Technische Universiteit Delft

Developing a Markov-Modulated Process Model for Mobility Processes

Liufei Yang

TU Delft

NAS

Developing a Markov-Modulated Process Model for Mobility Processes

by

Liufei Yang

to obtain the degree of Master of Science
in Electrical Engineering
Track Telecommunications and Sensing Systems
at the Delft University of Technology,
to be defended publicly on Monday June 28, 2021 at 1:30 PM.

Student number:	4906446	
Project duration:	August, 2020 – June, 2021	
Thesis committee:	Prof. dr. ir. Piet Van Mieghem,	TU Delft, supervisor
	Dr. Johan Dubbeldam,	TU Delft
	Dr. Mattia Sensi,	TU Delft, daily supervisor

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

TU Delft

Abstract

A major factor in the spreading of viruses is human-to-human transmission, and human mobility is clearly linked to the spreading process of epidemics. If we hope to understand the evolution of an epidemic, then we must also understand the underlying mobility process and the interaction between the two.

We propose the Markov modulated process (MMP) model as a tool for modeling mobility processes. We take a link-based approach to the MMP model, where modulating actions are a joint combination of adding and removing a number of links. We demonstrate that for such an approach, the states of the Markov process must encode not only a modulating action, but also the current number of links in the graph. We further show that in this case, the number of states will increase with $O(N^6)$ where N is the number of nodes. We refer to this as the "unaggregated" MMP model, and we introduce the concept of state aggregation to create the "aggregated" MMP model which only requires $O(N^2)$ states.

We demonstrate that both the aggregated and unaggregated MMP models are able to match the average number of links in the simulated mobility process, capturing the long-term dynamics of the mobility process with high accuracy. Both the aggregated and unaggregated MMP models are also able to match the average number of links added and removed between two time steps, capturing the short-term dynamics with high accuracy. Finally, we demonstrate that for certain mobility processes, the aggregated MMP model is able to produce results which are indistinguishable from the unaggregated MMP model while requiring just $O(N^2)$ states compared to $O(N^6)$.

Preface

This thesis project was carried out as part of the degree requirements for a Master of Science in Electrical Engineering at the Delft University of Technology. The project has been carried out at the Network Architectures and Services (NAS) group. I would like to express gratitude to Professor Piet Van Mieghem for providing me the opportunity to research this interesting topic and for the feedback and guidance along the way. I have thoroughly enjoyed my time working on my thesis project here at the NAS group.

For my thesis, I have been researching the modeling of mobility processes under the daily supervision of Dr. Mattia Sensi, Fenghua Wang, and Massimo Achterberg. I would like to thank my supervisors for the fruitful discussions and helpful advice during our weekly meetings. I am especially grateful for all of the proofreading and constructive feedback they have provided as I worked on writing this thesis. It has truly been a pleasure working with all of them.

I would like to thank Dr. Johan Dubbeldam for being a part of my thesis committee. I would also like to thank Dr. Marco Rinaldi for graciously sharing data with me so that I could test the model.

Finally, I am very thankful for my family and friends for their continued support and encouragement throughout my studies.

Liufei Yang
Den Haag, June 2021

Contents

1

Introduction

On March 11, 2020, the World Health Organization (WHO) officially declared the COVID-19 outbreak a pandemic. Since then, over one year has passed and the pandemic is still not under control in many countries. Worldwide, the WHO reports the cumulative number of confirmed cases has now exceeded 161 million and the number of deaths is over 3.3 million [1]. The spread of the virus is still not contained; at the time of writing, there are over 677,000 newly reported cases in the past 24 hours alone.

As a response to the pandemic, many countries have implemented measures such as travel restrictions, lockdowns, and curfews in an effort combat the spread of the virus. The link between restrictions on human mobility and the evolution of the pandemic is a topic which is under extensive research. Zhou et al. [2] investigated the effect of mobility restrictions on the spread of COVID-19 and found that by limiting the proportion of symptomatic individuals who could move freely to 20%, the peak of the epidemic was delayed by two weeks. Kishore et al. [3] found that in a simulated epidemic, the travel surge before a lockdown could actually lead to a faster and further initial spread of the epidemic. Espinoza et al. [4] showed that in some cases, allowing unrestricted mobility between a high-risk and low-risk community could result in an overall smaller epidemic size, whereas imposing a quarantine on the high-risk community reduced the number of infections in the low-risk community but led to an increase in the overall epidemic size.

Human mobility clearly has an impact on spreading of viruses and the evolution of the pandemic. If we hope to understand the evolution of the pandemic, then we must also understand the human mobility process on which the pandemic spreads and the interaction between the two. In this research, we investigate the potential application of the Markov modulated process (MMP) model as a tool for modeling human mobility. We consider a simulated mobility process represented as a series of adjacency matrices where the links in the adjacency matrix are the contacts between nodes. We create an MMP model which modulates a graph by applying actions on the graph. The goal is to design an MMP model which produces a modulated graph which, when expressed as a series of adjacency matrices, matches the behavior of the adjacency matrices of the mobility process.

The structure of this thesis is as follows. In Chapter 2, we define the simulated mobility process that we will be modeling. In Chapter 3, we define what an MMP model entails and we construct our first MMP model. We evaluate the performance of this first model and identify its limitations. In Chapter 4, we design two new MMP models to address the limitations of the first MMP model. In Chapter 5, we compare the performance of the two new MMP models. In Chapter 6, we consider the applications of our MMP model in the context of modeling a real mobility process and performing an SIR simulation. Finally, we present our conclusions and discuss the scope for future research in Chapter 7.

2

The Simulated Mobility Process

2.1. Initial Conditions

We start by considering a simulated mobility process. We consider a square region of size $W \times W$ units in which nodes are able to move. Nodes are dimensionless points and they are constrained to move within this region. We consider N nodes where N is a square number. Initially, the N nodes are arranged in a square lattice which has n rows and n columns where $n = \sqrt{N}$; in Figure 2.1, the initial positions of the nodes for $N = 25$ are shown. The nodes are represented by the blue circles and the border of the region is shown by the green square.

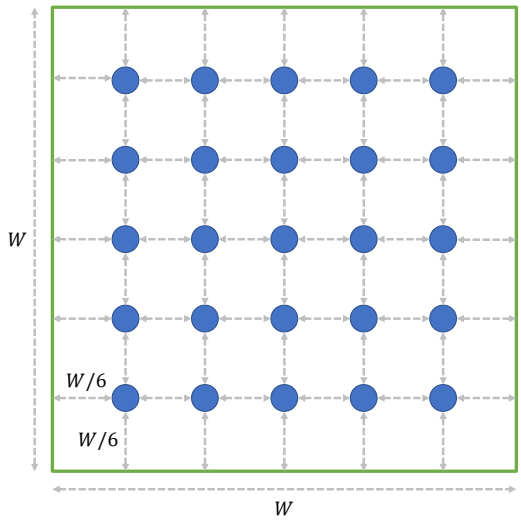


Figure 2.1: Initial positions of nodes for $N = 25$. The nodes are represented by the blue circles and the border of the region is shown by the green square. With $N = 25$, we have $n = 5$ and the spacing in the lattice is $W/6$.

The columns and rows of the starting lattice are spaced evenly within the region; the adjacent columns and rows are separated by a distance of $W/(n+1)$. The distance between the edges of the lattice and the border of the region is also $W/(n+1)$. Note that the figure is not to scale as the nodes have been enlarged for illustration purposes; in our simulations, the nodes are actually dimensionless points.

2.2. Node Movement

In the simulated mobility process, the positions of the nodes are evaluated at discrete time steps. At time $t = 0$, each node is initialized with a random direction. The direction is a random integer in the range $[0, 359]$

and is uniformly distributed. The direction represents an angle which is defined in the same manner as the unit circle (that is, East is defined as 0 degrees, and the angle increases counterclockwise).

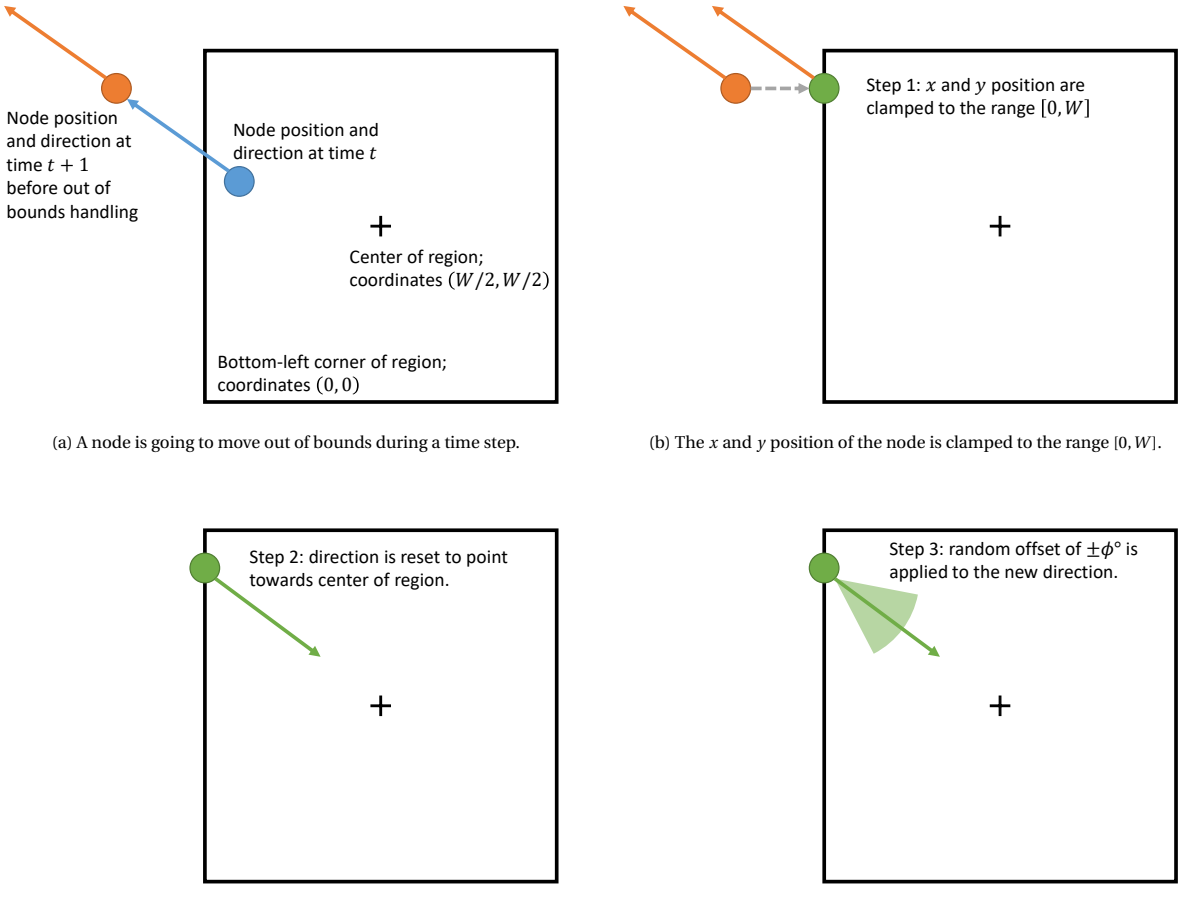
At each time step, each node can change its direction by up to $\pm\theta$ degrees. The random change in direction is an integer in the range $[-\theta, \theta]$ and is uniformly distributed. Each node then moves forward in a straight line along its current direction. The move distance is uniformly distributed in the range $[0, 2v]$ so that average move distance is v units. The move distance is not necessarily an integer and can take any value in the range $[0, 2v]$.

2.3. Constructing the Contact Network

At each time step, every node moves according to the procedure described in the previous section. After the nodes have moved, we create a network based on the current positions of the nodes. We consider two nodes to be connected when the distance between them is d units or less. All of the connections over all of time steps form the contact network G which is represented by the time-varying adjacency matrix $A(t)$. Each element $a_{i,j}(t)$ describes if there is a connection between node i and node j at time t. When there is a connection, $a_{i,j}(t) = 1$; otherwise, $a_{i,j}(t) = 0$.

2.4. Out of Bounds Handling

We constrain the movement of the nodes to the square region of size $W \times W$. We use Cartesian coordinates to represent the position of each node, with the bottom-left corner of the region having coordinates $(0, 0)$. An illustration of the out of bounds handling is shown in Figure 2.2.



(a) A node is going to move out of bounds during a time step.

(b) The x and y position of the node is clamped to the range $[0, W]$.

(c) The direction of the node is reset to point to the center of the region.

(d) A random offset of $\pm\phi°$ is applied to the new direction.

Figure 2.2: Illustration of the out of bounds handling in the simulated mobility process.

At each time step, if a node attempts to move out of bounds (Figure 2.2a), the following steps are applied:

1. The x and y coordinates of the node are both clamped to the range $[0, W]$, bringing the node's position back into the region (Figure 2.2b).

2. The direction of the node is reset to point towards the center of the region (Figure 2.2c). The direction can be viewed as a vector from the node's new position pointing towards the point $(W/2, W/2)$. The direction is an angle expressed in degrees and is rounded to the nearest integer.

3. A random offset is applied to the new direction of the node (Figure 2.2d). The offset is a random integer angle expressed in degrees and is uniformly distributed in the range $[-\phi, \phi]$.

2.5. Simulation Parameters

For the mobility process simulations, we choose $N = 25$ nodes and $W = 10$ for the size of the region. At each time step, nodes change their angle by a random integer uniformly distributed in the range $[-\theta, \theta]$ and we choose $\theta = 20$ degrees. After changing their angle, nodes move a random distance uniformly distributed in the range $[0, 2v]$; we choose $v = 1$ so the average distance moved in a time step is 1 unit. When a node attempts to move out of bounds, its angle is reset to point towards the center of the region, after which a random angle offset uniformly distributed in the range $[-\phi, \phi]$ is applied. We choose $\phi = 30$ degrees.

Each trial of the simulated mobility process runs for $T = 1,000$ time steps. The contact network produced by a single trial is represented by a time-varying series of adjacency matrices from $A(0)$ to $A(1,000)$, with each matrix $A(t)$ describing the contacts in the network at time t. At each time t, two nodes are considered to be connected if the distance between them is $d = 1.5$ units or less. We perform 100 independent trials of the simulated mobility process.

2.6. Mobility Process Simulation Results

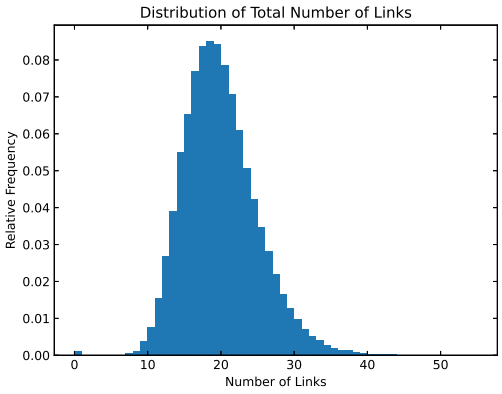
After performing 100 trials of 1,000 time steps each, we analyze the results of the mobility simulations. At each time step, the adjacency matrix of the graph has been recorded. To start with, we consider the distribution of the number of links in the graph at each time step, and for each trial, we calculate the following values:

- Total number of links at each time step.
- Number of links added between two time steps.
- Number of links removed between two time steps.

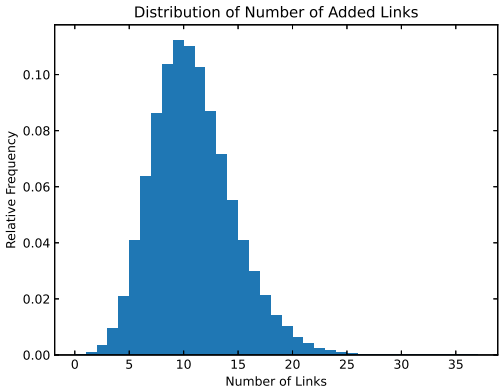
After calculating these values at each time step for each trial of the mobility process, we plot the distribution of these values in Figure 2.3. The average values are as follows:

- Average number of links at each time step: 19.466.
- Average number of links added between two time steps: 10.377.
- Average number of links removed between two time steps: 10.357.

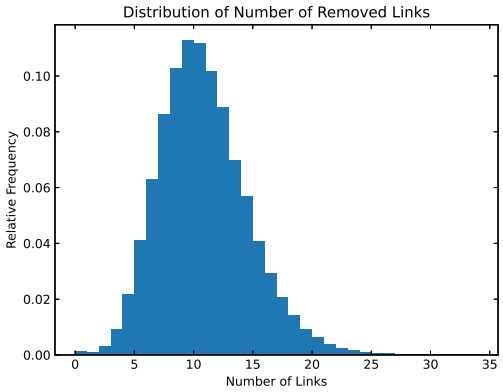
A plot of the number of links at each time step for one trial of the mobility process is shown in Figure 2.4.



(a) Distribution of the total number of links at each time step (bin width 1).



(b) Distribution of the number of links added between two time steps (bin width 1).



(c) Distribution of the number of links removed between two time steps (bin width 1).

Figure 2.3: Histograms of link statistics for the simulated mobility process with 25 nodes across 100 trials of 1,000 time steps.

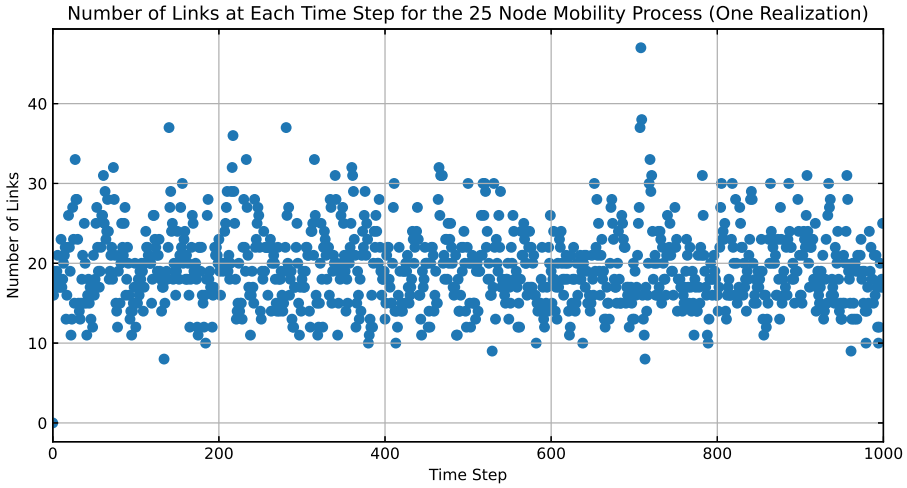


Figure 2.4: Number of links at each time step for one trial of the mobility process.

3

Modeling the Simulated Mobility Process Using an MMP Model

3.1. Definition of the MMP Model

We start with a formal definition of the Markov modulated process (MMP) model. The idea behind the model is a simple one: consider a temporal graph G where the links in the graph can be changed over time, and the changes in the temporal graph are controlled by a Markov process. Each state of the Markov process has an action associated with it, and when the Markov process enters that state, we apply the corresponding action to modulate the graph. In principle, the action associated with a state can be anything. Some example actions include:

- Remove 1 link from the graph.
- Add 10 links to the graph.
- Remove all links from the graph.
- Choose a random number integer between 1 and 10 and add that many links.
- Generate an Erdős-Rényi graph with link probability $p = 0.1$.
- Look outside your window. If it is sunny, add 1 link. If it is raining, remove 1 link.

The last example serves to illustrate that there is no restriction on what the actions can be; they are limited only by your creativity. This poses the first challenge of the MMP model: what should we define as our actions, and how many actions do we want to define?

The evolution of the Markov process is controlled by the probability transition matrix P where each element $p_{i,j}$ of the probability transition matrix is the probability of transitioning from state i to state j. Given N actions and hence N states, the size of the probability transition matrix P is $N \times N$. The probability transition matrix poses the second challenge of the MMP model: how should we choose our transition probabilities?

In this chapter, we will take the first step in applying the MMP model to model the simulated mobility process discussed in Chapter 2. In Section 3.2, we define the goal of our modeling process, and in Section 3.3, we define our criteria for evaluating the goodness of the model. In Section 3.4, we develop our first MMP model and devise a strategy for defining the actions and transition probabilities. In Section 3.5, we use the MMP model we have created to modulate a graph, and we compare the resulting modulated graphs with the graphs of the mobility process. We perform a steady state analysis of our MMP model in Section 3.6 to better understand the behavior of our model, and we conclude the chapter with a discussion in Section 3.7.

3.2. Goal of the Modeling Process

After the mobility process simulations have been performed, we can start modeling the mobility process using an MMP model. We have a series of adjacency matrices which are derived from the mobility process. There are 100 trials of the mobility process and each trial is performed for $T = 1,000$ time steps. The results

of each trial consists 1,001 adjacency matrices with each matrix representing the contacts in the graph at time t. Each time step produces one adjacency matrix and there is an additional adjacency matrix at time $t = 0$ representing the initial state of the graph. The adjacency matrix at time $t = 0$ is the same for every trial because the nodes are always initialized in the same evenly-spaced square lattice. The spacing in the lattice is $10/6 \approx 1.667$ which exceeds $d = 1.5$; therefore, there are no links in the graph at $t = 0$ and the adjacency matrix is always filled with zeros. The goal of the modeling process is to use a discrete Markov process to modulate a graph of 25 nodes such that the series of adjacency matrices corresponding to the modulated graph has properties which resemble those of the series of adjacency matrices produced by the mobility process.

3.3. Evaluating the Goodness of the Model

To evaluate whether the graph modulated by the Markov process is a good model for the mobility process, we compare some metrics of the graphs produced by the MMP model with the metrics of the simulated mobility process. To begin with, we consider same three metrics considered in Section 2.6.

- Average number of links at each time step.
- Average number of links added between two time steps.
- Average number of links removed between two time steps.

These metrics have been chosen because they capture the most fundamental properties that need to be modeled. The number of links at each time step captures the long-term behavior of the model. On average, the Markov modulated process should produce graphs which have the same number of links as the mobility process. If this condition is not satisfied, then the Markov modulated process is either producing graphs with too many or too few links, and it does not accurately model the mobility process.

The number of links added and the number of links removed between two time steps capture the short-term behavior of the model. These metrics describe the rate at which the graph is evolving from one time step to the next. In the mobility process, a link appears when two nodes move within a certain distance d of each other during a time step. If they move out of this range, the link disappears. If the Markov modulated process is a good model for the mobility process, the graphs it produces should be evolving at the same rate as the mobility process. Therefore, on average, the number of links added and removed between two time steps for the Markov modulated process should be the same as that of the mobility process.

3.4. Developing the First MMP Model

3.4.1. Deriving State and Actions

We would like the MMP model to follow the same time scale as the mobility process such that one step in the MMP model corresponds to one step in the mobility process. Otherwise, we will lose some of the structure of the mobility process. In the simulated mobility process, we observe that from one time step to the next, links can be added as well as removed. Therefore, in order to accurately model the mobility process, our MMP model should also be able to add and remove links in a single time step. We therefore define the actions as jointly adding and removing a certain number of links from the graph.

From the series of adjacency matrices produced by the mobility process simulations, we can see how many links are added and removed between each two time steps. Across the 100 trials, the maximum number of links added in a single time step was 36 and the minimum was 0. The maximum number of links removed in a single time step was 33 and the minimum was 0. If we consider all possible combinations of added and removed links, we would have $37 \times 34 = 1,258$ possible actions (and hence states). We foresee two issues with defining actions in this manner. Firstly, considering every possible combination of added and removed links results in a large number of states that would affect the scalability of the model. Secondly, if we are considering every case where we *exactly* add and *exactly* remove a certain number of links, our states become very specific and we might actually try to model the stochasticity in the mobility process instead of the underlying behavior.

We address both of these issues by binning the data and considering cases when the number of links added or removed falls within a certain *range* of values. To define our bins, we start by observing that although the maximum number of links added is 36, it can be seen from the histogram (Figure 2.3b) that the probability of this occurring is actually very small. In fact, even the probability that the number of links added exceeds 25 is less than 0.1%. For the number of links removed, we observe that the probability of more than 25 links

being removed is also about 0.1% (Figure 2.3c). We therefore considered the following 5 ranges for adding and removing links:

- $[0,5]$
- $[6,10]$
- $[11,15]$
- $[16,20]$
- $[21,25]$

For each combination of ranges of added and removed links, we assign a state and an action. This results in 25 states which are summarized in Table 3.1. The action corresponding to each state is denoted by a tuple which represents the number of links to add and remove from the graph. For example, in state 2, the action is $(+3,-8)$, which means to add 3 links and remove 8 links. The number of links to add and remove are calculated from the midpoint of the corresponding ranges used to define the state. In state 2, the range for added links is $[0,5]$ and the midpoint is 2.5, which is rounded to 3. The range for removed links is $[6,10]$ and the midpoint is 8.

Table 3.1: Summary of states and actions for the first MMP model with 25 nodes.

Added Links	**Removed Links**	**State**	**Action**
[0, 5]	[0, 5]	1	(+3, -3)
[0, 5]	[6, 10]	2	(+3, -8)
[0, 5]	[11, 15]	3	(+3, -13)
[0, 5]	[16, 20]	4	(+3, -18)
[0, 5]	[21, 25]	5	(+3, -23)
[6, 10]	[0, 5]	6	(+8, -3)
[6, 10]	[6, 10]	7	(+8, -8)
[6, 10]	[11, 15]	8	(+8, -13)
[6, 10]	[16, 20]	9	(+8, -18)
[6, 10]	[21, 25]	10	(+8, -23)
[11, 15]	[0, 5]	11	(+13, -3)
[11, 15]	[6, 10]	12	(+13, -8)
[11, 15]	[11, 15]	13	(+13, -13)
[11, 15]	[16, 20]	14	(+13, -18)
[11, 15]	[21, 25]	15	(+13, -23)
[16, 20]	[0, 5]	16	(+18, -3)
[16, 20]	[6, 10]	17	(+18, -8)
[16, 20]	[11, 15]	18	(+18, -13)
[16, 20]	[16, 20]	19	(+18, -18)
[16, 20]	[21, 25]	20	(+18, -23)
[21, 25]	[0, 5]	21	(+23, -3)
[21, 25]	[6, 10]	22	(+23, -8)
[21, 25]	[11, 15]	23	(+23, -13)
[21, 25]	[16, 20]	24	(+23, -18)
[21, 25]	[21, 25]	25	(+23, -23)

3.4.2. Calculating the Transition Probabilities

With the states now defined, we use the data from the mobility process to determine the transition probabilities. For each time step t in the simulation, we compare the graph at time t, represented by the adjacency matrix $A(t)$, with the graph of the next time step, represented by $A(t+1)$. We compare the adjacency matrices to determine how many links have been added and removed between these two time steps, and we then use

the ranges defined in Table 3.1 to assign a state to time step t. In the rare cases where the number of links added or removed exceeds 25, we consider it to be in the range $[21,25]$ and assign the state accordingly.

After assigning a state to each time step, we can then derive the probability transition matrix from the sequence of states. For estimating the transition probability from state i to state j, we count the number of times that we observe a transition from state i to j and divide it by the total number of observed transitions from state i to any state k. We can express this as

$$p_{i,j} = \frac{n_{i,j}}{\sum_{k \in X} n_{i,k}}, \tag{3.1}$$

where $p_{i,j}$ is the probability of going from state i to state j, $n_{i,j}$ is the number of times that the transition from state i to state j is observed, and X is set of all possible states.

Using all 100 trials of the mobility process, the transition matrix has been calculated and is shown in Figure 3.1. The probabilities have been rounded to two decimal places for legibility. Larger probabilities are highlighted by a darker shade of green, and gray cells are probabilities that are actually equal to 0 as opposed to small probabilities which appear as 0.00 due to rounding.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0.02	0.06	0.01	0.00	0.00	0.12	0.38	0.05	0.00	0.00	0.07	0.21	0.03	0.00	0.00	0.01	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.02	0.05	0.01	0.00	0.00	0.14	0.34	0.06	0.00	0.00	0.08	0.21	0.04	0.00	0.00	0.01	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.02	0.04	0.01	0.00	0.00	0.14	0.33	0.06	0.00	0.00	0.08	0.22	0.04	0.00	0.00	0.01	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	0.01	0.04	0.01	0.00	0.00	0.11	0.33	0.10	0.01	0.00	0.09	0.18	0.05	0.00	0.00	0.00	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.10	0.00	0.00	0.00	0.05	0.26	0.10	0.00	0.00	0.03	0.28	0.13	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	0.01	0.07	0.02	0.00	0.00	0.05	0.31	0.15	0.01	0.00	0.03	0.18	0.11	0.01	0.00	0.00	0.03	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
7	0.01	0.05	0.02	0.00	0.00	0.05	0.31	0.13	0.01	0.00	0.03	0.20	0.10	0.01	0.00	0.01	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	0.01	0.05	0.02	0.00	0.00	0.05	0.30	0.13	0.01	0.00	0.03	0.20	0.11	0.01	0.00	0.01	0.04	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00
9	0.01	0.04	0.02	0.00	0.00	0.05	0.27	0.15	0.01	0.00	0.02	0.19	0.13	0.02	0.00	0.00	0.04	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	0.00	0.06	0.02	0.00	0.00	0.04	0.24	0.19	0.04	0.00	0.02	0.16	0.11	0.04	0.00	0.01	0.03	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00
11	0.00	0.04	0.04	0.01	0.00	0.01	0.20	0.23	0.03	0.00	0.01	0.13	0.16	0.04	0.00	0.00	0.03	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00
12	0.00	0.04	0.03	0.00	0.00	0.02	0.19	0.22	0.04	0.00	0.01	0.13	0.18	0.04	0.00	0.00	0.03	0.04	0.01	0.00	0.00	0.00	0.01	0.00	0.00
13	0.00	0.03	0.03	0.01	0.00	0.01	0.18	0.21	0.04	0.00	0.01	0.12	0.19	0.05	0.00	0.00	0.03	0.05	0.01	0.00	0.00	0.00	0.01	0.00	0.00
14	0.00	0.03	0.03	0.01	0.00	0.01	0.15	0.22	0.05	0.00	0.01	0.12	0.20	0.06	0.00	0.00	0.02	0.05	0.02	0.00	0.00	0.00	0.00	0.00	0.00
15	0.00	0.02	0.02	0.01	0.00	0.00	0.12	0.25	0.09	0.00	0.00	0.09	0.22	0.09	0.01	0.00	0.01	0.04	0.02	0.00	0.00	0.00	0.01	0.00	0.00
16	0.00	0.02	0.04	0.02	0.00	0.01	0.14	0.17	0.10	0.02	0.00	0.05	0.17	0.11	0.02	0.00	0.02	0.05	0.04	0.00	0.00	0.01	0.01	0.01	0.00
17	0.00	0.02	0.03	0.01	0.00	0.00	0.08	0.20	0.11	0.01	0.00	0.05	0.20	0.11	0.02	0.00	0.02	0.06	0.04	0.01	0.00	0.00	0.01	0.01	0.00
18	0.00	0.01	0.03	0.02	0.00	0.00	0.06	0.18	0.11	0.02	0.00	0.05	0.19	0.13	0.02	0.00	0.02	0.07	0.05	0.01	0.00	0.00	0.01	0.01	0.00
19	0.00	0.01	0.02	0.01	0.00	0.00	0.05	0.18	0.13	0.02	0.00	0.03	0.17	0.15	0.04	0.00	0.02	0.08	0.06	0.02	0.00	0.00	0.01	0.01	0.00
20	0.00	0.01	0.03	0.01	0.00	0.00	0.03	0.18	0.14	0.06	0.00	0.02	0.15	0.18	0.05	0.00	0.00	0.04	0.05	0.03	0.00	0.01	0.01	0.00	0.00
21	0.00	0.00	0.06	0.00	0.00	0.02	0.06	0.08	0.02	0.06	0.00	0.06	0.13	0.19	0.04	0.00	0.02	0.06	0.10	0.04	0.00	0.00	0.04	0.00	0.00
22	0.00	0.00	0.02	0.02	0.01	0.00	0.02	0.08	0.15	0.05	0.00	0.02	0.11	0.20	0.09	0.00	0.00	0.04	0.06	0.07	0.00	0.00	0.01	0.03	0.02
23	0.00	0.00	0.01	0.03	0.00	0.00	0.00	0.09	0.15	0.07	0.00	0.01	0.11	0.19	0.13	0.00	0.01	0.04	0.08	0.04	0.00	0.00	0.00	0.02	0.01
24	0.00	0.00	0.02	0.04	0.01	0.00	0.02	0.07	0.14	0.07	0.00	0.00	0.09	0.18	0.14	0.00	0.00	0.05	0.06	0.07	0.00	0.00	0.01	0.03	0.01
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.12	0.15	0.17	0.00	0.00	0.05	0.15	0.15	0.00	0.00	0.00	0.10	0.07	0.00	0.00	0.00	0.02	0.02

Figure 3.1: Probability transition matrix of the first MMP model.

3.5. First MMP Model Simulation

Now that the states, actions, and probability transition matrix are defined, we are ready to use the Markov process to modulate a graph. The simulation procedure is as follows. To start with, the graph is initialized as an Erdős-Rényi graph with 25 nodes and link probability $p = 19/300$, producing a random graph with average of 19 links. This is chosen because the mobility process had about 19 links on average. The Markov process is initialized in state 7 because it has the highest steady state probability (discussed in further detail in Section 3.6).

In the first MMP model, the action is applied upon entering the state. To illustrate with an example, suppose that at time $t = 0$ the Markov process is in some state x_0. At time $t = 1$, the Markov process transitions to state x_1. The action corresponding to state x_1 is then immediately applied to modulate the graph, and the resulting graph after the action is applied gives us our adjacency matrix at $t = 1$.
We adhere to the following set of rules for choosing which links to add and remove links at each time step:

- The links to be removed are chosen randomly from the set existing links *before* any links have been added. That is is to say, we cannot remove a link that has been added in the current step.

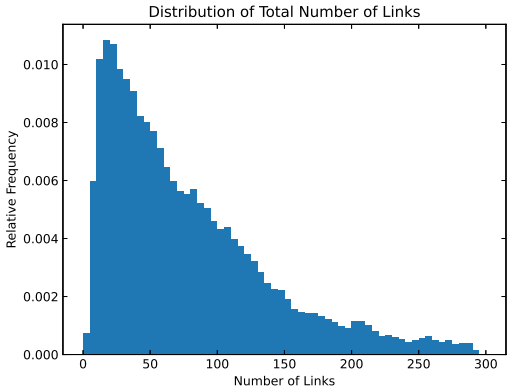
- The links to be added are chosen randomly from the set all non-existing links *before* any links have been removed. That is to say, we cannot add back a link that has been removed in the current step.
- If the action in the current time step calls for us to remove more links than exist in the graph, we remove as many links as possible and ignore the excess.
- If the action in the current time step calls for us to add more links than possible (because the graph will become a complete graph), we add as many links as possible and ignore the excess.

Each simulation is run for 1,000 time steps, and 100 trials are performed. The distributions of the total number of links at each time step, and the number of links added between two time steps, and the number of links removed between two time steps have been plotted in Figure 3.2. The averages are as follows:

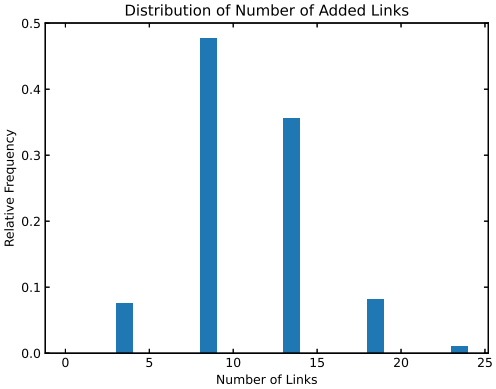
- Average number of links at each time step: 78.377.
- Average number of links added between two time steps: 10.376.
- Average number of links removed between two time steps: 10.275.

We observe in Figure 3.2b and Figure 3.2c that the distribution of added and removed links is concentrated at 3, 8, 13, 18, and 23 links; this is direct result of the way we have defined our actions, since we only ever add or remove these numbers of links.

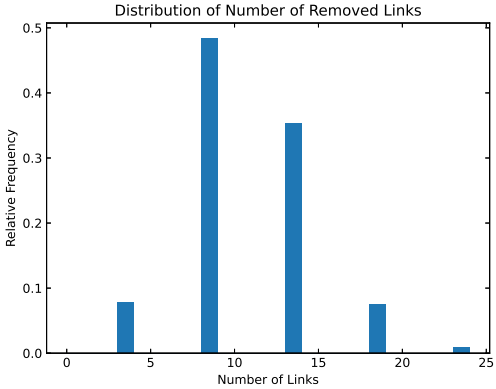
A plot of the number of links in the modulated graph at each time step for one trial of the first MMP model is shown in Figure 3.3a; for comparison, the plot of the number of links at each time step for one of the mobility process has been repeated in Figure 3.3b.



(a) Distribution of the total number of links at each time step (bin width 5).

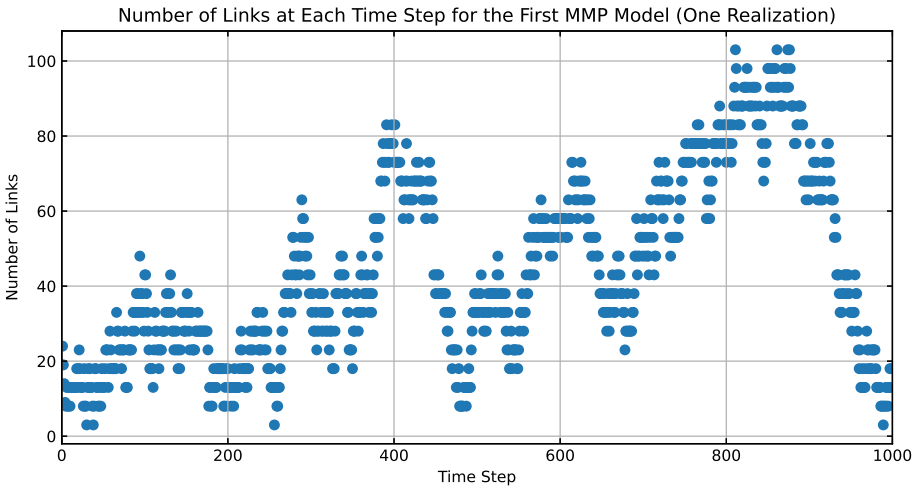


(b) Distribution of the number of links added between two time steps (bin width 1).

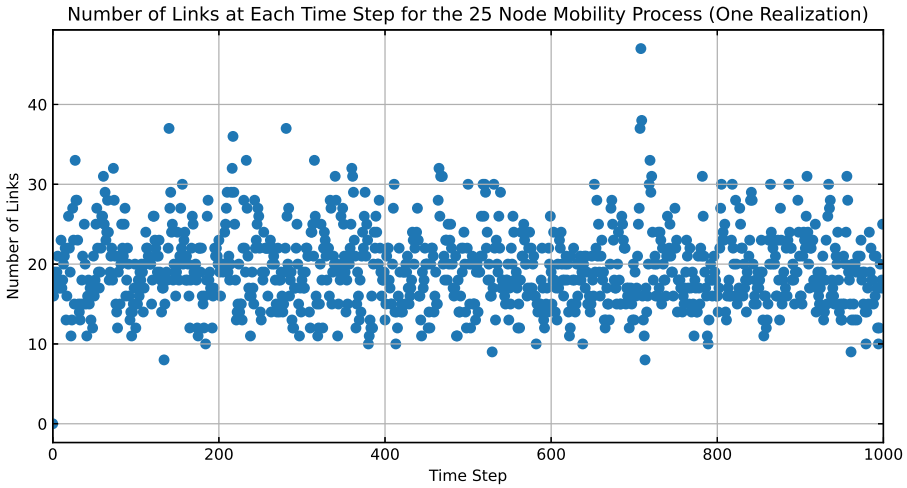


(c) Distribution of the number of links removed between two time steps (bin width 1).

Figure 3.2: Histograms of link statistics for the first MMP model across 100 trials of 1,000 time steps.



(a) Number of links at each time step for one trial of the first MMP model.



(b) Number of links at each time step for one trial of the mobility process.

Figure 3.3: Comparison of the number of links at each time step in the MMP modulated graph and the simulated mobility process.

3.6. Steady State Analysis of the First MMP Model

The average number of links in the graphs produced by this first Markov modulated process model was 78.377, which does not match the average number of links in the simulated mobility process which had an average of 19.466. However, the average number of links added between two time steps was 10.376 and is almost exactly the same as the value observed in the mobility process which had a value 10.377, a difference of less than 0.01%. The average number of links removed in the MMP model was 10.275, which is also very close to the value of 10.357 in the mobility process, differing by less than 1%. To better understand this behavior, we can analyze the steady state of the Markov process.

The steady state vector of the Markov process has been calculated by solving $P^T\pi^T = \pi^T$ [5, pp. 191-192]. The steady state probability for each state is shown in Table 3.2. The number of links added and removed in each state is also shown, and the net change in links Δ has been calculated. If we weight the number of added links in each state by the steady state probability and take the sum, we find 10.371. Therefore, in the long term, we expect 10.371 links to be added in each time step. Taking the weighted sum of the number of links removed, we find 10.356. Therefore, we expect 10.356 links to be removed on average in each time step. Taking the weighted sum of the net change in links, we find $\Delta_{\text{avg}} = 0.015$. This means that on average, we expect a net increase of 0.015 links in each time step. This makes sense because, as we have just calculated,

we are adding slightly more links on average than we are removing.

Table 3.2: Summary of steady state probabilities and actions.

State	Steady State Probability	Add	Remove	Net Change Δ
1	0.0078	3	3	0
2	0.0407	3	8	-5
3	0.0226	3	13	-10
4	0.0041	3	18	-15
5	0.0004	3	23	-20
6	0.0406	8	3	5
7	0.2407	8	8	0
8	0.1620	8	13	-5
9	0.0300	8	18	-10
10	0.0031	8	23	-15
11	0.0229	13	3	10
12	0.1607	13	8	5
13	0.1359	13	13	0
14	0.0326	13	18	-5
15	0.0045	13	23	-10
16	0.0041	18	3	15
17	0.0318	18	8	10
18	0.0329	18	13	5
19	0.0104	18	18	0
20	0.0019	18	23	-5
21	0.0003	23	3	20
22	0.0031	23	8	15
23	0.0043	23	13	10
24	0.0022	23	18	5
25	0.0004	23	23	0

In the modulated graph, we observed an average of 10.376 links added and 10.275 removed between two time steps, which is quite close to the expected values calculated from the steady state value. When combined with the fact that $\Delta_{\text{avg}} = 0.015$ is very close to 0, this explains the behavior of the number of links in the modulated graph in Figure 3.3a. On average, we are adding almost as many links as we are removing, so when looking at the graph of the number of links, it appears to behave like a random walk.

The fact that Δ_{avg} is positive means that on average, we have a net increase in links every time step. In theory, this would mean that over an infinite amount of time, the number of links will go to infinity. However, due to the fact that we have a finite number of nodes and we cannot add more links than the complete graph, this will not occur. When the simulation is repeated for $1,000,000$ time steps, we obtained the following results for the number of links:

- Average number of links at each time step: 156.996.
- Average number of links added between two time steps: 10.327.
- Average number of links removed between two time steps: 10.327.

We observe that even though we expect a net increase in links at each time step, the modulated graph does not converge to the complete graph. The number of links in the graph continues to behave like a random walk with a slight positive bias. This is due to the fact that we do not allow more links to be added when the maximum (complete graph) is reached, and Δ_{avg} is very close to 0 but slightly positive. If we had $\Delta_{\text{avg}} = 0$, we would expect the number of links to behave exactly like a random walk between the minimum and maximum number of links (0 and 300 respectively for a graph of 25 nodes) with an average of 150. However, since $\Delta_{\text{avg}} = 0.015$ is slightly positive, there is a positive bias and we observe that the average number of links is slightly higher than 150.

3.7. Discussion

In the first MMP model, we defined the actions as jointly adding and removing a certain number of links from the graph. We observed in our first MMP model that the average number of links in the graph does not match the simulated mobility process. In the mobility process, we observed that number of links at each time step oscillates around an average value of 19.466 over time. However, in the MMP model, the average number of links behaves like a random walk bounded by 0 and 300 links, the maximum and minimum number of links in a graph of 25 nodes. For a random walk bounded by 0 and 300, we would expect an average value of 150. In the MMP model, we observed an average 78.377 over 100 trials of 1,000 time steps, which is quite far from 150. The reason for this is that 1,000 time steps is not enough time for the random walk to stabilize. When we repeat the MMP model simulation for a single trial of 1,000,000 time steps, we observed an average of 156.996 links, which is quite close to the expected value of 150.

The reason for this random walk behavior is because we add and remove a random number of links at each time step, and on average, the number of links added and removed is almost the same. However, there is no feedback mechanism that forces the number of links in the graph to stay near a certain value. Therefore, the number of links just oscillates between the maximum and minimum number of links and behaves like a random walk. From the steady state, we calculated that we tend to add slightly more links than we remove by a small margin, and this manifests itself as a positive bias in the random walk. This is consistent with the fact that we observed an average of 156.996 links in the single MMP trial of 1,000,000 time steps, which is slightly more than the expected value of 150 in an unbiased random walk.

However, there are similarities between the MMP model and the mobility process that show promise. The average number of links added and the average number of links removed in the MMP model closely matched the values of the mobility process. The average number of links added in the MMP model differed from the average number of links added in the mobility process by less than 0.01%, while the average number of links removed differed by less than 1%. The MMP model is therefore producing graphs which are evolving at almost the exact same rate as the mobility process. Furthermore, when calculating the expected net change in links per time step Δ_{avg} in the MMP model, we find 0.015 which is very close to 0. So on average, the number of links in the graph is not really growing or shrinking, it is close to equilibrium. This matches the behavior that we observed in the mobility process: we see that the number of links is not increasing or decreasing over time; it is hovering around an average.

These results of this chapter illustrate a fundamental limitation of using an MMP model which is “blind” to the number of links in the graph. Without any knowledge of the number of links in the graph, it is not possible to keep the number of links close to an average. Without any form of feedback in the MMP model, the number of links drifts and behaves like a random walk. Additionally, at some time steps, we may also come close to the complete graph or empty graph, and because our MMP model does not know anything about the number of links in the graph, there is the possibility that it tries to apply an invalid action by adding or removing more links than possible. In the next chapter, we will address both of these issues by introducing the concept of a link-aware MMP model which has states that encode the current number of links in the graph.

4

Towards a Link-Aware MMP Model

4.1. Introduction

The results of the first MMP model indicate that a strategy of blindly adding and removing links causes the number of links in the graph to drift and behave like a random walk, which does not match the behavior of the number of links in the simulated mobility process where the average number of links is stable. In order to control the number of links, our model needs to have some kind of feedback to prevent the number of links from drifting. This is only possible if our model contains information about the number of links which exist in the graph.

Another disadvantage of the first MMP model is that the model can try to apply invalid actions to modulate the graph. The number of links which can exist in a finite graph of N nodes is also finite: the lower bound is $L_{\min} = 0$ and the upper bound is given by $L_{\max} = N(N-1)/2$, which is the number of links in the complete graph. Therefore, our actions are limited by the number of links which already exist in the graph, because we cannot remove more links than currently exist and we cannot add any more links once the complete graph is reached. In the first MMP model, we observed numerous instances where the model tried to add or remove more links than possible. In our simulations, we simply added or removed as many links as possible and ignored the excess. However, it would be much better if we could avoid this situation by having a model that only applies valid actions. The set of valid actions is dictated by how many links currently exist in the graph, and we again see that addressing this issue requires the MMP model to contain some information on the number of links in the graph.

In this chapter, we will introduce the concept of a link-aware MMP model; that is an MMP model in which the states of the Markov process also encode the number of links in the graph. As we will see in this chapter, the number of states required for a complete MMP model which models all possible number of links and all possible actions are massive; for a graph of N nodes, the number of states increases with $O(N^6)$. This model quickly becomes infeasible for large N, and we therefore introduce the concept of *state aggregation* to reduce the number of states in the MMP model. We refer to this reduced MMP model as the *aggregated MMP model*, and the complete link-aware MMP model as the *unaggregated MMP model*. Both of these models are link-aware, so we will not explicitly mention it in the name for the remainder of this thesis (unless it is relevant).

The structure of this chapter is as follows. In Section 4.2, we will introduce the unaggregated MMP model and calculate the number of required states. In Section 4.3, we will present a method for aggregating the states of the unaggregated MMP model. In Section 4.3.1, we will illustrate the unaggregated MMP model for $N = 3$. In Section 4.3.2, we will then present a method for state aggregation based on steady state probabilities. We will demonstrate that this method preserves the average number of links in the modulated graph, and further show that the rate at each action is applied in the aggregated MMP model will be the same as in the unaggregated MMP model. In Section 4.3.3, we provide a generalized proof of the described aggregation method by using the concept of *probability flux*. Finally, in Section 4.3.4, we will apply the proof to construct an aggregated MMP model for $N = 3$ nodes. We conclude this chapter with a discussion in Section 4.4.

4.2. The Unaggregated Link-Aware MMP Model

Let us consider the number of states required to create an MMP model in which each state models the number of links in the graph as well as an action. In a graph of N nodes, the maximum and minimum number

links which can exist are given by $L_{\max} = N(N-1)/2$ (the complete graph) and $L_{\min} = 0$ (the empty graph) respectively. We cannot add any more links once the complete graph is reached, and we cannot remove any more links once the graph is empty.

We continue to define actions as jointly adding and removing a certain number of links from the graph. At any time t, the set of valid actions is determined by the number of links which exist in the graph at that time L_t. The maximum number of links that can be removed is bounded by L_t because we cannot remove more links than currently exist in the graph. Similarly, the maximum number of links that can be added is bounded by $L_{\max} - L_t$ because we cannot add any more links once the complete graph is reached. The possible number of links that can be removed is therefore $\{0, 1, 2, ..., L_t\}$, and the possible number of links that can be added is $\{0, 1, 2, ..., L_{\max} - L_t\}$. Each combination of added and removed links is a different action, so there are $(L+1) \times (L_{\max} - L + 1)$ possible actions given L links in the graph. By summing the number of possible actions over all possible numbers of links, we find the number of states $\#X$ in the unaggregated MMP model:

$$\#X = \sum_{L=0}^{L_{\max}} (L+1)(L_{\max} - L + 1). \tag{4.1}$$

The number of states $\#X$ can be expressed exactly in terms of the number of node N. The derivation is shown in Appendix A, and the result is

$$\#X = \frac{N^6}{48} - \frac{N^5}{16} + \frac{5N^4}{16} - \frac{25N^3}{48} + \frac{7N^2}{6} - \frac{11N}{12} + 1. \tag{4.2}$$

It follows from (4.2) that the number of states in the unaggregated MMP model grows with $O(N^6)$. For 2 nodes, there are 4 states; for 3 nodes, there are 20 states; for 4 nodes, there are 84 states; for 25 nodes, there are 4,590,551 states. As the number of nodes increases, the number of states explodes and this model is no longer feasible. In the next section, we will discuss the concept of state aggregation to reduce the number of states.

4.3. The Aggregated Link-Aware MMP Model

As the number of nodes increases, the unaggregated MMP model becomes infeasible due to the massive amount of states. In this section, we will derive an aggregated MMP model which reduces the number of states through state aggregation. The aggregated MMP model we derive remains link-aware: the states of the aggregated MMP model encode the number of links in the graph. The key difference is that in the unaggregated MMP model, we decouple the actions from the states, and instead associate actions with state *transitions*. To illustrate this concept, we will start by considering the unaggregated MMP model for a graph with $N = 3$ nodes.

4.3.1. Unaggregated Link-Aware MMP Model for N=3

A diagram of the unaggregated MMP model for 3 nodes is shown in Figure 4.1. With 3 nodes, there are a maximum of 3 links in the graph. States 1-4 (shown in blue) encode all possible actions when there are 0 links in the graph, states 5-10 (shown in orange) encode all possible actions when there is 1 link in the graph, states 11-16 (shown in green) encode all possible actions when there are 2 links in the graph, and states 17-20 (shown in yellow) encode all possible actions when there are 3 links in the graph.[1]

The diagram has been organized into blocks to make it easier to read. In the blue block, the current number of links in the graph is $L_t = 0$. The grey dotted lines separate the blocks into columns which show the new amount of links L_{t+1} after the action is applied. For example, in state 1, the number of links at time t is 0 and the action to be applied is $(+0, -0)$; that is add 0 links and remove 0 links. The number of links at time $t+1$ will therefore still be 0. State 1 can therefore only transition to states which encode 0 links, and in the diagram it can be seen that the arrow leaving state 1 points to the blue block of states. This represents a transition probability to each one of the states in the blue block, but it is represented using a single arrow to make the diagram easier to read. For state 2, there are 0 links at time t and the action to be applied is $(+1, -0)$, after which there will be 1 link in the graph. Therefore, state 2 can transition to any of the states in the orange block (as the orange states encode the actions when there is 1 link in the graph).

[1] The dash notation will always be used to denote a range of states, e.g. "states 1-4". This is done in order to prevent states ranges from being confused with state transitions.

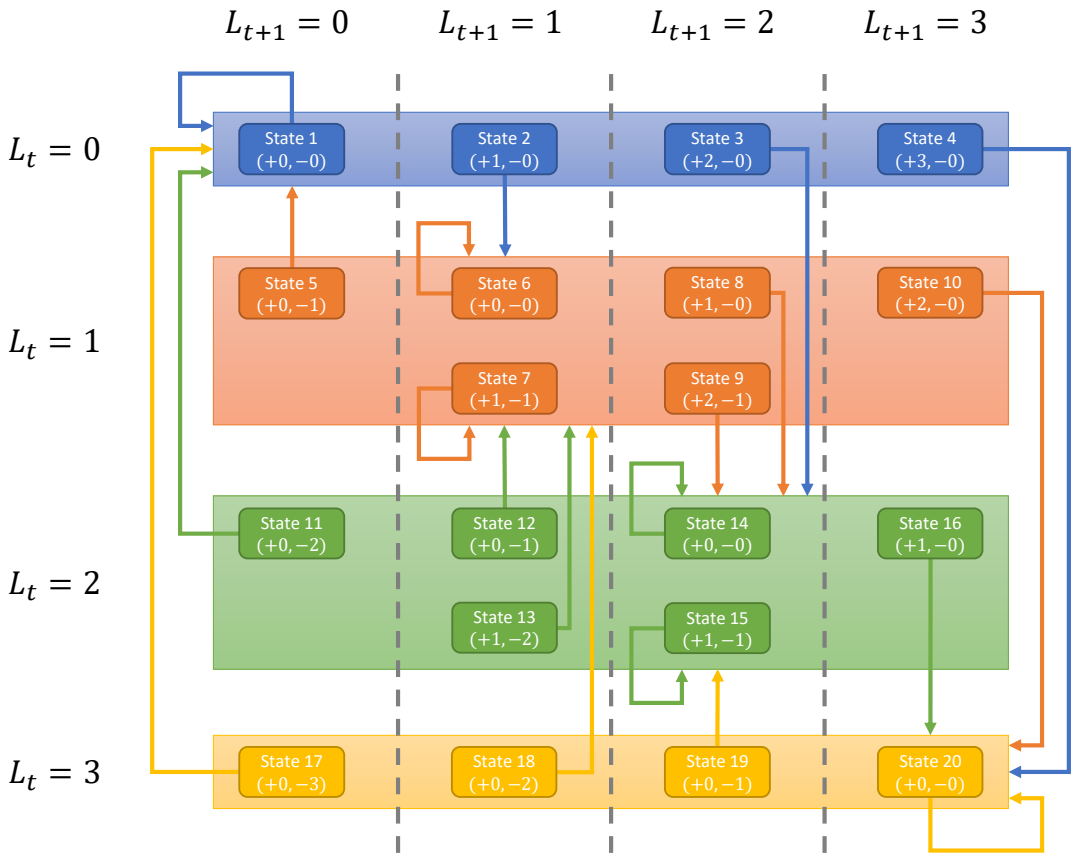


Figure 4.1: Unaggregated MMP model for $N = 3$ nodes.

The expected number of links in a graph modulated by the unaggregated MMP model follows from the steady state of the Markov process. The steady state probability of each state can be interpreted as the long-term fraction of time that the Markov process spends in that state, and each state encodes a certain number of links. Therefore, the sum of the steady state probabilities of states 1-4 is the fraction of time in which there are 0 links, the sum of the steady state probabilities of states 5-10 is the fraction of time in which there is 1 link, the sum of the steady state probabilities of states 11-16 is the fraction of time in which there are 2 links, and the sum of the steady state probabilities of states 17-20 is the fraction of time in which there are 3 links.

4.3.2. An Intuitive Description of State Aggregation

Continue to consider the unaggregated MMP model for $N = 3$ shown in Figure 4.1. The MMP model will be used to modulate a graph, and each state of the Markov process encodes the number of links in the graph and also a corresponding action to apply. Assuming that the steady state vector π of the Markov process exists, the expected value of the number of links in the modulated graph can be directly calculated by weighting each element of the steady state vector with the number of links encoded by the corresponding state. The average number of links is therefore given by

$$L_{\text{avg}} = \left(0 \times \sum_{i=1}^{4} \pi_i\right) + \left(1 \times \sum_{i=5}^{10} \pi_i\right) + \left(2 \times \sum_{i=11}^{16} \pi_i\right) + \left(3 \times \sum_{i=17}^{20} \pi_i\right). \tag{4.3}$$

Since each state also encodes an action, the steady state probability of each state can be interpreted as the fraction of time that the corresponding action is applied.

Recall that the purpose of this chapter is to derive an MMP model where the states encode the number of links in the graph. Therefore, if we want to aggregate states, the logical course of action is to aggregate states where the number of links is the same. That is, the set of states which encode 0 links (states 1-4), the set of states which encode 1 link (states 5-10), the set of states which encode 2 links (states 11-16), and the set of states which encode 3 links (state 17-20) will each become a single state. Observe that the number of states

is now equal to $L_{\max}+1$: each state encodes the number of links in the graph, and the number of links can range from 0 (empty graph) to $L_{\max} = N(N-1)/2$ (complete graph). The number of states in the aggregated MMP model grows with $O(N^2)$, which is greatly reduced compared to the unaggregated MMP model which grows with $O(N^6)$.

In the aggregated Markov process, the states encode only the number of links in the graph; the actions have been decoupled from the states and will instead be tied to state transitions. We will first consider how the transition probabilities for the aggregated states can be derived, and afterwards we will discuss the mechanism for applying actions in the unaggregated MMP model.

Aggregation of States Which Encode 3 Links

In Figure 4.2, the probability transition matrix of the unaggregated MMP model for $N = 3$ nodes is shown. Let us consider states 17-20, the set of states that encode 3 links. We want to aggregate these states into a single new state; we will remove states 17-20 from the original Markov process and replace it with a single state which we name L3.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$																
2					$p_{2,5}$	$p_{2,6}$	$p_{2,7}$	$p_{2,8}$	$p_{2,9}$	$p_{2,10}$										
3											$p_{3,11}$	$p_{3,12}$	$p_{3,13}$	$p_{3,14}$	$p_{3,15}$	$p_{3,16}$				
4																	$p_{4,17}$	$p_{4,18}$	$p_{4,19}$	$p_{4,20}$
5	$p_{5,1}$	$p_{5,2}$	$p_{5,3}$	$p_{5,4}$																
6					$p_{6,5}$	$p_{6,6}$	$p_{6,7}$	$p_{6,8}$	$p_{6,9}$	$p_{6,10}$										
7					$p_{7,5}$	$p_{7,6}$	$p_{7,7}$	$p_{7,8}$	$p_{7,9}$	$p_{7,10}$										
8											$p_{8,11}$	$p_{8,12}$	$p_{8,13}$	$p_{8,14}$	$p_{8,15}$	$p_{8,16}$				
9											$p_{9,11}$	$p_{9,12}$	$p_{9,13}$	$p_{9,14}$	$p_{9,15}$	$p_{9,16}$				
10																	$p_{10,17}$	$p_{10,18}$	$p_{10,19}$	$p_{10,20}$
11	$p_{11,1}$	$p_{11,2}$	$p_{11,3}$	$p_{11,4}$																
12					$p_{12,5}$	$p_{12,6}$	$p_{12,7}$	$p_{12,8}$	$p_{12,9}$	$p_{12,10}$										
13					$p_{13,5}$	$p_{13,6}$	$p_{13,7}$	$p_{13,8}$	$p_{13,9}$	$p_{13,10}$										
14											$p_{14,11}$	$p_{14,12}$	$p_{14,13}$	$p_{14,14}$	$p_{14,15}$	$p_{14,16}$				
15											$p_{15,11}$	$p_{15,12}$	$p_{15,13}$	$p_{15,14}$	$p_{15,15}$	$p_{15,16}$				
16																	$p_{16,17}$	$p_{16,18}$	$p_{16,19}$	$p_{16,20}$
17	$p_{17,1}$	$p_{17,2}$	$p_{17,3}$	$p_{17,4}$																
18					$p_{18,5}$	$p_{18,6}$	$p_{18,7}$	$p_{18,8}$	$p_{18,9}$	$p_{18,10}$										
19											$p_{19,11}$	$p_{19,12}$	$p_{19,13}$	$p_{19,14}$	$p_{19,15}$	$p_{19,16}$				
20																	$p_{20,17}$	$p_{20,18}$	$p_{20,19}$	$p_{20,20}$

Figure 4.2: Probability transition matrix of the unaggregated MMP model for $N = 3$ nodes. The probability $p_{i,j}$ is the probability of transitioning from state i to state j. Grey cells indicate transitions which cannot occur (probability 0).

We start by imposing some conditions on the steady state probabilities and transition probabilities:

1. The transition probabilities $p_{i,j}$ between all states $i, j = 1,...,16$ must remain unchanged. The aggregation of states 17-20 should not affect the dynamics between the remaining unaggregated states.

2. The steady state probabilities π_i for all states $i = 1,...,16$ must remain unchanged. This means that, after aggregation, the fraction of time in which the graph has 0 links, 1 link, or 2 links remains the same, and the fraction of time that the actions of these states are being applied is also the same.

3. From the second condition, it directly follows that the steady state probability of the new state L3 must be equal to $\sum_{i=17}^{20} \pi_i$. States 17-20 encode 3 links, and they will be replaced by a single state L3 which also encodes 3 links. The fraction of time during which the graph has 3 links is therefore unchanged.

The probability transition matrix after aggregating states 17-20 is shown in Figure 4.3. Under the first condition, we cannot change any of the transition probabilities $p_{i,j}$ for $i, j = 1,...,16$. Therefore, the transition probability $p_{i,\mathrm{L3}}$ for $i = 1,...,16$ must be equal to

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	L3
1	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$													
2					$p_{2,5}$	$p_{2,6}$	$p_{2,7}$	$p_{2,8}$	$p_{2,9}$	$p_{2,10}$							
3											$p_{3,11}$	$p_{3,12}$	$p_{3,13}$	$p_{3,14}$	$p_{3,15}$	$p_{3,16}$	
4																	$p_{4,L3}$
5	$p_{5,1}$	$p_{5,2}$	$p_{5,3}$	$p_{5,4}$													
6					$p_{6,5}$	$p_{6,6}$	$p_{6,7}$	$p_{6,8}$	$p_{6,9}$	$p_{6,10}$							
7					$p_{7,5}$	$p_{7,6}$	$p_{7,7}$	$p_{7,8}$	$p_{7,9}$	$p_{7,10}$							
8											$p_{8,11}$	$p_{8,12}$	$p_{8,13}$	$p_{8,14}$	$p_{8,15}$	$p_{8,16}$	
9											$p_{9,11}$	$p_{9,12}$	$p_{9,13}$	$p_{9,14}$	$p_{9,15}$	$p_{9,16}$	
10																	$p_{10,L3}$
11	$p_{11,1}$	$p_{11,2}$	$p_{11,3}$	$p_{11,4}$													
12					$p_{12,5}$	$p_{12,6}$	$p_{12,7}$	$p_{12,8}$	$p_{12,9}$	$p_{12,10}$							
13					$p_{13,5}$	$p_{13,6}$	$p_{13,7}$	$p_{13,8}$	$p_{13,9}$	$p_{13,10}$							
14											$p_{14,11}$	$p_{14,12}$	$p_{14,13}$	$p_{14,14}$	$p_{14,15}$	$p_{14,16}$	
15											$p_{15,11}$	$p_{15,12}$	$p_{15,13}$	$p_{15,14}$	$p_{15,15}$	$p_{15,16}$	
16																	$p_{16,L3}$
L3	$p_{L3,1}$	$p_{L3,2}$	$p_{L3,3}$	$p_{L3,4}$	$p_{L3,5}$	$p_{L3,6}$	$p_{L3,7}$	$p_{L3,8}$	$p_{L3,9}$	$p_{L3,10}$	$p_{L3,11}$	$p_{L3,12}$	$p_{L3,13}$	$p_{L3,14}$	$p_{L3,15}$	$p_{L3,16}$	$p_{L3,L3}$

Figure 4.3: Probability transition matrix of the MMP model for $N = 3$ nodes after aggregating states 17-20. The probability $p_{i,j}$ is the probability of transitioning from state i to state j. Grey cells indicate transitions which cannot occur (probability 0).

$$p_{i,\mathrm{L3}} = \sum_{j=17}^{20} p_{i,j}. \tag{4.4}$$

Out of states 1-16, only states 4, 10, and 16 have possible transitions to states 17-20, so only they will have non-zero transition probabilities to state L3. Since the rows of the probability transition matrix must sum to 1, it is clear that $p_{4,\mathrm{L3}}$, $p_{10,\mathrm{L3}}$, and $p_{16,\mathrm{L3}}$ are all equal to 1.

The transition probabilities $p_{\mathrm{L3},i}$ still need to be calculated. Intuitively, it can be understood that in order to preserve the dynamics of the remaining states 1-16, state L3 should fairly "combine" the effects of state 17-20. Therefore, the fraction of time L3 should "behave" as each of the original unaggregated states 17-20 should be proportional to the steady state probability of the original unaggregated state. We can calculate this fraction by dividing the steady state probability of the original unaggregated state by the steady state probability of L3; for example, the fraction of time that L3 should "behave" as state 17 is given by $\pi_{17}/\pi_{\mathrm{L3}}$. Therefore, to calculate the transition probability $p_{\mathrm{L3},i}$, we should take the weighted sum of $p_{17,i}$, $p_{18,i}$, $p_{19,i}$, and $p_{20,i}$, using the fraction of time that L3 should "behave" as each of the states 17-20 as the corresponding weight. We can express this as

$$p_{\mathrm{L3},i} = \frac{\pi_{17}}{\pi_{\mathrm{L3}}} p_{17,i} + \frac{\pi_{18}}{\pi_{\mathrm{L3}}} p_{18,i} + \frac{\pi_{19}}{\pi_{\mathrm{L3}}} p_{19,i} + \frac{\pi_{20}}{\pi_{\mathrm{L3}}} p_{20,i}, \tag{4.5}$$

for $i = 1,...,16$. From the probability transition matrix of the unaggregated MMP model in Figure 4.2, we observe that there is no overlap between the states which states 17-20 can transition to: state 17 can transition to states 1-4, state 18 can transition to states 5-10, state 19 can transition to states 11-16, and state 20 can transition to states 17-20. The transition probabilities can therefore be simplified depending on the target state i. For $i = 1,...,4$, we find

$$p_{\mathrm{L3},i} = \frac{\pi_{17}}{\pi_{\mathrm{L3}}} p_{17,i}. \tag{4.6}$$

For $i = 5,...,10$, we find

$$p_{\mathrm{L3},i} = \frac{\pi_{18}}{\pi_{\mathrm{L3}}} p_{18,i}. \tag{4.7}$$

For $i = 11,...,16$, we find

$$p_{\text{L3},i} = \frac{\pi_{19}}{\pi_{\text{L3}}} p_{19,i}. \tag{4.8}$$

Finally, the transition probability $p_{\text{L3,L3}}$ follows from the condition that the row of the probability transition matrix must sum to 1. We start by setting the sum of the final row to 1 and splitting index of the summation into three parts:

$$\begin{aligned} \sum_{i=1}^{16} p_{\text{L3},i} + p_{\text{L3,L3}} &= 1 \\ \sum_{i=1}^{4} p_{\text{L3},i} + \sum_{i=5}^{10} p_{\text{L3},i} + \sum_{i=11}^{16} p_{\text{L3},i} + p_{\text{L3,L3}} &= 1. \end{aligned} \tag{4.9}$$

We can then substitute (4.6), (4.7), and (4.8) into (4.9) and simplify:

$$\begin{aligned} \sum_{i=1}^{4} \frac{\pi_{17}}{\pi_{\text{L3}}} p_{17,i} + \sum_{i=5}^{10} \frac{\pi_{18}}{\pi_{\text{L3}}} p_{18,i} + \sum_{i=11}^{16} \frac{\pi_{19}}{\pi_{\text{L3}}} p_{19,i} + p_{\text{L3,L3}} &= 1 \\ \frac{\pi_{17}}{\pi_{\text{L3}}} + \frac{\pi_{18}}{\pi_{\text{L3}}} + \frac{\pi_{19}}{\pi_{\text{L3}}} + p_{\text{L3,L3}} &= 1. \end{aligned} \tag{4.10}$$

Since we know that $\pi_{\text{L3}} = \sum_{i=17}^{20} \pi_i$, it follows that

$$p_{\text{L3,L3}} = \frac{\pi_{20}}{\pi_{\text{L3}}}. \tag{4.11}$$

In Section 4.3.3, we will prove for the general case that choosing the transition probabilities in this manner will keep the steady state probabilities of the unaggregated states the same, that is π_i for all states i which are not part of the aggregation will remain the same. Additionally, we show that the transition probabilities between the unaggregated states will also remain the same, that is $p_{i,j}$ for all states i and j which are not part of the aggregation will remain the same. For the moment, we will take this result for granted and move on to define the actions.

Defining Action Sets

The new aggregated state L3 represents the state of the graph having 3 links, but it does not have any action associated with it. In the aggregated MMP model, we instead define *action sets* $J_{i,j}$ which are associated with a transition from i links to j links. The action set is the set of all allowable actions that result in the transition from i to j links, and each action set has its own probability distribution. When we transition from i links to j links, we apply one of the actions from the corresponding action set probabilistically. For example, consider the transition from 3 links to 3 links. The only allowable action which results in this transition is to add 0 and remove 0 links; we cannot add anymore links because 3 links is the complete graph, and we cannot remove any links or else we will have fewer than 3 links. The action set $J_{3,3}$ is therefore $\{(+0,-0)\}$, and since there is only one action, the probability distribution is simply $\Pr[(+0,-0)] = 1$.

After the first aggregation step, we define action sets for transition from 3 links to i links where $i = 0,...,3$.

- $J_{3,0} = \{(+0,-3)\}$
- $J_{3,1} = \{(+0,-2)\}$
- $J_{3,2} = \{(+0,-1)\}$
- $J_{3,3} = \{(+0,-0)\}$

In this case, since each of the action sets only has one possible action, the probability distribution of each action set is trivial; there is only one action so its probability is 1. When the Markov process transitions from state L3 to states 1-4 (i.e. from 3 links to 0 links), the action from the action set $J_{3,0}$ is applied. When the Markov process transitions from state L3 to states 5-10 (i.e. from 3 links to 1 links), the action from the action set $J_{3,1}$ is applied. When the Markov process transitions from state L3 to states 11-16 (i.e. from 3 links to 2 links), the action from the action set $J_{3,2}$ is applied. When the Markov process transitions from state L3 to state L3 (i.e. from 3 links to 3 links), the action from the action set $J_{3,3}$ is applied.

To give an example of a state transition which will have an action set with multiple actions, let us consider the transition from 1 link to 1 link. When there is 1 link in the graph, there are only two possible actions which can result in a transition to 1 link: $(+0,-0)$ and $(+1,-1)$, which are encoded by state 6 and state 7 respectively

in the unaggregated MMP model. Therefore, the action set $J_{1,1}$ is $(+0,-0),(+1,-1)$.[2] Since we have multiple actions, the probability distribution of the action set is non-trivial. The probability distribution of the action set is dictated by the steady state probability of the states corresponding to the actions; the probability of each action is proportional to the steady state probability of its corresponding state:

$$\begin{aligned}\Pr[(+0,-0)] &= \frac{\pi_6}{\pi_6+\pi_7},\\ \Pr[(+1,-1)] &= \frac{\pi_7}{\pi_6+\pi_7}.\end{aligned} \tag{4.12}$$

We will demonstrate in Section 4.3.4 that if we choose the probability distribution of the action set in this manner, the rate at which the actions are applied in the aggregated MMP model will be the same as in the unaggregated MMP model. In the next section, we will first prove that the method for state aggregation that we have just described will not affect the steady state probabilities of the unaggregated stats. Additionally, we will prove that the transition probabilities between the unaggregated states will not be affected either. The proof can then be applied to finish creating the aggregated MMP model for $N = 3$ nodes.

4.3.3. State Aggregation by Matching Probability Flux

In this section, we prove that given a Markov process with a set of states S, an arbitrary subset of states $M \subset S$ can be aggregated into a single state m such that both the transition probabilities between the states not in M, and the steady state probabilities of the states not in M, remain unchanged. The latter also implies that the steady state probability of the new state m is equal to the sum of the steady state probabilities of all states in M.

The steady state vector of the discrete time Markov process satisfies

$$P^T \pi^T = \pi^T. \tag{4.13}$$

We refer to [5, pp. 191-192] for a proof of this. From (4.13), it directly follows that

$$\pi_i = \sum_{j\in S} \pi_j p_{j,i}, \tag{4.14}$$

for all $i \in S$, where S is the set of states in the Markov process. Since $\sum_{j\in S} p_{i,j} = 1$, (4.14) can also be rewritten as

$$\pi_i \sum_{j\in S} p_{i,j} = \sum_{j\in S} \pi_j p_{j,i}, \tag{4.15}$$

for all $i \in S$. Equation (4.15) has the following interpretation. We define $\pi_i p_{i,j}$ as the probability flux from state i to state j. The left-hand side of (4.15) is the sum of the outward probability flux from state i going to all other states (including state i). The right-hand side of (4.15) is the sum of the inward probability flux from all states (including state i) and going to state i. The steady state equations require that the outward and inward probability fluxes be equal.

Let us now consider a strict subset of states $M \subset S$. The balance equation (4.15) can be rewritten as

$$\pi_i \sum_{j\in M} p_{i,j} + \pi_i \sum_{j\in S\setminus M} p_{i,j} = \sum_{j\in M} \pi_j p_{j,i} + \sum_{j\in S\setminus M} \pi_j p_{j,i}, \tag{4.16}$$

for all $i \in S$. On the left-hand side of (4.16), the sum of outward probability fluxes has been split into the sum of the probability fluxes going to states in M and states which are not in M (i.e. the rest of states). On the right-hand side of (4.16), the same has been done for the inward probability fluxes.

Suppose we now want to merge the states in M into a single state m such that the steady state probabilities π_i for $i \in S \setminus M$ remain unchanged. We also want the transition probabilities $p_{i,j}$ to be unchanged for all $i, j \in S \setminus M$. State m replaces all of the states in the set M, and since the steady state probability must sum to 1, this implies that the steady state probability of the new state m is given by

$$\pi_m = 1 - \sum_{i\in S\setminus M} \pi_i = \sum_{i\in M} \pi_i. \tag{4.17}$$

[2] Notice that the $(+2,-2)$ is not included: this is not a valid action because it would require us to remove more links than currently exist in the graph (in fact, $(+2,-2)$ is never a valid action for a graph with 3 nodes because it requires a total of 4 links to be changed, but only 3 links can exist in the graph).

For the states which are not in M, the new balance equations can be written as

$$\pi_i p_{i,m} + \pi_i \sum_{j \in S \setminus M} p_{i,j} = \pi_m p_{m,i} + \sum_{j \in S \setminus M} \pi_j p_{j,i}, \tag{4.18}$$

for all $i \in S \setminus M$. For the new state m, the balance equation is

$$\pi_m \sum_{i \in S \setminus M} p_{m,i} + \pi_m p_{m,m} = \sum_{i \in S \setminus M} \pi_i p_{i,m} + \pi_m p_{m,m}. \tag{4.19}$$

We require the steady state probabilities π_i for $i \in S \setminus M$ to be unchanged, and also the transition probabilities p_j for all $i, j \in S \setminus M$ to be unchanged. It thus follows from (4.16) and (4.18) that

$$\pi_i p_{i,m} = \pi_i \sum_{j \in M} p_{i,j}. \tag{4.20}$$

and

$$\pi_m p_{m,i} = \sum_{j \in M} \pi_j p_{j,i}. \tag{4.21}$$

Equation (4.20) reads that the outward flux from a state i to the new state m is equal to the sum of the outward fluxes to all of the states in M (which are now merged into the new state m). Equation (4.21) reads that the inward flux to a state i from the new state m is equal to the sum of the inward fluxes from all of the states in M.

From (4.20), the transition probabilities to state m immediately follow as

$$p_{i,m} = \sum_{j \in M} p_{i,j}, \tag{4.22}$$

for all $i \in S \setminus M$. From (4.21), we have the following transition probabilities from state m to the other states:

$$p_{m,i} = \frac{1}{\pi_m} \sum_{j \in M} \pi_j p_{j,i}, \tag{4.23}$$

for all $i \in S \setminus M$. Moreover, the transition probability $p_{m,m}$ follows from the condition that

$$p_{m,m} + \sum_{i \in S \setminus M} p_{m,i} = 1. \tag{4.24}$$

Finally, we can verify that the balance equation for the new state m is satisfied. Simplifying the term $\pi_m p_{m,m}$ from both sides of (4.19), we have

$$\pi_m \sum_{i \in S \setminus M} p_{m,i} = \sum_{i \in S \setminus M} \pi_i p_{i,m}. \tag{4.25}$$

From (4.22) and (4.23), we have expressions for $p_{i,m}$ and $p_{m,i}$ respectively. Substituting them into (4.25) we find

$$\pi_m \sum_{i \in S \setminus M} \frac{1}{\pi_m} \sum_{j \in M} \pi_j p_{j,i} = \sum_{i \in S \setminus M} \pi_i \sum_{j \in M} p_{i,j}. \tag{4.26}$$

The term π_m and $\frac{1}{\pi_m}$ on the left-hand side cancel out leaving us with

$$\sum_{i \in S \setminus M} \sum_{j \in M} \pi_j p_{j,i} = \sum_{i \in S \setminus M} \pi_i \sum_{j \in M} p_{i,j}. \tag{4.27}$$

Equation (4.27) reads that the sum of the inward flux to the states $i \in S \setminus M$ from the states $j \in M$ is equal to the sum of the outward flux to the states $j \in M$ from the states $i \in S \setminus M$. The set of all states $i \in S \setminus M$ is the complement of M. It has been shown by [6, pp. 15] that in the steady state of a Markov chain, the outward flux from any set of states A to its complement A^C is equal to the inward flux to A from A^C. Therefore, we have verified that the balance equation for the new state m is also satisfied.

4.3.4. Aggregated Link-Aware MMP Model for N=3

Using the result of Section 4.3.3, the MMP model for $N = 3$ has been aggregated into four states, with each state encoding a different number of links. The states in the aggregated model are numbered starting from zero so that the state number corresponds to the number of links in the graph, i.e. state L0 encodes 0 links, state L1 encodes 1 link, etc. The probability transition matrix of the aggregated MMP model is given by

$$P = \begin{bmatrix} \frac{\pi_1}{\pi_{\text{L0}}} & \frac{\pi_2}{\pi_{\text{L0}}} & \frac{\pi_3}{\pi_{\text{L0}}} & \frac{\pi_4}{\pi_{\text{L0}}} \\ \frac{\pi_5}{\pi_{\text{L1}}} & \frac{\pi_6+\pi_7}{\pi_{\text{L1}}} & \frac{\pi_8+\pi_9}{\pi_{\text{L1}}} & \frac{\pi_{10}}{\pi_{\text{L1}}} \\ \frac{\pi_{11}}{\pi_{\text{L2}}} & \frac{\pi_{12}+\pi_{13}}{\pi_{\text{L2}}} & \frac{\pi_{14}+\pi_{15}}{\pi_{\text{L2}}} & \frac{\pi_{16}}{\pi_{\text{L2}}} \\ \frac{\pi_{17}}{\pi_{\text{L3}}} & \frac{\pi_{18}}{\pi_{\text{L3}}} & \frac{\pi_{19}}{\pi_{\text{L3}}} & \frac{\pi_{20}}{\pi_{\text{L3}}} \end{bmatrix}, \tag{4.28}$$

where the steady state probabilities of the aggregated model are given by

$$\begin{aligned} \pi_{\text{L0}} &= \sum_{i=1}^{4} \pi_i, \\ \pi_{\text{L1}} &= \sum_{i=5}^{10} \pi_i, \\ \pi_{\text{L2}} &= \sum_{i=11}^{16} \pi_i, \\ \pi_{\text{L3}} &= \sum_{i=17}^{20} \pi_i. \end{aligned} \tag{4.29}$$

The derivation can be found in Appendix B. The average number of links in the aggregated MMP model is given by

$$L_{\text{avg}} = (0 \times \pi_{\text{L0}}) + (1 \times \pi_{\text{L1}}) + (2 \times \pi_{\text{L2}}) + (3 \times \pi_{\text{L3}}). \tag{4.30}$$

Substituting the steady state values in (4.29) into (4.30), we find

$$L_{\text{avg}} = \left(0 \times \sum_{i=1}^{4} \pi_i\right) + \left(1 \times \sum_{i=5}^{10} \pi_i\right) + \left(2 \times \sum_{i=11}^{16} \pi_i\right) + \left(3 \times \sum_{i=17}^{20} \pi_i\right). \tag{4.31}$$

We observe that (4.31) is equal to (4.3). Therefore, the average number of links in the modulated graph produced by the aggregated MMP model is the same as the unaggregated MMP model.

Each transition from state i to state j in the aggregated MMP model has an associated action set. Each action set has its own probability distribution, and when the Markov process transitions from state i to state j, one of the actions from the action set is probabilistically applied to modulate the graph. The action sets for all 16 possible transitions are as follows:

- $J_{0,0} = \{(+0,-0)\}$
- $J_{0,1} = \{(+1,-0)\}$
- $J_{0,2} = \{(+2,-0)\}$
- $J_{0,3} = \{(+3,-0)\}$
- $J_{1,0} = \{(+0,-1)\}$
- $J_{1,1} = \{(+0,-0),(+1,-1)\}$
- $J_{1,2} = \{(+1,-0),(+2,-1)\}$
- $J_{1,3} = \{(+2,-0)\}$
- $J_{2,0} = \{(+0,-2)\}$
- $J_{2,1} = \{(+0,-1),(+1,-2)\}$
- $J_{2,2} = \{(+0,-0),(+1,-1)\}$
- $J_{2,3} = \{(+1,-0)\}$
- $J_{3,0} = \{(+0,-3)\}$
- $J_{3,1} = \{(+0,-2)\}$
- $J_{3,2} = \{(+0,-1)\}$
- $J_{3,3} = \{(+0,-0)\}$

Many of the action sets contain only one possible action. In this case, the probability distribution of the action set is trivial: with only one action, the probability of that action being applied is 1. For the action sets which contain multiple actions, the probability distribution of the actions is weighted according to the steady

state probabilities of the original states corresponding to those actions before the aggregation. To revisit the example, action set $J_{1,1}$ contains $(+0,-0)$ and $(+1,-1)$ as actions. The action $(+0,-0)$ given there is currently one link in the graph corresponds to state 6, and the action $(+1,-1)$ given there is currently one link in the graph corresponds to state 7. The probability distribution of action set $J_{1,1}$ is therefore

$$\begin{aligned} \Pr[(+0,-0)] &= \frac{\pi_6}{\pi_6+\pi_7}, \\ \Pr[(+1,-1)] &= \frac{\pi_7}{\pi_6+\pi_7}. \end{aligned} \tag{4.32}$$

The steady state probability of L1 in the aggregated MMP model is π_{L1}. From the probability transition matrix of the aggregated MMP model, the probability of transition from state L1 to L1 is given by $(\pi_6+\pi_7)/\pi_{\text{L1}}$. When transitioning from state L1 to L1, the action set $J_{1,1}$ is applied. The probability of applying action $(+0,-0)$ is given by $\frac{\pi_6}{\pi_6+\pi_7}$. Therefore, in the aggregated MMP, the long-term probability of applying action $(+0,-0)$ when there is 1 link in the graph is given by

$$\pi_{\text{L1}} \frac{\pi_6+\pi_7}{\pi_{\text{L1}}} \frac{\pi_6}{\pi_6+\pi_7} = \pi_6. \tag{4.33}$$

For the action $(+1,-1)$, we find

$$\pi_{\text{L1}} \frac{\pi_6+\pi_7}{\pi_{\text{L1}}} \frac{\pi_7}{\pi_6+\pi_7} = \pi_7. \tag{4.34}$$

Therefore, in the aggregated MMP model, the rate at which actions are being applied is the same as in the unaggregated MMP model.

4.4. Discussion

In this chapter, we have introduced the concept of a link-aware MMP model in which the states of the Markov process encode the number of links in the graph. In the unaggregated link-aware MMP model, each state of the model corresponds to one action. The main drawback of the unaggregated model is that it requires a massive number of states which increases with $O(N^6)$. We therefore introduced the concept of state aggregation, in which states of the Markov process are aggregated based on the number of links encoded by the state. In the aggregated link-aware MMP model, the required number of states is bounded by the maximum number of links plus one, i.e. $L_{\max}+1 = N(N-1)/2+1$. Compared to the unaggregated MMP model, there is a huge reduction in the required number of states: the number of states in the aggregated MMP model grows with $O(N^2)$.

In the aggregated MMP model, the actions have been decoupled from the states and are instead tied to state transitions; an action set is defined for every possible transition of the Markov process. The action set $J_{i,j}$ for a transition from state i to state j is defined as the set of all possible actions, given that i links exist in the graph, that result in a transition to j links. The action set is bounded because we cannot remove more links than currently exist, and we cannot add more links than the complete graph. Each action set has a probability distribution, and each time the Markov process transitions from state i to state j, one of the actions from the corresponding action set $J_{i,j}$ is randomly applied based on the probability distribution of the action set.

The aggregated MMP model will produce a modulated graph which has the same average number of links as the unaggregated MMP model. Furthermore, the long-term rate at which actions are applied in the aggregated MMP model is the same as in the unaggregated MMP model. The key difference between the unaggregated and aggregated MMP model lies in the dependence between consecutive actions. In the unaggregated MMP model, each state encodes an action as well as a number of links. The next action is therefore dependent on the previous action as well as the previous number of links. Equivalently, we can interpret the next action as being dependent on the previous action as well as the *current* number of links, because the current number of links in the graph is fully determined by the the previous action and the previous number of links. This definition is more intuitive because the next action that we can apply is limited by the number of links which currently exist in the graph.

However, in the aggregated MMP model, each state encodes only the number of links, and the actions are chosen based on the probability distribution of the action set. The Markov states do not model actions, and the probability distribution of each action set is fixed (and therefore also independent of the previous action). Thus, the next action in the aggregated MMP model is only a function of the current number of links; the dependence on the previous action has been lost as a result of the aggregation. In the next chapter, we

will investigate the effects of the state aggregation by comparing the modeling performance of the aggregated and unaggregated MMP model.

5

Comparison of the Aggregated and Unaggregated MMP Model

5.1. Introduction

In the previous chapter, we developed MMP models which encode the number of links in the graph. In the unaggregated MMP model, each state encodes the number of links in the graph and an action to apply on the graph. All possible numbers of links and all possible actions are considered, resulting in a Markov process with a number of states equal to

$$\#X = \sum_{L=0}^{L_{\max}} (L+1)(L_{\max} - L + 1), \tag{5.1}$$

where $L_{\max} = N(N-1)/2$ and N is the number of nodes in the graph. In the aggregated MMP model, each state only encodes the number of links in the graph. The number of states is thus reduced to $L_{\max}+1$. Each state transition of the aggregated MMP model has an associated action set with a probability distribution, and an action is randomly chosen from the corresponding action set when a state transition occurs.

In the unaggregated MMP model, the number of states increases with $O(N^6)$, and the model is no longer feasible for large N. In the original mobility process simulation, the number of nodes was $N = 25$ which would result in 4,590,551 states for the unaggregated MMP model. We therefore perform a new mobility process simulation with 9 nodes. For a mobility process with 9 nodes, the unaggregated MMP model has 9,139 states which is still tractable. The aggregated MMP model has 37 states. In this chapter, we will simulate a mobility process with 9 nodes and model the mobility process using the aggregated MMP model as well as the unaggregated MMP model. We will compare the results of the two models to investigate the effect of the state aggregation on the resulting model. The results of this chapter show that the aggregated MMP model produces results which are indistinguishable from the unaggregated MMP model, and the state aggregation does not cause any observable loss in modeling performance.

5.2. Simulated Mobility Process for N=9 Nodes

5.2.1. Modified Simulation Parameters

We choose a 9 node mobility process because it is the largest square number for which the number of states in the unaggregated MMP model is still manageable for us. With 9 nodes, the number of states in the unaggregated MMP model is 9,139; for 16 nodes, the number of states would be 302,621. The mobility process simulation is based on the same mobility process described in Chapter 2 with some modified simulation process. Since the size of starting square lattice is now 3×3 instead of 5×5, the size of the region is also reduced from 10×10 to 6.66×6.66. This keeps the spacing between the nodes in the starting lattice the same as illustrated in Figure 5.1.

The remaining simulation parameters are unchanged. At the $t = 0$, each node is initialized with a random direction in the range $[0,359]$ representing an angle in degrees. At each time step, nodes can change their current direction by a random offset uniformly distributed in the range $[-20,20]$, where the offset is an integer angle expressed in degrees. After applying the angle offset, the node can then move linearly along its current

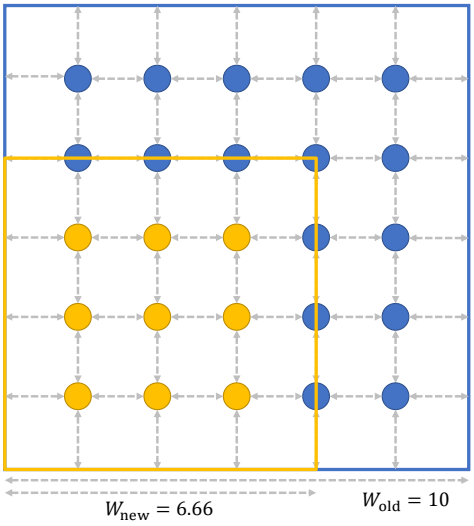


Figure 5.1: Initial positions of nodes for $N = 9$. The yellow circles and yellow border show initial positions of the nodes and the bounds of the region. The blue circles and the blue border show the original positions and bounds in the 25 node mobility process.

direction. The move distance is uniformly distributed in the range $[0, 2v]$ with $v = 1$, and the move distance does not have to be an integer. The nodes are constrained to move within the region, and out of bounds handling is performed using the process described in Section 2.4. Each realization of the mobility process is simulated for $T = 1{,}000$ time steps. At each time step, the adjacency matrix of the graph is calculated; two nodes are considered to be connected when the distance between them is less than or equal to $d = 1.5$ units. Due to the fact that the unaggregated MMP model has a very large number of states, the number of realizations of the mobility process is increased to provide more data to use to generate the transition matrix. The number of trials is increased from 100 to 10,000; this results in about 10 million data points that can be used to generate the probability transition matrix.

5.2.2. Simulation Results

After performing 10,000 trials of 1,000 time steps each, we analyze the results of the simulated mobility process with 9 nodes. To begin with, the same link metrics are considered, namely the total number of links at each time step and the number of links added and removed between two time steps. The plots of distributions of these values can be found in Appendix C.1 and the averages are as follows:

- Average number of links at each time step: 4.834.
- Average number of links added between two time steps: 2.491.
- Average number of links removed between two time steps: 2.487.

5.3. Constructing the MMP Models

5.3.1. Constructing the Unaggregated MMP Model

In the unaggregated MMP model, each state encodes the number of links as well as an action. In order to construct the probability transition matrix for the unaggegated MMP model, we first need to convert the sequences of adjacency matrices into sequences of states in the MMP model. An illustration of how this is done for a single realization of the mobility process is shown in Figure 5.2.

The blue squares represent the sequence of adjacency matrices. Each realization has 1,001 adjacency matrices, but for illustration purposes only 4 are shown. The first step is to calculate the number of links added and removed between each pair of adjacency matrices, which will be the action in the MMP model. Next, we calculate the number of links at each time step. Finally, we can assign a state to each unique combination of links and actions. Recall that in the unaggregated MMP model, the state number has no special meaning and

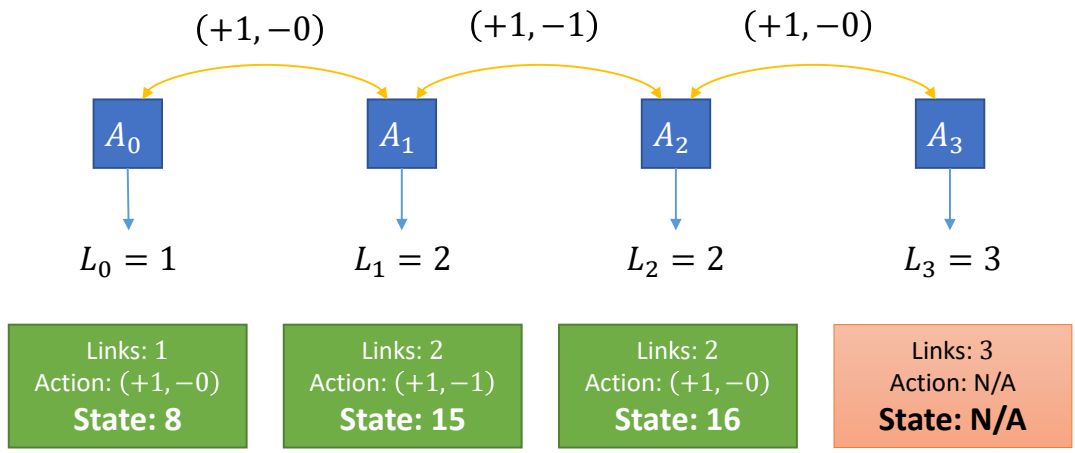


Figure 5.2: Illustration of the steps for mapping the adjacency matrices to states in the unaggregated MMP model.

serves only as an identifier to uniquely identify the number of links and actions. In Figure 5.2, we are using the state definitions from the unaggregated MMP model for $N = 3$ that was discussed in Section 4.3.1. Note that we cannot assign a state for the last adjacency matrix because there is no action after the last time step. Through this process, each sequence of 1,001 adjacency matrices is transformed into a sequence of 1,000 states of the unaggregated MMP model.

The probability transition matrix can then be derived from from the sequence of states using the same method described in Section 3.4.2. For estimating the transition probability from state i to state j, we count the number of times that we observe a transition from state i to state j and divide it by the total number of observed transitions from state i to any state k. We can express this as

$$p_{i,j} = \frac{n_{i,j}}{\sum_{k \in X} n_{i,k}}, \tag{5.2}$$

where $p_{i,j}$ is the probability of going from state i to state j, $n_{i,j}$ is the number of times that the transition from state i to state j is observed, and X is set of all possible states.

Using all 10,000 trials of the mobility process, the probability transition matrix is calculated. Each trial is a series of 1,001 adjacency matrices which is mapped to 1,000 states, resulting in 999 state transitions. Therefore, there are 9,990,000 observed state transitions which are used to construct the probability transition matrix. In principle, the unaggregated MMP model for 9 nodes consists of 9,139 states, but 6,216 states were not observed in the simulated mobility process. The unaggregated MMP model constructed from the mobility process can therefore be reduced to 2,923 states.

5.3.2. Constructing the Aggregated MMP Model

In the aggregated MMP model, the states only encode the number of links. The sequence of adjacency matrices can be directly transformed into a sequence of states by counting the number of links in each adjacency matrix (i.e. summing all elements of the adjacency matrix and dividing by 2). From the sequence of states, the probability transition matrix can be constructed using the same method described in the previous section; each transition probability can be calculated using (5.2).

For the aggregated MMP models, action sets also need to be defined for each state transition. An illustration of how the a single action set is derived is shown in Figure 5.3. In the illustrated example, we are interested in finding the action set $J_{1,2}$, which is the action set when we transition from 1 link to 2 links. To find an arbitrary action set $J_{i,j}$, the first the first step is to calculate the number of links at each time step in the series of adjacency matrices and find all instances where a transition from i links to j links is observed. For each transition from i links to j links, the action that was applied can be calculated by comparing the adjacency matrices at the two time steps. The set of all unique actions that are observed forms the action set $J_{i,j}$, and the probability distribution of $J_{i,j}$ is calculated by counting the number of times each action is observed and normalizing over the total number of observed transitions from i links to j links.

In the illustrated example, two unique actions were observed: $(+1, -0)$ and $(+2, -1)$. The action set is therefore $J_{1,2} = \{(+1, -0), (+2, -1)\}$. The action $(+1, -0)$ was observed twice, and in total, we observed three

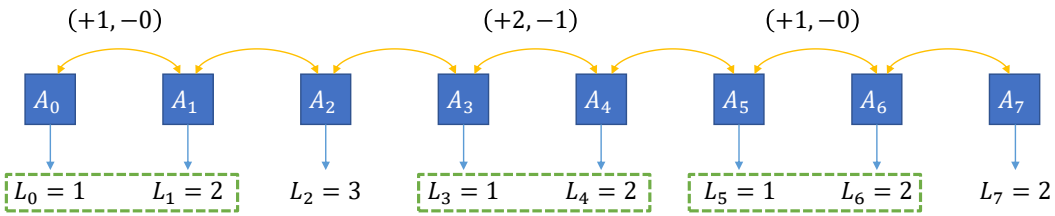


Figure 5.3: Illustration of the steps for constructing the action set in the aggregated MMP model. In this example, the action set we are interested in is $J_{1,2}$.

transitions from 1 link to 2 links. Therefore, the probability of applying the action $(+1,-0)$ is $2/3$. The action (+2, -1) was also observed once, and therefore its probability is $1/3$. In this example, there is only one realization of the mobility process; when constructing the action sets of the aggregated MMP model, all 10,000 trials are used. Each trial of the mobility process is a series 1,001 adjacency matrices which are directly mapped to states by calculating the number of links in each matrix. In a series of 1,001 states, there are 1,000 state transitions. Across all trials, there is thus a total of 10,000,000 state transitions which are used to construct the probability transition matrix.

In principle, the aggregated MMP model for 9 nodes has 37 states. The states encode the number of links in the graph and the maximum number of links is $9 \times (9-1)/2 = 36$, and there is one additional state for zero links. In the aggregated MMP model constructed from the mobility process, 4 states were not observed, so there are only 33 states.[1] With 33 states, the total number of possible state transitions is $33^2 = 1,089$ and each transition has its own action set. However, in the constructed MMP model, 463 state transitions were not observed, leaving us with 626 action sets.

5.3.3. MMP Simulation Methodology

Simulation Procedure for the Unaggregated MMP Model

The starting state of the mobility process simulations has 0 links, and therefore it makes sense to initialize the MMP models in a state with 0 links. In the unaggregated MMP model, each state encodes the current number of links in the graph as well as an action to apply. Therefore, there are multiple states corresponding to 0 links but with different actions. The starting state of the unaggregated MMP model is chosen randomly from the set of states which encode 0 links. The probability of each state being chosen is proportional to the steady state probability of that state.

In total, 100 trials will be performed and each simulation will run for 1,000 time steps. In the unaggregated MMP model, actions are applied at the start of the following time step. To illustrate with an example, suppose that the unaggregated MMP model is initialized in some state x_0 which represents 0 links and the action $(+1,-0)$. At time $t = 0$, the Markov process is in some state x_0 and the graph is empty. At time $t = 1$, the action $(+1,-0)$ is first applied to the graph, and the graph now has 1 link. The Markov process then transitions to some other state x_1, thereby choosing the next action which will be applied at time $t = 2$. The method for adding and removing links remains the same as in the first MMP model: the links to be added are chosen randomly from the set of all available links, while the links to removed are chosen randomly from the set of all links which exist in the graph (excluding the links that are added in the current time step).

Simulation Procedure for the Aggregated MMP Model

The aggregated MMP model is also initialized with 0 links. Since each state of the aggregated MMP model corresponds to a different number of links, the Markov process is simply initialized in state 0 which represents 0 links. At each time step, the Markov process transitions to the next state. One of the actions from the action set associated with the state transition is then applied immediately to modulate the graph.

To illustrate with an example, suppose that at time $t = 0$, the Markov process is in state 0, i.e. there are 0 links in the graph. At time $t = 1$, the Markov process transitions to the next state, and state 2 is chosen.

[1] These unobserved states are not actually removed from the probability transition matrix, but they will have no effect on the simulation because these states will never be entered. This is done because in the aggregated MMP model, the state number corresponds to the number of links. Removing these unused states will require the remaining states to be re-indexed and the one-to-one correspondence between state number and links would be lost, making it more difficult to process the simulation results.

One of the actions from the action set $J_{0,2}$ is then randomly selected based on the probability distribution of the action set, and the chosen action is applied to modulate the graph. The action set $J_{0,2}$ only contains one action which is $(+2,-0)$, because the only possible way to go from 0 links to 2 links is to add 2 links and remove 0 links. The probability distribution of the action set is $\Pr[(+2,-0)] = 1$, since there is only one action and the probability distribution must sum to 1. Therefore, the action (+2,-0) is chosen and is applied to the graph immediately. After applying the action, there will be 2 links in the graph, which corresponds to the state that was chosen (state 2).

In total, 100 trials of the aggregated MMP model will also be performed, and each simulation will be run for 1,000 time steps. The method for choosing which links to add and remove is the same as in the unaggregated MMP model.

5.4. Comparison of the Unaggregated and Aggregated MMP Models

5.4.1. Link Distributions

After performing 100 trials of 1,000 time steps for both the aggregated and unaggregated MMP models, the same link metrics are calculated for each MMP model. The distributions of the total number of links at each time step and number of links added and removed between two time steps for the unaggregated MMP model can be found in Appendix C.2. The averages are as follows:

- Average number of links at each time step: 4.850.
- Average number of links added between two time steps: 2.502.
- Average number of links removed between two time steps: 2.497.

The distributions of the total number of links at each time step and the number of links added and removed between two time steps for the aggregated MMP model can be found in Appendix C.3. The averages are as follows:

- Average number of links at each time step: 4.815.
- Average number of links added between two time steps: 2.486.
- Average number of links removed between two time steps: 2.481.

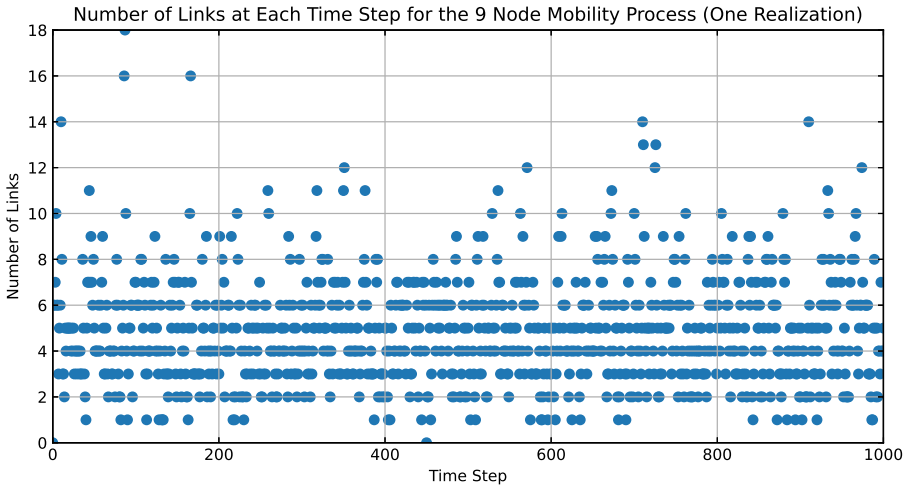
The distributions of the total number of links and the number of links added and removed in both MMP models closely match the real mobility process. The mean values of the total number of links and the number of links added and removed in both MMP models differ from the mean values of the mobility process by less than 1%. A plot of the number of links at each time step for a single realization for each MMP model as well as the mobility process is shown in Figure 5.4. We observe a similar behavior in all three cases, the number of links fluctuates around an average of about 5 links and we do not observe any kind of drift in the average number of links.

5.4.2. K-Step Link Retention Probability

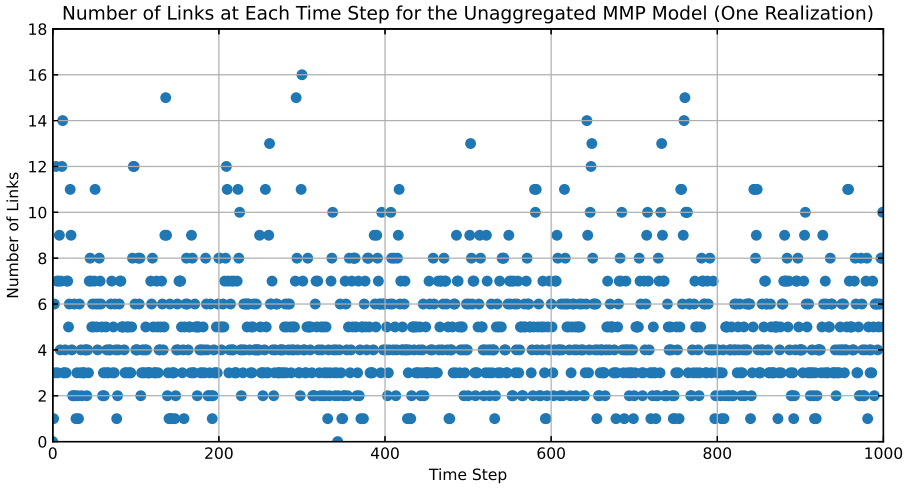
Both MMP models accurately model the average number of links in the graph, and the average number of links added and removed between two time steps is also accurately modeled. This means that on average, the graphs produced by the MMP models have the same amount of links as the mobility process, and the graphs evolve at the same rate. However, one aspect of the mobility process that the MMP models cannot capture is the *selection* of which links to remove. In the mobility process, a link is removed when nodes move away from each other and the distance exceeds $d = 1.5$ units. In the MMP models, there is no information about the positions of nodes; therefore, the selection of links to remove is completely random. To investigate the effect that this may have on the resulting graphs, we define a metric called the *K-step link retention probability*.

The K-step link retention probability is defined as the probability that a link which exists at time t will still exist at each subsequent time step up to and including $t+K$. So the 2-step link retention probability is the probability that given a link exists at time t, it also exists at time $t+1$ and $t+2$. The K-step link retention probability for 1 to 10 steps has been plotted in Figure 5.5; the probabilities are averaged over all realizations of the MMP model and mobility process.

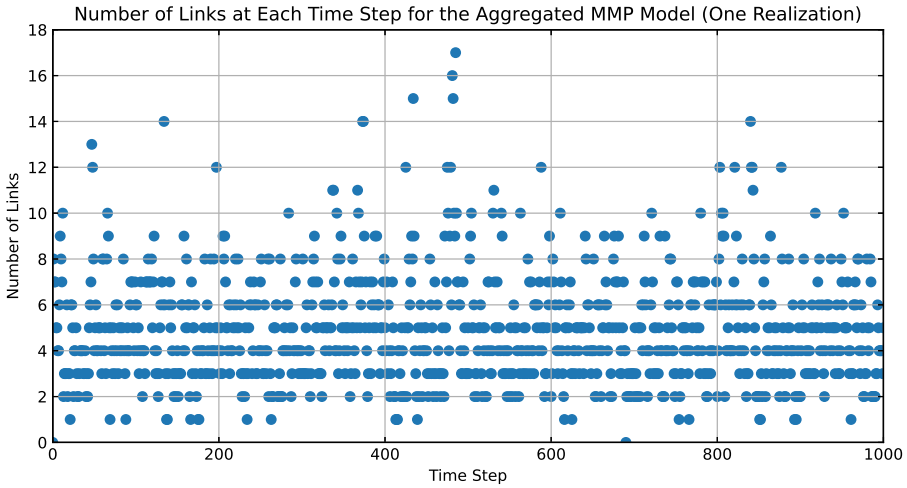
The K-step link retention probability for both the unaggregated and aggregated MMP model is almost exactly the same, which is to be expected because they use the same strategy for removing links. For both



(a) Number of links at each time step for one realization of the mobility process.



(b) Number of links at each time step for one realization of the unaggregated MMP model.



(c) Number of links at each time step for one realization of the aggregated MMP model.

Figure 5.4: Comparison of the number of links at each time step in the MMP modulated graphs and the simulated mobility process.

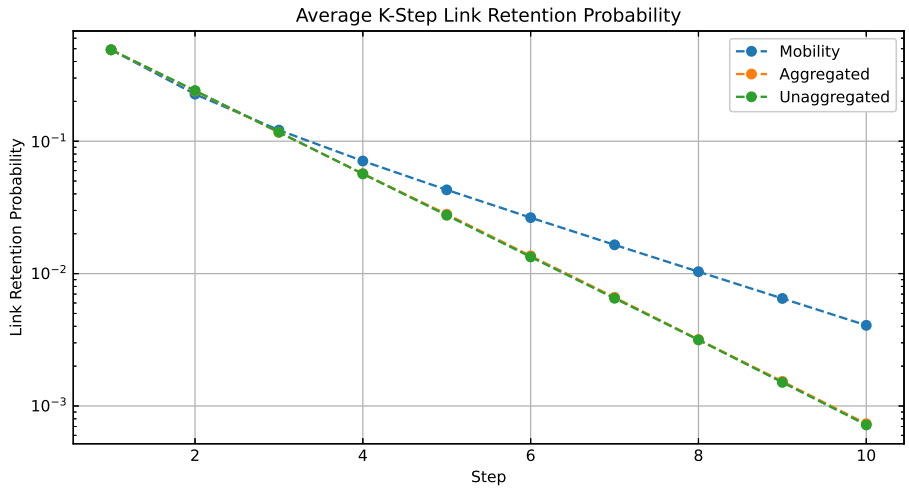


Figure 5.5: The K-step link retention probability for 1 to 10 time steps. Note that the results of the results of the two MMP models coincide almost exactly, and the orange line (aggregated MMP) is just barely visible underneath the green line (unaggregated MMP).

MMP models, the link retention probability decreases exponentially, and fitting with each model with an exponential trend line results in a R^2 value of 1.000 in both cases. At each step the link retention probability decreases by a little over one-half; this makes sense when we consider that the average number of links in the graph is about 4.8 for both models, and on average about 2.5 links are removed in a time step. The probability that a link is not removed is then $1-(2.5/4.8) \approx 0.48$, and since the removed links are chosen randomly, the exponential decay is expected.

Compared to the mobility process, however, the link retention probability of the MMP models is lower for larger steps K. After 3 time steps, the link retention probability of the MMP models and the mobility process begins to diverge. The decay in the link retention probability of the mobility process is approximately exponential and has a R^2 value of 0.994 when fit with an exponential trend line. When looking at the plot of the link retention probability, we observe that from K=1 to K=4, the decay is not quite exponential for the mobility process (though it is quite close). However, from K=4 to K=10, the decay starts to follow an exponential decrease, but it decays slower compared to the MMP models. When fitting the points from K=4 to K=10 of the mobility process with an exponential trend line, a R^2 value of 1.000 is found; the link retention probability of the mobility process follows an exponential decay almost exactly after K=4.

5.4.3. Unique Links Observed

We can also consider differences that may arise between the MMP model and the mobility process due to the different strategies for the adding of links. In the mobility process, links are created when nodes move within a distance of $d = 1.5$ units. On the other hand, the MMP models do not have any information about the positions of nodes, and can choose links at random from the set of all links which do not currently exist in the graph. To investigate the effect that this may have on the resulting graphs, we define another metric: the total number of unique links observed up to a certain time step. The number of unique links observed from 0 to 50 time steps has been plotted in Figure 5.6; the number of unique links is averaged over all trials of the MMP model and mobility process.

The average number of unique links observed at each time step for the aggregated and unaggregated MMP model are almost the same. This makes sense because both MMP models add and remove the same amount of links on average, and they use the same strategy for choosing links to add (random selection from the set of all available links). The number of unique links observed converges to 36 because in a graph of 9 nodes, there are 36 possible links. In the case of the mobility process, the average number of unique links observed at each time step tends to be slightly higher than the MMP models. This is due to the fact that the mobility process is constrained to a small region, so nodes frequently reach the border of the region. As described in Section 2.4, when a node reaches the border of a region, the out of bounds handling will first point the node back towards the center of the region and then apply a random angle offset, so it does not move exactly towards the center in the next time step. If nodes frequently reach the border of the region, they will often be redirected to head

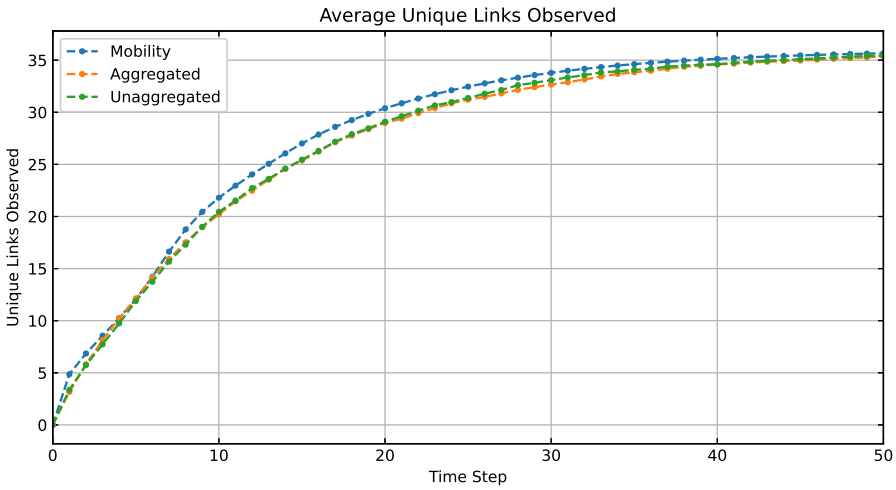


Figure 5.6: The number of unique links observed from 0 to 50 time steps.

roughly towards the center of the region, making it more likely for a new link to be observed. However, as seen in the graph, the difference does not appear to be very significant, the number of unique links observed quickly converges to 36 in all three cases. The maximum observed difference between the average unique links in the mobility process and MMP models does not exceed 1.7 links.

5.5. Quantifying Dependence on Previous Action

When comparing the distributions of the total number of links at each time step and the number of added and removed links between two time steps, we observe that the MMP models produce the same result as the mobility process. This is expected because these properties of the mobility process have been embedded in the MMP models, and the results show that the resulting graphs produced by the MMP models do indeed possess these properties. When considering the K-step link retention probability and the number of unique links observed, we see some minor differences between the MMP models and the mobility process. These differences arise from the fact that in the mobility process, links are added and removed based on the proximity of nodes, but the MMP models do not have this information and can only add and remove links at random. When comparing the results of the two MMP models, however, we are unable to observe any meaningful differences. The link distributions, link retention probability, and unique links observed are the same for both the aggregated and unaggregated MMP models. This is interesting because it suggests that the unaggregated MMP model performs just as well as the aggregated MMP model, even though the number of states is much less.

As previously discussed, the states of the unaggregated MMP model encode the number of links in the graph as well as an action to apply. Therefore, the next action in the unaggregated MMP model is dependent on the previous action as well as the current number of links. In the aggregated MMP model, the states only encode the number of links in the graph. When the Markov process transitions, an action is randomly chosen from the action set corresponding to the state transition. The probability distribution of each action set is fixed and is not affected by the previous action. Therefore, the dependence on the previous action is lost in the aggregated MMP model. This raises an interesting question: in the real mobility process, how much dependence is there between the previous action and the next action? If the dependence is negligible, then probability distributions of actions are almost entirely determined by the current number of links in the graph, and the aggregated MMP model should model the mobility process just as accurately as the unaggregated MMP model.

To understand how we can answer this question, let us first revisit the unaggregated MMP model for $N = 3$ introduced in Section 4.3.1 and illustrated in Figure 4.1. Recall that each of the arrows in the diagram points to a group of states and represents a transition probability to any states in that group.

Let us start by assuming there is currently 1 link in the graph. The six possible actions are $(+0,-1)$, $(+0,-0)$, $(+1,-1)$, $(+1,-0)$, $(+2,-1)$, and $(+2,-0)$; the actions are encoded by states 5-10 (the states illustrated in orange

in Figure 4.1). In the unaggregated MMP model, entering a state is equivalent to selecting the corresponding action. Therefore, to find the probability distribution of these actions, we need to look to the set of states which *transition to* states 5-10. The set of states which transition to states 5-10 are:

- State 2
- State 6
- State 7
- State 12
- State 13
- State 18

Let us first consider state 2. In state 2, the current number of links in the graph is 0, and the action of state 2 is $(+1,-0)$. After applying the action, there is 1 link in the graph. State 2 can then transition to any of the orange states 5-10 which encode 1 link, thereby choosing the next action to apply. The probability distribution of the actions encoded by states 5-10 corresponds to the transition probability from state 2 to the states 5-10. The same holds for the remaining states 6, 7, 12, 13, and 18: after the action of each of these states is applied, there will be 1 link in the graph, so these states will transition to one of the states 5-10 and choose the next action. Each of the states 6, 7, 12, 13, and 18 has its own probability distribution for choosing the next action. We thus have 6 different probability distributions for the next action when there is currently 1 link in the graph, with each probability distribution corresponding to a different previous action.

In the aggregated MMP model, the states encode only the number of links. Therefore, the next action depends only on the current number of links and not the previous action. There is a single probability distribution for each number of links, but the probability distribution is split into different action sets. The relevant action sets given that there is currently 1 link in the graph are as follows:

- $J_{1,0} = \{(+0,-1)\}$
- $J_{1,1} = \{(+0,-0),(+1,-1)\}$
- $J_{1,2} = \{(+1,-0),(+2,-1)\}$
- $J_{1,3} = \{(+2,-0)\}$

These are the same six actions in the unaggregated MMP model, but the action sets divide them based on the new number of links after the action is applied. In the aggregated MMP model, the states encode the number of links, and when transitioning from state i to state j, one of the actions from the corresponding action set $J_{i,j}$ is applied. The probability that action set $J_{i,j}$ is chosen is given by the transition probability from state i to state j in the transition matrix, $p_{i,j}$. The key observation is that choosing the next state in the Markov process (i.e. choosing the action set) and choosing an action from the action set are independent actions. Thus, for a given current number of links i, we can obtain the complete probability distribution of all possible actions by multiplying the probability distribution of each action set $J_{i,j}$ with the corresponding transition probability $p_{i,j}$ for all states j. This probability distribution is only dependent on the current number of links.

Continuing with the current example, in the unaggregated MMP model, we have six probability distributions for the next action given that there is current 1 link in the graph, and each probability distribution corresponds to a different previous state/action. In the aggregated MMP model, we do not have any dependence on the previous action, so there is only one probability distribution. To determine whether there is actually a dependence on the previous action, we need to compare the probability distributions. We quantify the similarity between the each of the unaggregated probability distributions with the aggregated probability distribution by calculating the coefficient of determination R^2, which is given by

$$R^2 = 1 - \frac{SS_{\text{res}}}{SS_{\text{tot}}}. \tag{5.3}$$

Here, SS_{tot} is the total sum of squares given by

$$SS_{\text{tot}} = \sum_i (y_i - \bar{y})^2, \tag{5.4}$$

where y is the probability distribution of the next action in the unaggregated MMP model. Instead, SS_{res} is the sum of squares of residuals given by

$$SS_{\text{res}} = \sum_i (y_i - f_i)^2, \tag{5.5}$$

where f is the probability distribution of the next action in the aggregated MMP model. If there is no dependence on the previous action, the probability distribution of the unaggregated and aggregated models should be the same, and an R^2 value of 1 will be observed.

We can now compare each of the probability distributions of the unaggregated MMP model with the aggregated MMP model. The unaggregated MMP model that we constructed has 2,923 states, meaning there are 2,923 probability distributions which need to be compared with the aggregated MMP model. A histogram of the calculated R^2 values is shown in Figure 5.7.

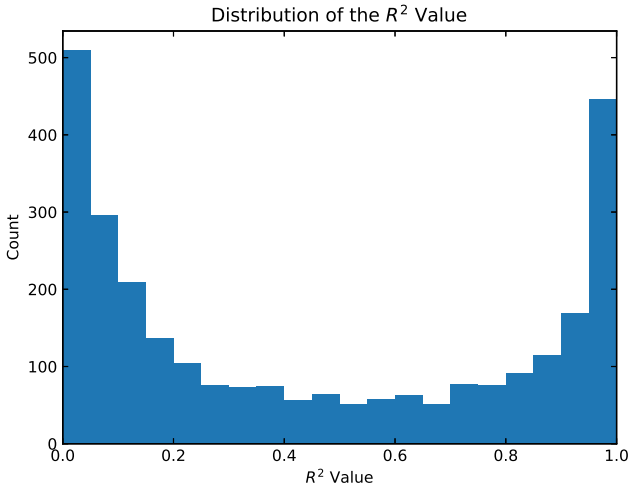


Figure 5.7: A histogram of the calculated R^2 values (bin width 0.05).

Each of the calculated R^2 values corresponds to a different state in the unaggregated MMP model. Directly taking the mean of the R^2 values results in 0.428. However, this does not accurately capture the behavior of the MMP model. First and foremost, when directly taking the average, all of the states contribute equally, and this is not fair because not all states are equally likely to be encountered. When we consider states which are rarely encountered in the MMP model, we see that there is a second problem: a state which is rarely encountered has a probability distribution which is constructed from very few data points, making the distribution less consistent and resulting in worse R^2 values. This leads to an even higher bias in the average R^2 value if the states are weighted evenly.

To address these issues, we instead weight each of the R^2 values by the steady state value of the corresponding state and take the sum. After doing so, we find the weighted R^2 value of 0.989. This weighted R^2 value takes into account the similarity between the unaggregated and aggregated probability distributions as well as the fraction of time that the given probability distribution is applied in the unaggregated MMP model. The weighted R^2 has a value which is very close to 1, meaning that probability distribution of the aggregated MMP model is almost the same as the as that of the unaggregated MMP model nearly all of the time. This strongly suggests that the next action has very little dependence on the previous action and is mostly determined by the current number of links in the graph. Moreover, this explains why the aggregated MMP model performs just as well as the unaggregated MMP model.

5.6. Discussion

In this chapter, we applied the unaggregated and aggregated MMP models to model a simulated mobility process. As discussed in the previous chapter, the unaggregated and aggregated MMP models are both link-aware; the states of the models encode the number of links in the graph. As a result, we observe that the number of links in both MMP models no longer drift, and the average number of links in the MMP models closely match the average number of links in the mobility process. The unaggregated MMP model's average number of links differs from the mobility process by about 0.32%, which is only marginally better the aggregated MMP which differs by about 0.39%. When considering the average number of links added and removed, the unaggregated MMP deviates from the mobility process by about 0.42% while the aggregated MMP model is slightly better at about 0.22%. In terms of the link metrics, both of the MMP models are able to model our

simulated mobility process with very high accuracy.

When considering the K-step link retention probability and the unique links observed, we observed some differences between the MMP models and the mobility process due to the different link removal strategies: the MMP model randomly choose links to add and remove in an unbiased way, while the links that are added and removed in the mobility process are based on the positions and movements of nodes. However, when comparing the aggregated MMP with the unaggregated MMP, we find that the results are indistinguishable. This is an important observation, because the aggregated MMP model requires far fewer states than the unaggregated MMP model and is far more preferable from a computational standpoint. Recall that the number of states in the unaggregated MMP model increases with $O(N^6)$, while the aggregated MMP model only increases with $O(N^2)$.

As discussed in Section 4.4, the information that is lost when performing the state aggregation is the dependence between consecutive actions: in the unaggregated MMP model, the next action is a function of the previous action as well as the current number of links, while in the aggregated MMP the next action is only a function of the current number of links. However, as we have just shown in Section 5.5, the amount of dependence between the next action and previous action is extremely low. Therefore, the probability distribution of the next action is predominantly determined by the number of links in the graph. This explains why we do not observe any difference in modeling performance between the aggregated and unaggregated MMP: if there is almost no dependence between the next action and previous action, then the information that is lost during the state aggregation is negligible, which appears to be the case here. In the next chapter, we consider some applications of the MMP model, and we will now focus only on the aggregated MMP model.

6

Applications of the MMP Model

6.1. Introduction

In this chapter, we consider two applications of the MMP model. In Section 6.2, we consider using the MMP model to model a real mobility process, and in Section 6.3 we consider simulating a SIR epidemic spreading process on the series of graphs generated by the MMP model. In the previous chapter, we have shown that the aggregated MMP performs just as well as the unaggregated MMP model while requiring far fewer states, and we therefore limit ourselves to considering the aggregated MMP model in this chapter. Due to the limited availability of real mobility data and time constraints, this chapter only provides a high-level overview; the intention is to provide some insight into how the MMP model behaves in these two applications, in the hopes that it may inspire directions for deeper research in the future.

6.2. Modeling a Real Mobility Process with the Aggregated MMP Model

6.2.1. Data Collection and Preprocessing

In this section, we consider the modeling of a real mobility process recorded in a library. The data has been collected as part of an ongoing project led by Prof. dr. ir. Serge Hoogendoorn, which involves various researchers at TU Delft. We would like to thank Dr. Marco Rinaldi for graciously agreeing to share the library mobility data and performing the data preprocessing.

The mobility data was recorded in a library over the course of a single day, from 08:00 to 18:00. In the tracking experiment, each participant wears a Bluetooth tracking device and there are fixed Bluetooth beacons inside the library. Each Bluetooth tracking device will record contact events with beacons, and each data record includes the time and duration of the contact, the distance from the beacon, and information to uniquely identify the device and beacon. There are 37 participants in the tracking experiment.

The list of all contact events is initially filtered based on proximity to the Bluetooth beacons; only contact events at distances classified as dangerous are considered. The distance is chosen to be 1.5 meters, and contact events at distances which exceed 1.5 meters are filtered out. The list of filtered contact events is then used to estimate the contacts between participants. When two participants i and j are in contact with the same beacon at the same time, the distance $d_{i,j}$ between the two participants is estimated as

$$\hat{d}_{i,j} = \sqrt{d_i^2 + d_j^2}, \tag{6.1}$$

where d_i and d_j are the distances of participant i and j from the beacon respectively. Two participants are considered to be in contact when the distance $\hat{d}_{i,j}$ is less than 1.5 meters.

A series of contact matrices is then generated by polling the data at a fixed interval of 30 seconds. Every 30 seconds, a contact matrix C is generated; the contact matrix is a 37×37 matrix which represents the number of contacts between participants (hereafter referred to as nodes) and the duration of each contact over the 30 second interval since the previous polling time. Each element $c_{i,j}$ in the contact matrix is equal to 0 if there is no contact between node i and node j during the last 30 seconds. Otherwise, element $c_{i,j}$ is a real number in the range $[0,30]$ which represents the duration (in seconds) of the contact between node i and node j. Polling the data from 08:00 to 18:00 with a polling interval of 30 seconds results in 1,201 contact matrices.

6.2.2. Building the MMP Model

We have a series of 1,201 contact matrices which we denote as $C(t)$. The time step t is an integer in the range $[0,1200]$; the physical interpretation of t is the number of 30 second intervals since 08:00, that is $t = 0$ represents 08:00 and $t = 1,200$ represents 18:00. Each contact matrix can be interpreted as a weighted adjacency matrix where the weight is the duration of the contact. For our modeling, we are only interested in whether a contact exists, and the duration of the contact is not considered. We therefore convert the series of contact matrices $C(t)$ into a series of unweighted adjacency matrices $A(t)$. At each time t, if there is no contact between node i and node j, that is the element $c_{i,j}(t) = 0$, then we set the element $a_{i,j}(t) = 0$. Otherwise, if there is a contact between node i and node j, that is the element $c_{i,j}(t) > 0$, then we set the element $a_{i,j}(t) = 1$.

After converting the contact matrices into adjacency matrices, we plot the number of links (i.e. contacts) at each time step in Figure 6.1. We observe that there are almost no contacts recorded from 08:00 to 11:04. From 11:04 to 12:54, there is a spike in the number of contacts. From 12:54 to 15:10, there are no contacts recorded. From 15:10 to 16:20, there is another spike in in the number of contacts. After 16:20, no more contacts are recorded.

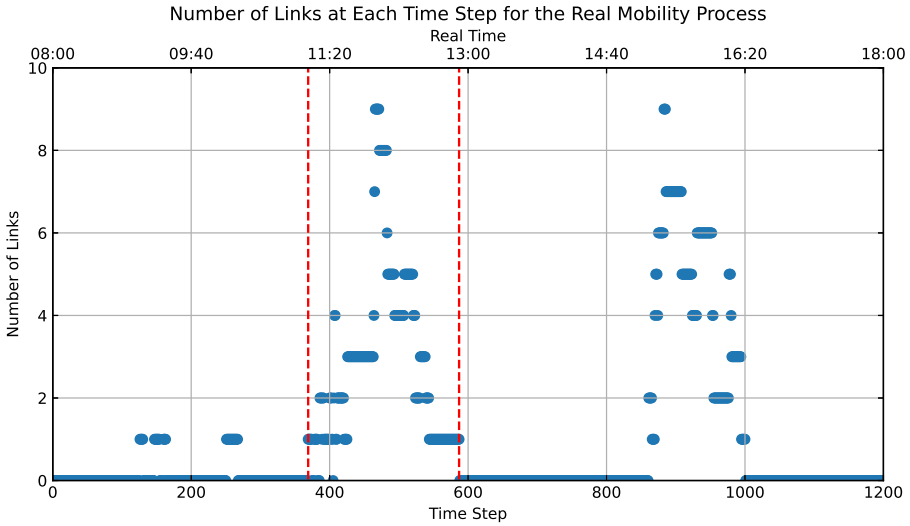


Figure 6.1: Number of links at each time step for the real mobility process. The dotted red lines show the start and end of the time window which will be modeled by the MMP model.

Since our MMP model is based on the number of links in the graph, we focus on modeling the first spike in contacts from 11:04 to 12:54; this time window is marked by the dotted red lines in Figure 6.1. We again calculate the total number of links at each time step and the number of links added and removed between two time steps. The distributions of these values can be found in Appendix C.4 and the averages are as follows:

- Average number of links at each time step: 2.845.
- Average number of links added between two time steps: 0.110.
- Average number of links removed between two time steps: 0.110.

From our investigation in the previous chapter, we found that the aggregated MMP model produces results which are indistinguishable from those of the unaggregated MMP model while requiring far fewer states. We therefore choose to use the aggregated MMP model to model the mobility process.[1] We use the adjacency matrices from time step 368 to 588 to generate the probability transition matrix and build the action sets. The modeled segment of the mobility process runs for a little over 200 time steps, so we choose to simulate the MMP model for 200 time steps; the number of trials is kept at 100 as in the previous models.

[1] Since we are not considering the unaggregated MMP model, we will simply refer to the aggregated MMP model as "the MMP model" for the remainder of this chapter for brevity.

6.2.3. Results of the MMP Model

The maximum number of links in a graph of 37 nodes is 666, and the MMP model will have a maximum of 667 states because one additional state is needed to encode 0 links. After constructing the MMP model, there are only 10 states because the maximum number of links we observed was 9. With 10 states, there are 100 possible state transitions, but we only observed 27 of them so there are 27 action sets.

In each trial, the Markov process is initialized in state 0 with an empty graph. After performing 100 simulations of 200 time steps, we calculate the total number of links at each time step and the number of links added and removed between two time steps. The distributions of these values can be found in Appendix C.5 and the averages are as follows:

- Average number of links at each time step: 2.663.
- Average number of links added between two time steps: 0.123.
- Average number of links removed between two time steps: 0.109.

In the MMP model, the average number of links observed at each time step is about 6.4% lower compared to the real mobility process. The average number of links added is about 12% higher, and the average number of links removed is about 1.5% lower. The larger percentage difference is partially due to the fact that the number of links in the graph is lower, and also because the model is run for fewer time steps.[2]

A plot of the number of links in the modulated graph at each time step for two trials of the MMP model is shown in Figure 6.2. Figure 6.2a shows the number of links at each time step in the mobility process and Figure 6.2b shows the number of links in the first trial of the MMP model. Figure 6.2c shows the number of links in the fourth trial of the MMP model; this result is included because in the first three trials of the MMP model, the maximum number of links was only 4 so we are unable to see how the model behaves at higher numbers of links.

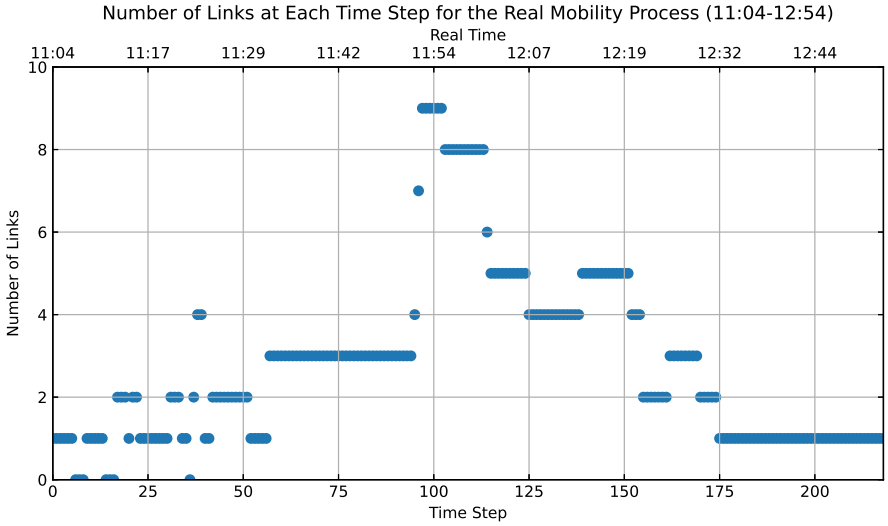
One immediate observation is that the real mobility process exhibits a pronounced increase in the number of links followed by a decrease, but this trend is not captured by the MMP model. When looking at Figure 6.2c, we observe that the modulated graph reaches 9 links at $t = 87$ and stays there until $t = 97$. At $t = 98$ it then transitions to 8 links, and stays there for until $t = 130$. To better understand this behavior, we can take a look at the probability transition matrix of the MMP model which is shown in Figure 6.3.

	0	1	2	3	4	5	6	7	8	9
0	0.57	0.14	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.04	0.90	0.04	0.01	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.16	0.77	0.03	0.03	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	0.02	0.96	0.02	0.00	0.00	0.00	0.00	0.00
4	0.00	0.05	0.05	0.00	0.80	0.05	0.00	0.05	0.00	0.00
5	0.00	0.00	0.00	0.00	0.09	0.91	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
8	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.00	0.91	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.17	0.83

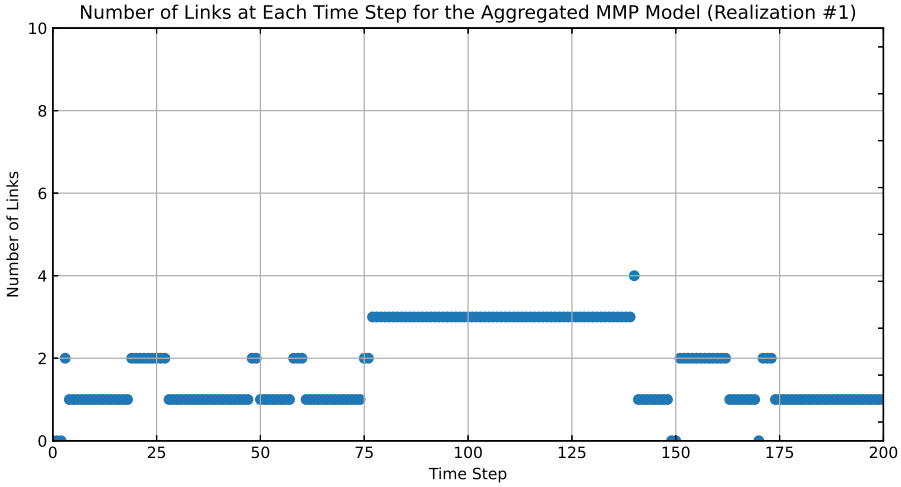
Figure 6.3: Probability transition matrix of the MMP model based on the real mobility process.

When looking at row 9 of the transition matrix, we see that there is a 83% probability of staying at 9 links and a 17% probability of transitioning to 8 links. Row 9 corresponds to state 9 which encodes 9 links in the graph. We can see how this probability is derived when we look at the graph of the number of links in the mobility process in Figure 6.2a. At time $t = 97$, the graph has 9 links. In the next 5 time steps, the graph stays at 9 links, i.e. we observe 5 transitions from 9 links to 9 links. Finally, at $t = 103$, the graph transitions to 8 links and for the remainder of the mobility process, we do ever reach 9 links in the graph again. In total we observed 5 transitions from 9 links to 9 links, and 1 transition from 9 links to 8 links. Apart from this, we did not observe 9 links transition to any other number of links. Therefore, in the probability transition matrix, we see that state 9 only transitions to state 8 and 9, and the probabilities are $1/6 \approx 0.17$ and $5/6 \approx 0.83$ respectively.

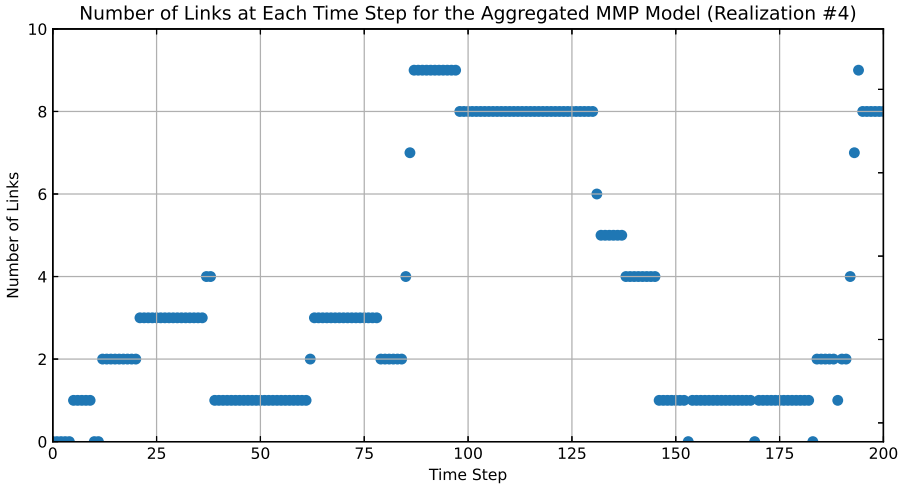
[2] When the MMP simulation was repeated for 100 trials of 1,000 time steps, the average number of links was 2.752 which is about 3.3% less than the mobility process. The average number of links added was 0.110 which is 0.3% higher than the mobility process, and the average number of links removed was 0.108 which is 3.1% lower than the mobility process.



(a) Number of links at each time step in the time window of the real mobility process used to create the MMP model. The time steps have been reindexed starting from 0.



(b) Number of links at each time step for the first realization of the MMP model.



(c) Number of links at each time step for the fourth realization of the MMP model.

Figure 6.2: Comparison of the number of links at each time step in the MMP modulated graph and the real mobility process.

Due to the limited amount of data, we observe that the MMP model very specific to the idiosyncrasies of this particular dataset and the model "overfits" the data. For example, let us consider how state 9 can be reached. From the probability transition matrix, we see that the only way to reach state 9 is from state 7, and the transition probability is 1. When we consider the original mobility process, the interpretation is as follows: the only way for there to be 9 contacts between people is if there were 7 contacts between them 30 seconds ago (one time step). Furthermore, if there are currently 7 contacts between people, there is a 100% probability that after 30 seconds, there will be 9 contacts. Of course we see that this is ridiculous, but the MMP model is only able to model transitions that have been observed in the mobility process, and transitions that have not been observed in the real mobility process have a probability of 0 in the MMP model as well.

6.3. SIR Simulations on the Aggregated MMP Model

6.3.1. SIR Epidemic Model Overview

In this section, we consider a SIR spreading process which runs on the time-varying graphs produced by the MMP model. We will investigate whether there are any differences between a SIR spreading process running on top of the MMP model compared to the same SIR spreading process running on top of the mobility process. The purpose of this investigation is to determine whether an MMP model may be a suitable model for the mobility process in the context of a SIR simulation, and whether there is potential for this idea to be explored further.

In the SIR model, nodes are in one of three states: susceptible, infectious, or recovered. A susceptible node who is neighbors with an infectious node can be infected with the infection probability β. If the susceptible node has multiple infectious neighbors, each neighbor can infect the susceptible node with probability β. An infectious node can become recovered with the curing probability δ. Once an infectious node has recovered, it can no longer infect others or be infected again.

6.3.2. Building the MMP Model

In our investigation, we are considering a mobility process in which the nodes are moving, and therefore a node can have different neighbors at each time step. For the purpose of the SIR simulation, we use the data of the simulated 25 node mobility process that was discussed in chapter 2. To reiterate, the link metrics of the 25 node mobility process are as follows:

- Average number of links at each time step: 19.466.
- Average number of links added between two time steps: 10.377.
- Average number of links removed between two time steps: 10.357.

From the data of the mobility process, we create an aggregated MMP model. The aggregated MMP model is used to modulate a graph for 1,000 time steps, and we perform 100 trials. For the aggregated MMP model, we calculate the following link metrics:

- Average number of links at each time step: 19.511.
- Average number of links added between two time steps: 10.395.
- Average number of links removed between two time steps: 10.377.

These link metrics of the aggregated MMP model are very close to the mobility process, differing by less than 1%.

6.3.3. SIR Simulation Process

We have 25 nodes, and at the start of the SIR simulation, we randomly choose 1 node to be infectious while the remaining 24 nodes are susceptible. Each trial of the mobility process and MMP model consists of 1,001 adjacency matrices representing the links between nodes. At each time step, a susceptible node that shares a link with an infectious node can be infected with probability β. If a susceptible node is connected to multiple infectious nodes, each infectious node can independently try to infect the susceptible node with probability β. If at least one of the infection attempts succeeds, then the susceptible node becomes infectious at the next time step; given m infectious neighbors, the probability that a susceptible node is not infected is therefore

$(1-\beta)^m$. At each time step, an infectious node can also recover with probability δ. If this happens, then the node becomes recovered at the next time step and can no longer infect any other nodes or be infected again.

For both the mobility process and MMP model, we have the adjacency matrices of 100 trials. On each series of adjacency matrices, we perform 10 trials of the SIR simulation. After performing the simulations, we average the number of susceptible, infectious, and recovered nodes at each time step over all 1,000 trials and plot the result. We perform SIR simulations for $\delta = 0.1, 0.3$ and $\beta = 0.1, 0.3, 0.5, 0.7$. The results for $\delta = 0.1$ and $\beta = 0.3$ are plotted in Figure 6.4. The remaining simulation results can be found in Appendix D.

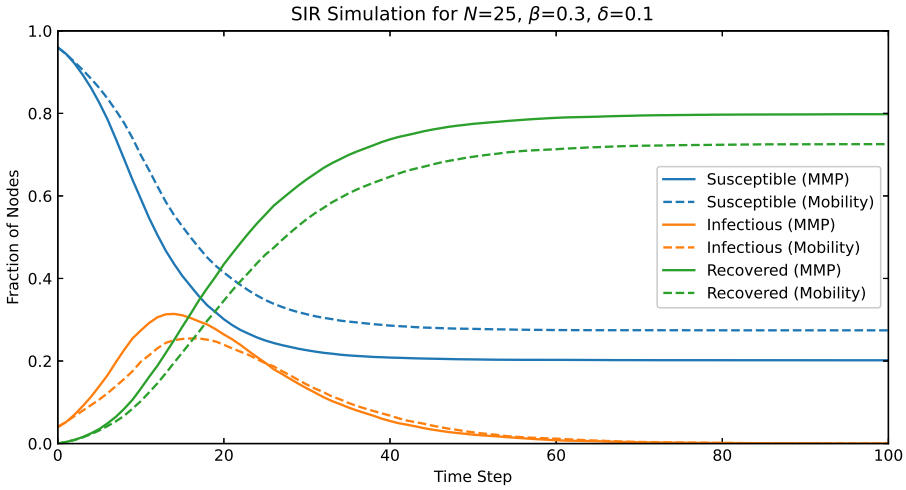


Figure 6.4: SIR simulation results for $\beta = 0.3$, $\delta = 0.1$.

Across all of the simulations, we observe similar curves for the fraction of susceptible, infectious, and recovered individuals in the MMP model and the mobility process. In all cases, the average fraction of infectious individuals in the MMP model tends to be slightly higher than in the mobility process. This could be due to the fact that the MMP model does not know anything about the positions of nodes so it adds links at random. There are no restrictions on which links can be added, whereas in the mobility process, the creation of links is tied to the positions of nodes so it is not possible for nodes that are far away from each other to suddenly become connected. Therefore, it is comparatively harder for the infection to spread in the mobility process.

7

Conclusion and Future Directions

7.1. Conclusion

In this research, we have investigated the modeling of a simulated mobility process using an MMP model. We took a link-based approach to this problem, wherein the actions of the MMP model are defined as jointly adding and removing a certain number of links from the graph. In our first MMP model described in Chapter 3, the states of the Markov process only encode an action, and the Markov process is blind to the number of links which exist in the graph. Although the first MMP model was able to capture the average number of links added and removed between two time steps, the average number of links in the modulated graph did not match the mobility process. The key observation of Chapter 3 is that if the MMP model does not have knowledge of how many links are in the graph, the number of links in the modulated graph will behave like a random walk. In the long term, the average number of links added and removed by the MMP model is the same, but without any form of feedback to control the number of links, the number of links will drift. Therefore, in order to be able to accurately model the mobility process with this link-based approach, the MMP model must contain some information about the number of links in the graph.

In Chapter 4, we designed a link-aware MMP model in which each state of the MMP model encodes an action as well as the number of links in the graph, and we refer to this as the unaggregated MMP model. The main disadvantage of the unaggregated MMP model is that the number of states grows with $O(N^6)$, which quickly becomes infeasible as the number of nodes N increases. We therefore introduced the concept of state aggregation to merge states of the unaggregated MMP model based on the number of links encoded by the states. This results in the aggregated MMP model, which only requires $O(N^2)$ states. We proved that the aggregated MMP model will produce a modulated graph which has the same average number of links as the unaggregated MMP model. The aggregated MMP model also applies actions at the same rate as the unaggregated MMP model, so average number of links added and removed between two time steps will also be the same in both models. The key difference between the aggregated and unaggregated MMP model is that the probability distribution of the next action in the unaggregated MMP model is dependent both on the current number of links in the graph and also the previous action. However, in the aggregated MMP model, the probability distribution of the next action is only dependent on the current number of links in the graph, and we lose the dependence on the previous action.

In Chapter 5, we chose a small number of nodes $N = 9$ and constructed both an aggregated and unaggregated MMP model to investigate the differences between their modeling performance. The results showed that both MMP models could accurately model the average number of links in the mobility process as well as the number of links added and removed between two time steps. The results produced by aggregated and unaggregated MMP models were effectively identical. As mentioned previously, the aggregated MMP model cannot take into account the dependence between consecutive actions. When we investigated the amount of dependence between actions in the $N = 9$ mobility process, we found that there was almost no dependence. Therefore, the aggregated MMP model could produce results which are just as good as the unaggregated MMP model while using $O(N^2)$ states compared to $O(N^6)$ in the unaggregated MMP model.

In Chapter 6, we applied the aggregated MMP model to model a real mobility process. Due to the fact that the amount of data was very limited, we observed that the MMP model overfits the mobility process. In spite of this, the MMP model was still able to fit the data relatively well when considering the average number of

links as well as the average number of added and removed links, though we did observe a larger percentage error in the averages compared to the previous chapters when we modeled a simulated mobility process. Due to the very limited amount of data, we cannot draw any definitive conclusions from this.

Finally, we simulated SIR spreading processes on the graphs produced by the simulated mobility process and the graphs produced by the aggregated MMP. Across all of the SIR simulations, we observed that the fraction of infectious individuals for the aggregated MMP model was higher than in the mobility process. We performed these simulations as an initial test to provide some insight as to whether or not the MMP model might be a suitable model for a mobility process in the context of studying epidemic processes. For most of the SIR simulations, the MMP and mobility process appear to produce similar results. Furthermore, in the context of an epidemic process, it is preferable to err on the side of caution and overestimate the fraction of infectious individuals, suggesting that this idea could be worth exploring.

7.2. Future Directions

This thesis project represents only the first step in exploring the idea of applying the Markov modulated process to model mobility processes. There are still many questions to be answered, and in this section, we propose some directions for future research.

Epidemics

In Section 6.3, we took a brief look at SIR simulations on the MMP model compared to the simulated mobility process from which the MMP model was derived. We observed that in the MMP model, the peak number of infectious individuals in the SIR simulation was consistently higher than in the simulated mobility process. The SIR dynamics of the MMP model has now been investigated in more detail by Chang [7]. Chang considered a larger number of nodes and more MMP models; in all of the simulations, the infection on the MMP model consistently spread faster and to more nodes than the mobility process. Chang concluded that the infection spreading in the mobility process is bottlenecked by the movement of nodes whereas the MMP model can randomly create links between many different parts of the graph and spread the infection much faster.

Changing Number of Nodes

In our investigation, we have considered a constant number of nodes, and in our simulated mobility process, we restricted the nodes to remain inside the region of interest. In a real mobility process, however, the number of nodes is unlikely to be constant, as people can leave or enter the area that is under observation. A possible direction for future research is to investigate how the MMP model can be applied in situations where the number of nodes is changing.

One interesting observation that we can immediately make is that the MMP model does not actually know how many nodes are in the graph. For example, consider an aggregated MMP model where the maximum number of links encoded by the model is 20. In principle, this MMP model can actually be used to modulate *any* graph with enough nodes to accommodate at least 20 links.

Dependence of Actions in a Real Mobility Process

As a result of the state aggregation, the aggregated MMP is unable to model the dependence between actions, and the next action in the aggregated MMP is only dependent on the current number of links in the graph. The key result of Chapter 5 was that if there is little to no dependence between actions, then the aggregated MMP produces results which are indistinguishable from the unaggregated MMP while using far fewer states. Therefore, it would be interesting to investigate the level of dependence between actions for a real mobility process. If the dependence is low, then the aggregated MMP is clearly the better option. However, if the dependence is non-negligible, then it raises the question of how much the accuracy of the aggregated MMP is reduced and whether the tradeoff in the computational complexity can justify it.

Dealing with Insufficient Data and Overfitting

As seen in Section 6.2, when there is a very limited amount of data, the MMP model can overfit the mobility process and produce a transition matrix that models the idiosyncrasies of the data. It could be interesting to explore how this issue can be dealt with. For example, perhaps the probability transition matrix could be perturbed to allow for state transitions that were not actually observed in the data. This could be investigated

in a controlled manner by performing many trials of a simulated mobility process, but only using the results of a single trial to generate the MMP model.

Bibliography

[1] World Health Organization, "WHO coronavirus (COVID-19) dashboard." [Online]. Available: https://covid19.who.int/table

[2] Y. Zhou, R. Xu, D. Hu, Y. Yue, Q. Li, and J. Xia, "Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data," *The Lancet Digital Health*, vol. 2, pp. e417–e424, 08 2020.

[3] N. Kishore, R. Kahn, P. Martinez, P. De Salazar, A. Mahmud, and C. Buckee, "Lockdowns result in changes in human mobility which may impact the epidemiologic dynamics of SARS-CoV-2," *Scientific Reports*, vol. 11, no. 1, p. 6995, Mar. 2021.

[4] B. Espinoza, C. Castillo-Chávez, and C. Perrings, "Mobility restrictions for the control of epidemics: When do they work?" *PLOS ONE*, vol. 15, p. e0235731, 07 2020.

[5] P. Van Mieghem, *Performance analysis of complex networks and systems*. Cambridge University Press, 2014.

[6] J. T. Chang, "Stochastic processes," 2007, unpublished manuscript. [Online]. Available: http://www.stat.yale.edu/~pollard/Courses/251.spring2013/Handouts/Chang-notes.pdf

[7] B. Chang, "Modeling the spread of epidemics: Visualizing the spread of COVID-19 and applying a Markov-modulated process model to mobility processes," Master's thesis, Delft University of Technology, 2021.

A

Number of States in the Unaggregated MMP Model as a Function of N

The number of states in the unaggregated MMP model is given by

$$\#X = \sum_{L=0}^{L_{\max}} (L+1)(L_{\max} - L + 1), \tag{A.1}$$

where $L_{\max}$ is the maximum number of links which can exist in the graph, i.e. the number of links in the complete graph. For a graph of N nodes $L_{\max}$ is given by

$$L_{\max} = N(N-1)/2. \tag{A.2}$$

We start be expanding the product in (A.1) and then collecting the terms of the sum, resulting in

$$\begin{aligned}
\#X &= \sum_{L=0}^{L_{\max}} (L_{\max}L - L^2 + L + L_{\max} - L + 1) \\
&= \sum_{L=0}^{L_{\max}} (L_{\max}L - L^2 + L_{\max} + 1) \\
&= -\sum_{L=0}^{L_{\max}} L^2 + \sum_{L=0}^{L_{\max}} L_{\max}L + \sum_{L=0}^{L_{\max}} (L_{\max}+1) \\
&= -\sum_{L=0}^{L_{\max}} L^2 + L_{\max}\sum_{L=0}^{L_{\max}} L + (L_{\max}+1)\sum_{L=0}^{L_{\max}} 1.
\end{aligned} \tag{A.3}$$

We first evaluate each of the sums individually. From Faulhaber's formula, the sum $\sum_{L=0}^{L_{\max}} L^2$ is equal to

$$\sum_{L=0}^{L_{\max}} L^2 = \frac{1}{6} L_{\max}(L_{\max}+1)(2L_{\max}+1). \tag{A.4}$$

The sum $\sum_{L=0}^{L_{\max}} L$ is the sum of an arithmetic series. The result is well-known and the sum is equal to

$$\sum_{L=0}^{L_{\max}} L = \frac{1}{2} L_{\max}(L_{\max}+1). \tag{A.5}$$

The sum $\sum_{L=0}^{L_{\max}} 1$ is trivial and is equal to

$$\sum_{L=0}^{L_{\max}} 1 = L_{\max} + 1. \tag{A.6}$$

Substituting (A.4), (A.5), and (A.6) into (A.3), we find

$$\#X = -\frac{1}{6}L_{\max}(L_{\max}+1)(2L_{\max}+1) + \frac{1}{2}L_{\max}^2(L_{\max}+1) + (L_{\max}+1)^2. \tag{A.7}$$

After expanding the products and collecting like terms, we find an expression for the number of states $\#X$ in terms of the maximum number of links $L_{\max}$:

$$\#X = \frac{1}{6}L_{\max}^3 + L_{\max}^2 + \frac{11}{6}L_{\max} + 1. \tag{A.8}$$

Finally, we substitute (A.2) into (A.8). A calculator is used to simply the result, and we obtain the expression for the number of states $\#X$ in terms of the number of nodes N:

$$\begin{aligned}\#X &= \frac{N^6 - 3N^5 + 15N^4 - 25N^3 + 56N^2 - 44N}{48} + 1 \\ &= \frac{N^6}{48} - \frac{N^5}{16} + \frac{5N^4}{16} - \frac{25N^3}{48} + \frac{7N^2}{6} - \frac{11N}{12} + 1.\end{aligned} \tag{A.9}$$

Therefore, in the unaggregated MMP model, the number of states grows with $O(N^6)$.

B

Derivation of the Aggregated MMP Model for N=3

B.1. A Simple Algorithm for State Aggregation

When aggregating a set of states M into a single state m, the transition probabilities $p_{i,m}$ are calculated as

$$p_{i,m} = \sum_{j \in M} p_{i,j}. \tag{B.1}$$

The transition probabilities $p_{m,i}$ are calculated as

$$p_{m,i} = \frac{1}{\pi_m} \sum_{j \in M} \pi_j p_{j,i}, \tag{B.2}$$

where π_m is the sum of the steady state probability of all of the states in m. Given the probability transition matrix of the Markov process, we can devise a simple two-step algorithm for calculating the probability transition matrix after the state aggregation:

1. Condense the columns corresponding to the states in M into a single column by taking the row sum of each column.

2. Condense the rows corresponding to the states in M into a single row by weighting each row with the steady state probability of the corresponding row and taking the column sum, then divide the result by the total steady state probability of all states in M.

B.2. First Aggregation Step: States 17-20

We aggregate states 17-20 into a single state L3 which encodes 3 links in the graph. L3 will have a steady state probability given by

$$\pi_{\text{L3}} = \sum_{i=17}^{20} \pi_i. \tag{B.3}$$

We first take the row sum of columns 17-20 resulting in Figure B.1. We then weight each row 17-20 by the corresponding steady state probability, take the column sum, and divide the result by the total steady state probability of states 17-20. This results in the transition matrix shown in Figure B.2 and yields the following transition probabilities for $p_{\text{L3},i}$ for $i = 1,...,16$:

$$p_{\text{L3},i} = \begin{cases} \frac{\pi_{17}}{\pi_{\text{L3}}} p_{17,i} & i = 1,...,4 \\ \frac{\pi_{18}}{\pi_{\text{L3}}} p_{18,i} & i = 5,...,10 \\ \frac{\pi_{19}}{\pi_{\text{L3}}} p_{19,i} & i = 11,...,16 \end{cases} \tag{B.4}$$

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	L3
1	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$													
2					$p_{2,5}$	$p_{2,6}$	$p_{2,7}$	$p_{2,8}$	$p_{2,9}$	$p_{2,10}$							
3											$p_{3,11}$	$p_{3,12}$	$p_{3,13}$	$p_{3,14}$	$p_{3,15}$	$p_{3,16}$	
4																	1
5	$p_{5,1}$	$p_{5,2}$	$p_{5,3}$	$p_{5,4}$													
6					$p_{6,5}$	$p_{6,6}$	$p_{6,7}$	$p_{6,8}$	$p_{6,9}$	$p_{6,10}$							
7					$p_{7,5}$	$p_{7,6}$	$p_{7,7}$	$p_{7,8}$	$p_{7,9}$	$p_{7,10}$							
8											$p_{8,11}$	$p_{8,12}$	$p_{8,13}$	$p_{8,14}$	$p_{8,15}$	$p_{8,16}$	
9											$p_{9,11}$	$p_{9,12}$	$p_{9,13}$	$p_{9,14}$	$p_{9,15}$	$p_{9,16}$	
10																	1
11	$p_{11,1}$	$p_{11,2}$	$p_{11,3}$	$p_{11,4}$													
12					$p_{12,5}$	$p_{12,6}$	$p_{12,7}$	$p_{12,8}$	$p_{12,9}$	$p_{12,10}$							
13					$p_{13,5}$	$p_{13,6}$	$p_{13,7}$	$p_{13,8}$	$p_{13,9}$	$p_{13,10}$							
14											$p_{14,11}$	$p_{14,12}$	$p_{14,13}$	$p_{14,14}$	$p_{14,15}$	$p_{14,16}$	
15											$p_{15,11}$	$p_{15,12}$	$p_{15,13}$	$p_{15,14}$	$p_{15,15}$	$p_{15,16}$	
16																	1
17	$p_{17,1}$	$p_{17,2}$	$p_{17,3}$	$p_{17,4}$													
18					$p_{18,5}$	$p_{18,6}$	$p_{18,7}$	$p_{18,8}$	$p_{18,9}$	$p_{18,10}$							
19											$p_{19,11}$	$p_{19,12}$	$p_{19,13}$	$p_{19,14}$	$p_{19,15}$	$p_{19,16}$	
20																	1

Figure B.1: Probability transition matrix of the MMP model for $N = 3$ nodes after condensing columns 17-20.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	L3
1	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$													
2					$p_{2,5}$	$p_{2,6}$	$p_{2,7}$	$p_{2,8}$	$p_{2,9}$	$p_{2,10}$							
3											$p_{3,11}$	$p_{3,12}$	$p_{3,13}$	$p_{3,14}$	$p_{3,15}$	$p_{3,16}$	
4																	1
5	$p_{5,1}$	$p_{5,2}$	$p_{5,3}$	$p_{5,4}$													
6					$p_{6,5}$	$p_{6,6}$	$p_{6,7}$	$p_{6,8}$	$p_{6,9}$	$p_{6,10}$							
7					$p_{7,5}$	$p_{7,6}$	$p_{7,7}$	$p_{7,8}$	$p_{7,9}$	$p_{7,10}$							
8											$p_{8,11}$	$p_{8,12}$	$p_{8,13}$	$p_{8,14}$	$p_{8,15}$	$p_{8,16}$	
9											$p_{9,11}$	$p_{9,12}$	$p_{9,13}$	$p_{9,14}$	$p_{9,15}$	$p_{9,16}$	
10																	1
11	$p_{11,1}$	$p_{11,2}$	$p_{11,3}$	$p_{11,4}$													
12					$p_{12,5}$	$p_{12,6}$	$p_{12,7}$	$p_{12,8}$	$p_{12,9}$	$p_{12,10}$							
13					$p_{13,5}$	$p_{13,6}$	$p_{13,7}$	$p_{13,8}$	$p_{13,9}$	$p_{13,10}$							
14											$p_{14,11}$	$p_{14,12}$	$p_{14,13}$	$p_{14,14}$	$p_{14,15}$	$p_{14,16}$	
15											$p_{15,11}$	$p_{15,12}$	$p_{15,13}$	$p_{15,14}$	$p_{15,15}$	$p_{15,16}$	
16																	1
L3	$p_{L3,1}$	$p_{L3,2}$	$p_{L3,3}$	$p_{L3,4}$	$p_{L3,5}$	$p_{L3,6}$	$p_{L3,7}$	$p_{L3,8}$	$p_{L3,9}$	$p_{L3,10}$	$p_{L3,11}$	$p_{L3,12}$	$p_{L3,13}$	$p_{L3,14}$	$p_{L3,15}$	$p_{L3,16}$	π_{20}/π_{L3}

Figure B.2: Probability transition matrix of the MMP model for $N = 3$ nodes after condensing rows 17-20.

The transition probability $p_{\text{L3,L3}}$ goes from an aggregated state to another aggregated state, so its transition probability is now fixed and will no longer change during the rest of the aggregation steps. It has been highlighted in green in the probability transition matrix shown in Figure B.2.

B.3. Second Aggregation Step: States 11-16

We aggregate states 11-16 into a single state L2 which encodes 2 links in the graph. L2 will have a steady state probability given by

$$\pi_{\text{L2}} = \sum_{i=11}^{16} \pi_i. \tag{B.5}$$

	1	2	3	4	5	6	7	8	9	10	L2	L3
1	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$								
2					$p_{2,5}$	$p_{2,6}$	$p_{2,7}$	$p_{2,8}$	$p_{2,9}$	$p_{2,10}$		
3											1	
4												1
5	$p_{5,1}$	$p_{5,2}$	$p_{5,3}$	$p_{5,4}$								
6					$p_{6,5}$	$p_{6,6}$	$p_{6,7}$	$p_{6,8}$	$p_{6,9}$	$p_{6,10}$		
7					$p_{7,5}$	$p_{7,6}$	$p_{7,7}$	$p_{7,8}$	$p_{7,9}$	$p_{7,10}$		
8											1	
9											1	
10												1
11	$p_{11,1}$	$p_{11,2}$	$p_{11,3}$	$p_{11,4}$								
12					$p_{12,5}$	$p_{12,6}$	$p_{12,7}$	$p_{12,8}$	$p_{12,9}$	$p_{12,10}$		
13					$p_{13,5}$	$p_{13,6}$	$p_{13,7}$	$p_{13,8}$	$p_{13,9}$	$p_{13,10}$		
14											1	
15											1	
16												1
L3	$p_{L3,1}$	$p_{L3,2}$	$p_{L3,3}$	$p_{L3,4}$	$p_{L3,5}$	$p_{L3,6}$	$p_{L3,7}$	$p_{L3,8}$	$p_{L3,9}$	$p_{L3,10}$	π_{19}/π_{L3}	π_{20}/π_{L3}

Figure B.3: Probability transition matrix of the MMP model for $N = 3$ nodes after condensing columns 11-16.

We first take the row sum of columns 11-16 resulting in Figure B.3. The transition probability $p_{\text{L3,L2}}$ follows from (B.4):

$$\begin{aligned} p_{\text{L3,L2}} &= \sum_{i=11}^{16} p_{\text{L3},i} \\ &= \sum_{i=11}^{16} \frac{\pi_{19}}{\pi_{\text{L3}}} p_{19,i} \\ &= \frac{\pi_{19}}{\pi_{\text{L3}}} \sum_{i=11}^{16} p_{19,i} \\ &= \frac{\pi_{19}}{\pi_{\text{L3}}}. \end{aligned} \tag{B.6}$$

The transition probability $p_{\text{L3,L2}}$ goes from an aggregated state to another aggregated state, so its transition probability is now fixed and will no longer change during the rest of the aggregation steps. It has been highlighted in green in the probability transition matrix.

We then weight each row 11-16 by the corresponding steady state probability, take the column sum, and divide the result by the total steady state probability of states 11-16. This results in the transition matrix shown in Figure B.4 and yields the following transition probabilities for $p_{\text{L2},i}$ for $i = 1, \ldots, 10$:

	1	2	3	4	5	6	7	8	9	10	L2	L3
1	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$								
2					$p_{2,5}$	$p_{2,6}$	$p_{2,7}$	$p_{2,8}$	$p_{2,9}$	$p_{2,10}$		
3											1	
4												1
5	$p_{5,1}$	$p_{5,2}$	$p_{5,3}$	$p_{5,4}$								
6					$p_{6,5}$	$p_{6,6}$	$p_{6,7}$	$p_{6,8}$	$p_{6,9}$	$p_{6,10}$		
7					$p_{7,5}$	$p_{7,6}$	$p_{7,7}$	$p_{7,8}$	$p_{7,9}$	$p_{7,10}$		
8											1	
9											1	
10												1
L2	$p_{L2,1}$	$p_{L2,2}$	$p_{L2,3}$	$p_{L2,4}$	$p_{L2,5}$	$p_{L2,6}$	$p_{L2,7}$	$p_{L2,8}$	$p_{L2,9}$	$p_{L2,10}$	$(\pi_{14}+\pi_{15})/\pi_{L2}$	π_{16}/π_{L2}
L3	$p_{L3,1}$	$p_{L3,2}$	$p_{L3,3}$	$p_{L3,4}$	$p_{L3,5}$	$p_{L3,6}$	$p_{L3,7}$	$p_{L3,8}$	$p_{L3,9}$	$p_{L3,10}$	π_{19}/π_{L3}	π_{20}/π_{L3}

Figure B.4: Probability transition matrix of the MMP model for $N = 3$ nodes after condensing rows 11-16.

$$p_{\mathrm{L2},i} = \begin{cases} \frac{\pi_{11}}{\pi_{\mathrm{L2}}} p_{11,i} & i = 1,\ldots,4 \\ \frac{1}{\pi_{\mathrm{L2}}}(\pi_{12} p_{12,i} + \pi_{13} p_{13,i}) & i = 5,\ldots,10 \end{cases} \tag{B.7}$$

The transition probabilities $p_{\mathrm{L2,L2}}$ and $p_{\mathrm{L2,L3}}$ go from an aggregated state to another aggregated state, so their transition probabilities are now fixed and will no longer change; they have been highlighted in green in the probability transition matrix shown in Figure B.4.

B.4. Third Aggregation Step: States 5-10

We aggregate states 5-10 into a single state L1 which encodes 1 link in the graph. L1 will have a steady state probability given by

$$\pi_{\mathrm{L1}} = \sum_{i=5}^{10} \pi_i. \tag{B.8}$$

	1	2	3	4	L1	L2	L3
1	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$			
2					1		
3						1	
4							1
5	$p_{5,1}$	$p_{5,2}$	$p_{5,3}$	$p_{5,4}$			
6					1		
7					1		
8						1	
9						1	
10							1
L2	$p_{L2,1}$	$p_{L2,2}$	$p_{L2,3}$	$p_{L2,4}$	$(\pi_{12}+\pi_{13})/\pi_{L2}$	$(\pi_{14}+\pi_{15})/\pi_{L2}$	π_{16}/π_{L2}
L3	$p_{L3,1}$	$p_{L3,2}$	$p_{L3,3}$	$p_{L3,4}$	π_{18}/π_{L3}	π_{19}/π_{L3}	π_{20}/π_{L3}

Figure B.5: Probability transition matrix of the MMP model for $N = 3$ nodes after condensing columns 5-10.

We first take the row sum of columns 5-10 resulting in Figure B.5. The transition probability $p_{\mathrm{L2,L1}}$ follows from (B.7):

$$\begin{aligned} p_{\mathrm{L2,L1}} &= \sum_{i=5}^{10} p_{\mathrm{L2},i} \\ &= \sum_{i=5}^{10} \frac{1}{\pi_{\mathrm{L2}}} (\pi_{12} p_{12,i} + \pi_{13} p_{13,i}) \\ &= \frac{1}{\pi_{\mathrm{L2}}} \left(\pi_{12} \sum_{i=5}^{10} p_{12,i} + \pi_{13} \sum_{i=5}^{10} p_{13,i} \right) \\ &= \frac{\pi_{12} + \pi_{13}}{\pi_{\mathrm{L2}}}. \end{aligned} \tag{B.9}$$

The transition probability $p_{\mathrm{L3,L1}}$ follows from (B.4):

$$\begin{aligned} p_{\mathrm{L3,L1}} &= \sum_{i=5}^{10} p_{\mathrm{L3},i} \\ &= \sum_{i=5}^{10} \frac{\pi_{18}}{\pi_{\mathrm{L3}}} p_{18,i} \\ &= \frac{\pi_{18}}{\pi_{\mathrm{L3}}} \sum_{i=5}^{19} p_{18,i} \\ &= \frac{\pi_{18}}{\pi_{\mathrm{L3}}}. \end{aligned} \tag{B.10}$$

	1	2	3	4	L1	L2	L3
1	$p_{1,1}$	$p_{1,2}$	$p_{1,3}$	$p_{1,4}$			
2					1		
3						1	
4							1
L1	$p_{L1,1}$	$p_{L1,2}$	$p_{L1,3}$	$p_{L1,4}$	$(\pi_6+\pi_7)/\pi_{L1}$	$(\pi_8+\pi_9)/\pi_{L1}$	π_{10}/π_{L1}
L2	$p_{L2,1}$	$p_{L2,2}$	$p_{L2,3}$	$p_{L2,4}$	$(\pi_{12}+\pi_{13})/\pi_{L2}$	$(\pi_{14}+\pi_{15})/\pi_{L2}$	π_{16}/π_{L2}
L3	$p_{L3,1}$	$p_{L3,2}$	$p_{L3,3}$	$p_{L3,4}$	π_{18}/π_{L3}	π_{19}/π_{L3}	π_{20}/π_{L3}

Figure B.6: Probability transition matrix of the MMP model for $N = 3$ nodes after condensing rows 5-10.

We then weight each row 5-10 by the corresponding steady state probability, take the column sum, and divide the result by the total steady state probability of states 5-10. This results in the transition matrix shown in Figure B.6 and yields the following transition probability for $p_{\mathrm{L1},i}$ for $i = 1,...,4$:

$$p_{\mathrm{L1},i} = \frac{\pi_5}{\pi_{\mathrm{L1}}} p_{5,i}. \tag{B.11}$$

The transition probabilities $p_{\mathrm{L1,L1}}$, $p_{\mathrm{L1,L2}}$, and $p_{\mathrm{L1,L3}}$ go from an aggregated state to another aggregated state, so their transition probabilities are now fixed and will no longer change; they have been highlighted in green in the probability transition matrix shown in Figure B.6.

B.5. Fourth Aggregation Step: States 1-4

We aggregate states 1-4 into a single state L0 which encodes 0 links in the graph. L0 will have a steady state probability given by

$$\pi_{\mathrm{L0}} = \sum_{i=1}^{4} \pi_i. \tag{B.12}$$

We first take the row sum of columns 1-4 resulting in (B.7). The transition probability $p_{\mathrm{L1,L0}}$ follows from (B.11):

	L0	L1	L2	L3
1	1			
2		1		
3			1	
4				1
L1	π_5/π_{L1}	$(\pi_6+\pi_7)/\pi_{L1}$	$(\pi_8+\pi_9)/\pi_{L1}$	π_{10}/π_{L1}
L2	π_{11}/π_{L2}	$(\pi_{12}+\pi_{13})/\pi_{L2}$	$(\pi_{14}+\pi_{15})/\pi_{L2}$	π_{16}/π_{L2}
L3	π_{17}/π_{L3}	π_{18}/π_{L3}	π_{19}/π_{L3}	π_{20}/π_{L3}

Figure B.7: Probability transition matrix of the MMP model for $N = 3$ nodes after condensing columns 1-4.

$$\begin{aligned} p_{\mathrm{L1,L0}} &= \sum_{i=1}^{4} p_{\mathrm{L1},i} \\ &= \sum_{i=1}^{4} \frac{\pi_5}{\pi_{\mathrm{L1}}} p_{5,i} \\ &= \frac{\pi_5}{\pi_{\mathrm{L1}}} \sum_{i=1}^{4} p_{15,i} \\ &= \frac{\pi_5}{\pi_{\mathrm{L1}}}. \end{aligned} \tag{B.13}$$

The transition probability $p_{\mathrm{L2,L0}}$ follows from (B.7):

$$\begin{aligned} p_{\mathrm{L2,L0}} &= \sum_{i=1}^{4} p_{\mathrm{L2},i} \\ &= \sum_{i=1}^{4} \frac{\pi_{11}}{\pi_{\mathrm{L2}}} p_{11,i} \\ &= \frac{\pi_{11}}{\pi_{\mathrm{L2}}} \sum_{i=1}^{4} p_{11,i} \\ &= \frac{\pi_{11}}{\pi_{\mathrm{L2}}}. \end{aligned} \tag{B.14}$$

The transition probability $p_{\mathrm{L3,L0}}$ follows from (B.4):

$$\begin{aligned} p_{\mathrm{L3,L0}} &= \sum_{i=1}^{4} p_{\mathrm{L3},i} \\ &= \sum_{i=1}^{4} \frac{\pi_{17}}{\pi_{\mathrm{L3}}} p_{17,i} \\ &= \frac{\pi_{17}}{\pi_{\mathrm{L3}}} \sum_{i=1}^{4} p_{17,i} \\ &= \frac{\pi_{17}}{\pi_{\mathrm{L3}}}. \end{aligned} \tag{B.15}$$

	L0	L1	L2	L3
L0	π_1/π_{L0}	π_2/π_{L0}	π_3/π_{L0}	π_4/π_{L0}
L1	π_5/π_{L1}	$(\pi_6+\pi_7)/\pi_{L1}$	$(\pi_8+\pi_9)/\pi_{L1}$	π_{10}/π_{L1}
L2	π_{11}/π_{L2}	$(\pi_{12}+\pi_{13})/\pi_{L2}$	$(\pi_{14}+\pi_{15})/\pi_{L2}$	π_{16}/π_{L2}
L3	π_{17}/π_{L3}	π_{18}/π_{L3}	π_{19}/π_{L3}	π_{20}/π_{L3}

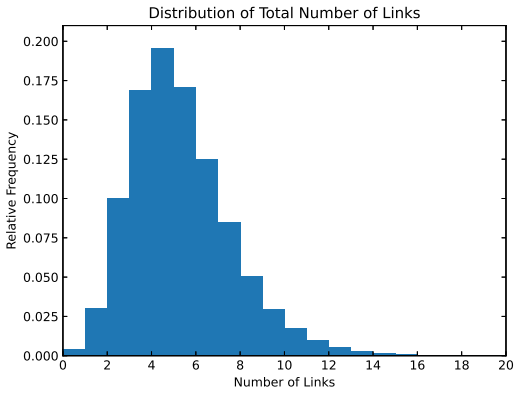
Figure B.8: Probability transition matrix of the MMP model for $N = 3$ nodes after condensing rows 1-4.

We then weight each row 1-4 by the corresponding steady state probability, take the column sum, and divide the result by the total steady state probability of states 1-4. This results in the transition matrix shown in Figure B.8, and the state aggregation is now complete.

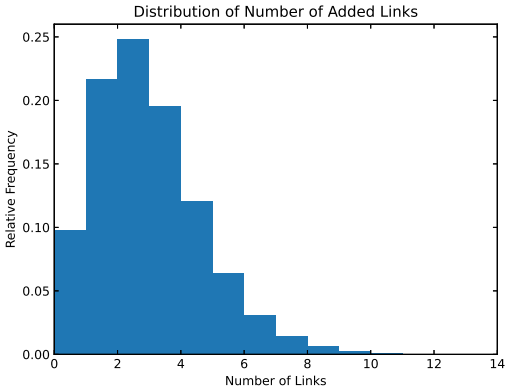
C

Additional Link Distribution Plots

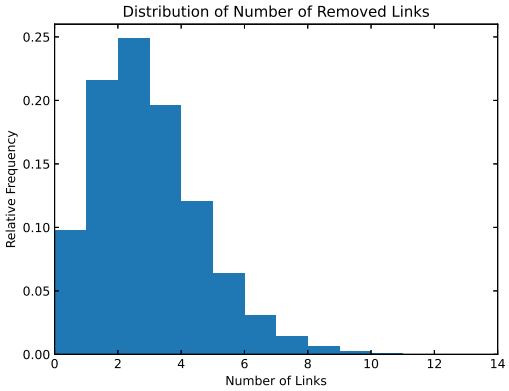
C.1. Link Distributions for the Mobility Process of 9 Nodes



(a) Distribution of the total number of links at each time step (bin width 1).



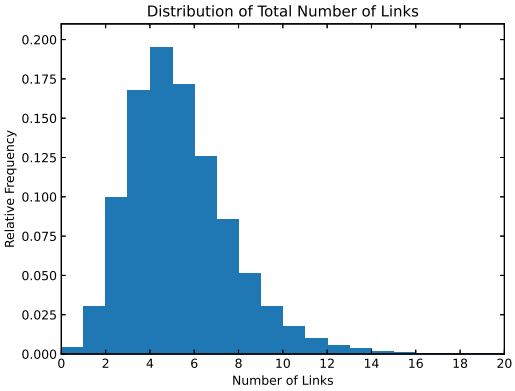
(b) Distribution of the number of links added between two time steps (bin width 1).



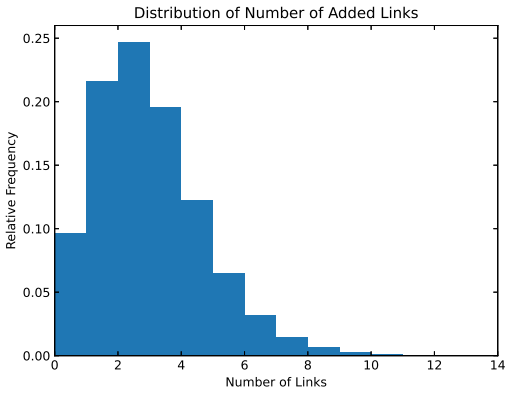
(c) Distribution of the number of links removed between two time steps (bin width 1).

Figure C.1: Histograms of link statistics for the simulated mobility process with 9 nodes across 10,000 trials of 1,000 time steps.

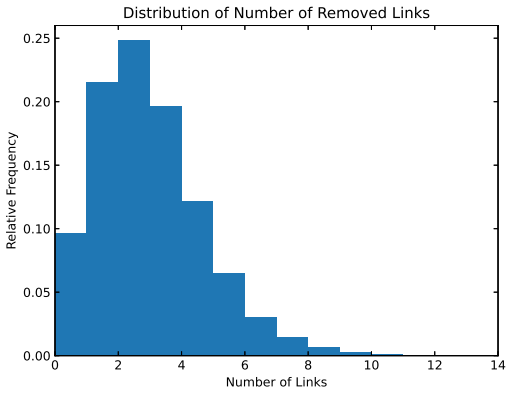
C.2. Link Distributions for the Unaggregated MMP Model of 9 Nodes



(a) Distribution of the total number of links at each time step (bin width 1).



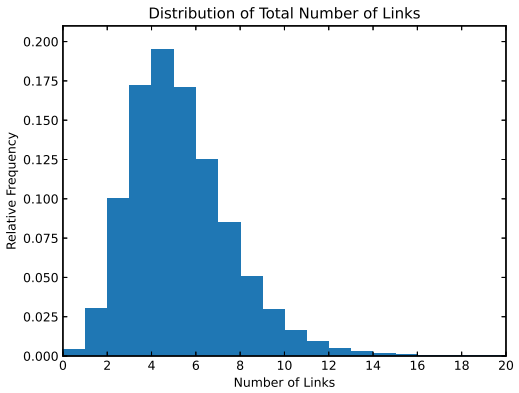
(b) Distribution of the number of links added between two time steps (bin width 1).



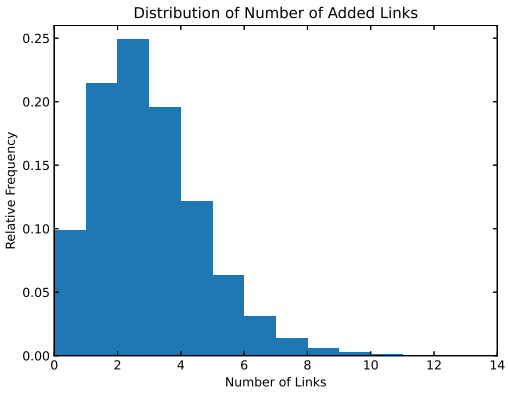
(c) Distribution of the number of links removed between two time steps (bin width 1).

Figure C.2: Histograms of link statistics for the unaggregated MMP model across 100 trials of 1,000 time steps.

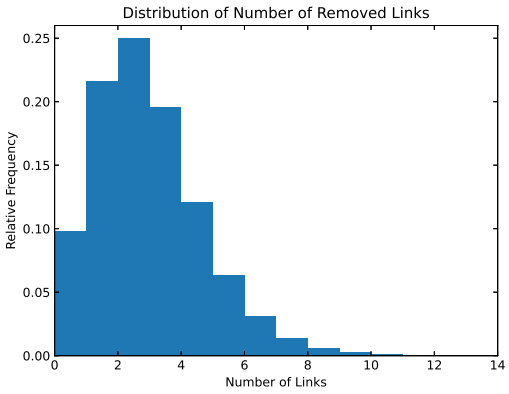
C.3. Link Distributions for the Aggregated MMP Model of 9 Nodes



(a) Distribution of the total number of links at each time step (bin width 1).



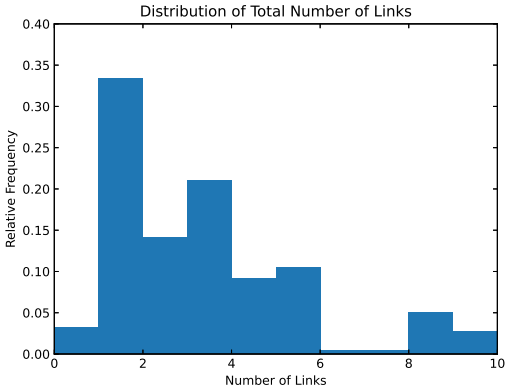
(b) Distribution of the number of links added between two time steps (bin width 1).



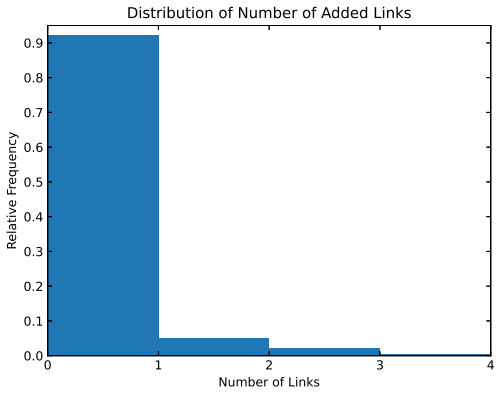
(c) Distribution of the number of links removed between two time steps (bin width 1).

Figure C.3: Histograms of link statistics for the aggregated MMP model across 100 trials of 1,000 time steps.

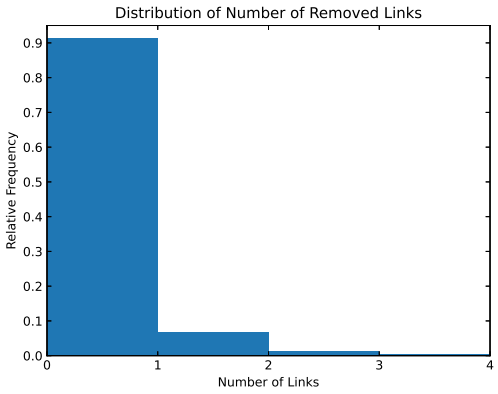
C.4. Link Distributions for the Real Mobility Process



(a) Distribution of the total number of links at each time step (bin width 1).



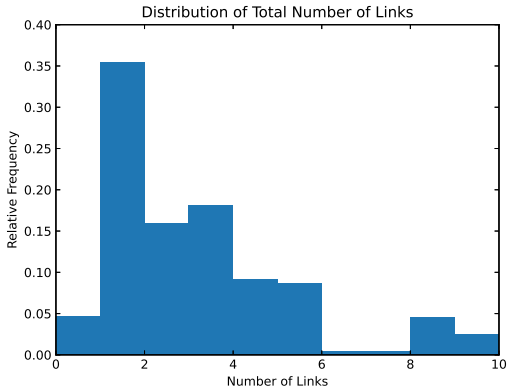
(b) Distribution of the number of links added between two time steps (bin width 1).



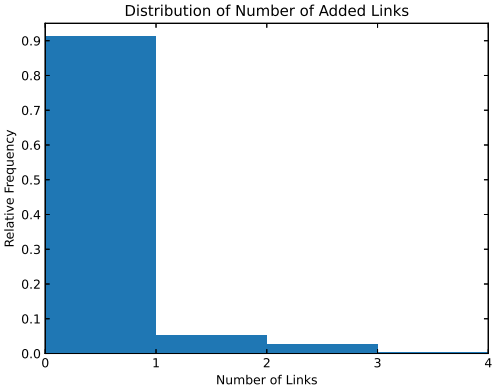
(c) Distribution of the number of links removed between two time steps (bin width 1).

Figure C.4: Histograms of link statistics for the real mobility process.

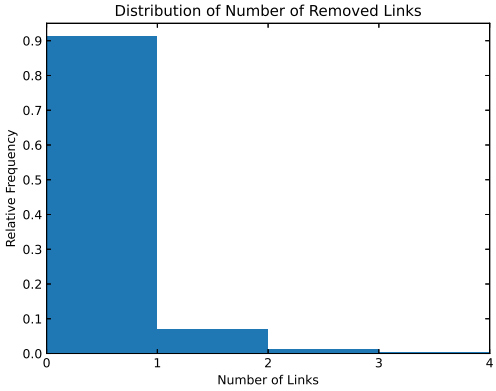
C.5. Link Distributions for the Aggregated MMP Model of the Real Mobility Process



(a) Distribution of the total number of links at each time step (bin width 1).



(b) Distribution of the number of links added between two time steps (bin width 1).



(c) Distribution of the number of links removed between two time steps (bin width 1).

Figure C.5: Histograms of link statistics for MMP model based on the real mobility process.

D

Additional SIR Simulation Results

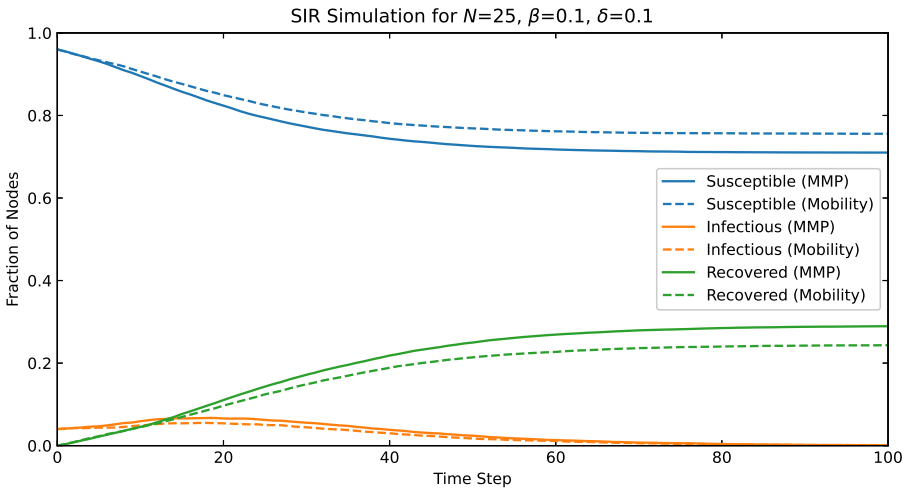


Figure D.1: SIR simulation results for $\beta = 0.1$, $\delta = 0.1$.

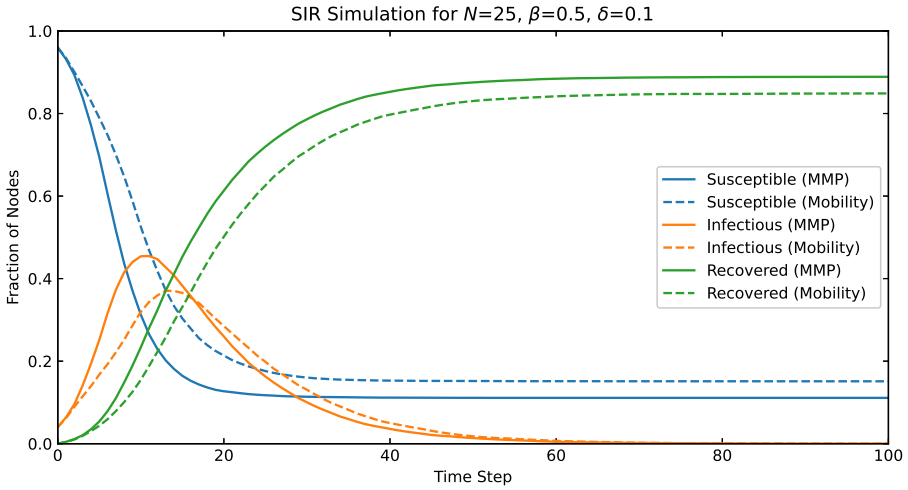


Figure D.2: SIR simulation results for $\beta = 0.5$, $\delta = 0.1$.

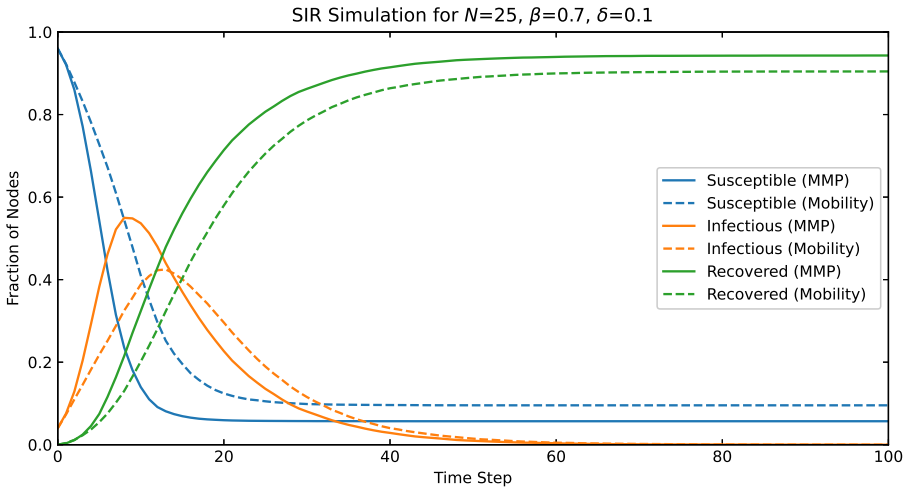


Figure D.3: SIR simulation results for $\beta = 0.7$, $\delta = 0.1$.

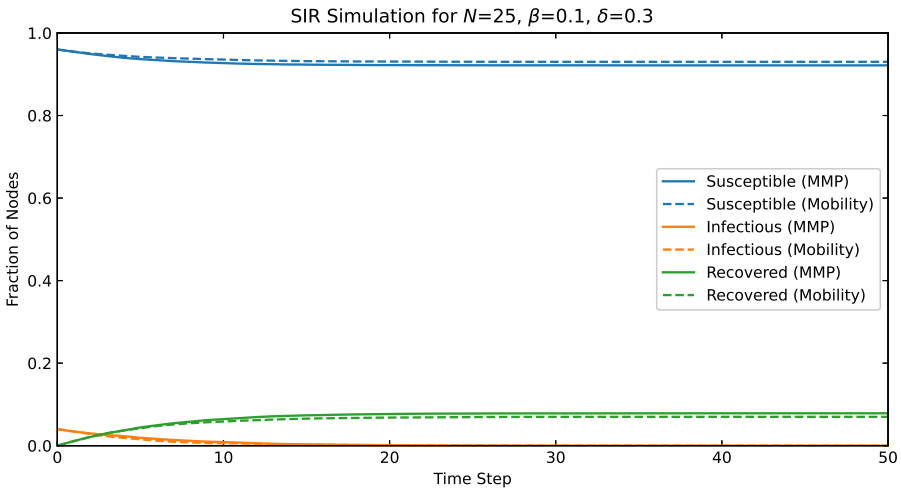


Figure D.4: SIR simulation results for $\beta = 0.1$, $\delta = 0.3$.

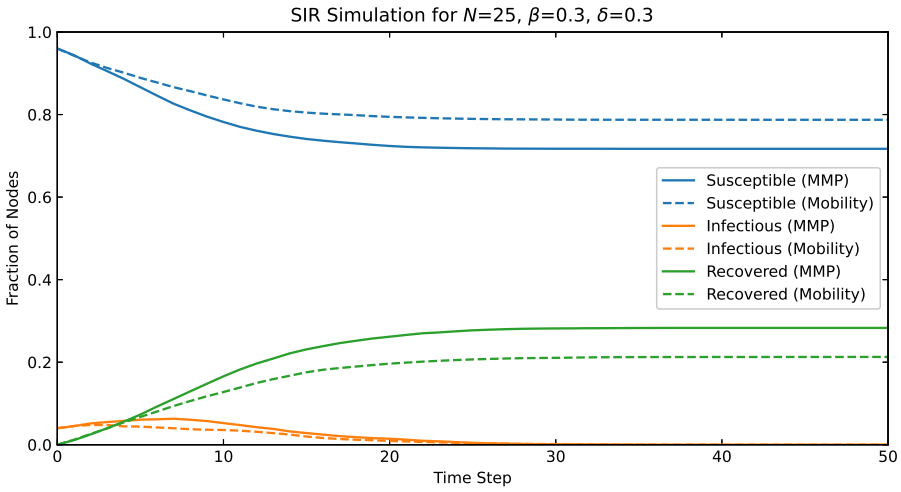


Figure D.5: SIR simulation results for $\beta = 0.3$, $\delta = 0.3$.

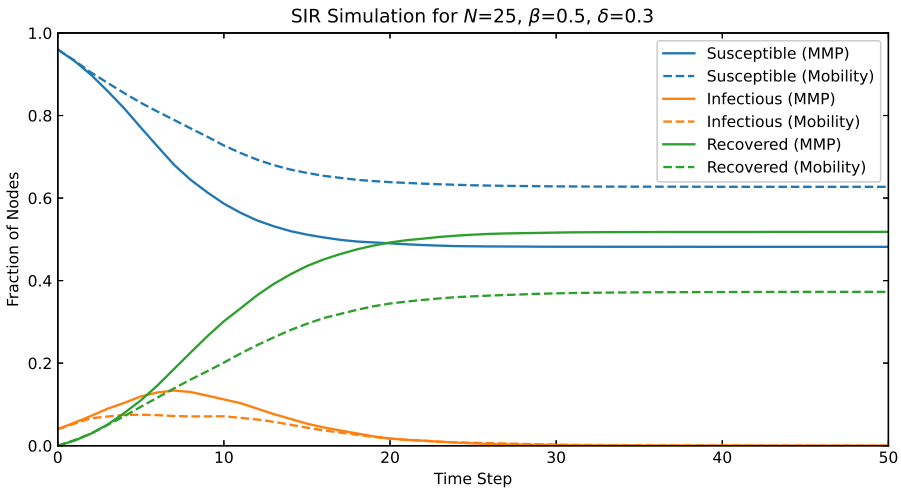


Figure D.6: SIR simulation results for $\beta = 0.5$, $\delta = 0.3$.

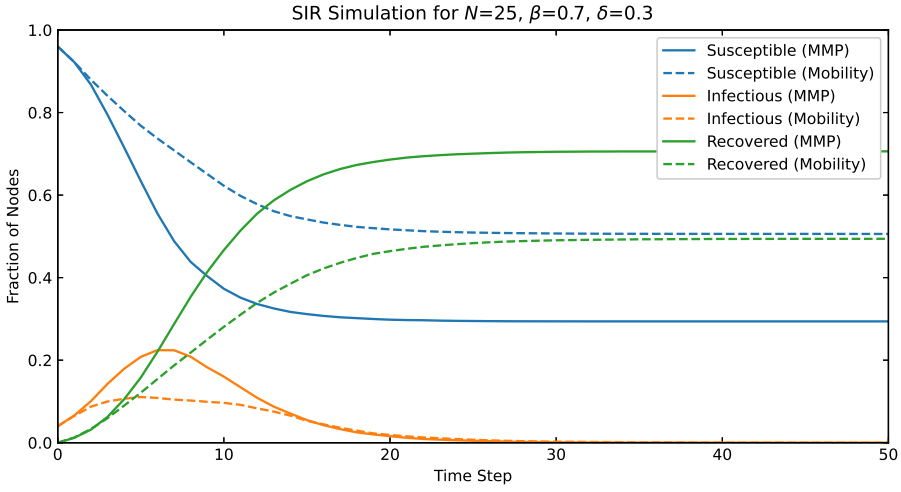


Figure D.7: SIR simulation results for $\beta = 0.7$, $\delta = 0.3$.