

**AN OPERATIONS RESEARCH APPROACH FOR
THE SINK NODE ANONYMITY IN A WIRELESS
SENSOR NETWORK**

BY

MOHAMMED ADLY KHALIL ALSAAFIN

A Thesis Presented to the
DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

INDUSTRIAL AND SYSTEMS ENGINEERING

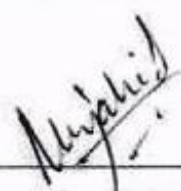
DECEMBER 2022

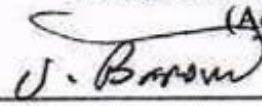
KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

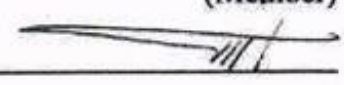
DHAHRAN- 31261, SAUDI ARABIA

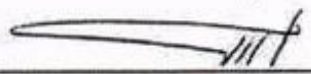
DEANSHIP OF GRADUATE STUDIES

This thesis, written by **Mohammed Al Saafin** under the direction of his thesis advisor and approved by his thesis committee, has been presented and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN INDUSTRIAL & SYSTEMS ENGINEERING**


Dr. Syed N. Mujahid
(Advisor)


Dr. Uthman Baroudi
(Member)


Dr. Anas Alghazi
(Member)


Dr. Anas Al-Ghazi
Department Chairman


Dr. Suliman S Al-Homidan
Dean of Graduate Studies

Date



© **Mohammed Al Saafin**

2022

Dedication

I dedicate this thesis to my parents who played a huge role in my life and success. Thank you for all the effort that you did in bringing me up to be a better person, your continuous motivation, your endless support, your continuing love, and encouragement to pursue higher education. I really hope this accomplishment will make you proud and fulfill part of the path that you've envisioned for me.

ACKNOWLEDGMENT

I'd like to thank my thesis advisor Dr. Syed Mujahid and express my appreciation for his guidance, support, perfect insight, and the time he dedicated to help me in the preparation and implementation of this thesis. Specifically, the idea of this thesis, model formulation and execution. I thank the committee members, Dr. Uthman Baroudi and Dr. Anas Alghazi for their valuable insight and support. I appreciate KFUPM and the industrial and systems engineering department for giving me this opportunity. I extend my appreciation to my graduate advisor Dr. Salih Duffuaa that guided and supported me heavily in my master's journey.

TABLE OF CONTENTS

ACKNOWLEDGMENT	v
LIST OF TABLES.....	vii
LIST OF FIGURES.....	viii
ABSTRACT	x
ملخص الرسالة.....	xi
THESIS OBJECTIVE	1
CHAPTER 1: INTRODUCTION.....	2
CHAPTER 2: LITERATURE REVIEW	8
CHAPTER 3: PROBLEM DEFINITION	16
CHAPTER 4: WSN MATHEMATICAL MODEL	17
4.1 Base Model Formulation	17
4.1.1 Indices Set.....	17
4.1.2 Base Model Parameters.....	18
4.1.3 Base Model Variables	18
4.1.4 Base Model Constraints.....	19
4.1.5 Base Model Objective Function	23
4.2 Extended model #1	23
4.2.1 Extended Model #1 Constraints.....	24
4.2.2 Extended Model #1 Objective Function.....	27
4.3 Extended model #2	28
4.3.1 Extended Model #2 Constraints.....	29
4.3.2 Modified model #2 Objective Function.....	33
CHAPTER 5: MODEL EXECUTION	34
5.1. GAMS Code Execution:	34
CHAPTER 6: RESULTS AND FINDINGS	37
6.1. Case Study of 21 nodes network:	37
6.1.1. Results for the 21 nodes using the base model:.....	38
6.1.2. Results for the 21 nodes Y-model:	44
6.1.3. Results for the 21 nodes V-model:	50
6.2. Results of 30 nodes base model network:	58
6.2.1. Sensitivity analysis of the 30 nodes base model	67
6.2.2. Results of the 30 nodes Y-model:	70
6.2.3. Results of the 30 nodes V-model:	76
CHAPTER 7: CONCLUSION.....	83
References.....	84
Vitae	87

LIST OF TABLES

Table 1:WSN Literature summary	14
Table 2:Cont. WSN Literature summary	15
Table 3:Placement of Nodes in Grid g.....	35
Table 4: Sample Feasibility Matrix.....	35
Table 5:Sample Listening nodes.....	36
Table 6:Sample Cost Matrix.....	36
Table 7:Summary of the models	38
Table 8: Number of Equations.....	82
Table 9:Number of Variables	82

LIST OF FIGURES

Figure 1: Optimal Shortest Path Value (BM-21).....	38
Figure 2: The Shortest Path (BM-21).....	39
Figure 3: Path Chart (BM-21).....	39
Figure 4: Beta Optimal Value (BM-21)	40
Figure 5:Beta Chart (BM-21)	40
Figure 6: tmax Optimal Value (BM-21)	41
Figure 7: The Path when tmax is minimized (BM-21)	41
Figure 8: Tmax Chart (BM-21)	42
Figure 9: The route for run 4 (BM-21).....	43
Figure 10 : Final Chart (BM-21).....	43
Figure 11: Variable Results (BM-21)	44
Figure 12: Optimal Shortest Path Value (M1-21)	44
Figure 13: The Shortest Path(M1-21)	45
Figure 14:Path Chart (M1-21)	46
Figure 15: Beta Optimal Value (M1-21).....	46
Figure 16:Beta Chart (M1-21).....	47
Figure 17: tmax Optimal Value (M1-21)	47
Figure 18: The Path when tmax is minimized (M1-21).....	48
Figure 19:Tmax Chart (M1-21)	48
Figure 20:Final Chart (M1-21)	49
Figure 21: The route for run 4(M1-21)	49
Figure 22:Variable Results (M1-21)	50
Figure 23:Optimal Shortest Path Value (M2-21)	51
Figure 24:The Shortest Path (M2-21)	51
Figure 25:Path Chart (M2-21)	52
Figure 26: Beta Optimal Value (M2-21).....	53
Figure 27: Sample Route (M2-21).....	54
Figure 28:Beta Chart (M2-21).....	54
Figure 29: tmax Optimal Value (M2-21)	55
Figure 30:Tmax Chart (M2-21)	55
Figure 31:Final Chart (M2-21)	56
Figure 32: The route for run 4(M2-21)	57
Figure 33:Variable Results (M2-21)	58
Figure 34: Optimal Shortest Path Value (BM-30).....	58
Figure 35: The Shortest Path (BM-30).....	59
Figure 36 Path Chart (BM-30).....	59
Figure 37: Beta Optimal Value (BM-30)	60
Figure 38 Beta Chart (BM-30)	60
Figure 39: tmax Optimal Value (BM-30).....	61
Figure 40: Sample Path when tmax is minimized (BM-30).....	61
Figure 41:Tmax chart (BM-30)	62
Figure 42: Optimal Shortest Path Value (BM-30).....	62

Figure 43:Variable results (BM-30).....	63
Figure 44: Final chart (BM-30)	63
Figure 45: The shortest path (BM-30)	64
Figure 46:Sample Alpha Values	65
Figure 47:Variable Values (beta=2).....	65
Figure 48:Sample Alpha Values 1	66
Figure 49:Pareto Front	67
Figure 50:Surface plot	69
Figure 51: Optimal Shortest Path Value (M1-30)	70
Figure 52:Shortest path chart(M1-30).....	70
Figure 53: The Shortest Path (M1-30)	71
Figure 54: Beta Optimal Value (M1-30).....	72
Figure 55:Beta Chart (M1-30).....	72
Figure 56: tmax Optimal Value (M1-30)	73
Figure 57: Sample Path when tmax is minimized (M1-30)	73
Figure 58:Tmax Chart (M1-30).....	74
Figure 59:Optimal Shortest Path details (M1-30).....	74
Figure 60: Final chart (M1-30).....	75
Figure 61: Sample route of run 4(M1-30)	75
Figure 62:Variable Results (M1-30).....	76
Figure 63: Optimal Shortest Path Value (M2-30)	76
Figure 64:Shortest path chart(M2-30).....	77
Figure 65: Beta Optimal Value (M2-30).....	77
Figure 66:Beta Chart (M2-30).....	78
Figure 67: tmax Optimal Value (M2-30).....	78
Figure 68: Tmax Chart (M2-30)	79
Figure 69:Optimal Shortest Path details (M2-30).....	79
Figure 70: Final chart (M2-30).....	80
Figure 71: Variable Results (M2-30)	80

ABSTRACT

Full Name : [Mohammed A. K. Al Saafin]
Thesis Title : [An Operations Research Approach for the Sink Node Anonymity in a Wireless Sensor Network]
Major Field : [Industrial and Systems Engineering (ISE)]
Date of Degree : [December 2022]

Wireless Sensor Networks (WSNs) have attracted significant interest from research community owing to its wide array of applications such as border defense, security surveillance, disaster warning and health care. WSN operates frequently in unattended environment and the network is vulnerable to attacks. This study focuses on protecting the sink node from attackers. We propose and develop a novel Mixed Integer Programming (MIP) model on protecting the sink node in a WSN with several source nodes. General Algebraic Modeling System (GAMS) which is a high-level modeling system for mathematical optimization software program was used with the built in CPLEX solver to tackle and solve the optimization models considered in this thesis. Phantom routing mechanism is developed as a location privacy protection method for the sink node to increase its anonymity. Moreover, the trade-off between the energy consumption and the location privacy protection of the sink node is studied. Different instances of the data are tested to check the validity of the base and extended models.

ملخص الرسالة

الاسم الكامل: محمد عدلي السعافين

عنوان الرسالة : نهج بحث العمليات لإخفاء هوية المحطة الأساسية في شبكة الاستشعار اللاسلكية

التخصص: هندسة النظم الصناعية

تاريخ الدرجة العلمية: ديسمبر 2022

جذبت شبكات الاستشعار اللاسلكية اهتمامًا كبيرًا من الباحثين نظرًا لمجموعتها الواسعة من التطبيقات مثل الدفاع عن الحدود والمراقبة الأمنية والتحذير من الكوارث والرعاية الصحية. غالبًا ما تكون شبكات الاستشعار اللاسلكية في بيئة قاسية وهذا يعرض الشبكة للهجمات الإلكترونية. تركز هذه الدراسة على حماية المحطة الأساسية من المهاجمين. هذه هي المحاولة الأولى لتطوير نموذج برمجة عدد صحيح مختلط (MIP) لحماية المحطة الأساسية في شبكة الاستشعار اللاسلكية مع عدة عقد مصدر. تم استخدام نظام النمذجة الجبرية العامة (GAMS) وهو نظام نمذجة عالي المستوى للتحسين الرياضي مع محلل CPLEX المدمج لمعالجة وحل نماذج التحسين التي تم النظر فيها. تم تطوير آلية التوجيه الوهمية كطريقة لحماية خصوصية الموقع للمحطة الأساسية مما يزيد من إخفاء هويتها. علاوة على ذلك ، تتم دراسة المفاضلة بين استهلاك الطاقة وحماية خصوصية الموقع للمحطة الأساسية. سيتم اختبار مثيلات مختلفة من البيانات للتحقق من صحة النماذج الأساسية والممتدة.

THESIS OBJECTIVE

The objective of this thesis is to develop a mathematical modeling approach for the sink node anonymity in Wireless Sensor Networks (WSN). The thesis will consider multiple objectives concerning the route length, traffic balance and total time needed to send all the packets of the network. The primary goal will be to achieve higher anonymity of the sink node, decrease the total time of the network and lower energy consumption. Exact solution techniques will be used to solve the suggested WSN formulation. Numerical examples will be studied to validate the proposed WSN mathematical model(s).

CHAPTER 1: INTRODUCTION

In the past years, WSNs have attracted significant researchers' interest owing to its wide array of applications such as border defense, security inspection, disaster warning, health care and object tracking. WSN is frequently comprised of multiple sensors that are arbitrarily distributed in the monitored environment to gather data then send the gathered data to the sink node by multi-hop routing. Multi-hop communications are used in the WSNs because of the small transmission range of each sensor node, this will lead to energy saving and therefore increase the lifespan of the entire network. The sink node collects and process the gathered data and sends it to the remote users for various needs. The sensor nodes are small devices that have limited recourses and lifespan. It is not necessary for every sensor node to directly communicate with the sink node. The sensor node consists of different parts such as microcontroller, transceiver, power source and one or multiple sensors that are used to capture data in the monitored environment.

Sensor nodes will send a packet when they detect some event in the monitored environment. WSNs are installed in tough and hostile locations and the concerned channel is accessible which will cause the network to be vulnerable to attacks that threatens the location privacy of the nodes but most importantly the sink node. The anonymity of the sink node is extremely important to protect it from the adversary's attack. In this work, a novel mixed integer linear programming model (MILP) is developed for a WSN with an objective of minimizing the total traffic in the network to protect the sink node from malicious attacks. The goal of the study is to protect the sink node from different attacks by using privacy protection mechanisms such as phantom routing and decrease the energy consumption of the network (since high energy consumption is not appropriate for WSNs) while finding the shortest path of the packets to flow from the sources to the sink node.

Mathematical modelling is used to translate problems from an application area into tractable mathematical formulations whose analysis will help provide answers to the concerned application. It falls under a term called operations research (OR) which deals with the applying of advanced analytical methods to enhance better decision making. Mathematical modelling is used in many complex decision-making problems and in various areas such as artificial intelligence, computer science, finance, electrical engineering, and neuroscience.

Different techniques were used in the literature to tackle problems in the computer science and telecommunications areas. In this thesis will use a novel mathematical model to tackle the problem of sink node anonymity in WSNs. Mathematical model formulations consist of objective functions, decision variables and constraints. The objective functions can contain linear, nonlinear or a mixture of both terms.

Mostly, the objective function of the model will either maximize the profit or minimize the cost and time. The decision variables can contain continuous, discrete or a mixture of both decision variables. The constraints are the part which the optimization problem must satisfy. There are three types of constraints which are equality constraints, inequality constraints and integer constraints. These constraints can be linear, nonlinear or a mixture of both in the complete mathematical formulation.

A brief overview of mathematical programming methods will be discussed in this section. The most basic type is Linear Programming (LP) which consists of linear objective function and constraints, but it deals with continuous variables only. There are different types of programming types that are quite like LP such as Integer programming (IP) and Mixed integer programming (MIP). The difference between IP and LP is only in the variables which are continuous in the LP

methods but are integers in IP. Mixed integer programming is a mixture of both LP and IP regarding the variables. MIP can have mixed variables in the formulation which consists of some integer and continuous variables.

LP problems can take many forms such as standard inequality form and the standard equality form with maximization or minimization of the objective function. We will consider the standard inequality form with minimization of the objective function and the general formulation of the LP.

Standard Inequality form

$$\text{Minimize } c^T x$$

Subject to

$$Ax \geq b$$

$$x \geq 0$$

General form

$$\text{Minimize } \sum_{j=1}^N c_j x_j$$

Subject to

$$\sum_{j=1}^N \alpha_{ij} x_j \geq b_i \quad i = 1, \dots, m_1$$

$$x_j \geq 0 \quad j = 1, \dots, n_1$$

Where A is an $m \times n$ matrix, c is constant n -vector (cost coefficient), b is an m -vector and m & n are scalars.

The second type of mathematical programming is IP. It differs from LP that all the variables are integer variables. It can be expressed in the general form as follows:

$$\text{Minimize } \sum_{j=1}^N c_j x_j$$

Subject to

$$\sum_{j=1}^N a_{ij} x_j \geq b_i \quad i = 1, \dots, m_1$$

$$x_j \geq 0 \quad j = 1, \dots, n_1$$

$$x_j \in \mathbb{Z}^n \quad j = 1, \dots, n_1$$

As seen from the above general form, the only difference is that the x variable belongs to the integer variables set.

We can generalize this further and take the case of MIP, where the variables will be mixed between integer and continuous. The objective function and the constraints won't change but the restriction on the variables won't be consisting of only integer variables but continuous variables too.

There is a special case of integer variables and that is binary variables. These variables have 2 options to take, and they are either 0 or 1 as value.

$$x_j \in \{0, 1\} \quad j = 1, \dots, n_1$$

Integer programming and mixed integer programming problems can be solved using various algorithms such as exact algorithms or heuristic methods when needed since integer linear programming models are NP-hard. One of the most popular algorithms to solve IP and MIP is branch and bound method and its variants.

There are multiple reasons that integer variables are used in modelling problems but there are two main and important reasons to translate real life applications and problems to a mathematical model. The first reason is that you can't represent quantities as a continuous variable in the model for example if you want to have a mathematical model that calculated the optimal number of pipes needed in a production line, the number of pipes can't be for example 17.5 pipes, but it can be either 17 or 18 pipes and that is an important usage of integer variables in mathematical models. The second reason is concerned with binary variables and their usage in representing decisions (e.g., whether a factory is producing units in a specific time). If it's producing, then the variable should be 1 and 0 otherwise.

The third type is non-linear programming problems which is abbreviated as NLP and can be expressed as follows:

Minimize $f(x)$

Subject to

$$g_i(x) \leq 0 \quad \forall i = 1, \dots, p$$

$$h_j(x) = 0 \quad \forall j = 1, \dots, q$$

$$x \in X$$

Where $f(x)$ is the objective function,

$g_i(x)$ are the inequality constraints, $h_j(x)$ are the equality constraints and $X \subseteq \mathbb{R}^n$

$f(x), g_i(x)$ and $h_j(x)$ are real valued functions on X with at least one of them being nonlinear.

Knowing that n, p, q are positive integers.

The most important aspect of dealing with NLP is to prove that the nonlinear functions are convex or non-convex. There are two types of problems that concern NLP's which are convex programming and non-convex programming problems. Minimizing a convex function is known as convex mathematical programming model and maximizing a concave function is like minimizing a convex function.

Dealing with NLP is extremely difficult when compared to LP due to various reasons. We can always optimize an LP and calculate the optimal solution since all of LP problems are convex but it's a completely different case when dealing with NLP problems since we don't know that the nonlinear functions are convex or not.

The following sections will discuss the related WSN problems in the literature and propose a novel multi-objective mathematical model that increase the sink node anonymity in a WSN with various number of nodes. The results section will show the findings and discussion of the two networks. The final section will conclude the thesis and include some recommendations for future research.

The objective of this thesis is to propose a multi-objective mathematical modeling approach for the sink node anonymity in WSN.

CHAPTER 2: LITERATURE REVIEW

WSNs are most of the time implemented in hostile and tough environments and their channels is open which make them extremely exposed to attacks from adversaries. The anonymity of important nodes in the network is extremely crucial to the normal functioning of the WSN. For example, the objective of the application in the famous Panda-hunter game model that was first suggested by Ozturk et al., (2004) is to protect the pandas from the attacks of the hunter while taking into consideration the normal functionality of observing the pandas in the network. The hunter's purpose is to try to trace back the packets in the network to reach the source nodes which are close to the protected animals such as the pandas in this case. If the hunter found out the real position of the nodes, then eventually the pandas will get caught, and their position will be compromised. By saying that, we are implying that the location privacy of the nodes is crucial in huge applications of WSNs. Srikant et al., (2003) and Latanya, (2002) developed two algorithms; anonymous and encryption to protect the information security, but they weren't efficient in handling attacks that threatened the location privacy of the nodes. There are great challenges in studying the location privacy algorithms and techniques since different types of attacks from adversaries need specific types of protection algorithms. Two types of attacks were discussed in the literature against the privacy of the location of the sensor nodes in the network. Passive attacks and initiative attacks. In the passive attacks, the adversary observes and investigates the movement to reach the node they are attacking most commonly a source or a sink node. On the other hand, the initiative attacks are much more dangerous to the network since they can manipulate the typical network tasks such as altering packets and damaging source/sink nodes.

The first paper to tackle the problem of location privacy in WSNs is by Ozturk et al, (2004). The popularity of the topic increased in the past years and some algorithms were developed to protect against a local attacker, for example see George et al., (2012). Some major problems in protecting the location in a WSN was studied by Liang et al., (2013). A portion of the papers in the past years studied the protection of the location of some important nodes in the network such as source nodes by Conti et al., (2013) or merged algorithms to protect both sink and source nodes in Dhivya et al., (2014). The location privacy protection techniques can be listed as three types which are source location privacy, sink node location privacy and location privacy protection for both the sink and source nodes. Different mechanisms were proposed and developed in the literature to tackle the source location privacy protection problem. Phantom routing was first suggested by Ozturk et al., (2004) which is a mechanism that choose some selected phantom nodes and some packets from the resource nodes, then these packets are sent to the phantom nodes and after that the packets are transmitted to the base station through strategies such as single path routing or flooding. Ozturk's model was in two phases which are the random walk stage which is sending packets from the source node in an arbitrary manner then flooding is the 2nd stage, which is that the packets are flooded to the sink node using a specific strategy. The energy of the flooding stage is quite high for WSNs, so Kamat et al. (2005) suggested phantom single path routing (PSPR) which instead of flooding the packets it transferred through the shortest paths and this attains a tradeoff between the usage of energy and the safety of the concerned location. Some extensions to the proposed model were done by Xi et al., (2006) which created a new algorithm named greedy random walk. Another mechanism was developed in 2009 in the paper by Kang, (2009) which suggested the fake packets injection mechanism, this mechanism sends a fake packet each time a source node senses a message. In the paper by Tang et al. (2015) a random walk mechanism was developed which

clustered the network into small parts while picking the shortest path strategy in sending the packets to reduce the energy utilization then the shortest path is altered to routing path to increase the anonymity of the source nodes. Kazatzopoulos et al., (2006) created a brand-new method called information holding in disturbing environment which create routes named bus and the upper end of these new routes is the base station. A merge of the phantom routing and face packet injection is suggested in the paper of Mahmoud and Shen, (2012) which has the strategy of sending dummy packets to fake sources to baffle the attacker. Multi-path routing mechanism was used by Wang et al., (2009) where they developed a random parallel route that sends packets to disjoint paths. Some other mechanisms were developed such as data mule mechanism and cryptography mechanism by Li et al. (2012,2014). There are advantages and disadvantages of the previous mechanisms. For example, the phantom routing technique to protect the source nodes is good for local attacks but it's not efficient to handle global attacks. Rimjhim et al., (2020) proposed a method to choose the places of phantom sources or ingenuine ones in the WSN. The method applies the idea of geometry for choosing the location of the false and phantom nodes. Lilian et al., (2020) proposed two fake packet routing protocols and two new phantom-based source location privacy routing protocols. There are two phases of confusion to the attacker when they perform traffic analysis strikes on the routes. Wang et al., (2019) proposed two new grids based SLP in WSNs which uses a strong base station to help the source nodes to decide a phantom node candidate set from which the source node randomly chooses a phantom node behaving as a false source node.

The second category of location privacy protection is the protection of an important node in the WSNs which is the sink node. The sink node is used to collect data from the entire network and send it to remote users to satisfy specific needs. If the attacker identifies the location of the base station, then they can launch a devastating attack that will cripple the functionality of the WSN or

get their hands on sensitive information and will cause outrageous losses. As mentioned before WSNs are usually in tough and hostile environments and their wireless channels are open. Protecting the sink node by increasing its anonymity is one of the crucial tasks when handling WSN's. The sink node protection criteria when concerned with the place of the nodes can be categorized into five types. Random walk, ring routing, phantom routing, fake packet injection and phantom routing merged with fake packets mechanism. In the random walk type a random data scheme is suggested in the paper by Ngai et al., (2013) where the packets are transferred from a source node in a random path. These packets are duplicated and stored at arbitrarily chosen nodes on the same route. The base stations in this model move until they reach the nodes that are collecting packets. The benefit of this strategy is that the attackers won't be able to know the base stations movements in the network so it will be hard to locate the sink nodes which will increase the anonymity. The disadvantage of the random walk mechanism is that it has a high energy consumption and minimal data gathering rate. In the paper by Liu et al., (2017) a new mechanism was developed to surmount the downsides of the previous mechanism. Liu designed a semi arbitrary circle forwarding for a flexible sink to improve the rate of the data collection. This mobile sink will move in a predefined way not in a random manner and packets are transferred from the regular sources to the flexible sink. Long et al., (2015) proposed a model that implements many dummy sinks in the WSN to make it tricky for the attacker to discover the genuine base station. The first proposed model that implemented the k-anonymity knowledge to defend the base stations which is formulated as a nonlinear constrained mathematical model was done by Chai et al., (2013) which is categorized in the phantom base station mechanism. The mechanism will create k nodes that are there to hide the genuine sink node and mislead the attackers. Ying et al., (2014) developed a method that will make each node to produce fake packets to stabilize the traffic to mislead the

attackers, since in regular networks the traffic near the base station is greater than the other parts of the system. Chen et al., (2015) proposed a model in the fake packet injection type that deals with a collaborative global attack from multiple attackers, which will share the information they have and generate the global network traffic. They proposed a model named sending rate adjustment that unifies the sending rate of each node in the network and by doing that it balances the traffic in the WSN which will hide the genuine sink spot. Baroutis and Younis, (2017) also suggested another fake packet injection model that balances the traffic in the network which will make it harder for the adversaries to discover the base station, but they studied the tradeoff between location privacy and energy consumption by Corona-based load balanced routing tree that aims to balance the traffic density across the network and to avoid worsening the overall network performance. Yao et al., (2013) used the fake packets and fake sink node with the sole purpose of concealing the position of the sink node. A mixture of real and fake messages was sent to confuse the attacker and increase the time of finding the genuine sink node. A consistent insertion for base station safety with node compromise defense mechanism was developed by Rios et al., (2015) which generates a predefined number of dummy packets to conceal the genuine packets. Chen et al., (2014) proposed a model to defend against an attack on the sink node called based attack scheme in which the attacker can eavesdrop in the network or alter the sensor nodes to collection important information. The proposed protection to this attack is a way of altering the routes to a collection of parent and child sets. Wang et al., (2017) suggested a random walk and fake packets method to hide the real packets while sending fake packets to random locations to hide the sink location. Mehdi et al., (2020) created an approach called formation of illusive void (FIVA) that exploits the geographic routing to mislead the attacker about the topology. FIVA creates a fake

void around the sink node and forward packets to encounter this void. The goal of this method is to exclude the sink node from the routes to mislead the attacker away from the sink node area.

Finally, the 3rd type which is protecting both the sink and the source node simultaneously which is done either in fake packet injection or anonymous communication mechanisms. Chen and Lou, (2015) proposed four methods to deal with protecting the sink and the sources nodes by a combination of sending fake packets or having a proxy sink which will help in hiding the sink node while concealing the location of the source nodes. Lee et al., (2015) developed a steady rate broadcast method where nodes in the system transmit the same size of packets including genuine and fake packets to its neighboring nodes. The method can be efficient against global attacks and compromised nodes. Chen et al., (2012) developed a method named effective unidentified communication for WSN. The sink nodes in the WSN always have an ID for dealing with packets. This method creates anonymous identities which will make it harder for the attacker to classify them as the important nodes. Abuzneid et al., (2015) suggested to use pseudonyms instead of actual IDs to protect both the sink and the source nodes if the attacker examined the id info. A summary of the main WSN problems in the literature is shown in Tables 1 and 2. The papers are classified based on the problem that is tackled and what algorithm was used to solve it.

Different methodologies were used to tackle the issue of the sink node anonymity in a WSN but in this thesis we propose a novel method that uses an operations research approach to tackle it. A multi-objective mixed integer programming mathematical model will be showcased in the following sections with results and discussion to validate the findings.

Year	Author	Problem	Anonymity Criteria
2002	Sweeney, Latanya	Protecting the information security	The k-anonymity protection model
2004	Ozturk et al.	Location privacy	Famous Panda-hunter game model/Phantom Routing
2005	Kamat et al.	Energy of the flooding stage	Phantom single path routing (PSPR)
2006	Xi et al.	Energy of the flooding stage	Greedy random walk (GROW)
2006	Kazatzopoulos et al.	Hiding sources of information in WSN	Information Hiding in Distributing Environments (iHIDE)
2009	Kang et al.	Protecting location privacy in large-scale WSNs	Fake packets injection (FPI)
2009	Wang et al.,	Maximize source location privacy	Phantom Routing with Locational Angle (PRLA)
2012	Mahmoud and Shen	Efficient source location privacy-preserving scheme for WSNs	Merge of the phantom routing and face packet injection (PRFPI)
2012	Li et al	Increase the anonymity of the source nodes	Mules-Saving-Source (MSS)
2012	Chen et al.	Anonymous communication protocol for wireless sensor networks	Effective unidentified communication for WSN (EAC)
2013	Ngai et al.	Providing location privacy for mobile sinks	Random data collection scheme (RDACS)
2013	Chai et al.	Enhancing sink-location privacy	Nonlinear constrained mathematical model
2013	Yao et al.	Concealing the position of the sink node	Fake packets and the fake base station (FPFBS)
2014	Dhivya et al.	Protecting both sink and source nodes	Merged algorithm

Table 1: WSN Literature summary

Year	Author	Problem	Anonymity Criteria
2014	Ying et al.	Anti-traffic analysis attack for location privacy	Anti-traffic analysis (ATA)
2014	Chen et al.	Protecting end-to-end location privacy against local eavesdropper	Parent based attack scheme (PAS)
2015	Long et al.	Protecting the base station	Ring Based Routing (RBR)
2015	Rios et al.,	Concealing the genuine packets	Homogenous Injection for Sink Privacy with Node Compromise protection (HISP-NC)
2015	Tang et al.	Increasing the anonymity of the source nodes	Random walk/ Cost aware Secure routing (CASER)
2015	Chen and Lou	Hiding the sink node while concealing the location of the source nodes.	Combination of sending fake packets or having a proxy sink (LPSS)
2015	Lee et al.,	Protecting both the sink and source nodes	constant-rate broadcast scheme (CRBS)
2015	Abuzneid et al.,	Location privacy is WSN	Fortified anonymous communication (FAC)
2017	Liu et al.,	Sink node location privacy protection	Semi Random Circle routing for mobile Sink joint Ray Routing for data (SRCRR)
2017	Baroutis and Younis	Increase the anonymity of the base station for the adversaries	Preserve Location Anonymity scheme through Uniform Distribution of Traffic volume (PLAUDIT)
2017	Wang et al.,	Hiding the sink node location	All-direction random routing algorithm (ARR)
2019	Wang et al.,	Protecting the source node by acting as a false source node	SPS and DPS: Grid-based source location privacy scheme
2020	Rimjhim et al.,	Source node location privacy	Select the location of fake and phantom sources based on geometry concept
2020	Lilian et al.,	Source/sink node location privacy	Two fake packet routing protocols and two new phantom-based source location privacy routing
2020	Mehdi et al.,	Excluding the sink node from the routes	Formation of illusive void (FIVA)

Table 2:Cont. WSN Literature summary

CHAPTER 3: PROBLEM DEFINITION

WSN is a network of small devices known as sensors which communicate via an ad hoc communication manner. This study will develop a mixed integer programming model (MIP) for a WSN that has several source nodes and exactly one sink node. The sink node collects and processes the gathered data and sends it to the remote users for various needs. The sink node has an important and unique role in the entire network which will attract unwanted attention from the attackers because of the high transmission rate near it. The study focuses on protecting the sink node from attackers and increasing its anonymity by different methods. As far as the study's on protecting the sink node in a WSN, this study is the first attempt to develop a mathematical model to represent the WSN with several source nodes and one sink node while focusing on increasing the anonymity of the sink node. A basic model for a WSN with the objective of balancing/decreasing the traffic in the entire grid will be discussed in the coming section. The objective of the model will decrease the probability of finding the sink node by the attacker since the traffic is decreased in the whole grid. Two extended models that monitor and control the packets in detail with implementing a phantom routing mechanism will be presented in the coming sections. The tradeoff between the energy consumption and the location privacy of the sink node will be studied.

CHAPTER 4: WSN MATHEMATICAL MODEL

This section will discuss and present the novel WSN mathematical model that was discussed in previous chapters.

4.1 Base Model Formulation

In the following subsections the base model Sets, Parameters, Variables, Constraints, and Objective Function will be explained thoroughly.

4.1.1 Indices Set

In this subsection the sets that are representing the indices of the decision variables and the parameters will be discussed:

Set $M = \{1, 2, 3, \dots, M\}$ = Set of Nodes

Set $S = \{1, 2, 3, \dots, S\}$ = Set of Source Nodes

Set $D = \{1, 2, 3, \dots, D\}$ = Set of Sink Nodes

Set $G = \{1, 2, 3, \dots, G\}$ = Set of Grid Size

Set $T = \{1, 2, 3, \dots, T\}$ = Set of Time

Set $Z = \{1, 2, 3, \dots, Z\}$ = Complement Set of S

Set $D1 = \{1, 2, 3, \dots, D1\}$ = Complement Set of D

Set $SS = \{1, 2, 3, \dots, SS\}$ = Set of Regular Nodes excluding the sink and the source nodes

Set $PN = \{1, 2, 3, \dots, PN\}$ = Set of Phantom Nodes

4.1.2 Base Model Parameters

a_{ig} = 1 If node i belongs to grid g , 0 otherwise.

b_{ij} = 1 If node i is transmitting a packet to node j , 0 otherwise.

q_s = The number of packets that are sent from source node s

c_{ij} = The listening nodes for each (i,j) pair of nodes

4.1.3 Base Model Variables

a_{gk} = The maximum difference between any two grids during any time

$\hat{P}_{gt}$ = The traffic in grid g at time t .

X_{ijt} = 1 if a packet is sent from node i to node j at time t , and zero otherwise

P_{gt} = Probability of base station being in grid g at time t .

β = Maximum difference in the entire network M

t_{max} = The maximum amount of time that is needed to send all the packets

4.1.4 Base Model Constraints

Let us assume that the network has M nodes, and the network is divided into $N \times N$ grid. Our N is 3 and the grid size is 3 x 3 with 9 cells. Our M is 18 nodes, one of them is the sink node.

The model constraints and objective function will be explained in this section. There is a proportional relationship between the probability and the traffic in each grid at time t which is given by this relation:

$$P_{gt} \propto \sum_{i=1}^m a_{ig} \left[\sum_{j=1}^M X_{ijt} + \sum_{j=1}^M X_{jit} \right]$$

Constraints 1,2,3

$$\alpha_{gk} \geq |\hat{P}_{gt} - \hat{P}_{kt}| \quad \forall t, g, k$$

$$\hat{P}_{gt} = \sum_{i=1}^m a_{ig} \left[\sum_{j=1}^M X_{ijt} + \sum_{j=1}^M X_{jit} \right] \quad i \neq j, \forall t, g$$

$$\beta \geq \alpha_{gk} \quad \forall g, k$$

These constraints will calculate the traffic in each grid g in the WSN at time t . The number of packets transferred to or generated by node i will be calculated if node i belongs to a specific grid g . The variable β is used in constraint 3 to minimize it in the objective function and to take the highest value that α_{gk} takes.

Constraint 4

$$\sum_j^M \sum_t^T X_{sjt} = q_s \quad \forall i \in S, s \neq j$$

This constraint means that the sources will send all the specified number of packets based on the parameter q_s to the j th node in all times.

Constraint 5

$$\sum_{j=1}^M X_{ij1} = 0 \quad \forall i \in Z, t$$

This constraint means that no packet will be sent at time 1 from any node that doesn't belong to the sources set.

Constraint 6

$$\sum_{i=1}^M X_{ijt} = \sum_{i=1}^M X_{ji(t+1)} \quad \forall j \in ss, t < |T|$$

This constraint is concerned with the flow of the packets in the network. The packet that is received at node j from node i should be sent to another node i at time $t+1$.

Constraint 7

$$X_{ijt} \leq b_{ij} \quad \forall t, i, j$$

The connection between node i and node j will be fulfilled if they satisfy the predefined flow of packets.

Constraint 8

$$\sum_i^M \sum_t^T X_{idt} = \sum_{s=1}^M q_s \quad \forall d$$

This constraint means that the number of packets that needs to be received in the sink node is total of all the packets that are send from all sources.

Constraint 9

$$X_{ist} = 0 \quad \forall i, s, t$$

The purpose of this constraint is that the source nodes can't receive any packet at any time.

Constraints 10, 11, 12

$$\sum_{j=1}^M X_{ijt} \leq 1 \quad \forall i, t$$

$$\sum_{i=1}^M X_{ijt} \leq 1 \quad \forall j, t$$

$$\sum_{i=1}^M X_{izt} + \sum_{i=1}^M X_{zit} = 1 \quad \forall z, t$$

These constraints make sure that each node can only send or receive one packet at most for each time t .

Constraint Set 13

$$\beta \geq 2$$

$$\alpha_{gk} \leq 2 \quad \forall g, k$$

These constraints make sure to restrict the beta value to a minimum of 2 and alpha to a maximum of 2.

Constraint 14

$$tmax \geq ord(t) * X_{ijt} \quad \forall i, j, t$$

This constraint means that the maximum time of the entire network should be larger than or equal to the last ordered value of time if the packet is sent from node i to node j .

Constraint 15

$$\sum_j^J \sum_m^M c_{ij} * X_{jmt} + \sum_j^J \sum_m^M c_{ij} * X_{mjt} \leq 1 + 10000 * (1 - \sum_{m=1}^M X_{imt}) \quad \forall i, t$$

This constraint serves the purpose of not sending any packet to the nodes which are in the listening area of the i th node which is the sending node.

4.1.5 Base Model Objective Function

$$\text{Minimize } \sum_{i,j,t} W1 X_{ijt} + W2 \beta + W3 t_{\max}$$

The objective function will minimize the traffic in the entire network to decrease the probability that the attacker can find the sink node and will try to get the shortest path of the packets. The variable β should be minimized because it's related to the variable α_{gk} in Constraints 1,2,3 which will calculate the traffic in the WSN at a specific grid cell at a certain time. The last part of the objective function will minimize the value of $t_{\max}$ which is the maximum time needed to send all the packets needed from the sources to the sink. So basically, the objective function is trying to get the shortest path for the packets that are sent from the source nodes through other nodes until they reach the sink node, minimizing the maximum difference between two grids in the entire WSN and minimizing the overall time that it needs for the packets to arrive at the sink.

4.2 Extended model #1

This section will discuss the changes that are made to the model by adding 2 extra indices to observe the packets for each source and destination pair.

$Y_{sdijt} = 1$ if a packet is sent from node i to node j at time t for the source and destination pair (s, d) , and zero otherwise

4.2.1 Extended Model #1 Constraints

Constraints 1,2,3

$$\alpha_{gk} \geq |\hat{p}_{gt} - \hat{p}_{kt}| \quad \forall t, g, k$$

$$\hat{p}_{gt} = \sum_{i=1}^m a_{ig} \left[\sum_s^S \sum_d^D \sum_i^I \sum_j^J Y_{sdi jt} + \sum_s^S \sum_d^D \sum_i^I \sum_j^J Y_{sdj it} \right] \quad \forall t, g, i \neq j$$

$$\beta \geq \alpha_{gk} \quad \forall g, k$$

These constraints will calculate the traffic in each grid g in the WSN at time t . The number of packets transferred to or generated by node i will be calculated if node i belongs to a specific grid g . The variable β is used in Constraint 3 to minimize it in the objective function and to take the highest value that α_{gk} takes.

Constraint 4

$$\sum_d^D \sum_t^T \sum_j^J Y_{sdsjt} = q_s \quad \forall i \in S, s \neq j$$

This constraint means that the sources will send all the specified number of packets based on the parameter q_s to the j th node in all times.

Constraint 5

$$\sum_s \sum_d \sum_j Y_{sdsjt} = 0 \quad \forall i \in Z, t = 1$$

This constraint means that no packet will be sent at time 1 from any node that doesn't belong to the sources set.

Constraint 6

$$\sum_i \sum_d Y_{sdisst} = \sum_i \sum_d Y_{sdssi(t+1)} \quad \forall j \in SS, s$$

This constraint is concerned with the flow of the packets in the network. The packet that is received at node j from node i should be sent to another node i at time $t+1$.

Constraint 7

$$Y_{sdijs} \leq b_{ij} \quad \forall s, d, t, i, j$$

The connection between node i and node j will be fulfilled if they satisfy the predefined flow of packets.

Constraint 8

$$\sum_s \sum_i \sum_t Y_{sdijs} = \sum_{s=1}^M q_s \quad \forall d$$

This constraint means that the number of packets that needs to be received in the sink node is total of all the packets that are send from all sources.

Constraint 9

$$Y_{sdist} = 0 \quad \forall d, i, s, t$$

The purpose of this constraint is that the source nodes can't receive any packet at any time.

Constraints 10&11&12

$$\sum_s^S \sum_d^D \sum_j^J Y_{sdijt} \leq 1 \quad \forall i, t$$

$$\sum_s^S \sum_d^D \sum_i^I Y_{sdijt} \leq 1 \quad \forall j, t$$

$$\sum_s^S \sum_d^D \sum_i^I Y_{sdizt} + \sum_s^S \sum_d^D \sum_i^I Y_{sdzjt} = 1 \quad \forall z, t$$

These constraints make sure that each node can only send or receive one packet at most for each time t .

Constraint 13

$$\beta \leq 2$$

$$\alpha_{gk} \leq 2 \quad \forall g, k$$

These constraints restrict the beta and alpha values.

Constraint 14

$$tmax \geq ord(t) * Y_{sdijt} \quad \forall s, d, i, j, t$$

This constraint means that the maximum time of the entire network should be larger than or equal to the last ordered value of time if the packet is sent from node i to node j .

Constraint 15

$$\begin{aligned} \sum_s^S \sum_d^D \sum_j^J \sum_m^M c_{ij} * Y_{sdjmt} + \sum_s^S \sum_d^D \sum_j^J \sum_m^M c_{ij} * Y_{sdmjt} \\ \leq 1 + 10000 * (1 - \sum_s^S \sum_d^D \sum_m^M Y_{sdimt}) \quad \forall i, t \end{aligned}$$

This constraint serves the purpose of not sending any packet to the nodes which are in the listening area of the i th node which is the sending node.

4.2.2 Extended Model #1 Objective Function

$$\text{Minimize} \sum_{s,d,i,j,t} W1Y_{sdijt} + W2\beta + W3tmax$$

The 1st part of the objective function is the number of hops needed to send the last packet to the sink node and it's denoted by Y_{sdijt} . There is a relationship between the number of hops and the energy consumption of a WSN. If the number of hops increases, then the energy consumed to send all these hops increases.

The 2nd part of the objective function is the global traffic difference between two grids, and it's denoted by *beta*. This objective plays a huge role in our model and in a real WSN's since it's an anonymity criterion to increase the anonymity of the sink node by balancing traffic in all grids.

The 3rd part of the objective function is the total time needed to send all the packets to the sink node and it's denoted by *tmax*. This objective is important in a WSN since it impacts the energy consumption of the network. If the network needs a higher time to send all the packets to the sink node, then it will consume higher total energy.

The constraints are needed to structure the restrictions in a WSN. For example, the sink node can only receive packets and not allowed to send any packets at any time. These restrictions can be formulated as constraints to capture the real behavior of the packets in a WSN.

4.3 Extended model #2

This section will discuss the changes that are made to Model 2 by adding an extra index that is concerned with the packets sent from the sources while implying the phantom routing mechanism for each source and destination pair. A new parameter co_{ij} was added which will help to randomize the packets sent near the sink node by creating a zigzag pattern to distract the adversary.

$V_{sdrijt} = 1$ if the packet r is sent from node i to node j at time t for the source and destination pair (s, d) , and zero otherwise

Set $PN = \{1, 2, 3, \dots, PN\}$ = Set of Phantom Nodes

4.3.1 Extended Model #2 Constraints

Constraints 1,2,3

$$\alpha_{gk} \geq |\hat{P}_{gt} - \hat{P}_{kt}| \quad \forall t, g, k$$

$$\hat{P}_{gt} = \sum_{i=1}^m a_{ig} \left[\sum_s^S \sum_d^D \sum_i^I \sum_j^J \sum_r^R V_{sdrijt} + \sum_s^S \sum_d^D \sum_i^I \sum_j^J \sum_r^R V_{sdrjit} \right] \quad \forall t, g, i \neq j$$

$$\beta \geq \alpha_{gk} \quad \forall g, k$$

These constraints will calculate the traffic in each grid g in the WSN at time t . The number of packets transferred to or generated by node i will be calculated if node i belongs to a specific grid g . The variable β is used in constraint 3 to minimize it in the objective function and to take the highest value that α_{gk} takes. These constraints will capture the traffic and the traffic difference between two grids in a WSN, and this will help in applying the anonymity criteria which is traffic balancing.

Constraint 4

$$\sum_t^T \sum_j^J V_{sdrsjt} * b_{sj} = 1 \quad \forall i \in S, s \neq j, r \in R, d \in D$$

This constraint means that each packet should be assigned to at most one source in that specific source and destination pair. In a WSN, the beginning of the packet starts from at most one source node.

Constraint 5

$$\sum_s^S \sum_d^D \sum_j^J \sum_r^R V_{sdrzjt} = 0 \quad \forall i \in Z, t = 1$$

This constraint means that no packet will be sent at time 1 from any node that doesn't belong to the sources set. This captures the behavior of packets at time 1 in a WSN.

Constraint 6

$$\sum_i^I V_{sdrisst} = \sum_i^I V_{sdrssi(t+1)} \quad \forall j \in ss, s, r, d, t$$

This constraint is concerned with the flow of the packets in the network. The packet that is received at node j from node i should be sent to another node i at time $t+1$. The packets will flow to the nodes that are in their range for a WSN. This constraint is based on the flow of the packets with the assumption that there is no loss.

Constraint 7

$$V_{sdrijt} \leq b_{ij} \quad \forall s, d, r, t, i, j$$

The connection between node i and node j will be fulfilled if they satisfy the predefined flow of packets. The packets will be sent to nodes that are in range for a WSN.

Constraint 8

$$\sum_t^T \sum_i^I V_{sdridt} = 1 \quad \forall s, d, r$$

This constraint means that each packet for a specific source and destination pair should be received in the sink node. Each packet that is sent needs to end up in the sink node for a WSN.

Constraint 9

$$V_{sdrist} = 0 \quad \forall d, i, s, r, t$$

The purpose of this constraint is that the source nodes can't receive any packet at any time. Source nodes in a WSN can't receive any packet.

Constraints 10&11&12

$$\sum_s^S \sum_d^D \sum_j^J \sum_r^R V_{sdrijt} \leq 1 \quad \forall i, t$$

$$\sum_s^S \sum_d^D \sum_i^I \sum_r^R V_{sdrijt} \leq 1 \quad \forall j, t$$

$$\sum_s^S \sum_d^D \sum_i^I V_{sdrizt} + \sum_s^S \sum_d^D \sum_i^I V_{sdrzit} = 1 \quad \forall z, r, t$$

These constraints make sure that each node can only send or receive one packet at most for each time t . In a WSN, each intermediate node can store at most one packet at a time. It's assumed that there is no collision due to coordination between the users.

Constraint 13

$$\beta \geq 1$$

$$\alpha_{gk} \leq 1 \quad \forall g, k$$

These constraints make sure to restrict the beta value to a minimum of 1 and alpha to a maximum of 1. This constraint is used when the constructor of the model wants to restrict the values of alpha and beta.

Constraint 14

$$tmax \geq ord(t) * V_{sdrijt} \quad \forall s, d, r, i, j, t$$

This constraint means that the maximum time of the entire network should be larger than or equal to the last ordered value of time if the packet is sent from node i to node j . In a WSN, this constraint captures the last time needed to receive a packet in the sink node.

Constraint 15

$$\sum_s^S \sum_d^D \sum_j^J \sum_m^M \sum_r^R c_{ij} * V_{sdrjmt} + \sum_s^S \sum_d^D \sum_j^J \sum_m^M \sum_r^R c_{ij} * V_{sdrmjt} \\ \leq 1 + 10000 * (1 - \sum_s^S \sum_d^D \sum_m^M \sum_r^R V_{sdrimt} \quad \forall i, t$$

In a WSN, this constraint serves the purpose of not sending any packet to the nodes which are in the listening area/range of the i th node which is the sending node.

Constraint 16

$$\sum_{pn}^{PN} \sum_t^T \sum_i^I V_{sdripnt} \geq 1 \quad \forall s, d, r$$

This constraint means that at least one phantom node must be visited in the path of each packet for each source and destination pair. In a WSN, this constraint captures the packets that are visiting the predefined phantom nodes.

4.3.2 Modified model #2 Objective Function

$$\text{Minimize} \sum_{s,d,i,j,r,t} co_{ij} * W1V_{sdrijt} + W2\beta + W3t_{max}$$

This objective function serves the same three purposes that were mentioned in the base model's objective function section with an extra cost parameter that is multiplied by the variable v . The cost parameter has a cost of 1 for any connection in the network except the nodes 21-25. The cost to send to or receive from these specific nodes is zero. This will create a random zigzag movement near the sink node which serves the purpose of distracting the adversary from detecting the sink node. There is a trade-off between creating the zigzag pattern and the total energy consumed in the network. The total energy consumed by the network will increase if the number of packets increased even if some of the nodes are multiplied by zero cost. The v -variable will calculate the value of the hops that aren't multiplied by zero. We added a parameter in the results that calculates the real number of hops and that will help us in knowing how many packets traveled through the nodes that we chose to have zero cost.

CHAPTER 5: MODEL EXECUTION

The software General Algebraic Modelling System (GAMS) was used to build the basic model and the extended versions of the basic model. All the models were solved using the CPLEX solver on an Intel core i9-9900k-3.60GHz CPU and 16 GB RAM under Win 10 Pro Operating System.

5.1. GAMS Code Execution:

The parameters in the model are being called from an excel spreadsheet file. There is an excel file for each of the runs for the three main models. The parameters contain the following for the base model and the first extended model:

$a_{ig} = 1$ If node i belongs to grid g , 0 otherwise.

$b_{ij} = 1$ If node i is transmitting a packet to node j , 0 otherwise.

$q_s =$ The number of packets that are sent from source node s

$c_{ij} =$ The listening nodes for each (i,j) pair of nodes

We don't need the q parameter in the V-model, because it's defined as a set to help control the packets more efficiently. The following tables will show samples of the tables for the last run of the V-Model.

Grid	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
9	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
10	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
11	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
12	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
13	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
14	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
15	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
16	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
21	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
22	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
24	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
26	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

Table 3: Placement of Nodes in Grid g

Packets	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0
2	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1
3	0	1	0	1	0	0	0	0	0	0	0	0	0	0	1
4	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0
5	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
8	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
9	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
10	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0
11	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
12	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0
13	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0
14	0	1	0	0	0	0	0	0	0	0	0	0	1	0	1
15	0	1	1	0	0	0	0	0	0	0	0	0	0	1	0
16	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1
17	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
18	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0
19	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0
24	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 4: Sample Feasibility Matrix

List Nodes	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0
2	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1
3	0	1	0	1	0	0	0	0	0	0	0	0	0	1	1
4	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1
5	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0
6	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
7	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
8	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0
9	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0
10	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0
11	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
12	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0
13	1	1	0	0	0	0	0	0	0	0	0	1	0	1	0
14	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1
15	0	1	1	1	0	0	0	0	0	0	0	0	0	1	0
16	0	0	1	1	1	0	0	0	0	0	0	0	0	0	1
17	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0
18	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0
19	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
20	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0
21	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
22	1	0	0	0	0	0	0	0	0	0	1	1	1	0	0
23	0	1	0	0	0	0	0	0	0	0	0	1	1	1	0
24	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1
25	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0
26	0	1	0	0	0	0	0	0	0	0	0	1	0	1	0
27	0	0	1	0	1	0	0	0	0	0	0	0	0	0	1
28	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0
30	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0

Table 5:Sample Listening nodes

Cost	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1
5	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1
6	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1
7	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1
8	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1
9	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1
12	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1
13	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
15	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0
16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1
22	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1
23	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1
24	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0
25	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
28	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Table 6:Sample Cost Matrix

CHAPTER 6: RESULTS AND FINDINGS

The results section will be split into sections for each of the models starting with the base model (X-model), Extended Model #1 (Y-Model) and Extended Model #2 (V- model). The section will include the 21 nodes network then the 30 nodes network results.

6.1. Case Study of 21 nodes network:

The model will be run four times by modifying the objective function and adding constraints in the final model to check the optimal value of each variable in the objective function. The four runs will be in the following order: optimal value of the shortest path (X-variable), the total time of the network ($tmax$), the global maximum traffic difference between two grids (beta) and the final run will be only for the x-variable with adding two extra constraints for beta and $tmax$.

The criteria for the two extra constraints will be the variable less than or equal to 20% tolerance of the optimal value that we previously got. For example, if we got the optimal value for the Beta variable alone is 2 then the added constraint will be $\beta \leq 2.4$. In the fourth run we will multiply beta and $tmax$ by zero and run the model for the X-variable with the two added constraints that are concerned with $tmax$ and beta.

This strategy was chosen to decrease the solving time of the fourth run and to get the optimal values for each variable in the objective function separately. A summary of the models' advantages and disadvantages is shown in Table 3. The computational time of the four runs for the 30 nodes network models is calculated in the third column (Computational Time).

Model	Evaluation Method	Anonymity Criteria	Computational Time	Controlling Packets	Phantom Routing Mechanism	Cost Parameter
Base (X)	Mathematical modelling	Traffic Balancing	10 Min	No	No	Yes
Extended #1 (Y)	Mathematical modelling	Traffic Balancing	73 Min	Yes	No	Yes
Extended #2 (V)	Mathematical modelling	Traffic Balancing	180 Min	Yes	Yes	Yes

Table 7: Summary of the models

6.1.1. Results for the 21 nodes using the base model:

The results of the X-variable only:

The following Figure shows the optimal solution when $tmax$ and $beta$ are multiplied by zero and the model is only minimizing the shortest path separately. The model got an optimal solution in 1 second with a value of 20 hops.

```
Proven optimal solution.
MIP Solution:          20.000000
Final Solve:          20.000000

Best possible:         20.000000
Absolute gap:          0.000000
Relative gap:          0.000000
```

Figure 1: Optimal Shortest Path Value (BM-21)

Figures 2 and 3 show the route that the packets take from a source to a sink and at what time. For example, the first packet is sent from source 1 to intermediate node 7 then at time 2 the packet is sent from node 7 to the sink node number 21.

----- 122 VARIABLE x.L						
	1	2	3	4	5	6
1 .7	1.000		1.000		1.000	
7 .21		1.000		1.000		1.000
	+	7	8	38	39	40
						41
1 .7	1.000					
1 .12						1.000
4 .15			1.000		1.000	
7 .21		1.000				
15.21				1.000		1.000
	+	42	43	44	45	46
						47
4 .15	1.000		1.000		1.000	
12.21	1.000					
15.21		1.000		1.000		1.000

Figure 2: The Shortest Path (BM-21)

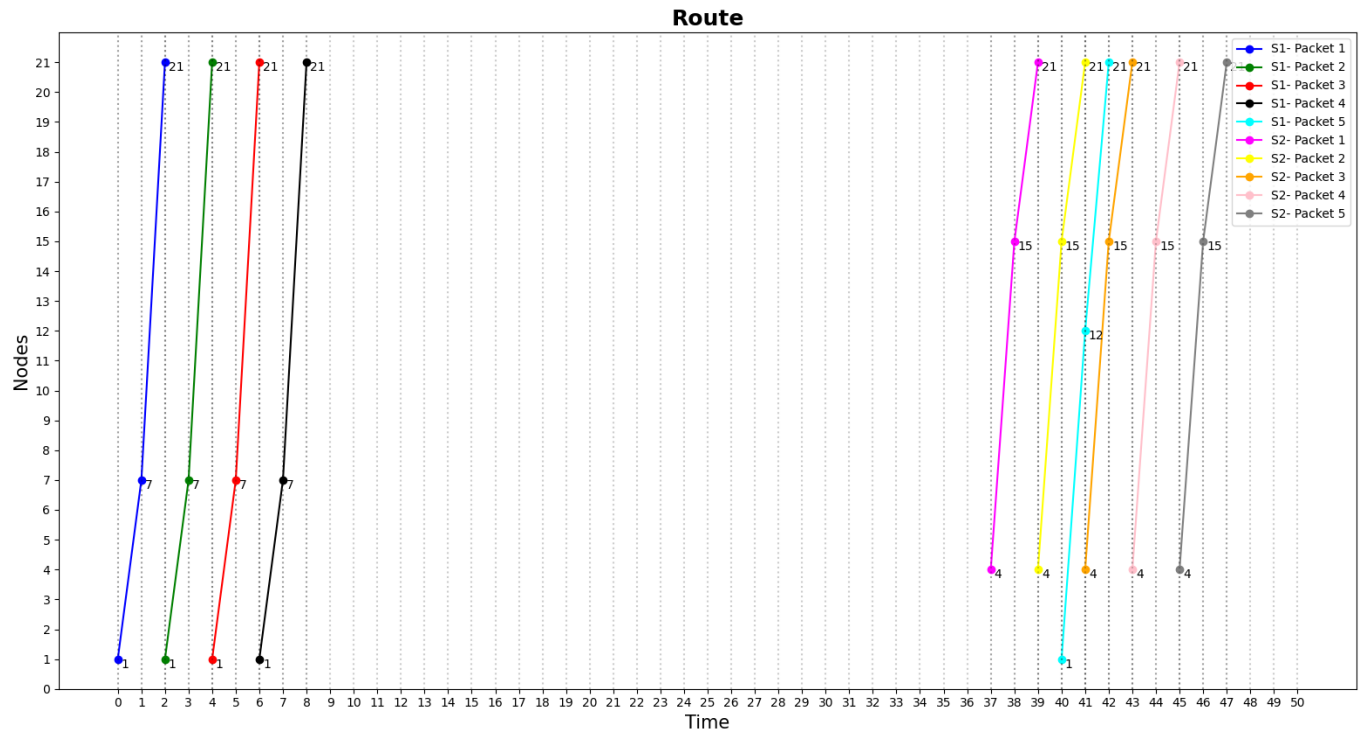


Figure 3: Path Chart (BM-21)

The results of the beta only:

The optimal value is found to be 1 which is the lowest it can be for the global maximum difference in traffic between two grids. The solver calculated the value in 1 second. The path of the packets is shown in Figure 5 when beta is minimized.

```
MIP Solution:          1.000000
Final Solve:           1.000000

Best possible:         1.000000
Absolute gap:           0.000000
Relative gap:           0.000000
-----
--- Restarting execution
--- beta-only.gms (119) 4 Mb
```

Figure 4: Beta Optimal Value (BM-21)

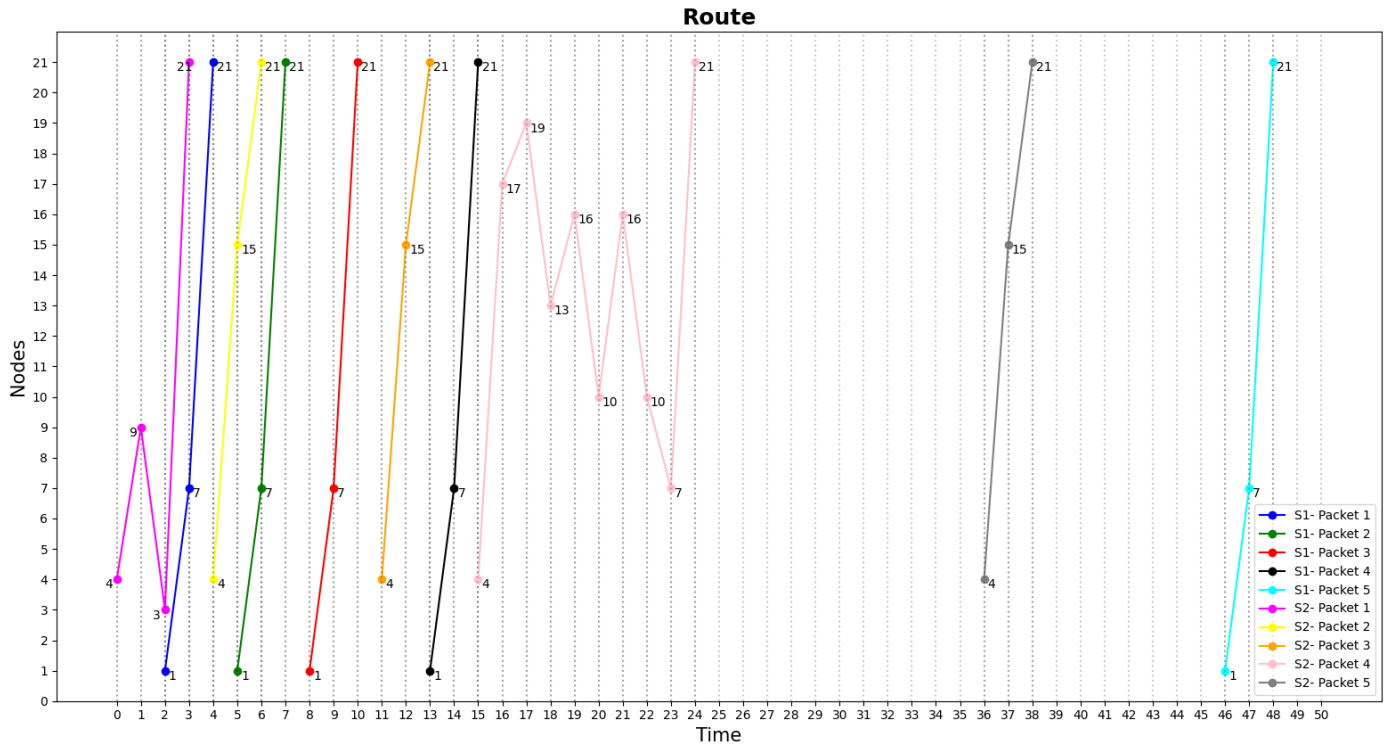


Figure 5: Beta Chart (BM-21)

The results of the tmax only:

The lowest time to send the 10 packets for both sources needed 11 units of time to finish as it can be seen in Figure 6. Figures 7 and 8 show the route of the packets while tmax is minimized. The elapsed time was 7 seconds to find the optimal tmax value.

Proven optimal solution.

```
MIP Solution:          11.000000
Final Solve:          11.000000

Best possible:         11.000000
Absolute gap:          0.000000
Relative gap:          0.000000
```

```
--- Restarting execution
```

```
--- tmax-only.gms (119) 4 Mb
```

Figure 6: tmax Optimal Value (BM-21)

```
----      122 VARIABLE x.L

              1          2          3          4          5          6

1 .7          1.000          1.000          1.000
3 .21         1.000          1.000
4 .9          1.000
4 .15         1.000          1.000          1.000
7 .21         1.000          1.000          1.000
9 .3          1.000
15.21         1.000

+          7          8          9          10         11

1 .7          1.000          1.000
4 .15         1.000          1.000          1.000
7 .21         1.000          1.000          1.000
15.21         1.000          1.000          1.000
```

Figure 7: The Path when tmax is minimized (BM-21)

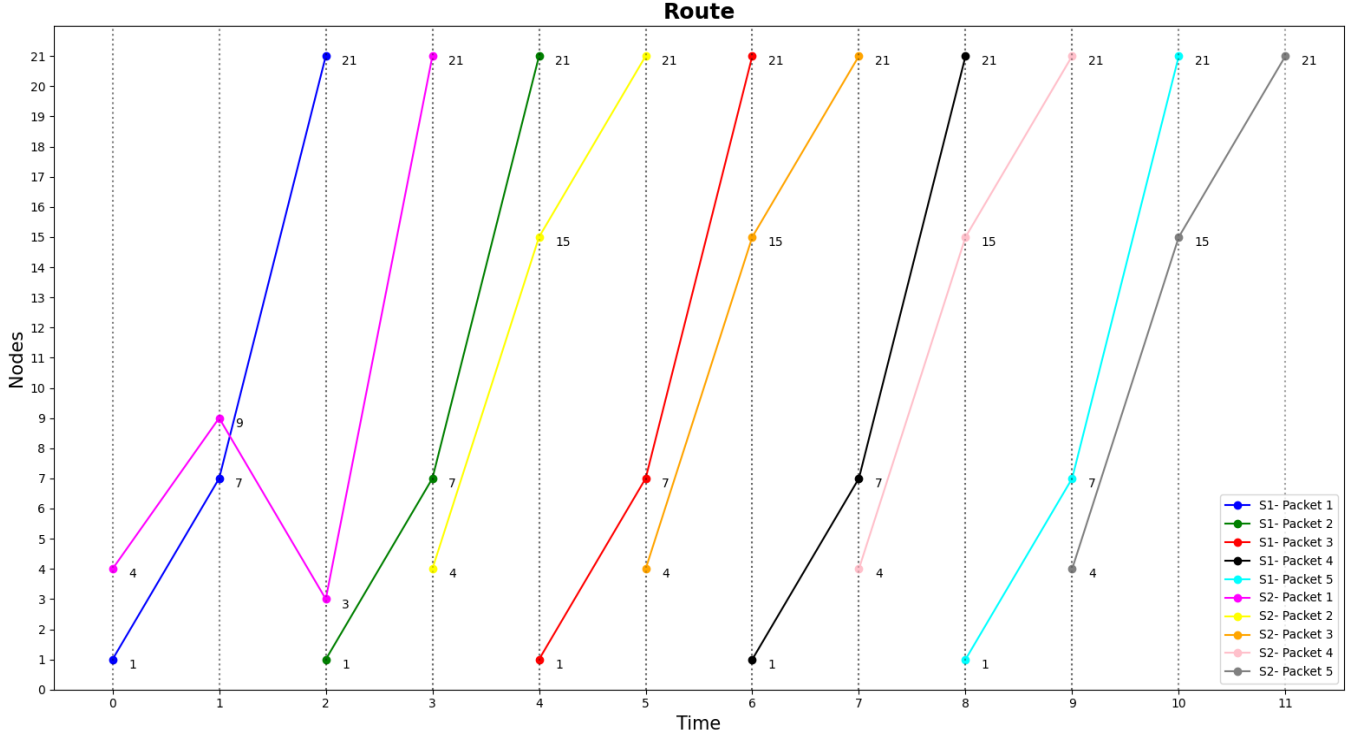


Figure 8: Tmax Chart (BM-21)

The results of the x-variable only while adding the two extra constraints:

In this part we will minimize the x-variable while giving an upper bound to the tmax and the beta. The optimal beta and tmax values respectively were 1 and 11, the two extra constraints added will be as follows: $\beta \leq 1.2$ and $t_{\max} \leq 13.2$. Figures 9 and 10 show the route of the whole network.

----- 124 VARIABLE x.L						
	1	2	3	4	5	6
1 .12		1.000			1.000	
4 .15	1.000			1.000		
12.21			1.000			1.000
15.21		1.000			1.000	
+						
	7	8	9	10	11	12
1 .12	1.000		1.000		1.000	
4 .15		1.000		1.000		1.000
12.21		1.000		1.000		1.000
15.21			1.000		1.000	
+						
	13					
15.21	1.000					

Figure 9: The route for run 4 (BM-21)

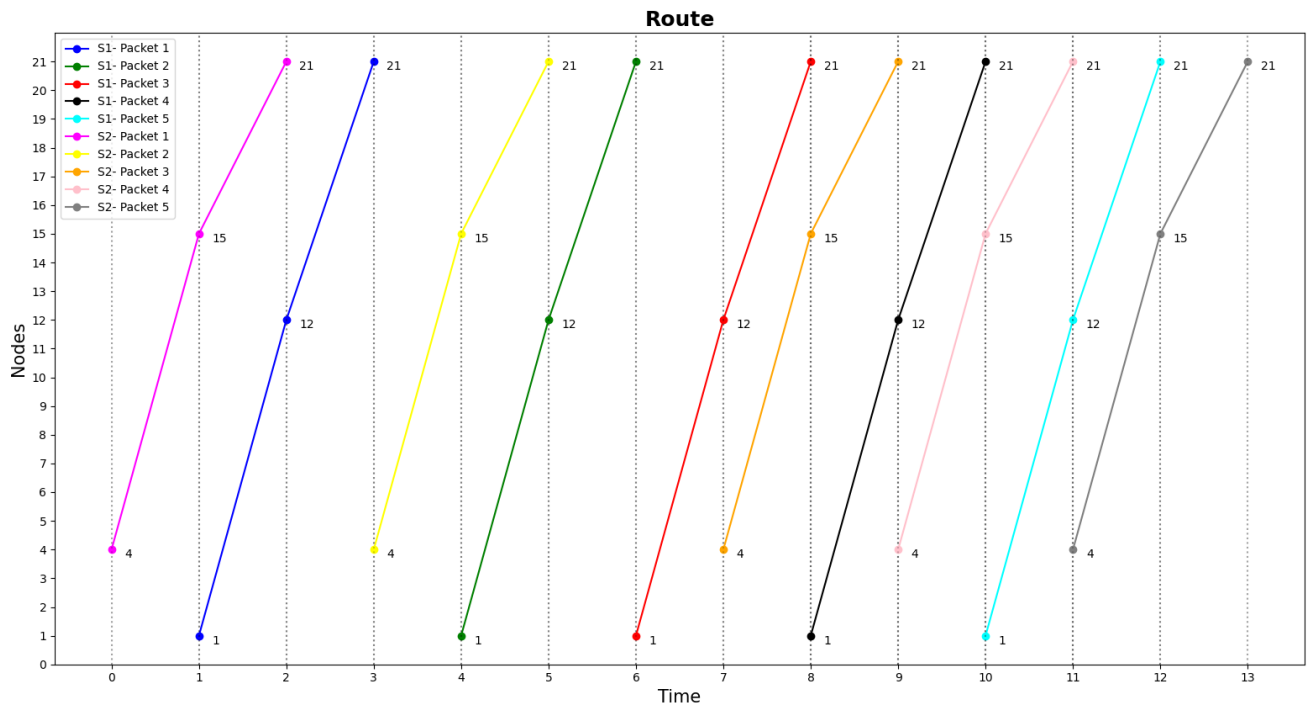


Figure 10 : Final Chart (BM-21)

It can be seen from Figure 11 that the number of hops stayed the same as 20 since the model is minimizing that specific variable but the t_{max} increased by 2 units of time to 13. The beta stayed the same as the optimal of 1. One can argue why did the t_{max} change and that is simply because

the model is minimizing the number of hops to finish the specified packets and to minimize the number of hops its relaxing the $tmax$ for the tolerance that was chosen (20%). The elapsed time was 3 seconds to find the optimal solution.

```

----      124 VARIABLE beta.L           =          1.000
           VARIABLE bn.L               =          20.000
           VARIABLE tmax.L              =          13.000

```

Figure 11: Variable Results (BM-21)

6.1.2. Results for the 21 nodes Y-model:

The results of the Y-variable only:

The following Figure shows the optimal solution when $tmax$ and $beta$ are multiplied by zero and the model is only minimizing the shortest path separately. The model got an optimal solution in 5 second with a value of 20 hops.

```

| Proven optimal solution.
|
| MIP Solution:           20.000000
| Final Solve:           20.000000
|
| Best possible:         20.000000
| Absolute gap:           0.000000
| Relative gap:           0.000000

```

Figure 12: Optimal Shortest Path Value (M1-21)

Figures 13 and 14 shows the route that the packets take from a source to a sink and at what time.

For example, the first packet is sent from source 4 to intermediate node 15 then at time 2 the

packet is sent from node 15 to the sink node number 21. The difference between the base model and the modified model is that we are observing the packets for each source and destination pair.

```

|---- 131 VARIABLE y.L
INDEX 1 = 1 INDEX 2 = 21

```

	22	23	26	27	28	29
1 .12	1.000		1.000		1.000	
12.21		1.000		1.000		1.000
+	30	31	36	37		
1 .12	1.000		1.000			
12.21		1.000		1.000		

```

INDEX 1 = 4 INDEX 2 = 21

```

	1	2	34	35	39	40
4 .15	1.000		1.000		1.000	
15.21		1.000		1.000		1.000
+	46	47	49	50		
4 .15	1.000		1.000			
15.21		1.000		1.000		

Figure 13: The Shortest Path(M1-21)

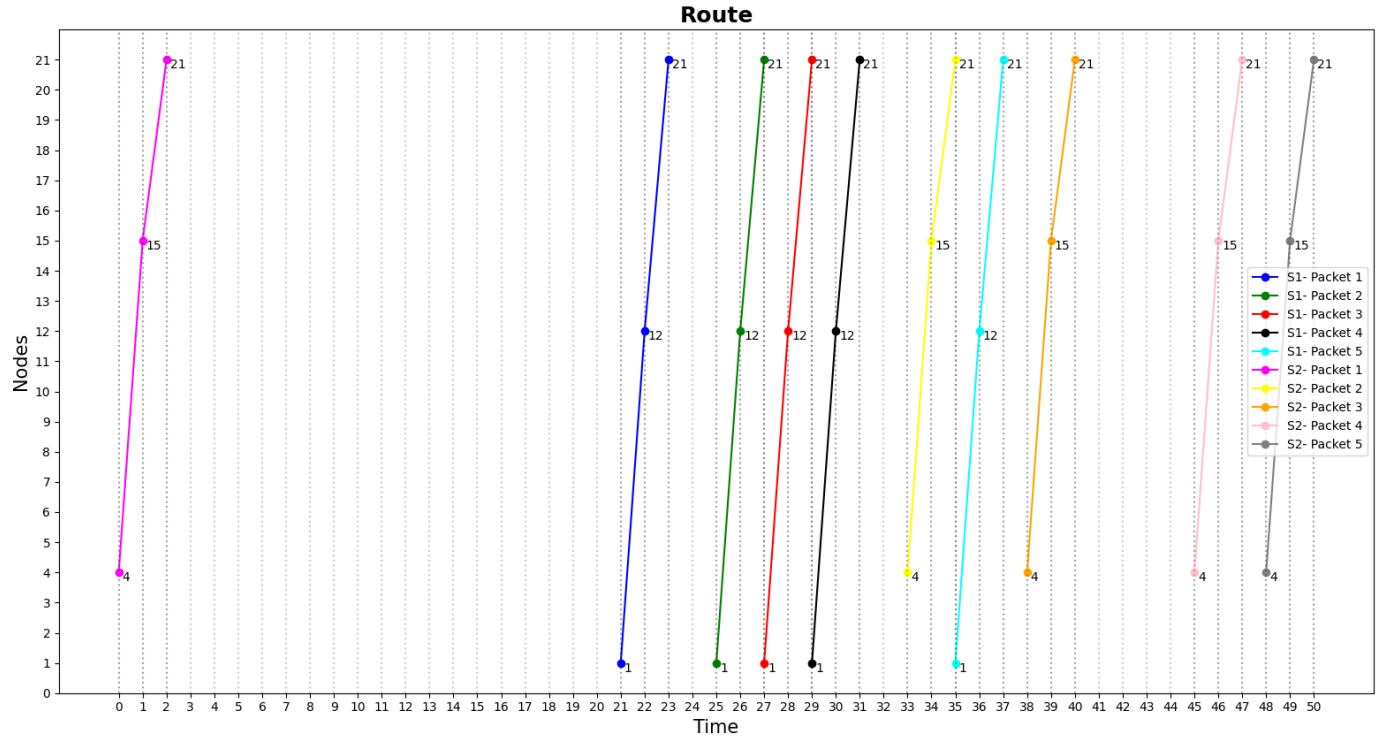


Figure 14: Path Chart (M1-21)

The results of the beta only:

The optimal value is found to be 1 which is the lowest it can be for the global maximum difference in traffic between two grids. The solver calculated the value in 4 seconds. Figure 16 shows the route of the packets when the beta is the only objective function being minimized.

```

MIP Solution:          1.000000
Final Solve:           1.000000

Best possible:          1.000000
Absolute gap:           0.000000
Relative gap:           0.000000
-----
--- Restarting execution
--- beta-only.gms (119) 4 Mb

```

Figure 15: Beta Optimal Value (M1-21)

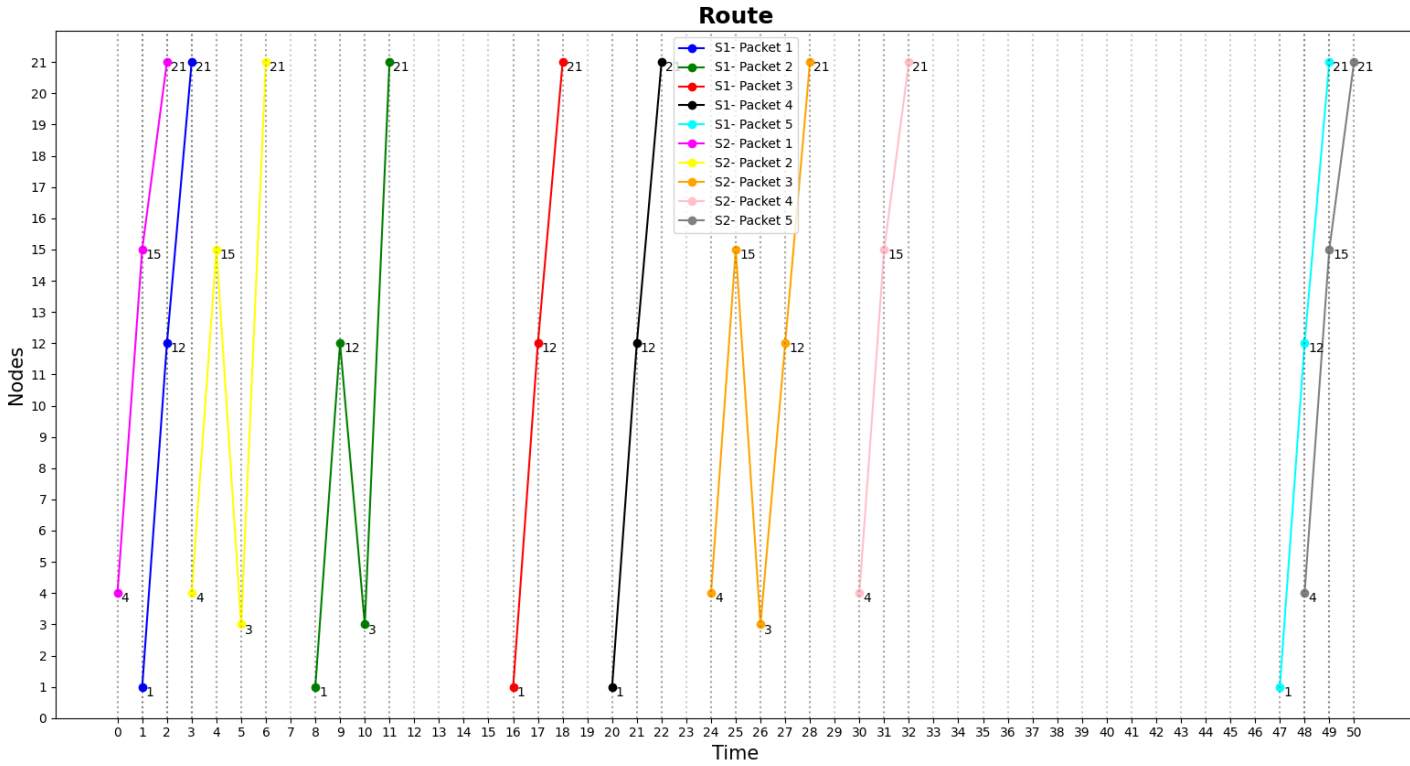


Figure 16: Beta Chart (M1-21)

The results of the t_{max} only:

The lowest time to send the 10 packets for both sources needed 11 units of time to finish as it can be seen in Figure 17. Figures 18 and 19 shows the route of the packets while t_{max} is minimized. The elapsed time was 1 min and 20 seconds to find the optimal t_{max} value.

Proven optimal solution.

MIP Solution: 11.000000
Final Solve: 11.000000

Best possible: 11.000000
Absolute gap: 0.000000
Relative gap: 0.000000

--- Restarting execution
--- tmax-only.gms (119) 4 Mb

Figure 17: t_{max} Optimal Value (M1-21)

```

----- 131 VARIABLE y.L

INDEX 1 = 1  INDEX 2 = 21

          2          3          4          5          6          7

1 .7          1.000          1.000          1.000
1 .12
7 .21          1.000          1.000          1.000
12.21          1.000

+          8          9          10          11

1 .7          1.000
1 .12          1.000
7 .21          1.000
12.21          1.000

INDEX 1 = 4  INDEX 2 = 21

          1          2          3          4          5          6

4 .15          1.000          1.000          1.000          1.000
15.21          1.000          1.000          1.000          1.000

+          7          8          9          10

4 .15          1.000          1.000          1.000
15.21          1.000          1.000

```

Figure 18: The Path when t_{max} is minimized (M1-21)

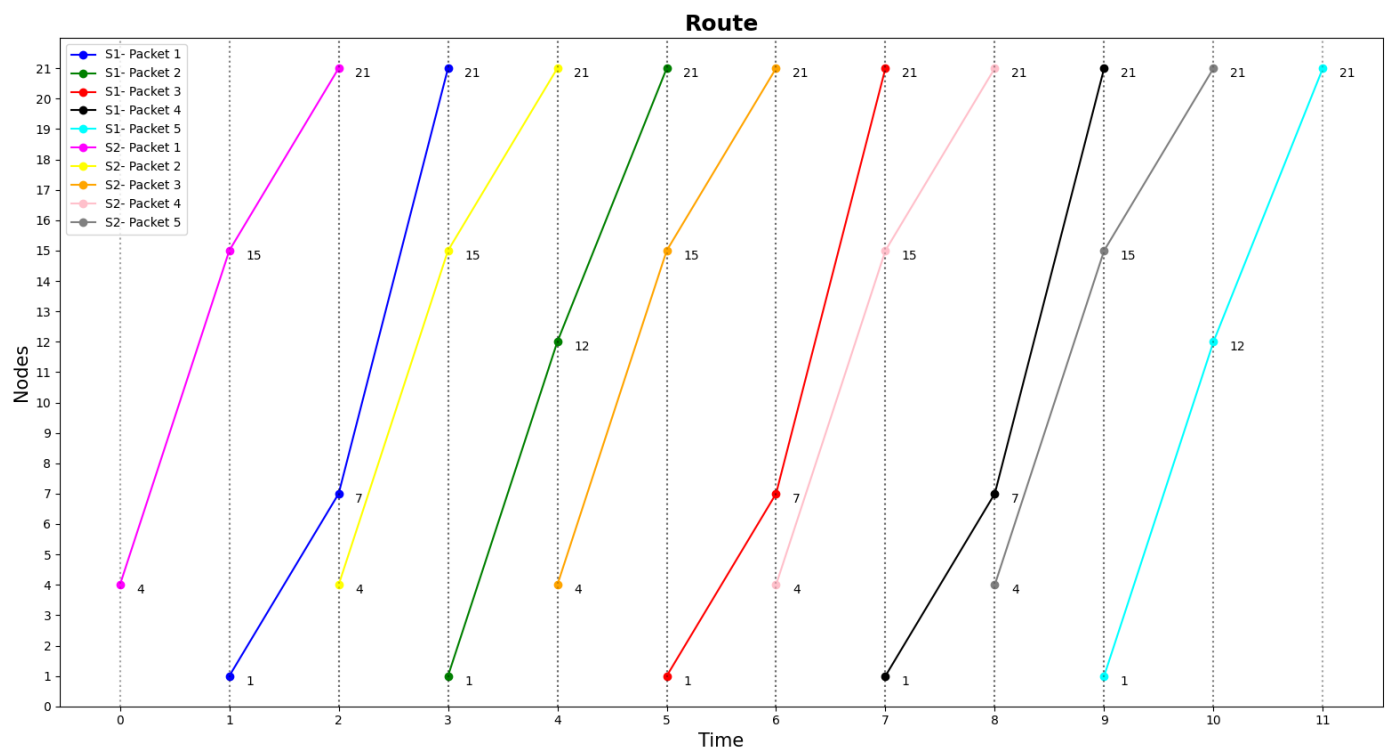


Figure 19: T_{max} Chart (M1-21)

The results of the Y-variable only while adding the two extra constraints:

In this part we will minimize the X-variable while giving an upper bound to the $tmax$ and the beta. The optimal beta and $tmax$ values respectively were 1 and 11, the two extra constraints added will be as follows: $\beta \leq 1.2$ and $tmax \leq 13.2$

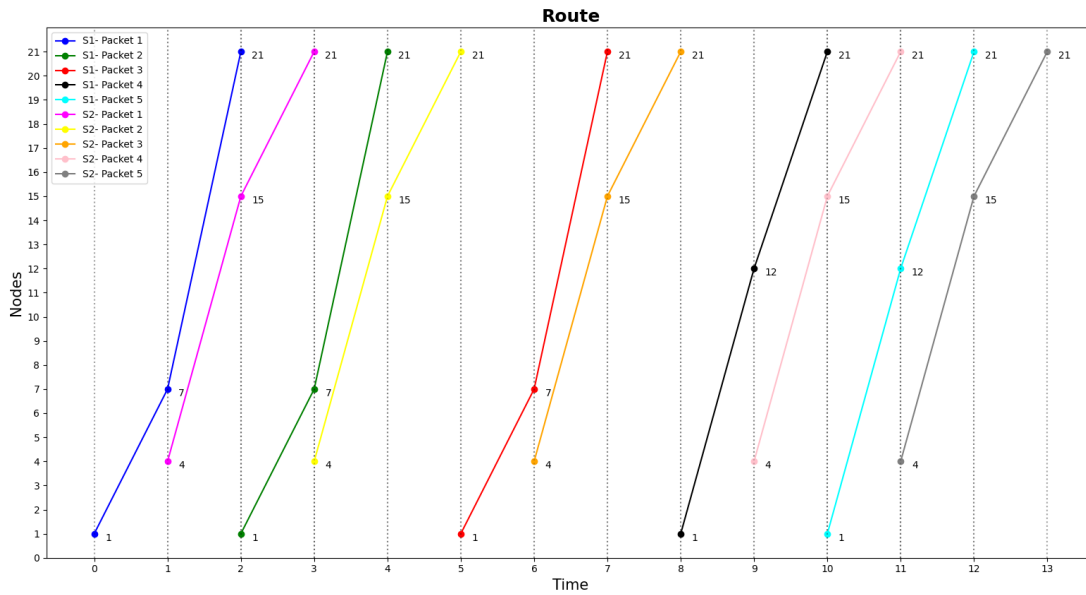


Figure 20: Final Chart (M1-21)

```

|---- 133 VARIABLE y.L
INDEX 1 = 1 INDEX 2 = 21

```

	1	2	3	4	6	7
1 .7	1.000		1.000		1.000	
7 .21		1.000	1.000	1.000		1.000
+	9	10	11	12		
1 .12	1.000		1.000			
12.21		1.000		1.000		

```

INDEX 1 = 4 INDEX 2 = 21

```

	2	3	4	5	7	8
4 .15	1.000		1.000		1.000	
15.21		1.000		1.000		1.000
+	10	11	12	13		
4 .15	1.000		1.000			
15.21		1.000		1.000		

Figure 21: The route for run 4(M1-21)

Figures 20 and 21 show the route of the final chart when adding the two extra constraints and minimizing the number of hops.

It can be seen from Figure 22 that the number of hops stayed the same as 20 since the model is minimizing that specific variable but the $tmax$ increased by 2 units of time to 13. The beta remained the same as the optimal of 1. The run time was 4 seconds to find the optimal solution.

133	VARIABLE	beta.L	=	1.000
	VARIABLE	bn.L	=	20.000
	VARIABLE	tmax.L	=	13.000

Figure 22: Variable Results (M1-21)

The objective function is notated as bn in Figure 22 which means that the shortest path in the fourth run need 20 hops at least to send all the packets required in the network.

6.1.3. Results for the 21 nodes V-model:

The results of the V-variable only:

The following Figure shows the optimal solution when $tmax$ and beta are multiplied by zero and the model is only minimizing the shortest path separately. The model got an optimal solution in 16 seconds with a value of 20 hops.

```

Proven optimal solution.

MIP Solution:          20.000000      (356 iter:
Final Solve:          20.000000      (0 iterat:

Best possible:          20.000000
Absolute gap:           0.000000
Relative gap:           0.000000

--- Restarting execution
--- V-only.gms(130) 23 Mb
--- Reading solution for model Modell
--- V-only.gms(130) 56 Mb
--- Executing after solve: elapsed 0:00:16.580

```

Figure 23: Optimal Shortest Path Value (M2-21)

Figures 24 and 25 show the route that each packet takes from a source to a sink and at what time. The indices are as follows; index 1 is the source, index 2 is the sink node and index 3 is the packet. For example, at time 9 packet number 4 for the 1st source and the sink node pair will be sent from the source node 1 to the intermediate node 12.

```

----- 133 VARIABLE v.L
INDEX 1 = 1  INDEX 2 = 21  INDEX 3 = 1
          33          34
1 .12      1.000
12.21      1.000
INDEX 1 = 1  INDEX 2 = 21  INDEX 3 = 2
          49          50
1 .12      1.000|
12.21      1.000
INDEX 1 = 1  INDEX 2 = 21  INDEX 3 = 3
          47          48
1 .12      1.000
12.21      1.000
INDEX 1 = 1  INDEX 2 = 21  INDEX 3 = 4
          9          10
1 .12      1.000
12.21      1.000
INDEX 1 = 1  INDEX 2 = 21  INDEX 3 = 5
          2          3

```

Figure 24: The Shortest Path (M2-21)

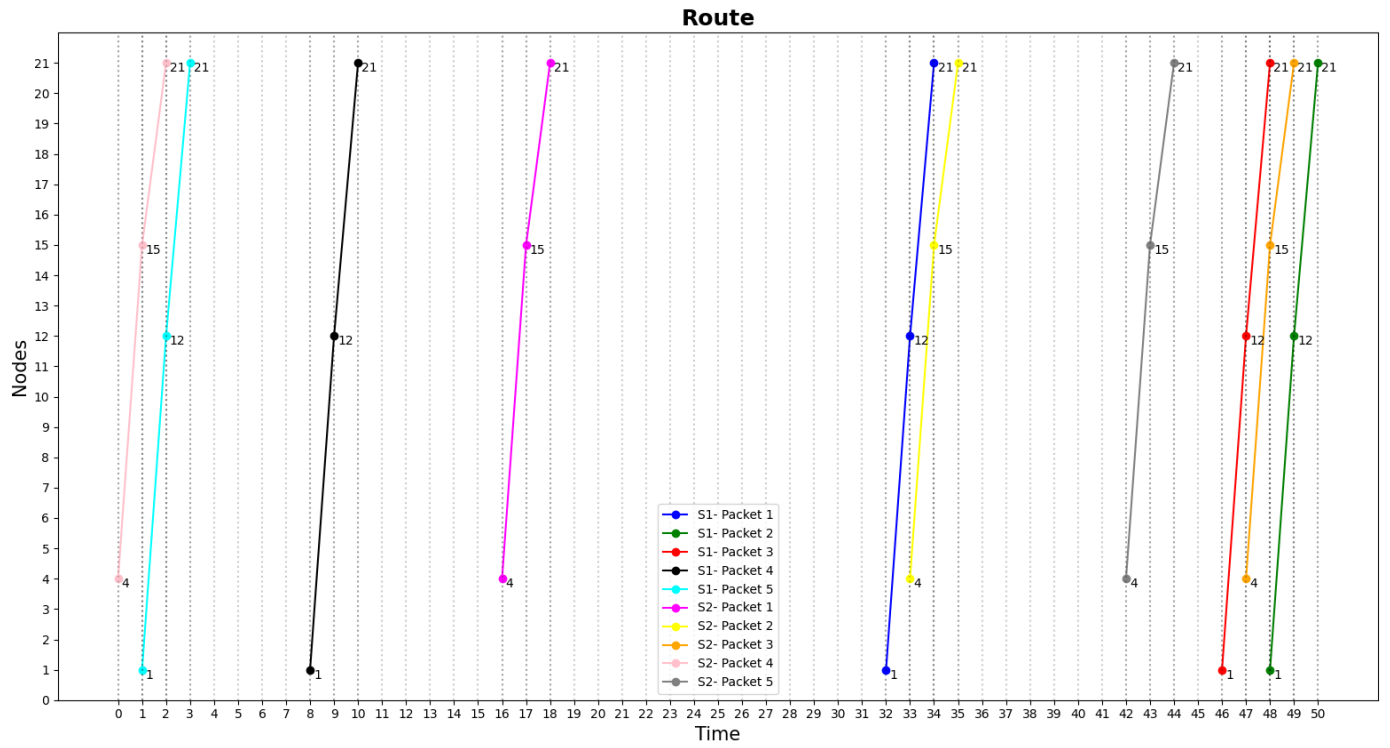


Figure 25:Path Chart (M2-21)

The results of the beta only:

The optimal value is found to be 1 in Figure 26 which is the lowest it can be for the global maximum difference in traffic between two grids. The solver calculated the value in 10 seconds. Figure 27 shows the sample route of two packets when beta is minimized. Figure 28 shows the complete path when beta is minimized. We can observe that the red path is a bit lengthy in Figure 28 and that's because the solver is only minimizing beta and doesn't have any weight for the route length and that changes completely when minimizing t_{max} or the final model.

```
Proven optimal solution.
MIP Solution:          1.000000      (2447 ite
Final Solve:           1.000000      (0 iterat
Best possible:         1.000000
Absolute gap:          0.000000
Relative gap:          0.000000

--- Restarting execution
--- Vmodel-beta-only.gms (132) 15 Mb
--- Reading solution for model Model1
--- Vmodel-beta-only.gms (132) 35 Mb
--- Executing after solve: elapsed 0:00:09.958
```

Figure 26: Beta Optimal Value (M2-21)

Figure 27: Sample Route (M2-21)



The lowest time to send the 10 packets for both sources needed 11 units of time to finish as it can be seen in Figure 29. Figure 30 show the route of the packets while t_{max} is minimized. The elapsed time was 1 min and 55 seconds to find the optimal t_{max} value.

Proven optimal solution.

MIP Solution: 11.000000 (181853 it
Final Solve: 11.000000 (0 iterati

Best possible: 11.000000
Absolute gap: 0.000000
Relative gap: 0.000000

--- Restarting execution
--- Vmodel-tmax-only.gms(132) 23 Mb
--- Reading solution for model Modell
--- Vmodel-tmax-only.gms(132) 56 Mb
--- Executing after solve: elapsed 0:01:55.113

Figure 29: tmax Optimal Value (M2-21)

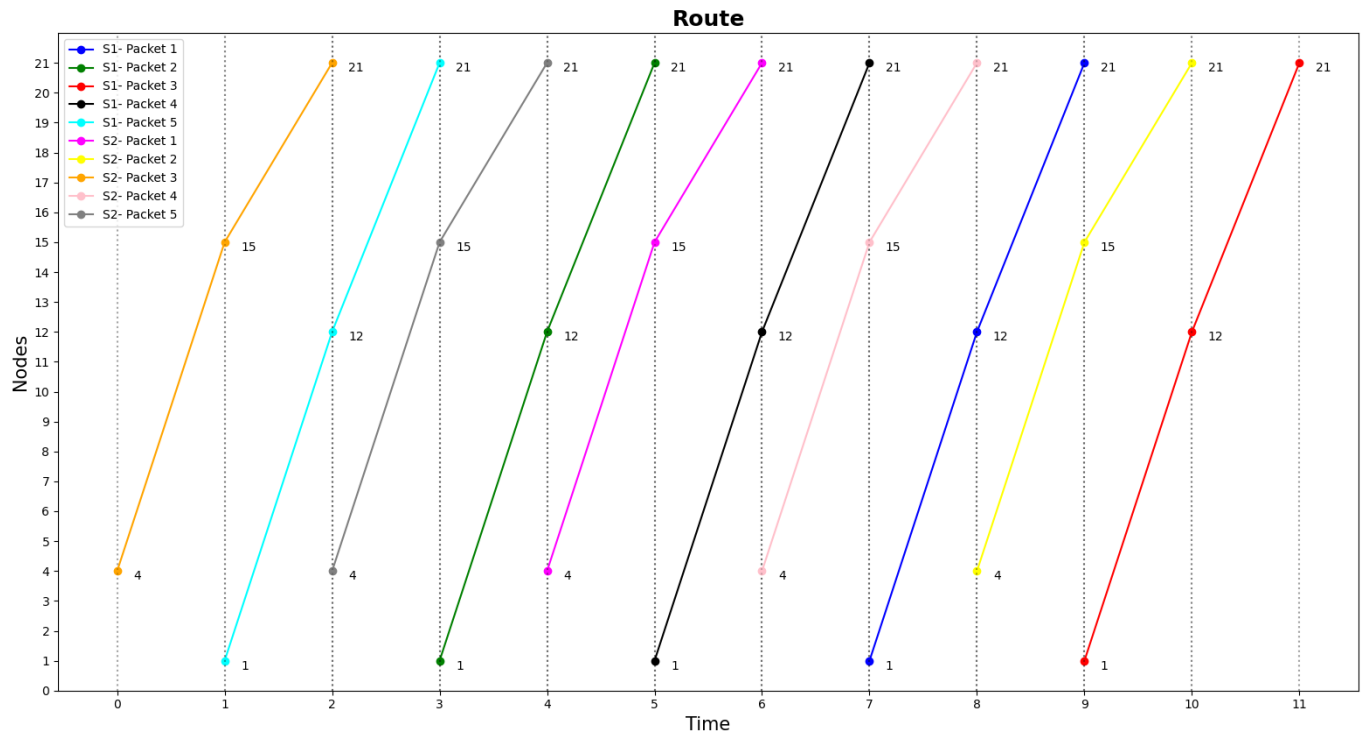


Figure 30: Tmax Chart (M2-21)

The results of the v-variable only while adding the two extra constraints:

In this part we will minimize the V-variable while giving an upper bound to the $tmax$ and the beta. The optimal beta and $tmax$ values respectively were 1 and 11, the two extra constraints added will be as follows: $\beta \leq 1.2$ and $tmax \leq 13.2$

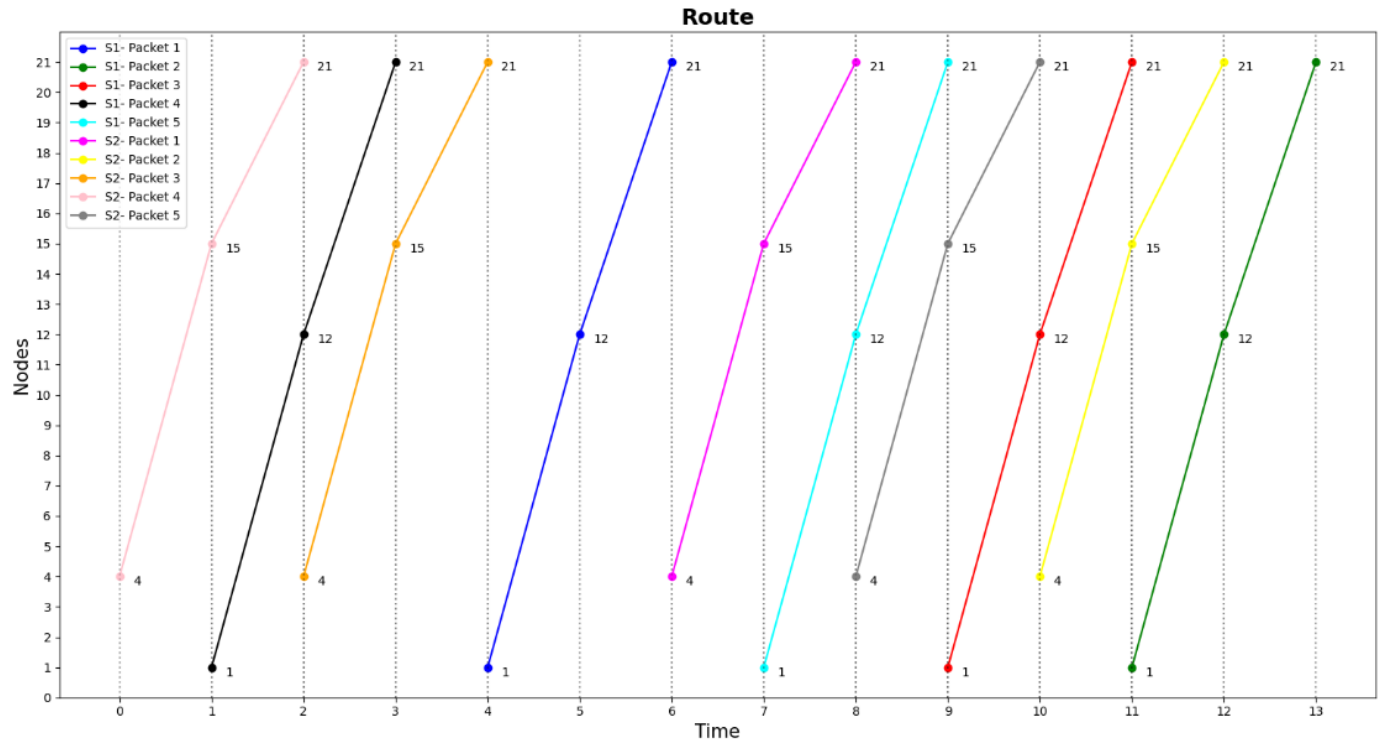


Figure 31:Final Chart (M2-21)

INDEX 1 = 1	INDEX 2 = 21	INDEX 3 = 1
	5	6
1 .12	1.000	
12.21		1.000
INDEX 1 = 1	INDEX 2 = 21	INDEX 3 = 2
	12	13
1 .12	1.000	
12.21		1.000
INDEX 1 = 1	INDEX 2 = 21	INDEX 3 = 3
	10	11
1 .12	1.000	
12.21		1.000
INDEX 1 = 1	INDEX 2 = 21	INDEX 3 = 4
	2	3
1 .12	1.000	
12.21		1.000
INDEX 1 = 1	INDEX 2 = 21	INDEX 3 = 5
	8	9
1 .12	1.000	
12.21		1.000

Figure 32: The route for run 4(M2-21)

It can be seen from Figures 31 and 32 that the number of hops stayed the same as 20 since the model is minimizing that specific variable but the $tmax$ increased by 2 units of time to 13. The beta remained the same as the optimal of 1. The run time was 16 seconds to find the optimal solution. The objective function is notated as bn in Figure 33 which means that the shortest path in the final run needs 20 hops at least to send all the packets required in the network.

```

133 VARIABLE beta.L           =          1.000
    VARIABLE bn.L             =         20.000
    VARIABLE tmax.L           =         13.000

```

Figure 33: Variable Results (M2-21)

6.2. Results of 30 nodes base model network:

6.2.1. The results of the X-variable only:

The following Figure shows the optimal solution when $tmax$ and $beta$ are multiplied by zero and the model is only minimizing the shortest path separately. The model got an optimal solution in 6 seconds with a value of 40 hops.

```

Proven optimal solution.

MIP Solution:          40.000000
Final Solve:           40.000000

Best possible:         40.000000
Absolute gap:           0.000000
Relative gap:           0.000000

--- Restarting execution
--- xmodel-xvar.qms(119) 7 Mb

```

Figure 34: Optimal Shortest Path Value (BM-30)

Figures 35 and 36 show the route that the packets take from a source to a sink and at what time. For example, the first packet is sent from source 1 to intermediate node 13 then at time 2 the packet is sent from node 13 to the intermediate node 23 until it reaches the sink node (#30) at time 4 from node 27.

----- 122 VARIABLE x _i .L						
	1	2	3	4	10	11
1 .13	1.000					
6 .17					1.000	
13.23		1.000				
17.25						1.000
23.27			1.000			
27.30				1.000		
	+	12	13	15	16	17
6 .17				1.000		
17.25					1.000	
25.27						1.000
25.28	1.000					
27.30						1.000
28.30		1.000				

Figure 35: The Shortest Path (BM-30)

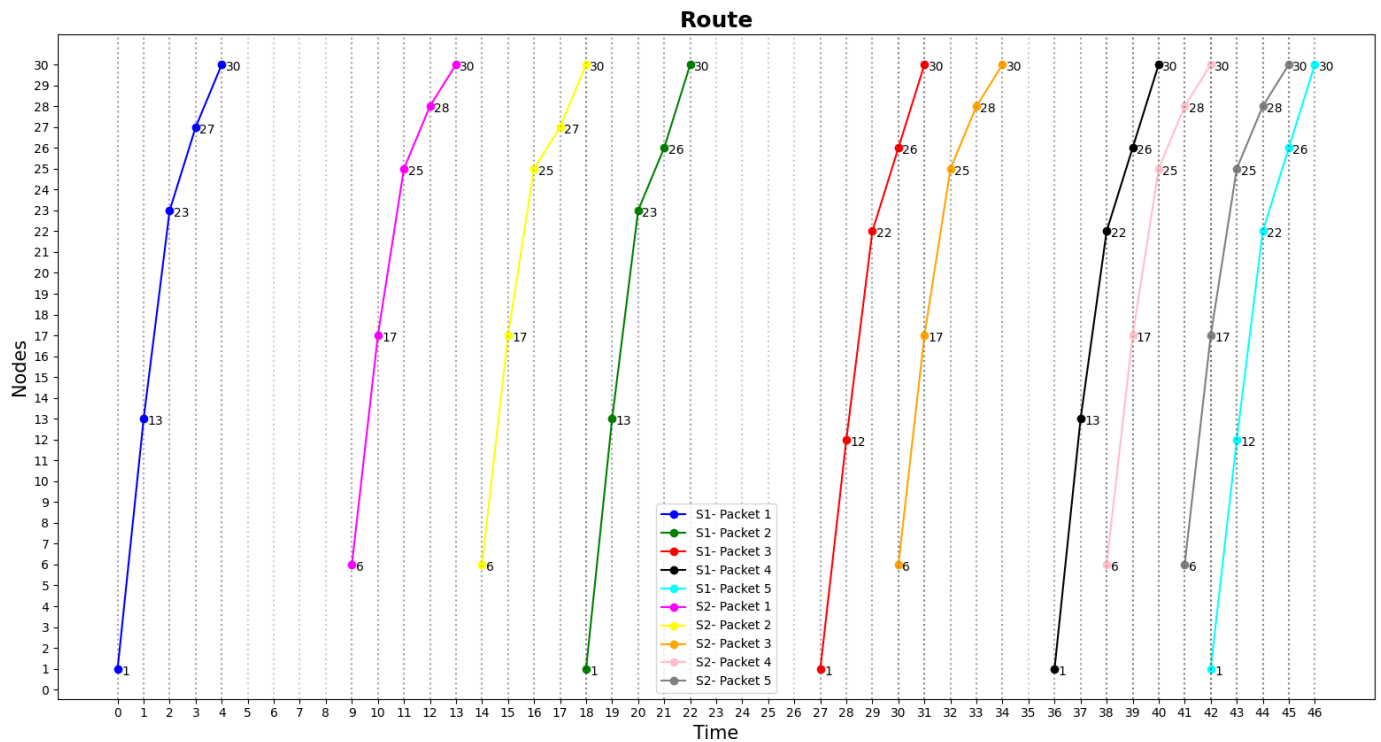


Figure 36 Path Chart (BM-30)

The results of the beta only:

The optimal value is found to be 1 which is the lowest it can be for the global maximum difference in traffic between two grids. The solver calculated the value in 14 seconds. Figure 38 shows the route when beta is the only objective function being minimized.

```

Proven optimal solution.

MIP Solution:          1.000000
Final Solve:           1.000000

Best possible:         1.000000
Absolute gap:          0.000000
Relative gap:          0.000000

--- Restarting execution
--- xmodel-beta.qms(119) 7 Mb

```

Figure 37: Beta Optimal Value (BM-30)

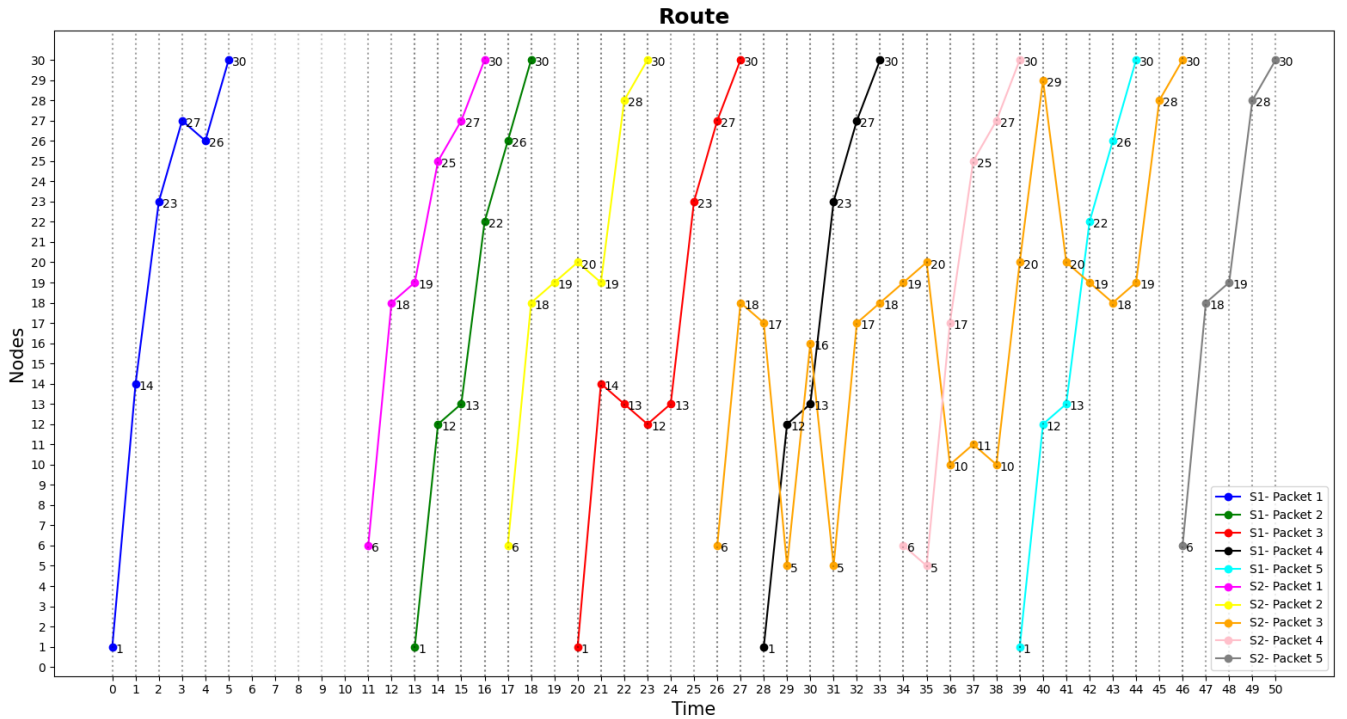


Figure 38 Beta Chart (BM-30)

The results of the t_{max} only:

The lowest time to send the 10 packets for both sources needed 17 units of time to finish as it can be seen in Figure 39. Figures 40 and 41 shows the route of the packets while t_{max} is minimized.

The elapsed time was 422 seconds to find the optimal t_{max} value.

Proven optimal solution.

MIP Solution: 17.000000 (5989973 i
Final Solve: 17.000000 (0 iterati

Best possible: 17.000000
Absolute gap: 0.000000
Relative gap: 0.000000

--- Restarting execution
--- xmodel-tmax.gms(119) 7 Mb
--- Reading solution for model Modell
--- xmodel-tmax.gms(119) 15 Mb
--- Executing after solve: elapsed 0:07:21.637

Figure 39: $t_{\max}$ Optimal Value (BM-30)

122 VARIABLE x.L						
	1	2	3	4	5	6
1 .12	1.000					
1 .13				1.000		
6 .7	1.000					
6 .18					1.000	
7 .8		1.000				
8 .19			1.000			
12.22		1.000				
13.22					1.000	
18.19						1.000
19.28				1.000		
22.26			1.000			1.000
26.30				1.000		
28.30					1.000	

Figure 40: Sample Path when $t_{\max}$ is minimized (BM-30)

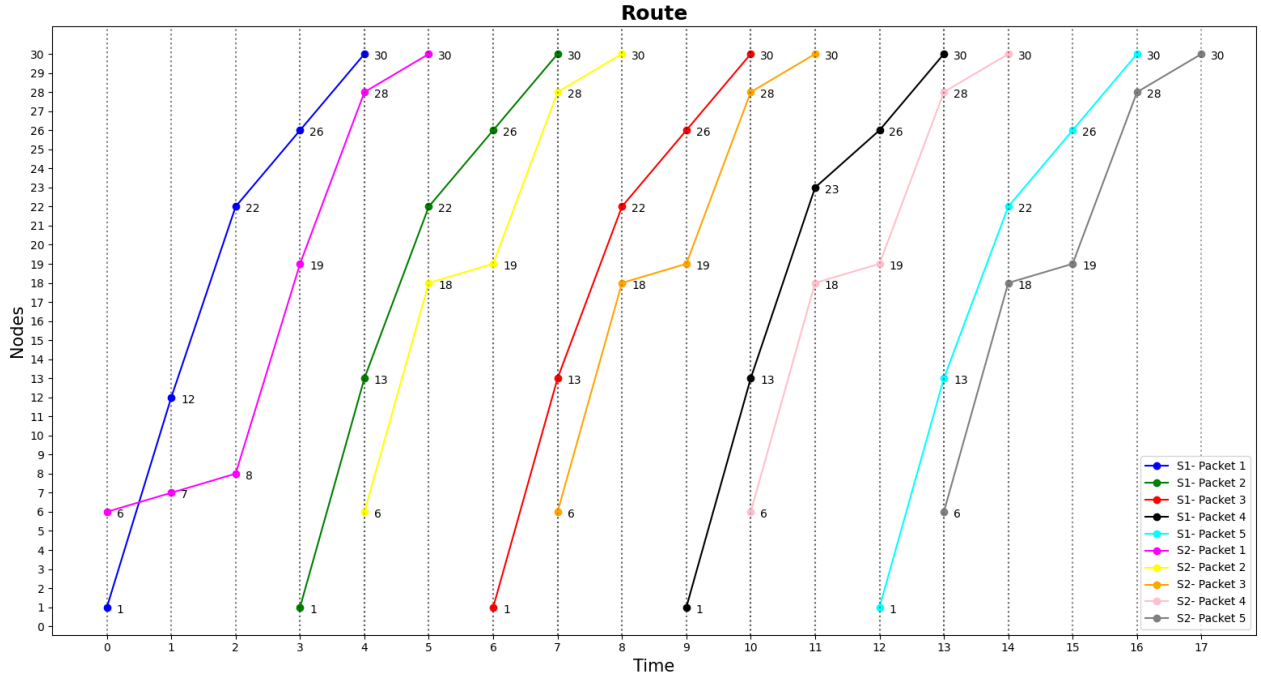


Figure 41: Tmax chart (BM-30)

The results of the x-variable only while adding the two extra constraints:

In this part we will minimize the x-variable while giving an upper bound to the tmax and the beta. The optimal beta and tmax values respectively were 1 and 17, the two extra constraints added will be as follows: $\beta \leq 1.2$ and $tmax \leq 20$.

```
Proven optimal solution.

MIP Solution:          48.000000      (520772 it
Final Solve:           48.000000      (0 iterati

Best possible:         48.000000
Absolute gap:          0.000000
Relative gap:          0.000000

--- Restarting execution
--- xmodel-30 nodes final.gms(121) 8 Mb
--- Reading solution for model Modell
--- xmodel-30 nodes final.gms(121) 18 Mb
--- Executing after solve: elapsed 0:01:40.737
--- xmodel-30 nodes final.gms(128) 19 Mb
```

Figure 42: Optimal Shortest Path Value (BM-30)

It can be seen from Figure 42 that the number of hops increased by 8 to 48 hops. The run time to find the optimal solution is 101 seconds.

Figure 43 shows that the beta stayed the same as the optimal value of 1. The total time increased from 17 to 20. One can argue why did the t_{max} change and that is simply because the model is minimizing the number of hops to finish the specified packets and to minimize the number of hops its relaxing the t_{max} within the tolerance that was chosen (20%).

124	VARIABLE beta.L	=	1.000
	VARIABLE bn.L	=	48.000
	VARIABLE tmax.L	=	20.000

Figure 43: Variable results (BM-30)

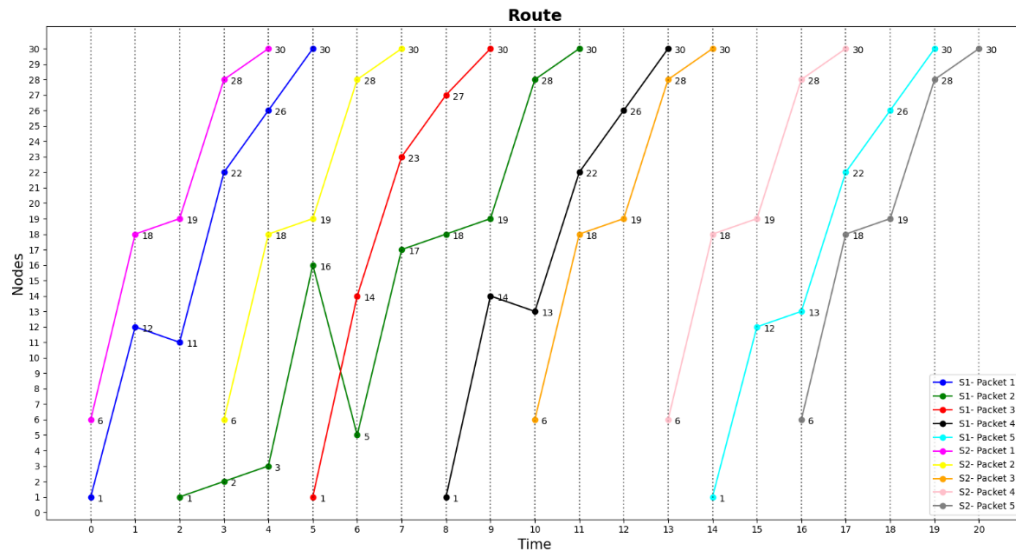


Figure 44: Final chart (BM-30)

Figures 44 and 45 show the shortest path after adding the 2 constraints to the model and minimizing the X variable which is basically finding the shortest path of the 30 nodes network.

124 VARIABLE x.L						
	1	2	3	4	5	6
1 .2			1.000			
1 .12	1.000					
1 .14						1.000
2 .3				1.000		
3 .16					1.000	
6 .18	1.000			1.000		
11.22			1.000			
12.11		1.000				
16.5						1.000
18.19		1.000			1.000	
19.28			1.000			1.000
22.26				1.000		
26.30					1.000	
28.30				1.000		
+						
	7	8	9	10	11	12
1 .14			1.000			
5 .17	1.000					
6 .18					1.000	
13.22					1.000	
14.13				1.000		
14.23	1.000					
17.18		1.000				
18.19			1.000			1.000
19.28				1.000		
22.26						1.000
23.27		1.000				
+						
	13	14	15	16	17	18
1 .12			1.000			
6 .18		1.000			1.000	
12.13				1.000		
13.22					1.000	
18.19			1.000			1.000
19.28	1.000			1.000		
22.26						1.000
26.30	1.000					
28.30		1.000			1.000	
+						
	19	20				
19.28	1.000					
26.30	1.000					
28.30		1.000				

Figure 45: The shortest path (BM-30)

```

---- 124 VARIABLE alpha.L

```

	1	2	3	4	5	6
1		2.000	2.000	2.000	2.000	2.000
2	2.000				2.000	2.000
3	2.000				2.000	2.000
4	2.000				2.000	2.000
5	2.000	2.000	2.000	2.000		2.000
6	2.000	2.000	2.000	2.000	2.000	
7	2.000				2.000	2.000
8	2.000				2.000	2.000
9	2.000				2.000	2.000
10	2.000	1.000	1.000	1.000	1.000	2.000
11	2.000	2.000	2.000	2.000	1.000	2.000
12	2.000	2.000	2.000	2.000	2.000	2.000
13	2.000				2.000	2.000
14	2.000				2.000	2.000
15	2.000	1.000	1.000	1.000	2.000	2.000
16	2.000	1.000	1.000	1.000	1.000	2.000

Figure 46: Sample Alpha Values

Figure 46 shows the alpha values in all grids for the 30 nodes network when beta isn't minimized. Some values can be seen to be 2 and that means that the traffic difference between two grids was 2 and that isn't the optimal value. This happens in the runs that isn't prioritizing to minimize the beta in the network, but instead it just minimizes the number of hops or the total time to send all the packets in the network.

This means that the probability of having the sink node in the grids that have a maximum difference of 2 will be higher than the grid cells with a value of 1 and that's why we need to minimize the beta in the model, so we fool the adversary by balancing the traffic difference in the entire network. Figure 47 shows the value of beta when at least one value of alpha is 2.

```

---- 124 VARIABLE beta.L = 2.000
      VARIABLE bn.L = 40.000
      VARIABLE tmax.L = 20.000

```

Figure 47: Variable Values (beta=2)

----	124 VARIABLE alpha.L					
	1	2	3	4	5	6
1		1.000	1.000	1.000	1.000	1.000
2	1.000		1.000	1.000	1.000	1.000
3	1.000	1.000		1.000	1.000	1.000
4	1.000	1.000	1.000		1.000	1.000
5	1.000	1.000	1.000	1.000		1.000
6	1.000	1.000	1.000	1.000	1.000	
7	1.000	1.000	1.000	1.000	1.000	1.000
8	1.000	1.000	1.000	1.000	1.000	1.000
9	1.000	1.000	1.000	1.000	1.000	1.000
10	1.000	1.000	1.000	1.000	1.000	1.000
11	1.000	1.000	1.000	1.000	1.000	1.000
12	1.000	1.000	1.000	1.000	1.000	1.000
13	1.000	1.000	1.000	1.000	1.000	1.000
14	1.000	1.000	1.000	1.000	1.000	1.000
15	1.000	1.000	1.000	1.000	1.000	1.000
16	1.000	1.000	1.000	1.000	1.000	1.000

Figure 48: Sample Alpha Values 1

Figure 48 shows the alpha values when we are minimizing the beta in specific runs such as the fourth run when adding two constraints and one of them is related to minimizing the beta. Since beta is the global maximum difference between two grids then the value of all the alphas in the grid will be 1, and that is the optimal value of beta that we can reach in this network. By doing so, we are balancing the traffic difference between two grids in the whole network to make it harder for the attacker to reach the sink node. The variance of the Alpha values in Figure 48 is 0.05859375.

6.2.1. Sensitivity analysis of the 30 nodes base model

This subsection will show the impact of changing the tolerance of the different objective functions on each other. A pareto curve is constructed to show the values of the routing cost (X-variable) when changing the values of the total time needed to finish all the packets in the network (t_{max}) for two values of the global maximum traffic difference between two grids (β).

Pareto Front

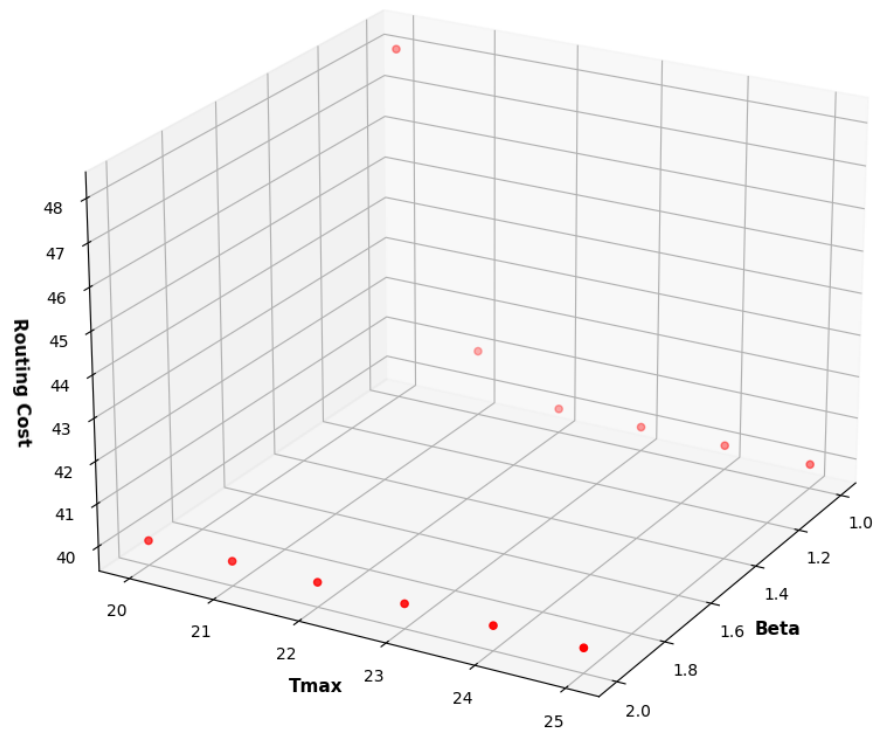


Figure 49: Pareto Front

Two curves can be seen in Figure 49 that represent the routing cost for different β and t_{max} values. It can be seen from the Figure that the value of the routing cost is 48 when the t_{max} value is 20 and β is 1.

The value decreases to 41 for an increasing value of t_{max} and that is caused by giving a higher tolerance to the t_{max} . The value of the routing cost decreases to the optimal when t_{max} is greater than or equal to 22.

The routing cost is constant for any value of β and t_{max} . This happens because the model isn't pressured to minimize the value of β , so the routing path that it gives for the solution is less tight than the path when β is equal to 1. The value of the routing cost for t_{max} greater than or equal to 22 is always 40 which is the optimal value. The lowest value of t_{max} is 20 which is the optimal value when minimizing all the objective functions from the previous section that's why it's the first value in the previous figure since a value of t_{max} that is equal to 19 will give an infeasible solution.

The behavior of the different objective functions will be the same in the two extended models since we have the same network. If we have different parameters for a completely different WSN then the values will differ.

3D Surface Plot

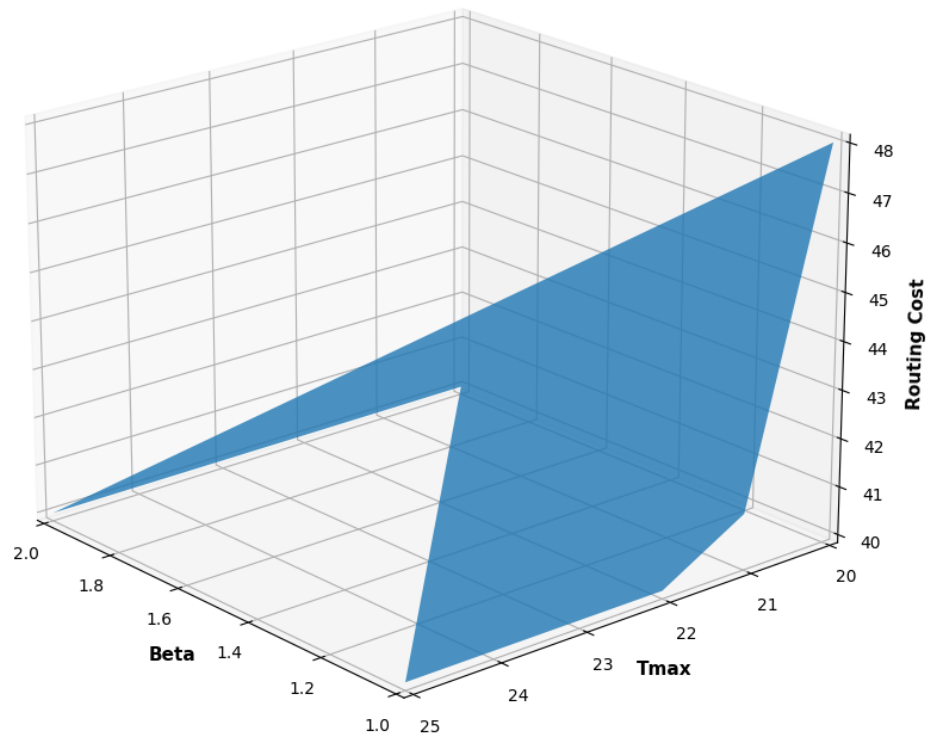


Figure 50: Surface plot

Figure 50 shows the 3D surface plot of the objective functions and their impact on each other in a detailed manner. For example, the value of the routing cost is 48 when t_{max} is 20 and the beta is 1.

6.2.2. Results of the 30 nodes Y-model:

The results of the Y-variable only:

The following Figure shows the optimal solution when $tmax$ and $beta$ are multiplied by zero and the model is only minimizing the shortest path separately. The model got an optimal solution in 7 seconds with a value of 40 hops.

```
Proven optimal solution.

MIP Solution:          40.000000      (1171 itera
Final Solve:           40.000000      (0 iteratic

Best possible:         40.000000
Absolute gap:          0.000000
Relative gap:          0.000000

--- Restarting execution
--- ymodel-yvar.gms(128) 11 Mb
--- Reading solution for model Modell1
--- ymodel-yvar.gms(128) 26 Mb
--- Executing after solve: elapsed 0:00:07.736
--- ymodel-yvar.gms(135) 26 Mb
```

Figure 51: Optimal Shortest Path Value (M1-30)

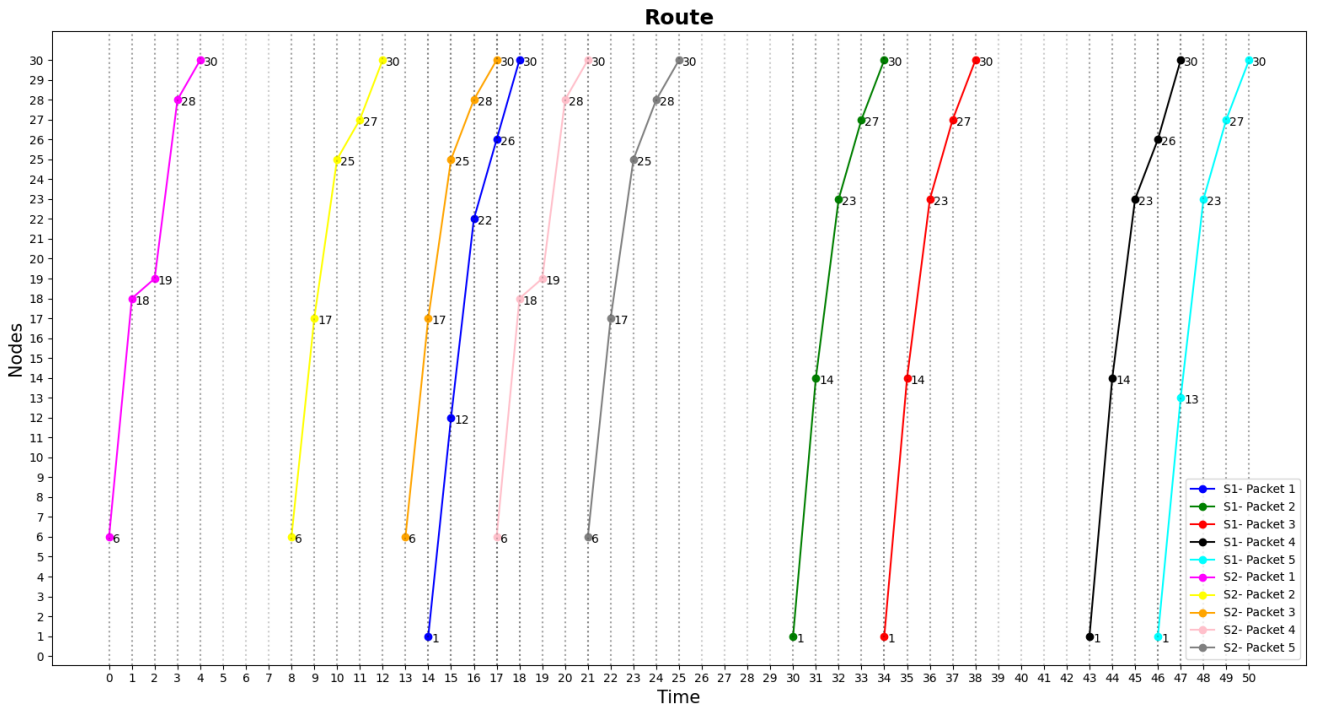


Figure 52: Shortest path chart (M1-30)

Figures 52 and 53 show the route that the packets take from source number 1 to the sink and at what time. For example, the first packet is sent from source 1 to intermediate node 12 at time 15 then the packet is sent from node 12 to the intermediate node number 22 and finally the sink node receives the packet at time 18 from node 26. The difference between the base model and the modified model is that we are observing the packets for each source and destination pair.

INDEX 1 = 1		INDEX 2 = 30					
		15	16	17	18	31	32
1 .12	1.000						
1 .14						1.000	
12.22			1.000				
14.23							1.000
22.26				1.000			
26.30					1.000		
+		33	34	35	36	37	38
1 .14				1.000			
14.23					1.000		
23.27	1.000					1.000	
27.30			1.000				1.000
+		44	45	46	47	48	49
1 .13					1.000		
1 .14	1.000						
13.23						1.000	
14.23			1.000				
23.26				1.000			
23.27							1.000
26.30					1.000		
+		50					
27.30	1.000						

Figure 53: The Shortest Path (M1-30)

The results of the beta only:

The optimal value is found to be 1 which is the lowest it can be for the global maximum difference in traffic between two grids. The solver calculated the value in 243 seconds. Figure 55 shows the path when beta is the only objective function minimized.

```
Proven optimal solution.

MIP Solution:          1.000000      (244047 iterations)
Final Solve:           1.000000      (0 iterations)

Best possible:         1.000000
Absolute gap:          0.000000
Relative gap:          0.000000

--- Restarting execution
--- ymodel-beta.gms(128) 11 Mb
--- Reading solution for model Modell1
--- ymodel-beta.gms(128) 26 Mb
--- Executing after solve: elapsed 0:04:05.703
--- ymodel-beta.gms(135) 26 Mb
```

Figure 54: Beta Optimal Value (M1-30)

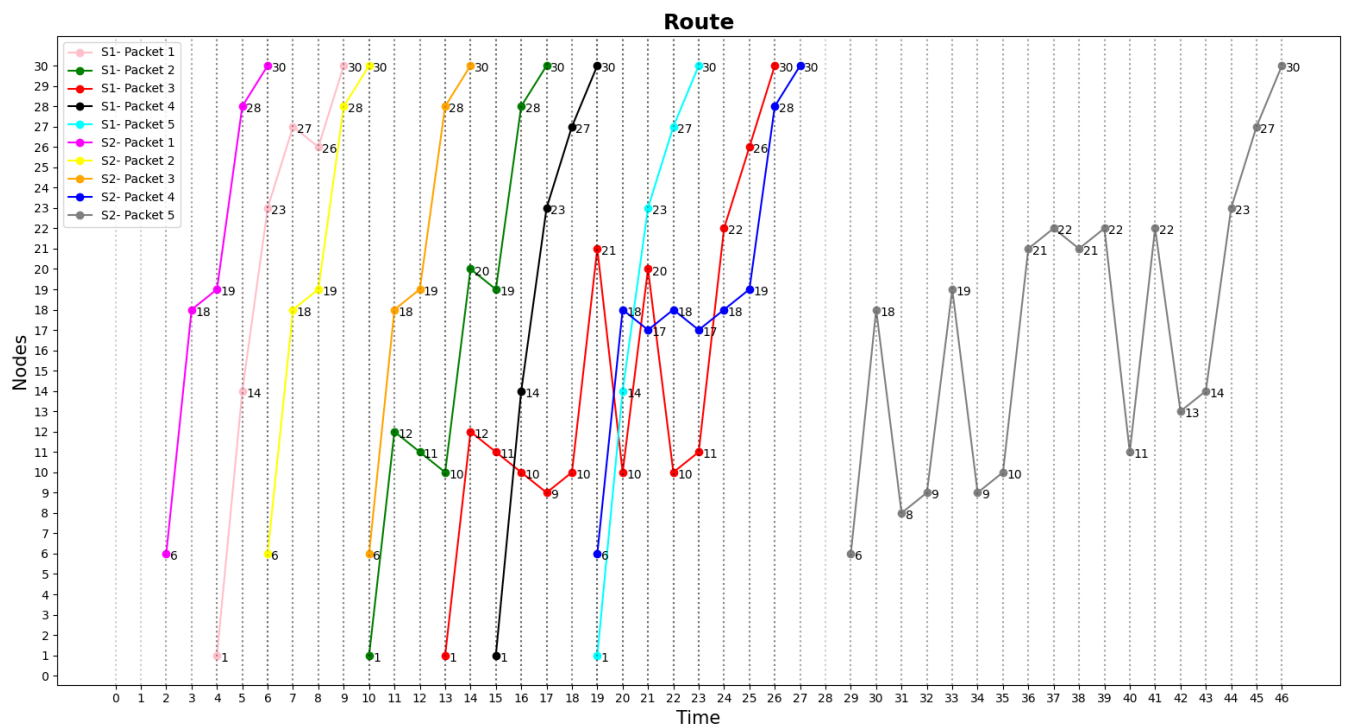


Figure 55: Beta Chart (M1-30)

The results of the t_{max} only:

The lowest time to send the 10 packets for both sources needed 17 units of time to finish as it can be seen in Figure 56. Figure 58 shows the route of the packets while t_{max} is minimized. The elapsed time was 58 min and 6 seconds to find the optimal t_{max} value.

```

Proven optimal solution.
MIP Solution:          17.000000      (31263954
Final Solve:          17.000000      (0 iterati
Best possible:          17.000000
Absolute gap:           0.000000
Relative gap:           0.000000

--- Restarting execution
--- ymodel-tmax.gms (128) 11 Mb
--- Reading solution for model Modell1
--- ymodel-tmax.gms (128) 26 Mb
--- Executing after solve: elapsed 0:58:06.181
--- ymodel-tmax.gms (135) 26 Mb

```

Figure 56: t_{max} Optimal Value (M1-30)

INDEX 1 = 1 INDEX 2 = 30							
		1	2	3	4	5	6
1 .12	1.000						
1 .13					1.000		
12.22			1.000				
13.23						1.000	
22.26				1.000			
23.26							1.000
26.30					1.000		
+	7	8	9	10	11	12	
1 .13	1.000			1.000			
13.22		1.000					
13.23					1.000		
22.26			1.000				
23.26							1.000
26.30	1.000			1.000			
+	13	14	15	16			
1 .13	1.000						
13.23		1.000					
23.26			1.000				
26.30	1.000			1.000			

Figure 57: Sample Path when t_{max} is minimized (M1-30)

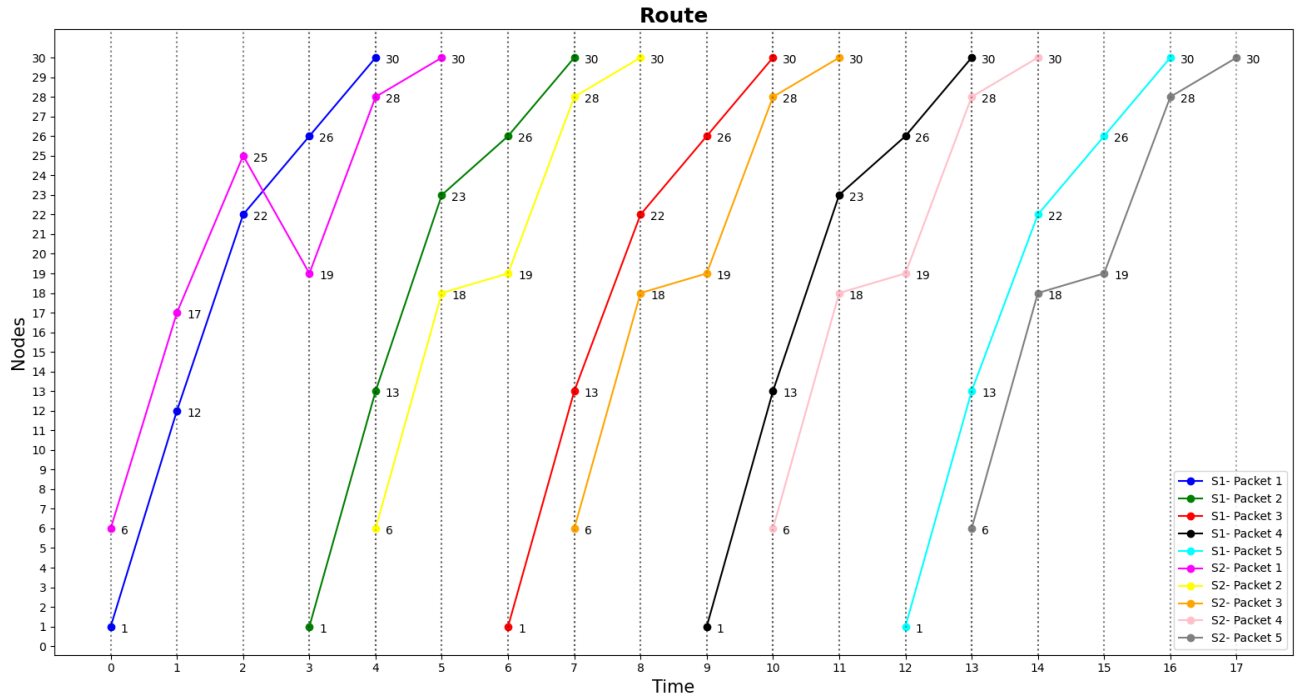


Figure 58: Tmax Chart (M1-30)

The results of the Y-variable only while adding the two extra constraints:

In this part we will minimize the X-variable while giving an upper bound to the $tmax$ and the beta. The optimal beta and $tmax$ values respectively were 1 and 17, the two extra constraints added will be as follows: $\beta \leq 1.2$ and $tmax \leq 20.4$

```
Proven optimal solution.

MIP Solution:          48.000000      (3512323 :
Final Solve:           48.000000      (0 iterat:

Best possible:         48.000000
Absolute gap:          0.000000
Relative gap:          0.000000

--- Restarting execution
--- ymodel-final.gms(130) 14 Mb
--- Reading solution for model Modell1
--- ymodel-final.gms(130) 33 Mb
--- Executing after solve: elapsed 0:10:05.478
--- ymodel-final.gms(137) 33 Mb
```

Figure 59: Optimal Shortest Path details (M1-30)

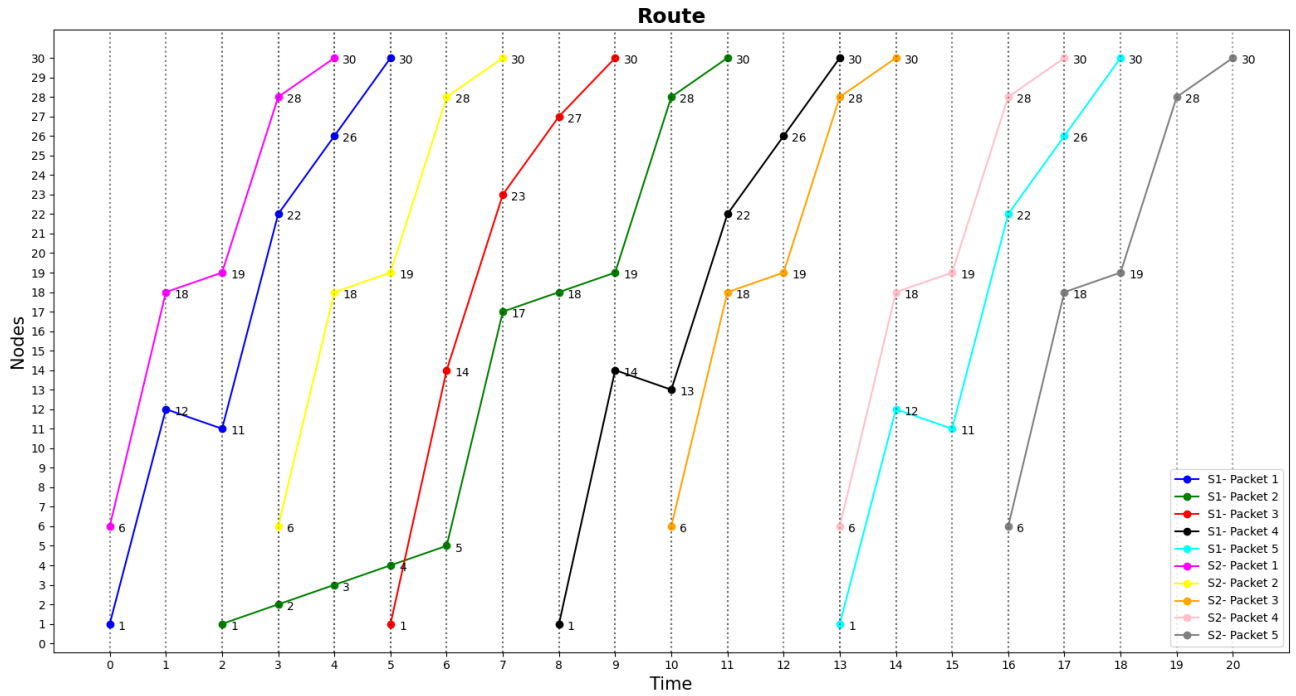


Figure 60: Final chart (M1-30)

---- 133 VARIABLE y.L

INDEX 1 = 1 INDEX 2 = 30

	1	2	3	4	5	6
1 .2			1.000			
1 .12	1.000					
1 .14						1.000
2 .3				1.000		
3 .4					1.000	
4 .5						1.000
11.22			1.000			
12.11		1.000				
22.26				1.000		
26.30					1.000	
+	7	8	9	10	11	12
1 .14			1.000			
5 .17	1.000					
13.22					1.000	
14.13				1.000		
14.23	1.000					
17.18		1.000				
18.19			1.000			
19.28				1.000		
22.26						1.000
23.27		1.000				
27.30			1.000			
28.30					1.000	

Figure 61: Sample route of run 4(M1-30)

It can be seen from Figures 60, 61 and 62 that the number of hops changed from 40 to 48 because we added the two constraints that are used to minimize the total time and the beta. Since we added these constraints the number of hops increased by 8 hops and the total time of the network increased by 3 units of time.

```

----      133 VARIABLE beta.L           =          1.000
            VARIABLE bn.L              =          48.000
            VARIABLE tmax.L            =          20.000

```

Figure 62: Variable Results (M1-30)

The objective function is notated as bn in Figure 61 which means that the shortest path in the fourth run need 48 hops at least to send all the packets required in the network.

6.2.3. Results of the 30 nodes V-model:

The results of the V-variable only:

Figure 63 shows the optimal solution when *tmax* and *beta* are multiplied by zero and the model is only minimizing the shortest path separately. The model got an optimal solution in 49 seconds with a value of 20. The number of hops isn't 20 in this case because there are some hops that cost zero. Figure 64 shows the complete shortest route.

```

Proven optimal solution.

MIP Solution:          20.000000      (2153 iter.
Final Solve:           20.000000      (0 iterati.

Best possible:          20.000000
Absolute gap:           0.000000
Relative gap:           0.000000

--- Restarting execution
--- Vvarmodel-vvar.gms (130) 59 Mb
--- Reading solution for model Model1
--- Vvarmodel-vvar.gms (130) 145 Mb
--- Executing after solve: elapsed 0:00:49.582

```

Figure 63: Optimal Shortest Path Value (M2-30)

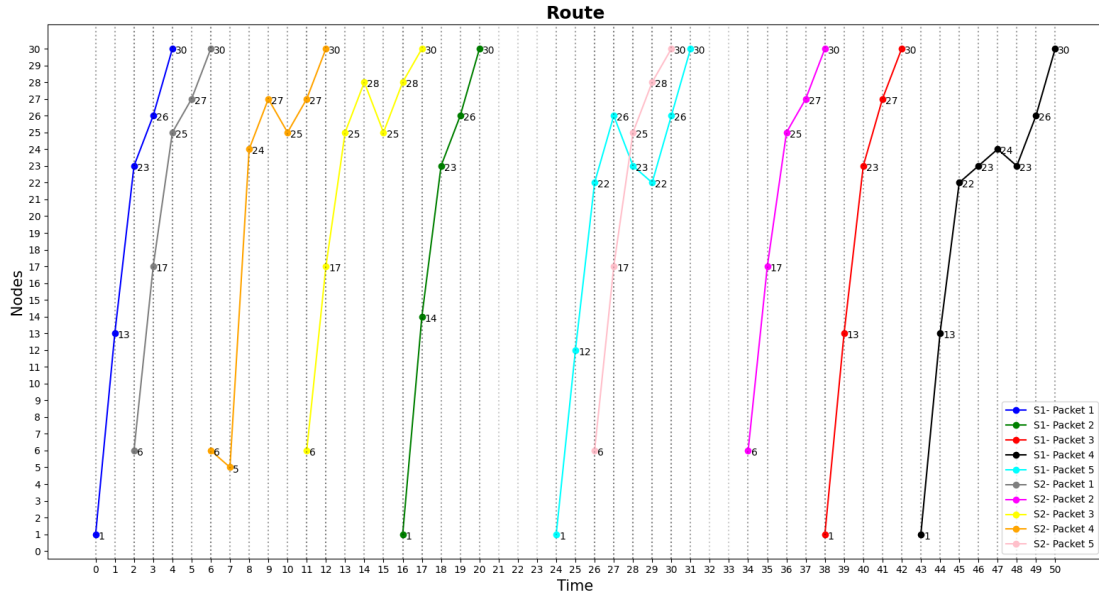


Figure 64: Shortest path chart(M2-30)

The results of the beta only:

The optimal value is found to be 1 which is the lowest it can be for the global maximum difference in traffic between two grids. The solver calculated the value in approximately 1 hr and 30 min. Figure 66 shows the path when beta is the only objective function minimized.

135	VARIABLE beta.L	=	1.000
	VARIABLE bn.L	=	31.000

Figure 65: Beta Optimal Value (M2-30)

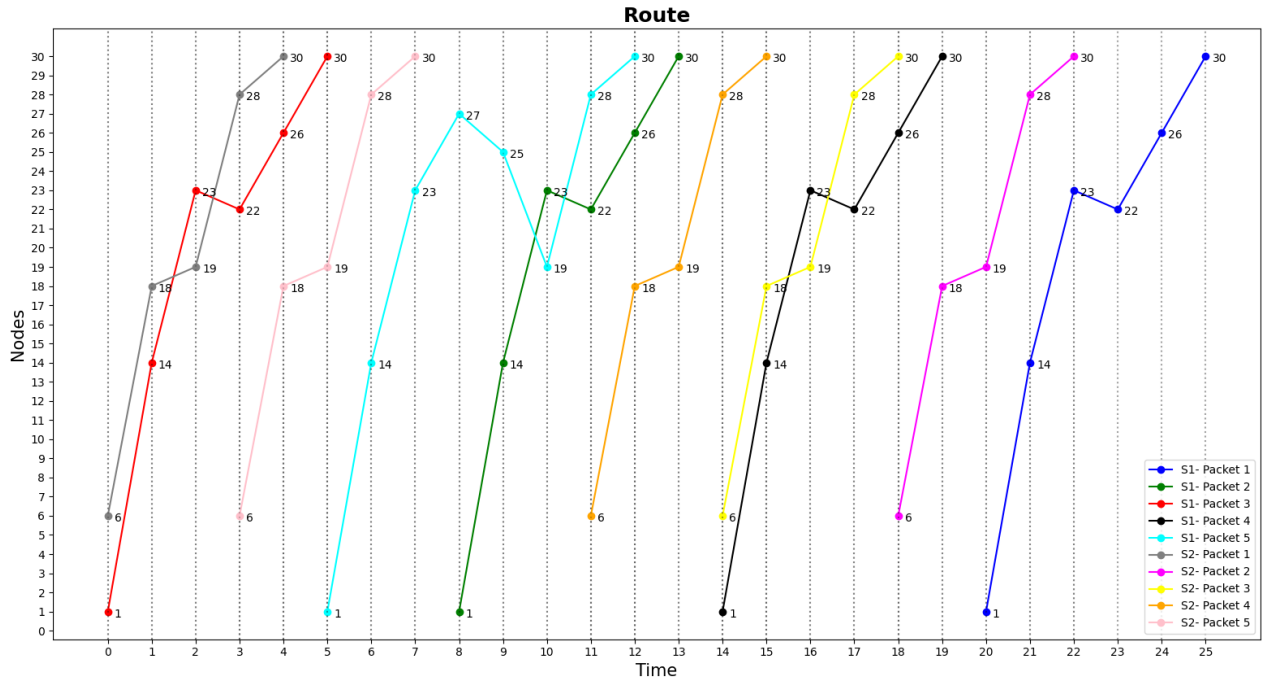


Figure 66: Beta Chart (M2-30)

The results of the t_{max} only:

The lowest time to send the 10 packets for both sources needed 17 units of time to finish as it can be seen in Figure 67. Figure 68 shows the route of the packets while t_{max} is minimized. The elapsed time was 8 min and 19 seconds to find the optimal t_{max} value.

Proven optimal solution.

MIP Solution: 17.000000 (848117 it
Final Solve: 17.000000 (0 iterati

Best possible: 17.000000
Absolute gap: 0.000000
Relative gap: 0.000000

--- Restarting execution
--- Vvarmodel-tmax.gms (132) 19 Mb
--- Reading solution for model Modell
--- Vvarmodel-tmax.gms (132) 47 Mb
--- Executing after solve: elapsed 0:08:18.090
--- Vvarmodel-tmax.gms (139) 47 Mb

Figure 67: t_{max} Optimal Value (M2-30)

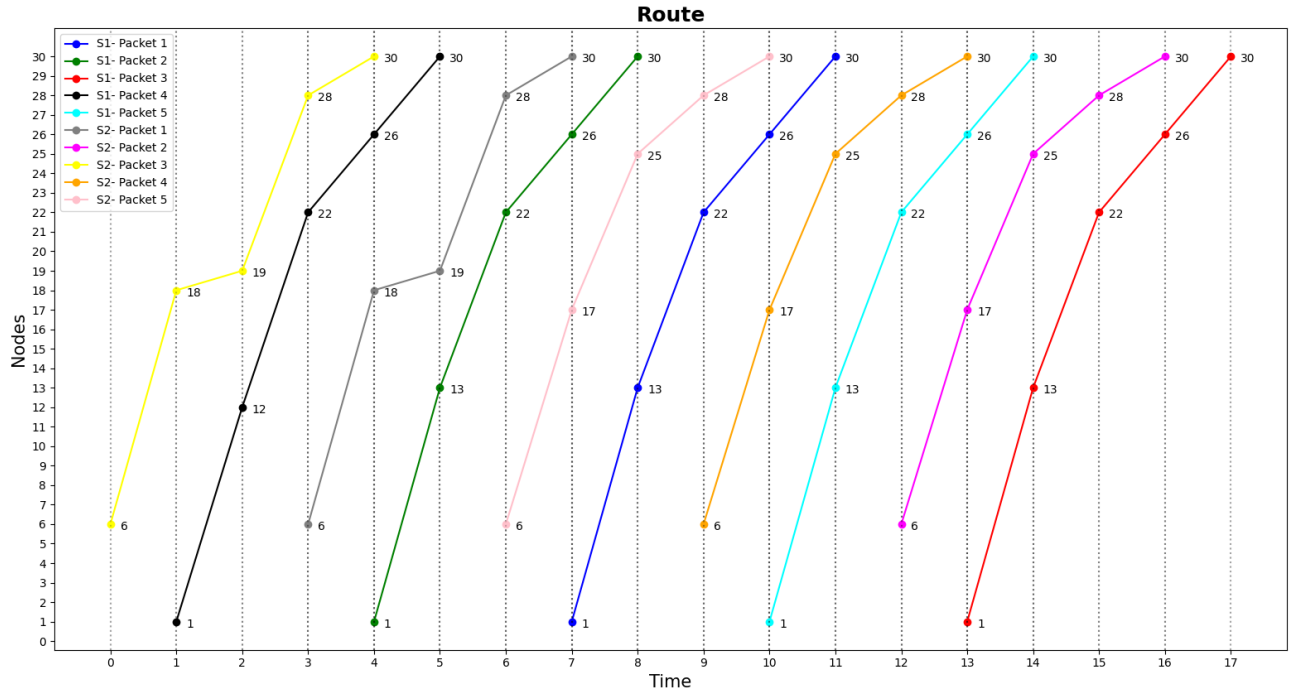


Figure 68: Tmax Chart (M2-30)

The results of the V-variable only while adding the two extra constraints:

In this part we will minimize the V-variable while giving an upper bound to the $tmax$ and the beta. The optimal beta and $tmax$ values respectively were 1 and 25, the two extra constraints added will be as follows: $\beta \leq 1.2$ and $tmax \leq 25$

Proven optimal solution.

MIP Solution: 30.000000 (12187313
Final Solve: 30.000000 (0 iterati

Best possible: 30.000000
Absolute gap: 0.000000
Relative gap: 0.000000

--- Restarting execution
--- Vvarmodel-final.gms(132) 24 Mb
--- Reading solution for model Modell
--- Vvarmodel-final.gms(132) 58 Mb
--- Executing after solve: elapsed 2:11:51.998
--- Vvarmodel-final.gms(139) 58 Mb

Figure 69: Optimal Shortest Path details (M2-30)

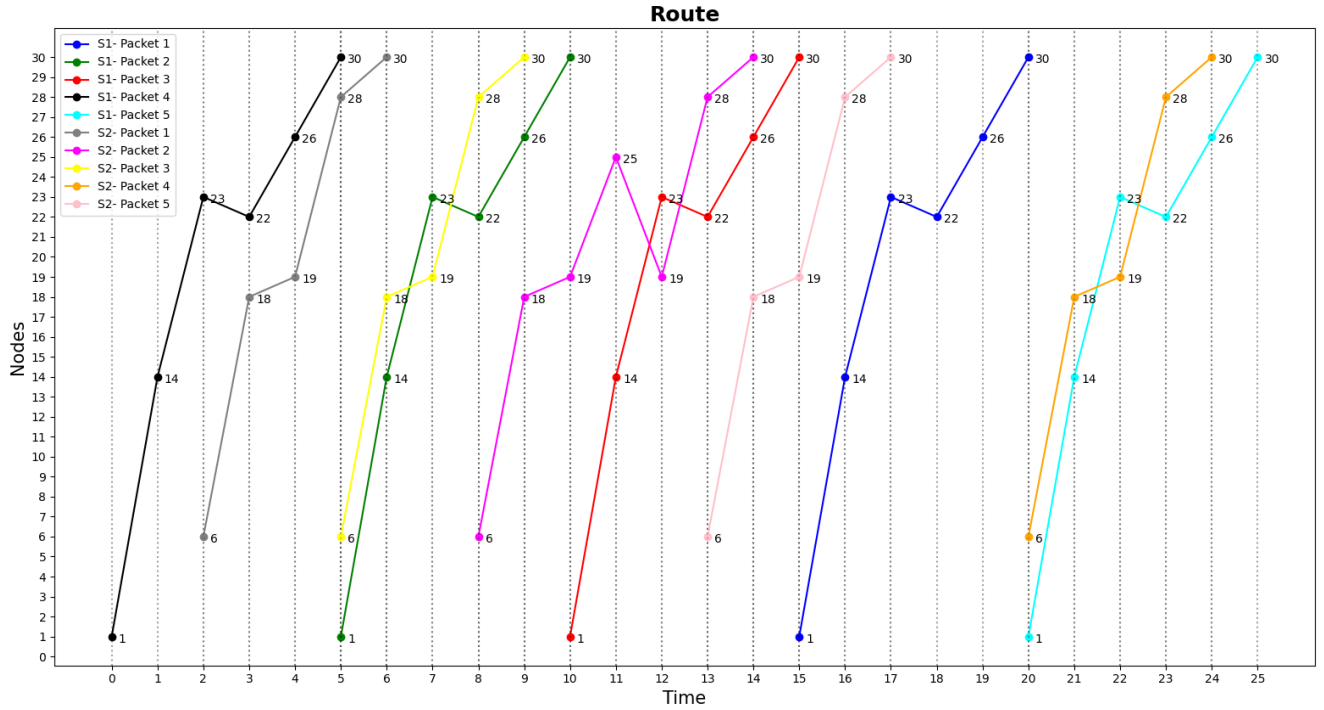


Figure 70: Final chart (M2-30)

It can be seen from Figures 69, 70 and 71 that the number of hops changed from 40 to 47 because we added the two constraints that are used to minimize the total time and the beta. Since we added these constraints the number of hops increased by 8 hops and the total time of the network increased by 5 units of time.

----	135	VARIABLE	beta.L	=	1.000
		VARIABLE	bn.L	=	30.000
		VARIABLE	tmax.L	=	25.000
----	147	PARAMETER	tes	=	47.000

Figure 71: Variable Results (M2-30)

The objective function is notated as bn in Figure 71 which means that the shortest path in the fourth run need 47 hops at least to send all the packets required in the network. The bn value is 30 and that doesn't represent the real number of hops because some of them are multiplied by zero to create a pattern near the sink node to distract the adversary. The real number of hops is shown in the parameter tes which is 47 hops. The tmax increased from 20 to 25 when compared

to the past models, and that's simply because of the pattern that we are creating near the sink node with the packets that doesn't cost anything to send or receive. This pattern will increase the total time of the entire network and that is the main reason why it increased by 5 units of time. The phantom routing technique is a second reason why it increases, because the model is forced to visit nodes that are not in the shortest path to satisfy the phantom routing condition. There is a tradeoff between the total time of the network and the pattern near the sink node that distracts the adversary.

The complexity of the model can be shown in terms of the number of variables and equations it has. The number of equations can be captured for each constraint by multiplying the cardinality of all the sets. The general formula to capture the number of equations for 2 sets A and B is $|A| \times |B|$.

For example, constraint 7 is $\forall s, d, r, t, i, j$ and to capture the number of equations in this constraint we need to multiply the cardinality of each set by the other. The number of equations is equal to $|S| \times |D| \times |R| \times |T| \times |I| \times |J| = 2 \times 1 \times 5 \times 50 \times 30 \times 30 = 450000$ equation. The number of equations of all constraints and variables for the 30 nodes network V model will be shown in the following tables.

# Constraint	# Of Equations
1	25600
2	800
3	256
4	10

5	28
6	13500
7	450000
8	10
9	15000
10	1500
11	1500
12	7000
13	1
14	256
15	450000
16	1500
17	10
SUM	966971

Table 8: Number of Equations

Variable	# Of variables
V_{sdrijt}	450000
a_{gk}	256
$\hat{P}_{gt}$	800
β	1
$tmax$	1

Table 9: Number of Variables

The number of variables and equations will differ slightly between the models, since we add two extra constraints to the final run of each model. The 21 nodes network will have different number of variables and constraints since the sets differ in number. This problem is NP hard since it has binary variables. Mathematical models that have integer variables are always NP-

hard. The number of variables of the first variable in Table 9 were captured by the following formula $|S| \times |D| \times |R| \times |T| \times |I| \times |J| = 2 \times 1 \times 5 \times 50 \times 30 \times 30 = 450000$ Variable. It will be calculated by multiplying the cardinality of the sets in the indices of each variable. If a variable doesn't have any indices, then it's considered as a single variable like *tmax* in Table 9.

CHAPTER 7: CONCLUSION

A literature review of WSN, and different privacy mechanisms in WSN is presented in this thesis. The focus of our work is on the sink node anonymity. We proposed a novel multi-objective mathematical model which deals with the route path length, total time of the network, transmission rate and energy consumption. Phantom routing techniques were embedded in the final extended model to help increase the sink node location anonymity by having the freedom to control the packets. In the final model, sending or receiving packets for specific nodes near the sink node doesn't influence the objective function by adding a cost parameter to create a pattern near the sink node to distract the adversary. Two networks were constructed as numerical case studies to validate the mathematical model that was created to increase the sink node anonymity. Twenty and thirty nodes' networks were created to showcase that the operations research approach is valid.

For future research recommendations, different optimization methodologies can be used to tackle the sink node anonymity in a WSN problem to reach a solution in a reasonable amount of time. This will be extremely helpful when dealing with larger networks and it will decrease the computational time drastically.

References

- [1] Ozturk, Celal, Yanyong Zhang, and Wade Trappe. "Source-location privacy in energy-constrained sensor network routing." In *Proceedings of the 2nd ACM workshop on Security of Ad hoc and Sensor Networks*, pp. 88-93. 2004.
- [2] Sweeney, Latanya. "k-anonymity: A model for protecting privacy." *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 10.05 (2002): 557-570.
- [3] Agrawal, Rakesh, Alexandre Evfimievski, and Ramakrishnan Srikant. "Information sharing across private databases." *Proceedings of the 2003 ACM SIGMOD international conference on Management of data*. 2003.
- [4] George, Chinnu, and Dhinakaran Nathaniel. "Protecting Location Privacy in Wireless Sensor Networks against a Local Eavesdropper-A Survey." *International Journal of Computer Applications* 56.5 (2012).
- [5] Liang, Zong-Wen, et al. "The survey of location privacy protection." *2012 International Conference on Wavelet Active Media Technology and Information Processing (ICWAMTIP)*. IEEE, 2012.
- [6] Conti, Mauro, Jeroen Willemsen, and Bruno Crispo. "Providing source location privacy in wireless sensor networks: A survey." *IEEE Communications Surveys & Tutorials* 15.3 (2013): 1238-1280.
- [7] Kamat, Pandurang, et al. "Enhancing source-location privacy in sensor network routing." *25th IEEE international conference on distributed computing systems (ICDCS'05)*. IEEE, 2005.
- [8] Xi, Yong, Loren Schwiebert, and Weisong Shi. "Preserving source location privacy in monitoring-based wireless sensor networks." *Proceeding's 20th IEEE International Parallel & Distributed Processing Symposium*. IEEE, 2006.
- [9] Kang, Lei. "Protecting location privacy in large-scale wireless sensor networks." *2009 IEEE International Conference on Communications*. IEEE, 2009.
- [10] Tang, Di, et al. "Cost-aware secure routing (CASER) protocol design for wireless sensor networks." *IEEE Transactions on Parallel and Distributed Systems* 26.4 (2014): 960-973.
- [11] Kazatzopoulos, Leonidas, et al. "ihide: Hiding sources of information in wsns." *Second International Workshop on Security, Privacy and Trust in Pervasive and Ubiquitous Computing (SecPerU'06)*. IEEE, 2006.

- [12] Mahmoud, Mohamed Elsalih, and Xuemin Shen. "Secure and efficient source location privacy-preserving scheme for wireless sensor networks." *2012 IEEE International Conference on Communications (ICC)*. IEEE, 2012.
- [13] Wang, Haodong, Bo Sheng, and Qun Li. "Privacy-aware routing in sensor networks." *Computer Networks* 53.9 (2009): 1512-1529.
- [14] Li, Jian, et al. "Hop-by-hop message authentication and source privacy in wireless sensor networks." *IEEE transactions on parallel and distributed systems* 25.5 (2013): 1223-1232.
- [15] Raj, Mayank, et al. "Using data mules to preserve source location privacy in wireless sensor networks." *Pervasive and Mobile Computing* 11 (2014): 244-260.
- [16] Rimjhim, Pradeep Kumar Roy, and Jyoti Prakash Singh. "Finding location of fake and phantom source for source location privacy in wireless sensor network." *International Journal of Communication Networks and Distributed Systems* 26.1 (2021): 114-134.
- [17] L. C. Mutalemwa and S. Shin, "Energy Balancing and Source Node Privacy Protection in Event Monitoring Wireless Networks," *2021 International Conference on Information Networking (ICOIN)*, 2021, pp. 792-797, doi: 10.1109/ICOIN50884.2021.9333901.
- [18] Wang, Qihua, et al. "SPS and DPS: Two new grid-based source location privacy protection schemes in wireless sensor networks." *Sensors* 19.9 (2019): 2074.
- [19] Ngai, Edith C-H., and Ioana Rodhe. "On providing location privacy for mobile sinks in wireless sensor networks." *Wireless networks* 19.1 (2013): 115-130.
- [20] Liu, Anfeng, et al. "Preserving smart sink-location privacy with delay guaranteed routing scheme for WSNs." *ACM Transactions on Embedded Computing Systems (TECS)* 16.3 (2017): 1-25.
- [21] Long, Jun, et al. "An energy-efficient and sink-location privacy enhanced scheme for WSNs through ring-based routing." *Journal of parallel and Distributed computing* 81 (2015): 47-65.
- [22] Chai, Guofei, et al. "Enhancing sink-location privacy in wireless sensor networks through k-anonymity." *International Journal of Distributed Sensor Networks* 8.4 (2012): 648058.
- [23] Di Ying, Bi, Dimitrios Makrakis, and Hussein T. Mouftah. "Anti-traffic analysis attack for location privacy in WSNs." *EURASIP Journal on Wireless Communications and Networking* 2014.1 (2014): 1-15.
- [24] Chen, Honglong, and Wei Lou. "On protecting end-to-end location privacy against local eavesdropper in wireless sensor networks." *Pervasive and Mobile Computing* 16 (2015): 36-50.
- [25] Baroutis, Nikolaos, and Mohamed Younis. "Load-conscious maximization of base-station location privacy in wireless sensor networks." *Computer Networks* 124 (2017): 126-139.

- [26] Yao, Lin, et al. "Protecting the sink location privacy in wireless sensor networks." *Personal and Ubiquitous Computing* 17.5 (2013): 883-893.
- [27] Rios, Ruben, Jorge Cuellar, and Javier Lopez. "Probabilistic receiver-location privacy protection in wireless sensor networks." *Information Sciences* 321 (2015): 205-223.
- [28] Wang, Jian, et al. "Sink location privacy protection under direction attack in wireless sensor networks." *Wireless Networks* 23.2 (2017): 579-591.
- [29] Boulaiche, Mehdi, and Mohamed Younis. "Increasing base-station anonymity through illusive void formation." *International Journal of Communication Networks and Distributed Systems* 25.4 (2020): 433-460.
- [30] Lee, Sangho, Jong Kim, and Yoonho Kim. "Preserving source-and sink-location privacy in sensor networks." (2016).
- [31] Chen, Juan, Xiaojiang Du, and Binxing Fang. "An efficient anonymous communication protocol for wireless sensor networks." *Wireless Communications and Mobile Computing* 12.14 (2012): 1302-1312.
- [32] Abuzneid, Abdel-Shakour, et al. "Fortified anonymous communication protocol for location privacy in WSN: a modular approach." *Sensors* 15.3 (2015): 5820-5864.

Vitae

Name: Mohammed A. K. Al Saafin

Nationality: Palestinian

Date of Birth: 2nd of November 1992

Email: mohammed.alsaafeen1@gmail.com

Address: Saudi Arabia – Eastern Province – Khobar Rawabi 34421

Academic Background:

- Bachelor of Science in Industrial & Systems Engineering from King Fahd University of Petroleum & Minerals (KFUPM) - Dhahran, KSA in 2017.
- Master of Science in Industrial & Systems Engineering from King Fahd University of Petroleum & Minerals (KFUPM) - Dhahran, KSA in 2022.

Research Papers:

No	Type	Paper Title	Status
1	Conference	Location Privacy Preservation Mathematical Model for the Sink Node Anonymity in Wireless Sensor Networks	Under Preparation