations and realistic scenarios. However, it dismissed one crucial factor that many infectious diseases delay during transmission and that there is a time threshold in which the group gradually transitions from susceptible to infected ones. Considering only the direct transition between susceptibility and infection would lead to overestimating the infection rate. A new transformed model SEIR was built to satisfy this requirement. (MULDOWNEY, 1994) was introducing the mathematic model in detail. The basic SEIR model has four different groups which are the three groups introduced above and the extra exposed group.

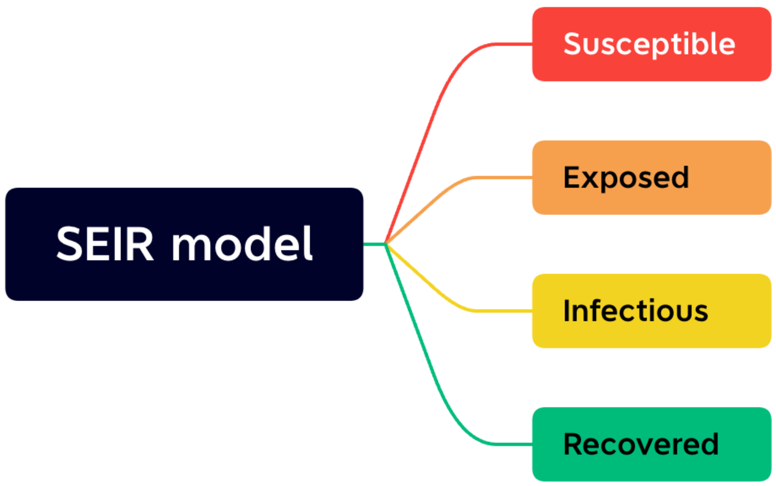

Figure 2.1 SEIR Model

In the *Figure 2.1* SEIR Model, the evolution sequence of the different groups in the model is explained. Assume that the death rate and birth rate are kept the same, which helps to control the total number is constant. The new model is defined in equation (2.3):

$$\frac{dS}{dt} = \mu N - \mu S - \frac{\beta IS}{N}$$

$$\frac{dE}{dt} = \frac{\beta IS}{N} - (\mu + \alpha)E$$

$$\frac{dI}{dt} = \alpha E - (\gamma + u)I$$

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

(2.3)

The average spread latency is α and is under exponential distribution. $\frac{dS}{dt}$ and $\frac{dR}{dt}$ are the same as the above SIR model, while the $\frac{dE}{dt}$ is the number of exposed people except for the people who died and transformed into the infected group. The $\frac{dI}{dt}$ is the people who have been exposed to the virus for α minutes except for the ones who have recovered and died. SEIR thoroughly considered the situations of the exposure latency, which has helped to build the simulations of the epidemic model in real lives.

The SEIR described above provides an excellent solution to the problem at the macro level, for example, by collecting data on infections in a country or a city for the last month and then going on to predict the number of infections for the following 1-3 months, as well as determining whether the growth of the epidemic will be dynamically balanced based on the calculated R-value. However, when explicitly exploiting the development of the outbreak over a short period in an enclosed space, more aspects need to be considered, and the model needs to be adjusted accordingly. Airborne infection is the most common portal of entry for viruses into the human body. A resting human inhales around 2 gallons of fresh air every minute, which may contain the virus from the cough or sneeze of an infected individual (Louten, 2016). Because of the confinement in indoor environments, air mobility has become a key point for researchers, and the SIER model has been reorganized afterwards. Combining ventilation-based models of indoor airborne transmission with the SIR model (Laura Gammaitoni, 1997)is a valuable way to evaluate the virus spread in a closed area. The equation (2.4) associates the average virus quanta exchange rate q with the room size V, pulmonary ventilation rate of susceptible p, and ventilation rate A (AC/h) are as follows:

$$\frac{dS}{dt} = \frac{-p}{V} QS$$

$$\frac{dQ}{dt} = -AQ + qI$$

(2.4)

Q is the total amount of air that contains the virus. Here $\frac{dS}{dt}$ is the average infection rate of the susceptible, which corresponds to the amount of the virus quanta. At the same time, the $\frac{dQ}{dt}$ is the total virus quanta produced by infected people except for the escaping viral air. Assuming continuous generation of quanta by the infectors and a steady state of ventilation ($\frac{dQ}{dt} = 0$) (C. J. NOAKES, 2006), the basic SIR model is adapted to equation (2.5):

$$\frac{dS}{dt} = -\frac{pq}{VA} IS$$

$$\frac{dI}{dt} = \frac{pQ}{VA} IS - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

(2.5)

In the normal indoor environment, birth and mortality rates do not need to be considered, but incubation period is still a main factor. SEIR models that do not take into account vital dynamics is defined in equation (2.6) (C. J. NOAKES, 2006):

$$\frac{dS}{dt} = -\frac{pq}{VA} IS$$

$$\frac{dE}{dt} = \frac{pq}{VA} IS - \alpha E$$

$$\frac{dI}{dt} = \alpha E - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

(2.6)

Here α is the exposure rate of the infectors and the reproduction number is $\frac{pq}{VA}\frac{N}{\gamma}$. This reproduction number has fully considered the influence of air ventilation and room size.

The ventilation based SEIR model has successfully shown the role of air mobility in the spread of the epidemic, but because the population in the mall has high mobility, i.e., move in and out. The Recovery group is difficult to define in the mall scenario. In (Fabian Ying, 2021), a new virus transmission model is introduced. The model is in the format of SEI and removes the recovery people. Every customer has individual exposure time to infectious people, and the transmission rate is β. When the probability βE is large than some threshold number, it can be concluded that this customer is infected. It simplified the SEIR model and provided a straightforward way to measure the spread of the epidemic indoors.

The above materials have given many insights into the epidemic modelling from the global view and enclosed area; the next section will review some of the documents related to the behaviours of the indoor customers.

2.2 DTN model

Delay-tolerant Networks (DTN) is a wireless network which operates in a challenging environment. Compared with the other ad-hoc network, DTN has high error rates and long latency. It overcomes the obstacles and high uncertainty of the nodes in the network by using the standard switching method: store-carry-forward. Once getting the message, the node will keep the messages and broadcast them to the others in the valid range.

The history of DTN can date back to the 1970s; with the invention of more personnel computers, researchers started looking for the routing protocol between moving objects. (Fall, 2003) first created the definition of DTN and pointed out some drawbacks in the traditional TCP/IP architecture. Although the packet drop rate is low, instability can occur under certain conditions. For example, nodes moving at high speeds, high transmission latency in satellite systems, and CPU bottlenecks in sensor systems. It tried to attach the nodes to the network’s edge using agents and use the data carry-transit mode to deal with the challenging network. It operated in an overlay architecture and provided a store-and-forward messaging system between networks with various communication protocols. The DTN route forwarding should be determined in advance based on the predictability and latency of transmitting messages to contacts. It also introduced a flow control plan with priority queue architecture to pick the data carrying and forwarding order.

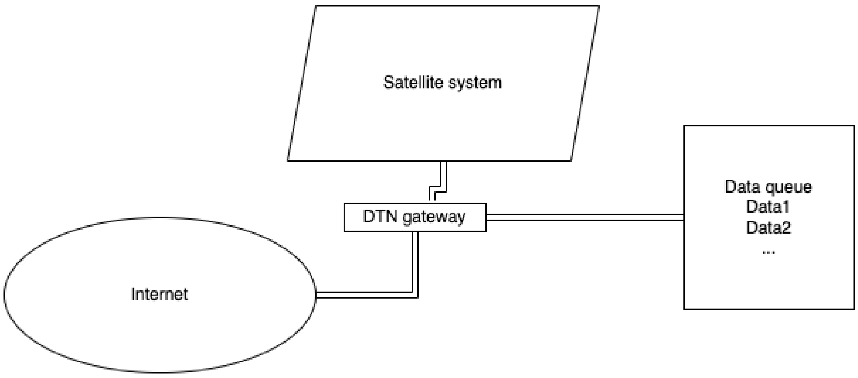

Figure 2.2 DTN among dissimilar network

The *Figure 2.2* DTN among dissimilar network shows how DTN works between the edges of two dissimilar network architectures and how to manage the data flow.

The routing protocol for DTN is a crucial point to be considered due to the unreliable messaging mechanism in the challenging network architecture. (Tamer Abdelkader, 2016) has pointed out that the routing protocol should achieve a balance between maximising packet delivery ratio and minimising the number of transmissions in the network. Four protocols were introduced in the document: epidemic routing, spray and wait (SnW)protocol, PROPHET routing, and MAXDROP routing. Epidemic routing is the first protocol designed for DTN and in the format of blinding routing and full blooding. In the epidemic routing, every data packet has one unique ID associated with all the copies it generates. When two nodes meet, they exchange the list of all packets' IDs they have in the buffer. The epidemic routing ensures the lowest latency between the data transmission between nodes due to the flooding strategy. SnW protocol is similar to epidemic routing but limits the number of data copies in the network. It keeps the data copies only before the buffer overflow and lifetime expiry. Compared with the epidemic routing, SnW has less memory and CPU consumption during the data flooding.

The above findings have given many insights into the DTN and routing. However, the application to the area of epidemic spreading in the public space is still under exploitation. In (Gabriela Moutinho de Souza Dias, 2012), the epidemic routing was applied to the DTN architecture and the SIR model. The process of forwarding messages in the DTN model is like the virus spreading among human groups. An infected individual can transmit the virus to others through droplets or physical contact; after the incubation period, the infectious people can send it to others. In the same way, the source node in Epidemic DTN transmits copies of messages to all nodes it has contact with and continues propagating until the destination nodes. In the network, each node has one of the following states:

Susceptible: the nodes which available to receive the copies of messages.

They were infected: the nodes which have the ability to transmit messages.

Recovered: The unavailable nodes to receive and transmit messages were recovered.

When the infected nodes are in the communication area of other susceptible nodes, the message can be flooding to the susceptible ones, and the time for the node's connection and transmissions is negligible. When the nodes are in a recovery state, they cannot transmit or receive any message from other nodes.

They can be treated as permanent death in the network. Once epidemic DTN is contructed, 98% confidence intervals can be achieved by available and

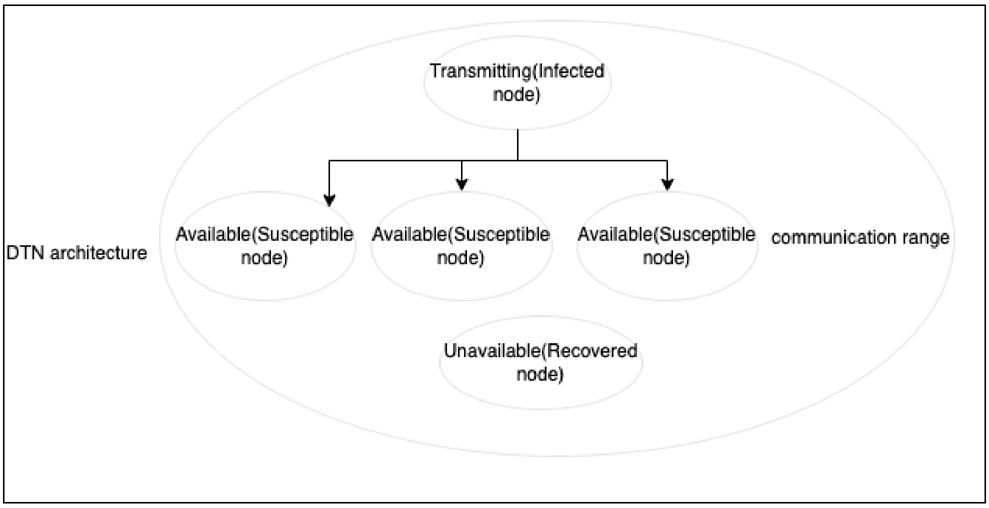

Figure 2.3 Epidemic DTN

transmitting nodes (Gabriela Moutinho de Souza Dias, 2012).

Figure 2.3 Epidemic DTN shows the Epidemic DTN's architectures and the routing rules between the different nodes.

To be concluded, these findings give some excellent ideas to simulate the supermarket layout in network architecture and combine with the epidemic spread mathematical model (SIR) and routing protocol. The following section will review some research on supermarket queuing systems.

2.3 Supermarket Queuing System

Queuing systems is a typical discrete event problem, and some research has been conducted to exploit the highest efficiency of the waiting modes. In (Hong Lian, 2007), queuing system is applied to the computer simulation. The data units are considered as customers, and channels are like queues. The queuing time includes both entity arrival time and service processing time. It requires the system to make suitable allocations for the different resources to the idle servers. The waiting time and queue length are uncertain and random. Poisson distribution is often adapted to the entity arrival problem, which is suitable to describe the number of times an unexpected event occurs per unit of time. This ensures that the arrival time is random and independent of one another (Stallings, 2005). Equation (2.7) defines the Poisson distribution:

$$P(k) = \frac{e^{-\lambda s}(\lambda s)^k}{k!}$$

(2.7)

Where λ is the number of arrivals, and k is the target number of arrivals after period S. The average arrival time is $\frac{1}{\lambda}$.

The service time is random or particular during the time period. Normal distribution is an excellent model to indicate the service time because the vast amount of random variables will eventually converge to this distribution. Equation (2.8) defines normal distribution:

$$P = \frac{1}{\sqrt{2\pi}\sigma} e - \frac{(x-\mu)^2}{2\sigma^2}$$

(2.8)

Where μ is the mean value, and the variance is σ^2.

The two models above have satisfied the single server model, which assumes that there is only one service desk and one queue in the system. While the queuing system in the supermarket always has several lines to control the population flow.

However, the single queue system is very low efficiency in the supermarket scenario since many customers will continue to enter the supermarket, and the size is limited. The double queue system was referred to solve this problem. There are two different ideas for implementing a dual queue system according to the past findings, one is to assign the customer to the least longer queue (Hong Lian, 2007), and the other one is to determine the order of queuing by the number of purchased goods (Xing Wenjie, 2015). Customers will select the shortest queue to finish the checking-out service for the first situation. If all the columns are of the same length, the customer will choose the line randomly. The selection of the tills should be optimised since too large several tills will lead to idle resources and fewer tills will cause a long waiting time. Assume that the average service time is h, the number of services is n, the waiting time for each queue is w, and the length for each service is L(n). Time-consuming in the double-queuing system is expressed in equation (2.9):

$$T = hn + wL(n)$$

(2.9)

Where T should be smaller than T_{n-1} and T_{n+1}.

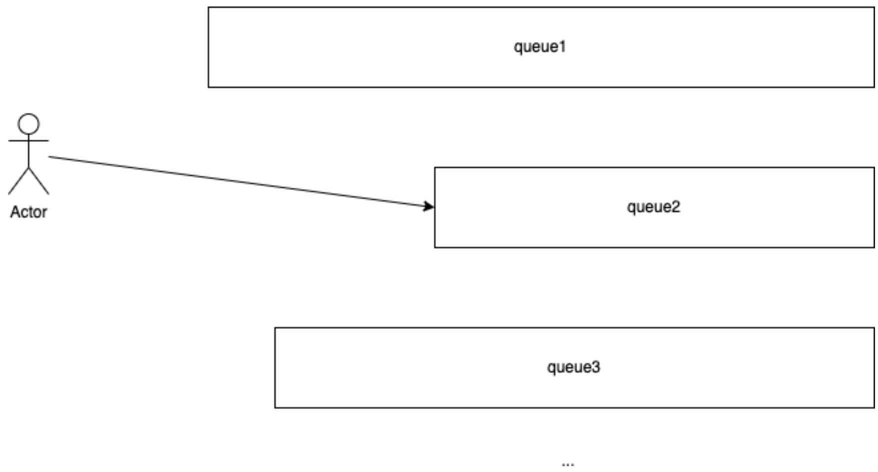

Figure 2.4 Double Queuing system-by queue length

Figure 2.4 Double Queuing system-by queue length shows the queuing system, which allocates users to the minor long queue.

Two kinds of tills are defined for the second solution of the double queuing system. The customers who have more than X goods will go to the Till 1 (T1), while those who have less than X goods will go to the Till 2 (T2). The total time consuming for this system is in equation (2.10):

$$T = n1(C_w\theta_1 + C_s) + n2(C_w\theta_2 + C_s) \tag{2.10}$$

Where n_1, n_2 are the number in the different queues, C_w is the waiting time for one customer in the row, C_s is the service time for one customer, and θ_1, θ_2 are the arrival rate for different lines, which the Poisson distribution can measure.

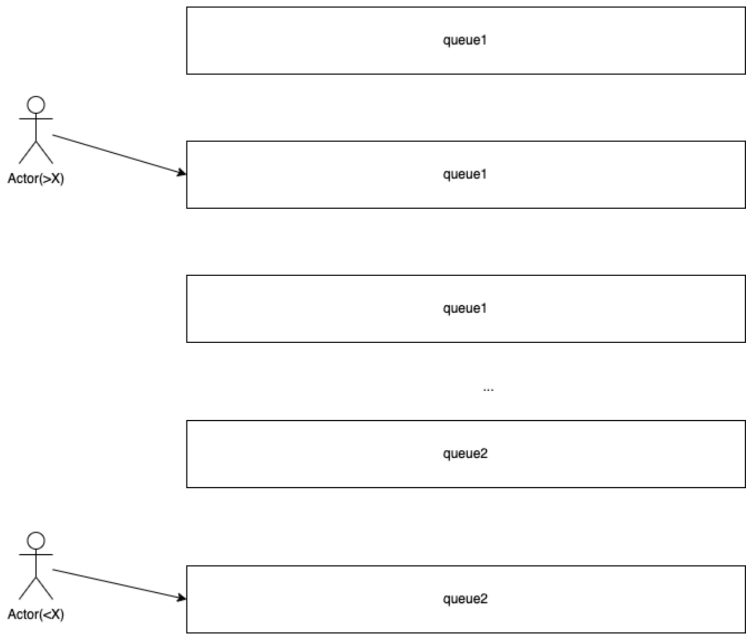

Figure 2.5 Double Queueing system-by number of goods

Figure 2.5 Double Queueing system-by number of goods shows that the double system allocates users by the number of bought goods.

Both the double queuing systems have detailed research into optimising the time efficiency and till usage. While the first one has given standard solutions to most of the scenarios, and the second reduces time from 57.57s in a single queuing system to 54.84s when customers have massive differences in the number of goods purchased. The following section will review some findings about the simulation of supermarket layout, customer behaviour, and mobilities.

2.4 Customer mobilities

Graph theory is a standard tool for simulation where any mathematical object involving points and connections between them can be called a graph. The links can be unidirectional or bidirectional, which serve for different situations (Kenneth H. Rosen, 2003). Building models for animated objects is much easier with graph data structures. Since nodes in the graph can represent the various pieces of stuff, the connections can alter their distances. Some researchers put effort into applying graph theory to architecture design. In (MUKHLISH FUADI, 2015), they tried to create a web-based virtual store layout to enhance efficiency and flexibility. The traditional grid layout design was simulated by the graph elements: nodes, edges, vertex etc.

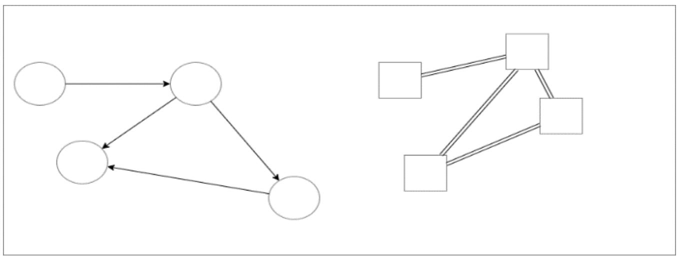

Figure 2.6 Unidirection graph and bidirectional graph

Figure 2.6 Unidirection graph and bidirectional graphshows the unidirectional graph and bidirectional graph.

The critical point for the supermarket simulation is tracking the customers' mobilities. Customer behaviours have been researched a lot throughout the last half-century. According to (Herb Sorensen, 2017), quick in-store trips are ordinary, and people usually tend to buy things according to their initial plans. In the graph model, it equals finding the optimal path between the different nodes. (Brown, 2005) tried to design a grocery and assume that every customer wishes to purchase their items as soon as possible. Two algorithms were given in the paper to realise the customers' behaviour: the Floyds algorithm and the Heuristic algorithm. For the Floyds solution, a 2D adjacency matrix is defined to measure the distances between different nodes, where the not-connected ones are measured as ∞. The shortest path between vertex i and vertex j allows any vertex or combination of vertices from 1 to k to be the intermediate points. This is defined in equation (2.11):

$$D_{i,j,k} = D_{i,k,k-1} + D_{k,j,k-1}, if\ k\ in\ path$$
$$D_{i,j,k} = D_{i,j,k-1}, if\ k\ not\ in\ path$$
$$(2.11)$$

Where k is the nodes except i and j, and the D is the distance between i and j. Therefore, the shortest path between node i and j is in (2.12):

$$D_{i,j,k} = \min\left(D_{i,j,k-1}, D_{i,k,k-1} + D_{k,j,k-1}\right)$$
$$(2.12)$$

This heuristic algorithm is to select the shortest path to an unvisited vertex in the adjacency matrix and to make another array to store the shortest paths.

The depth-First-Searching algorithm is also a standard solution for the shortest path problem. (I. Stojmenovic, 2000) shows how DFS worked in the wireless network. DFS created a path from the source node to the destination without jumping from one node to another node that is not its neighbour. The nodes in the wireless network are in grey; while the jumps were made, the visited node would turn yellow. If no choice can be made when reaching a specific node, the route would trace back to the last node. Once the destination and source node were in the same colour, the searching terminated. This can be applied to customer movement since every step should have four directions to forward, and customers will choose the relevantly optimal one in every step. This process is like what the DFS algorithm does.

The shortest path routing algorithm has proved very helpful in the simulation of individual movement predictions but measuring the movement trends, and routes of groups will be different. (Zipf, 1946) first introduced a model for people's intercity movement, the gravity model. (Piovani Duccio, 2018) researched the bus transportation system in Teresina and tried to find out the accessibility and congestion on different roads with this model and radiation model. For the gravity model, the flow from city i to city j is correlated to the city j's employment rate E_j, the demand P_i in the city i, and travelling time which is measured by the exponential decay of form $e^{-\beta c_{ij}}$, where β is a hyperparameter to control the driving time. The equation (2.13) can calculate the population flow between two cities:

$$T_{ij} = A_i B_j P_i E_j e^{-\beta c_{ij}}$$

(2.13)

Where the A_i and B_j are constant normalizations. Because the inflow and outflow are equivalent to P_i, E_j. The normalizations are in (2.14):

$$A_i = \frac{1}{\sum_k B_k E_k e^{-\beta c_{ik}}}$$

$$B_j = \frac{1}{\sum_k A_k P_k e^{-\beta c_{kj}}}$$

(2.14)

The formulas above can calculate the estimated flow in the bus transportation system in a specific period.

(Fabian Ying, 2022) used the gravity model in the supermarket scenario. Each node in the supermarket has two types of populations similar to the research above: departure population O_k and arrival population D_k. The departure location and arrival population are identical and can be used to measure the flow from and to node k. The gravity model is defined in (2.15):

$$F_{ij} = O_i D_j (d_{ij}/l)^{-\gamma}$$

(2.15)

The last item $(d_{ij}/l)^{-\gamma}$ is an exponential function and used to measure the distance between the node i and node j.

The shortest-circuit and gravity algorithms summarize and analyze the movement of customers and the movement trends of groups respectively. These kinds of agent-based algorithms are very commonly used in epidemic spread scenarios (C M Macal, 2010) and have been explored in depth in conjunction with customer behavior. The next section will summarize all the findings above and suggests areas that have not been exploited and research plans that will conduct in this paper.

2.5 Summary and improvement

The above findings focus on the construction of the flow disease model (SEIR model), the simulation of the supermarket scenario (DTN), the analysis of the queuing system (single & double-queue queuing system) and the mobility of the customers (shortest path & gravity model). Different epidemic models, including SIR, vital dynamics SIR, SEIR, and SEI models were reviewed, and finally, the SEI model is more suitable as a guide model for this topic because there are no recovered people in the supermarket and there will be some incubation periods. In Section 2, the Epidemic DTN model was reviewed, combining a high latency network model with various genetic disease models to simulate the spread of an epidemic in a closed area. The double queuing system has several different implementations (queue length, number of items purchased) and can be chosen specifically for different scenarios. The graph model and the shortest-circuit problem can effectively address the specificity of customer mobilities, and the gravity model can measure the movement of groups. These documents provide theoretical and mathematical support for various aspects of the covid propagation experiments in supermarkets, with particular reference to (Fabian Ying, 2021) and (C. J. NOAKES, 2006). These two papers discussed covid propagation under the movement of customers by zones in supermarkets and the genetic disease model of airborne propagation in closed environments, respectively. However, (Fabian Ying, 2021) is somewhat incomplete in its determination of infectivity, as it used the SEI model and chose total exposure time to assess covid transmission between populations. However, the variation between populations can be considerable, and the whole exposure time may only be present in a few specific customers. What's more, the importance of airborne transmission is not taken into account in it, so in combination with (C. J. NOAKES, 2006), it can obtain the R0 value of transmission over time for the population, which can be used as an indicator to make an accurate assessment of the outbreak's influence in the long term. Also, the mobilities of customers have a relative disadvantage in the supermarket model, as the starting point of customers is the same, which leads to repeated path searches. In the present experiment, the mathematical models described above have to be improved accordingly.

In the next chapter, the various mathematical models and tools mentioned above are integrated and used to simulate the spread of covid in supermarkets, including queuing systems, the spread of airborne epidemics and the evaluation of the space.

3 Experiment Design & Methodologies

The project is to study the propagation of covid in the supermarket. This chapter will design a typical supermarket layout in the graph data structure, including goods, paths and till areas. The customers' shopping behaviour model mainly includes queuing to enter the supermarket, buying one target item and queuing to exit the supermarket. The queuing and customer movement will be controlled using the double queuing system and customer mobility modelling from Chapter 2. The virus transmission process will be implemented concerning the epidemic DTN and SEI models in the studied findings, where the number of infected people is determined and the change in the number of new infections over a certain period is observed. The development of the epidemic in the long term is also modelled in conjunction with the indoor airborne. The following subsections will elaborate on modelling the supermarket layout, the queuing checkout system, the movement path selection and the virus propagation.

3.1 Supermarket layout

The supermarket model is based on Lidl's construction (Lidl, 2022) and contains a shopping area, a waiting area and a checkout area, with a total size of $108 * 72m^2$. It includes entrances, exits, food, fruit, dairy stuff, discounted items, checkout counters and paths. The supermarket model is modelled by graph model, where the different areas are equivalent to nodes, and the

Figure 3.1 Supermarket layout

vertexes between them are some already fixed paths.

Figure 3.1 Supermarket layout shows the supermarket's indoor design. This design is based on the grid layout, which is in a conventional retailing store style and easy to measure the occupied area (Adam P. Vrechopoulos a, 2004). The simulation supermarket graph is 1080*720 pixels in total, which means 10 pixels will represent 1 meter in the simulation. Five different till stations are available, and customers can choose any till station for checkout. Table 3-1 Shopping area location shows the location information for each area:

Table 3-1 Shopping area location

	Start X	End X	Start Y	End Y
Entrance	720	795	0	50
Exit	795	870	0	50

Tills	975	1025	160	450
Fruit	300	660	0	50
Discount	660	710	100	240
Food Area 1	100	660	100	240
Food Area 2	0	200	0	50
Food Area 3	0	50	50	720
Food Area 4	50	920	670	720
Dairy	100	780	290	430
Drink	100	780	480	620

The food is divided into four different areas for the convenience of path selection simulation. The checking-out area contains five tills, each of which is 5m * 5m in size and one meter apart. The supermarkets are passable except for the items mentioned above, each road in the shopping area is 5m wide, and all roads are acceptable in both directions, so by default, there are no blockages in the supermarket.

This subsection describes the construction of the virtual supermarket and the graph structure representation. The following subsection will introduce the access queuing system and checkout system.

3.2 Customer Arrival and Checkout System

Based on the description in (Fabian Ying, 2021), it should be 2.55 customers who enter the supermarket every minute based on the population in UK supermarkets over three months. Queuing system will use the Poisson distribution mentioned in the previous chapter and set the value of λ at 2.55. As discussed in equation (2.7), $P(k)$ is the possibility of k customers entering the supermarket per unit time t. When setting $t = 1$, $P(k)$ will represent the possibility of k customers' arrival in every minute. The number of arrival customers is randomly selected based on the results of P(1), P(2), P(3), P(4), P(5). The range of arrivals is chosen from one to five because these numbers have the greatest chance to represent the number of customer occurrences. To control the spread of the epidemic, a max capacity is also set to control the number of people indoors.

The checkout system refers to the double queuing system in the previous chapter (Hong Lian, 2007) and arranges the checkout locations for customers according to the length of the queue. The waiting and processing times are represented by a Poisson distribution, similar to that of an artificial system. The number of checkouts per unit time is adjusted according to the value of the adjusted λ. In this experiment, the five queues are evaluated at the same time, and the deque of each line is not considered individually.

This subsection describes the customer arrival system and the checkout system, as well as the essential theoretical part of the implementation (Poisson distribution). In the next section, customer mobilities and path selection will be covered.

3.3 Customer Mobilities and path Selection

The classic customer behavior movement is to choose the shortest path to find the target items with algorithms like DFS and Floyds, but there are more drawbacks in this experiment. (Brown, 2005) assumed that customers would find the shortest path in the grocery and realised the simulation with Floyds. However, this algorithm presupposes that the start and end points are relatively random. Otherwise, the shortest circuit would be fixed. In the current experiment, all customers start with the location near the entrance, which means that if the shortest-path greedy algorithm is used, customers with the same target will follow the same path. However, this is not reasonable, as in most cases, customers will choose random paths to buy items based on the population in front of them. Moreover, in traditional search algorithms, such as DFS, the default endpoint and start point have no length or width so that the target can be searched accurately. But in this scenario, all the items are in a rectangular area, and the customer can choose an arbitrary side to buy the item, which has led to more difficulties.

The approach used in this project is to design multiple deterministic paths for different shopping areas, with the routes covering all the locations that the customer might reach, ensuring that all the circumstances of the user's shopping are taken into account. Customers should walk straight through the aisles, passing through corners or bends at random locations, known as intermediate points, which increase the randomness of the user's movement. The next few sub-sections will show in detail the route options for each area in the supermarket.

3.3.1 Path selection for Food Area 1

There are three different routes for food area 1, which are designed for the locations where customers are likely to arrive.

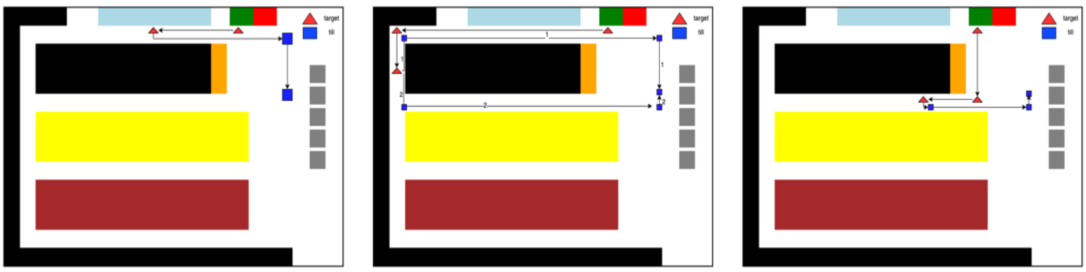

Figure 3.2 Path design for Food Area 1

In the second sub-diagram of *Figure 3.2* Path design for Food Area 1, because the two paths to the checking-out area are equidistant, the path to tills has been designed as two routes, which are distinguished by the label numbers.

3.3.2 Path Selection for Fruit Area

Figure 3.3 Path design for Fruit Area shows the route directed to the fruit area, which is relatively simple with only one option.

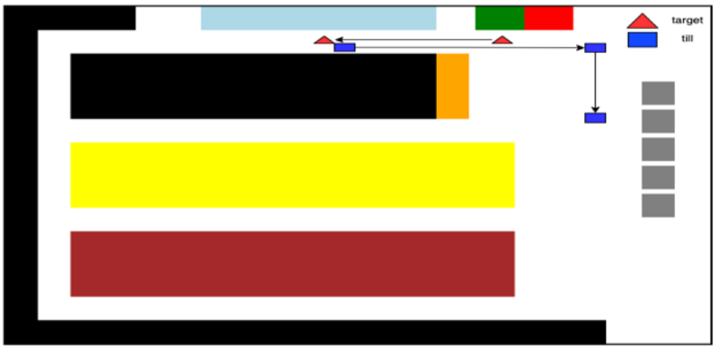

Figure 3.3 Path design for Fruit Area

3.3.3 Path selection for Food Area 2

There is only one route to food area 2, but two different paths have been designed to the checking area, which is marked on *Figure 3.4* Path design for Food Area 2.

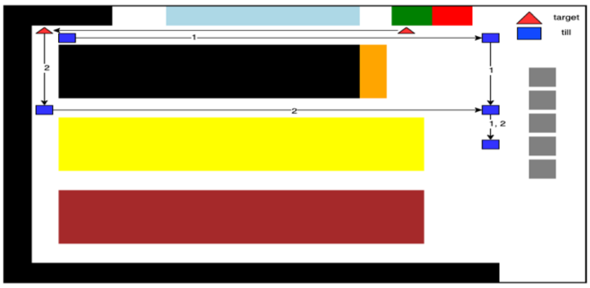

Figure 3.4 Path design for Food Area 2

3.3.4 Path selection for Food Area 3

Four different routes to reach the food area 3 and four different ways to reach the checking out area have been given in the *Figure 3.5* Path design for Food Area 3. The allocation of roads in this experiment is random to maintain the diversity of customer behaviour.

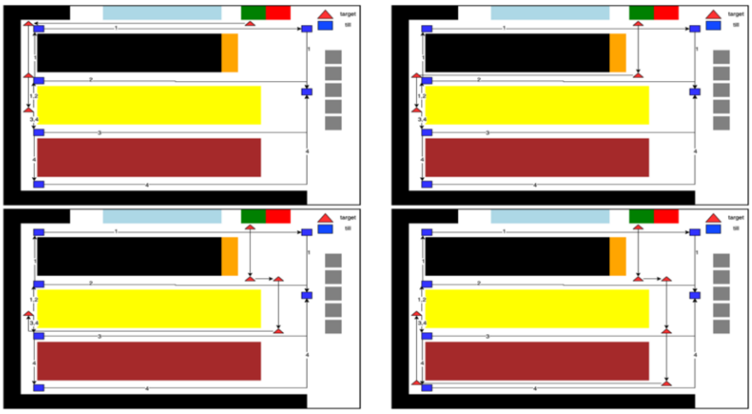

Figure 3.5 Path design for Food Area 3

3.3.5 Path selection for Dairy Area

A total of six routes have been designed to reach the dairy shopping area, covering the four sides of the area.

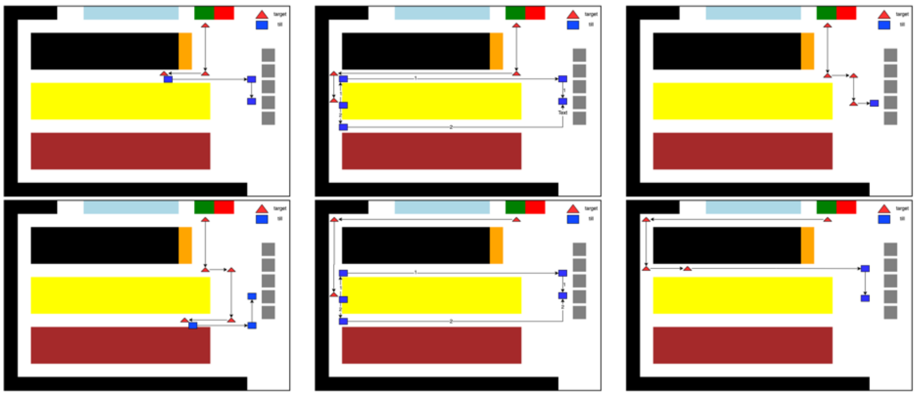

Figure 3.6 Path design for dairy area

In the second and fifth sub-diagram of *Figure 3.6*, two different marked routes have been designed for the checking-out area. At the same time, the other sub-diagrams have only one way covered.

3.3.6 Path selection for Food Area 4

Figure 3.7 shows three routes to the drink Food Area 4, with the paths forwarding to the checking-out tills.

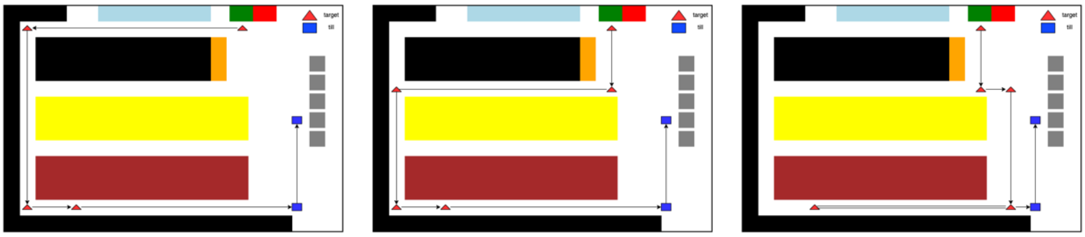

Figure 3.7 Path design for Food Area 4

The route to the checking-out area in the sub-diagrams is the same, as it is significantly shorter than the other options.

3.3.7 Path selection for Drink Area

A total of 8 different routes have been designed to reach the drink shopping area, covering four sides of the rectangular area.

The third and fifth subplots in *Figure 3.8* have two different routes to the target till, while the other subplots have only one option.

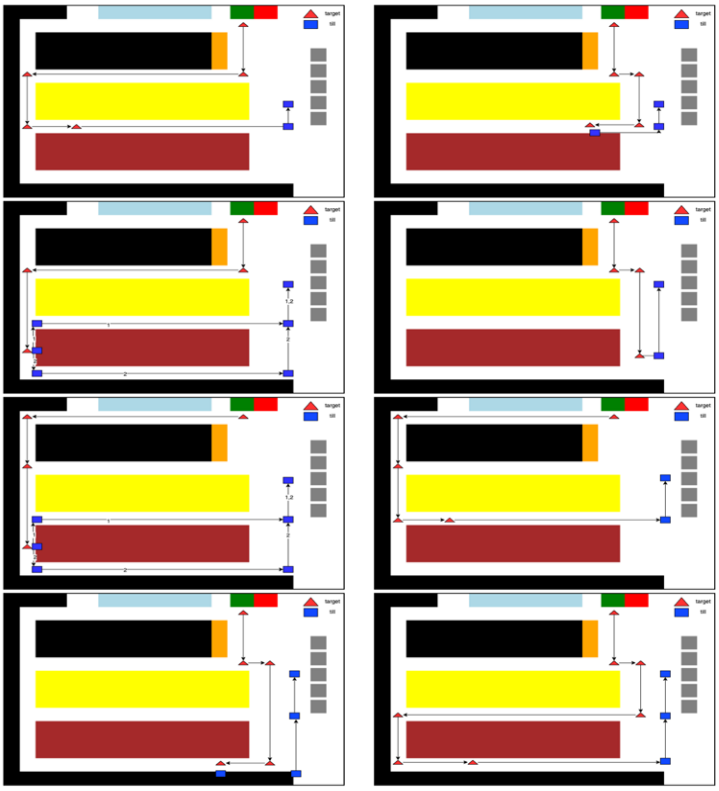

Figure 3.8 Path design for Drink Area

3.4 Virus Propagation model

The viral propagation model is mainly based on the epidemic DTN concept mentioned in (Gabriela Moutinho de Souza Dias, 2012), where customers in a supermarket are simulated by moving nodes with different characteristics and visualised in the constructed supermarket layout. The occupied area of the customers is ignored, and the case of congestion in the path is not considered. The SEI model (Fabian Ying, 2021) will also be used as an epidemic model for the simulation of virus transmission. The following subsections will describe the customer mobilities in the supermarket and the virus spread model in detail.

3.4.1 Customer Simulation

According to the design of the delay tolerant network, this experiment assumes that the supermarket is a network architecture and customers are mobile node items. Customers are initially positioned anywhere near the entry, and their movement in the supermarket follows the same pattern as in the previous section After reaching the discount area or the target area, the customer stays for a certain amount of time. The DTN was chosen as the theoretical architecture because of the randomness of the customer's movement in the supermarket and the pause at some points in time. DTN is very compatible with high latency and unstable network architectures and can ensure the reliability of the infected people virus propagation simulation. Depending on the infection of customers, customers are represented in this experiment with two different nodes for simulation. One is an infected node, which can infect other nodes within a specific range, and the other is a susceptible node, which transforms into an infected node after a certain exposure period. The ratio of the two nodes is defined before the experiment starts and changes state as the experiment progresses.

3.4.2 Virus Spread Model

In this experiment, instead of applying the traditional SIR and SEIR models to model the epidemic DTN, the SEI model was used because the average shopping time of customers was around 37 minutes (The sun, 2022), but the recovery time from covid usually took 1-2 weeks (Masson, 2022). For the SEI model, the customer population is divided into three different groups, namely susceptible, exposed and infected, and the exposed group may become infected after a certain period of exposure time, which takes into account the latency of the transmission process. This takes into account the latency of the transmission process.

The characteristics of the SEI model combined with the rules of customer mobilities were used in this experiment to construct the virus spread model. A transmission range was defined in the project. Any susceptible person entering the transmission range of infected people becomes an exposed person and increases the exposure time. When the exposure time reaches an upper limit, they become infected and have the ability to spreading the virus as well. This is similar to the broadcast mechanism in graph theory.

3.5 Summary

This chapter reviewed in detail the simulation design and methodology of this project, including the modelling of the indoor layout, the simulation of customer behaviour and mobilities, and the selection and implementation of an epidemic spreading model for the supermarket scenario. The following section discusses the implementation of the epidemic spreading for the designed supermarket layout and the modelling of the long-term influence of the epidemic in the context of indoor airborne.

4 Implementation

The previous chapter detailed the design of the experiment, including layout, customer mobilities and epidemic spreading. This chapter presents the corresponding implementations of all the above models and simulations. In particular, customer arrival, virus transmission and movement, and customer dequeuing will be implemented in python 3.9 (2020-10-05), where the values are calculated using NumPy 1.21.2 (2021-08-15). A long connection to the front-end display page was established using Flask framework 2.0.2 (2021-10-04) and Flask-socketio 4.3.1 (2020-07-02), which transmitted all customer location information to the website every 0.2s. The display page used React framework and react-konva 18.1.1 (2022-05-02) for real-time rendering of the supermarket layout and customer locations.

In the following few subsections, the application of the data visualization is described accordingly, followed by a detailed description of the ideas and parameters required for the implementation of the various stages of epidemic spreading. Finally, a summary of the implementation will be given, and in the next chapter, hyperparameter tuning and result analysis will be performed.

4.1 Data visualization

The supermarket layout is drawn using the react-konva package, using different colours to indicate the different areas. After establishing a socket connection with the backend Flask, the location and health status of the customer are obtained every 0.2s and rendered in real-time in the supermarket layout. *Figure 4.1* shows the different customers in the supermarket.The red colour represents the infected state and the green colour the susceptible state.

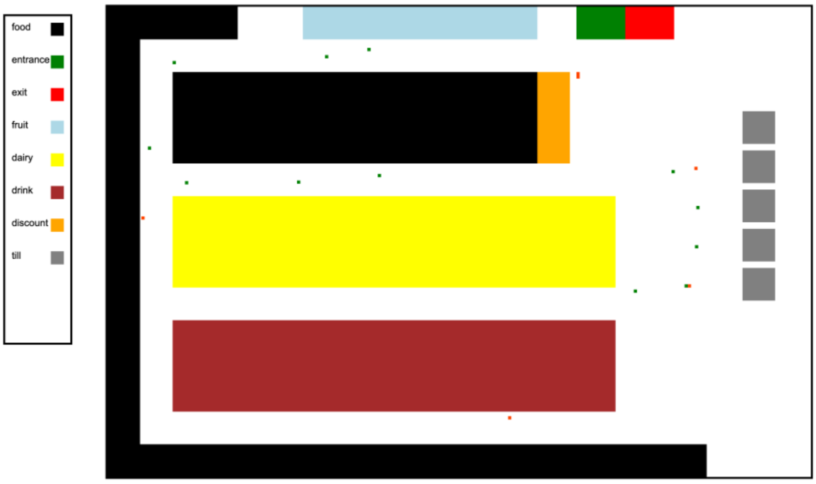

Figure 4.1 Customer mobility in Supermarket

4.2 Epidemic Spreading Process

The data processing and analysis of the epidemic spreading process is divided into the following stages: creating the customer model with location and status, building the incoming and outgoing queue simulator, constructing the customer movement simulator and analysing the infection situation. The implementation of these stages is described in detail in the next few sections.

4.2.1 Customer model

The customer is constructed by class in Python which contains the following parameters x, y, exposed_time, waiting_time, discount_waiting_time, till_x, till_y, destination_x, destination_y, infected, speed, target and path where x, and y indicates the current location of the customer, exposed_time shows the time the customer is exposed to infected people, till_x, till_y suggests the place of the checking-out till, which will determine by the queue length, destination_x and destination_y are the random sites of the shopping area reached, target is the item which customer want to purchase and infected is the situation of infection.

According to (Zyl, 2015), humans tend to walk at about 1.4 metres per second. Given that shoppers need to push a shopping cart or carry a shopping bag, the speed of the customer is set to 0.8m/s. Waiting_time and discount_waiting_time are used as hyperparameters to control virus spreading during the data analysis. The path represents the route choice, described in detail in the next section on customer mobilities.

By constructing the customer class, the customer's location, movement and infection can be fully controlled. In the next section, a systematic description of the queuing system in and out of the supermarket will be given.

4.2.2 Queuing system

As described in the previous chapter, the customer arrival system will use a Poisson distribution and set λ to 2.55. The project takes the five most significant values from the Poisson distribution as the probability of the number of people arriving in one minute, which in the arrival model are P(1), P(2), P(3), P (4) and P(5). The system will generate a random number of customers based on the probability values and place them in the simulator.

The checkout system is similar to the customer arrival system, where λ here is used as a hyper-parameter to control the flow of people. The maximum five probability values are calculated according to the Poisson distribution and the number of people checking out in a minute is chosen randomly according to these values. At the same time, the customer will move forward to the shortest one among the waiting queues in order to reduce the chance to get exposed.

4.2.3 Customer Movement and Epidemic Spread

In this project, the upper limitation of 50 is set for the number of people in the supermarket. When the number of people has not reached the upper limit, customers can enter normally and continue moving inside. As mentioned in the previous chapter, each shopping area has various paths to reach, and the system will randomly assign routes to customers based on the probability values in the Table 4-1 and their target items.

Table 4-1 path allocation probability for different areas

	Path 1	Path 2	Path 3	Path 4	Path 5	Path 6	Path 7	Path 8
Food Area 1	0.5	0.25	0.25					
Food Area 2	1							
Fruit	1							
Food Area 3	0.4	0.2	0.2	0.2				
Dairy	0.15	0.15	0.15	0.15	0.25	0.15		
Drink	0.1	0.1	0.1	0.1	0.1	0.1	0.25	0.15
Food Area 4	0.5	0.25	0.25					

The customer then moves along a defined route until it reaches the destination and updates its position every 0.2 seconds, which is passed to the data visualisation module by the socket communication. The time and distance in the project differ from the actual scenario, where 10 seconds in the project is equivalent to 1 minute, and 10 pixels are equal to 1 meter. Therefore, the customer will move 9.6 pixels every 0.2 seconds in the simulation.

According to the epidemic spreading model defined in the previous chapter, each infected individual has an infectious range. (Authority, 2020) quoted that each country has different criteria for social distancing, the shortest distance in Australia is 1.5 meters, and the longest stretch in Singapore is 3 meters. In order to fully take into account the possibility of transmission, the exposure range is defined as 3 meters. Whenever a customer updates their location, the system counts whether infected people are within 3 meters of each susceptible person with the *NumPy linalg* package. Each time a customer updates their site, the system will matter whether there are infected people within 3 meters of each susceptible person and update their exposure time property and infection status. The 15-minute rule (Cleveland Clinic, 2020) was introduced to indicate that the average person becomes infected after 15 minutes of exposure to a newly crowned patient. But this data is inaccurate, according to (Crist, 2020), out of 21 patients sampled, eight patients were infected within 5 minutes, or 38%. Also cited (Gupta, 2020), A high school student in Jeonju contracted the virus after 5 minutes of brief contact with a saleswoman. In this project, as with the selected value for social distance, the limiting case for infection with the virus was considered, and the threshold for exposure time was set at 5 minutes. A susceptible person will be considered infected after this time of exposure to infected people.

4.2.4 Infection Analysis

The above subsections introduced the implementation of the covid spread in supermarkets, including the choice of parameters and methods. The analysis of the infection situation will be considered from two perspectives, including the duration of exposure and the number of new infections. In (Fabian Ying, 2021), quantitative analysis is carried out mainly for the average and total exposure time. However, this is not rigorous enough because there is variance in virus transmission in different areas, and certain areas can affect the overall statistical results. Moreover, in the SEI model, the length of exposure does not directly affect the increase in the number of infections due to the exposure latency. In this project, the change in the number of people infected was also considered to help determine the transmission rate of the virus. To evaluate the covid spread situation in different areas, the supermarket was also divided into ten zones to make the comparison. In the next section, the details of the analysis are presented.

4.2.4.1 Overall Analysis

Once the customer arrival system has been set up, every 200 seconds (20 minutes in reality), the system will record the total number of customers, infected customers, new infected customers and full exposure time. (three hours in reality). The outbreak trend can be exploited by comparing the number of new infected customers with the exposure time every 200 seconds. The total number of new infected customers and exposure time will also allow the comparison of the supermarket infection results for various situations, i.e., different choices of average checking out time.

4.2.4.2 Zone Analysis

Based on the distribution of the shopping area, the supermarkets were divided into ten different zones, one of which corresponds to the checking-out area and the others to 2-3 shopping areas.

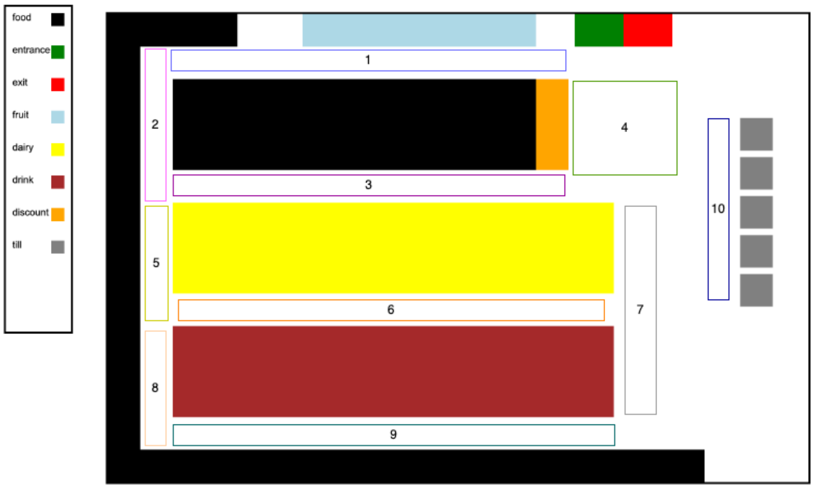

Figure 4.2 Supermarket zones

Table 4-2 Zone locations

	X	Y	Shopping area
Zone 1	100-660	50-100	Food 1, Food2, Fruit
Zone 2	50-100	50-290	Food 3, Food 1
Zone 3	100-660	240-290	Food 1, Dairy
Zone 4	710-870	100-240	
Zone 5	50-100	290-480	Food 3, Dairy
Zone 6	100-660	430-480	Dairy, Drink
Zone 7	780-880	290-620	Dairy, Drink
Zone 8	50-100	480-670	Food 3, Drink
Zone 9	100-660	620-670	Food 4, Drink
Zone 10	890-975	185-425	Checking out area

During the simulator running process, the system records the location where each exposure occurs, the location of the infection, and the target location of each newly infected person. In the subsequent data analysis, these data are aggregated and analysed, counting the exposure time and new infected customers for each zone.

4.3 Summary

This chapter describes the implementation of the customer model and the covid spread simulator, including the choice of parameters and the simulation of scenarios. The methods of data analysis and the categories of statistical data are also described. The next chapter will analyse the factors influencing covid spread from different perspectives through adjusting hyperparameters and reducing the infection rate.

5 Evaluation

The implementation of the experiment was described in the previous chapter, where several hyperparameters were mentioned, including the time of shopping, λ in the Poisson distribution for checking out process and the initial number of infected people. In the following, the effect of different factors on the infection rate and exposure of susceptible individuals will be evaluated by comparing the base model with other models with different hyperparameters. This will be followed by an assessment of the long-term epidemic influence using the SEIR model, taking into account the indoor airborne as well. Finally, a summary of the various factors to covid spread will be presented.

5.1 Base model

Since the customers move once they enter the supermarket, and the route selection is very diverse, it is not easy to get infected during the mobility. Virus exposure occurs mainly during the shopping period and in the check-out area. The impact of these factors on the epidemic spread will be discussed by adjusting for the time spent shopping for target items, time spent in the discount area, the number of people checking out in one minute, and the initial proportion of infected people. In the base model, the purchasing time for each customer is a random value between 20 - 30 seconds (2-3 minutes in real life), and the time spent in the discount area is a random number between 5 – 10 seconds (30 seconds - 1 minute in real life). The number of people checking out in one minute is 10, and assumed that 30% of customers get infected initially.

In the base model, a total of 446 customers entered the supermarket, and 113 were infected. After the experiment, there were 18 new patients and the total exposure time was 4226 seconds. The *Figure 5.1* shows the change in average exposure time and the number of newly infected people within 1800 seconds in the simulation. Because 10 seconds in the simulation is equivalent to 1 minute in reality, values on the x-axis represent the time (minute) in real life.

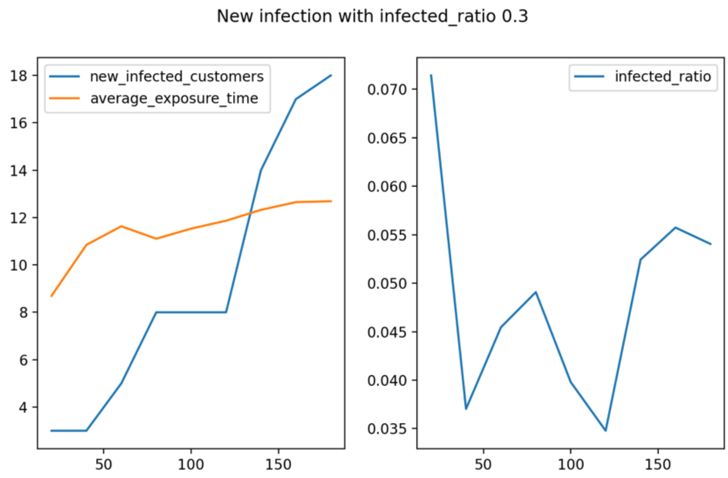

Figure 5.1 base model new case and exposure time tendency

The average exposure time reaches 12 seconds at 60 minutes and remains stable. There was no new case from 75 to 125 minutes, which increased rapidly during other periods. The infection rate was highest in the initial phase at approximately 7%, and the average value during the experiment was 5.4%. *Figure 5.3* shows the distribution of exposure times in the different zones and compares virus exposure in the checking-out and shopping areas. In the checking-out area, the virus transmission was the most severe, with a total exposure time of 1864 seconds for susceptible customers, accounting for 51.7%. In the shopping area, customers were exposed to the infected people for 725 seconds, followed by zones 2 and 4 with 233 and 23 seconds respectively.

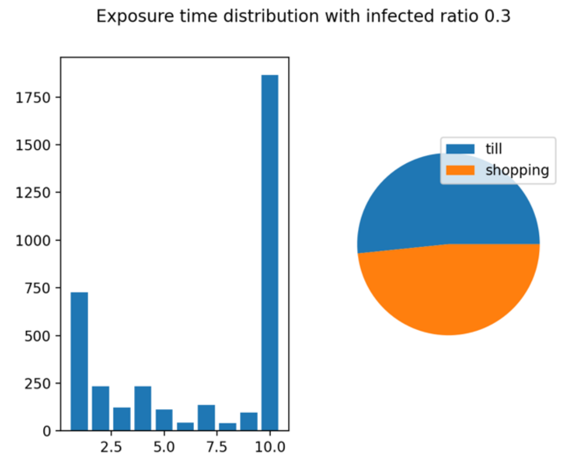

Figure 5.3 Base model exposure time distribution

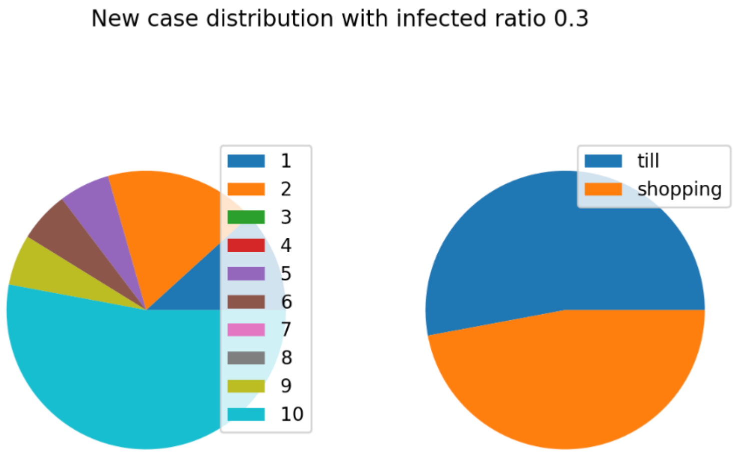

Figure 5.2 Base model new case distribution

Figure 5.2 shows the distribution of new customers in different zones, with the highest infection rate in the checking-out area, accounting for 53% of the total customers. The following areas with high infection ratios were zone 2 and zone 1, accounting for 17% and 12% respectively.

5.2 Tuning Proportion of infected people

The number of initial infections is an essential factor in the spread of the epidemic, and the percentage of infected people will be adjusted to 20% and 40% to see how the results changes.

When the initial infection rate was 20%, there were 489 customers, 113 of whom were infected. The average exposure time was 9.3 seconds, 0.1 seconds less than the base model, and the average infection rate was 5.8%, 0.4% higher than the base model, which is not a significant difference.

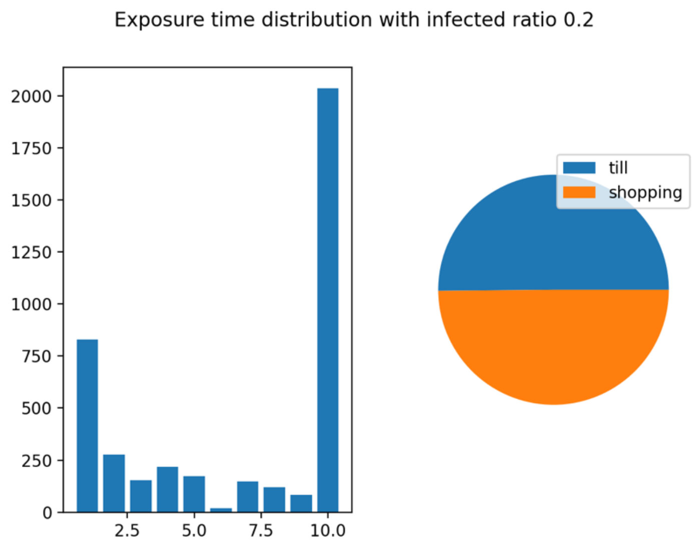

Figure 5.4 Expoure time distribution when infection ratio=0.2

Figure 5.4 shows the exposure of different zones. Like the base model, the exposure times of zone 1 and checking out is longer. However, this time the percentage of exposure in the shopping area increased to 49.8%.

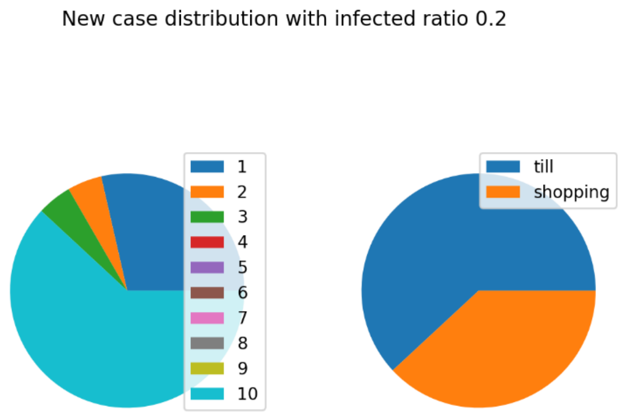

Figure 5.5 New case distribution when infection ratio=0.2

Figure 5.5 shows the infection situation in different zones, similar to the base model. Checking out areas, zone 1 and 2 have high infection rates.

After adjusting the initial infection rate to 40%, there were 469 observations, of which 188 were infected. There were 29 new cases during the experiment, and the total exposure time was 4592.4 seconds. The infection rate reached 10.3%, nearly double that of the base model, and the average exposure time of customers was 9.79 seconds, 0.4 seconds longer than that of the base model. After adjusting the initial infection rate to 40%, there were 469 observations, of which 188 were infected. There were 29 new cases during the experiment, and the total exposure time was 4592.4 seconds. Figure 5.6 describes the changes in exposure time and infected customers during the experiment. The average exposure time stabilised at 15 seconds after 50 minutes, much higher than the 12 seconds in the base model.The infection ratio peaked at 14% and remained above 10%.

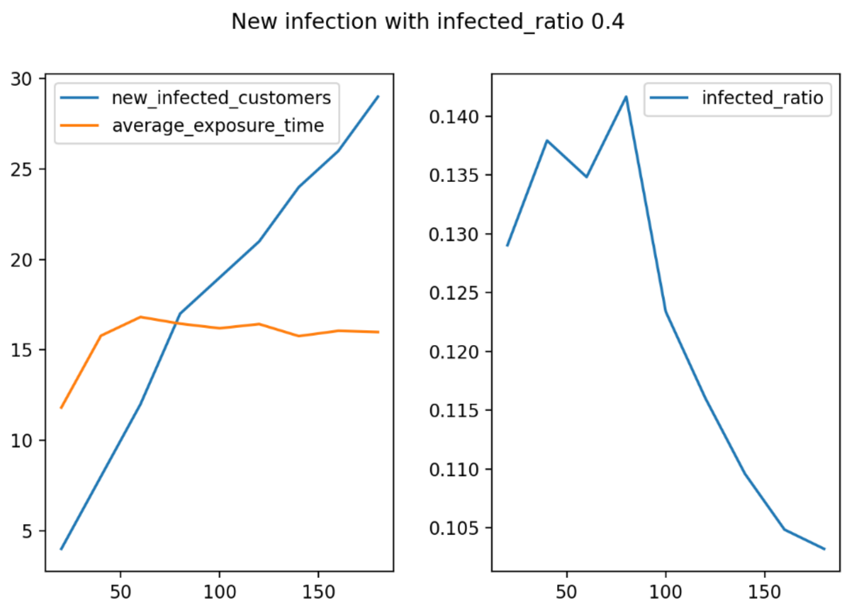

Figure 5.6 new infection and exposure time change with infection ratio 0.4

Figure 5.7 shows the distribution of infected people in zone 9, with zone 1 having more infected people, unlike the base model. The checking-out area still accounts for more than half of the infected people.

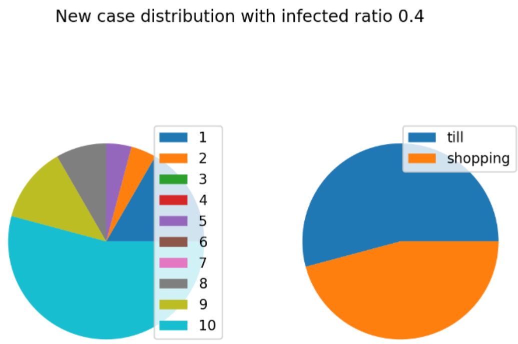

Figure 5.7 New case distribution with infection ratio 0.4

To be concluded, when the infection ratio is 0.2, the infection situation is not much different from the original model. However, when the initial number of infected people is 40%, the infection rate and exposure time increase significantly, with the infection rate rising from 5% to 10%. This indicates that the number of infected customers does affect the spread of the epidemic and that there is a spike above the threshold (40%).

5.3 Tuning purchasing time

It is tough to achieve an exposure time of more than 50 seconds during movement unless the customers take the same route. In the base model, exposure occurs most often during shopping, with a default shopping duration of 20-30 seconds (2-3 minutes in reality). In the following experiments, the interval of shopping time will be adjusted to 15-25 seconds and 10-20 seconds to observe the exposure and infection in the shopping area.

When the shopping time range was reduced to 15-25 seconds, 446 customers entered the supermarket, and 143 were infected with the virus. There were 12 new cases after the experiment and the total exposure time was 3706 seconds, with six fewer infected people and 520 seconds less exposure time than the base model. The infection rate was 3.9%, much lower than the previous 5.4%. *Figure 5.9* exposure time distribution with shopping time 15-25shows the exposure in the different zones. It can be seen that the exposure time of susceptible people in each shopping area has decreased, with zone 1 and zone 2 having a decrease of 145 and 80 seconds compared to the base model.

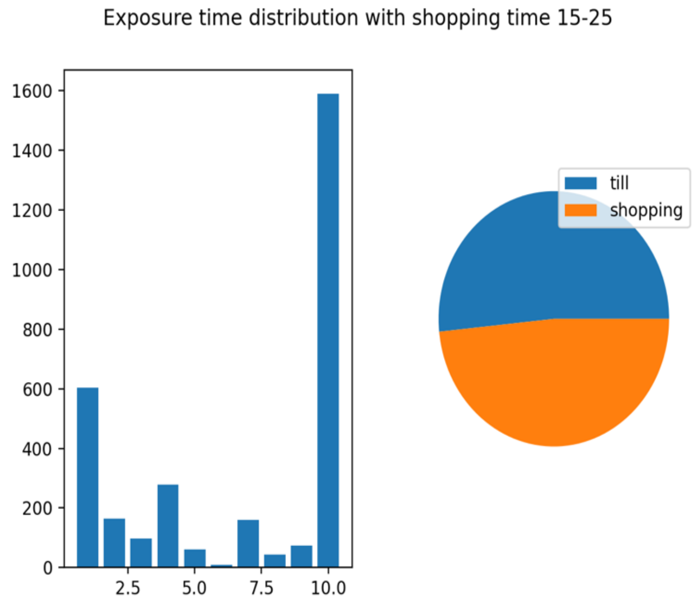

Figure 5.9 exposure time distribution with shopping

time 15-25

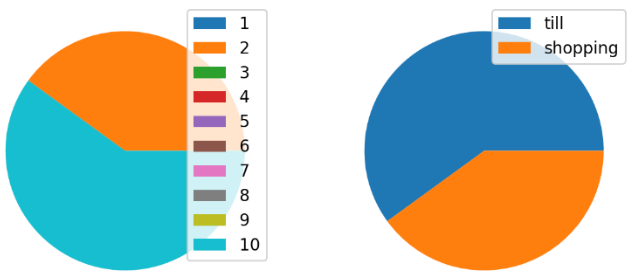

Figure 5.8 New case distribution with shopping time 15-25

Figure 5.8 New case distribution with shopping time 15-25 shows the infection situation. Infection only occurs in zone 1 and zone 2, as most routes need to pass through these two zones, but the decrease in shopping time effectively reduces the spread of the epidemic in the other zones. The number of infections in the shopping area is only 40% of the total, down 7% from 47% in the base model.

When the shopping time became 10-20 seconds, there were 459 customers, 130 of whom were infected. Ten new cases were added at the end of the exper, and the total exposure time was 4215.4 seconds. Eight fewer people were infected than in the base model, and the infection rate was 3%, another 0.9% decrease from the model with 15-25 seconds purchasing time.

Figure 5.10 shows the exposure in each zone. It can be seen that although the total exposure time has risen, it is concentrated in the checking out area, the exposure time in the shopping area zone has dropped again compared to the previous model, and the customer exposure time in The exposure time in the

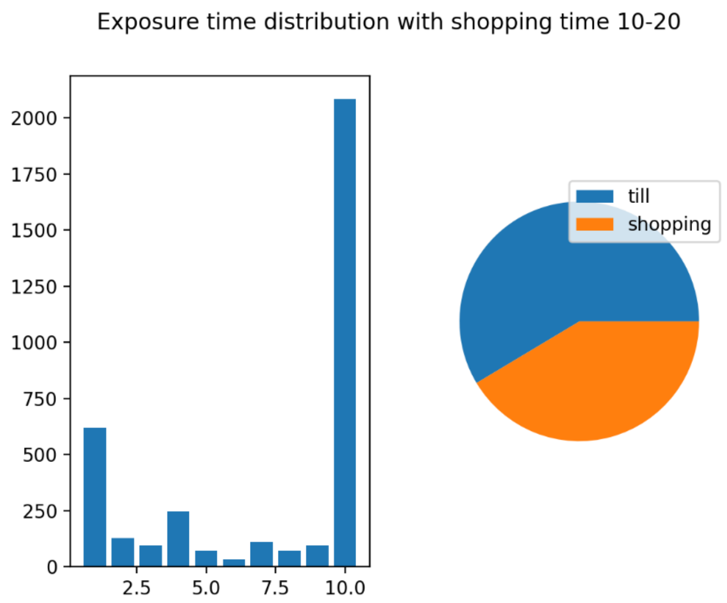

Figure 5.10 exposure distribution with shopping time 10-20

shopping area only accounts for 41.4% of the total exposure time, which is a significant decrease compared to the 48.3% in the base model.

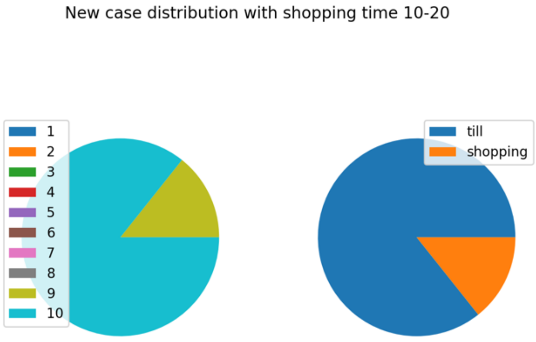

Figure 5.11 New case distribution with shopping time 10-20

Figure 5.11 shows the infection situation in different zones, and 85.7% of new cases are in the checking-out area.

To summarise, controlling the shopping time can effectively reduce the spread of the epidemic in the shopping area. After reducing the shopping time, zone 1, zone 2 and zone 4, which have the longest exposure time, all significantly decrease. When the shopping time was reduced to 10-20 seconds, the exposure time in the shopping area decreased from nearly 50% to 41.4%. At the same time, new cases were also reduced, from 18 in the base model to 12 and 10, with the new patients mainly in the checking-out area. Zone 1 and zone 2 had some new cases because of the many routes passing through.

5.4 Tuning Purchasing time in the discount area

The base model specifies that customers will stay in the discount area for 5-10 seconds. This interval was adjusted to 2.5-7.5 seconds in the project to observe changes in the spread of the epidemic in the associated area, Zone 4, and evaluate the overall infection situation.

After the purchasing time dropped to 2.5-7.5 seconds, a total of 494 customers entered the supermarket, of which 138 were infected. The rate of infection and the average exposure time is shown in *Figure 5.12*. It indicates that the model's overall performance didn't improve a lot.

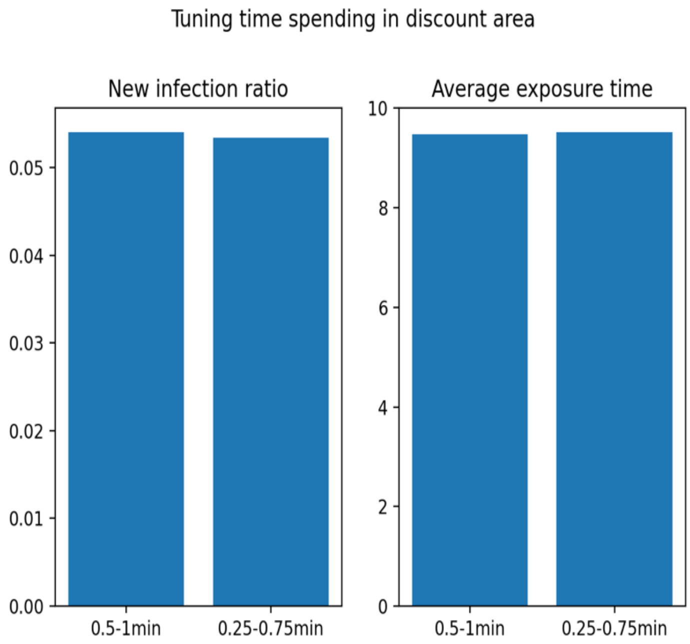

Figure 5.12 new case and exposure time when tuning time spent in discount area

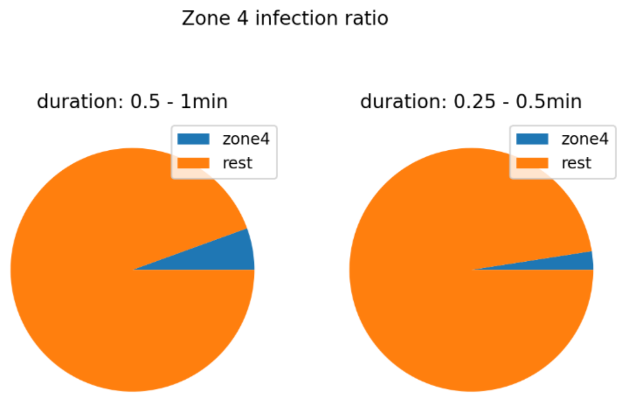

Figure 5.13 zone 4 infection ratio

In the base model, the exposure time in zone 4 accounted for 13.4% of the entire shopping area, and after reducing the stay time, the exposure time dropped to 6.1%, as shown in the Figure 5.13.

To sum up, changing the length of stay in the discount area can effectively reduce the risk of exposure to infected people in Zone 4, but it has little impact on the total exposure time and infection rate compared to controlling the shopping time directly.

5.5 Tuning speed of checking out

The average number of checkouts per minute in the base model is 10, i.e. the λ in the Poisson distribution is 10. As can be found in all the models above, the exposure time in the checking-out area is the longest, accounting for more than 50% of the total. The following step will be to adjust the number of checkouts to 15 and 20 people per minute and to study the changes in the development of the epidemic in the checking-out area.

After increasing the number of checkouts per minute to 15, 455 samples were observed, of which 140 were infected. Eighteen new cases were added afterwards, with a total exposure time of 4176 seconds. The infection rate was 5.7%, and the average exposure time was 9.17s, which was not significantly different from the base model.

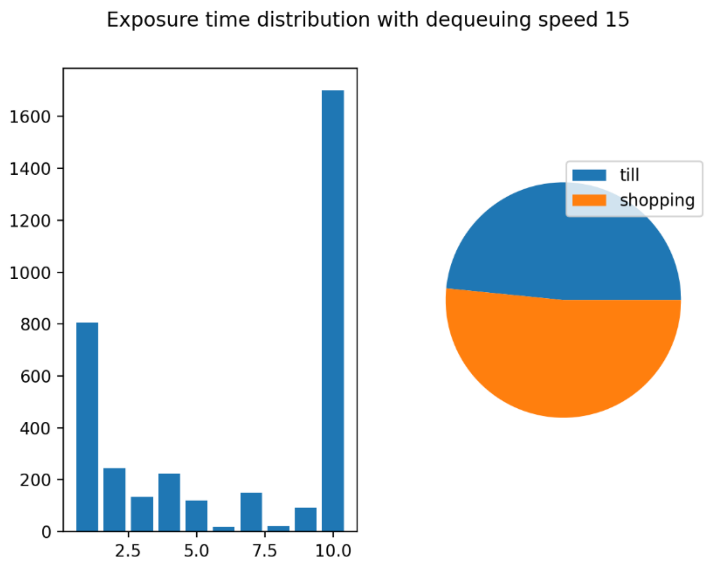

Figure 5.14 exposure time with dequeuing speed 15

Figure 5.14 shows the exposure time distribution in different zones. In the base model, the exposure time in till areas was 1760 seconds, but this time it dropped to 1700 seconds, and the exposure time in checking out dropped from 52% to 48%, which is lower than the one in the shopping area.

When the number of dequeuing people became 20 per minute, there were 499 customers, 144 of whom were infected. After the experiment, 23 new patients were added, and the total exposure time was 4696 seconds. The infection rate reached 6.4%, 1% higher than the base model, and the average exposure time was 9.4s, which was consistent with the base model.

Figure 5.16 plots the distribution of exposure time in different zones. Although the overall performance is slightly lower than that of the base model, the percentage of exposure time in the checking-out area is again lower, at 46%.

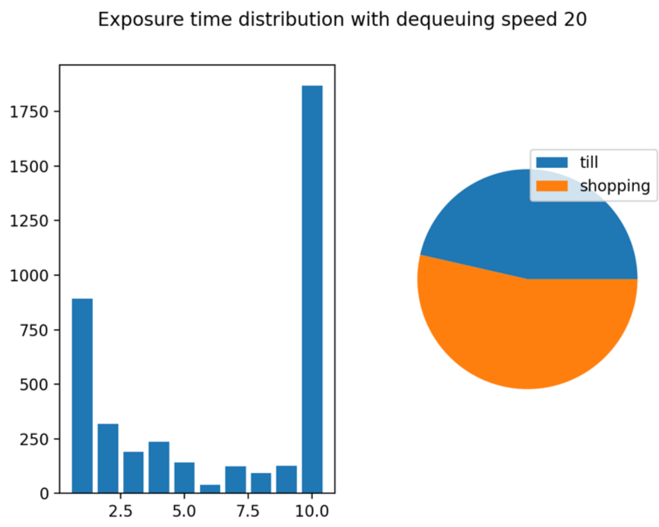

Figure 5.16 Exposure time with dequeuing speed 20

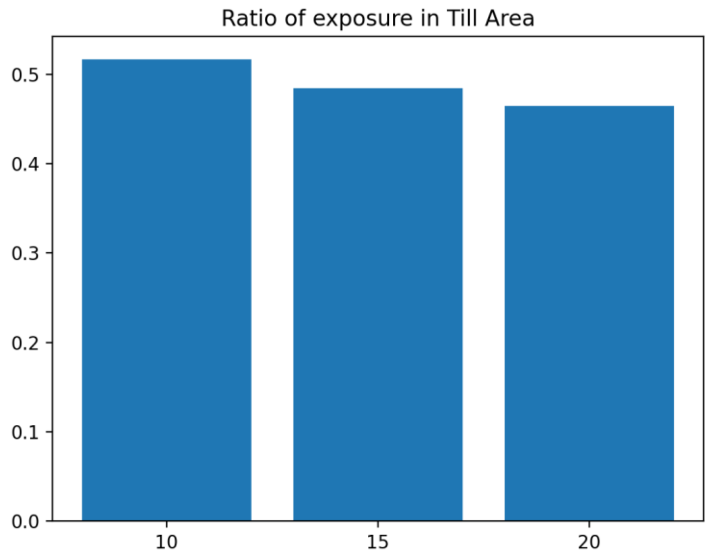

Figure 5.15 ratio of exposure in checking out area

Figure 5.15 shows the percentage of exposure time in the checking out area at different checking out speeds, which gives a decreasing trend.

In general, adjusting the number of checkouts per minute effectively slows down the spread of the epidemic in the till area, with the percentage dropping gradually from 52% to 46%. However, it was not effective in improving the overall performance. Although the average exposure time remained almost unchanged, the infection rate increased from 5.4% to 6.4%.

5.6 Long-term Impact of the Epidemic

The above subsections adjust the hyperparameters to analyse several essential factors that influence the spread of the epidemic in supermarkets. The SEIR model with indoor airborne mentioned in (C. J. NOAKES, 2006) will be used in this section to predict changes in the epidemic over three months. The choice of parameters will be presented first, and then the model will be applied for evaluation.

5.6.1 Parameter selection

According to the mathematical model in chapter 2, the following parameters have to be determined: number of people (N), number of infected people (I), number of exposed people (E), number of recovered people (R), ventilation rate (A), infected people's quanta production rate (q), supermarket size (V), average pulmonary ventilation rate (p), incubation time (α) and recovery time (γ) and experiment length (T).

Assume that the number of people is 200, and there are 30% of them get infected. Cited (Colin J.Axon, 2020), supermarkets are usually ventilated at a rate of up to 6 ACH, while (C. J. NOAKES, 2006) says that the ventilation range is between 3 and 8 ACH. Combining the data from these two sources, a value of 4.5 for the ventilation rate was determined for this topic. The supermarket size is 108 * 72 * 3.5 m^2. According to (Carroll, 2007), the pulmonary ventilation rate is 6000 ml/min, which is equivalent to 0.36 m^2/h. The incubation time is typically 2-5 days (HSE, 2022), and the mean value of this range was applied to this project. In (Feldman, 2022), it said that the typical isolation period is 5 days for covid, which is similar to the recovery time. The quanta production rate varies widely and according to (Jin Li, 2021) the value within the classroom is 194.4, but the median value in the gym is 454.87.While (Miller, 2021) indicates a quanta value of 970. In this experiment, I used 200 as the lower limit and 970 as the upper limit to observe the development of the epidemic with different quanta rate. The following table shows the value of the parameters:

Table 5-1 Parameters for SEIR with indoor airborne

	Value	measurement
N	200	
I	60	
E	0	
R	0	
A	4.5	AC/h
p	0.36	m^3/h
V	27216	m^2
q	200-970	$quanta/h$
α	0.33	1/day
γ	0.5	1/day
T	90	day

5.6.2 Results

In the experiment, the quanta rates were set to 200, 400, 600, 800 and 970 respectively. *Figure 5.17* shows the changes in Susceptible, Expose, Infected and Recovered people.

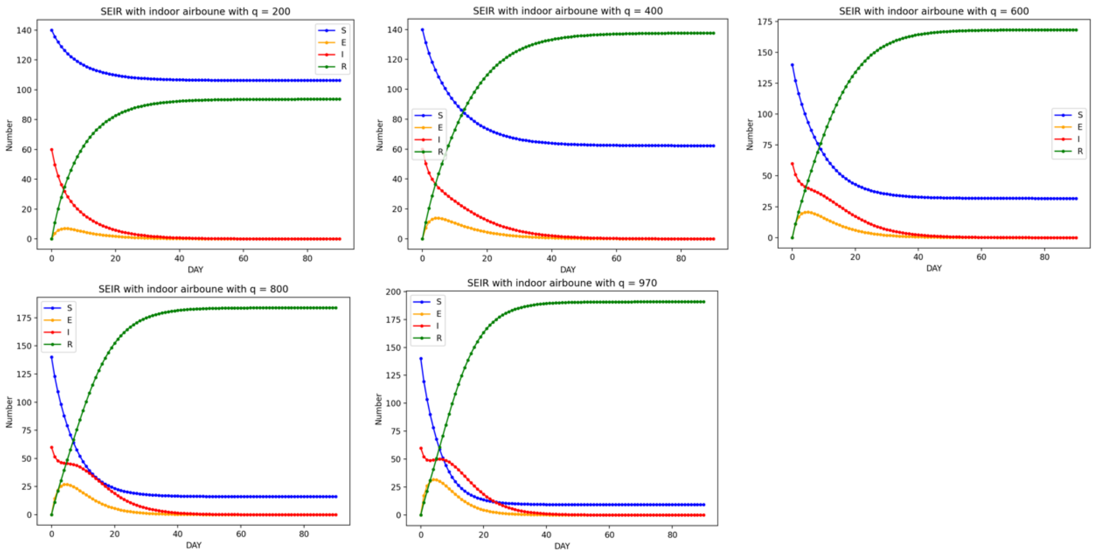

Figure 5.17 change in different group of people

The graph shows that as q increases, the number of infections increases. When q is 200, the number of susceptible people finally stabilises at 110. After q rises to 800 and 970, the final number of susceptible people is under 25. At the same time, the epidemic cycle extended. The number of infected people becomes zero in about 35 days when the q is 200, but in about 40 days after the q increases.

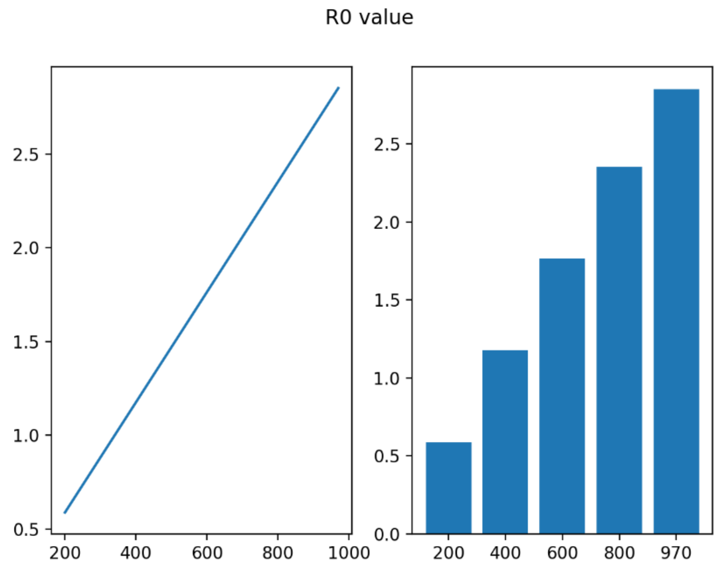

Figure 5.18 R0 value

Figure 5.18 shows the variation of the R0 value, which increases linearly and exceeds the critical value (R0 = 1) after the quanta rate reaches 400 ACH and becomes 2.35 and 2.85 after rising to 800 and 970 ACH. According to the report in (Virginia Department of Health, 2022), the R0 value of flu is between 1-2. When the quanta rate is over 800, it will propagate much quicker than the regular flu in the supermarket scenario.

5.7 Summary

This chapter analyses the factors affecting the spread of the epidemic and the long-term development of the epidemic. The initial number of infected people affects the infection rate and exposure time, and when the proportion of infected people reaches 40%, the infection ratio rises from 5% to 10.4%. The change in shopping time reduced the infection rate from 5% to 3% and effectively reduced the number of new cases in the shopping area. The percentage of exposure time in the shopping area was reduced from 48% to

41%, and most new cases only occurred in congested zones 1 and 2. Increasing the number of people checking out per minute did not significantly reduce the number of infections, but it decreased the exposure time ratio in the checking-out area from 52% to 46%, effectively controlling the spread of the epidemic in that place. This was followed by an analysis of the long-term influence using the SEIR model, where a rapid epidemic spread occurred when the quanta rate exceeded 800 and the R0 value exceeded 2. The next chapter will summarise the whole project and suggest future directions for the work.

6 Conclusion and future work

In this experiment, a standard supermarket model was constructed using a grid layout. Customer mobilities were simulated using DTN, and the SEI model was applied to analyse the spread of the virus. The experiment went smoothly, as the risk of exposure and infection were reduced by controlling factors such as shopping time and checking-out speed. The long-term epidemic development was analysed by the SEIR model as well.

In the future, more complex supermarket layouts will be designed, such as adding more shopping areas or non-rectangular areas. What's more, the path design was not included when taking into account the factors affecting the spread of the epidemic. In the future, the customer flow of each route should be considered to avoid the situation of multiple paths intertwining in zone 1 and zone 2. Finally, the impact of masks on exposure should also be felt in the epidemic model, which will require a lot of data collection in the future to find out how long customers will get infected with masks.

Bibliography

Adam P. Vrechopoulos a, R. M. O. b. G. I. D. a. G. J. S. c., 2004. Virtual store layout: an experimental comparison in the context of grocery retail. *Journal of Retailing*, pp. 13-22.

Authority, H. I. a. Q., 2020. *Rapid review of public health guidance on physical distancing in the context of COVID-19.* [Online]
Available at: https://www.hiqa.ie/sites/default/files/2020-06/Rapid-review-of-public-health-guidance-for-physical-distancing_COVID-19_0.pdf
[Accessed 8 July 2022].

Brown, V., 2005. *Grocery Store Desig,* s.l.: maa.

C M Macal, M. J. N., 2010. Tutorial on agent-based modelling and simulation. *Journal of Simulation*, pp. 151-162.

C. J. NOAKES, C. B. B. P. A. S. K. G. K., 2006. Modelling the transmission of airborne infections in enclosed spaces. *Epidemiol Infect*, 14 Feburary.

Carroll, R. G., 2007. *Elsevier's Integrated Physiology.* s.l.:Elsevier.

Cleveland Clinic, 2020. *What is the Coronavirus 15-Minute Rule?*. [Online]

Available at: https://health.clevelandclinic.org/what-is-the-coronavirus-15-minute-rule/

[Accessed 8 August 2022].

Colin J.Axon, M. K. E. R. S. A. T. J. C. T. &. J. G. W., 2020. *Airflow in Supermarkets and Similar Retail Stores: a rapid survey on infection transmission.* [Online]

Available at: https://bura.brunel.ac.uk/bitstream/2438/20857/1/FullText.pdf

[Accessed 10 July 2022].

Crist, C., 2020. *Study: COVID Can Be Spread in Less Than 15 Minutes.* [Online]

Available at: https://www.webmd.com/lung/news/20210126/study-says-covid-can-spread-in-less-than-15-minutes

[Accessed 8 July 2022].

EPA, 2021. *Indoor Air and Coronavirus (COVID-19).* [Online]

Available at: https://www.epa.gov/coronavirus/indoor-air-and-coronavirus-covid-19

[Accessed 1 July 2022].

Fabian Ying, A. O. G. W. M. B.-D. M. A. P. a. S. D. H., 2022. Customer mobility and congestion in supermarkets. *PHYSICAL REVIEW E*.

Fabian Ying, N. O., 2021. Modelling COVID-19 transmission in supermarkets using an agent-based model. *PLOS ONE*, 29 April.

Fall, K., 2003. *A Delay-Tolerant Network Architecture for Challenged Internets.* Intel Research, Berkeley, sigcomm.

Feldman, N., 2022. *How long does it take to recover from COVID?.* [Online] Available at: https://whyy.org/articles/how-long-does-it-take-to-recover-from-covid/ [Accessed 10 July 2022].

Gabriela Moutinho de Souza Dias, R. M. S., 2012. *Epidemic SIR Model Applied to Delay-Tolerant Networks.* BRASILIA, BRAZIL, 30th BRAZILIAN TELECOMMUNICATIONS SYMPOSIUM.

Gupta, A., 2020. *Study highlights COVID-19 spread indoors: 5 minutes exposure, 20 feet away and still infected.* [Online]
Available at: https://www.timesnownews.com/health/article/study-highlights-covid-19-

spread-indoors-5-minutes-exposure-20-feet-away-and-still-infected/692889

[Accessed 8 July 2022].

Herb Sorensen, S. B. K. A. G. T. A. S. R. K. B. P. M. W., 2017. Fundamental patterns of in-store shopper behavior. *Journal of Retailing and Consumer Services,*, Volume 37, p. 37.

Hong Lian, Z. W., 2007. The Computer Simulation for Queuing System. *World Academy of Science*, August.

HSE, 2022. *5a COVID-19.* [Online]

Available at: https://www.hse.ie/eng/health/immunisation/hcpinfo/guidelines/covid19.pdf

[Accessed 10 July 2022].

HSE, 2022. *Protect yourself and others from COVID-19.* [Online]

Available at: https://www2.hse.ie/conditions/covid19/preventing-the-spread/protect-yourself-and-others/

[Accessed 11 August 2022].

I. Stojmenovic, M. R. a. B. V., 2000. *Depth first search and location based localized routing and QoS routing in wireless networks.* s.l., Proceedings 2000 International Conference.

Ian Cooper 1, A. M. 2. C. G. A. 2., 2020. A SIR model assumption for the spread of COVID-19 in different communities. *Chaos Solitons Fractals*, 28 Janurary.

Jin Li, Z. C. Y. Z. N. M. S. G. Q. W. L. Z. &. E. L., 2021. Evaluation of infection risk for SARS-CoV-2 transmission on university campuses. *Science and Technology for the Built Environment*, 21 July, pp. 1165-1180.

Kenneth H. Rosen, P., 2003. *HANDBOOK OF GRAPH THEORY.* 1 ed. s.l.:DISCRETE MATHEMATICS AND ITS APPLICATION.

Laura Gammaitoni, M. C. N., 1997. Using a Mathematical Model to Evaluate the Efficacy of TB Control Measures. *Emerging Infectious Diseases*, July, pp. 342, vol.3.

Lewis, D., 2021. *Why indoor spaces are still prime COVID hotspots.* [Online] Available at: https://www.nature.com/articles/d41586-021-00810-9
[Accessed 1 July 2022].

Lidl, 2022. *THE STANDARD STORE MODERN LOCAL SHOPPING FACILITY.* [Online] Available at: https://property.lidl.ie/location_13346/standard-store [Accessed 2 July 2022].

Louten, J., 2016. Virus Transmission and Epidemiology. *Elsevier Public Health Emergency Collection*, 6 May.

Masson, G., 2022. *What we know about omicron symptoms, recovery: A clinical timeline.* [Online] Available at: https://www.beckershospitalreview.com/public-health/what-we-know-about-omicron-symptoms-recovery-a-clinical-timeline.html [Accessed 2 July 2022].

McKendrick, W. O. K. a. A. G., 1927. A contribution to the mathematical theory of epidemics. *Journal*, 1 August.

Miller, S. L. N. W. W. J. J. L., 2021. Transmission of SARS-CoV-2 by inhalation of respiratory aerosol in the Skagit Valley Chorale superspreading event.. *Indoor air*, 1 March.

MUKHLISH FUADI, I. M., 2015. THE DESIGN OF HYBRID VIRTUAL STORE LAYOUT (HVSL): A SIMULATION EXPERIMENT OF INDONESIAN CUSTOMERS' CONTEXT. *Journal of Basic and Applied Research*, 12 November.

MULDOWNEY, M. Y. L. A. J. S., 1994. Global Stability for the SEIR Model in Epidemiology. *ELSEVIER*, 6 April.

PhD, R. G. C., 2007. *Pulmonary System.* 1 ed. s.l.:s.n.

Piovani Duccio, A. E. U. G. W. A. a. B. M., 2018. Measuring accessibility using gravity and radiation models. *Royal Society Open Science*, 12 September.

Ross Beckley, C. W. M. A. M. C. A. J. a. G. S. B., 2013. *Modeling epidemics with differential equations.* [Online]
Available at: https://www.tnstate.edu/mathematics/mathreu/filesreu/GroupProjectSIR.pdf
[Accessed 1 July 2022].

Saina Abolmaali, S. S., 2021. A comparative study of SIR Model, Linear Regression, Logistic Function and ARIMA Model for forecasting COVID-19 cases. *AIMS Public Health*, 26 August.

Stallings, W., 2005. *QUEUING ANALYSIS.* [Online]

Available at: http://csci.csusb.edu/ykarant/courses/w2005/csci531/QueuingAnalysis.pdf

[Accessed 1 July 2022].

Tamer Abdelkader, K. N. A. N. N. G. V. S., 2016. A performance comparison of delay-tolerant network routing protocols. *IEEE Network,* 30(2).

The sun, 2022. *WAIT IN LINE Brits spend eight and a half months in supermarkets throughout their lives.* [Online]

Available at: https://www.thesun.co.uk/money/17325634/brits-spend-eight-and-a-half-months-in-supermarkets/

[Accessed 12 July 2022].

Virginia Department of Health, 2022. *COVID-19 and Influenza Surveillance.* [Online]

Available at: https://www.vdh.virginia.gov/coronavirus/2022/01/07/covid-19-and-influenza-surveillance/

[Accessed 12 July 2022].

WHO, 2022. *Coronavirus disease (COVID-19): How is it transmitted?.* [Online]

Available at: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted

[Accessed 1 August 2022].

Wikipedia, 2022. *Compartmental models in epidemiology.* [Online]

Available at: https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology

[Accessed 1 July 2022].

Xiang Zhou, X. M. ,. N. H. L. S. Y. M. J. H. H. J. C. L. G. S. ,. Z. S. Z. Y. L., 2020. *Forecasting the Worldwide Spread of COVID-19 based on Logistic Model and SEIR Model.* [Online]

Available at: https://www.medrxiv.org/content/10.1101/2020.03.26.20044289v2.full.pdf

[Accessed 10 July 2022].

Xing Wenjie, L. S. H. L., 2015. *Simulation model of supermarket queuing system.* s.l., 2015 34th Chinese Control Conference.

Zipf, G. K., 1946. The P1 P2/D Hypothesis: On the Intercity Movement of Persons.. *American Sociological Review,* Volume 11.

Zyl, L. V., 2015. *Fit & Proper: What is the ideal walking speed for you?.* [Online] Available at: business-standard.com/article/current-affairs/fit-proper-what-is-the-ideal-walking-speed-for-you-115100900029_1.html [Accessed 5 July 2022].