

**COST-EFFECTIVE NETWORK PLANNING AND
OPERATION FOR RURAL COMMUNITIES**

A DOCTOR OF PHILOSOPHY THESIS SUBMITTED TO THE
DEPARTMENT OF ELECTRONIC ENGINEERING AND
COMPUTER SCIENCE
QUEEN MARY UNIVERSITY OF LONDON

BY

NAZRUL ISLAM

June 2021

I, Nazrul Islam, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with or supported by others, that this is duly acknowledged below, and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the University has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

The copyright of this thesis rests with the author, and no quotation from it or information derived from it may be published without the prior written consent of the author.

Signature:

Date: 25/06/2021

ABSTRACT

Broadband Internet access is central to the regeneration of remote communities and reducing the digital divide between rural and urban regions. This thesis focuses on rural communities with limited financial resources, environmental issues including long reach from conurbations, and mountainous or otherwise adverse terrain, typically with limited access to a wired power supply. As such, regular access technologies based on cable or fibre optics are not financially viable. To overcome this challenge, we consider the deployment of a Free-Space Optical (FSO) based relay network as the primary technology, using diversity to provide resilience to atmospheric effects. The aim of this research is to design and evaluate a rural network planning and traffic engineering framework employing FSO communication using light emitting diodes/lasers to construct backhaul rural network infrastructures.

FSO systems are relatively cheap and easy to implement [1]. Various proof-of-concept technologies already exist [2] [3] [4]. However, the focus of this work is on the design of a flexible network-planning tool together with a robust management framework that is designed to operate over such an infrastructure to ensure it functions efficiently despite changes in load or communication channel outages. Although the work concentrates on an FSO based infrastructure, this could be extended to support heterogeneous networks employing a combination of technologies.

More precisely, this research first describes a novel network planning tool with an intelligent resource management system based on a Multi-Objective Evolutionary Algorithm (MOEA) that determines the suitable location of FSO relay nodes, taking into account end-to-end link speed which is bitrate of user data and the degree of path diversity coupling with battery power. This MOEA approach can account for Line-of-Sight occlusions and allows various compromises to be selected from a Pareto front to suit individual needs. We provide suitable results to show the satisfactory operation of the tool and outline avenues for future development. Following on from this, we design and evaluate an intelligent traffic-engineering framework to make the best use of the deployed infrastructure that can adapt to environmental changes. This aims to ensure a good service is maintained at all times by suitable reconfiguration.

Glossary of Terms

ADSL	-	Asymmetric Digital Subscriber Line
AND	-	Average Node Degree
ANN	-	Artificial Neural Networks
BDUK	-	Broadband Delivery UK
BER	-	Bit Error Rate
BS	-	Base Station
CAPEX	-	Capital Expenditure
CD	-	Crowding Distance
DSL	-	Digital Subscriber Loop
EA	-	Evolutionary Algorithms
E2E	-	End-to-End
FCC	-	Federal Communications Commission
FEC	-	Forward Error Correction
FSO	-	Free-Space Optical
FTTC	-	Fibre to the Curb
GA	-	Genetic Algorithm
GDP	-	Gross Domestic Product
GIS	-	Geographical Information System
ICT	-	Information Communication Technology
ILP	-	Integer Linear Programs
IrDA	-	Infrared Data Association
ISP	-	Internet Service Provider
LEO	-	Low Earth Orbit
LOS	-	Line of Sight
LTE	-	Long-Term Evolution
MOEA	-	Multi-Objective Evolutionary Algorithm
MOEC	-	Multi-Objective Evolutionary Computation
MPLS	-	Multi-Protocol Label Switching
NN	-	Neural Network
NOS	-	Network Operating System
NSGA	-	Nondominated Sorting Genetic Algorithm
OPEX	-	Operational Expenditure
PF	-	Pareto Front

PS	-	Pareto Set
PON	-	Passive Optical Network
PPP	-	Public Private Partnership
PSO	-	Particle Swarm Optimization
QoS	-	Quality of Service
RF	-	Radio Frequency
ROI	-	Return on Investment
SDN	-	Software-Defined Networking
SNR	-	Signal to Noise Ratio
SPEA	-	Strength Pareto Evolutionary Algorithm
TE	-	Traffic Engineering
TVWS	-	TeleVision Whitespace
UHF	-	Ultra High Frequency
VHF	-	Very High Frequency
VOIP	-	Voice Over Internet Protocol

Table of Contents

CHAPTER 1. INTRODUCTION	15
1.1 RESEARCH AIM AND MOTIVATION	17
1.2 SUMMARY OF NOVELTY	18
1.3 STRUCTURE OF THE THESIS.....	19
CHAPTER 2. STATE OF THE ART.....	20
2.1 CHALLENGES AND REQUIREMENTS	22
2.1.1 <i>Technological Aspects</i>	22
2.1.2 <i>Economical Aspects</i>	23
2.1.3 <i>Demographic Aspects</i>	24
2.1.4 <i>Other Constraints and Requirements</i>	25
2.2 CURRENT TECHNOLOGIES IN RURAL AREAS	27
2.2.1 <i>ADSL</i>	27
2.2.2 <i>Mobile Internet</i>	28
2.2.3 <i>Wireless and Satellite Broadband</i>	29
2.3 RECENT ACTIVITY WITH RURAL BROADBAND	29
2.4 FREE-SPACE OPTICAL (FSO) COMMUNICATION.....	30
2.5 NETWORK PLANNING TOOLS.....	31
2.6 GENETIC ALGORITHMS (GA)	31
2.7 MULTI-OBJECTIVE EVOLUTIONARY ALGORITHMS (MOEA).....	34
2.8 STRENGTH PARETO EVOLUTIONARY ALGORITHM (SPEA)	36
2.9 PARETO FRONT (PF).....	39
2.10 DIJKSTRA’S SHORTEST PATH ALGORITHM	41
2.11 NSGA CROWDING DISTANCE.....	42
2.12 SOFTWARE-DEFINED NETWORKING (SDN).....	43
CHAPTER 3. NETWORK PLANNING TOOL.....	45
3.1 INITIAL DESIGN WITH A GENETIC ALGORITHM (GA).....	45
3.2 REQUIREMENTS	47
3.3 FITNESS FUNCTION.....	47
3.4 IMPLEMENTATION	49
3.5 FITNESS FUNCTION EVALUATION	49
3.6 LINE OF SIGHT (LOS).....	50
3.7 CHROMOSOME SELECTION.....	51
3.8 RANKING.....	51
3.9 FINAL OUTPUT	52

3.10	VALIDATION	53
3.11	PERFORMANCE VALIDATION.....	54
3.11.1	<i>Impact of Population Size</i>	55
3.11.2	<i>Impact of Generation Count</i>	56
3.11.3	<i>Impact of Mutation Rate</i>	58
3.11.4	<i>Multipoint (2-Point) Crossover</i>	60
3.11.5	<i>New Chromosome Insertion</i>	60
3.12	DISCUSSION	61
CHAPTER 4. NETWORK PLANNING TOOL - MULTIPLE OBJECTIVES		63
4.1	PROBLEM DESCRIPTION.....	63
4.2	FLOWCHART OF SPEA-II	64
4.2.1	<i>Initialisation</i>	64
4.2.2	<i>Evaluation</i>	65
4.2.3	<i>Fitness Assignment</i>	70
4.2.4	<i>Environmental Selection</i>	72
4.2.5	<i>Mating Selection</i>	72
4.2.6	<i>Reproduction</i>	72
4.2.7	<i>Final Solution</i>	73
4.3	SIMPLE FLOWCHART OF SPEA-II	74
4.4	HYBRID MULTI-OBJECTIVE EVOLUTIONARY ALGORITHM.....	74
4.4.1	<i>Line of Sight</i>	75
4.4.2	<i>Dijkstra's Shortest Path Algorithm</i>	76
4.4.3	<i>NSGA Crowding Distance</i>	78
4.4.4	<i>Final Hybrid Algorithm Flow-Chart</i>	79
4.5	FINAL OUTPUT	80
4.6	VERIFICATION OF THE NETWORK PLANNING TOOL.....	81
4.6.1	<i>Experimental Verification of the Tool</i>	81
4.6.2	<i>Performance Validation</i>	85
4.6.3	<i>Injecting Fresh Chromosomes</i>	87
4.6.4	<i>Increasing Mutation Rate</i>	87
4.6.5	<i>NSGA Crowding Distance</i>	89
4.7	DISCUSSION	91
CHAPTER 5. TRAFFIC ENGINEERING SYSTEM		94
5.1	TRAFFIC ENGINEERING DESIGN PARAMETERS	95
5.1.1	<i>End-to-End (E2E) Speed</i>	96

5.1.2	<i>Minimum Residual Battery Life (RBLm)</i>	98
5.1.3	<i>Transmit Power Setting Impactor (TPSI)</i>	98
5.2	TE UNDER GENETIC ALGORITHM	99
5.2.1	<i>Fitness Score</i>	99
5.3	END-TO-END (E2E) SPEED	100
5.3.1	<i>Expectation</i>	101
5.3.2	<i>Result</i>	101
5.3.3	<i>Analysis</i>	102
5.4	MINIMUM RESIDUAL BATTERY LIFE (RBLM)	103
5.4.1	<i>Expectation</i>	103
5.4.2	<i>Result</i>	103
5.4.3	<i>Analysis</i>	105
5.5	TRANSMIT POWER SETTING (TPS)	106
5.5.1	<i>Expectation</i>	107
5.5.2	<i>Result</i>	107
5.5.3	<i>Analysis</i>	107
5.6	COMBINED OBJECTIVES	107
5.6.1	<i>Expectation</i>	108
5.6.2	<i>Result</i>	108
5.6.3	<i>Analysis</i>	110
5.7	EFFECT OF WEATHER EVALUATION	110
5.7.1	<i>Expectation</i>	111
5.7.2	<i>Result</i>	111
5.7.3	<i>Analysis</i>	113
5.8	TRAFFIC ENGINEERING TOOL UNDER MOEA	114
5.8.1	<i>Improved Strength Pareto Evolutionary Algorithm</i>	114
5.8.2	<i>Final Flowchart</i>	116
5.8.3	<i>Design Parameters</i>	117
5.8.4	<i>Implementation</i>	118
5.8.5	<i>Output</i>	119
5.8.6	<i>Analysis</i>	123
5.8.7	<i>Final Generation Pareto Set</i>	126
5.8.8	<i>Final Output Evaluation</i>	127
5.8.9	<i>Simulation versus Theoretical Solutions</i>	128
5.8.10	<i>Performance Evaluation</i>	129

5.8.11	<i>Impact of Weather</i>	138
5.8.12	<i>Results</i>	140
5.8.13	<i>Real-life Scenario</i>	143
5.8.14	<i>Performance over Multiple Epochs</i>	148
5.9	DISCUSSION	150
CHAPTER 6.	CONCLUSIONS AND FUTURE WORK	152
CHAPTER 7.	REFERENCES	156

List of Figures

Figure 1: Rural Broadband Design Framework.....	18
Figure 2: Percentage of “Superfast” Broadband by Nation [45]	25
Figure 3: Possible Chromosome Structure (N_i is Number of the Node)	32
Figure 4: Simple Genetic Algorithm Flowchart	33
Figure 5: Flowchart of SPEA-II [98]	37
Figure 6: Crossover Demonstration	39
Figure 7: Process of Mutation Operator.....	39
Figure 8: MOEA with Pareto Front [27].....	40
Figure 9: Pareto Optimal Front Demonstration	40
Figure 10: NSGA-II Crowding Distance Process [102]	42
Figure 11: Crowding Distance Calculation [102].....	43
Figure 12: SDN Overall Architecture	44
Figure 13: Network Planning Scenario	45
Figure 14: Simple Genetic Algorithm Process	46
Figure 15: Initial Output of the GA Networking Planning Tool	47
Figure 16: LOS Evaluation	50
Figure 17: Example LOS Topological Representation	52
Figure 18: Topological Representation with an Obstacle (Blue Region)	53
Figure 19: The Optimal Solution	53
Figure 20: Output experiencing Premature Convergence.....	55
Figure 21: Assessing Increased Population Size	56
Figure 22: Output with Increased Generation Count.....	58
Figure 23: Increased Mutation Scenario	60
Figure 24: SPEA-II Flow-Chart.....	64
Figure 25: Can-See Representation	66
Figure 26: Complex Can-See Representation.....	67
Figure 27: Can-See Tree Formation	68
Figure 28: Process of Diversity.....	69
Figure 29: Graphical Representation of Range Profile.....	70
Figure 30: Dominance Rank [121]	71

Figure 31: Dominance Count [121]	71
Figure 32: Overall SPEA-II Flowchart	74
Figure 33: Dijkstra’s Algorithm Representation.....	77
Figure 34: Flowchart of Modified Dijkstra’s Algorithm	78
Figure 35: Crowding Distance (CD) Process Flow	79
Figure 36: Final Algorithm Flowchart	80
Figure 37: Typical Final Output	81
Figure 38: Initial Fitness Test	82
Figure 39: Node Placement Experiment	83
Figure 40: Range Profile for Experiment.....	83
Figure 41: Example 3 Relay-Node Placements for Differing Constraint Preferences	84
Figure 42: Example 4 Relay-Node Placements for Differing Constraint Preferences	84
Figure 43: Node Placement with Obstacles	85
Figure 44: Typical Final Output	86
Figure 45: Output for Convergence Test	89
Figure 46: Pareto Front in Generation 10	90
Figure 47: Pareto Set without CD	91
Figure 48: Typical Rural Topology with TE-Enabled Relay Nodes	94
Figure 49: Typical Topographical Topology	95
Figure 50: Range Profile for Different Transmit Power Settings	96
Figure 51: Scenario using New Range Profile.....	97
Figure 52: Flowchart of WEP	100
Figure 53: Topology for Evaluating Transmit Power Setting	107
Figure 54: Impact of Weather on Rural Broadband.....	111
Figure 55: Flow Diagram of the Traffic Engineering MOEA	115
Figure 56: Final Process Flowchart of the Improved SPEA.....	116
Figure 57: Solutions Found in Generation 1	120
Figure 58: Solutions Found in Generation 2.....	122
Figure 59: Solutions Found in Generation 3.....	123
Figure 60: Feasible Solutions and the Pareto Set.....	125
Figure 61: Solutions in the Final Generation	126
Figure 62: Graphical Representation of all Possible Solutions	128

Figure 63: A Complex Rural Topology to Provide Broadband	130
Figure 64: Performance Evaluation Range Profile	130
Figure 65: Final Output after 100 Generations	131
Figure 66: Pareto Set of the Final Generation	132
Figure 67: Execution Time versus the Number of Nodes.....	137
Figure 68: Execution Time versus Average Node Degree	138
Figure 69: Rural Broadband Impacted by Weather	139
Figure 70: Clear Weather Final Generation Secondary Population.....	141
Figure 71: Light Weather Final Generation Secondary Population.....	142
Figure 72: Bad Weather Final Generation Secondary Population.....	143
Figure 73: Weather Conditions across Different Epochs	149
Figure 74: Comparative Performance with TE Mechanism Enabled/Disabled under Dynamic Weather Conditions	150
Figure 75: SDN Controlled Network Infrastructure	154

List of Tables

Table 1: Cost of Different Types of Broadband Technology [65].....	26
Table 2: Chromosome Evaluation	50
Table 3: Best Chromosomes Preserved	51
Table 4: Ranked Chromosomes	52
Table 5: Increasing the Number of Generations	57
Table 6: Final Output with Single Gene Mutation.....	59
Table 7: Output with Chromosome Injection	61
Table 8:Chromosome Representation	65
Table 9: Primary Population with Chromosome Injection	87
Table 10: Increased Mutation Rate Scenario	88
Table 11:Power Setting Chart.....	98
Table 12: All Viable Solutions with their Performance Measures	117
Table 13: All Viable Solutions with their Performance Measures with WEP.....	119
Table 14: Different Solution Colour Codes	119
Table 15: Tabular Form of Unique Solutions in Generation 1	120
Table 16: Tabular Form of Unique Solutions in Generation 2	121
Table 17: Tabular Form of Unique Solutions in Generation 3	122
Table 18: All Theoretically Viable Paths/Settings Discovered over All Generations	124
Table 19: All Theoretical Solutions with their Objectives Values	125
Table 20: Secondary Population in the Final Generation	126
Table 21: All Possible Solutions with their Objective Values.....	127
Table 22: Unique Chromosomes from all Possible Solutions	128
Table 23: All Feasible Solutions from the Simulation.....	129
Table 24: Tabular Output of the Temporary Population	132
Table 25: Final Output with 100 Generations.....	133
Table 26: Final Output with 200 Generations.....	134
Table 27: Final Output with 500 Generations.....	134
Table 28: Final Output with 1000 Generations.....	135
Table 29: All Possible Paths in the given Topology	135
Table 30: All Paths with High TPSI and the corresponding E2E Speed	136

Table 31: All Paths with Low TPSI and the Corresponding E2E Speed	136
Table 32: Range Profile considering the Impact of Weather	140
Table 33: Final Secondary Population under Clear Weather Conditions	141
Table 34: Final Secondary Population under Light Weather Conditions	142
Table 35: Final Secondary Population under Bad Weather Conditions	143
Table 36: Weather Conditions for Successive Epochs	144
Table 37: Final Output of the 1st Simulation.....	145
Table 38: Final Output of the 2nd Simulation	145
Table 39: Final Output of the 3rd Simulation	146
Table 40: Final Output of the 4th Simulation	147
Table 41: Final Output of the 5th Simulation	147
Table 42: Final Output of the 6th Simulation	148
Table 43: E2E Speed Across Multiple Epochs	150

Chapter 1. Introduction

The importance of the Internet in our lives today cannot be overstated. In recent years there has been much research and development devoted to high-capacity backbone networks, as described in [5], [6]. The Internet has been declared a fundamental right of citizens in several countries such as Finland, Spain and Greece. The demand for this high-capacity broadband is increasing significantly as modern applications appear [7]. Internet-based services are no longer indulgences but essential life activities for urban and rural residents. These services require a communication network. However, communication networks have not been well developed in rural areas.

Communication and new means of accessing the Internet are increasingly being developed around the world [8]. Despite a large number of Internet users now enjoying high bitrate speed access, there are still geographic regions where broadband services are either excessively expensive or simply inaccessible at any price. Broadband can be made available anywhere with a phone line. However, fibre optic connections are usually required for broadband that can reach fast speeds, which need a slightly more advanced infrastructure. This sort of infrastructure can be set up in urban and residential areas reasonably easily. For example, roads can be dug up to install cables, and a single fibre-connected street cabinet can connect to many homes. However, this is not the case in rural areas. Broadband Internet access deployment is a crucial factor for revitalising rural areas and reducing the existing digital divide between rural and urban areas [9]. Lack of broadband access in rural areas can contribute to unfairness, where access to education and information is more impoverished than in cities and where it becomes less favourable to establish industries and businesses [10]. Therefore, a robust and effective rural network planning framework is required to construct backhaul rural network infrastructures, possibly using resilient and wireless communication. Thus, our goal was to design a flexible network-planning tool together with a robust management framework designed to function over the infrastructure to ensure it operates efficiently despite various conditions.

Different factors affect rural broadband access, including the services, infrastructure, demography, and topology. Many rural areas worldwide, especially in developing countries, do not have reasonable connectivity solutions that are economically feasible. Rural areas/residential areas are organised around services provided by rural Internet infrastructure. These infrastructures require a significant upgrade or a better solution for a rural network to offer the current need. However, broadband Internet access deployment is a key factor for the revitalisation of rural areas and the reduction of the existing digital divide between rural and urban areas [11]. In urban areas with high population densities, a broadband services market is hugely competitive nowadays, and user requirements are therefore well catered for. In low-populated rural areas, the Return on Investment (RoI) is the main obstacle for broadband service in the foreseeable future, and consequently, rural users cannot access high-bitrate broadband services at all.

Our research considers a robust and effective rural network-planning framework using cheap Free-Space Optical (FSO) communication links to construct backhaul rural network infrastructures. While Radio Frequency (RF) communication is assumed between relays and

users, the RF or fibre link between relays and Base Station (BS) can be replaced by FSO links [12] [13]. FSO has the following advantages:

- Ease of deployment
- License-free long-range operation (in contrast with radio communication)
- High bitrates attainable
- Immunity to electromagnetic interference
- Increased security when working with a narrow beam(s)

The major challenge is the impact of atmospheric effects that can significantly impact on the Signal to Noise Ratio (SNR), particularly in fog.

Due to the terrain and sparse populations involved, traditional cable-based delivery systems are unable to justify the capital expenditure involved. Similar financial issues apply to satellite systems. Therefore, this thesis looks to low-cost incremental systems where we assume the network comprises several FSO relay nodes strategically placed across the environment. Although the range of these links may only be several kilometres due to dispersion and environmental effects (i.e., weather), they have the advantage of not requiring spectrum licence to operate and can provide meaningful transmission rates. Furthermore, by creating topologies with path diversity, some measure of resilience is possible.

In this research, we first consider the placement of the relay nodes subject to various, conflicting constraints. Once this is achieved, we refocus our attention on how best to maintain service delivery by selecting the most appropriate forwarding path considering factors such as relay node transmission power, residual battery life and End-to-End (E2E) transmission bitrate. As these factors are influenced by the environment, we need to provide a mechanism for dynamic path selection and configuration of the transceivers.

Thus, our work comprises the design of a flexible network-planning tool together with a robust management framework that is designed to function over an FSO infrastructure to ensure it operates efficiently despite various consequences. For example, poorer SNR can be offset by increasing the transmission power or suitable path selection.

This thesis includes a critical analysis of the research that has been done, and a solution is presented for rural communities to facilitate broadband access. FSO has been selected as the initial technology for the points listed above. However, we confine ourselves to scenarios where the distance from the rural community to the Internet Service Provider's point-of-presence is typically ≤ 20 Km.

The initial network planning research was undertaken using a Genetic Algorithm (GA), but due to the nature of the problem with competing objectives of diversity and end-to-end speed, a hybrid Multi-Objective Evolutionary Algorithm (MOEA) is developed. This algorithm determines the topological node placement to reach the rural access backhaul taking into account line of sight positioning, range, link speed and path diversity.

1.1 Research Aim and Motivation

The modern economy is highly dependent on the Internet. A lack of high-bitrate broadband access in rural areas can contribute to unfairness, where access to Information Communication Technology (ICT) is more impoverished than in cities. Thus, it becomes less favourable to establish industries and businesses [10]. Setting up a broadband infrastructure is more expensive for rural subscribers. Providing wireless broadband access to rural and remote areas is challenging for Internet Service Providers (ISPs) due to the necessity of financial viability. By laying cables, broadband can reach the highest bitrates through fibre optic connections. However, the Capital Expenditure (CAPEX) for buried cabling is often cost-prohibitive in rural areas. Houses tend to be further apart in the countryside, so a single street cabinet cannot connect as many homes, and it is expensive to maintain or improve that infrastructure. Therefore, many rural exchanges and cabinets have not been upgraded to provide faster broadband speeds. Moreover, geographical features of rural areas, such as hills and rivers, impact the cost of setting up a broadband infrastructure. Therefore, our research first considers relay node placement for a rural area bearing in mind the end-to-end communication channel bitrate. For most users, availability and bitrate are the most critical factors and thus is the primary consideration in this research. Our aim is to provide reliable rural broadband access with reasonable bitrate and resilience through path diversity.

Our research is concerned with improving broadband Internet access to the rural community. More precisely, our research goals are as follows:

- To design a network-planning tool that can determine where to place communication relay nodes such that a range of near-optimal combinations of path diversity and end-to-end speed is determined. This tool should cope with line-of-sight (LOS) obstacles
- To further enhance the network planning tool to take into account other trade-offs such as the overall bitrate of data and energy consumption, given that the transmission power is adjustable according to current conditions.
- The tool will be modified to allow it to explore questions such as "how many nodes are enough?" to achieve a certain level of availability and "what would be the additional benefit of increasing the number of nodes further?" Our tool will be flexible with a number of nodes to be placed in the rural region to achieve a degree of flexibility and availability of node placement. Therefore, the broadband provider has the flexibility to observe how many nodes they want to place in order to provide rural broadband by adding a number of nodes. This observation offers broadband ISP providers greater flexibility to control CAPEX for rural broadband. Again, the provider will be able to see the benefits of increasing or decreasing nodes. This flexibility and control through our tool provide financial and operational benefits to ISP providers, thus controlling CAPEX and OPEX.
- Having created the network topology, to design a Traffic Engineering (TE) framework that can make fair use of this infrastructure under differing conditions. This will include the use of multi-path redundancy adjustable transmission power to maintain connectivity when the Bit Error Rate (BER) would otherwise impede communication.

This research focuses on rural communities with environmental issues, including long reach from conurbations and hilly or otherwise hostile terrain, typically limited access to a wired power supply. Poor or variable weather conditions are also assumed. The aim is to create a network-planning algorithm that can assess alternative deployments and determine the expected operational performance as far as is practical. Once deployed, the management of such an infrastructure is then considered. Thus, the system will encompass monitoring and management functionality to ensure the network performs effectively in varying conditions, including the possibility of outages, by appropriate Traffic Engineering (TE) of alternative paths. A typical scenario is shown in Figure 1.

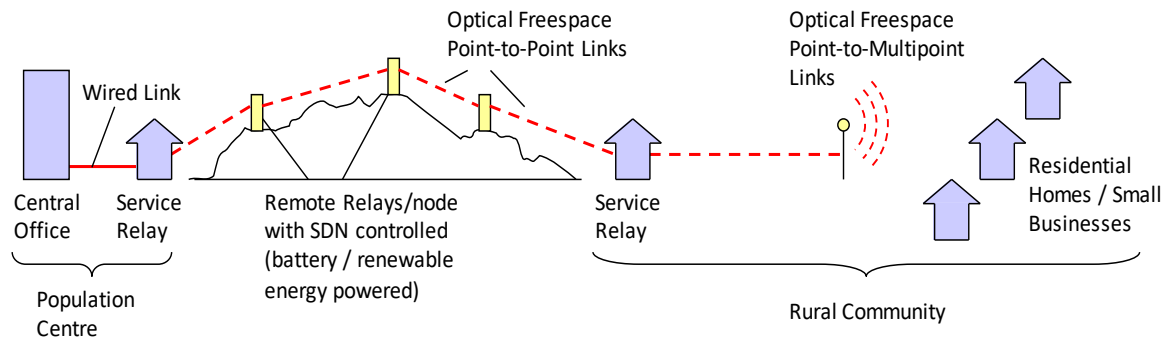


Figure 1: Rural Broadband Design Framework

1.2 Summary of Novelty

The novelty of this research relates to both the network-planning tool and traffic engineering application for the configuration and operation of a rural community Internet infrastructure. Although a number of commercial network planning and simulation tools exist, such as OPNET [14], NETSIM [15], these are proprietary and not tailored to the deployment of FSO links or take into account rural factors. Furthermore, no published work has considered the trade-off between cost, reliability, and the supported bitrate of user data, which we also refer to as the end-to-end speed, when placing relay nodes. Our research concentrates on providing a communication network with limited resources that are low maintenance. Furthermore, once constructed, we explore how to manage this infrastructure, given potential outages and variable transmission link quality, in order to maintain a reasonable service to end-users.

The novel contributions can be summarised as follows:

- A new network-planning tool is designed to determine the appropriate deployment of relay nodes to provide remote access whilst meeting a set of constraints. Initially, a Genetic Algorithm (GA) has been used and modified to employ a Multi-Objective Evolutionary Algorithm (MOEA). To our knowledge, no research has previously used an MOEA in this context.
- The network-planning tool is used to evaluate the feasibility of low-cost network deployment in a rural context, initially using Free Space Optical (FSO). Our network-planning tool considers the environment and trade-offs to discover suitable node locations.

- Our attention then turns from network planning to Traffic Engineering (TE), designing and implementing a control plane manager that (re)direct the traffic flows across the rural access infrastructure in order to meet varying needs. This includes altering the transmission bitrate of links and/or path steering to allow specific relay nodes to "sleep" when conditions permit.
- An MOEA-based TE algorithm has been developed where three objectives are considered to identify suitable trade-offs in order to find the best possible solution for rural broadband. To our best knowledge, no one designed a TE tool for use in a rural context considering multiple objectives with an MOEA.
- The TE has been extended to take into account changing weather conditions on the satisfactory operation of the network using periodic reconfiguration.

1.3 Structure of the Thesis

This research focuses on rural broadband from the perspective of network planning/deployment followed by traffic engineering, and therefore will be addressed in two parts. The thesis is organised as follows after this first introductory chapter. Chapter 2 describes the background knowledge needed to design network planning and traffic engineering tools. This includes some contextual material as well as a review of current technologies in rural areas, focusing on FSO. Initially, a primary network planning tool was designed with a Genetic Algorithm (GA), described in Chapter 3 followed by its implementation in 3.4. Following on from this, we introduce a Multi-Objective Evolutionary Algorithm (MOEA) to enhance this networking tool in 4. In 4.1, the design of this tool is described, along with the implementation of this tool. Next, in 4.6 we present the program verification along with performance results. Chapter 5 then refocuses attention on the traffic engineering of a network that would typically have been created in accordance with the output of the network planning tool. It describes the design parameters of the intelligence and initially concentrates on a GA-based Traffic Engineering (TE), providing a brief performance evaluation. This includes the impact of weather effects. In 5.8, the intelligent TE system is modified to employ an MOEA. A performance evaluation explores how various design parameters affect the overall efficacy of the tool. Chapter 6 presents some overall conclusions and outlines possible ideas for future work.

Chapter 2. State of the Art

Current broadband technologies for rural areas, including low-speed broadband, fixed wireless access, and satellite networks, do not provide enough capacity for the future[1]. Several papers [2] [4] [16] have been published that propose new approaches for rural Internet access, and some individual projects [17] [18] [19] [20] aim to improve Internet infrastructure and provide high-speed Internet for specific rural areas. However, to see widespread deployment requires re-engineering the rural Internet infrastructure model to offer high Quality of Service (QoS) and reliability at a cost-effective price. QoS is the measurement of the overall performance of a service, typically in relation to metrics such as delay and loss probability.

Commerce in the modern world is highly dependent on the Internet. A number of studies have examined mechanisms for providing broadband access to rural communities in developed countries [1] [10] [21]. Similar broadband initiatives for rural areas in developing countries are a relevant part of economic development plans and overall modernisation strategies [8]. Even so, providing broadband access to rural and remote areas is challenging for Internet Service Providers (ISPs) due to costs relative to the anticipated revenue.

It is essential to take into account that many real-world problems are multi-objective in nature, and so is the network planning problem, such as the shortest path routing problem. Extensive research has been done on shortest path routing problems. It includes Dynamic Programming for directed networks, Dijkstra's labelling algorithm and the Bellman-Ford successive approximation algorithm [22]. These algorithms have major deficiencies, such as they can only find the shortest/least-cost route, and they demonstrate high computational complexity for real-time communications. Artificial Neural Networks (ANN) have been examined to solve the shortest path problem relying on their parallel architecture to provide a fast solution [23]. However, the ANN approach has several limitations, such as the complexity of the hardware, which increases considerably with an increasing number of network nodes. Evolutionary algorithms such as Genetic Algorithms (GA) [24] and Particle Swarm Optimization (PSO) [25] have been utilized. However, these approaches are designed to perform single-objective optimization, mostly a weighted cost metric. So, there is a desire to design a mechanism that gives multi-objective trade-off solutions involving two or more objectives.

Our research focuses on finding a range of solutions given different constraints and optimisation characteristics through the application of multi-objective optimization. Much of the research that investigates optimisation when there are conflicting objectives exploits Multi-objective Optimization Evolutionary Computation (MOEC) methods [26] [27] [28]. However, other strategies for the optimization of network layouts are based on Integer Linear Programming (ILP) described in [29] or Evolutionary Algorithms (EAs) [30]. Most of the known ILP problem approaches [31] [32], as well as the majority of EA approaches [33] [34], are restricted to predefined topologies like tree networks or ring layouts. Additionally, these ILPs are mainly limited to relatively small problems and restricted to the optimization of a single linear function.

Design methods for broadband access networks are extensively studied in [35] [36] [37] [38] [39] [40]. There are special algorithms and techniques that allow optimization and automation of the design process for access networks based on optical or copper cabling at every stage of planning during network transformation. Optical access networks, in addition to the cable and associated linear devices, contain active network elements such as routers, access switches, DSLAMs, which have different performances and, accordingly, a high price [41] [42]. Highly developed countries often lay multi-fibre optic cables for the “last mile” prior to their eventual use and can forget the “last mile” problem for a long period [43]. Other countries can use this method as one way of accessing network upgrade with limited financing [44].

The digital divide between rural or remote and urban areas impedes socio-economic development: with poor Internet access, or no access at all, citizens in rural and remote areas are disadvantaged both in terms of communication and economic opportunities. The social, economic and environmental conditions of rural and remote areas, however, mean that there are very few incentives for the private sector to invest in broadband provision there. This is one of the main challenges that rural policymakers and local authorities are facing today.

The UK Government defines “superfast broadband” as bitrates of user data greater than 24 Mbps, whereas Ofcom defines it as more than 30Mbps. Ofcom’s Connected Nations 2020 report [45] states that superfast broadband is available to 96% of UK homes. It is now projected that around 60% of premises that are able to provide superfast broadband now take up this service. However, a small but significant number of properties are still struggling to get connected. It is estimated that 43,000 premises cannot access either a decent fixed broadband service or good 4G coverage indoors.

Despite the increase in coverage of superfast services, many homes and small businesses still are unable to receive broadband speeds that are adequate to perform a range of everyday online activities reliably. One Telegraph report [46] suggests that in Welsh villages such as Abererch, Gwynedd, where broadband is so slow, locals get up at 4 am to send emails. Therefore, the problem remains.

A number of factors influence the attractiveness of rural areas for the introduction of new technologies and digital infrastructure. The first is low population density: rural areas are generally sparsely populated, and this hampers the achievement of economies of scale for the production of goods and, in the case of digital innovation, services. Rural areas are also experiencing two important demographic phenomena: firstly, many regions face an ageing population which poses serious challenges in terms of the economic sustainability of various services. Secondly, there is an exodus of young people out of rural regions representing a brain drain that influences the digital skill level of the areas. Furthermore, rural areas have varying types of terrain. Depending on the geography of the region, anything from rock to sand can be found, making planning and deploying a fibre build-out difficult. Additionally, many rural areas show an average per-capita income generally lower than that of cities, and this has a significant influence on the propensity of rural consumers to purchase digital services.

One of the main reasons for poor broadband speeds in rural areas is the length of the connection to the property. Rural properties are often further from the exchange or street cabinet than in urban areas. For copper-based telephony networks, the longer the connection,

the slower the bitrates that are likely due to the broadband signal's attenuation. People who live in less densely populated parts of the UK are more likely to live further from the exchange and obtain lower broadband speeds.

The resistance of the copper wire increases with the length of the wire, so bitrates decay as the distance between the premises and the exchange increases. Bit rates typically start to decrease between 1 and 2km from the exchange and are reduced considerably at distances more than 3.5 km. Fibre to the Curb (FTTC) based broadband that supports superfast bitrates up to 80Mbit/s uses optical fibre to the cabinet and reduce the length of copper wire. People living more than 300m from a cabinet can expect their bitrates to be less than half the maximum possible. However, most people who live too far from the cabinet to receive superfast broadband may still benefit from the upgrade at the cabinet, as the reduction in the length of the copper access line will still improve their broadband speeds.

2.1 Challenges and Requirements

2.1.1 Technological Aspects

When it comes to closing the connectivity gap and ensuring ultrafast services can reach most rural communities, there are several issues to be overcome, both in terms of the physical challenges of deployment and broader socio-economic factors. For example, low population densities mean that many economies of scale that enable large investments in new technology are reduced. Meanwhile, people in rural areas typically have lower incomes per capita than urban areas, affecting their willingness and ability to pay for more advanced services. On top of this, rural areas often face more challenging terrain, from the uneven and rocky ground to, at the other extreme, sandy ground, all of which can make deploying fibre cabling more difficult and expensive. The European Union (EU) handbook [2] noted that one way of reducing the impact of terrain issues is to reuse existing infrastructure deployments. For example, it highlighted in Spain [3] efforts that use water mains for optical fibre, deploying a small, flexible, high-density polyethylene duct within this infrastructure, allow any resident with access to the municipal water network can also gain access to this fibre network.

Some rural areas do not have direct Internet access. They rely on dial-up technology to transmit data over copper lines, unable to access online services even at elementary download and upload speeds. The development of alternatives to fibre optic cable has been slow and uneven. Satellite broadband can be a solution in very sparsely populated areas, but it often suffers from high latency¹, lack of significant bandwidth, and high data costs.

Scientists and engineers have proposed new technologies to reduce infrastructure costs and address topographical challenges such as mountains, long distances in rural communities. Relevant technologies include a rural extension for WiFi [5], a point-to-multipoint wireless distribution system that integrates renewable energy technologies for off-grid siting [6], and a method to leverage underutilised TV white space frequencies [7]. This work aims to assess the

¹ Latency is the time it takes for data to be transferred between its original source and its destination, measured in milliseconds.

factors that impact broadband infrastructure investment in rural communities; however, we defer to the corresponding references for a more in-depth analysis of the functionality.

The development of alternatives to fibre optic cable has been slow and uneven. While mobile telecommunications technologies such as 3G and 4G have provided customers with broadband bitrates through mobile devices such as smartphones, the cellular model of this technology is designed, economically speaking, for densely populated areas and faces the same connectivity and capacity gaps as traditional broadband [47]. Free-Space Optical (FSO) point-to-point links provide an interesting alternative technology, capable of bitrates from 100Mb/s to 10Gb/s that can replace the RF or fibre links in the backhaul network [12] [48].

In terms of deployment algorithms, a number of researchers have considered evolutionary approaches for network planning, but these mostly concentrate on improving the Quality of Service (QoS) and CAPEX. In [48], the authors use a Multi-Objective Evolutionary Algorithm (MOEA) to improve QoS and CAPEX using an existing topology rather than considering relay node placement. In addition, some machine learning techniques such as Neural Networks (NN) [50] have been proposed. However, none of these consider the constraints or characteristics of FSO.

2.1.2 Economical Aspects

Broadband access in urban areas is already provided by Passive Optical Networks (PON) and Digital Subscriber Loop (DSL) technologies [1]. However, the much higher cost of network deployment in rural areas [51] means such approaches are usually uneconomical. In urban areas with high population densities the broadband services market is extremely competitive; user requirements are therefore well catered for. In low-populated rural areas, the Return on Investment (ROI) is the main obstacle for broadband service deployment. As it becomes less favourable to establish industries and businesses [10], this contributes to the lack of fairness. In rural areas, houses tend to be further apart, so a single street cabinet does not connect as many homes, and it is not economically viable to improve that infrastructure laying costly lines for a few customers. Moreover, geographical features, such as hills and rivers, impact further on the cost of deployment.

High-speed rural Internet access is associated with increased incomes and reduced unemployment via increased opportunities for remote work and the ability to expand brick and mortar enterprises online [52]. Econometrics analysis suggests that a 10% increase in fixed broadband access increases Gross Domestic Product (GDP) in developed countries by approximately 1.2% [53]. As a result, governments around the world have invested funds to deploy broadband infrastructure in underserved areas. The US Federal Communications Commission (FCC) has established download/upload bitrate benchmarks of 25 Mbps / 3 Mbps for fixed services and 10 Mbps / 3 Mbps (median) for mobile services [54]. A recent study by the FCC shows that rural and tribal communities still halt behind in broadband deployment. In rural areas, only 68.6% of Americans have access to both fixed and mobile LTE broadband services (at any bitrate), compared to 97.9% in urban areas [55]. However, FCC is determined that "mobile services are not currently full substitutes for fixed services".

Embracing broadband can benefit the economy in rural areas, including increasing income, lowering unemployment rates and creating jobs. Broadband can reach fast bitrates through fibre optic connections, by laying cables. In rural areas, the Capital Expenditure (CAPEX) for buried cabling is often cost-prohibitive. Houses tend to be further apart in the countryside, so a single street cabinet cannot connect as many homes, and it is expensive to improve that infrastructure. It is not always economically viable for broadband companies to lay such costly lines for a few customers. Therefore, many rural exchanges and cabinets have not been upgraded to provide faster broadband speeds. Moreover, geographical features of rural areas, such as hills and rivers, impact the cost of setting up broadband infrastructure.

Capital expenditures associated with hardware costs are often identified as a primary barrier for deploying broadband in rural communities, which is why most developed countries have focused on providing funds to address this barrier. [56] surveyed fibre, fixed wire, wireless, and satellite costs to conclude that commercial provisioning is unfeasible in markets where profitability is low. Rural areas tend to have a lower population density, which provides fewer customers to absorb the fixed costs of broadband infrastructure. [57] found that weather patterns (e.g., increased storms) and topographical conditions (e.g., mountains) also increased deployment costs in Malaysia's rural areas. [58] examined the use of analogue television waves in specified ranges to provide wireless broadband service to Canada's rural communities. One of the primary challenges was the high cost of equipment, \$600-700 per household [55]. The specific frequency band was only available in Canada, and economies of scale could not reduce hardware costs. However, there are opportunities to develop non-traditional business models, such as co-operatives or public-private partnerships, that are better suited to ensuring affordability by reducing the need for high-profit margins. Besides, there may be opportunities to leverage existing infrastructure (e.g., electrical poles) or coordinate deployment activities (e.g., bury cables when constructing a new road) to reduce the required investment.

Developed countries like America and South Korea had put direct investment as well as guaranteed loans into the rural communities to reduce the digital divide via reliable broadband infrastructure and services [59]. Discounts or direct funding to end users as incentive-based subsidies may be more effective via the provision of broadband vouchers to economically marginalized people in rural areas [60]. Service providers will be stimulated to expand their infrastructure and services to rural areas only if they are guaranteed some grants like tax exemptions on the network equipment that is introduced for rural broadband development. Hence, exemption from local taxes and or taxes while importing equipment as well as direct subsidies through public private partnerships (PPP) can be considered as an important avenue for social development and reduce the digital divide by means of reliable and sustainable broadband services [61]. As a priority, subsidies need to be guaranteed first for building a new network rather than upgrading the existing network [62].

2.1.3 Demographic Aspects

Although a large number of Internet users now experience high-speed access, there are still large geographic regions where broadband services are either prohibitively expensive or simply unavailable at any price [21]. According to the Q4-2018 Ofcom report [45], the average UK download bitrate was just 34 Mbps in November 2018 compared to 53% in urban areas,

although that is a bit of an improvement from the previous report. In 2018, 7 per cent of premises in rural areas were not able to access a basic broadband. This relates with just 1% of premises in urban areas. Nevertheless, some rural villages and remote areas are receiving Internet access at a bitrate of a third of the current UK broadband average. Others have no access at all. Through government investment of over £780 million, superfast broadband is now available to over 80% of homes and businesses. The proportion of superfast broadband coverage by the nation at the year 2020 is shown in Figure 2. Lower coverage levels are largely due to a higher proportion of rural terrain with a more challenging topography.

	Superfast	Urban	Rural
England	96%	98%	84%
Northern Ireland	89%	99%	66%
Scotland	94%	98%	72%
Wales	94%	98%	78%
UK	96%	98%	81%

Figure 2: Percentage of “Superfast” Broadband by Nation [45]

Superfast broadband (defined as delivering a download bitrate of at least 30Mbps) is available for 83% of premises in rural areas, compared with 98% in urban areas [63].

In urban and suburban areas, illustrated by high population densities, mobile operators have been willing to invest in new technologies and to deploy advanced wireless broadband solutions. Similar cost-benefit and the techno-economic analysis of key factors affecting rural broadband access are essential [64]. However, although mobile access companies have invested considerably in setting up fibre-based backhaul, they remain reluctant to deploy anything similar in rural areas due to high deployment costs relative to the anticipated return on investment.

2.1.4 Other Constraints and Requirements

The copper wires that are used to carry standard Asymmetric Digital Subscriber Line (ADSL) Internet broadband slow connection speeds the further they have to travel from the telephone exchange to the residential dwellings. This is not so much of a problem in urban areas where exchanges serve many properties as users tend not to be far from the exchange. However, in rural regions many homes are several miles from their local exchange. Therefore, the “broadband” speed is slow.

Building an infrastructure that can provide at least 20Mbps with current technology suggests the use of fibre optics. Often, the capital expenditure (CAPEX) for building such infrastructure is not economically feasible. Moreover, operational expenditure (OPEX) directly influences the long-term sustainability of operating a network, particularly in terms of equipment maintenance costs. The CAPEX and OPEX of different broadband technologies are given in Table 1.

Table 1: Cost of Different Types of Broadband Technology [65]

Technology	Main Costs	Licencing cost	Notes
FSO	155MbPS to 1.2Gbps from £13k to £19k up to 3.6km. For up to 5.3km from £22k to £33k in the same bitrate	Is not needed a licence	Fast installation but radio back up needed to improve QoS
Wi-Fi	From £10k to £20k depending on the distance (1-10km) 18, 26 or 38 GHz can transmit up to 300 Mbps	Licence needed Cost from £5k while setting up enterprise network	Fast installation but limited bandwidth.
ZigBee	Less than £20 per node but depends on the network model	Typically, doesn't need any license	Bit rate is only 250 kBit/s at 2.4 GHz, and transmission distance range from 10 to 100 meters [66]
Bluetooth	Very low cost on setting up cost	Not needed but for new product need \$8000 membership charge [67]	Bluetooth 4.0 promises a 25Mbps data rate
Fibre	In the case of a P2MP system that is the most cost-effective in CAPEX concerning a pure FTTH. A standard GPON OLT (with 4G Ethernet ports), a 1:16 splitter, an ONU serving 48 VDSL2 end users, 15 ONTs with 2 G Ethernet ports and VOIP [Everything compliant with the ITU G.984 standard] has a total cost of around 26k€. (Source: vendor). Around 70€/m for the dig and 10€/m for the cable in an urban environment. In the river basin, the installation plus the cable (120 fibres) can go down at around 16€/m. (Source: PA).	Dark fibres can be rented on average from 5 to 15 €/m with a 15-year based contract. Free ducts can also be rented at around 12,3€/m	Slow installation, dig is needed with or without a trench.
xDSL	The costs can vary from 14k€ for an IP DSLAM serving 64 users (modems included) up to 40k€ for an IP DSLAM serving up to 1024 end users (modem included). In this last case, the average bit/rate per client is at around 10-15 Mbps over a 1.2 km – access network segment.	Copper line/client in unbundling regime is on average at 7,5€ (source: operator). A twisted pair can vary from €80 to €225, depending on the cable quality. (source: operator). Installation costs are on average at 80-85€/m	Maximum available bandwidth is limited by interferences. Major costs due to copper maintenance Other voices (CO Apparatus rental, CO apparatus maintenance, et al.)

The United Nations recently declared Internet Access a Human Right [68]. Lack of broadband denies rural areas the human right to communicate. Without broadband, rural communities are further isolated from economic and civic participation. However, there is no one-size-fits-all broadband solution for connecting rural communities. We consider an FSO-based solution. Current ADSL broadband can provide an average 8.4 Mbit/s download speed and only 0.8 Mbit/s upload speed [69]. FSO offers much higher transmission bitrates without the need for cables. Even so, given the potentially adverse effects of weather, one additional factor that needs to be taken into account is path diversity. Thus, two parameters of interest that we consider in the next chapter are end-to-end speed and the extent of topological diversity, which in turn has a direct impact on resilience.

Typically, fast broadband needs an infrastructure that equip with all most recent and high specification technologies, but this isn't cheap, and network operators will consider what return they can receive on their investment. This, unfortunately, can mean that more remote locations miss out on upgrades as they're not considered economically viable [70].

The geographical location determines the terrain, with challenges defined by topological characteristics such as flatlands, hilly areas, dense forest areas, etc. A hilly and densely forested area may have more fading that reduces signal intensity due to the propagation of signal along multiple paths and interference leading to more attenuation compared to a relatively flat area with less tree cover [71]. The location influences the cost of developing infrastructure and transportation of all necessary equipment as well as troubleshooting and maintenance of the network.

Long-range FSO links are degraded by weather phenomena and atmospheric turbulence across significantly different time scales. As an example, the performance of an operational link can degrade sharply in thick clouds, severe fog, or dust storms. Atmospheric turbulence is also of concern because even in clear weather, local temperature gradients, pressure variations, and scattering by airborne particles produce a varying refractive index along the transmission path [72].

Suppose the GA population is not chosen intelligently. In that case, it becomes challenging to find the desire solution to the problem, whether in the case of initial population selection or the selection of a population for the next generation. Factors that can affect the performance include search space, fitness function, diversity, problem difficulty, selection pressure and a number of individuals [73].

2.2 Current Technologies in Rural Areas

Currently only a few technologies are available in rural areas. These are:

2.2.1 ADSL

Asymmetric Digital Subscriber Line (ADSL) technology delivers broadband using the standard copper telephone lines – the connection speed will depend on which type of ADSL is being used and the quality and length of the line from the telephone exchange to the premises. Getting high-speed ADSL broadband delivered over existing copper cables is not possible in

many areas, as the distance from the exchange or the street cabinet is so far that the broadband signal degrades, and bitrates drop.

However, bitrate is not the only limiting factor in rural areas. Connection reliability is also a setback, with connections failing randomly for no clear reason, or due to weather such as heavy rain, snow or wind.

2.2.2 Mobile Internet

Fixed wireless uses specific frequencies of the radio spectrum to transmit signals through the air in a similar way to mobile phone networks. Most wireless ISPs offer limited coverage in specific areas such as rural villages. A 4G mobile network connection could potentially provide 100Mbps bitrates. Unfortunately, the areas failed by poor fixed line broadband provision are often the same areas with poor mobile phone networks, particularly rural areas. While 2G/3G network coverage is better, it is far slower. Without unlimited data plans, users also face monthly caps on use as part of their contract. Weather conditions can also adversely affect the service.

It is identified that more than half of the world's population were using mobile internet by the end of 2020, but 6% of the population is still left without mobile Internet access [146]. This report also states that 4G network coverage continues to expand, while a number of countries have also seen significant increases in 5G coverage. People in rural areas are increasingly using mobile Internet as there is no other option available, but a significant rural-urban gap remains. Data from a consumer survey 2020 [147] show that those who are unconnected in rural areas, are due to:

- Geography – Most unconnected people live in rural areas. For example, in Nigeria, 63% of those living in rural areas are unconnected, compared to 40% of people living in urban areas.
- Age – People over 35 years old are more likely to be unconnected than those who are younger. For example, 58% of people over 35 years old are unconnected in Algeria, whilst only 13% of the younger population are unconnected.
- Education – A significant proportion of unconnected people only completed primary school. For example, 93% of those with only primary school education or below are unconnected in Bangladesh, whilst 56% of those who completed secondary education remain unconnected.
- Literacy – Those reporting low levels of literacy are also more likely to be unconnected. For example, in Guatemala, 78% of people who are illiterate are unconnected, compared to only 31% of those who are literate.

4G coverage remains largely unchanged by all operators in the UK whereas 5G services are being rolled out and are now operating in over 40 towns and cities across the UK. 91% of the UK has good 4G coverage but this leaves 9% of the UK without acceptable outdoor 4G coverage, predominantly in rural areas [74].

2.2.3 Wireless and Satellite Broadband

Satellite broadband is an option for those who live in rural areas where traditional fixed line based broadband services aren't available. It uses a satellite dish to provide access to broadband services. The main advantage of satellite broadband is that it can be provided virtually anywhere in the world, as long as there is a clear line of sight. But despite providing up to around 20Mbps bandwidth, satellite connections have high latency due to the far larger distances involved between satellites and the ground. High latency connections make it very difficult to use some applications such as Voice Over IP (VOIP), Skype or streaming video. It is not always an option in mountainous regions as the satellite is typically low on the horizon. Geostationary Earth Orbit satellites are comparatively simple and cost effective compared with other constellations [62]. Focusing on the unserved and underserved communities, Low Earth Orbit (LEO) satellites stipulate low latency and high-speed broadband services to rural communities. The LEO satellite system is becoming the global communication service provider's counterpart to local terrestrial paging, cellular, and fibre networks. Telesat, OneWeb, SpaceX [75], [76], and Amazon [77] have instigated LEO satellites to establish low cost, affordable, and sustainable global communications.

In the mountainous regions where a wired network is not possible, wireless communication such as Wi-MAX and 3G/LTE/4G technologies and the use of TVWS for long distance wireless coverage is possible.

2.3 Recent Activity with Rural Broadband

Wireless communications allow users of mobile devices to access the Internet at any time and many locations. For instance, IEEE 802.11 (Wi-Fi), Bluetooth, and the Infrared Data Association (IrDA) protocol are intended for short-range wireless data communications [48], while Long-Term Evolution (LTE) is for long-range wireless communication for both mobile phones and data terminals [78].

Broadband Delivery UK (BDUK), part of the Department for Culture, Media and Sport, is delivering superfast broadband and local full fibre networks to the nation. The Government is supporting investment and aims to provide superfast broadband (bitrates of 30 Mbps or more) coverage to 95% of the UK by December 2017. Also, the Government is determined to provide access to basic broadband (2Mbps) for all. This stimulates private investment in full fibre connections by 2021. The Government provided Broadband Connection Voucher Scheme to support across 50 UK cities. The Government Mobile Infrastructure Project supported investment in mobile infrastructure to improve coverage for voice calls and text messages for the final 0.3-0.4% of UK premises that were unable to receive a service.

The United States of America, sponsored by Microsoft, set a clear and ambitious, but achievable goal that aims to eliminate the rural broadband gap within the next five years by 2022 through the Rural Broadband Strategy. A combination of new technologies can substantially reduce the cost of extending broadband coverage across the nation. A technology model that uses a combination of wireless technology leveraging the TV white space spectrum, LTE fixed wireless, and satellite coverage can reduce the initial capital and operating costs by roughly 80% compared with the cost of using fibre cables alone and by approximately 50%

compared with the cost of current LTE fixed wireless technology [79]. Wireless technologies that utilise TV white space are designed to transmit in the VHF and UHF spectrum that was traditionally allocated for broadcast television. In addition, due to the characteristics of the television spectrum signals being able to travel long distances, it is ideally suited for bringing broadband to rural America. Overall, TV "white space" technologies appear to be a good solution for about 80% of rural America.

2.4 Free-Space Optical (FSO) Communication

Free-Space Optical (FSO) communication is a line of sight (LOS) wireless technology supporting high data transmission rates along a narrow beam of light through the air. FSO is a license-free transmission technology. FSO communication propagates in free space to wirelessly transmit data. It provides high data rates without the necessity of long cables for connecting the rural community to the network infrastructure [80]. FSO systems have some outstanding advantages such as high bandwidth, the ability to concentrate beams, no need for license fees to use radio channels, outstanding security against wiretapping and better environmental compatibility such as no electromagnetic pollution. There are also some disadvantages with FSO such as stronger atmospheric influences like rain, snow, fog, and direct sunlight.

FSO transmission uses the same principle as fibre optic transmission except that it employs air as the transmission medium rather than a waveguide. This LOS technology normally requires the signal to be transmitted in a narrow beam between the transmitter and the receiver. The quality of the FSO transmission is contingent on the weather conditions of the transmission medium, thus, FSO transmission is normally better in cold, dry, clear weather. Normally, FSO data rates range from 100Mbps to 10Gbps [81]. A typical optical communication system comprises a transmitter that encodes a message into an optical signal, a channel that carries the signal to its destination, and a receiver that decodes the message from the received optical signal. The main limitation of an FSO link is low cloud/fog attenuation that provides poor FSO link availability limiting the transmission signal to noise ratio to 100dB/km [82].

Now-a-days FSO systems are used for the quick setup of short-range inner-city links (point-to-point), broadband links over motorways, rail tracks or rivers, and for secure connections between company buildings and corresponding institutions. More applications could include the connection of mobile emergency response services, the fire brigade, and armed forces to a backbone network during crisis operations. The transmission rates may reach gigabit-per-second over few kilometres using the unlicensed frequencies in the hundreds of Terahertz.

Consideration towards FSO is enhanced by several factors. First and foremost, more and more bandwidths are needed by the user. The way e-commerce is getting popular, service providers need to offer more bandwidth to the customers. Secondly, FSO is cheap. Service providers are always in pressure to give service to customers at the lowest price. In these cases, FSO is not only cheaper than fibre optic cabling but is also cheaper than other popular

technologies like the Digital Subscriber Line (DSL) or cable modem services. Finally, FSO technology brings new possible services such as resilient delivery of data to remote areas.

2.5 Network Planning Tools

Network planning is an essential part of any network deployment. The prices of commercial network planning tools are expensive, making them inaccessible to many researchers and organizations with small networks. Developing a network planning tool consists of different activities such as programming, algorithm implementation and so forth. A network planning tool can add intelligence, automation and efficiency to multilayer optical networks, from design to commissioning. It provides a comprehensive set of engineering, planning and analysis functions for multi-technology networks, reducing total cost of ownership by ensuring the deployment of cost-efficient configurations. The current backbone networks are built according to a multilayer concept, where Multi-Protocol Label Switching (MPLS) traffic flows are routed over optically switched light-paths with typical rates of 10/40/100 Gbps. Multilayer optical networks have prompted the development of new cross-layer approaches to network planning, combining optimization of the optical and electronic layers [47].

We design and implement a network-planning tool to evaluate the feasibility of low-cost network deployment in a rural context. Although there are some networking-planning tools available, most of them are commercial. More than that, our network-planning tool will consider the environment and trade-offs to discover suitable FSO node locations.

2.6 Genetic Algorithms (GA)

Genetic Algorithms (GA) are adaptive heuristic algorithms based on the ideas of natural selection and genetics. GA is one of the most familiar categories of evolutionary algorithm. GA is global search method that mimics the metaphor of natural biological evolution, which was initially developed by John Holland, University of Michigan in 1970s and further developed by De Jong and Goldberg. GA operates on a population of potential solutions applying the principle of survival of the fittest to produce best approximations to a solution. Although randomised, GA is by no means random, instead it exploits historical information to direct the search into the region of better performance within the search space [84]. In simple words, GA provides a global search and optimisation technique through coding and special operators.

A GA works with a number of solutions in each iteration which are often coded in binary strings. At each generation, a new set of probable solutions is created by the process of selecting individuals according to their level of fitness in the problem domain and breeding them together using operators borrowed from natural genetics. This process leads to the evolution of populations of individuals that are better suited to their environment than the individuals that they were created from.

Every probable solution or individual from population is assigned a fitness which is directly related to the objective function of the search and optimisation problem. After that, the population of individuals is modified to form a new population by applying three core

operators: reproduction, crossover, and mutation. Therefore, GA is the method of solving problems by utilising the processes of selection, crossover and mutation.

A GA maintains a population of candidate solutions such as shown in Figure 3 where each candidate solution is usually coded as a string called a chromosome. In this instance, each gene could be the X-Y coordinates of a node location on a map. A set of chromosomes forms a population, which is evaluated and ranked by a fitness function. This plays a critical role in GAs because it provides information concerning how good each chromosome is, which in turn determines its likelihood of mating to ensure its genes survive into the next generation. The initial population is usually generated at random. The evolution from one generation to the next one involves three main steps: fitness evaluation, selection and reproduction [85].

N_1	N_2	N_3	N_4	N_{n-3}	N_{n-2}	N_{n-1}	N_n
XY	XY	XY	XY	XY	XY	XY	XY

Figure 3: Possible Chromosome Structure (N_i is Number of the Node)

Each chromosome represents a possible solution. Solutions from one population are taken and used to form a new population which may be better than the old one [86]. Solutions that are selected to form new solutions are selected randomly but based on their fitness – the more suitable they are the more chance they have to reproduce and survive. This is a repeated process until some condition is met which is called the termination criterion. A simple GA working principle is shown through the flowchart in Figure 4.

Once an initial population of chromosomes has been created the chromosomes are evaluated based on their fitness in terms of line-of-sight to their peers, the range to their peers, etc. The fitness function is used to assess the "goodness" of the solution and provides a score with which chromosomes can be compared. As a result of evaluation, all the chromosomes are thus ranked based on their fitness score. Environment selection permits the fittest solutions from the current population to be carried forward to the mating pool and the lower scoring ones are killed off, losing the chance to reproduce.

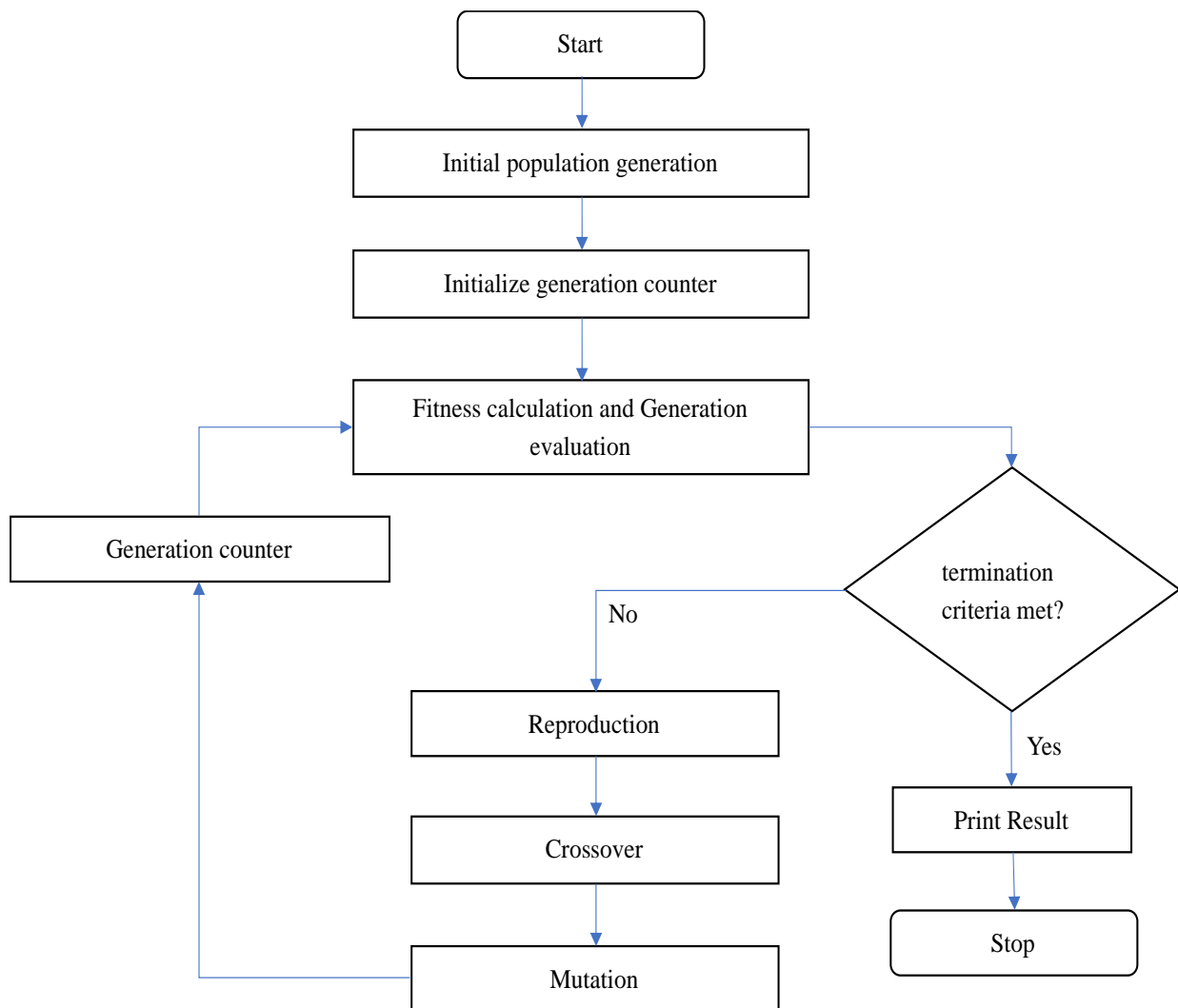


Figure 4: Simple Genetic Algorithm Flowchart

The mating pool retains desirable solutions. Pairs of solutions are selected randomly from the mating pool to reproduce. Reproduction covers two actions: crossover and mutation. Crossover involves the splicing together part of the chromosome from one parent with the residual part of another. This can be single point crossover or a more complex mechanism. This is followed by mutation. The mutation mechanism is used to maintain diversity between generations by randomly altering genes with a certain probability. This process leads to the formation of the next generation of chromosomes, which are then evaluated in terms of fitness and the cycle repeats until a terminal condition is reached as shown in Figure 4.

The steps of simple Genetic Algorithm are described below:

1. Start: Generate random population of n chromosomes.
2. Fitness: Evaluate the fitness $f(x)$ of each chromosome x in the population. Rank the results from best to worst.
3. Termination condition: If the end condition is satisfied, stop, and return the best solution in current population, else move to Step 4.
4. Place the best solutions in the mating pool. Discard the weaker solutions.

5. New population: Create a new population by repeating following Steps 6-9 until the new population is complete.
6. Selection: select two parent chromosomes from the mating pool. This can be according to their fitness: the better the fitness, the bigger chance to be selected by using various selection schemes.
7. Crossover: With a given crossover probability, crossover the parents to form new children. If no crossover is performed, the children are an exact copy of the parents.
8. Mutation: With a given mutation probability, mutate new child at each gene position in the chromosome.
9. Replace: place children in a new population.
10. Once new population is complete return to Step 2.

In contrast to a local search method, GA is based on a set of independent operators such as selection, crossover, and mutation controlled by a probabilistic strategy. GA is an example of an evolutionary algorithm which is used to solve problems that do not have well-defined solutions.

2.7 Multi-Objective Evolutionary Algorithms (MOEA)

The presence of multiple objectives in a problem gives rise to a set of optimal solutions instead of a single optimal solution. One of these Pareto-optimal solutions cannot be said to be better than the other. This demands a user to find as many Pareto-optimal solutions as possible. Classical optimization methods suggest converting the multi objective optimization problem to a single-objective optimization problem by emphasizing one particular Pareto-optimal solution at a time. Again, over the past decade, a number of Multi Objective Evolutionary Algorithms (MOEAs) have been suggested [27] [87] [88] [89]. The primary reason for this is their ability to find multiple Pareto-optimal solutions in one single simulation run.

An evolutionary algorithm uses mechanisms inspired by biological evolution, such as reproduction, mutation, recombination, and selection to solve problems. Genetic Algorithms are the most popular type of evolutionary algorithm. A potential solution to a problem is typically encoded as a set of properties or numbers to form a chromosome. Each candidate solution can then be mutated and crossed with other potential solutions to create new solutions that may be better, or more fit, than the parent(s) based on a specific fitness function. Better solutions have a greater chance of reproduction, effectively leading to a guided search for the “best” overall solution. However, for multi-objective problems the trade-off between fitness objectives must be addressed. There are many methods to solve multi-objective programming problems, and they can be classified into two types [90] as follows:

1. Switching to single-objective programming: This method integrates all the objectives into a single objective by developing a scalar equation using a weighting factor
2. Finding Pareto optimality: this involves the construction of a Pareto front by searching a set of potential solutions and creating a set of non-dominated cases according to specific criteria.

Multi-objective optimization is a procedure for optimizing two or more conflicting objectives simultaneously subject to certain constraints [91]. The conventional way to solve

multi-objective optimization problems is to use weighted sum technique or constraint-based technique. One of the ways to solve multi-objective optimization problems is to use multi-objective evolutionary algorithms (MOEA). Most MOEAs in the literature are based on Genetic Algorithms (GA). These include the Nondominated Sorting Genetic Algorithm (NSGA), Strength Pareto Evolutionary Algorithm (SPEA) etc.

Evolutionary based algorithms use the concept of GA to solve the multi-objective optimization problem. There are two types of multi-objective Evolutionary Algorithm that mostly discussed, those are: non-elitist and elitist. The steps of an MOEA are following:

- 1) Initialization: initialize a random population
- 2) Fitness assignment: assigning a rank to each individual of the population
- 3) Variation: apply crossover, mutation on the mating pool
- 4) Environmental selection: select the best solutions according to the mating pool
- 5) Termination criteria: repeat above steps until termination criteria met.

MOEAs are based on a population in which the iterations are performed on a set of solutions and after each generation it returns more than one solution. MOEAs are popular for solving multi-objective optimization problems due to the following reasons:

1. They are easy to implement
2. They typically return more than one optimal solution
3. There is less chance of the algorithm getting stuck in local minima
4. They are flexible and robust
5. They do not require any a prior knowledge of the problem [92]

In MOEAs and even in GA, many times good solutions get lost when crossover or mutation results in children which are weaker than their parents. Often the MOEAs re-discover these lost improvements in subsequent generations, but there is no guarantee. To solve this issue, researchers use a mechanism known as elitism [92]. Elitism is a mechanism to preserve the best solution(s) at every generation. There are two main goals in every MOEA: convergence towards the optimal Pareto front and maintaining diversity in the Pareto solution set.

Most of the difficulties associated with classical optimization methods as summarized below are eliminated when using evolutionary algorithms:

- An algorithm has to be applied many times to find multiple Pareto-optimal solutions.
- Most algorithms demand some knowledge about the problem being solved.
- Some algorithms are sensitive to the shape of the Pareto-optimal front.
- The spread of Pareto-optimal solutions depends on efficiency of the single objective optimizer.

In general, the goal of a multi-objective optimization algorithm is not only guiding the search towards the Pareto-optimal front but also maintaining population diversity in the set of the nondominated solutions.

2.8 Strength Pareto Evolutionary Algorithm (SPEA)

Strength Pareto EA is an approach to multi objective optimization that uses a mixture of established and new techniques in order to find multiple Pareto-optimal solutions in parallel. SPEA stores the nondominated solutions found so far externally as done in [93] [94] [95] and uses the concept of Pareto dominance in order to assign scalar fitness values to individuals. It also performs clustering to reduce the number of nondominated solutions stored without destroying the characteristics of the tradeoff front [96]. SPEA uses the same technique as Genetic Algorithm (GA) but the fitness of an individual is determined only from the solutions stored in the external nondominated set and all solutions in the external nondominated set participate in the selection. This algorithm-based approach has the following features:

- It uses an improved fitness assignment scheme which for each individual it takes into account how many individuals it dominates and is dominated by.
- Archive truncation method guarantees the preservation of boundary solutions.
- A nearest neighbor density estimation technique is incorporated which allows a more precise guidance of the search process.

The flow of the algorithm is as follows:

- 1) Initialization: Generate an initial population P_0 with N size and create an empty archive Pareto optimal set (of max size M).
- 2) The following steps are used to update the external Pareto-optimal set
 - a. Search the population for the nondominated individuals and copy them to the external Pareto set.
 - b. Search the external Pareto set for the nondominated individuals.
 - c. If the number of the individuals in the external set is $<M$, then, the best dominated individuals are copied into the new external set
 - d. If the number of the individuals in the external set is $>M$, then an archive truncation technique is invoked which iteratively removes individuals.
- 3) Fitness assignment: Calculate the fitness values of individuals in the external Pareto set and the population P_t as follows:
 - a. Assign Strength value for each individual in the external Pareto set and the population P_t .
 - b. The raw fitness is determined by the strength of its dominators in both the archive and population.
- 4) Selection: select two individuals at random from the updated external set and compare their fitness. Select the better one and copy it to the mating pool.
- 5) Crossover and Mutation: Perform the crossover and mutation operations according to their probabilities to generate the new population.
- 6) Termination: Check for stopping criteria.

SPEA-II is an extension of the original Genetic Algorithm [97], for multiple objective optimization problems. The objective of the algorithm is to identify and preserve a set of nondominated solutions, ideally a set of Pareto optimal solutions. The Pareto optimal solutions,

called the Pareto optimal set, are made up of the best nondominated solutions in the objective space. A computational flowchart of the SPEA approach is presented in Figure 5.

At generation 0, an initial primary population and an empty secondary population are generated. The initial population has a certain size of candidate solutions. Each new candidate solution is generated randomly so it needs to pass the Line of Sight (LOS) test explained on Chapter 3.6 to confirm its viability. If the new candidate solution passes the LOS test, it is put into initial population. Otherwise, it is eliminated, and the generation process repeats until initial population is full.

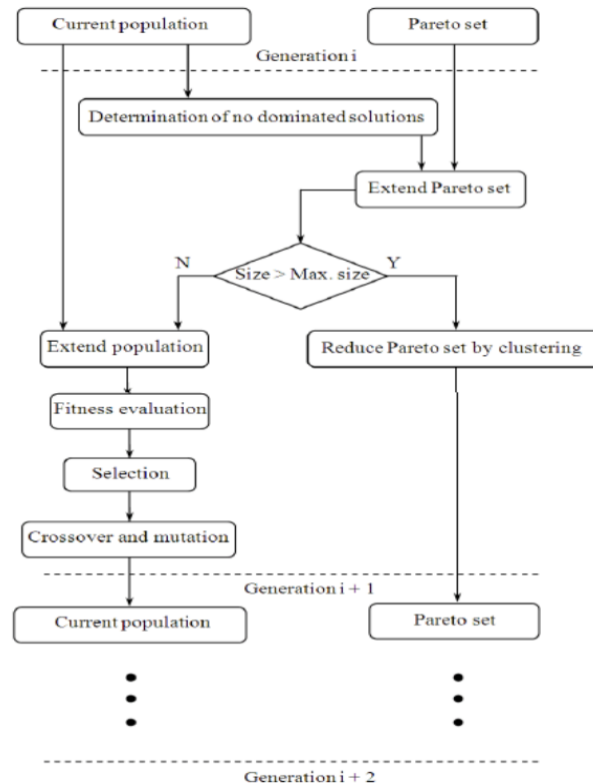


Figure 5: Flowchart of SPEA-II [98]

In the evaluation stage, each candidate solution in the primary population is evaluated and then assigned multiple objective values. In terms of our research, end-to-end speed and diversity are the two objectives considered.

When the evaluation step completes, the primary population and the secondary population are merged into a temporary population for fitness assignment. The size of this temporary population is the sum of primary population size and secondary population size except at the first generation as the initial secondary population is empty. Fitness assignment in SPEA-II [99] is on the basis of Pareto dominance. SPEA considers two methods: dominance rank and dominance count. Dominance rank records the number of candidate solutions by which a solution is dominated, and a strength value is assigned to each solution according to the number of solutions it is dominated by. Every solution in the temporary population has a strength value. After obtaining the strength value of solutions, a dominance count method is used for determining the fitness value. The fitness value of a certain solution is calculated by

summing up the strength values of its dominators. In this step, fitness with a large value means the solution is dominated by many solutions and conversely the fitness value of nondominated solutions is zero. Hence, it is important to note that a higher fitness value actually corresponds to the solution being less fit in the SPEA fitness assignment.

Environmental selection determines which solutions in the primary and secondary populations form the next generation secondary population. To update the secondary population, a truncation method [99] of selection is used to select a constant number of solutions. All the nondominated solutions in the temporary population are copied into the next generation secondary population. One of three actions is then taken depending on the secondary population size and the size of this nondominated solution set. Those are:

- If the secondary population size is the same as size of the nondominated solution set, then the secondary population is updated straight way.
- If the size of the nondominated solution set is smaller than the secondary population size, then the best dominated solutions from the temporary population are used to make up the shortfall.
- If the size of nondominated solutions exceeds the secondary population size, then some solutions are eliminated based on their density information given by the K-nearest neighbours [99].

After the secondary population update completes as above, mating selection is performed to determine which solutions are chosen from the secondary population to form the mating pool to reproduce offspring. The mating pool is cleared before the mating selection at every generation. Tournament selection with replacement is used. This means that two solutions are randomly selected from the secondary population and after comparing their fitness, the better solution is copied into the mating pool. The size of the mating pool is determined with a nominal value and can be similar to the size of secondary population.

In the reproduction stage, offspring are created by two genetic operations: crossover and mutation, both controlled by certain rate. Before reproduction, the primary population is purged ready for holding the offspring. Initially, two chromosomes from the mating pool are selected randomly to be the two parent chromosomes, then two children are generated according to crossover and mutation operators which is described below. Before entering the primary population, the children go through a LOS test, and failures are eliminated. This process repeats until the primary population is full and generation counter incremented.

Crossover produces offspring by exchanging gene information from the two parent chromosomes, that permits the combination of part of each parent to potentially generate better offspring. In the binary representation, there are many kinds of crossover such as: one-point crossover, two-point crossover and uniform crossover. Only one-point crossover is considered here. In simple one-point crossover, the position of crossover point is firstly randomly selected. Then, the two parent chromosomes are broken into two parts according to the crossover point. Finally, the first part of parent 1 and the second part of parent 2 are combined to become the first offspring. Similarly, the second offspring is generated by combining the second part of parent 1 and the first part of parent 2. Figure 6 shows a demonstration of the crossover process.

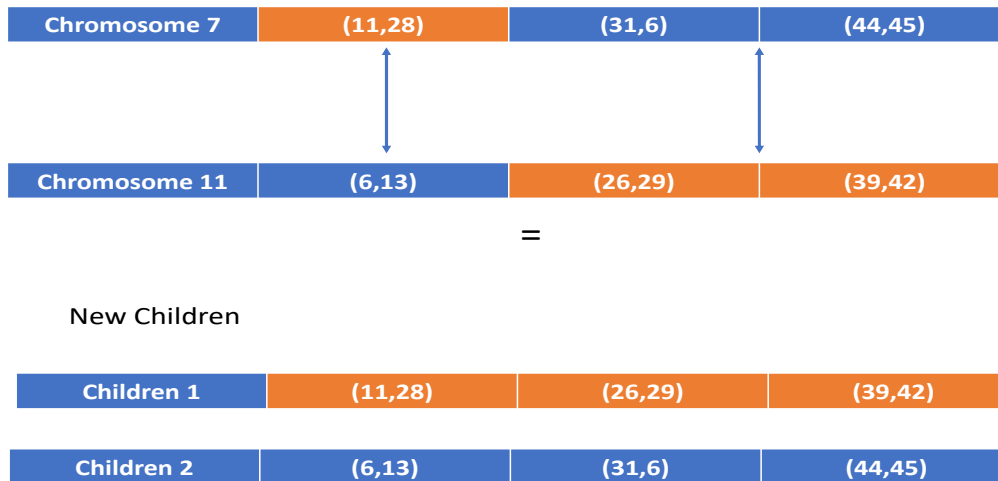


Figure 6: Crossover Demonstration

Mutation is used for maintaining the genetic diversity from one generation to another by potentially altering some gene values randomly. For each gene in a chromosome, a uniform random variable is firstly generated in the interval $[0, 1]$. An example of mutation is shown in Figure 7.

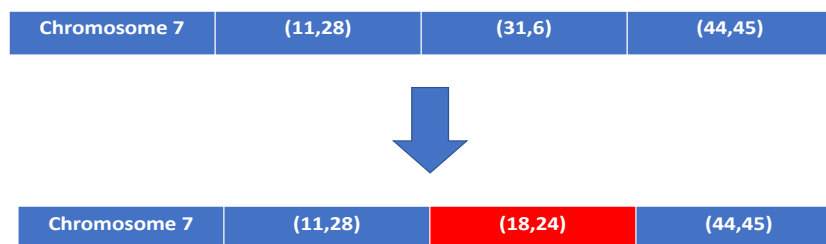


Figure 7: Process of Mutation Operator

When the terminal condition, which is the maximum number of generations, is reached the simulation stops. Then, the secondary population in the final generation is the outcome.

2.9 Pareto Front (PF)

In MOEAs, a solution is called a non-dominated solution if it is better in all objectives than other solutions or it is strictly better than other solutions in at least one objective. When we plot these non-dominated solutions on a graph, such a graph is called a Pareto front, which is shown in Figure 8, and solutions on the Pareto front are called the Pareto Set (PS). A set of non-dominated solutions found at the termination of the MOEA produce the (near) optimal Pareto front. These solutions are called the optimal Pareto solutions.

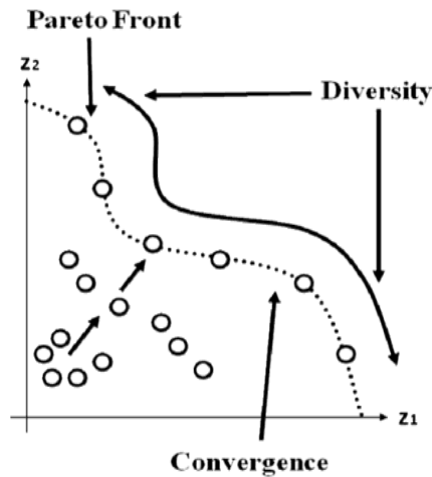


Figure 8: MOEA with Pareto Front [27]

There are many methods to solve multi-objective programming problems, and it can be classified as two types [90] as follows:

- a) Switching to single-objective programming: This method integrates all the objectives into a single objective by developing a scalar equation using a weighting factor
- b) Finding Pareto optimality: This constructs a Pareto front by searching a set of efficient solutions for the best non-dominated ones according to specific objectives.

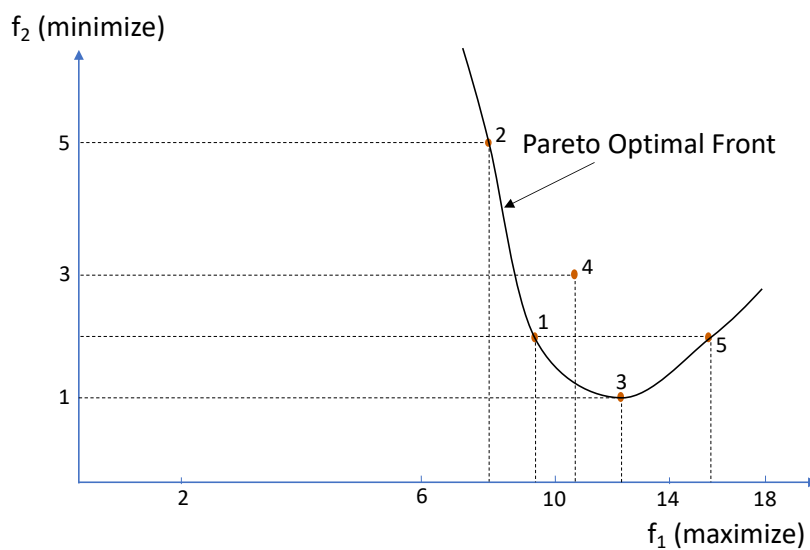


Figure 9: Pareto Optimal Front Demonstration

Let us compare solution 3 and 5 in Figure 9, because this comparison reveals a significant observation. It is clear that solution 5 is better than solution 3 in the first objective, while solution 5 is worse than solution 3 in the second objective. This simply suggests that we cannot conclude that solution 5 dominates solution 3, nor can we say that solution 3 dominates solution 5. This is because we aim to maximise f_1 while f_2 needs to be minimised. When this happens, we say that solution 3 and solution 5 are non-dominated with respect to each other. When each objective is essential, it cannot be said which of the two solutions 3 or 5 is better.

Many problems have multiple objectives, leading to set of optimal solutions, which are considered effective solutions. MOEA is highly effective for finding (near) optimal solutions in a simulation environment. Multiple trade-offs can be considered when deciding on a particular preferred solution. The aim is to obtain a set of Pareto optimal solutions/non-dominated solutions where none of the objective function values can be improved without degrading some of the other objective function values. In our case, each solution is evaluated by three objectives: reliability, cost, and speed.

Evolutionary based techniques use the concept of genetics and can solve multi-objective optimization problems. Due to the trade-offs involved, this sort of problem usually has no unique, perfect solution. Instead, they recognise that there are a set of equally valid, or non-dominated, alternative solutions known as the Pareto-optimal set. Due to multiple trade-offs to be considered, a Pareto ranking scheme, also known as non-dominated sorting, consists of identifying the non-dominated individuals by assigning them a suitable rank. When additional information or preferences are available for the objectives, it is then possible to discriminate between non-dominated solutions to select the most appropriate compromise.

According to the rural problem explained in Chapter 1.1, the presence of multiple objectives such as E2E speed, battery drain, and residual battery life gives rise to multiple optimal solutions based on the importance of these competing performance metrics instead of a single optimal solution. One of these Pareto-optimal solutions cannot be said to be better than the other. This depends on the preferences of the user/operator; it is thus convenient to find many Pareto-optimal solutions to construct a Pareto front. Despite this, in MOEAs and even in GAs, good solutions often get lost when crossover or mutation results in weaker children than their parents. Thus, there are two main goals in every MOEA:

1. Convergence towards the optimal Pareto front
2. Maintain diversity in the Pareto set

In general, the goal of an MOEA is to guide the search towards the Pareto-optimal front whilst maintaining population diversity in the set of nondominated solutions.

2.10 Dijkstra's Shortest Path Algorithm

Shortest Path calculations play an important role in networking. Often, a communication channel needs to be setup with few resources and cost, and the shortest path may give the highest bitrate. By using Dijkstra's Algorithm, these problems can be solved through an iterative process. The shortest path and the alternative paths can be computed based on traffic conditions. To solve the shortest path problem of a graph with nonnegative edge costs, Dijkstra's Algorithm is used to form the shortest path tree. This algorithm is used in routing and other network related protocols. For a given vertex in the graph, the algorithm finds the least-cost path from a source to any destination. [100] provides a summary of its operation. For example, the shortest path algorithm is used in traffic engineering in IP networks and to improve Intelligent Transportation Systems [101]. Dijkstra's algorithm finds the least cost path between two vertices in a graph. The algorithm works as follows:

1. Choose the source vertex

2. Define a set S of vertices and initialize it to empty set. As the algorithm progresses, the set S will store those vertices to which a shortest path has been found.
3. Label the source vertex with o , and insert it into S .
4. Consider each vertex, not in S connected by an edge from the newly inserted vertex. Label the vertex not in S with the label of the newly inserted vertex + the length of the edge. But if the vertex not in S was already labelled, its new label will be $\min(\text{label of newly inserted vertex} + \text{length of an edge}, \text{old label})$
5. Pick a vertex not in S with the smallest label, and add it to S .
6. Repeat from step 4, until the destination vertex is in S or there are no labelled vertices not in S . If the destination is labelled, its label is the distance from source to destination. If it is not labelled, there is no path from the source to the destination.

2.11 NSGA Crowding Distance

Non-dominated Sorting Genetic Algorithm (NSGA) is a popular MOEA that employs binary tournament selection, an elite preserving strategy, nondominated sorting and a crowding distance mechanism to obtain a good quality and uniform spread nondominated solution set. Enhancing a set of non-dominated solutions is a common task for MOEAs such as SPEA [99] and NSGA [27]. As mentioned earlier, diversity, convergence and robustness of Pareto solutions are important factors in the suitable operation of MOEAs. It is of great importance to produce a Pareto Front (PF) with good diversity and that can achieve efficient and varied solutions for decision-makers. To maintain diversity in a PF, a diversity maintenance strategy is often used. In this thesis, a Crowding Distance (CD) diversity maintenance strategy is implemented, similar to that used in NSGA-II. The NSGA-II CD mechanism is illustrated in Figure 10.

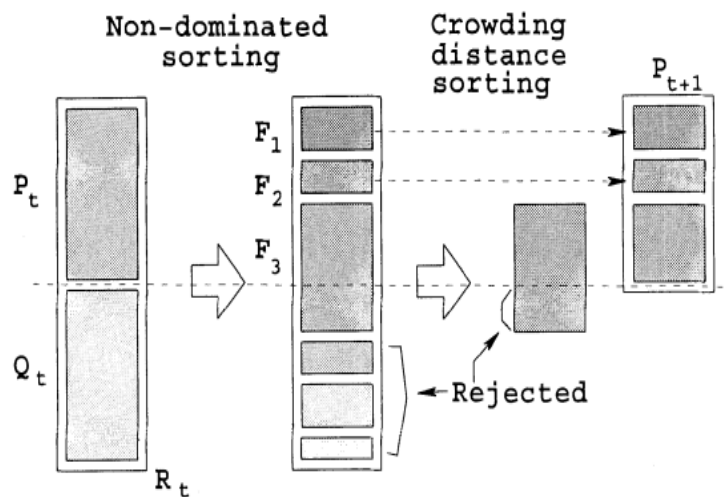


Figure 10: NSGA-II Crowding Distance Process [102]

The CD method was first suggested for use with NSGA-II in [102]. This method works by estimating the average side-length of cuboid surrounding an individual as shown in Figure 11. In this method, the population is first ranked according to the dominance relation to identify the non-dominated front. The CD is then calculated by summing the average distance of the two nearest solutions on either side of the solution corresponding to each objective function.

Finally, the population is sorted according to the dominance rank and crowding distance and the top-ranking individuals are picked and inserted into the next population until a desired number of individuals are selected.

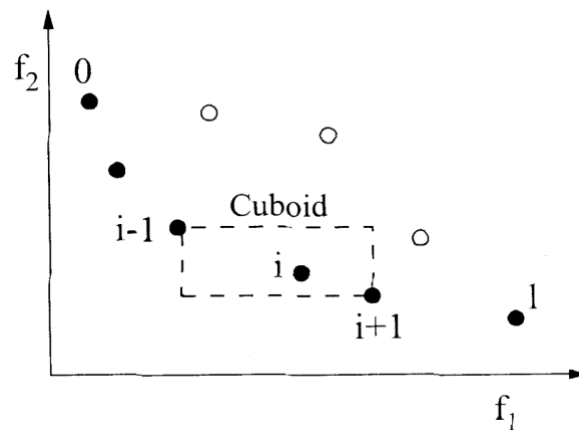


Figure 11: Crowding Distance Calculation [102]

This method can be used to maintain the population diversity without any extra user-defined parameters. It works by directly eliminating all redundant individuals, which have the lowest crowding distance from the non-dominated solutions. However, this may cause some loss of population diversity, both vertically and horizontally. The vertical diversity loss is due to many individuals in the population becoming non-dominated solutions after a few generations. The number of non-dominated and non-elitist solutions reduces quickly leading to vertical diversity loss. The horizontal diversity loss can result from two scenarios. Firstly, a solution which is near to one of its neighbours in a large cuboid can have a poorer diversity than another solution which is equidistantly located in a small cuboid. Assume that a solution needs to be removed from non-dominated solutions as seen in Figure 11. In this research horizontal vertices have been used. There are several vertical techniques described in [103] and [104]. This is the second scenario.

2.12 Software-Defined Networking (SDN)

The Software-Defined Networking (SDN) [105] [106] [107] [108] [109] architecture has emerged in response to the restrictions of traditional networking architectures due to the complexity of network devices and appliances and difficulty of installation and configuration. SDN proposes to decouple the control plane and data plane. This separation enables network operators/administrators to efficiently use network resources simplifying provisioning. Also, SDN provides ease of programmability [110] representing the characteristics of whole networks in a common manner. This streamlines the management of the network since it is decoupled from the data plane. Therefore, network operators can easily and quickly manage, configure, and optimize network resources with dynamic, automated and proprietary-free programs with the SDN architecture [111]. Additionally, since the network is logically centralized in SDN, controllers have global visibility, unlike conventional networking. Hence, they can dynamically optimize flow-management and resources.

The principal concept behind SDN is based on the separation of the control plane from the data plane in the architecture. This paradigm allows for a logically centralized controller, a

central program acting as the Network Operating System (NOS), controlling and managing the overall behaviour of the network. The network devices such as sensor nodes and switches exist in the data plane and simply forward data based on flow instructions provided by the controller. The overall architecture is shown in Figure 12.

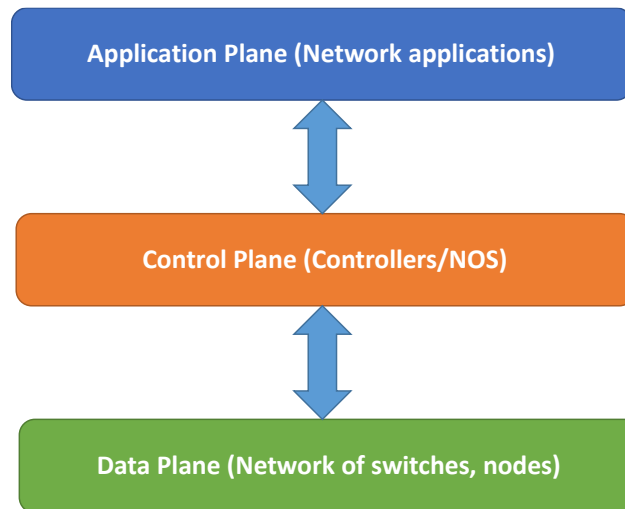


Figure 12: SDN Overall Architecture

SDN has emerged as a promising solution to revolutionize network deployment, operations and economic growth. This paradigm aims to address management and configuration complexities in legacy networks to reduce the total cost associated with deploying and running telecommunication infrastructures. Conventionally, network control and data planes are tightly coupled and deployed within the same proprietary network device. SDN presents an alternative paradigm by decoupling the control plane from the data plane, abstracting lower-level functionality of underlying hardware and enabling network programmability through a centralized controller. As the “brain” of the network, the controller can promptly and proficiently process and respond to changes in the data plane.

Although not a focus of our research SDN is envisioned to provide the support necessary for the traffic engineering framework we propose in later chapters of this thesis, enabling reconfiguration of the FSO relay nodes.

Chapter 3. Network Planning Tool

This research aims to improve the rural broadband infrastructure by developing a new network planning tool that will enable the construction of a low cost Free-Space Optical (FSO) based infrastructure. This technology can overcome many of the limitations of traditional cabling. It also provides a cheap alternative to mobile telecommunications technologies such as 4G LTE. The cellular model of this technology is designed for densely populated areas and face the same connectivity and capacity gaps as traditional broadband [79]. Satellite broadband can be a solution in very sparsely populated areas, but it often suffers from high latency, lack of significant bandwidth, and high data costs. Despite being cheap, FSO data rates range from 100Mbps to 10Gbps and with suitable planning, path redundancy can be provided for resilience.

Some rural areas have geographical disadvantages, being situated with many obstacles. The deployment costs associated with standard copper cable or fibre optic means these technologies cannot feasibly reach these areas. So before setting up the network infrastructure, intelligent network planning is necessary, and thus, we have designed a networking planning tool that is intended for rural areas. The network-planning tool is used to determine the suitable placement of relay nodes using FSO technology taking into account the terrain topography, as shown in Figure 13.

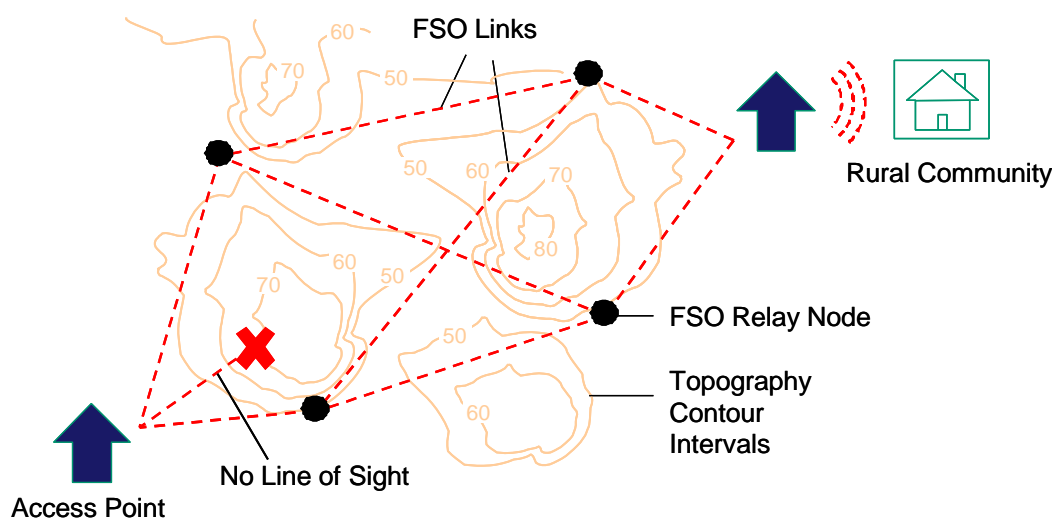


Figure 13: Network Planning Scenario

Due to hills etc., the tool can place nodes only where a line of sight is possible. The search and optimization are done through a Genetic Algorithm (GA) to start with. The GA is implemented in Java and used within our research to determine the placement of relay nodes for FSO-based rural broadband.

3.1 Initial Design with a Genetic Algorithm (GA)

A GA has been developed with specific parameters, determined based on various experiments. A fitness function has been designed for finding the node placement in rural terrain areas. Chapter 3.3 describes our fitness function. Apart from the fitness function, the GA parameters we have used are:

- No of Chromosomes: 20
- No of place-able nodes: variable (typically 3)
- Additional fixed nodes: 2 (source and destination)
- Crossover percentage: 70%
- Mutation percentage: 30%
- No of Generations: 100

A classic GA was described in Chapter 2.6, and a flow-chart of its structure is shown in Figure 14. As Line of Sight (LOS) is one of the requirements of FSO, our GA tool incorporates obstacles and tests for LOS.

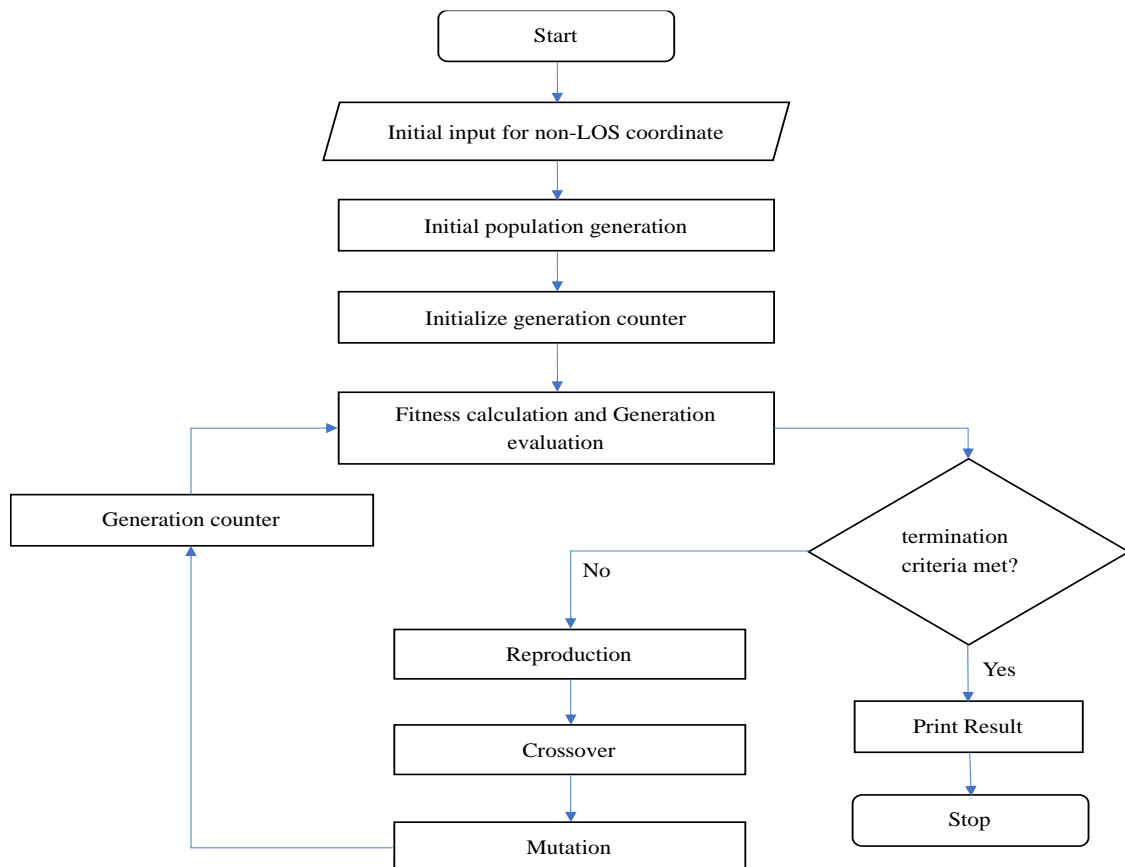


Figure 14: Simple Genetic Algorithm Process

With these GA parameters, together with the fitness function, a simple network planning tool has been designed for a prototype rural network infrastructure. An initial output from the tool is shown in Figure 15. In Figure 15, example relay node placement is shown with latitude and longitude coordinates.

According to our problem requirements, the tool searches in an attempt to optimize the relay node placement. Initially, Total fitness is determined through the shortest distance, which calculated as:

$$\text{Path distance} = \frac{\beta \cdot \text{Straight Path}}{\text{Actual Path}}$$

Where β = constant contributor (initial value 100)

$$\text{Straight Path} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

For our cases, the distance is calculated straight from the Source to Destination. Assuming these are located at diagonal opposites on a 50 x 50 grid, we have:

$$\text{Straight Path} = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$$

$$\begin{aligned} \text{Straight Path} &= \sqrt{(49 - 0)^2 + (49 - 0)^2} \\ &= \sqrt{2401 + 2401} \\ &= \sqrt{4802} \\ &= 69.3 \text{ km} \end{aligned}$$

The actual path of all nodes gives a comparison of the distance relative to the ideal situation, and we divide the straight path by the actual path. That means we are giving more credit if the actual path is closer to a straight path.

Now, the actual path is a summation of all inter-node distances along the path as follows:

$$\text{Actual Path} = \sum_1^{N-1} D_1 + D_2 + D_3 + D_4$$

Where...

$$\begin{aligned} D_1 &= \sqrt{(S_x - N_1^x)^2 + (S_y - N_1^y)^2} \\ D_2 &= \sqrt{(N_2^x - N_1^x)^2 + (N_2^y - N_1^y)^2} \\ D_3 &= \sqrt{(N_3^x - N_2^x)^2 + (N_3^y - N_2^y)^2} \\ D_4 &= \sqrt{(N_3^x - N_D^x)^2 + (N_3^y - N_D^y)^2} \end{aligned}$$

The mean absolute deviation of a data set is the average distance between each data value and the mean. The mean absolute deviation is a way to describe variation in a data set. This helps us get a sense of how uneven the distance between adjacent relay nodes along the path is. The deviation from the average is expressed as:

$$\text{Deviation} = \sqrt{\frac{1}{N-1} \sum_1^{N-1} (x_i - \mu)^2}$$

Where N = total number of nodes, and x_i = the distance from one node to another, including Source and Destination, in the actual path. For example:

$$x_1 = \sqrt{(S_x - N_1^x)^2 + (S_y - N_1^y)^2}$$

$$x_2 = \sqrt{(N_2^x - N_1^x)^2 + (N_2^y - N_1^y)^2}$$

$$x_3 = \sqrt{(N_2^x - N_3^x)^2 + (N_2^y - N_3^y)^2}$$

$$x_4 = \sqrt{(N_3^x - N_D^x)^2 + (N_3^y - N_D^y)^2}$$

We set μ = Average of straight-line distance. In our case, the average of straight-line distance, μ is calculated by computing distance from source coordinate (0, 0) to destination coordinate (49, 49) divided by the number of nodes from the source.

$$\mu = \frac{\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}}{N}$$

$$\mu = \frac{69.3}{4}$$

$$\mu = 17.325 \text{ km}$$

Finally, deviation from mean provide a means to ensure suitably spaced node locations that put equal distances between adjacent nodes, the including source and destination. This ensures the best end-to-end speed is achieved.

$$\text{Deviation from average} = \frac{\gamma * \text{Deviation}}{\mu}$$

Where, γ = constant contributor

3.4 Implementation

The primary focus at the initial stage is to implement a tool that can position relay nodes using Free Space Optical (FSO) technology. We create a GA-based implementation of design described in Chapter 3 for providing deployment information as shown in Figure 13 The network planning tool is implemented in accordance with the flow chart shown in Figure 14. In the GA process, the word ‘‘Chromosome’’ is encoded with positioning information of the relay nodes that will be placed in rural areas, with the possibility of obstacles. The initial role of this tool is to identify the correct node placement in a rural area under the constraint explained in Chapter 3.3.

The tool runs on a search space, assuming the rural broadband needs to be delivered from a fixed Source to Destination. The source has access to unlimited energy and so can always transmit at high power. The Destination is a rural receiver to provide broadband services to rural people. Initially, the tool randomly generates chromosomes representing different relay node placements.

3.5 Fitness function Evaluation

The first step of evaluation or selection is to assess the fitness of each chromosome and assign it a score. This process varies depending on the process. This can be based on simple

line of sight or range/distance between each node. The location of the nodes can be used to calculate inter-node distances and so determine a score based on the fitness function. A data-structure as shown in Table 2 is used to store the relevant information.

Table 2: Chromosome Evaluation

Chromosome No	Fitness score	Node 1 coordinate	Node 2 coordinate	Node 3 coordinate
1		N_1^1	N_2^1	N_3^1
2		N_1^2	N_2^2	N_3^2
3		.	.	.
4		.	.	.
5		.	.	.
6		.	.	.
7		.	.	.
8		.	.	.
9		.	.	.
10		N_1^{10}	N_2^{10}	N_3^{10}

Table 2 shows that each chromosome stores coordinates of three nodes. The tool creates a total of 10 chromosomes for this problem.

3.6 Line of Sight (LOS)

The fitness score can be based on the distance between each node to then decide what ranking it has relative to other chromosomes. It can also include a test for line of sight (LOS) of the individual node. As we chosen FSO for this research, it is necessary to have LOS between two nodes to form a valid transmitting link. This LOS test checks for any obstacle directly between two nodes. If nodes fail the LOS test, then there is no valid link between those nodes.

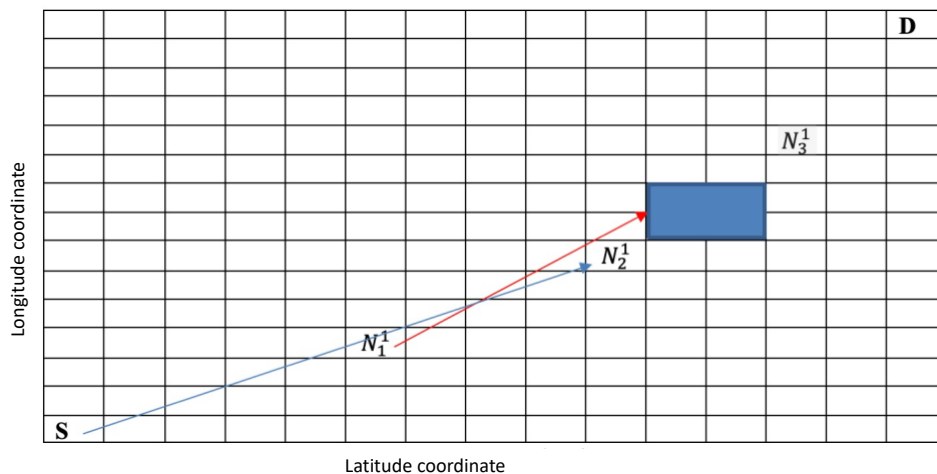


Figure 16: LOS Evaluation

Figure 16 shows a blue line that can see both N_1^1 & N_2^1 and the red line can see N_2^1 but can't see N_3^1 . Therefore, we apply certain rules as follows:

- According to Figure 16, any two node connections can't overlap the blockage in the search space. Otherwise, LOS is broken.
- Any node between source and destination has to connect with the closest node. The source or destination cannot skip the nearest node and connect with another node.
- In terms of geography, especially areas of hilly terrain, all nodes have to connect considering the height of land as the sender nodes need LOS to communicate with receiver nodes.

To do this line of sight, the procedure below is followed:

If a node finds any node or S or D in range, then it will receive 10 points straight away and then follow the below equation:

$$\text{Score} = \text{fitness function} + \text{line of sight factor.}$$

3.7 Chromosome Selection

The initial GA process starts with generating chromosomes randomly and evaluates them with fitness function evaluation explained in Chapter 3.5. The chromosomes showed in Table 2 are going through the LOS test and given fitness score using the fitness function explained in Chapter 3.3. After ranking them, the best chromosomes are stored before placing them into the mating pool. This is part of the evaluation/selection process but can be done separately. In Figure 16, if we take the best ranked 70% of chromosomes and delete 30% chromosomes, seven chromosomes will survive in the mating pool and so go on to the process of mating. For example, the best chromosomes are in Table 3.

Table 3: Best Chromosomes Preserved

Chromosome No	Fitness score	Node 1 coordinate	Node 2 coordinate	Node 3 coordinate
1	23	N_1^1	N_2^1	N_3^1
2	28	N_1^2	N_2^2	N_3^2
3	25	.	.	.
4
7
8
9
10				

3.8 Ranking

In this step, all chromosomes are sorting according to their fitness score as represented in Table 4. The most fit chromosome stays on the top, and potentially represents our final output if the tool has reached the final generation.

Table 4: Ranked Chromosomes

Chromosome No	Fitness score	Node 1 coordinate	Node 2 coordinate	Node 3 coordinate
2	28	N_1^1	N_2^1	N_3^1
3	25	N_1^2	N_2^2	N_3^2
1	23	.	.	.
4
7
8
9
10				

3.9 Final Output

The tool gives an output of the coordinate of nodes in search space. These outputs are represented in different forms. The first example output, as shown in Figure 17, represents a topological representation of a rural network without considering obstacles. The output shows the actual node placement.

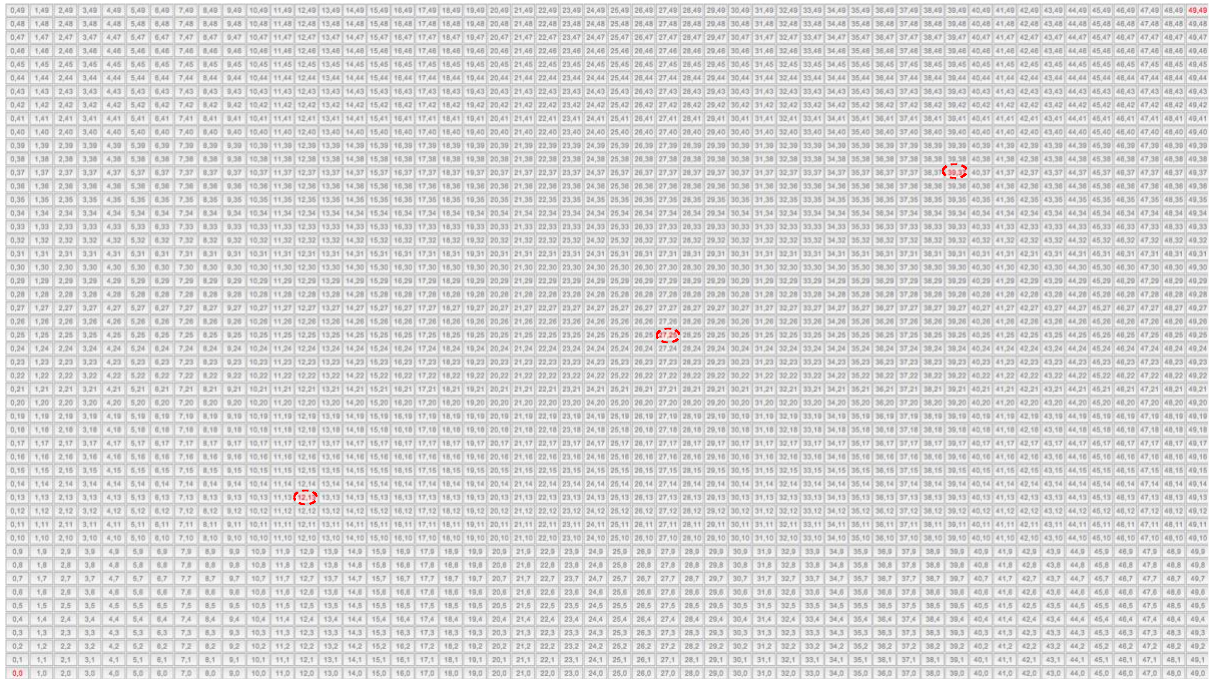


Figure 17: Example LOS Topological Representation

Another output considers an obstacle. The blue coordinates in Figure 18 represent the obstruction. Rural areas may have plenty of obstacles, such as hills. Therefore, LOS needs to consider in this tool. This tool should be able to detect any blockage in the LOS between nodes.

We run our tool to see whether it yields the same output as shown in Figure 19, the ideal placement of relay nodes. The actual output from tool, given in Figure 17, looks very similar to the ideal solution. This provides confirmation that our tool is functioning correctly. The slight variation in the placement is a result of quantization error as our scoring mechanism is based on a discretised range profile.

3.11 Performance Validation

One crucial issue for the success of genetic algorithms, especially for complex problems, is to avoid the premature convergence to suboptimal regions of the search space. The premature convergence of a genetic algorithm arises when the chromosomes of some high rated individuals quickly attain dominance in the population, constraining it to converge to a local optimum. In this case, the genetic operators cannot produce any more descendants better than the parents [112]; the algorithm's ability to continue the search for better solutions is, therefore, substantially reduced.

In a genetic algorithm, to avoid premature convergence is imperative to preserve the population diversity during the evolution. Among the methods used for this, we can enumerate limited selection, dynamic application of mutation, constraints for crossover and mutation probabilities, stochastic universal sampling, variable fitness assignment, population partial reinitialization, individuals grouping methods, restricted mating, elitism, chromosome conservation techniques, ranking sort based on Pareto dominance, local searches based on diversity. All these methods are heuristic by definition, and their effects vary for different problems.

In the real world, there are difficult instances where no single method is adequate. Again, another critical aspect is the double character of the population diversity:

- Diversity in the objective space
- Diversity in the parameter space

For some problems, it is sufficient to preserve the diversity in one space; for others, a good GA must maintain population diversity in both areas. Therefore, there are two imperatives related to premature convergence:

- Identify the occurrence of the premature convergence through various measures
- Evaluate its extent

The measures such as population size, generation count, mutation rate etc., to identify the premature convergence are, in fact, measures for the level of population degeneration. [113] uses as measure the difference between the average fitness and the best fitness in the population. Depending on this difference, they adaptively vary the crossover and mutation probabilities. The authors of [114] propose a statistical measure of the average Hamming distance between individuals and the variance of Hamming distances, both independent of chromosome number and population size.

3.11.1 Impact of Population Size

To assess and eliminate premature convergence from GA, tests have been undertaken to determine the impact on the final output. Therefore, we explored increasing the population size; for example, I have set Generation = 15, Crossover = 70%, Mutation = 5%, which is our initial value. And our number of chromosomes, that population size is increased to 20. This is one strategy to attempt to eliminate premature convergence. The final generation output is shown in Figure 20.

Generation 15			
Chromosome No	Node 1	Node 2	Node 3
1	(19,15)	(23,30)	(25,47)
2	(19,15)	(23,30)	(25,47)
3	(19,15)	(23,30)	(25,47)
4	(19,15)	(23,30)	(25,47)
5	(19,15)	(23,30)	(25,47)
6	(19,15)	(23,30)	(25,47)
7	(19,15)	(23,30)	(25,47)
8	(19,15)	(23,30)	(25,47)
9	(19,15)	(23,30)	(25,47)
10	(19,15)	(23,30)	(25,47)
11	(19,15)	(23,30)	(25,47)
12	(19,15)	(23,30)	(25,47)
13	(19,15)	(23,30)	(25,47)
14	(19,15)	(23,30)	(25,47)
15	(19,15)	(23,30)	(25,47)
16	(19,20)	(23,30)	(25,47)
17	(19,15)	(23,30)	(25,47)
18	(19,15)	(23,30)	(25,47)
19	(19,15)	(23,30)	(25,47)
20	(19,15)	(23,30)	(25,47)

Figure 20: Output experiencing Premature Convergence

The results show that premature convergence has arisen. The population simply consists of the same repeated chromosome after generation 12. According to Figure 21, the graphical output also indicates that the output is far from the ideal situation. Despite increasing population number, there is no sign of eliminating premature convergence, and the final graphical output shows the result remains far from ideal. This is a common problem with GA. The solution is to provide a mechanism to maintain population diversity. This is considered later in Section 3.11.5.

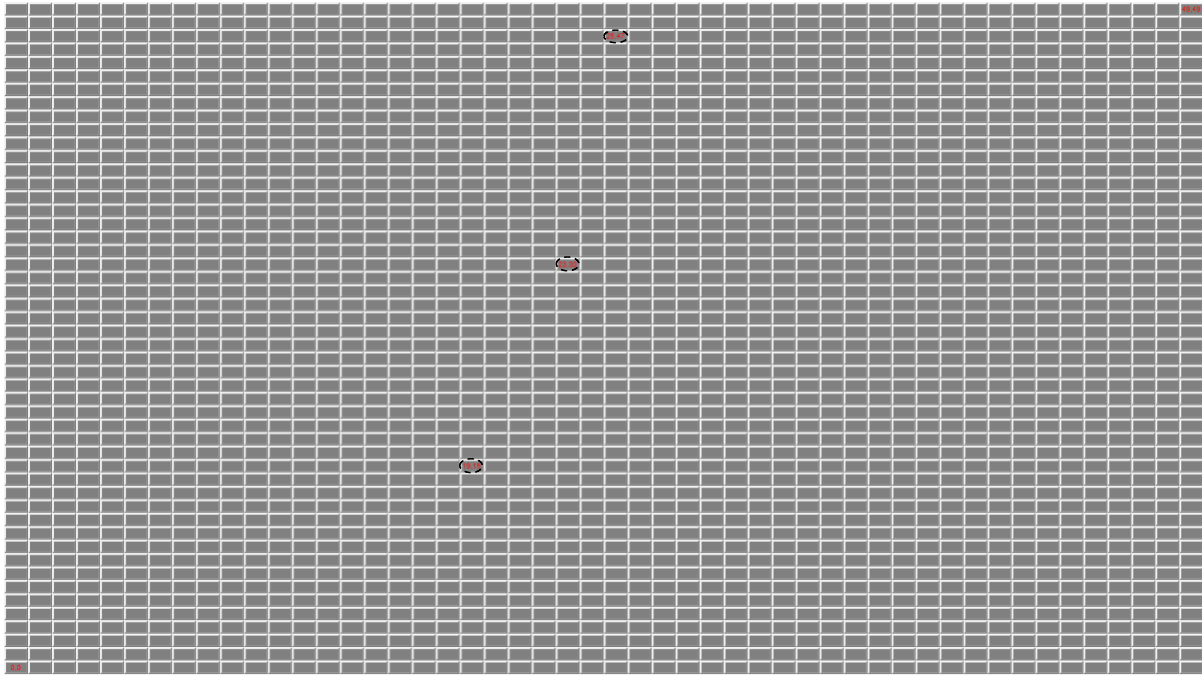


Figure 21: Assessing Increased Population Size

3.11.2 Impact of Generation Count

Another set of experiments increased the number of generations when running the GA to find a (near) optimal result whilst eliminating premature convergence. Increasing the generation number gives GA more generation/time to explore the search space for good results. For this purpose, I set the GA parameters as Crossover = 70%, Mutation = 5% and Population Size = 20. More importantly, I increased GA generation count to 25, 50 and 100 in turn. A sample output at 100 generations is given in Table 5.

Table 5: Increasing the Number of Generations

Generation 100			
Chromosome No	Node 1	Node 2	Node 3
1	(7,21)	(24,28)	(34,39)
2	(7,21)	(24,28)	(34,39)
3	(7,21)	(24,28)	(34,39)
4	(7,21)	(24,28)	(34,39)
5	(7,21)	(24,28)	(34,39)
6	(7,21)	(24,28)	(34,39)
7	(7,21)	(24,28)	(34,39)
8	(7,21)	(24,28)	(34,39)
9	(7,21)	(24,28)	(34,39)
10	(7,21)	(24,28)	(34,39)
11	(7,21)	(24,28)	(34,39)
12	(7,21)	(24,28)	(34,39)
13	(7,21)	(24,28)	(34,39)
14	(7,21)	(24,28)	(34,39)
15	(7,21)	(24,28)	(34,39)
16	(7,21)	(24,28)	(34,39)
17	(7,21)	(24,28)	(34,39)
18	(7,21)	(24,28)	(34,39)
19	(7,21)	(24,28)	(34,39)
20	(7,21)	(24,28)	(34,39)

This change does not provide any improvement as premature convergence still exists. The output remains the same after generation 13. There is no impact changing the generation count subsequently from 25 to 100 because premature convergence has already taken place and the mutation operator is insufficient to reintroduce diversity. On the other hand, the graphical output in Figure 22 shows that the output is a bit closer to the ideal situation due to mutations prior to convergence.

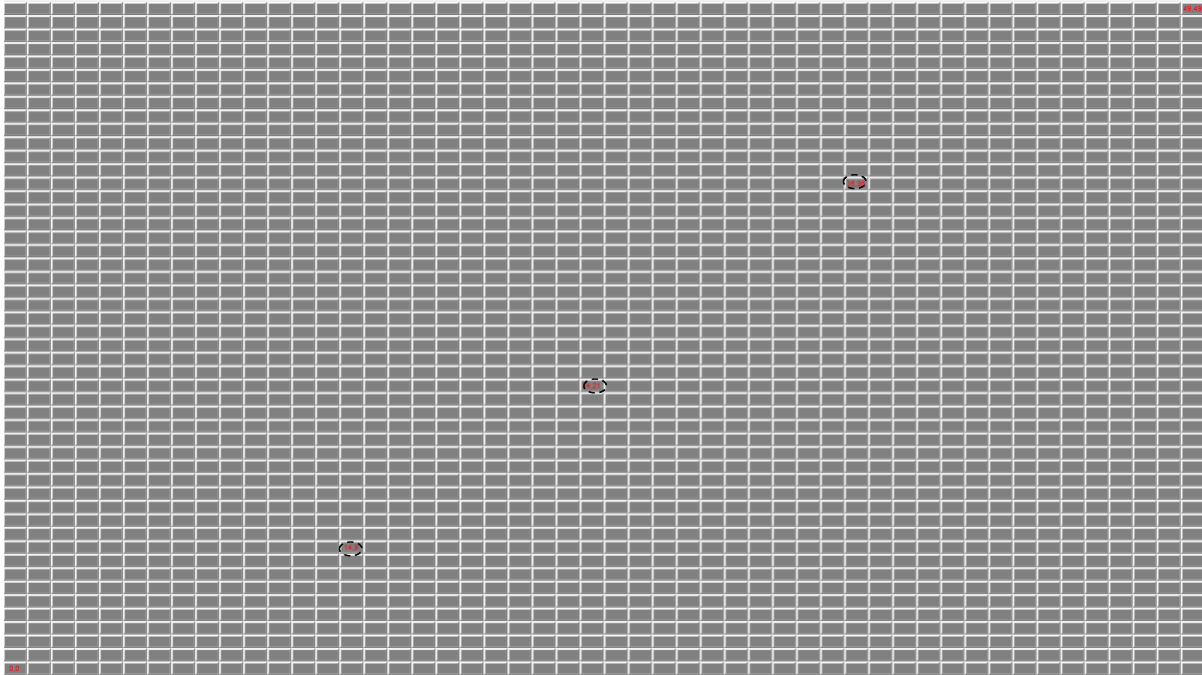


Figure 22: Output with Increased Generation Count

3.11.3 Impact of Mutation Rate

It is essential to allow the GA to explore all of the desired search space to find an optimal solution. This is less likely to happen when the initial population formed for the GA covers a small proportion of the search space; there is a risk of getting trapped in local minima. Therefore, some mechanism is needed to create fresh randomly generated chromosomes or genes within the chromosomes during the iterative search process. I have used mutation to create diversity in the GA that replaces one or more nodes in selected chromosomes. Thus, I initiate the number of Generations = 100, crossover in 70%, Population Size 20, and I increased the mutation rate to 30%. This will increase the diversity of solutions during the mating process. There are a couple of ways to mutate any chromosome. One way is to mutate a single node of a particular chromosome. Another method can be to mutate the whole chromosome by changing all three nodes. The sample output is shown in Table 6.

Table 6: Final Output with Single Gene Mutation

Generation 100			
Chromosome No	Node 1	Node 2	Node 3
1	(14,13)	(22,28)	(37,39)
2	(14,13)	(22,28)	(18,39)
3	(14,13)	(22,28)	(35,31)
4	(14,13)	(22,28)	(18,9)
5	(14,13)	(22,28)	(37,39)
6	(14,13)	(22,28)	(37,39)
7	(14,13)	(22,28)	(24,5)
8	(14,13)	(22,28)	(1,33)
9	(14,13)	(22,28)	(34,38)
10	(14,13)	(22,28)	(37,39)
11	(35,4)	(22,28)	(37,39)
12	(14,13)	(22,28)	(37,39)
13	(14,13)	(22,28)	(37,39)
14	(14,13)	(21,24)	(37,39)
15	(14,13)	(22,28)	(37,39)
16	(14,13)	(7,5)	(37,39)
17	(14,13)	(22,28)	(34,38)
18	(14,13)	(22,28)	(37,39)
19	(14,13)	(22,28)	(37,39)
20	(14,13)	(22,28)	(37,39)

If the GA mutates a single node from 3 nodes, then we see less diversity as premature convergence exists after generation 15. This shows this method is not compelling enough. Therefore, we try another way which is mutating all three nodes without increasing mutation rate and this way put GA more diverse as GA able to search more in our search space. This plan still shows that premature convergence still exists. According to the Genetic Algorithm and output, the repeated output is showing after generation 35. The graphical output in Figure 23 also indicates that the output is closer to the ideal situation. But at the output, even though higher mutation provides extra chromosomes, which covered enough search area. The mutate chromosomes occur only single node rather than a whole chromosome.

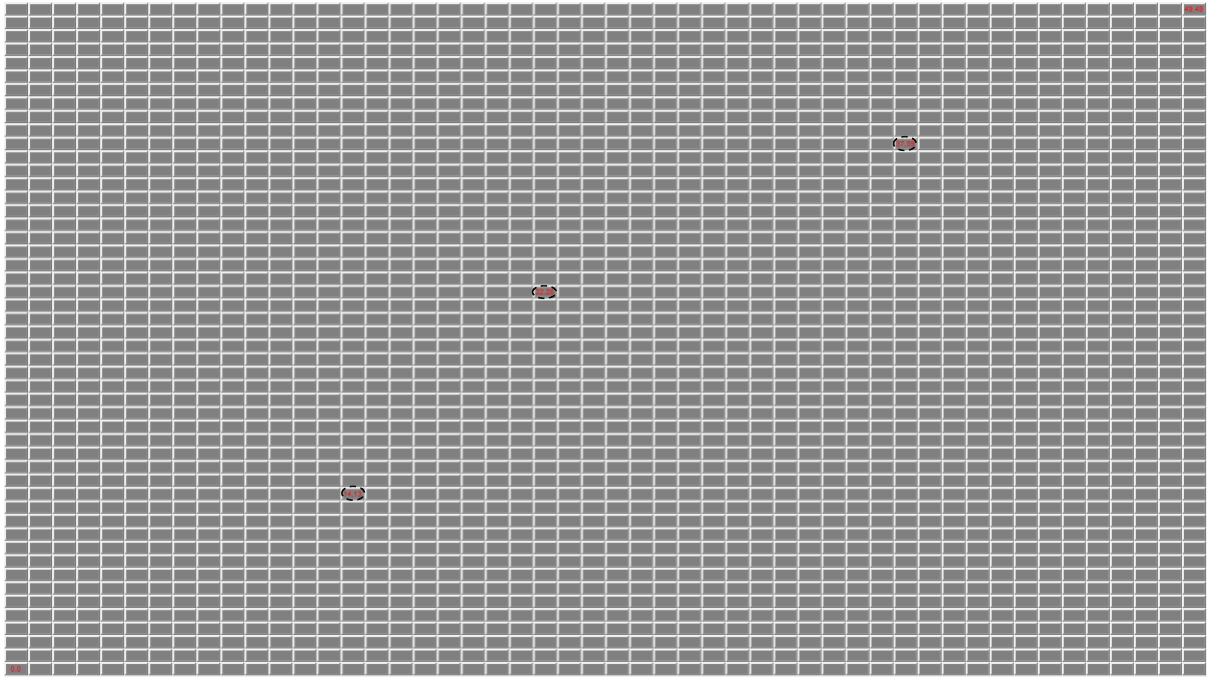


Figure 23: Increased Mutation Scenario

3.11.4 Multipoint (2-Point) Crossover

So far, the tool employed single-point crossover, but it can be implemented as multipoint or two-point crossover. This provides more diversity and therefore, I set the tool as Generation = 100, Crossover = 70%, Mutation = 30%, Population Size = 20 and repeated the experiment. After setting up the tool, premature convergence still exists as the impact of crossover enough to sufficient diversity in the GA.

3.11.5 New Chromosome Insertion

Another method to eliminate premature convergence is the introduction of new randomly generated chromosome(s) into every generation that allows more coverage in the search space and provides more diversity in the chromosome population. I consider the case of a single fresh chromosome per iteration. I assess this method with Generation = 100, Crossover = 70%, Mutation = 30%, Population Size = 20. This method provides a method for maintaining diversity in the GA population. The newly injected chromosomes improve the performance by randomly sampling more of the search space, thus decreasing the likelihood of premature convergence. In this experiment convergence arises after 70 generations but the output has been much improved. The output is near to perfect solution we are looking for. Sample output is shown in Table 7.

Table 7: Output with Chromosome Injection

Chromosome No	Node 1	Node 2	Node 3
1	(10,15)	(26,35)	(47,31)
2	(10,15)	(26,35)	(47,31)
3	(6,22)	(26,35)	(47,31)
4	(10,15)	(26,35)	(47,31)
5	(10,15)	(26,35)	(47,31)
6	(10,15)	(26,35)	(47,31)
7	(10,15)	(26,35)	(47,31)
8	(10,15)	(26,35)	(47,31)
9	(6,22)	(26,35)	(47,31)
10	(10,15)	(26,35)	(47,31)
11	(6,22)	(26,35)	(47,31)
12	(6,22)	(26,35)	(47,31)
13	(10,15)	(26,35)	(47,31)
14	(10,15)	(26,35)	(47,31)
15	(10,15)	(26,35)	(47,31)
16	(6,22)	(26,35)	(4,28)
17	(10,15)	(26,35)	(47,31)
18	(10,15)	(26,35)	(47,31)
19	(10,15)	(26,35)	(47,31)
20	(12,23)	(28,44)	(34,49)

3.12 Discussion

In this chapter a GA-based network planning tool is implemented and evaluated in scenarios with and without obstructions. This network planning tool is able to determine the topological placement of FSO relay nodes in rural areas. Although the tool operates correctly, one limitation is the need to aggregate factors into a weighted sum when calculating the fitness of each potential solution. This means that exploring trade-offs is only possible by adjusting the weighting given to the various fitness parameters. Going forward, I consider an Evolutionary Algorithm. As mentioned in Chapter 2.7, several types of Evolutionary Algorithm exist, including Multi-Objective Evolutionary Algorithms (MOEA). The presence of multiple objectives in a problem gives rise to a set of Pareto-optimal solutions. One example is an elitist non-dominated sorting based MOEA, called Non-Dominated Sorting Genetic Algorithm (NSGA-II) in [102]. Another Pareto based evolutionary algorithm that concentrates on multiple objectives to find an optimal solution is called the Strength Pareto Evolutionary Algorithm (SPEA) [115]. I thus propose an MOEA similar to these examples.

Network planning consists of software applications called network-planning tools. Different planning tools [116, 117, 118, 119] exist spanning a wide range of platforms, systems,

languages etc. In [47], commercial tools are discussed along with their features, before introducing Net2Plan. Net2Plan is an open-source network-planning tool, along with a repository of planning resources, that is not constrained to any specific network technology and is adaptable to any configurable network. Net2Plan is Java-based software, publicly and freely available to download from its website [83]. Although license free and other advantages, Net2Plan is not suitable for our project because Net2Plan is not tailored to FSO-based rural areas considering multiple factors.

Another network planning tool is Mantis [49]. This tool enables the design of an optical network considering a number of constraints, but again the tool does not create any topology. Moreover, similar authors have implemented a tool called DEPLOY [120], where they tried to avoid vendor hardware and software dependencies whilst designing a network planning tool. This tool does not actually create any new topology; instead, that tool uses generic vendor-agnostic models. It enhances the given topology with the parameters such as traffic, link parameters, node parameters etc.

Finally, our network planning tool can offer ISPs an easy-to-use tool where they need to consider a number of parameters when deploying physical nodes in rural areas. The tool has provided some promising results and is shown to function satisfactorily. Applying GA to FSO scenarios has hitherto not been considered. This tool has also addressed a major drawback of GA, which is premature convergence, through the injection of fresh chromosomes throughout the evolutionary process. Furthermore, a number of experiments are conducted to explore the performance of the tool as various design parameters are adjusted.

Chapter 4. Network Planning Tool - Multiple Objectives

As discussed above, part of our research is to design a new network planning tool for the placement of relay nodes in a rural setting. We focus on the deployment of Free-Space Optical technology (FSO), as it is relatively simple and cheap whilst offering high data rates.

4.1 Problem Description

In rural areas, the development of alternatives to traditional copper wire of narrow-band services has been slow and uneven. Mobile telecommunications technologies such as 4G LTE have given customers broadband-like bitrates through mobile devices such as smartphones; however, the cellular model of this technology is designed for densely populated areas and face the same connectivity and capacity gaps as traditional broadband [47, 89]. As our research is based on rural broadband and FSO is used to build the rural network, where data rates range from 100Mbps to 10Gbps. The topological design of communication networks, particularly rural networks, is a typical multi-objective problem involving simultaneous optimization of various factors such as end-to-end speed and diversity of node placement. The design of a network planning tool to determine the suitable placement of network nodes using the above technology and data concerning the terrain topography is shown in Figure 13. However, as this is based on GA, the various performance factors are aggregated to provide a weighted sum of the fitness of each possible solution that is considered.

Due to having multiple constraints in this problem, optimization of one or more of such factors, which makes the network efficient, is the main objective of design in most cases. End-to-end speed and diversity of node placement are two factors that are considered to design this network planning tool. The problem can be stated as: given a set of node locations and the traffic between the nodes, it is required to design the layout of links between the nodes, it is required to design the layout of links between the nodes while optimizing those criteria. There are several situations that can arise while considering both objectives in the tool. These are as follows:

- Increasing end-to-end speed gives high-speed Internet in a rural area, but node placement might be close to each other to achieve high bitrates.
- Spreading nodes over the topography yields better diversity. Considering high bitrate, this might not give a better solution as it might not provide the highest bitrate between nodes due to the distance/range between those nodes.
- Finally, we seek a proper balance between end-to-end speed and diversity while placing the number of nodes in search space. A nominal high-speed data rate should be achieved between nodes while balancing the good distribution of nodes in search space.

Considering the problem description, a classic GA is not possible to work efficiently in this situation as diversity and high end-to-end speed are conflicting requirements. GA works with multiple objectives by specifying a weighting factor between the cost components. However, this problem has various objectives that need to be considered concurrently when making any decision. Therefore, a Multi-Objective Evolutionary Algorithm (MOEA) is required in this

case. MOEA has been described in Chapter 2.7 and for this problem purposes, I have based my tool on the Strength Pareto Evolutionary Algorithm (SPEA-II), which described in Chapter 2.8.

4.2 Flowchart of SPEA-II

Initially, a basic SPEA-II scheme is designed and implemented to work on our objectives. As Line of Sight (LOS) is one of the requirements of FSO, a test for LOS is included in the tool. A step-by-step process diagram of SPEA-II is shown in Figure 24.

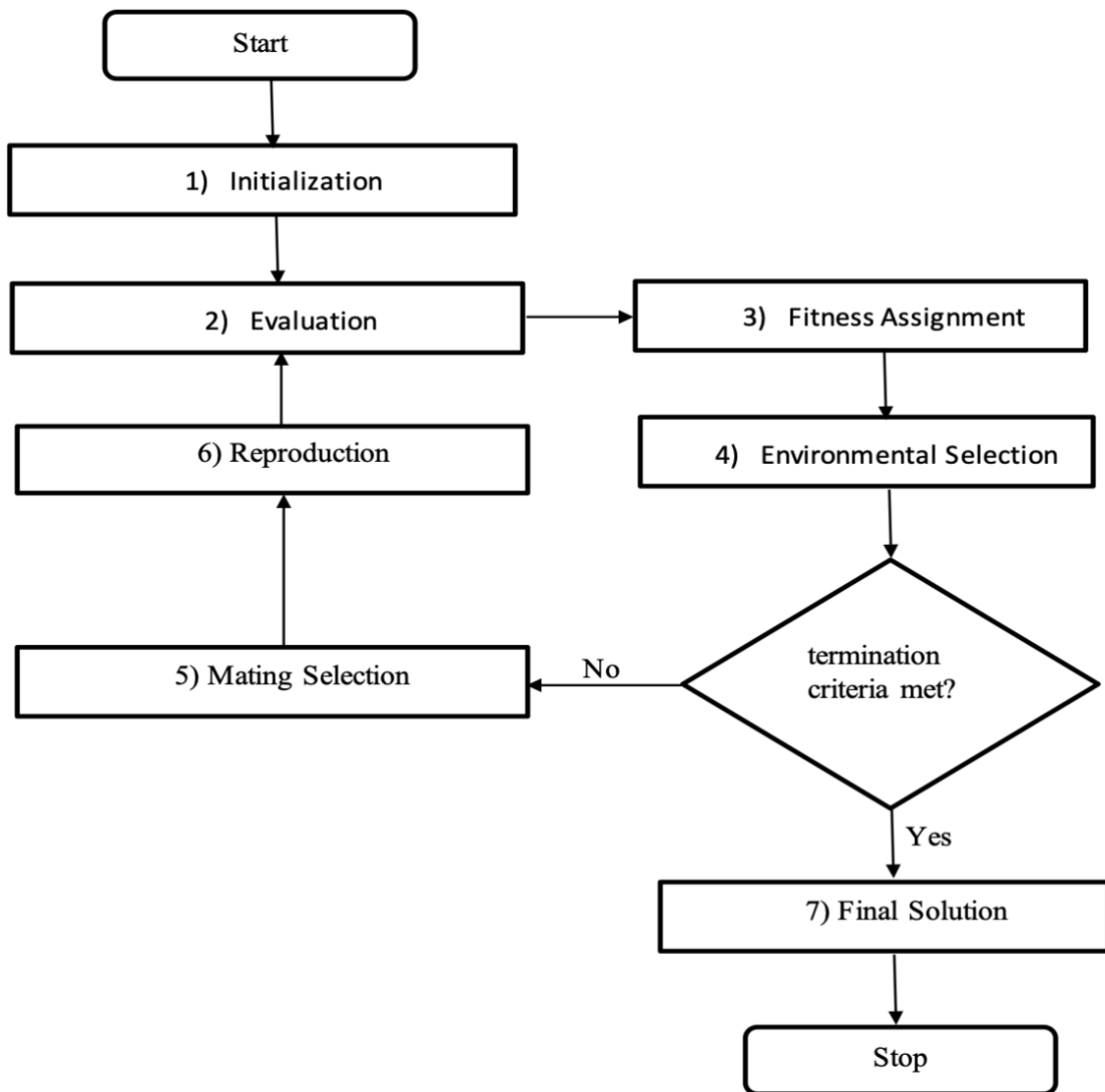


Figure 24: SPEA-II Flow-Chart

The overall process of SPEA-II has been followed, and each step of this algorithm has been designed with consideration of the rural broadband context.

4.2.1 Initialisation

The first step is to initialize the workspace or search space. Initially, a 50×50 grid is created. Initializing chromosomes is the second step; these are randomly by calling a random function called Primary population. For example:

Total Node = 3 [except source and destination, because they are fixed]

Total number of chromosomes, $T_{size} = 20$

So, the representation of chromosomes will be as follows:

Table 8:Chromosome Representation

Chromosome No	Node 1 coordinate	Node 2 coordinate	Node 3 coordinate
1	N_1^1	N_2^1	N_3^1
2	N_1^2	N_2^2	N_3^2
3	.	.	.
4	.	.	.
5	.	.	.
6	.	.	.
7	.	.	.
8	.	.	.
9	.	.	.
10	N_1^{10}	N_2^{10}	N_3^{10}

Above we represent 10 chromosomes with an example fitness rank, with each chromosome comprising three node coordinates, where the node representation N is identified as:

$$N_{Node\ sequence\ number}^{Chromosome\ number\ that\ the\ node\ belongs\ to}$$

The tool initializes the Secondary population as an empty set.

4.2.2 Evaluation

In Generation 0 and once the initial population is generated, all chromosomes go through an evaluation process according to their objectives. In our cases, two objectives have been selected for consideration. These are:

- a. Diversity: “Can see” situation
- b. End-to-End Speed

All chromosomes in the primary population are evaluated and then assigned two objective cost values using a defined mathematical formula.

4.2.2.1 Diversity

A cost value needs to be set for the can-see nodes between the source and destination to assess the fitness of all chromosomes. This calculation is based on a tree-based system. If nodes of a chromosome can reach between the source and destination, the tool starts counting the number of paths that can reach between the source to destination. Two measures are needed to be considered before calculating diversity cost. These are:

- Range
- Distance

The first measure is the range that determines whether two nodes can see each other or not. The communication rate is determined by the range between two nodes. Now, this range

is calculated through Euclidean distance calculation. Initially, I set the maximum FSO range as 25 units.

The second measure is distance calculation between two nodes. This tool stores coordinates of (x, y) value of a node. Therefore, the tool calculates the distance between coordinates of two nodes using the Euclidean distance formula. According to this formula, the distance between two points/nodes in the plane with coordinates (x₁, y₁) and (x₂, y₂) is as follows:

$$\text{Distance } ((x_1, y_1) (x_2, y_2)) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

As an example, the distance between nodes (5, 12) and (19, 34) is found to be

$$\begin{aligned} \text{Distance } ((5, 12) (19, 34)) &= \sqrt{(5 - 19)^2 + (12 - 34)^2} \\ &= \sqrt{(-14)^2 + (-22)^2} \\ &= \sqrt{196 + 484} \\ &= \sqrt{680} \\ &= 26.08 \end{aligned}$$

Once the tool calculates distance, it checks whether the nodes are within range of each other. This allows a communication graph to be created, indicating which nodes “can see” each other. If a single path exists between the source and destination, it would be assigned a diversity value of 1. However, as multiple paths may exist, we use a branching tree search to discover them all and so determine the overall diversity. Consider a simple node distribution as shown in Figure 25.

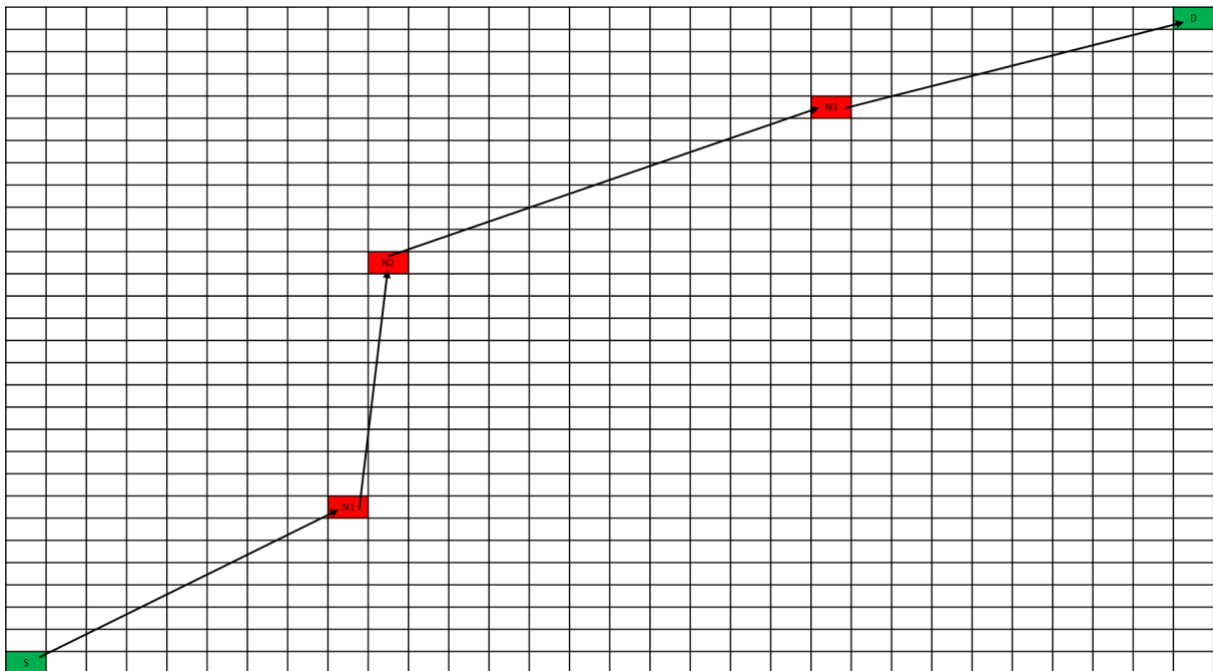


Figure 25: Can-See Representation

We start the diversity search by listing all the nodes, including source and destination, that can see each other. The can-see list goes forward if an individual node can see the next node. One rule that has been set in this process is a that the node “can see” list is only updated with the next node if the node is not already in the path list. We also do not backwards transitions. This prevents the inclusion of loops. According to Figure 25 the can-see list is as below. The nodes in the black path are selected and forwarding the final path. The red nodes are already chosen, and the newly selected node is joined with red nodes.

- | | |
|------------------------|-------------------|
| | <u>Final Path</u> |
| a) S can see: N1 | S→N1 |
| b) N1 can see: N2 & S | S→N1→N2 |
| c) N2 can see: N1 & N3 | S→N1→N2→N3 |
| d) N3 can see: N2 & D | S→N1→N2→N3→D |
| e) D can see: N3 | |

Thus, the Final Path is:

S→N1→N2→N3→D

Now, a more realistic scenario is set up to see whether the tool is able to find and deal with such a situation where an individual node has more than one forwarding nodes. Thus, this makes the can-see list complex and challenging to discover. The complex scenario is shown in Figure 26.

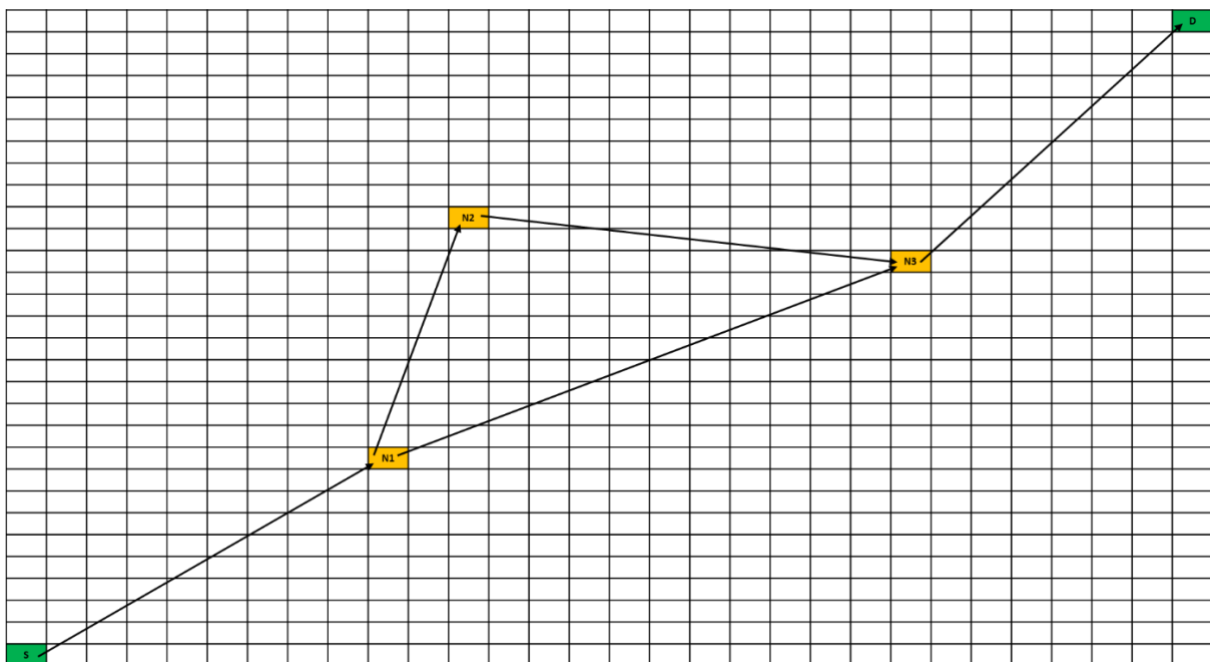


Figure 26: Complex Can-See Representation

A can-see list is created by using Figure 26. This process similar to the previous situation; the only difference is final path can be two paths. For example, (b) shows that node N1 can see N2 and N3. Therefore, N1 can transmit data via two nodes. The can-see situation looks like below:

- a. S can see: N1
- b. N1 can see: S, N2 & N3

Final Path

S→N1

1. S→N1→N2

2. S→N1→N3

- c. N2 can see: N1 & N3

1. S→N1→N2→N3

2. S→N1→N3→N2 (Not valid)

3. S→N1→N3

- d. N3 can see: N2, N1 & D

1. S→N1→N2→N3

2. S→N1→N3→N2

3. S→N1→N3→D

- e. D can see: N3

1. S→N1→N2→N3→D

2. S→N1→N3→N2 (Not possible because no destination)

3. S→N1→N3→D

Final Path:

There are two paths found in this case, as follows:

Path 1:

S→N1→N2→N3→D

Path 2:

S→N1→N3→D

Now, if the above situation, is drawn as a tree, it looks as follows:

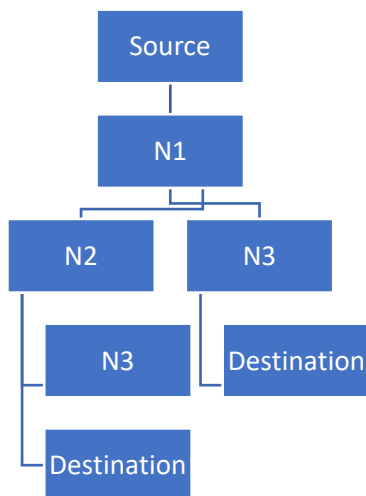


Figure 27: Can-See Tree Formation

As Figure 27 shows, the tool can find two branches in this tree; therefore, the tool gives the diversity value of the chromosome = 2. One thing to note is that none of the pathways allows a node to appear more than once, avoiding loops.

4.2.2.1.1 Flowchart

The process of how diversity works have been outlined in a flowchart shown in Figure 28. The tool calculates the distance between two nodes starting from the source and checks with our given range profile. This process continues until the tool reaches the destination node. Once the process finishes, it counts the number of times the tool reaches from source to destination throughout all given nodes. Finally, it assigns the value as the final fitness value.

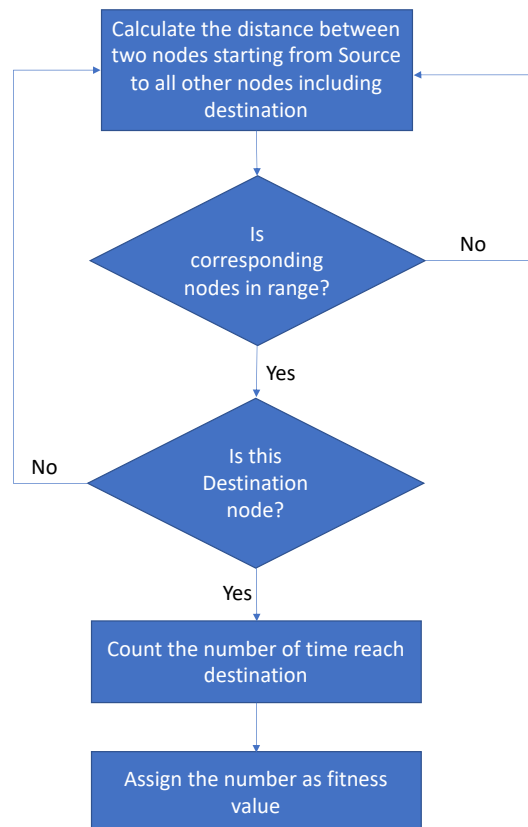


Figure 28: Process of Diversity

4.2.2.2 End-to-End (E2E) Speed

End-to-End (E2E) speed is the actual bitrate of user data in the channel. The tool calculates the E2E speed of each chromosome based on the distance between all the nodes along the path. Firstly, each individual link speed between adjacent nodes along the path is determined. The End-to-End speed is then the minimum of these link speeds. A graph given in Figure 29, shows that when distance increases, speed decreases and vice versa. In this case, tool has linear decay for representing link speed against distance. Although unrealistic, this linear decay profile has been implemented to enable simple calculation of the problem for verification. The calculation is fairly straightforward and provides confidence before moving onto a more complex version of the tool using realistic range characteristics.

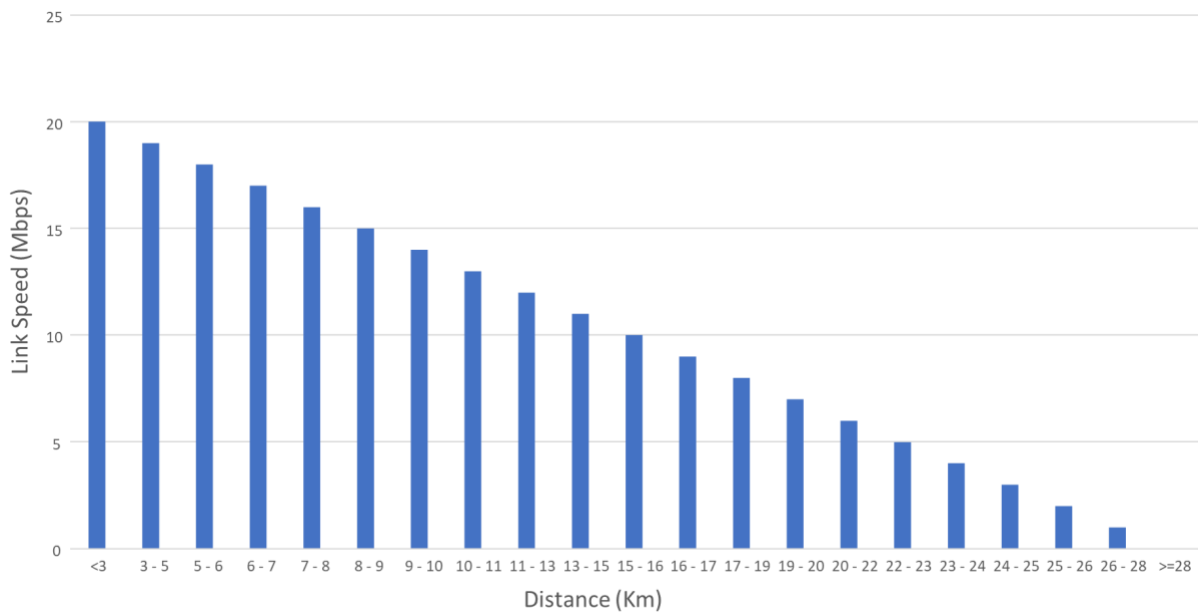


Figure 29: Graphical Representation of Range Profile

4.2.3 Fitness Assignment

Once the evaluation step is completed, the primary population and the secondary population are merged into a temporary population for fitness assignment.

$$P_{temp} = P_{pop} + S_{pop}$$

In all generations, chromosomes go through an evaluation process according to their objectives. In our cases, two objectives have been selected are considered:

- a. Diversity: Can see situation
- b. End-to-end Speed

All chromosomes in the temporary population are evaluated and then assigned two objectives' values. SPEA2 then considers both the number of dominating and dominated solutions of each solution using two methods:

1. Dominance rank
2. Dominance count

4.2.3.1 Dominance Rank

Dominance rank is by how many candidate solutions is an individual dominated and records with a strength value. This strength value is assigned to each solution according to the number of solutions it is dominated by. Example of strength values of candidate solutions shown below:

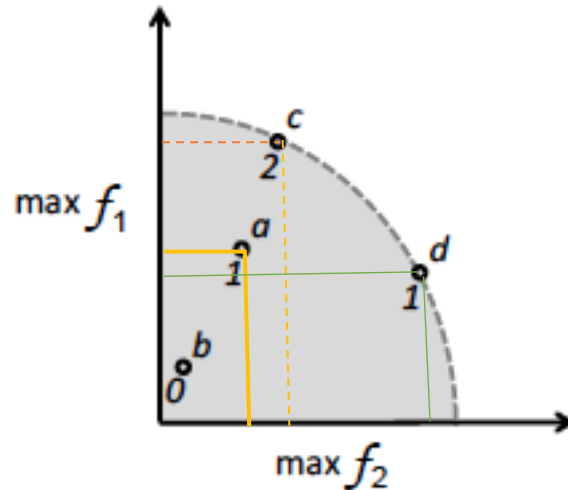


Figure 30: Dominance Rank [121]

An example of strength values is shown in Figure 30. For points a, b, c, and d, their strength values are determined by the number of solutions they dominate. For example, c dominates a and b, so its strength value is 2.

4.2.3.2 Dominance Count

In some cases, dominance rank is not enough to determine the fitness value of solutions. Although both a and d have the same strength value in the above example, d is better than a because there is no solution that can dominate d whilst a is dominated by c.

After obtaining the strength value of solutions, a dominance count method is used for determining the fitness value. The fitness value of a specific solution is calculated by summing up the strength values of its dominators.

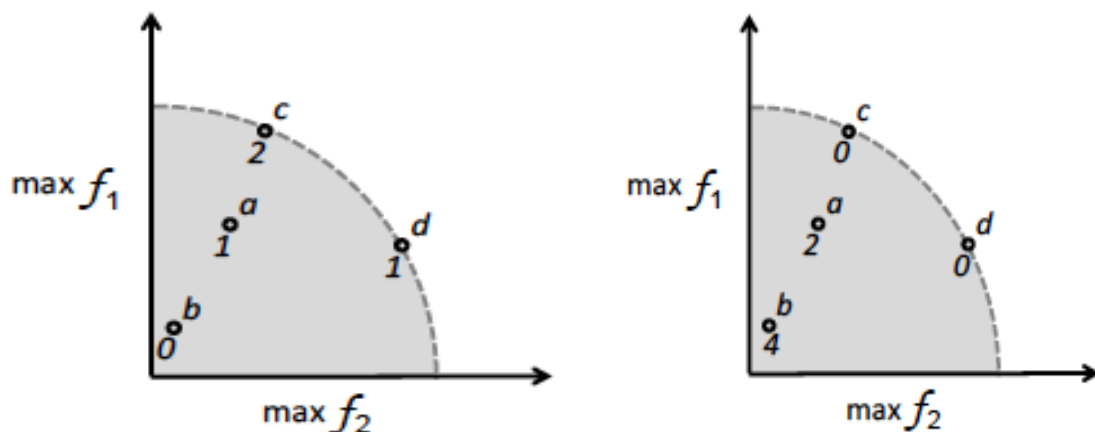


Figure 31: Dominance Count [121]

In Figure 31, b is dominated by a, c and d, and its fitness value is the sum of the strength values of a, c and d. That means that if a node has a larger fitness value, then the solution is dominated by many solutions and conversely, the fitness value of non-dominated solutions is zero. In Figure 31, b is dominated by a, c, d and its fitness value is the sum of strength values of a, c and d.

$$b = c + a + d$$

$$b = 2 + 1 + 1 = 4$$

Hence, it is essential to note that a higher strength value corresponds to the solution being less fit in the SPEA2 fitness assignment.

4.2.4 Environmental Selection

Firstly, all the nondominated solutions in the temporary population are copied into the next generation secondary population. The size of this nondominated solution set is denoted as T_{size} . Secondly, one of three actions are taken depending on the secondary population size S_{size} and T_{size} , namely:

- If Secondary population size $S_{size} = T_{size}$, the secondary population update completes.
- If $S_{size} < T_{size}$, the remaining places in the secondary population are occupied by the best-dominated solutions from the temporary population, as required.
- If $S_{size} > T_{size}$, The Crowding Distance mechanism is used to prune the secondary population

4.2.5 Mating Selection

After the secondary population update completes in Step 4, mating selection is performed to determine which solutions are chosen from the secondary population to form a mating pool to reproduce offspring. Two solutions are randomly selected from the secondary population. Their fitness is compared and the better one is copied to a space in the mating pool. If both solutions have the same fitness value, one of them is randomly selected. The process repeats until the mating pool is full.

4.2.6 Reproduction

In the Reproduction process, offspring are created by two operations:

1. Crossover
2. Mutation

Before reproduction, the primary population is emptied, ready for holding offspring.

4.2.6.1 Crossover

In the crossover process, the position of the crossover point is firstly randomly selected. Then, the two parent chromosomes are broken into two parts according to the crossover point. Finally, the first part of parent 1 and the second part of parent 2 are combined to form the first offspring. By symmetry, the second offspring is generated by combining the second part of parent 1 and the first part of parent 2. This process is called single or one-point crossover [122].

The first step is a crossover with all those chromosomes. The tool typically uses a probability of 70% or 0.7 as the crossover rate for all mating chromosomes. This probability rate is applied stochastically. Once the tool places all selected chromosomes in the mating pool, we invoke the crossover process, as appropriate, to create new children or chromosomes for the next generation. For example:

Chromosome: 2	Ranking: 9	N_1^2	N_2^2	N_3^2
---------------	------------	---------	---------	---------

+

Chromosome: 4	Ranking :7	N_1^4	N_2^4	N_3^4
---------------	------------	---------	---------	---------

Chromosome 2 & 4 together create a new child that can be called Chromosome 1 (this number for the next generation)

Chromosome: 1	N_1^4	N_2^2	N_3^4
---------------	---------	---------	---------

The above shows that crossover happens in this instance by taking one node from Chromosome 2 and 2 nodes from Chromosome 4. It can be another way round or vice versa. Chromosome selection has to be random, and crossover should be random as well. This crossover process will create new child chromosomes, which will go to the following procedure. By default, this is used to form 19 new chromosomes.

4.2.6.2 Mutation

Before the algorithm places all the offspring chromosomes into the next generation, one last process to consider is mutation. Mutation is used to maintain genetic diversity from one generation to another by randomly altering some gene values. This mechanism is applied stochastically as with crossover. A uniform random variable is firstly generated in the interval [0,1]. Mutation has a number rate, typically 50% or 0.5, to decide to change or create a new node of a chromosome. The mutation rate is set as very high number due to avoid premature convergence where it is ensured that the solutions has been searches throughout the given search space. From the above example: if we mutate the chromosome, we see that Node 3 has been updated.

Chromosome: 1	N_1^4	N_2^2	(6, 12)
---------------	---------	---------	---------

4.2.7 Final Solution

If the maximum number of generations is reached, the simulation stops. Then, the chromosomes of the secondary population in the final generation are the outcome. Again, there is no single solution, and final solutions can be found from secondary population of the final generation. Therefore, the secondary population is sorted by the fitness value of chromosomes. After sorting those stays on the top of the secondary population, they are non-dominated solutions. The non-dominated solutions are the final solutions of the simulation.

4.3 Simple Flowchart of SPEA-II

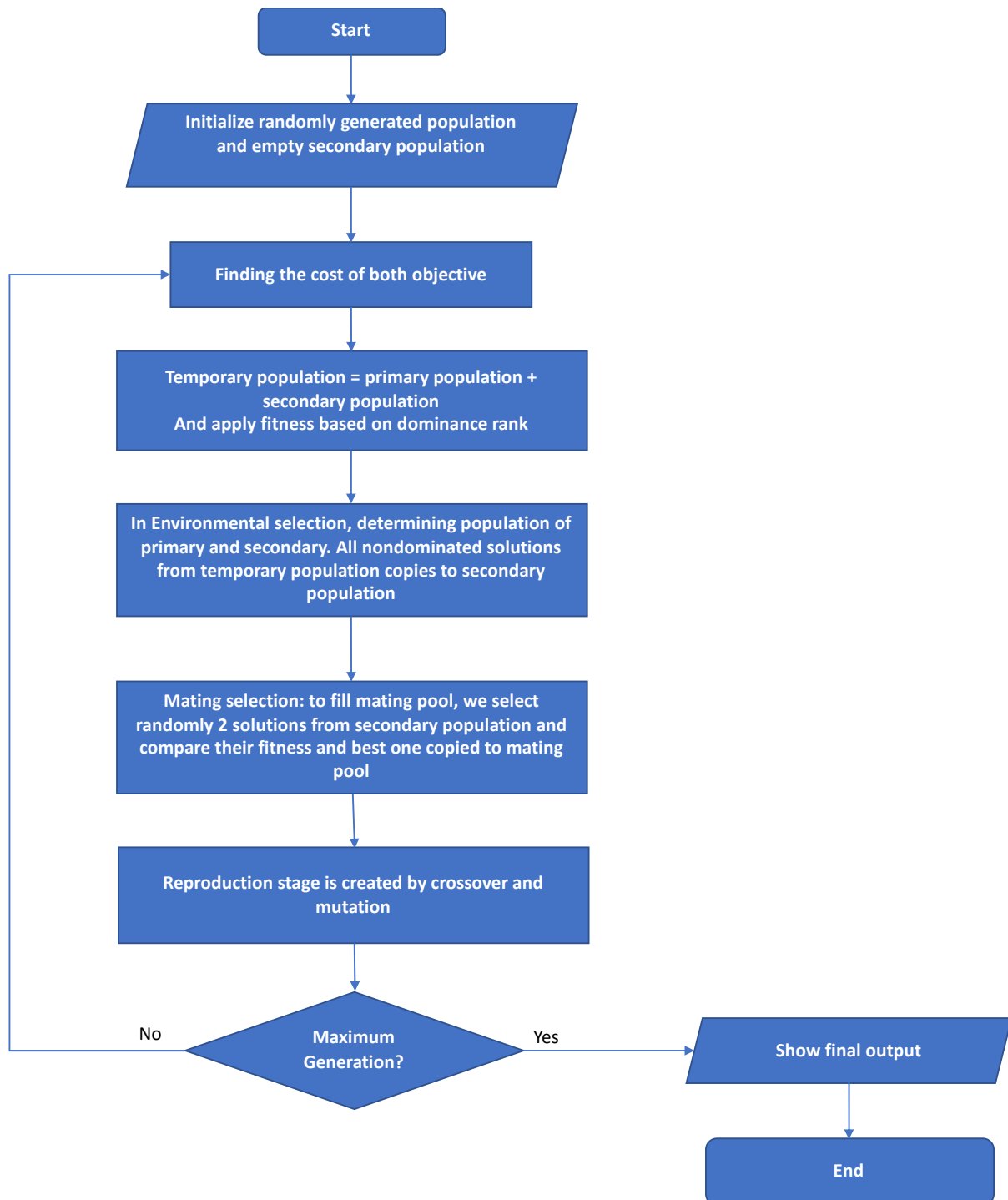


Figure 32: Overall SPEA-II Flowchart

4.4 Hybrid Multi-Objective Evolutionary Algorithm

Once the basic SPEA-II scheme has been designed and implemented under our constraints/objectives, there are a few changes needed to be made to ensure improved results can be achieved. These methods apply to different steps of SPEA-II and bring about a hybrid-MOEA. These are described as follows:

4.4.1 Line of Sight

As this tool is designed for rural areas with different geographical features such as hills, and FSO technology needs Line of Sight (LOS) between adjacent relay nodes, then it is essential to incorporate a LOS test within the tool. An equation has been built to determine whether nodes are in LOS or not. Below is the description of this equation:

Let, there be three-points: A (x_1, y_1), B (x_2, y_2) and C (x_3, y_3). Consider, B point is our blockage point. If A, B and C in the same line, it means A and C can't see each other, and B is an obstruction between them. To make the above statement true, the tool calculates the slope of AB, BC and AC.

Also, that slope equation is:

$$\text{The slope of any two-point} = \frac{y_2 - y_1}{x_2 - x_1}$$

For example, to start this, the tool needs to calculate the point-slope of any defined 3 points. In this 3-point arrangement, one point will be the blockage point. For example: A (13,8) B (23,23) C (24,24).

First, calculate the slope of the first two points, A (13,8) and B (23,23).

$$AB = \frac{y_2 - y_1}{x_2 - x_1}$$

$$AB = \frac{23 - 8}{23 - 13}$$

$$AB = \frac{15}{10}$$

$$AB = \frac{3}{2}$$

AB will be calculated as an integer variable which will eliminate the fraction value of the output. The fraction of the output has been truncated for remove the marginal error while calculating the slope. The direction of the slope is complicated due to the topological space has been designed with number of rectangular boxes. Therefore, there is some marginal error exist. Thus,

$$AB = 1$$

BC and AC are calculated using the same equation as above; Thus, calculating the slope of the second two points B (23,23) and C (24,24), we have:

$$BC = \frac{y_2 - y_1}{x_2 - x_1}$$

$$BC = \frac{24 - 23}{24 - 23}$$

$$BC = \frac{1}{1}$$

$$BC = 1$$

and...

$$AC = \frac{y_2 - y_1}{x_2 - x_1}$$

$$AC = \frac{24 - 8}{24 - 13}$$

$$AC = \frac{16}{11}$$

$$AC = 1$$

This means these points are colinear, and if point B (23,23) is a blockage point, then the remaining 2 points aren't in Line of Sight (LOS).

This way the tool checks if there are obstructions between nodes. If a blockage point exists between two nodes, then the search along the intersection line between the nodes should cease. Then tool suspend the search between those two nodes but carry on searching where LOS is established.

4.4.1.1 Implementation of LOS

LOS has been implemented as described in Chapter 4.2.2. This plays a significant role when creating the initial population in the tool. LOS is the first thing the tool checks after populating each chromosome. If a chromosome fails the LOS test, then the entire chromosome will be deleted and regenerated again. This way the tool makes sure no chromosome passes to the next step if the LOS evaluation fails.

4.4.2 Dijkstra's Shortest Path Algorithm

As this tool is designed to provide high-speed Internet to rural people, there is a relationship between distance and speed (bitrate). The minimum distance will provide the highest bitrate along a link, and therefore, a minimum distance route between source and destination is preferred. Thus, Dijkstra's Algorithm has been introduced to find the shortest distance route from the source to the destination to achieve the highest speed (bitrate). Dijkstra's Algorithm determines the shortest path from one node to every other node within the same graph data structure. This is necessary to determine best possible path with highest bitrates within same graph data structure. The algorithm runs until all vertices in the graph have been visited. This means that the shortest path between any two nodes can be saved and referred to as needed.

4.4.2.1 Implementation of Dijkstra's Algorithm

The implemented Dijkstra's Algorithm is invoked after the secondary population is created, which is during Initialization and after Environmental Selection according to Figure 24. Dijkstra's Algorithm is used to find the chromosome representing the shortest route, and this is tagged as our Elite Chromosome. The current Elite Chromosome is transferred to the primary population as the first chromosome in each subsequent iteration.

There is a logical constraint when applying Dijkstra's Algorithm. Usually, Dijkstra's Algorithm finds the shortest path using a least cost mechanism. In this tool, Dijkstra's Algorithm needs to find the shortest path between source and destination nodes but with the

highest End-to-End speed. For this reason, a unique equation has been created to transform all the highest E2E speeds into smaller values. This is a cost inverting process so that smaller cost is better. The new value is used with Dijkstra's Algorithm as described in Chapter 2.10. We made all the highest value into the smallest value, but every chromosome gets the same credit as its original value. A graphical representation is shown in Figure 33.

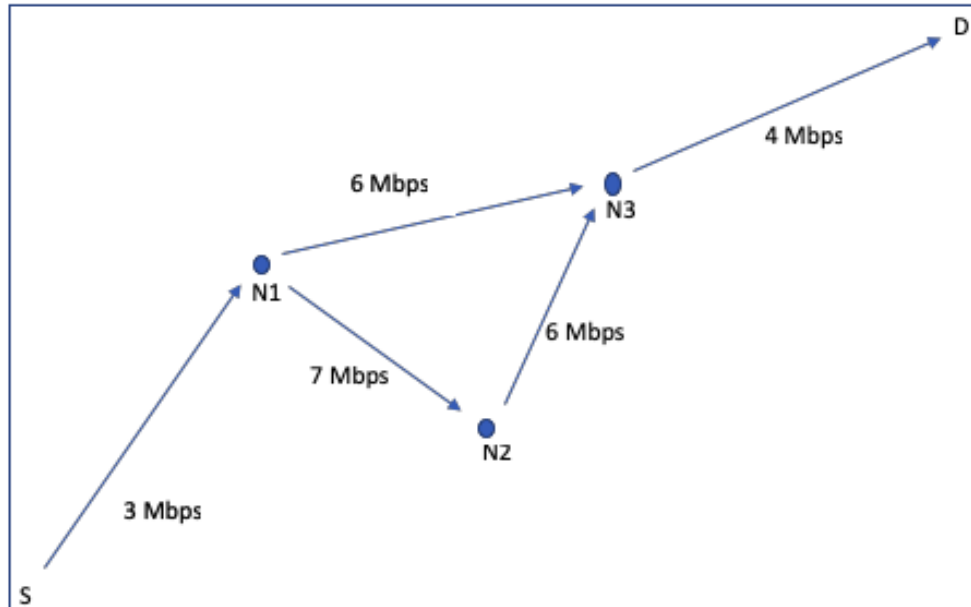


Figure 33: Dijkstra's Algorithm Representation

At first, the tool finds the fastest End-to-End speed of all the chromosomes, and then the equation α is used to find the corresponding value by using that fastest E2E Speed and All link speeds between adjacent nodes. This new cost is stored by calling a new variable α to develop this equation for Dijkstra's Algorithm.

$$\alpha = \frac{\textit{Fastest speed}}{\textit{Link speed between two nodes}}$$

4.4.2.2 Flow Chart of the Dijkstra's Algorithm Process

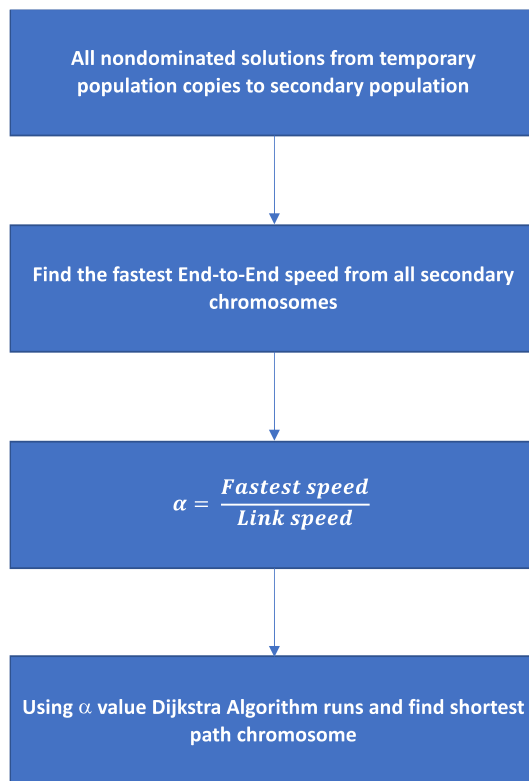


Figure 34: Flowchart of Modified Dijkstra's Algorithm

4.4.3 NSGA Crowding Distance

Multi-objective optimization aims to produce a Pareto front, which is the set of non-dominated solutions for problems having more than one objective. A solution is said to be non-dominated if it is difficult to improve one component of the solution without worsening the value of at least one other component of the solution. The goals of multi-objective optimization are:

- To guide the search toward the true Pareto front (non-dominated solutions) or approximate the Pareto optimal set, and
- To generate a well-distributed Pareto front

So, it is essential to maintain the diversity of solutions of the Pareto front. It is also important to maintain extreme solutions of the Pareto front. So, it is necessary to improve the diversity of solutions and preserve extreme solutions on the Pareto front in any MOEA. From all features mentioned above, NSGA-II is one of the most popular MOEAs.

In our tool the process of evaluating crowding distance is similar to NSGA as it also works with non-dominated solutions. We do not implement NSGA, but we use aspects of the crowding distance mechanism to increase and maintain a diverse population distribution along the Pareto front. A full description of NSGA and crowding distance explained in Chapter 2.11. In our tool, the original crowding distance approach is not used, but we employ a mechanism inspired by it. We check whether any chromosome along the Pareto front has any coincident

neighbours. If so, the tool deletes that chromosome and regenerates a new randomly generated chromosome.

4.4.3.1 Flow Chart of NSGA’s Crowding Distance

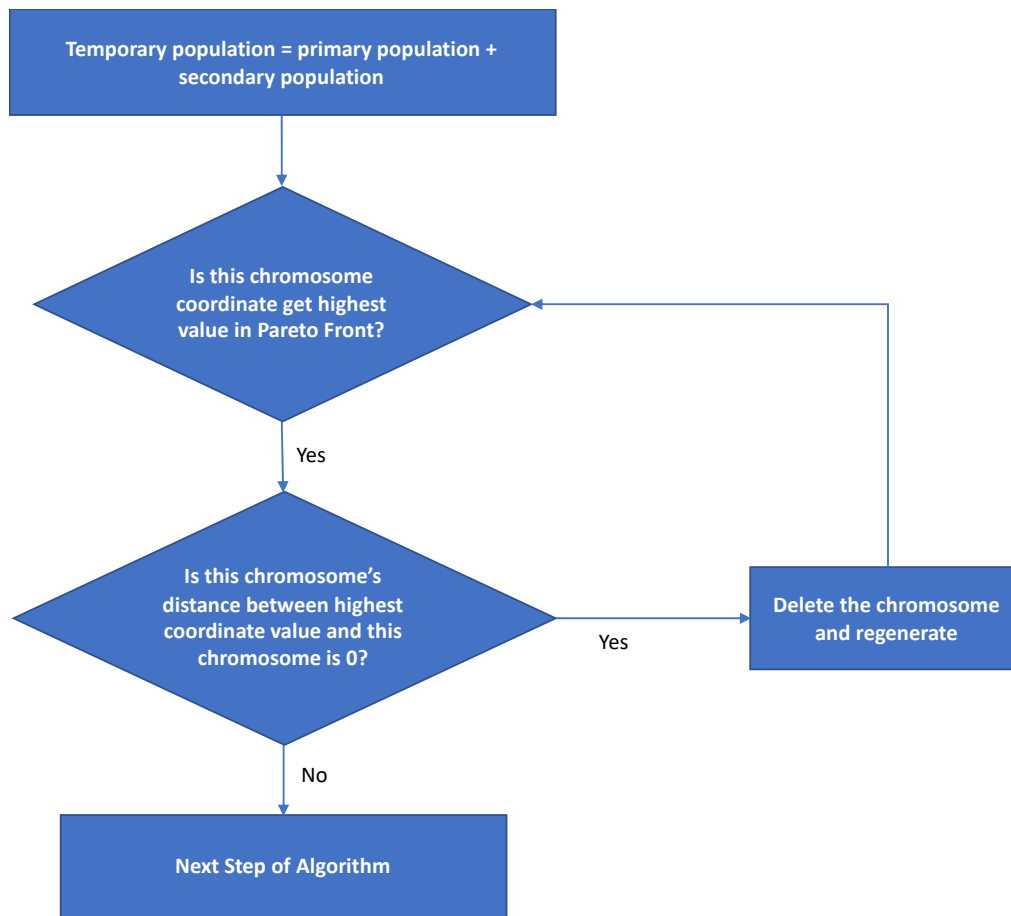


Figure 35: Crowding Distance (CD) Process Flow

4.4.3.2 Implementation of CD

This process, described in Chapter 4.2.3, ensures good distribution of solutions along the Pareto Front. First, the tool determines whether a chromosome’s objective function coordinates are at the highest point of the Pareto Set search space against the x-axis value. Then the tool collects all the x-axis values of the Pareto Set and checks the chromosome against the highest y-axis value of the Pareto Set. The tool deletes those close to the highest y-axis value, typically the exact coordinates according to our current tool settings. This way tool preserves the best non-dominated solution in PF and provides better distribution in PS. So, the tool checks all x-axis and preserve/upgrade all best non-dominated solution.

4.4.4 Final Hybrid Algorithm Flow-Chart

The final hybrid algorithm outlined with a flowchart in Figure 36. The blue part of the flowchart is the original SPEA implemented under the tool, where the green part is the hybrid part. The hybrid part has been added to address the context that needed for implementing topology under rural context.

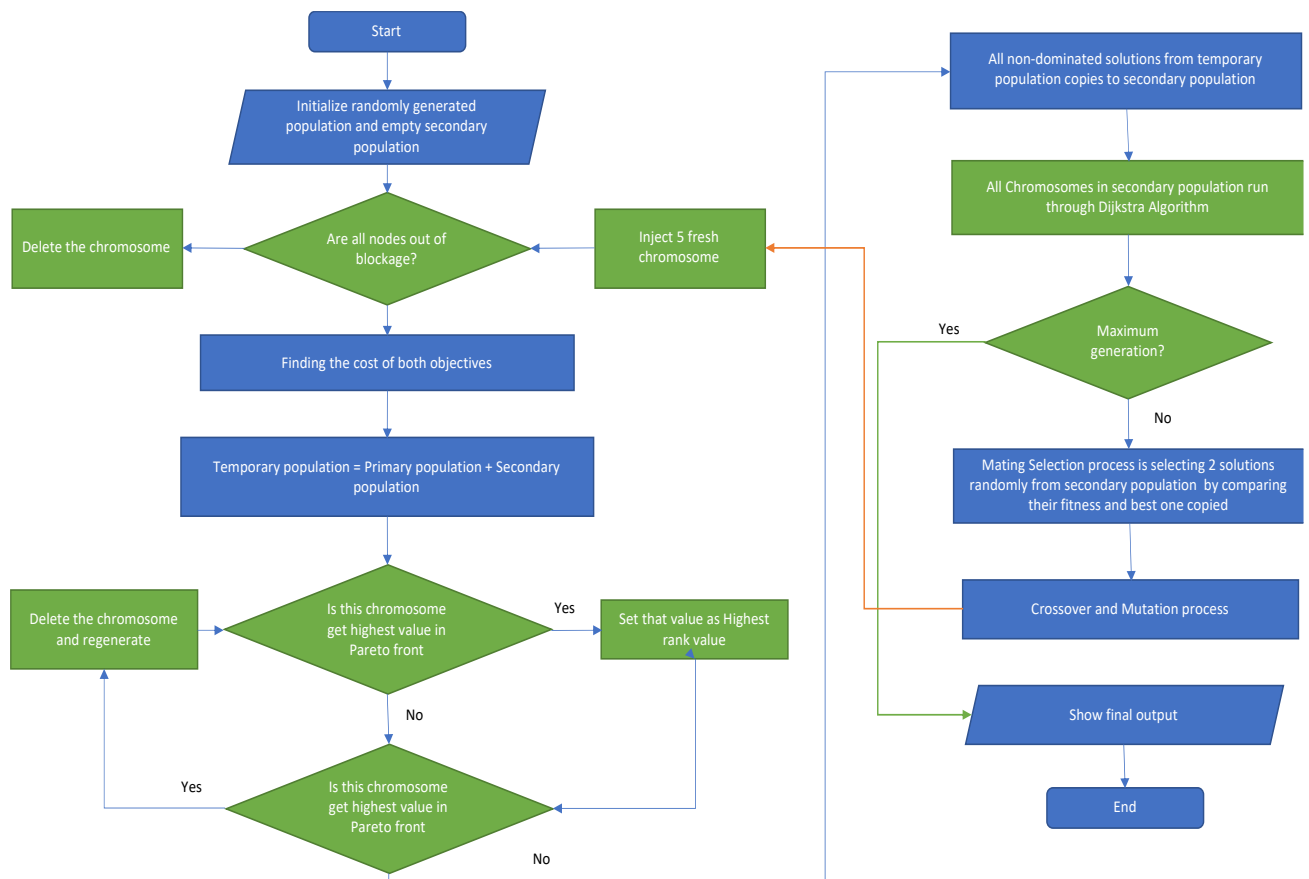


Figure 36: Final Algorithm Flowchart

The above flow-chart employs two colours that represent the following:

- Original Strength Pareto Evolutionary Algorithm (SPEA-II)
- Hybrid Section of SPEA-II Algorithm

4.5 Final Output

All chromosomes go through a process called crossover to generate new individuals. There are several techniques of crossover, but single-point crossover has been used. Once the crossover process finishes, mutation is applied to preserve diversity in the population. This operator is applied to each chromosome in the population with a predefined probability. After the reproduction stage, all chromosomes go through possible crossover and mutation until the maximum number of generations is reached, thus terminating the simulation. Due to considering multiple constraints, there are several outputs as it depends on the balance between objectives. The non-dominated chromosomes of the Pareto Front in the last generation are the final solution, as illustrated in Figure 37, although some dominated solutions may be present in the secondary population representing other compromises between the conflicting objectives.

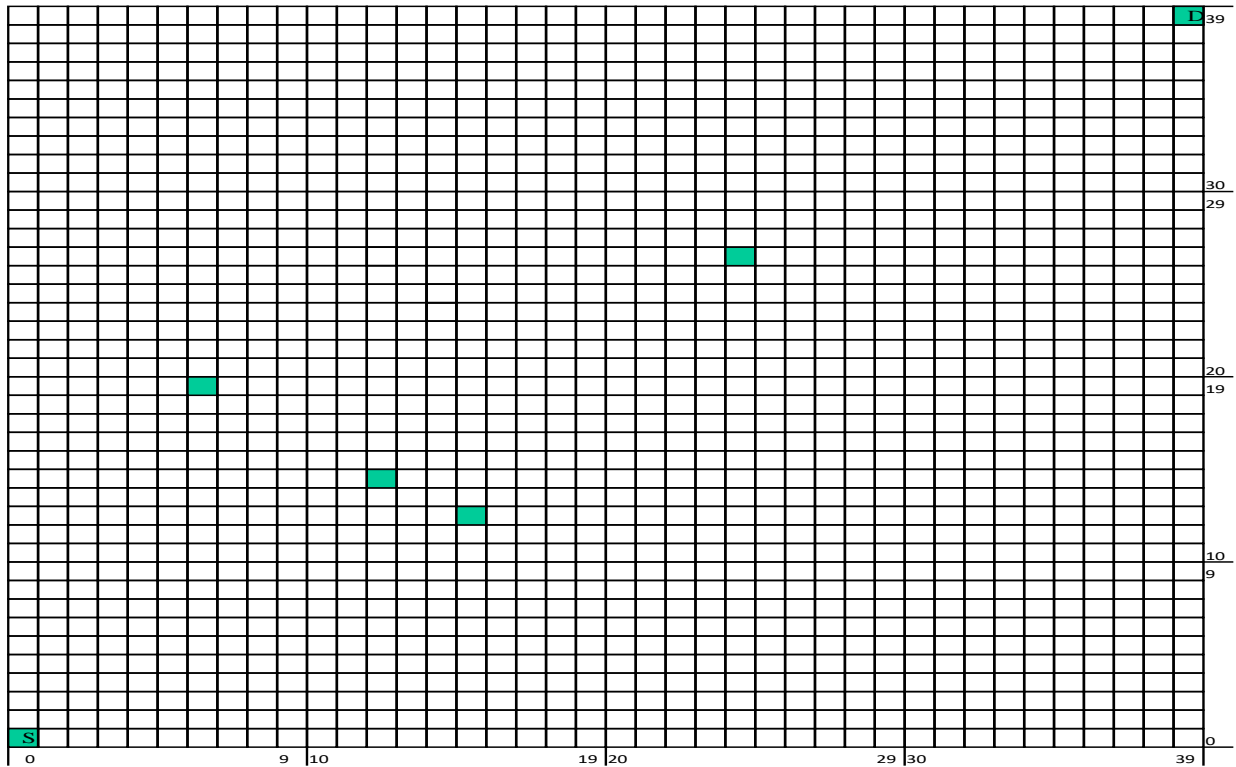


Figure 37: Typical Final Output

4.6 Verification of the Network Planning Tool

4.6.1 Experimental Verification of the Tool

We have conducted a number of experiments that verify that the Hybrid MOEA is operating appropriately. These experiments actually provide topological solutions for a rural community, taking into account the end-to-end speed and the degree of path diversity. The tool currently takes as inputs the number of relay nodes to be placed between two fixed points and a transmission range profile. The range profile is such that the further apart are the two nodes that comprise a point-to-point link, the lower will be the transmission speed/bitrate of user data. Beyond a specific range, no communication is possible.

4.6.1.1 Initial Test to Check Fitness

The algorithm we designed aims to find a suitable topological node placement for rural broadband, trading off two objectives, which are End-to-End speed and Diversity. There is no single solution to this problem; the “best” result depends on the relative importance of these two objectives to the network operator.

A Monte-Carlo method is introduced that uses repeated random sampling on search-space to find the best compromise solutions between both objectives to test whether the algorithm we developed can identify suitable node placements. Based on our objectives, the best outputs are shown in Figure 38. Basically, 1000 randomly chromosomes are generated where each chromosome contains three nodes fed through our objectives to observe whether our hybrid MOEA algorithm described in 4.4.4, can determine and store the best outcome. Using the output entitled ‘End-to-End Speed’ we see that the algorithm was able to find a

solution with good end-to-end speed, but it has compromised diversity. Equally, the output entitled ‘Diversity’ shows that relay nodes obtain better diversity, which means more paths exist between the source and destination by sacrificing end-to-end speed. The output entitled ‘Both’ shows a reasonable compromise between End-to-End speed and Diversity. The purpose of this experiment is to find whether the algorithm is able to find and store the best-compromise outcomes from the search space. In Figure 38, the red dots are source and destination, and blue dots represent relay nodes in the search space.

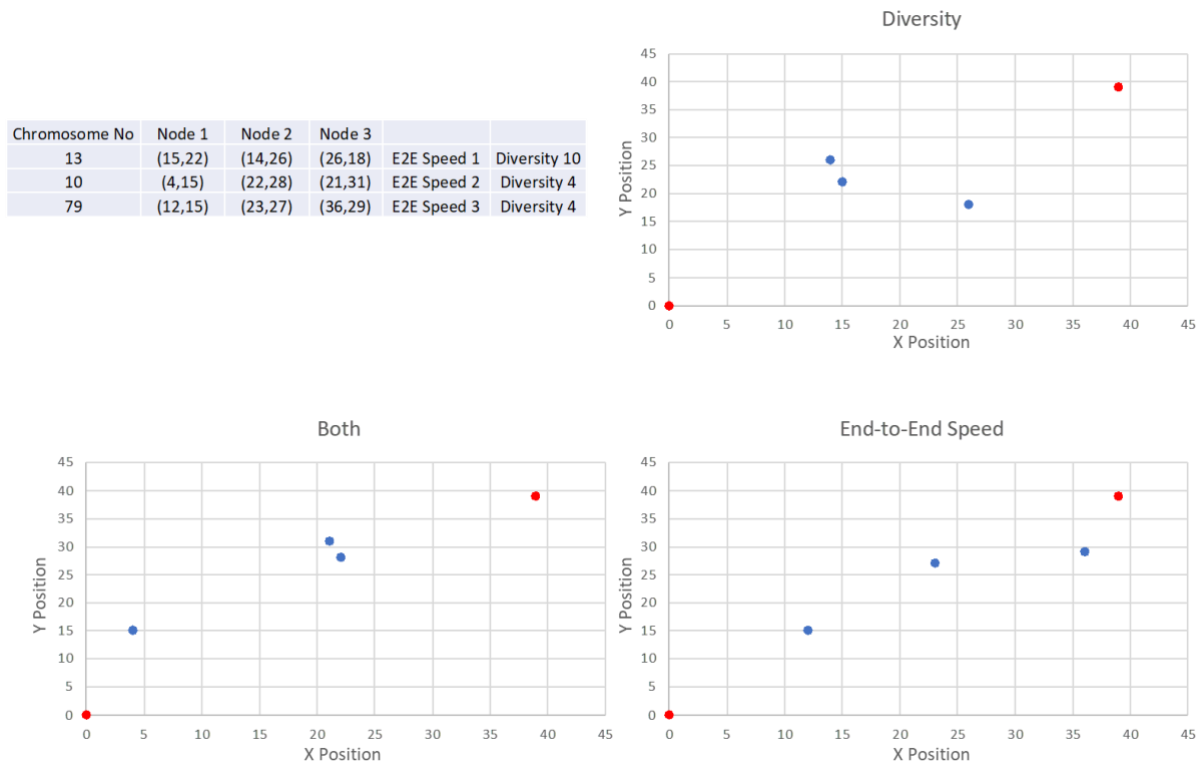


Figure 38: Initial Fitness Test

Another important factor is the range profile which is needed to calculate both objectives: End-to-End Speed and Diversity. The range profile is calculated speed over distance. Due to the discretised nature of the range profile, placing a node anywhere within a given region, such as the box shown in Figure 39, results in no change in the fitness score. This means placing a node anywhere inside the box receives the same score². An example of the discretised range profile is shown in Figure 40. It roughly follows a negative exponential characteristic with a finite tail.

² The benefit of using a discrete function is that it is easier to perform hand-calculations to verify the results. In addition, it may also prove useful in due course, by allowing the installation engineers some scope for adjusting the specific location of each relay node.

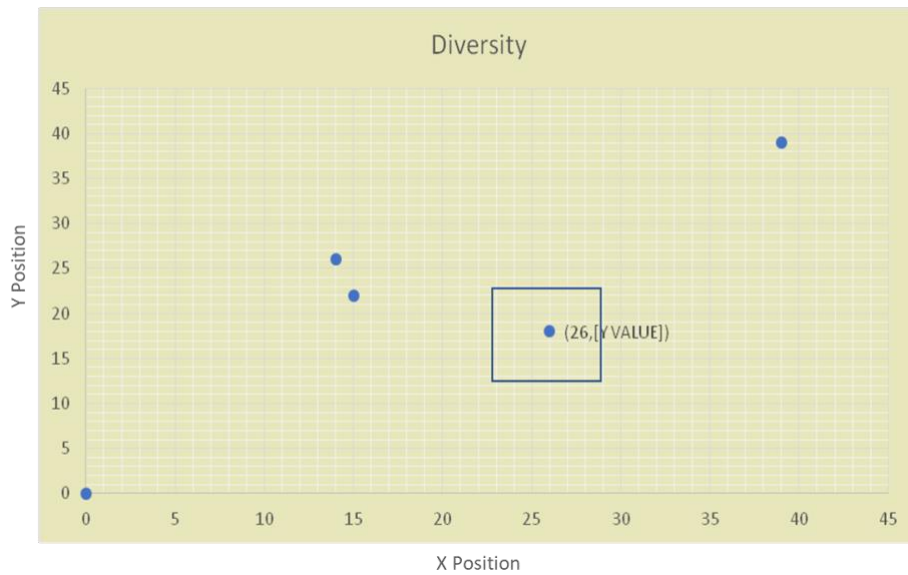


Figure 39: Node Placement Experiment

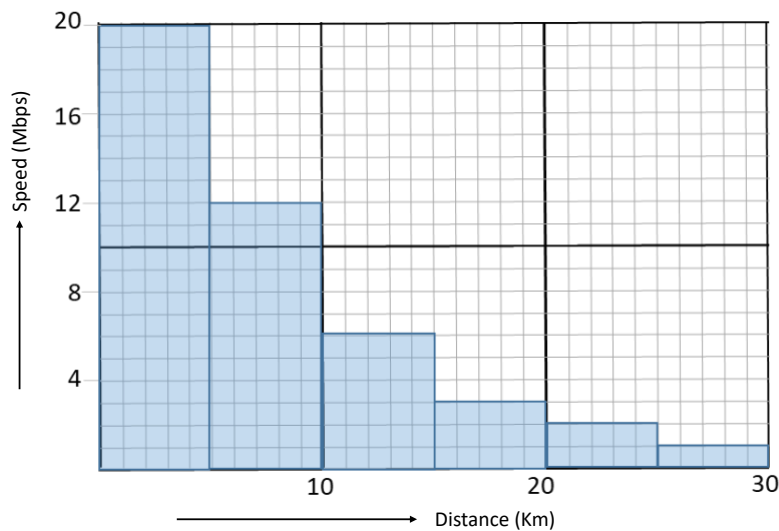


Figure 40: Range Profile for Experiment

4.6.1.2 Relay Node Placement Without Obstacles

This experiment illustrates the placement of three and four relay nodes after 500 generations, as shown in Figure 41 and Figure 42, respectively. The algorithm's Pareto Front forms as we expect. One of the purposes of the experiment is to confirm the flexibility of the tool for finding alternative solutions for node placement based on different compromise weightings of the two objectives. This experiment has been conducted with both 3 and 4 relay nodes in the environment shown with two different range profiles. The graph entitled 'End-to-End Speed' indicates that the algorithm was able to find a solution with good end-to-end speed but compromised diversity. Conversely, the graph entitled 'Diversity' shows that relay nodes obtain better diversity, which means more paths exist between the source and destination. However, this is achieved by sacrificing end-to-end speed. The graph entitled 'Both' shows a reasonable compromise between End-to-End speed and Diversity. There are some dominated

solutions that shows that tool has considered all possible solutions despite they are dominated by nondominated solutions in PF. But these solutions are also regarding as alternative solutions.

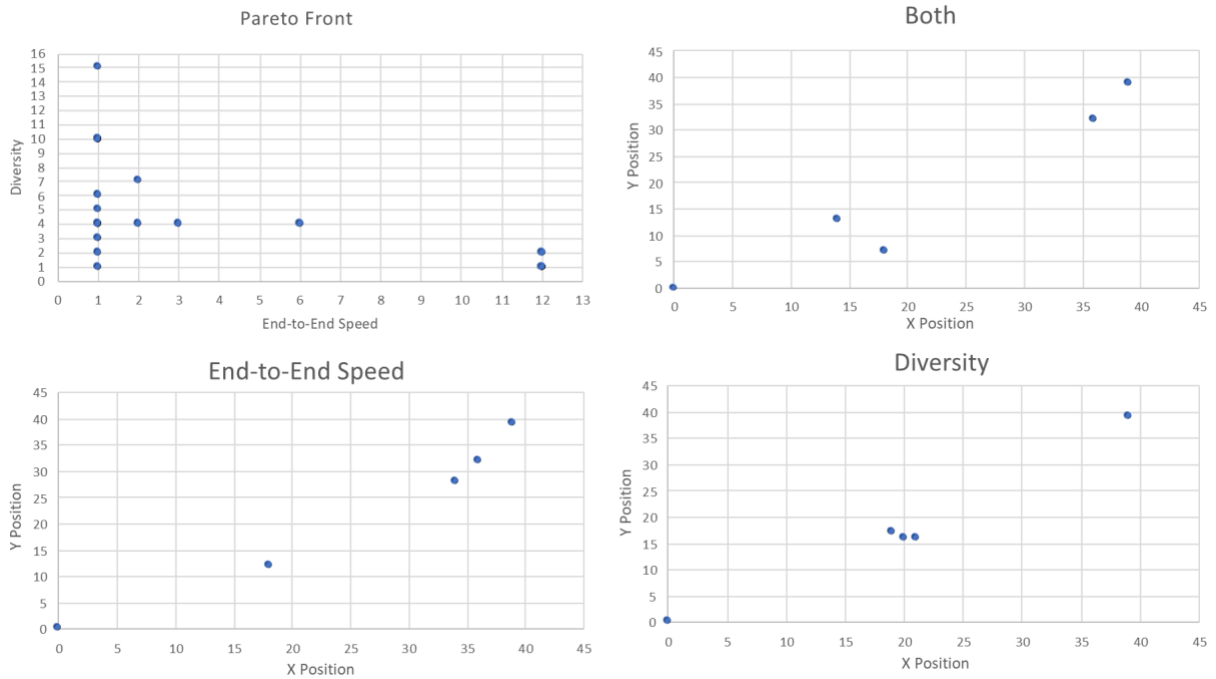


Figure 41: Example 3 Relay-Node Placements for Differing Constraint Preferences

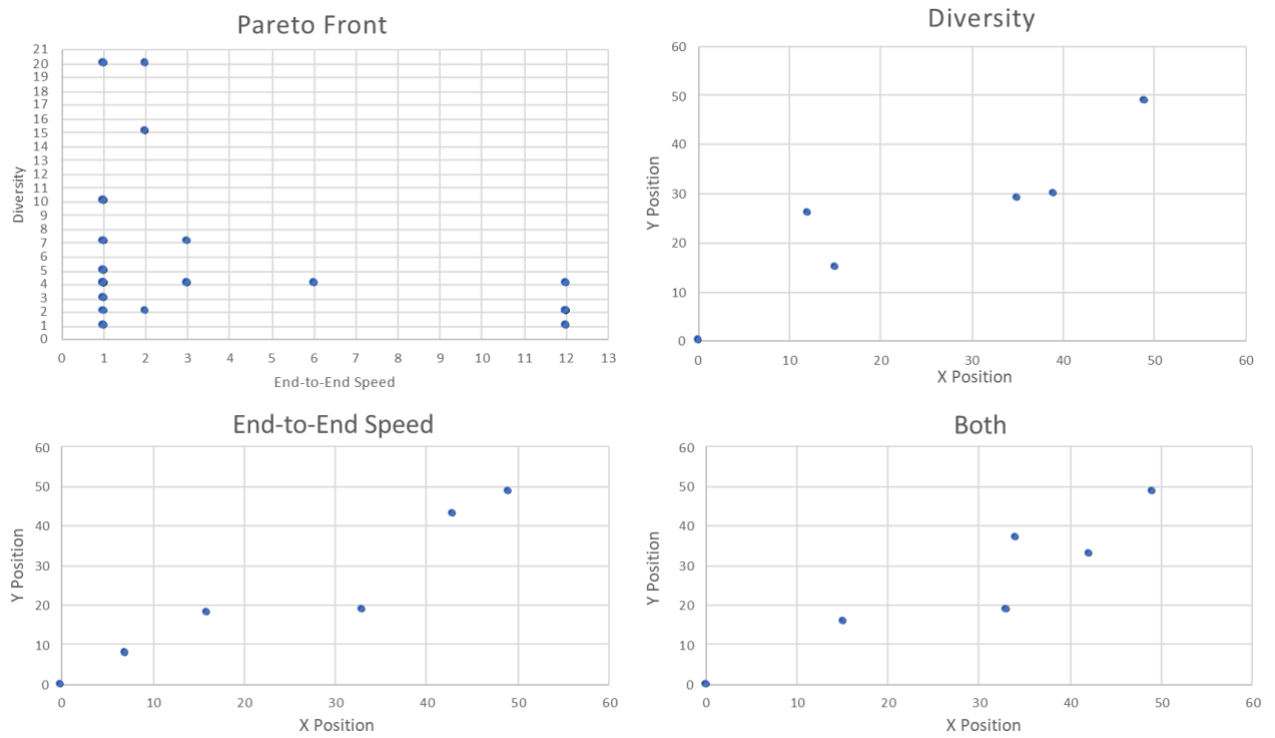


Figure 42: Example 4 Relay-Node Placements for Differing Constraint Preferences

4.6.1.3 Relay Node Placement with Obstacles

An experiment has been run with obstacles, keeping in mind that rural areas tend to have hills and other obstructions such as woodland. Therefore, the algorithm has been implemented in such a way as to eliminate one of the important drawbacks of rural placement. Figure 43 shows that our algorithm is able to bypass obstacles by performing line-of-sight checks. Thus, even if nodes are geographically close, if line-of-sight is not available, they are considered out-of-range. This experiment shows that our tool is able to deal with obstacle or line of sight problem. Since FSO technology will be used to implement, the LOS is important factor and therefore this experiment gives us solution of the LOS.

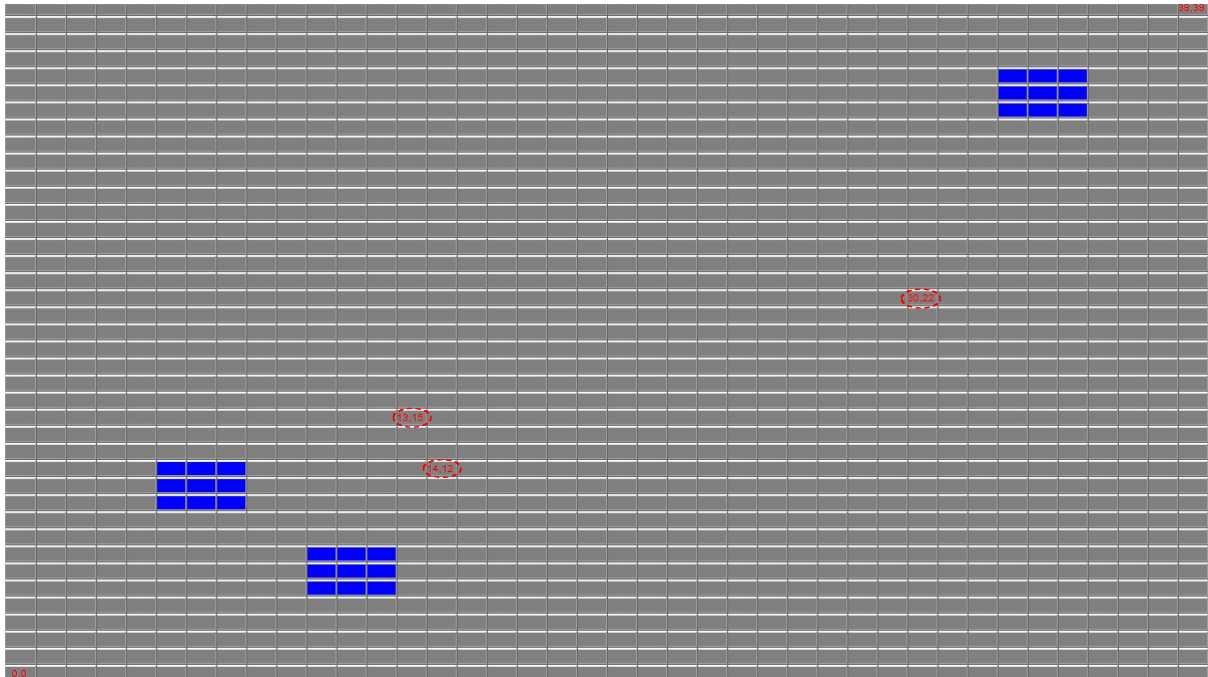


Figure 43: Node Placement with Obstacles

4.6.2 Performance Validation

Initial assessment of our algorithm has confirmed that the tool able to find feasible solutions for the rural problem. We introduce a new evolutionary approach to multicriteria optimization, based on SPEA, that combines several features of previous multi-objective EAs in a unique manner [123]. From Figure 44, the initial output shows encouraging and achieve our objective.

0.49	1.49	2.49	3.49	4.49	5.49	6.49	7.49	8.49	9.49	10.49	11.49	12.49	13.49	14.49	15.49	16.49	17.49	18.49	19.49	20.49	21.49	22.49	23.49	24.49	25.49	26.49	27.49	28.49	29.49	30.49	31.49	32.49	33.49	34.49	35.49	36.49	37.49	38.49	39.49	40.49	41.49	42.49	43.49	44.49	45.49	46.49	47.49	48.49	49.49	50.49
0.48	1.48	2.48	3.48	4.48	5.48	6.48	7.48	8.48	9.48	10.48	11.48	12.48	13.48	14.48	15.48	16.48	17.48	18.48	19.48	20.48	21.48	22.48	23.48	24.48	25.48	26.48	27.48	28.48	29.48	30.48	31.48	32.48	33.48	34.48	35.48	36.48	37.48	38.48	39.48	40.48	41.48	42.48	43.48	44.48	45.48	46.48	47.48	48.48	49.48	50.48
0.47	1.47	2.47	3.47	4.47	5.47	6.47	7.47	8.47	9.47	10.47	11.47	12.47	13.47	14.47	15.47	16.47	17.47	18.47	19.47	20.47	21.47	22.47	23.47	24.47	25.47	26.47	27.47	28.47	29.47	30.47	31.47	32.47	33.47	34.47	35.47	36.47	37.47	38.47	39.47	40.47	41.47	42.47	43.47	44.47	45.47	46.47	47.47	48.47	49.47	50.47
0.46	1.46	2.46	3.46	4.46	5.46	6.46	7.46	8.46	9.46	10.46	11.46	12.46	13.46	14.46	15.46	16.46	17.46	18.46	19.46	20.46	21.46	22.46	23.46	24.46	25.46	26.46	27.46	28.46	29.46	30.46	31.46	32.46	33.46	34.46	35.46	36.46	37.46	38.46	39.46	40.46	41.46	42.46	43.46	44.46	45.46	46.46	47.46	48.46	49.46	50.46
0.45	1.45	2.45	3.45	4.45	5.45	6.45	7.45	8.45	9.45	10.45	11.45	12.45	13.45	14.45	15.45	16.45	17.45	18.45	19.45	20.45	21.45	22.45	23.45	24.45	25.45	26.45	27.45	28.45	29.45	30.45	31.45	32.45	33.45	34.45	35.45	36.45	37.45	38.45	39.45	40.45	41.45	42.45	43.45	44.45	45.45	46.45	47.45	48.45	49.45	50.45
0.44	1.44	2.44	3.44	4.44	5.44	6.44	7.44	8.44	9.44	10.44	11.44	12.44	13.44	14.44	15.44	16.44	17.44	18.44	19.44	20.44	21.44	22.44	23.44	24.44	25.44	26.44	27.44	28.44	29.44	30.44	31.44	32.44	33.44	34.44	35.44	36.44	37.44	38.44	39.44	40.44	41.44	42.44	43.44	44.44	45.44	46.44	47.44	48.44	49.44	50.44
0.43	1.43	2.43	3.43	4.43	5.43	6.43	7.43	8.43	9.43	10.43	11.43	12.43	13.43	14.43	15.43	16.43	17.43	18.43	19.43	20.43	21.43	22.43	23.43	24.43	25.43	26.43	27.43	28.43	29.43	30.43	31.43	32.43	33.43	34.43	35.43	36.43	37.43	38.43	39.43	40.43	41.43	42.43	43.43	44.43	45.43	46.43	47.43	48.43	49.43	50.43
0.42	1.42	2.42	3.42	4.42	5.42	6.42	7.42	8.42	9.42	10.42	11.42	12.42	13.42	14.42	15.42	16.42	17.42	18.42	19.42	20.42	21.42	22.42	23.42	24.42	25.42	26.42	27.42	28.42	29.42	30.42	31.42	32.42	33.42	34.42	35.42	36.42	37.42	38.42	39.42	40.42	41.42	42.42	43.42	44.42	45.42	46.42	47.42	48.42	49.42	50.42
0.41	1.41	2.41	3.41	4.41	5.41	6.41	7.41	8.41	9.41	10.41	11.41	12.41	13.41	14.41	15.41	16.41	17.41	18.41	19.41	20.41	21.41	22.41	23.41	24.41	25.41	26.41	27.41	28.41	29.41	30.41	31.41	32.41	33.41	34.41	35.41	36.41	37.41	38.41	39.41	40.41	41.41	42.41	43.41	44.41	45.41	46.41	47.41	48.41	49.41	50.41
0.40	1.40	2.40	3.40	4.40	5.40	6.40	7.40	8.40	9.40	10.40	11.40	12.40	13.40	14.40	15.40	16.40	17.40	18.40	19.40	20.40	21.40	22.40	23.40	24.40	25.40	26.40	27.40	28.40	29.40	30.40	31.40	32.40	33.40	34.40	35.40	36.40	37.40	38.40	39.40	40.40	41.40	42.40	43.40	44.40	45.40	46.40	47.40	48.40	49.40	50.40
0.39	1.39	2.39	3.39	4.39	5.39	6.39	7.39	8.39	9.39	10.39	11.39	12.39	13.39	14.39	15.39	16.39	17.39	18.39	19.39	20.39	21.39	22.39	23.39	24.39	25.39	26.39	27.39	28.39	29.39	30.39	31.39	32.39	33.39	34.39	35.39	36.39	37.39	38.39	39.39	40.39	41.39	42.39	43.39	44.39	45.39	46.39	47.39	48.39	49.39	50.39
0.38	1.38	2.38	3.38	4.38	5.38	6.38	7.38	8.38	9.38	10.38	11.38	12.38	13.38	14.38	15.38	16.38	17.38	18.38	19.38	20.38	21.38	22.38	23.38	24.38	25.38	26.38	27.38	28.38	29.38	30.38	31.38	32.38	33.38	34.38	35.38	36.38	37.38	38.38	39.38	40.38	41.38	42.38	43.38	44.38	45.38	46.38	47.38	48.38	49.38	50.38
0.37	1.37	2.37	3.37	4.37	5.37	6.37	7.37	8.37	9.37	10.37	11.37	12.37	13.37	14.37	15.37	16.37	17.37	18.37	19.37	20.37	21.37	22.37	23.37	24.37	25.37	26.37	27.37	28.37	29.37	30.37	31.37	32.37	33.37	34.37	35.37	36.37	37.37	38.37	39.37	40.37	41.37	42.37	43.37	44.37	45.37	46.37	47.37	48.37	49.37	50.37
0.36	1.36	2.36	3.36	4.36	5.36	6.36	7.36	8.36	9.36	10.36	11.36	12.36	13.36	14.36	15.36	16.36	17.36	18.36	19.36	20.36	21.36	22.36	23.36	24.36	25.36	26.36	27.36	28.36	29.36	30.36	31.36	32.36	33.36	34.36	35.36	36.36	37.36	38.36	39.36	40.36	41.36	42.36	43.36	44.36	45.36	46.36	47.36	48.36	49.36	50.36
0.35	1.35	2.35	3.35	4.35	5.35	6.35	7.35	8.35	9.35	10.35	11.35	12.35	13.35	14.35	15.35	16.35	17.35	18.35	19.35	20.35	21.35	22.35	23.35	24.35	25.35	26.35	27.35	28.35	29.35	30.35	31.35	32.35	33.35	34.35	35.35	36.35	37.35	38.35	39.35	40.35	41.35	42.35	43.35	44.35	45.35	46.35	47.35	48.35	49.35	50.35
0.34	1.34	2.34	3.34	4.34	5.34	6.34	7.34	8.34	9.34	10.34	11.34	12.34	13.34	14.34	15.34	16.34	17.34	18.34	19.34	20.34	21.34	22.34	23.34	24.34	25.34	26.34	27.34	28.34	29.34	30.34	31.34	32.34	33.34	34.34	35.34	36.34	37.34	38.34	39.34	40.34	41.34	42.34	43.34	44.34	45.34	46.34	47.34	48.34	49.34	50.34
0.33	1.33	2.33	3.33	4.33	5.33	6.33	7.33	8.33	9.33	10.33	11.33	12.33	13.33	14.33	15.33	16.33	17.33	18.33	19.33	20.33	21.33	22.33	23.33	24.33	25.33	26.33	27.33	28.33	29.33	30.33	31.33	32.33	33.33	34.33	35.33	36.33	37.33	38.33	39.33	40.33	41.33	42.33	43.33	44.33	45.33	46.33	47.33	48.33	49.33	50.33
0.32	1.32	2.32	3.32	4.32	5.32	6.32	7.32	8.32	9.32	10.32	11.32	12.32	13.32	14.32	15.32	16.32	17.32	18.32	19.32	20.32	21.32	22.32	23.32	24.32	25.32	26.32	27.32	28.32	29.32	30.32	31.32	32.32	33.32	34.32	35.32	36.32	37.32	38.32	39.32	40.32	41.32	42.32	43.32	44.32	45.32	46.32	47.32	48.32	49.32	50.32
0.31	1.31	2.31	3.31	4.31	5.31	6.31	7.31	8.31	9.31	10.31	11.31	12.31	13.31	14.31	15.31	16.31	17.31	18.31	19.31	20.31	21.31	22.31	23.31	24.31	25.31	26.31	27.31	28.31	29.31	30.31	31.31	32.31	33.31	34.31	35.31	36.31	37.31	38.31	39.31	40.31	41.31	42.31	43.31	44.31	45.31	46.31	47.31	48.31	49.31	50.31
0.30	1.30	2.30	3.30	4.30	5.30	6.30	7.30	8.30	9.30	10.30	11.30	12.30	13.30	14.30	15.30	16.30	17.30	18.30	19.30	20.30	21.30	22.30	23.30	24.30	25.30	26.30	27.30	28.30	29.30	30.30	31.30	32.30	33.30	34.30	35.30	36.30	37.30	38.30	39.30	40.30	41.30	42.30	43.30	44.30	45.30	46.30	47.30	48.30	49.30	50.30
0.29	1.29	2.29	3.29	4.29	5.29	6.29	7.29	8.29	9.29	10.29	11.29	12.29	13.29	14.29	15.29	16.29	17.29	18.29	19.29	20.29	21.29	22.29	23.29	24.29	25.29	26.29	27.29	28.29	29.29	30.29	31.29	32.29	33.29	34.29	35.29	36.29	37.29	38.29	39.29	40.29	41.29	42.29	43.29	44.29	45.29	46.29	47.29	48.29	49.29	50.29
0.28	1.28	2.28	3.28	4.28	5.28	6.28	7.28	8.28	9.28	10.28	11.28	12.28	13.28	14.28	15.28	16.28	17.28	18.28	19.28	20.28	21.28	22.28	23.28	24.28	25.28	26.28	27.28	28.28	29.28	30.28	31.28	32.28	33.28	34.28	35.28	36.28	37.28	38.28	39.28	40.28	41.28	42.28	43.28	44.28	45.28	46.28	47.28	48.28	49.28	50.28
0.27	1.27	2.27	3.27	4.27	5.27	6.27	7.27	8.27	9.27	10.27	11.27	12.27	13.27	14.27	15.27	16.27	17.27	18.27	19.27	20.27	21.27	22.27	23.27	24.27	25.27	26.27	27.27	28.27	29.27	30.27	31.27	32.27	33.27	34.27	35.27	36.27	37.27	38.27	39.27	40.27	41.27	42.27	43.27	44.27	45.27	46.27	47.27	48.27	49.27	50.27
0.26	1.26	2.26	3.26	4.26	5.26	6.26	7.26	8.26	9.26	10.26	11.26	12.26	13.26	14.26	15.26	16.26	17.26	18.26	19.26	20.26	21.26	22.26	23.26	24.26	25.26	26.26	27.26	28.26	29.26	30.26	31.26	32.26	33.26	34.26	35.26	36.26	37.26	38.26	39.26	40.26	41.26	42.26	43.26	44.26	45.26	46.26	47.26	48.26	49.26	50.26
0.25	1.25	2.25	3.25	4.25	5.25	6.25	7.25	8.25	9.25	10.25	11.25	12.25	13.25	14.25	15.25	16.25	17.25	18.25	19.25	20.25	21.25	22.25	23.2																											

4.6.3 Injecting Fresh Chromosomes

The first modification of the tool is brought about by injecting new chromosomes in every generation. The tool injects five new random chromosomes in the primary population except for generation 0. This has a good impact on the final output as it increases exploration of the search space and tries to find a more efficient chromosome. That retains diversity in the population and has a beneficial impact on the final output. A sample primary population is shown in Table 9.

	Current Population			
Chromosome No	Node 1	Node 2	Node 3	Node 4
1	(18,6)	(12,21)	(29,25)	(43,34)
2	(18,6)	(12,21)	(29,25)	(43,34)
3	(18,6)	(12,21)	(29,25)	(43,34)
4	(18,6)	(12,21)	(29,25)	(43,34)
5	(25,48)	(12,21)	(29,25)	(43,34)
6	(18,6)	(12,21)	(29,25)	(43,34)
7	(18,6)	(18,31)	(29,25)	(43,34)
8	(18,6)	(12,21)	(29,25)	(43,34)
9	(18,6)	(12,21)	(29,25)	(43,34)
10	(4,33)	(12,21)	(29,25)	(43,34)
11	(18,6)	(37,32)	(29,25)	(43,34)
12	(18,6)	(12,21)	(29,25)	(43,34)
13	(18,6)	(12,21)	(29,25)	(43,34)
14	(18,6)	(12,21)	(29,25)	(43,34)
15	(18,6)	(12,21)	(29,25)	(43,34)
16	(31,24)	(36,18)	(42,22)	(23,42)
17	(2,4)	(22,11)	(36,25)	(29,36)
18	(1,6)	(10,6)	(42,9)	(44,40)
19	(10,26)	(27,39)	(40,32)	(40,41)
20	(12,11)	(18,41)	(24,38)	(46,26)

Table 9: Primary Population with Chromosome Injection

After observing Table 9, it is noticed that there are lots of repeated nodes, which means the tool is experiencing a convergence problem. But chromosome no 16-20 are new randomly generated chromosomes which have been injected. This technique has given renewed diversity to the population.

4.6.4 Increasing Mutation Rate

Mutation is the usual technique for maintaining diversity in populations to ensure the problem space is adequately searched. Given that the population size is finite, crossover by itself provides little opportunity to search many of the possible solutions. It is only producing variations of a limited set of genes. Mutation, however, introduces fresh genes into the chromosomes. I set the number of Generations = 100, Crossover = 70%, the number of chromosomes is 20, and I increase the mutation rate to 50%. This relatively high value helps

create fresh chromosomes. Given that this is coupled with elitism, we can allow for this without the risk of losing particularly good solutions, unless better ones are found.

There are a couple of ways to mutate any chromosome. One way is to mutate a single node of a particular chromosome. Another method can be to mutate the whole chromosome by changing all three nodes. Currently, I use the single node mutation technique. The sample output is shown in Table 10.

	Before Mutation					After Mutation			
Chromosome No	Node 1	Node 2	Node 3	Node 4	Chromosome No	Node 1	Node 2	Node 3	Node 4
1	(13,11)	(29,15)	(30,28)	(40,36)	1	(13,11)	(29,15)	(30,28)	(40,36)
2	(13,11)	(29,15)	(30,28)	(40,36)	2	(13,11)	(46,6)	(30,28)	(40,36)
3	(13,11)	(29,15)	(30,28)	(40,36)	3	(13,11)	(29,15)	(30,28)	(40,36)
4	(13,11)	(29,15)	(30,28)	(40,36)	4	(13,11)	(29,15)	(30,28)	(15,6)
5	(13,11)	(29,15)	(30,28)	(40,36)	5	(29,18)	(29,15)	(30,28)	(40,36)
6	(13,11)	(29,15)	(30,28)	(40,36)	6	(13,11)	(29,15)	(30,28)	(40,36)
7	(13,11)	(29,15)	(30,28)	(40,36)	7	(13,11)	(29,15)	(30,28)	(40,36)
8	(13,11)	(29,15)	(30,28)	(40,36)	8	(13,11)	(29,15)	(30,28)	(40,36)
9	(13,11)	(29,15)	(30,28)	(40,36)	9	(13,11)	(29,15)	(30,28)	(3,42)
10	(13,11)	(29,15)	(30,28)	(40,36)	10	(13,11)	(29,15)	(30,28)	(40,36)
11	(13,11)	(29,15)	(30,28)	(40,36)	11	(13,11)	(29,15)	(30,28)	(40,36)
12	(13,11)	(29,15)	(30,28)	(40,36)	12	(13,11)	(29,15)	(30,28)	(40,36)
13	(13,11)	(29,15)	(30,28)	(40,36)	13	(13,11)	(29,15)	(45,47)	(40,36)
14	(13,11)	(29,15)	(30,28)	(40,36)	14	(1,7)	(29,15)	(30,28)	(40,36)
15	(13,11)	(29,15)	(30,28)	(40,36)	15	(1,46)	(29,15)	(30,28)	(40,36)
16	(13,11)	(29,15)	(30,28)	(40,36)	16	(46,7)	(29,15)	(30,28)	(40,36)
17	(13,11)	(29,15)	(30,28)	(40,36)	17	(13,11)	(29,15)	(30,28)	(40,36)
18	(13,11)	(29,15)	(30,28)	(40,36)	18	(13,11)	(29,15)	(30,28)	(40,36)
19	(13,11)	(29,15)	(30,28)	(40,36)	19	(13,11)	(29,15)	(30,28)	(40,36)
20	(13,11)	(29,15)	(30,28)	(40,36)	20	(13,11)	(29,15)	(30,28)	(40,36)

Table 10: Increased Mutation Rate Scenario

There is a significant change in the population, and thus, it improves the diversity. Though there is some improvement, the result still shows that premature convergence exists. So, this technique is beneficial but not enough. If the tool mutates a single node from 4 nodes, then the tool is only making minor adjustments to the chromosomes in each generation and the risk of premature convergence persists. This shows this way isn't effective enough.

From Table 10 some evidence of premature convergence is present due to the repeated chromosomes. In this experiment it arises after generation 30. Even so, the graphical output in Figure 45 shows that the output is closer to the ideal situation.

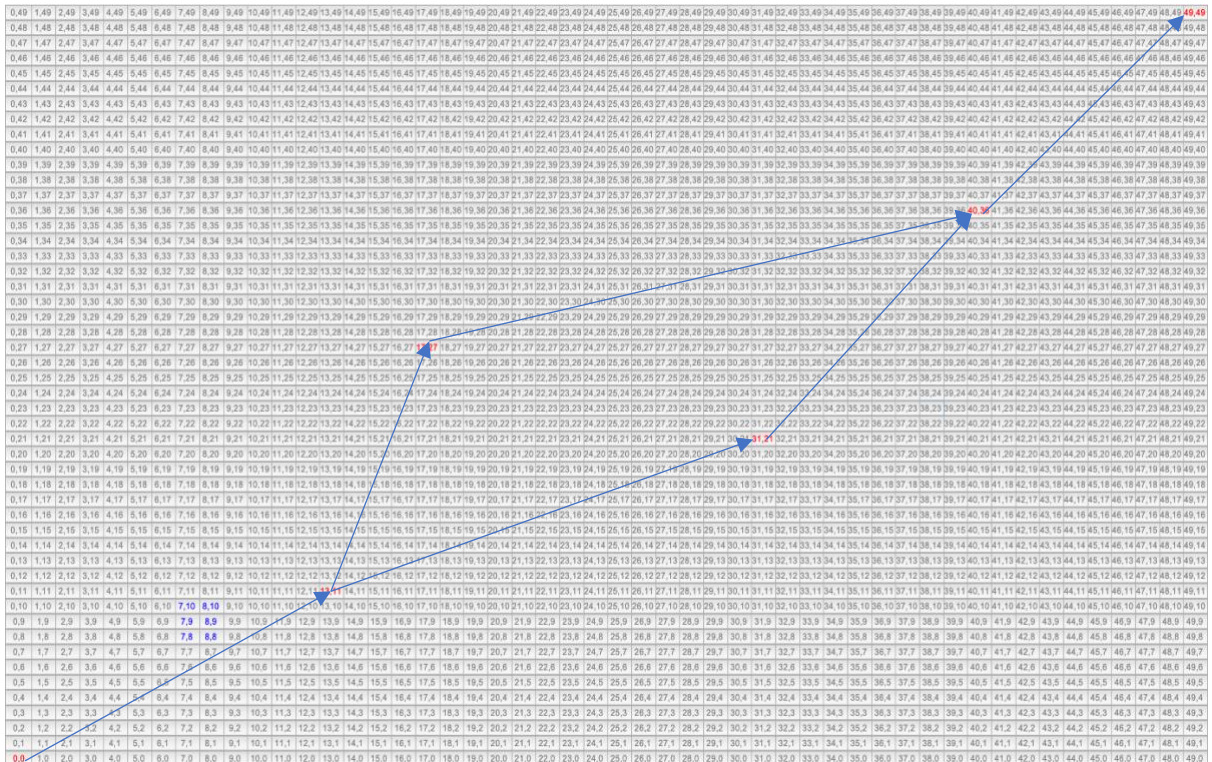


Figure 45: Output for Convergence Test

4.6.5 NSGA Crowding Distance

It is desirable for an Evolutionary Algorithm (EA) to maintain a good spread of solutions along with convergence to the Pareto-optimal set. The original NSGA used the well-known sharing function approach, which was designed to maintain sustainable diversity in a population with the appropriate setting of its associated parameter. This parameter is related to the distance metric chosen to calculate the proximity measure between two population members. This is called the crowding distance. The parameter denotes the largest value that distance metric within which any two solutions share each other's fitness. That is, between two solutions with differing non-domination ranks, the solution with a higher value of fitness, in other words, the lower rank, is preferred.

Diversity among non-dominated solutions is introduced by using the crowding comparison procedure used in the fitness assignment while developing the Pareto front. Since solutions compete with their crowding distance, no extra niche parameter is required. However, the crowding distance is calculated in the objective function. While plotting the Pareto Set, the x-axis is set as end-to-end (E2E) speed, and the y-axis is set as Diversity. Therefore, if the Pareto Set provides the highest y-axis value for any individual x-axis value, it receives the lowest dominance rank and vice versa. So, the aim is to receive the highest y-axis value once for any specific x-axis value. This way, the algorithm converges to the true Pareto front whilst finding diverse solutions.

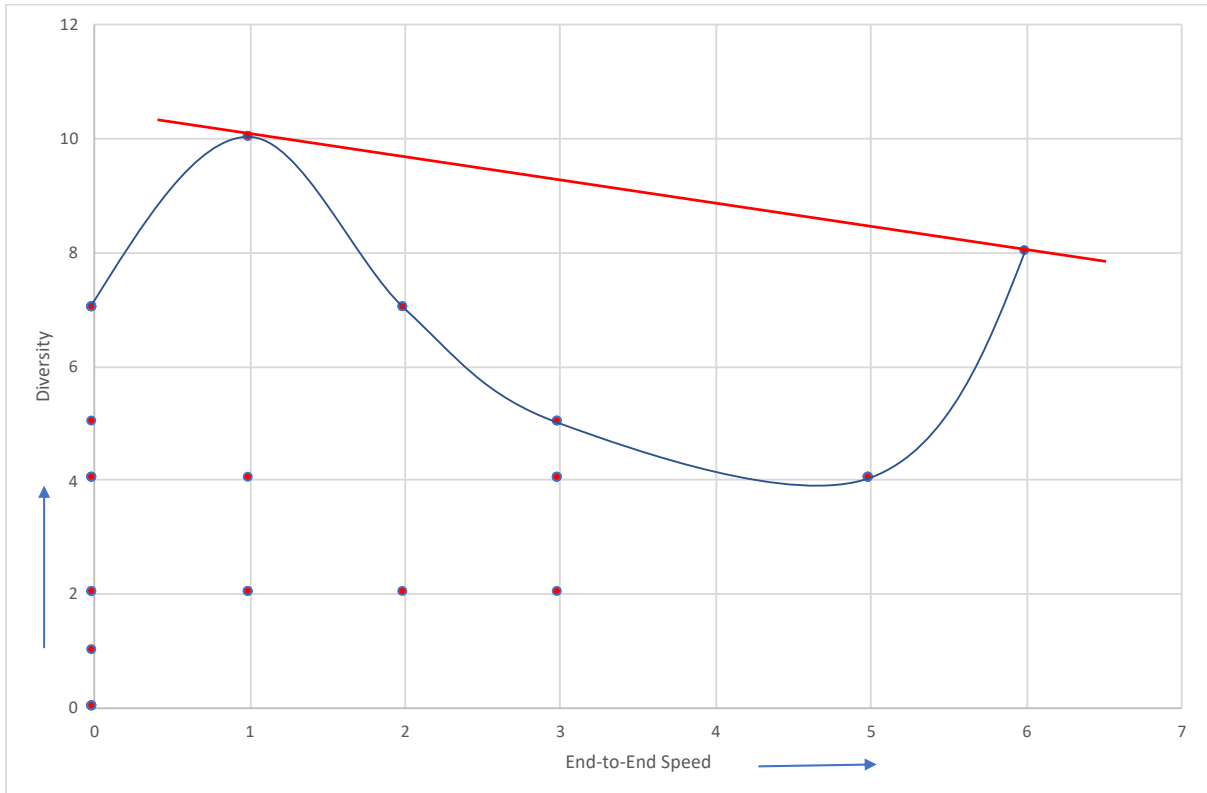


Figure 46: Pareto Front in Generation 10

Figure 46 displays Pareto Front and all solutions obtained with SPEA plotted in Pareto Set. We plot the solutions found from Generation 10 using their E2E Speed and Diversity as XY coordinates in the graph. SPEA performs better with crowding distance (CD) in this problem than without CD in both aspects of convergence and distribution of solutions. Figure 46 shows that the Pareto Set has been plotted where End-to-End speed is on the x-axis, and diversity is on the y-axis. Once a temporary population is created, the fitness cost of all chromosomes in the temporary population has been plotted in the Pareto Set. The blue line in Figure 46, represents the distribution of chromosomes and this remark clear indication of better distribution compared to Figure 47. Also, the red line represents the non-dominated solutions from PS. Undoubtedly, the final outputs are going to be much improved and under our expectations.

The Pareto Set without CD is drawn in Figure 47. This graph shows a less diverse population in the Pareto Set compared to Figure 46. Again, the blue line represents the distribution of the solution on the PS in Figure 47.

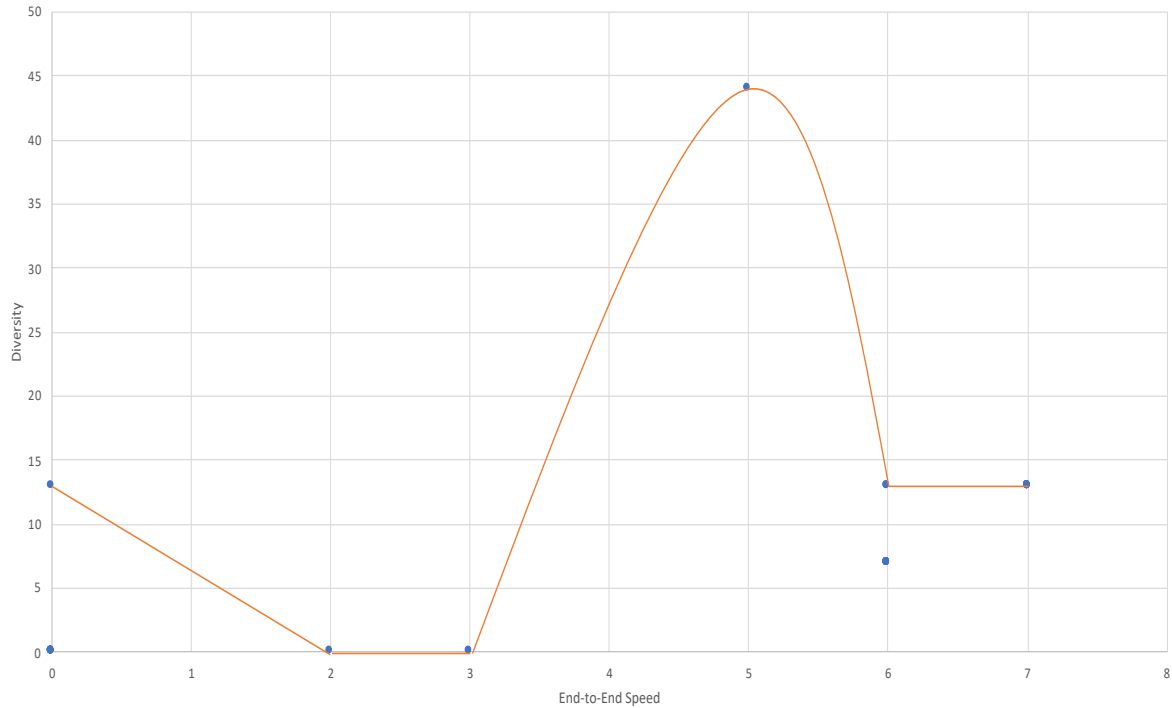


Figure 47: Pareto Set without CD

In terms of a better result, Figure 46 shows a more diverse result. More importantly, the Pareto front enables us to find a better solution based on the compromise we wish to make between both objectives. With CD, the distribution in solutions is better with SPEA. Though Figure 47 shows that when the End-to-End speed is 5, diversity reached the highest point, but there are too less distribution of solution to justify this result, because the PS isn't populated with proper distribution of solutions. It is clear that the solutions were duplicated in temporary population, thus, premature convergence exist.

4.7 Discussion

This tool is an improved version of the previous network planning tool described in Chapter 3 where the tool was designed with a Genetic Algorithm (GA) for the placement of relay nodes to form an FSO-based topology. Our problem has multiple constraints when considering a potential solution to improve network efficiency. Therefore, two objectives: E2E speed and diversity has been considered initially. However, there are other constraints, such as battery power, and FSO interference, can be considered in future. The final topologies we receive from the tool is the optimal or near optimal solution for the given objectives according to Section 4.2.2.

One of the novel contributions of this project is in the field of network planning. Our network-planning tool can be used to evaluate the feasibility of low-cost network deployment in a rural context. More than that, our network-planning tool considers environmental trade-offs to discover suitable relay node placement. An initial implementation was developed using a Genetic Algorithm (GA). The purpose of using GA is to traverse a search space for possible answers to a problem using previous "good" guesses to influence subsequent ones. This typically leads to a near-optimal solution being found quicker than would be the case with

random searches. However, a GA only optimises for a single objective, although this could be based on a weighted sum of multiple objectives.

There remains a need for a network planning tool for determining suitable node placement in a rural context, particularly for the case of FSO deployment. The reasons for slow Internet speeds in rural areas mostly relate to location and infrastructure. Most rural people with Internet access, obtain the service via the copper wires from a traditional telephone exchange. The further connection has to travel over these copper lines, the slower it gets, so the distance people live from nearest telephone exchange or street cabinet, the slower the service will be. To make matters worse, upgrades to the infrastructure often only occur where it is economically viable.

There are number of evolutionary techniques, however, we decided to use Strength Pareto Evolutionary Algorithm (SPEA-II), described in Section 2.8. MOEAs are a population-based optimization approach taking advantage of the principles of biological evolution. The optimization is done iteratively in two alternating steps: generation and selection. In the generation phase, a new solution/generation is created from a set of existing solutions in the population. This is done by crossover and mutation operators. Correspondingly, the selection filters out the worst solutions to improve the likelihood of convergence to optimal solutions. EA is a search technique used in computing to find exact or approximate solutions to optimization and search problems. EAs are typically used to provide good approximate solutions to problems that cannot be solved easily using other techniques. For the rural broadband problem, where we quantify the worth of solution, EA can perform a directed search of the solution space. On the other hand, Neural Networks (NN) are non-linear statistical data modelling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data. NN can ‘learn’ to classify items it hasn’t encountered before with complex mathematical methods [126].

I have implemented a tool with an Evolutionary Algorithm that can solve multiple objective problems. There are several types of Evolutionary Algorithm apart from Multi Objective Evolutionary Algorithm (MOEA) as mentioned earlier in background Section. The presence of multiple objectives in a problem, gives rise to a set of Pareto-optimal solutions. One approach is a fast and elitist Multi Objective Genetic Algorithm employing non-dominated sorting, called Non-Dominated Sorting Genetic Algorithm (NSGA-II) in [102] which explained in Section 2.11. Another Pareto based evolutionary algorithm that considers multiple objectives to find optimal solutions is called the Strength Pareto Evolutionary Algorithm (SPEA) [115]. We based our approach on SPEA based on its performance and suitability to our research. More precisely, we created a hybrid MOEA based on SPEA-II which implements line of sight (LOS) checking, crowding distance to avoid solution bunching, and Dijkstra Algorithm when quantifying the goodness of a solution. LOS is a primary constraint of FSO as the relay nodes must have LOS between neighbours in order to transmit the data. Dijkstra’s Algorithm is implemented to find minimum cost path, in other words, best possible path that provides highest E2E speed, as described in Section 4.4.2. One of the problems with MOEA is maintaining solution diversity along the Pareto Front (PF). Therefore, we implement NSGA’s Crowding Distance (CD) to eliminate this problem.

Our hybrid MOEA algorithm is able to provide optimal solutions where network operators are presented with various planning solutions according to their requirements. The final output as shown in Section 4.5 confirms that our tool able to find an optimal solution that provides path diversity with high E2E speed. In addition, we ran number of experiments to ensure our tool is working fine including scenarios with obstacles. The tool was shown to consistently provide good feasible solutions for the placement of FSO relay nodes is a rural area.

Chapter 5. Traffic Engineering System

Traffic Engineering (TE) is a method of optimizing a communication network's performance by dynamically analysing and controlling the behaviour of data transmitted over that network resources. It is a procedure used to adapt the traffic routing, and possibly other parameters, to the changes in the network conditions. The aim of TE is to improve and maintain network performance, Quality of Service (QoS) and user experience typically by providing the best E2E speed by efficient use of resources, which can lower the operational costs.

In this research, we assume that the placement of the relay nodes has already been determined (refer back to Chapter 30). The first part of the tool described in Chapter 30, has determined the placement of nodes in order to provide the best possible route to transmit user data to rural community. Therefore, the focus of our interest now is how best to maintain service delivery by selecting the most appropriate forwarding path considering factors such as relay node transmission power, residual battery life and End-to-End (E2E) transmission speed. As the environment influences these factors, we need to provide a mechanism for dynamic path selection and configuration of the transceivers. This is Traffic Engineering (TE). The difference between NP tool and TE tool is the tool decides the topology using NP and TE tool used to maintain and provide stability to the infrastructure built with that topology.

TE reduces the service degradation due to congestion and failure, e.g., links impacted by weather. It aims to ensure that if a transmitting node failure occurs in the network, the transmitted data can still be delivered to the destination. Another goal of TE is to balance the QoS against the cost of operating and maintaining the network. Such a TE framework could be based on Software Defined Networking (SDN), as introduced in Chapter 2.12, whereby an intelligent control plane can be used to (re)configure the underlying data plane in response to changing circumstances. We assume the TE system will be applied to rural FSO networks, such as shown in Figure 48. Each relay node is connected via FSO links to form the underlying topology. These links could then be used to interconnect the control plane functionality in each relay node via in-band signalling or other means.

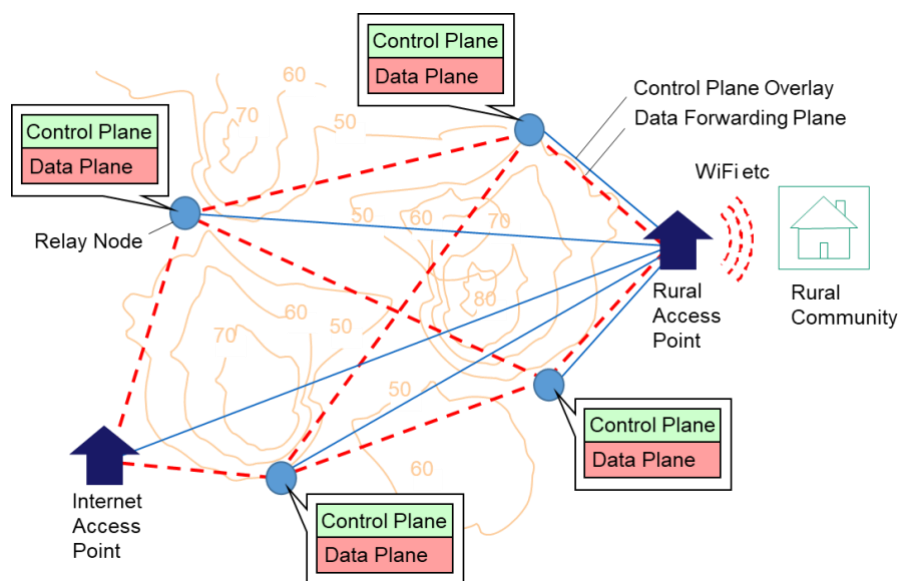


Figure 48: Typical Rural Topology with TE-Enabled Relay Nodes

Although it is beyond the scope of this thesis, we envision probe messages going back and forth between nodes to provide information about the link bitrate and battery power of individual nodes. All this information would be stored in an appointed Manager node. This special node would collate the monitored information and support the TE control software, such as the algorithm we have created, which decides the course of action (i.e. path selection and power setting) according to the current conditions.

The TE maintains the path that provides the highest E2E Speed. Therefore, it is essential to monitor the overall network constantly. Thus, the tool might need to decide over time to keep the network running. The frequency of running the tool is up to the ISP provider, who chooses how often TE decisions need to be made. However, the probe messages that transmit monitoring information to the manager node could be used to trigger the tool to run in order to make a decision immediately under certain circumstances, such as a link failure. However, it is normally expected to be run periodically, such as once an hour, in order to respond to progressively changing weather conditions.

Three experiments have been executed to develop the traffic engineering system that will collaborate with our network planning tool explained in Chapter 3, which determines node placement within a rural area. The TE system takes into account various performance measures and uses this information to (re)configure the network so that it operates efficiently.

5.1 Traffic Engineering Design Parameters

In this research, we start by assuming a network planning tool has already selected the location of the relay nodes from [127], as illustrated in Figure 49. These relay nodes are considered to employ a combination of renewable and battery power sources. Each relay is equipped with several transceivers that support FSO communication. Furthermore, the power of these transceivers is selectable, with greater power allowing for a higher transmission bitrate at a given range. The Internet Access Point (i.e., Point of Presence) is considered to be connected to the high-speed Internet and has access to a limitless power source. As such, its transmission links can continuously transmit at maximum power. Communication is via line-of-sight optical links, and the overall aim is to maintain a high bitrate path between the Internet Access Point (IAP) and the Rural Access Point (RAP). This can be achieved by a combination of path selection and the transmit power of individual links along the path.

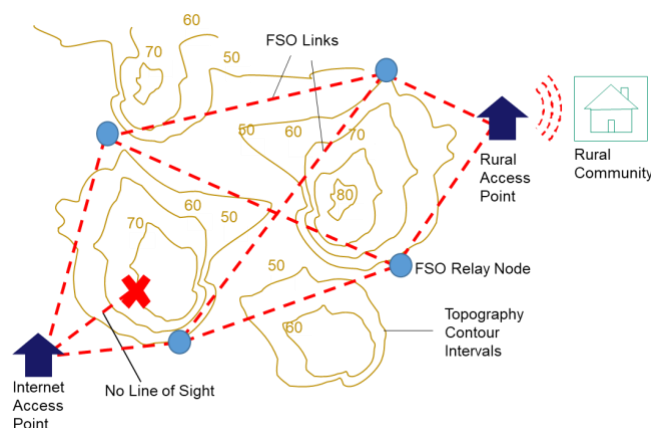


Figure 49: Typical Topographical Topology

We create a TE system using three parameters to ensure the rural broadband operates in an efficient manner. The three parameters we consider are:

1. E2E speed: From a consumer perspective the broadband delivery bitrate of data is perhaps the most significant parameter that determines the quality of experience. In a multi-hop system this is limited by the bitrate of the slowest link along the path. We thus aim to achieve a high value.
2. Transmit Power: The FSO relay nodes are assumed to be capable of adjusting their transmit power. For a given inter-node distance, increasing the power increases the communication bitrate. This parameter also allows the bitrate to be maintained to some degree despite changing weather conditions. However, increasing the power increases the energy usage, and thus drain on the battery (assuming this is the power source).
3. Residual Battery Life: This is an indication of the remaining charge within a node's battery. This is useful when trying to extend the lifetime of the network. Although we assume the batteries are rechargeable, if the charge is relatively low for the nodes along the transmission path, it may be desirable to switch to an alternative path to reduce the risk of a node running out of charge and thus causing a disconnect.

Although our network planning tool and traffic engineering tool are similar in approach, the optimisation parameters and the overall objectives differ. The overall objective of NP tool is to place the relay node in efficient way and TE tool aims to make most efficient use of given resources taking into account of environmental conditions. The TE tool assumes a fixed topology to find the best possible path given various objectives. The TE tool requires a time-sensitive decision, whereas the NP tool is used to decide the overall topology placement, which is expected to remain the same once deployed, so the NP tool isn't time-sensitive. Therefore, simulation time is more important for the TE tool than the NP tool.

5.1.1 End-to-End (E2E) Speed

To calculate the End-to-End (E2E) speed of a path, the distance between every pair of adjacent nodes along the path is obtained. The reference range profile, given in Figure 50, is consulted to determine the effective bitrate of the link for the selected transmit power setting.

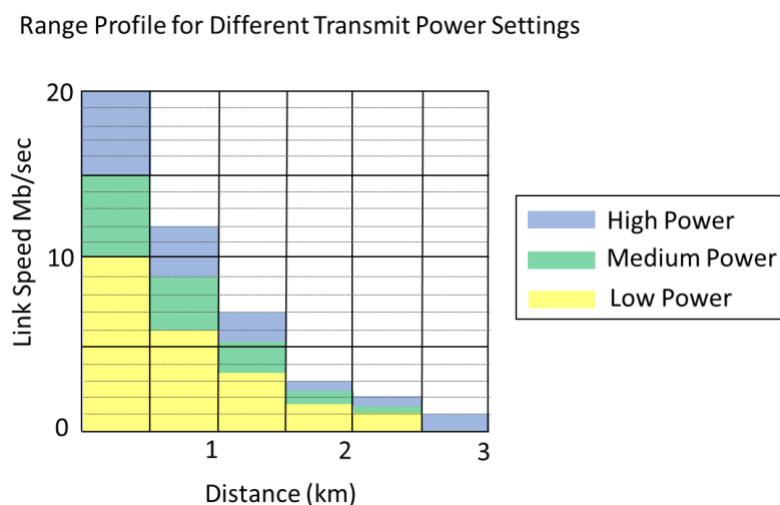


Figure 50: Range Profile for Different Transmit Power Settings

The E2E speed is then calculated as the minimum link bitrate of all the links along the given path. Interestingly, this link acts as a bottleneck such that it may be desirable to either increase the node’s transmit power, if possible, or to potentially reduce the power of other links along the path provided this has no impact on the E2E performance.

As an example, consider the scenario shown in Figure 51. The link bitrates for high, medium and low transmit power values are shown in red, green and black, respectively. The links from node “S” are only shown in red as it is assumed these transceivers would always operate on high power due to the node having access to a reliable power source. The purple trapezoid represents a terrain obstacle. As such only two paths are possible: S-1-2-D and S-3-4-D. However, different power settings are possible for Nodes 1 through to 4 as shown by the various link bitrates.

When encoding a chromosome, representing a possible solution, the power setting of each node along the associated path is recorded. This allows the overall E2E speed to be calculated simply as the bitrate of the slowest link along the path. If a chromosome for a path has node transmit power setting(s) such that the link would not be able to operate, we consider the chromosome as non-viable. Thus, returning to Figure 51, we see the bitrate from Node 3 to Node 4 shows only one speed (in red) which corresponds to high power. As we do not accept link bitrates of less than 1 Mbps, no values are shown for medium or low power settings.

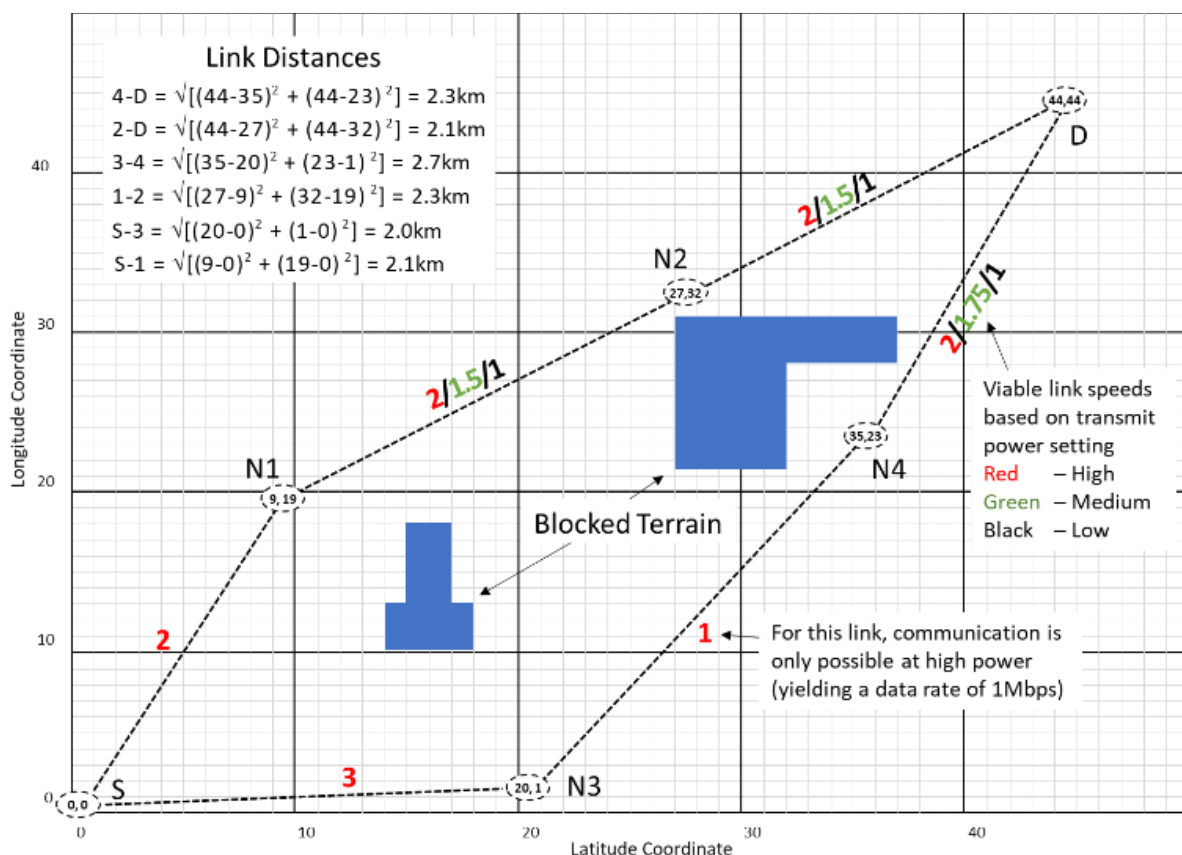


Figure 51: Scenario using New Range Profile

All chromosomes/paths go through the above process to determine their E2E speed and thus their fitness in terms of this parameter, where the goal is to maximize it.

5.1.2 Minimum Residual Battery Life (RBLm)

In a battery constrained environment this parameter is an essential variable before considering the final path/result. Every node has its own Residual Battery Life (RBL) value and its own Transmit Power Setting (TPS), which is the power of the downstream data-path transmitter. The RBL is calculated as:

$$RBL = \text{Current Battery Life} - \text{Energy Drain} \quad (1)$$

$$RBLm = |RBL|_{\min} \quad (2)$$

The Energy drain is directly related to the TPS such that the higher the transmit power, the greater the drain. The minimum Residual Battery Life (RBLm) is the lowest RBL of all the nodes along the data path under consideration. To calculate RBLm, the subsequent RBL will be calculated using the Power Setting table shown in Table 11 to determine the energy drain in Eqn 1. From this the revised RBL of each node on the path is calculated and then the node with the lowest revised battery life, RBLm, is identified.

Table 11: Power Setting Chart

Power Setting	Residual Battery Life (Reduction)
High Power (HP)	10 units
Medium Power (MP)	5 units
Low Power (LP)	2 units

5.1.3 Transmit Power Setting Impactor (TPSI)

There is a relation between the operational transmission speed and transmit power. At a given range, increasing the transmit power of a laser diode, for example, enables a higher data rate to be sustained. In this research we consider three possible power settings: Low Power (LP), Medium Power (MP) and High Power (HP). The range profile of the link using each of these power settings is shown in Figure 50. The corresponding energy drain on the battery in notional units is shown in Table 1. In this research, we define the Transmit Power Setting Impactor (TPSI) by first taking into account the Wasted Excess Power, WEP, of each node along the considered path. This overall WEP is calculated as follows:

$$WEP = \sum_{i=1}^N WEP_i \quad (3)$$

where N is the number of nodes in the corresponding path.

The WEP of a given node is assigned zero if there is no possibility to lower the current power setting whilst preserving the E2E speed along the path. It is assigned a value of 1 if the node power could be lowered by one step and still maintain the E2E speed, and it is assigned the value 2, if it is possible to lower the power by two steps (i.e., from HP to LP) without impacting the E2E speed along the path. The overall WEP is the summation of these values on the path. Thus, the higher the score, the greater degree of wasted power along the path.

The WEP is then used to calculate what we term as the Transmit Power Setting Impactor (TPSI), which is a normalised measure as shown in Equation (4):

$$TPSI = \frac{1}{(WEP+1)} \quad (4)$$

If there is no wastage along the path, i.e., any reduction in the power setting of a node would negatively impact on the E2E speed, then WEP would be zero and the TPSI value would be 1. However, if, for example a path comprises nodes S-1-2-3-D, and the relay Node 1 could drop 1 power step and Node 3 could drop 2 steps, without reducing the overall E2E speed, respectively, then the WEP would be $1+0+2 = 3$, and the TPSI would be $1 / (3 + 1) = 0.25^3$. Thus, the higher the TPSI value the better.

5.2 TE under Genetic Algorithm

We start by implementing the intelligent Traffic Engineering (TE) system using a genetic algorithm, and the GA's fitness score is calculated as the weighted sum of the three parameters given in Chapter 5.1. This weighting factor provides the priority of specific components that are given importance while calculating fitness score.

5.2.1 Fitness Score

This tool has been designed through a Genetic Algorithm that is a classic algorithm that is well known and efficient for linear problem-solving. This algorithm worked with a fitness value that is assigned to all chromosomes. These values called fitness score. The ranking/final result receive through this fitness score. In this tool, the fitness score as below:

$$score = (\alpha * E2E \text{ speed}) + (\beta * \text{Minimum Residual Battery Life}) - \gamma * \text{Total WEP}$$

5.2.1.1 Total WEP under GA

WEP is a reasonably new concept to score the objective and need a numeric expression to calculate. To express and calculate Total WEP under GA, both TPSI and WEP have been introduced in Chapter 5.1.3. The WEP is the amount of power that gets wasted due not to contribute any impact on final E2E speed. WEP is calculated through numeric value where WEP value increase if TPS can be lower. Every step that the TPS can get lower, WEP gets a value of 1, which adds up while checking all nodes corresponding to that path. The final value is the WEP value. The equation is:

$$Total \ WEP = \sum_1^n WEP$$

The process of WEP has been designed in a flowchart which is shown in Figure 52.

³ The Nodes S and D are assumed to have access to mains power and are omitted from the calculation.

5.2.1.2 Flowchart

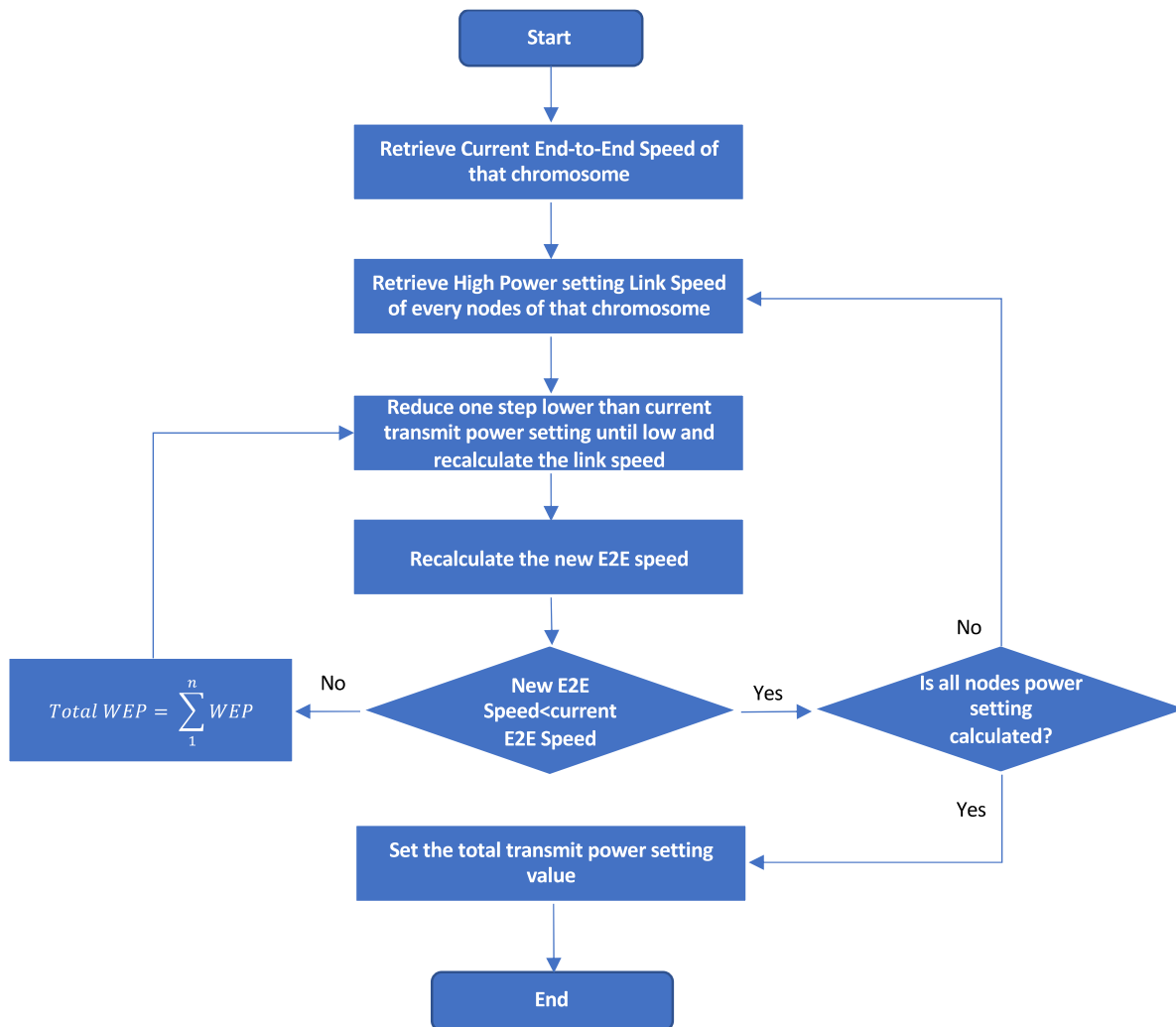


Figure 52: Flowchart of WEP

5.3 End-to-End (E2E) Speed

The purpose of this experiment is to find a valid path that provides the highest E2E speed. This experiment has been done with a unique topology shown in Figure 51, designed for where the priority is given to E2E speed. This means the fitness score in Chapter 5.2.1 will be calculated with higher priority to E2E speed. As it is noticed that the fitness score is calculated with the weighting factor, the score will be calculated in an extreme case such as $\alpha = 1$ and the rest of the weighting factors will be zero. This will provide a score to chromosomes based solely on their E2E speed.

The settings used for this evaluation are in accordance with the scenario given in Figure 51. These specific settings for the simulation are as follows:

- We set the number of generations to 100. This value also applies to subsequent GA evaluations unless otherwise stated.
- The topographical space is set to 45 x 45 square units.
- The chromosome population size is set to 20.

- Fresh five chromosomes are injected in every generation to maintain diversity, and these are randomly generated chromosomes.
- Crossover rate is 30% and mutation rate is 70%.
- To assess how the TE mechanism copes with changes over time, where appropriate, we run it for 5 epochs such that the residual battery life from one epoch becomes the initial value of the next epoch.

To observe the behaviour under dynamic circumstances, where the network evolves over time, without running the tool for many epochs, we set up one of the nodes with smaller Residual Battery Life (RBL) and, thus, Node 1's initial Residual Battery Life set to 30 units intentionally (instead of 100 units for all the remaining relay nodes).

5.3.1 Expectation

As mentioned above, this experiment is to observe the selected path, where the fitness function only considers the E2E speed with an elevated E2E speed weighting factor $\alpha = 1$ and the rest of the weighting factor set to zero. This way, the tool should favour the chromosome representing highest overall speed. Under these circumstances, the tool should select a specific path with the highest speed until the battery of any node along the current path runs out. This is because the fitness score is only based on E2E speed. Therefore, when the Battery Life of any node is exhausted, the link will fail, and thus the tool will switch to another path. According to Figure 51, the initial selected path should be the red one, but once the battery of any node on that path runs out, the traffic will be diverted through the green one. The red path E2E speed is 2 Mbps, and the green path provides 1 Mbps using the range profile in Figure 50. This speed will be accomplished if the nodes are in a high-power setting. As this experiment is based on E2E speed, the tool should allocate all the nodes to a high-power setting, assuming it improves the overall speed.

5.3.2 Result

In order to verify that our tool is functioning appropriately, we run the tool using the settings described in Chapter 5.3. The tool is run for five times in sequence with similar settings but the final RBL from one run becoming the initial RBL of the next. After running the TE tool, the results we found are as follows:

1st GA run: (Time X)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	H	20
N2	(28,32)	H	90
N3	(20,1)		100
N4	(35,23)		100
D	(44,44)		100

2nd GA run: (Time X+1)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	H	10
N2	(28,32)	H	80
N3	(20,1)		100
N4	(35,23)		100
D	(44,44)		100

3rd GA run: (Time X+2)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	H	0
N2	(28,32)	H	70
N3	(20,1)		100
N4	(35,23)		100
D	(44,44)		100

4th GA run: (Time X+3)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)		0
N2	(28,32)		70
N3	(20,1)	H	90
N4	(35,23)	M	95
D	(44,44)		100

5th GA run: (X+4)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)		0
N2	(28,32)		70
N3	(20,1)	H	80
N4	(35,23)	M	90
D	(44,44)		100

5.3.3 Analysis

From the results in Chapter 5.3.2 the tool has selected the path which provides the best End-to-End speed. The path remains the same path until the Residual Battery Life of N1 reaches zero. Since N1 runs out of battery life at the 3rd epoch, the current path switches to another path bypassing this node. Also, it has been noticed that most of the nodes are running with a higher (H) power setting, which provides the best E2E speed. Technically, there is either no impact of power setting or RBLm while calculating score due to the zero value of the weighting factor of power setting and RBLm. Theoretically, the reason behind the higher power setting is that the End-to-End speed is higher when the Power Setting is in high (H) mode.

One observation of the result in 4th epoch, which shows one of the nodes assigned a Low (L) power setting, which may appear surprising. However, in this instance the link speed can go

down to the Low (L) power setting without impacting the End-to-End speed (E2E speed). With reference to Chapter 5.3.1, our tool behaves as expected.

5.4 Minimum Residual Battery Life (RBLm)

Similar to the process described in Chapter 5.3, the tool will now prioritize the Minimum Residual Battery Life (RBLm) of the nodes in the chromosomes. Therefore, the weighting factor of Minimum Residual Battery Life (β) is set as 1, and the remaining weighting factors (α & γ) are set to zero. The fitness score will thus be determined by the RBLm alone. The behaviour we are expecting is that, as the Residual Battery Life (RBL) of the nodes along the active path is reducing over time, the transmitting path will alternate to preserve the minimum residual battery life. Hence, as the fitness score reduces an alternate the current path will be selected.

The simulation settings are the same as for the E2E speed experiment described in Chapter 5.3. However, in this instance we run the tool for 10 epochs to observe the selected path over time. We also set up couple of the nodes with a smaller initial RBL in order to see whether tool is able to switch the alternative path when RBL of one of those nodes runs out. Thus, Node 1 and Node 3's Residual Battery Life are initialised to 30 units. The remaining relay nodes have an initial battery life of 100 units.

5.4.1 Expectation

As mentioned above, the weighting factor of Minimum Residual Battery Life (β) is set at the highest value of one, and the rest of the weighting factors are set to zero. The tool should therefore make path selections in such a way as to preserve the battery life of the relay nodes. This happens because every time the GA run finishes, the Residual Battery Life of the current path's nodes gets reduced according to their transmit power setting. The tool chooses the path that will consume less battery life (as the RBLm will remain higher). Thus, a lower power setting will be selected, if possible, as it lowers RBLm less. As stated, the weighting factor is such that a better score is assigned to chromosomes that result in a higher RBLm. This is the reason the current path will stay the same until the energy cost for the forthcoming cycle will cause the RBLm to be lower than the value along alternate paths.

Observing Figure 51, we can see that the red path will be favoured first as its relay nodes are able to operate at a lower power setting. This will cause less drain on their batteries. However, a point will come when the drain is sufficient that the alternate green path is selected. However, as one of the nodes along this path must operate at high power, the green path is only used for a short time before the red path is again selected.

5.4.2 Result

As stated above, the tool runs with the same setting as stated in Chapter 5.3 except putting a higher weighting to the Residual Battery Life so that only this parameter is taken into account when calculating the fitness score. Thus, there is no impact from the power setting and the end-to-end speed. The outputs are shown as follows:

1st GA run: (Time X)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	28
N2	(28,32)	L	98
N3	(20,1)		30
N4	(35,23)		100
D	(44,44)		100

2nd GA run: (Time X+1)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	26
N2	(28,32)	L	96
N3	(20,1)		30
N4	(35,23)		100
D	(44,44)		100

3rd GA run: (Time X+2)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	24
N2	(28,32)	H	86
N3	(20,1)		30
N4	(35,23)		100
D	(44,44)		100

4th GA run: (Time X+3)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	22
N2	(28,32)	H	76
N3	(20,1)		30
N4	(35,23)		100
D	(44,44)		100

5th GA run: (Time X+4)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)		22
N2	(28,32)		76
N3	(20,1)	H	20
N4	(35,23)	H	90
D	(44,44)		100

6th GA run: (Time X+5)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	20
N2	(28,32)	L	74
N3	(20,1)		20
N4	(35,23)		90
D	(44,44)		100

7th GA run: (Time X+6)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	18
N2	(28,32)	M	69
N3	(20,1)		20
N4	(35,23)		90
D	(44,44)		100

8th GA run: (Time X+7)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	16
N2	(28,32)	L	67
N3	(20,1)		20
N4	(35,23)		90
D	(44,44)		100

9th GA run: (Time X+8)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	14
N2	(28,32)	L	65
N3	(20,1)		20
N4	(35,23)		90
D	(44,44)		100

10th GA run: (Time X+7)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	12
N2	(28,32)	H	55
N3	(20,1)		20
N4	(35,23)		90
D	(44,44)		100

5.4.3 Analysis

As displayed, the tool selects the path that has more Residual Battery Life. The tool has also assigned the power setting according to the fitness function such that communication is

maintained but that energy drain is reduced (to maximise RBLm). This means there is a relation between RBLm and the transmit power. It is essential to allocate the correct power setting in order to preserve battery life. From the output, it is noticed that Node 1 has permanently assigned a low power setting in order to save some battery life whilst maintaining communication. This shows that the tool is behaving the way we expected. On the 5th epoch, the current path flips to the path S-3-4-D, and this time, Node 3 allocated a high (H) power setting. According to our expectation, we might expect it to be allocated a lower power setting on Node 3. However, the link from Node 3 to 4 can only persist if Node 3 operates at high power.

One of the crucial things is that the path stays the same until battery life is reduced significantly compared to another path, according to Figure 51. As the RBLm designed, the fitness value always assigned the lowest RBLm for a specific path/chromosome. Therefore, the current path does not change until the RBLm value has the lowest Battery life in any node of that path. This makes the RBLm in such a way that the battery life decreases quickly for current path.

On the 6th epoch, the path flips again, favouring the red path in terms of the minimum impact on RBLm. Finally, the results confirm that the tool operates as expected given the fitness function configuration.

5.5 Transmit Power Setting (TPS)

This experiment is set up with a topology designed to verify whether the transmit power setting parameter of this tool is working correctly, as shown in Figure 53. Therefore, the parameter is calculated through total WEP described in Chapter 5.2.1.1 for this version of the tool. The power setting has an impact on both the energy drain (and thus the RBL) and the end-to-end speed. The impact on the bitrate is derived from the minimum link speed along the path. The link speed is obtained from the range profile for the given transmit power setting.

This Transmit Power Setting uses data range from Figure 50 that will be utilized to implement this experiment. This data range also shows that there is no end-to-end speed or zero E2E speed if the distance is more than 3 km.

The topology shown in Figure 53 deliberately has Node 1 close to Node 2, but Node 2 is relatively far from the destination D. This arrangement is such that Node 2 must operate on high power (H) to maintain a link to D. However, Node 1 only needs to use low power (L) to communicate with Node 2. Using a higher value confers no benefit and thus wastes energy.

The simulation is configured as follows:

- Number of generations = 100 generations.
- Grid size 50x50 using the topology in Figure 53.
- The chromosome population size is set to 20.
- Fresh chromosomes are injected in every generation to maintain diversity.

5.5.1 Expectation

In this topology, Node 1 placed near to Node 2, which shows the highest link speed but the Source to Node 1 can only operate at the minimum link speed. Similarly, Node 2 to the Destination is at the maximum range and so will provide the minimum link speed even when the power setting of Node 2 is High (H). If our tool is working, then it should show that Node 1 will be assigned a Low (L) power setting as increasing it beyond this value yields no E2E speed benefit but just wastes energy.

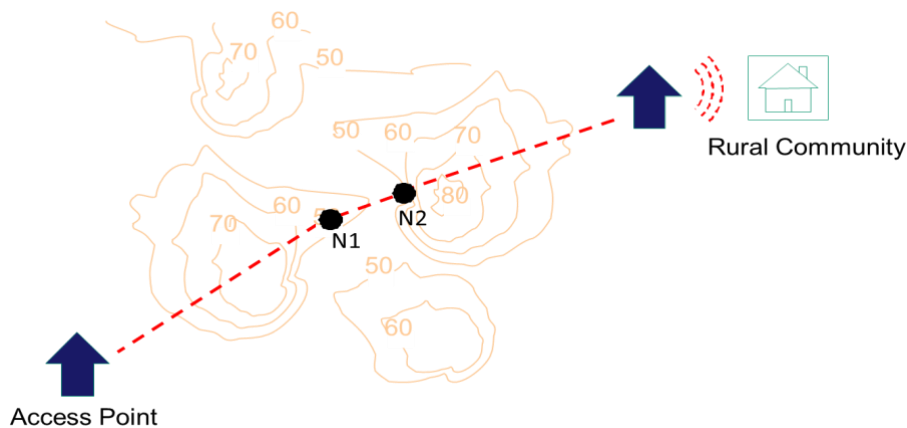


Figure 53: Topology for Evaluating Transmit Power Setting

In this experiment we adjust the Transmit Power Setting weighting which is total WEP to one and set the other two weighting values to zero. Thus, the fitness is based solely on TPS.

5.5.2 Result

For this experiment, we run the tool and obtain the following output (from the best chromosome in the final generation):

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(20,20)	L	98
N2	(30,30)	H	90
D	(49,49)		100

5.5.3 Analysis

As stated, the output clearly shows that Node 2 is at a high-power setting as the tool determines, correctly, that the link is only viable at this setting, operating at 1 Mbps. Moreover, the tool doesn't allow any speed less than 1 Mbps. Another thing to note is that Node 1 has a power setting of L, which is ideal. If Node 1 were to adopt a higher power setting, then the link speed between Node 1 and Node 2 would increase, but the overall E2E speed would remain 1 Mbps. This would be a waste of Residual Battery Life. This shows is functioning as anticipated.

5.6 Combined Objectives

This experiment concentrates on a real-life situation where we consider all three objectives equally when finding a solution. This means the weighting factor of all three

objectives (α, β, γ) set as equivalent to the value of 0.33. This will provide equal preference while calculating the fitness score. the tool settings are similar setting to Chapter 5.3 with the following amendments:

- We run the experiment over 10 epochs.
- The residual battery life of all nodes is set to 100 units initially.

5.6.1 Expectation

The expectation of this experiment doesn't have any definite answer, but we can observe the overall behaviour of the tool and confirm that it functions as expected. However, the equal weighting on fitness value will mean that a compromise between the objectives will be sought in each epoch.

5.6.2 Result

With the setting mentioned in Chapter 5.6, the tool yields the following results:

1st GA run: (Time X)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	98
N2	(28,32)	L	98
N3	(20,1)		100
N4	(35,23)		100
D	(44,44)		100

2nd GA run: (Time X+1)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	96
N2	(28,32)	L	96
N3	(20,1)		100
N4	(35,23)		100
D	(44,44)		100

3rd GA run: (Time X+2)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	94
N2	(28,32)	L	94
N3	(20,1)		100
N4	(35,23)		100
D	(44,44)		100

4th GA run: (Time X+3)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	92
N2	(28,32)	L	92
N3	(20,1)		100
N4	(35,23)		100
D	(44,44)		100

5th GA run: (Time X+4)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)		92
N2	(28,32)		92
N3	(20,1)	H	90
N4	(35,23)	L	98
D	(44,44)		100

6th GA run: (Time X+5)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	90
N2	(28,32)	L	90
N3	(20,1)		90
N4	(35,23)		98
D	(44,44)		100

7th GA run: (Time X+6)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	88
N2	(28,32)	L	88
N3	(20,1)		90
N4	(35,23)		98
D	(44,44)		100

8th GA run: (Time X+7)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	86
N2	(28,32)	L	86
N3	(20,1)		90
N4	(35,23)		98
D	(44,44)		100

9th GA run: (Time X+8)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	84
N2	(28,32)	L	84
N3	(20,1)		90
N4	(35,23)		98
D	(44,44)		100

10th GA run: (Time X+9)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	82
N2	(28,32)	L	82
N3	(20,1)		90
N4	(35,23)		98
D	(44,44)		100

5.6.3 Analysis

When the combination of all three objectives have been used to find the solution, the tool is able to find a compromise between all objectives. The results over time shown in Chapter 5.6.2 show that, at time X, the tool has tried to provide a reasonable E2E speed with a Low (L) power setting to conserve battery life. With a hand calculation, it is clear that the red path in Figure 51 has a higher potential E2E speed when the power setting is High (H). However, considering battery consumption, the tool has opted to be more cautious. Again, from time X+1 to X+3, the path stays the same path to preserve enough battery life to provide satisfactory E2E speed. The tool chooses this to make no change because the E2E speed along both paths is the same, but the red path from Figure 51 consuming less energy through a lower Transmit Power Setting than the green path. At time X+4, the path switches due to higher battery life, and the E2E speed is highest although the power setting of Node 3 must be High (H). Interestingly, Node 4 operates at Low (L) power as increasing it confers no benefit to the overall speed along the green path. The drain on the energy at Node 3 then causes the path to change back to the red one from time X+5 to X+9 and stays the same. The red path is chosen as it is more energy efficient. This shows that the tool is able to find a compromise between all of the objectives and provide a reasonable solution.

5.7 Effect of Weather Evaluation

This tool has the ability to model the effect of weather on the transmitting nodes. At each epoch, the weather conditions at each node can be ascertained the appropriate range profile consulted to determine the expected impact on the link speed. The TE tool thus addresses one of the drawbacks of rural broadband. By effectively checking the weather at regular intervals (epochs) the tool is able to adjust the power setting of nodes and even the selected path in order to maintain a reasonable service (if possible). A scenario has been set up as shown in Figure 54.

Figure 54 shows that the bad weather impacts the network where the red path has been broken due to bad weather, and therefore the green one is the only active path capable of transmitting data. The figure shows that weather has impacted Node 1, and it is unable to transmit data at any power setting.

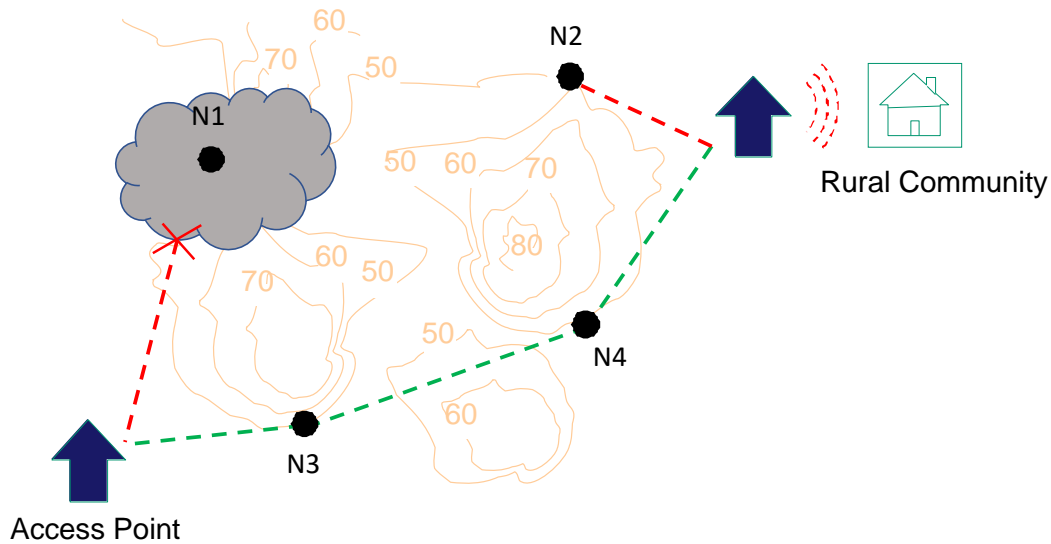


Figure 54: Impact of Weather on Rural Broadband

The basic experimental setup is the same as described in Chapter 5.6. However, in addition a random number generator is used to determine whether the nodes are experiencing any adverse weather effect and, if so, which nodes. The chances of a node experiencing bad weather was set to 60%. This percentage of bad weather was set arbitrarily to allow us to evaluate behaviour of the tool over multiple epochs; the bad weather is assumed to be so severe that the node(s) concerned are unable to transmit data.

5.7.1 Expectation

The topology is configured as shown in Figure 54. This experiment should show that if a node on a specific path is experiencing bad weather, the node becomes inaccessible and so it is unable to transmit data on its links. Therefore, the tool will be forced to choose another path. If the weather improves, the original path may then be again favoured.

5.7.2 Result

The results of the experiment are shown as follows:

1st GA run: (Time X)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	98
N2	(28,32)	L	98
N3	(20,1)		100
N4	(35,23)		100
D	(44,44)		100

2nd GA run: (Time X+1)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	96
N2	(28,32)	L	96
N3	(20,1)		100
N4	(35,23)		100
D	(44,44)		100

3rd GA run: (Time X+2) (bad weather at N1)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)		96
N2	(28,32)		96
N3	(20,1)	H	90
N4	(35,23)	L	98
D	(44,44)		100

4th GA run: (Time X+3)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	94
N2	(28,32)	L	94
N3	(20,1)		90
N4	(35,23)		98
D	(44,44)		100

5th GA run: (Time X+4) (bad weather at N4)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	M	89
N2	(28,32)	M	89
N3	(20,1)		90
N4	(35,23)		98
D	(44,44)		100

6th GA run: (Time X+5)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	87
N2	(28,32)	L	87
N3	(20,1)		90
N4	(35,23)		98
D	(44,44)		100

7th GA run: (Time X+6) (bad weather at N1)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)		87
N2	(28,32)		87
N3	(20,1)	H	80
N4	(35,23)	L	96
D	(44,44)		100

8th GA run: (Time X+7)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	85
N2	(28,32)	L	85
N3	(20,1)		80
N4	(35,23)		96
D	(44,44)		100

9th GA run: (Time X+8)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)	L	83
N2	(28,32)	L	83
N3	(20,1)		80
N4	(35,23)		96
D	(44,44)		100

10th GA run: (Time X+9) (bad weather at N1)

	Coordinate	Power Setting	RBL
S	(0,0)		100
N1	(9,19)		83
N2	(28,32)		83
N3	(20,1)	H	70
N4	(35,23)	L	94
D	(44,44)		100

5.7.3 Analysis

In this experiment, the tool proceeds with similar results that the tool provided in Chapter 5.6 as we use an equal weighting for the fitness function objectives. However, this time we concentrate on whether the tool can divert the path when bad weather blocks the original path. As shown in the results, when a node is affected by the weather, the tool avoids using the impacted path. From the results, in the 3rd epoch Node 1 is experiencing weather. Therefore, the red path from Figure 51 is unavailable, and our tool chooses the alternative path. The same thing happened in the 7th and 10th epochs. In the 5th epoch, although bad weather affects Node 4, as this is not on the transmission path, no change is required. The tool able to identify to divert traffic to alternative paths when appropriate and works as expected.

5.8 Traffic Engineering Tool under MOEA

According to the rural scenario we are considering, the presence of multiple objectives such as E2E speed, the impact of transmit power setting, and residual battery life give rise to multiple optimal solutions based on the importance of these competing performance metrics instead of a single optimal solution. One of these Pareto-optimal solutions cannot be said to be better than the other. This depends on the preferences of the user/operator; it is thus convenient to find many Pareto-optimal solutions to construct a Pareto front.

Our tool can execute and provide a 3-dimensional set of solutions. The problem has three objectives described in section 5.1, and the MOEA is used to find (near) final solutions where all three objectives are considered, and no single objective can be further improved without the detriment of one or more others. Our tool takes the form of an improved and tailored version of the Strength Pareto Evolutionary Algorithm (SPEA-II) where a Pareto front is created of all the non-dominated solutions. It is then up to the operator to determine which of these solutions best meets their needs / preferences at a given time.

5.8.1 Improved Strength Pareto Evolutionary Algorithm

In this research, we employ a modified form of SPEA. The steps involved in the algorithm can be summarised as listed in Figure 55. During initialisation, a primary population of viable solutions is randomly created. A secondary population of non-dominated solutions is created as a null population. Then for each iteration (i.e., generation) the primary and secondary populations are combined to form a temporary population. These are scored in terms of fitness for each objective and then ranked based on dominance. The non-dominated solutions are copied to the revised secondary population. The weakest solutions are culled, and the remainder participate in the mating process to partially complete the new primary population. Population diversity is maintained as random chromosomes are added to the residual portion of the primary population.

In order to design the TE tool under MOEA, an improvement has been made according to the problem we need to address. The improvement has made changes, and some additions to the existing SPEA-II described in Chapter 2.8 in order to maintain a balanced distribution in a population. The flow of the modified SPEA algorithm is as follows:

1. **Initialize:** Generate an initial primary population of N random viable chromosomes, P_p . Create a null secondary population, P_s , that will hold the non-dominated solutions when found. Also, initialize the iteration count i to 0.
2. **Merge:** Combine the primary and secondary populations to create a temporary population P_t . Purge P_s .
3. **Fitness and Rank:** Calculate the fitness of the chromosomes in P_t based on the dominance strength. Rank them based on dominance strength.
4. **Export:** Search and copy the non-dominant solutions into P_s to re-create a secondary population of size N .
 - a. If the number of non-dominated solutions exceeds N , truncate P_s to size N (typically pruning clusters), else if the number of non-dominated solutions is less than N , then copy the best of the dominated solutions to P_s until it is size N .

5. **Termination:** Check the termination condition(s). Typically, this is when the iteration count equals a particular value i_{max} . If $i < i_{max}$ then increment i and continue, else terminate.
6. **Cull:** Remove the weakest chromosomes from P_t (typically 30%) and place the remainder in the mating pool.
7. **Select and Mate:** Select chromosome pairs from the mating pool (using tournament selection etc.) and potentially invoke crossover and mutation (based on chosen crossover and mutation rates) to determine child chromosomes that form part of the new primary population P_p .
8. **Random Injection:** Once the new primary population is partially full (typically 80%), make up the remainder with fresh randomly generated viable chromosomes till P_p is size N . Then return to **Step 2**.

A flowchart of the process is provided in Figure 55. The advantage of this approach is that multiple objectives are used. In a given chromosome, each objective receives its own fitness score and then the chromosomes are ranked in terms of dominance. In our case three objectives are considered. These are the E2E speed along a potential path, the residual battery life of the “weakest node” along the path, and the TPSI which is a measure of the wasted excess power that is being used. These objectives are conflicting, as we are trying to maximize the bitrate whilst limiting the power needed and also taking into account the residual life of the nodes.

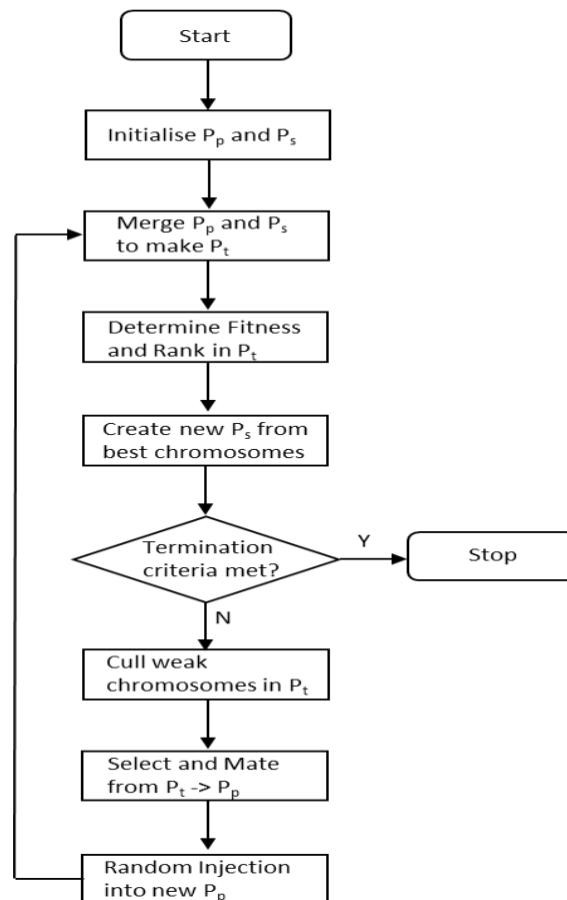


Figure 55: Flow Diagram of the Traffic Engineering MOEA

5.8.1.1 Culling

According to Figure 55, the algorithm culls weak chromosomes from the temporary population. This means the algorithm removes a percentage of the worst chromosomes from the temporary population. This is done by sorting chromosome according to their dominance rank and the weakest ones are then deleted. Typically, 30% of the chromosomes are culled.

5.8.1.2 Dijkstra Algorithm

As mentioned earlier, this TE tool is a modified version of SPEA-II. The computational method employs a similar process to the NP tool. This tool has been used to discover the best possible path after considering all the objectives. However, the TE tool does not employ Dijkstra Algorithm as the NP already has already employed this to help determine where to locate the relay nodes such that end-to-end speed is achievable.

5.8.1.3 Crossover and Mutation

The chromosome crossover is done by selecting a particular point of parent 1 and parent 2 and swapping their TPS. This swap does not happen according to traditional crossover. Similarly, the mutation does not take place as a traditional mutation process would. The process of mutation only happens by changing any particular node's TPS.

5.8.2 Final Flowchart

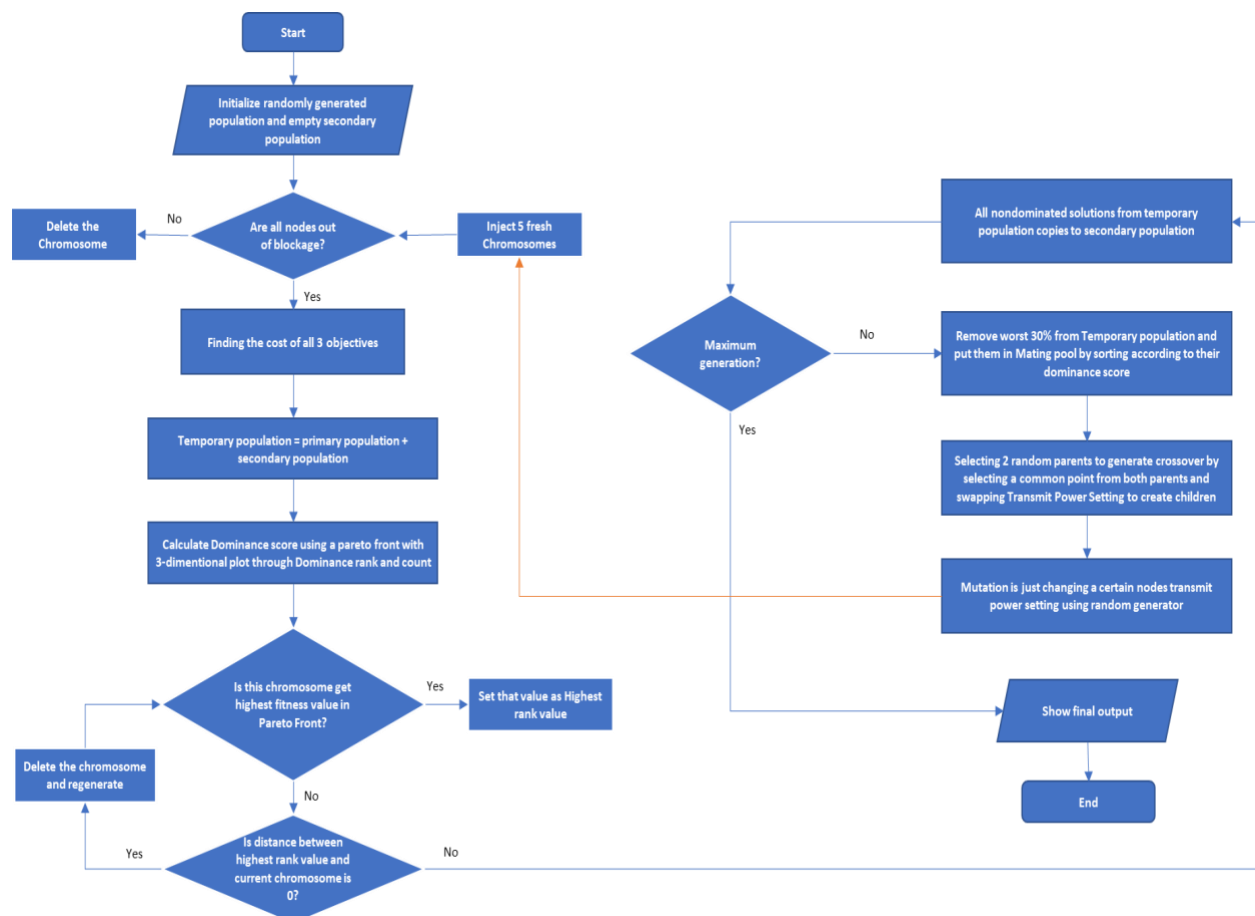


Figure 56: Final Process Flowchart of the Improved SPEA

5.8.3 Design Parameters

In this study, we start by assuming a network planning tool has already selected the location of the relay nodes [127], as illustrated in Figure 51. These relay nodes are considered to employ a combination of renewable and battery power sources. Each relay is equipped with several transceivers that support FSO communication. Furthermore, the power of these transceivers is selectable, with greater power allowing for a higher transmission bitrate at a given range. The Internet Access Point (i.e., Point of Presence) is considered to be connected to the high-speed Internet and has access to a limitless power source. As such, its transmission links can always transmit at maximum power. Communication is via line-of-sight optical links, and the overall aim is to maintain a high-speed path between the Internet Access Point (IAP) and the Rural Access Point (RAP). This can be achieved by a combination of path selection and the transmit power of individual links along the path.

Three parameters are currently taken into consideration when deciding the preferred path and power configuration. These are the:

1. End-to-End speed from the IAP to the RAP
2. Residual battery life of the relay nodes
3. TPSI of the FSO transceivers.

The tool is based on SPEA-II in order to discover the best possible outputs considering all three objectives. This topology has 12 different possible solutions that can be discovered by the tool, or, in other words, data can transmit from Source to destination through 12 different combinations of paths and power settings. Our tool will discover all the non-dominated solutions over these objectives. All viable solutions are shown in Table 12.

Table 12: All Viable Solutions with their Performance Measures

Node 1	Node 2	Node 3	Node 4	RBLm	Energy consumption	E2E Speed
H	H			90	10 (20)	2
H	M			90	10(15)	1.5
H	L			90	10(12)	1
M	H			90	10(15)	1.5
M	M			95	5(10)	1.5
M	L			95	5(7)	1.5
L	H			90	10(12)	1
L	M			95	5(7)	1
L	L			98	2(4)	1
		H	H	90	10(20)	1
		H	M	90	10(15)	1
		H	L	90	10(12)	1

Table 12 shows all the possible chromosomes that exist and should thus appear over multiple generations. Residual Battery Life (RBL) depends on the power settings in an individual chromosome. Regardless of RBL and current circumstances, the energy

consumption provides a predictor of the final or non-dominated solutions. In addition to energy consumption, total RBL provided together with RBLm which is calculated through summation of both node's RBL based on their TPS. Through analysis of Table 12, it can be predicted that the best possible solution could be the chromosome with either power setting of Node 1 (M) and Node 2(M) or Node 1 (M) and Node 2(L). One of the considerable solutions is possibly Node 1 (M) and Node 2 (L) though the final solution depends on operator preference.

One question that arises is why Node 3 and Node 4 have got only three viable path/power setting alternatives. This is because Node 3 is only able to communicate when it employs a high-power setting. Any other power setting results in a non-viable E2E speed which is not permissible in the tool. Power setting is a significant factor that has an impact on all parameters. However, RBL is also essential. The non-dominated solutions might change over time as the RBL changes.

The tool runs with our defined parameters and configuration. However, in this experiment, the essential factor is to discover whether the tool can explore all possible paths/power configurations throughout its run over many generations. Understandably, some solutions will be dominated by others. However, throughout the simulation, given the choices are limited, it should have explored all of those 12 viable solutions confirming that tool is correctly finding all the feasible solutions and then identifying the best ones.

5.8.4 Implementation

Our TE tool is based on SPEA-II in order to discover the best possible downstream path considering three parameters: end-to-end speed, residual battery life of the relay nodes, and wasted excess energy resulting from unnecessarily transmitting at too high a power. The structure of the tool is given in Chapter 5.8.1 and it was implemented in Python.

To confirm the validity of the tool, we start by considering the scenario shown in Figure 51, comprising four battery-powered FSO relay nodes (1...4) in addition to the mains powered source (S) and destination (D). In this topology, in accordance with the range profile shown in Figure 50, only 12 viable solutions exist, as listed in Table 13. The lowest Residual Battery Life along the path (RBLm) assumes that all nodes start with a battery charge of 100 units, and then the per-node RBL is calculated based on the transmit power setting. The RBLm is then identified. The value is shown in generic units. The WEP is calculated in accordance with Eqn 3 and is simply listed as an integer value, the higher, the more wasted energy. Finally, the E2E speed is the bitrate of the path expressed in Mbps in accordance with the range profile.

Table 13: All Viable Solutions with their Performance Measures with WEP

Node Number				RBLm	WEP	E2E Speed
1	2	3	4			
H	H			90	0	2
H	M			90	1	1.5
H	L			90	2	1
M	H			90	1	1.5
M	M			95	0	1.5
M	L			95	1	1
L	H			90	2	1
L	M			95	1	1
L	L			98	0	1
		H	H	90	2	1
		H	M	90	1	1
		H	L	90	0	1

From Table 13, it can be seemed that only two paths exist but that different transmit power setting Impactor are possible. However, as seen from the S-3-4-D path, Node 3 must operate at high power for the link to be operational and thus allow for the overall path to function.

5.8.5 Output

The configuration we used to perform this verification experiment is as follows:





- a) Number of generations = 100
- b) Topology grid size = 45 x 45; each grid square represents an area of 100m x 100m
- c) Range = 30 squares (3km)
- d) New chromosomes injected (per iteration) = 5
- e) Population size = 20

Once the tool is run with the above configuration, it provides output as expected, which is the following:

5.8.5.1 Colour Coded Solutions

A colour code system has been adopted to provide a clear description of chromosome representation to illustrate the Pareto Set (PS). The colour code describes how an individual chromosome performs on the PS. Also, these solutions are marked in different colours based on how good the solution is. In order to differentiate between the various solutions, we use four colour codes to show how good the solutions are as shown in Table 14.

Table 14: Different Solution Colour Codes

-  These chromosomes are superior by less than 3 other chromosomes
-  These chromosomes are non-dominated solutions, and they dominate to all other chromosomes
-  These chromosomes are dominated by non-dominated solutions
-  These chromosomes are dominated by all other chromosomes

The red-coloured chromosomes are primarily non-dominated solutions. These solutions are superior or uniform to all other solutions by at least two objectives. The Blue coloured chromosomes are defined as practical solutions, but they are not as worthy as red coloured. These solutions are superior, mainly non-dominated solutions (red) in terms of our objectives. The green ones are dominated by only those that are non-dominated solutions, typically red-coloured chromosomes. All chromosomes dominate the brown ones, which are worse in all objectives than the others.

5.8.5.2 Generation 1

Table 15: Tabular Form of Unique Solutions in Generation 1

Generation 1									
Chromosome No	Source	Node 1	Node 2	Destination	E2E Speed	RBLm	TPSI	Dominance Rank	Dominance Count
1	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(L)(1.0)	(44,44)	1	98	1	7	0
2	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(H)(2.0)	(44,44)	2	90	1	7	0
3	(0,0)(H)(2.0)	(9,19)(M)(1.5)	(27,32)(M)(1.5)	(44,44)	1.5	95	1	7	0
4	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(H)(2.0)	(44,44)	1	90	0.33	2	39
5	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(M)(1.5)	(44,44)	1	95	0.5	4	21
6	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(L)(1.0)	(44,44)	1	90	1	6	21
7	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(M)(1.5)	(44,44)	1	90	0.5	4	35
8	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(M)(1.5)	(44,44)	1.5	90	0.5	4	21

Generation 1 has 20 chromosomes in its temporary population as this is the initial generation, but only seven of them are used to plot the PS as they are unique. These chromosomes are the different solutions for this generation; in other words, these chromosomes have distinct characteristics. In consideration of superior compromises between all three objectives, chromosomes 1, 2 and 3 are non-dominated solutions (at this time). These non-dominated solutions are identified in red in Figure 57 and form the PF.

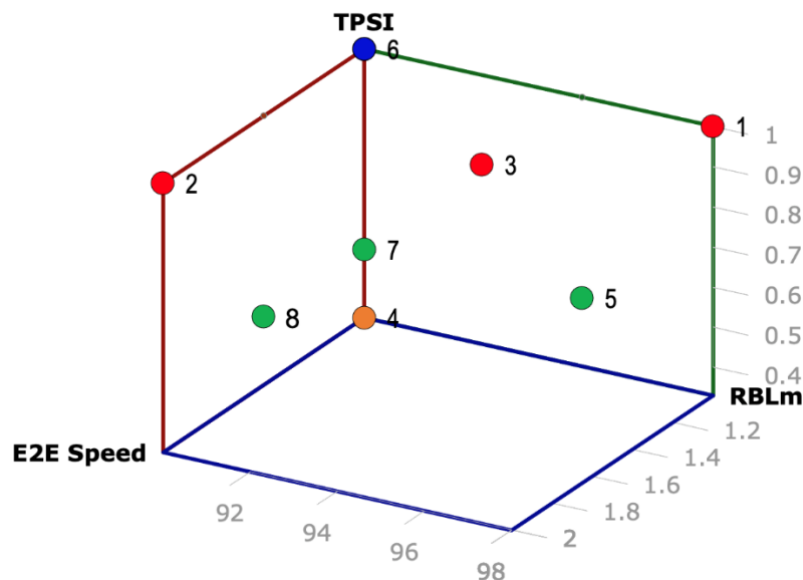


Figure 57: Solutions Found in Generation 1

From Figure 57, it is also seen those chromosomes 1, 2 and 3 are non-dominated. This leads to all other chromosomes has been superior by at least one chromosome. According to Table 15, chromosomes that have a zero (0) value of dominance count/fitness are the non-

dominated solutions. Again, the rule the tool uses to decide if a chromosome is dominated is to see if any other chromosome superiors by at least two objectives.

Chromosome 1 from Table 15 has the lowest E2E speed but better/similar to best RBLm than others and best TPSI. Again, Chromosome 2 might seem viable as it has the highest E2E speed and better TPSI, but it has the worst RBLm. Chromosome 6 also represents similar aspects, but it has the worst values in all objectives, so it is not one of the preferred outcomes. The final non-dominated solution, Chromosome 3 has the right balance of all objectives where it has decent E2E speed, RBLm and best TPSI. There is one more condition that the tool tests, that is if one objective of a chromosome is similar to another one, then whether that chromosome is dominating the other one. Chromosomes 1, 2 and 3 are dominating all chromosome by at least two objectives.

According to Table 14, chromosomes 1, 2 and 3 represent non-dominated solutions as they are primarily dominating all other chromosomes. According to the current problem, these are the best solutions, and therefore, they are highlighted in red. The blue colour chromosomes represent almost the best solution as they are almost superior to all other chromosomes. Similarly, chromosome 4 is marked as brown as this is the worst solution and is dominated by approximately all different chromosomes. Chromosomes 5, 7 and 8 are green as they are superior by half of the chromosomes in PF though they are better than brown ones.

Therefore, Chromosomes no 1, 2 and 3 represent the non-dominated solution in PF, and they represent the best solutions to our problem.

5.8.5.3 Generation 2

Table 16: Tabular Form of Unique Solutions in Generation 2

Generation 2									
Chromosome No	Source	Node 1	Node 2	Destination	E2E Speed	RBLm	TPSI	Dominance Rank	Dominance Count
1	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(L)(1.0)	(44,44)	1	98	1	7	0
2	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(H)(2.0)	(44,44)	2	90	1	7	0
3	(0,0)(H)(2.0)	(9,19)(M)(1.5)	(27,32)(M)(1.5)	(44,44)	1.5	95	1	7	0
4	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(H)(2.0)	(44,44)	1	90	0.33	2	39
5	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(M)(1.5)	(44,44)	1	95	0.5	4	21
6	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(M)(1.5)	(44,44)	1.5	90	0.5	4	21
7	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(L)(1.0)	(44,44)	1	90	1	6	21
8	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(M)(1.5)	(44,44)	1	90	0.5	4	35

Generation 2 acquires 40 chromosomes in the temporary population, but only eight are unique. These chromosomes are the distinct viable solutions for that generation; in other words, these chromosomes have unique fitness values. In consideration of all three objectives, chromosomes 1, 2 and 3 are the non-dominated solutions. These non-dominated solutions form the PF, which is shown in Figure 58.

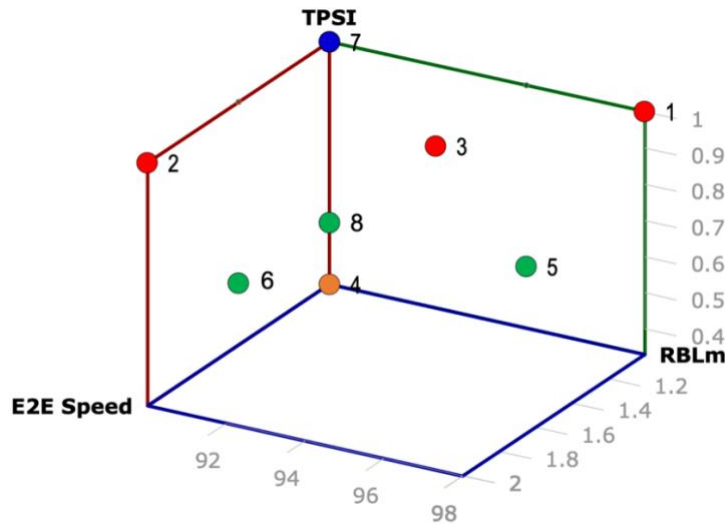


Figure 58: Solutions Found in Generation 2

All the chromosomes are colour coded according to Table 14, which represents how good the solutions are. Red chromosomes are the non-dominated solutions, and those are chromosomes 1, 2 & 3. Chromosome 4, in brown, is the worst solution in this scenario. Chromosome 7 is blue coloured and is a better solution. Finally, green chromosomes are superior primarily by better solutions though they superior other solutions. These are chromosomes 5, 6 and 8.

A visualisation of the chromosomes in generation 2 is given in Figure 58. Chromosomes 1, 2 and 3 dominate all other chromosomes and represent the non-dominated solutions. By analysing the PF, the non-dominated chromosomes are dominating a similar way as generation 1. Therefore, the final superior output is chromosomes 1, 2 and 3, as expected.

5.8.5.4 Generation 3

Table 17: Tabular Form of Unique Solutions in Generation 3

Generation 3									
Chromosome No	Source	Node 1	Node 2	Destination	E2E Speed	RBLm	TPSI	Dominance Rank	Dominance Count
1	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(H)(2.0)	(44,44)	2	90	1	7	0
2	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(L)(1.0)	(44,44)	1	98	1	7	0
3	(0,0)(H)(2.0)	(9,19)(M)(1.5)	(27,32)(M)(1.5)	(44,44)	1.5	95	1	7	0
4	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(H)(2.0)	(44,44)	1	90	0.33	2	39
5	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(M)(1.5)	(44,44)	1	90	0.5	4	21
6	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(M)(1.5)	(44,44)	1.5	90	0.5	4	21
7	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(L)(1.0)	(44,44)	1	90	1	6	21
8	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(M)(1.5)	(44,44)	1	95	0.5	4	35

Generation 3 has 40 chromosomes in the temporary population, but only eight of them are unique. These chromosomes represent the viable solutions for that generation. It is also identified that chromosomes 1 and 2 from Table 17 have better E2E speed, TPSI and RBLm. Again, Chromosome 3 might seem desirable as it has a decent E2E speed, very good RBLm and TPSI. The non-dominated solutions forming the PF. are identified in red as shown in Figure 59.

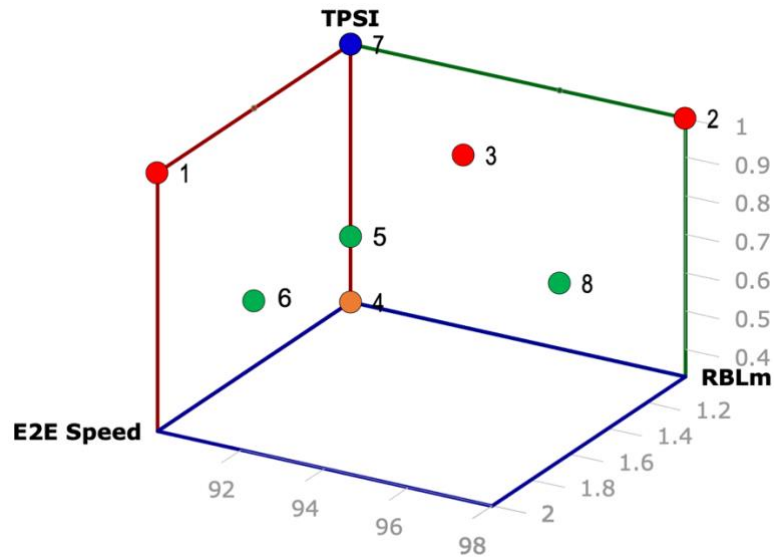


Figure 59: Solutions Found in Generation 3

In Figure 59 chromosomes 1, 2 and 3 dominate all other chromosomes. It is noticed that feasible chromosomes become more dominated over evolving generations. The final best output is represented by chromosomes 1, 2 and 3.

5.8.6 Analysis

In Chapter 5.8.5, the output from the temporary population of the tool is outlined, and these three generations reveal all the chromosomes that have been used to create Pareto Set (PS). The creation of PS shows the non-dominated solutions called PF from the temporary population, which are shown in Figure 57-Figure 59. Preserving these non-dominated solutions is crucial to finding a potentially better result. Therefore, there is one more experiment needed, which will confirm that our tool has considered all possible solutions before choosing the final output. This is not possible with complex scenarios; however, in this instance the task is feasible as the number of possible solutions is rather small.

We combine the data from all the generations to create a new table, Table 18, showing all the solutions found. This table represents each path/power setting selected, and in what generation they were first discovered.

Table 18: All Theoretically Viable Paths/Settings Discovered over All Generations

Node 1	Node 2	Node 3	Node 4	Discover in Generation
H	H			Generation 1
H	M			Generation 1
H	L			Generation 4
M	H			Generation 4
M	M			Generation 1
M	L			Generation 4
L	H			Generation 1
L	M			Generation 1
L	L			Generation 1
		H	H	Generation 2
		H	M	Generation 1
		H	L	Generation 1

The power setting plays a crucial role when considering E2E speed and RBL. The chosen paths/power settings in Table 18 show that the tool has the ability to locate all chromosomes in the Pareto front while evolving over a number of generations to better solutions (typically four generations in this case). The tool has the ability to eliminate duplicate outputs based on the three objectives using the crowding distance (CD) mechanism in order to avoid any premature convergence while developing the PF. This is essential to ensure good non-dominated solutions are not lost and diversity is preserved.

It is notable that the secondary population in Figure 57, Figure 58 and Figure 59, only contain eight distinct chromosomes used for plotting the PS despite the temporary population comprising 40 chromosomes. These are determined through their objectives value according to Table 19. Although there are actually 12 viable solutions, in terms of the three performance objectives some of them have the same scores. Thus, the CD mechanism will prevent these performance duplicates from appearing in the secondary population. Even so, the tool is able to discover every single unique chromosome, which is shown in Table 18 and confirms the tool is working as expected.

Table 19 has been built considering the E2E speed and RBLm, taking into account the node TPS values, and the TPSI. TPS affects the E2E speed, with reference to the range profile, and what it will be the RBL for the next period. Another thing to note is that the path formed via Node 3 and Node 4 exists only if Node 3 is at High TPS.

Table 19: All Theoretical Solutions with their Objectives Values

Node 1	Node 2	Node 3	Node 4	E2E Speed	RBLm	TPSI
H	H			2	90	1
H	M			1.5	90	0.5
H	L			1	90	0.33
M	H			1.5	90	0.5
M	M			1.5	95	1
M	L			1	95	0.5
L	H			1	90	0.33
L	M			1	95	0.5
L	L			1	98	1
		H	H	1	90	0.33
		H	M	1	90	0.5
		H	L	1	90	1

Table 19 shows three sets of chromosomes marked purple, yellow and green. These have similar objectives values, which means only three out of seven chromosomes (purple, yellow and green) can stay in the secondary population. This does not have any impact regardless of what chromosome the tool selects between these similar chromosomes. Therefore, the number of the unique chromosome will be eight, according to theory and our expectation. The results of the experiment are shown in Figure 60.

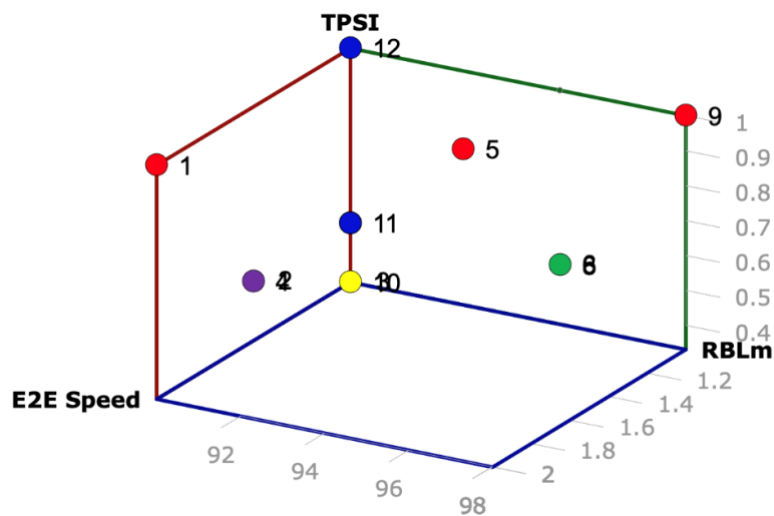


Figure 60: Feasible Solutions and the Pareto Set

Although there are 12 viable solutions, several of them have the same score/fitness in the three parameters we consider, so they are coincident. The red dots show the dominated solutions in the PS, and the red dots are the final Pareto Front. The results are in complete agreement with the theoretical predictions. It is also worth noting that the red non-dominated solutions represent the best compromises between the three parameters of interest (i.e., E2E speed, TPSI and RBLm) depending on preferences.

5.8.7 Final Generation Pareto Set

This experiment is to observe the visual representation of how the Pareto Set looks with the data provided by the tool. The Pareto Front provides us with all the non-dominated solutions found, preserving the best chromosomes in every iteration. It is an essential process for the MOEA, where the final outputs are all the non-dominated solutions in the final generation. This is another validation process for our tool. Tabulated and graphical results taken from the final generation are shown in Table 20 and Figure 61, respectively.

Table 20: Secondary Population in the Final Generation

Generation 100									
Chromosome No	Source	Node 1	Node 2	Destination	E2E Speed	RBLm	TPSI	Dominance Rank	Dominance Count
1	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(H)(2.0)	(44,44)	2	90	1	7	0
2	(0,0)(H)(2.0)	(9,19)(M)(1.5)	(27,32)(M)(1.5)	(44,44)	1.5	95	1	7	0
3	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(L)(1.0)	(44,44)	1	98	1	7	0
4	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(H)(2.0)	(44,44)	1	90	0.33	2	39
5	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(M)(1.5)	(44,44)	1	95	0.5	4	21
6	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(M)(1.5)	(44,44)	1	90	0.5	4	35
7	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(M)(1.5)	(44,44)	1.5	90	0.5	4	21
8	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(L)(1.0)	(44,44)	1	90	1	6	21

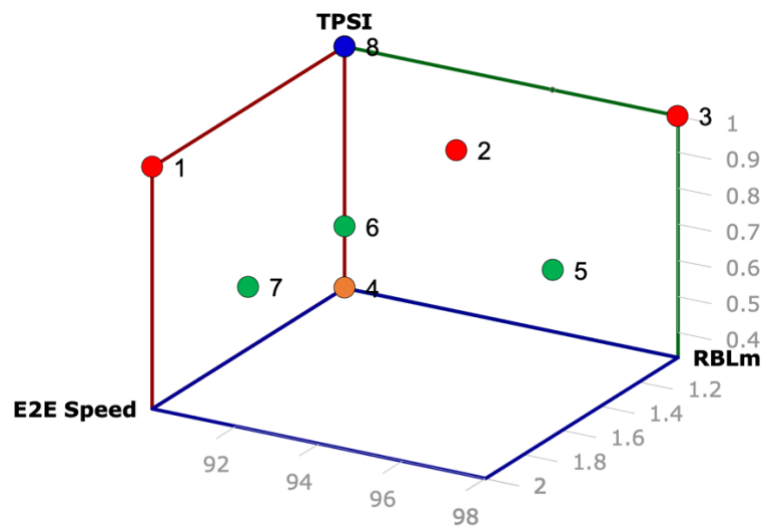


Figure 61: Solutions in the Final Generation

According to the Pareto Set, the three objectives have been compared and the non-dominated solutions are shown in red. By analysing the PF in Figure 61, chromosome 1 has excellent, in fact, the best E2E speed but with the highest TPSI value and the worst RBLm compared to all other chromosomes. Chromosome 3 has the lowest E2E speed compared to other chromosomes but has the best RBLm and TPSI. Conversely, chromosome 2 has a balanced E2E speed and RBLm with the best TPSI. It is noticed that feasible chromosomes typically become more dominated over evolving generations. A few non-dominated solutions tend to supersede the solutions found in early generations, as the MOEA has evolved to find the optimal results. The final best outputs are chromosome 1, 2 and 3.

5.8.8 Final Output Evaluation

This evaluation is based on a comparison between the final generation (100th generation) and all possible solutions shown in Table 21. This comparison is to confirm the satisfactory functioning of the tool. This experiment primarily identifies whether the tool can deliver the expected outcome from a simple and straightforward network where the best results can be manually calculated. This is essential before we use the tool and present it with more compiled, realistic situations.

All possible solutions are presented in Table 21, where the value of all the performance objectives.

Table 21: All Possible Solutions with their Objective Values

Node 1	Node 2	Node 3	Node 4	E2E Speed	RBLm	TPSI
H	H			2	90	1
H	M			1.5	90	0.5
H	L			1	90	0.33
M	H			1.5	90	0.5
M	M			1.5	95	1
M	L			1	95	0.5
L	H			1	90	0.33
L	M			1	95	0.5
L	L			1	98	1
		H	H	1	90	0.33
		H	M	1	90	0.5
		H	L	1	90	1

Please note that the TPS might contrast depending on the network topology. However, Table 21 is particularly useful on the current simple topology outlined in Figure 51, where we know the possible solutions. These possible solutions can be computed manually and find a solution. Again, Node 3 only exists when that is in High TPS, and therefore the TPS with Node 3 and Node 4 is different. So, the comparison between the PS of all unique points from Table 21 and the solution of generation 100 in Table 20 will provide further clarification. The identical unique solutions from Table 20 exist in Table 21. Again, the comparison between Figure 61 and Figure 62 is showing a similar case where both PS is highly identical. The unique set of chromosomes from all possible solutions that can be used to build a Pareto Set is shown in Table 22.

Table 22: Unique Chromosomes from all Possible Solutions

Chromosome no	Node 1	Node 2	Node 3	Node 4	E2E Speed	RBLm	TPSI
1	H	H			2	90	1
2	H	M			1.5	90	0.5
3	H	L			1	90	0.33
4	M	M			1.5	95	1
5	L	M			1	95	0.5
6	L	L			1	98	1
7			H	M	1	90	0.5
8			H	L	1	90	1

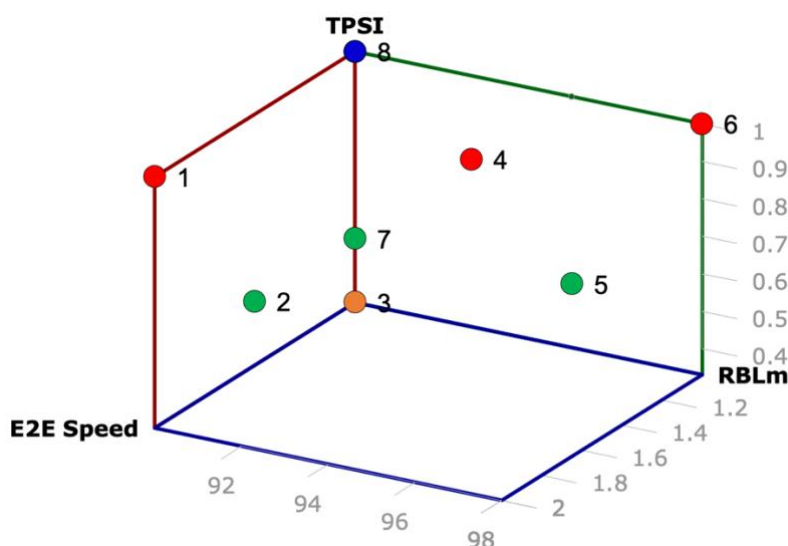


Figure 62: Graphical Representation of all Possible Solutions

Comparison of the PS result from Figure 61 and the PS of the final generation of Figure 62 shows that both have arrived at the same eight possible solutions. Although the tool has identified 12 attainable solutions, only 8 of them are unique in terms of the values of the objectives. As expected, the CD process eliminates duplicate chromosomes where the scores are identical. Thus, our tool is able to find every possible solution.

5.8.9 Simulation versus Theoretical Solutions

In Table 22, there are eight unique solutions intrigued in PS regardless of the non-dominated solution marked in “red”. These solutions should have appeared on simulation PF while evolving the number of generations. From Figure 60, the PF also shows few solutions have overlapped by other solutions. Those are outlined in Table 19, and the paths have been overlapped each other marked in yellow, green and purple. Again, these points are overlapped by a number of chromosomes where similar value are matched. Those chromosomes are overlapped are marked with similar colour in Table 19. In addition to that, Figure 60 shows their corresponding overlapped solutions in PS with their colour.

In comparison to Figure 61, we have a similar set of solutions except chromosome 6 in our final generation PS. Again, those solutions overlap with different TPSI; only one chromosome with specific TPS can survive in the final generation. Currently, the tool does not have any control over which chromosome will survive between those with duplicate objective values.

Table 23: All Feasible Solutions from the Simulation

Node 1	Node 2	Node 3	Node 4	Discover in Generation	E2E Speed	RBLm	TPSI
H	H			Generation 1	2	90	1
H	M			Generation 1	1.5	90	0.5
H	L			Generation 4	1	90	0.33
M	H			Generation 4	1.5	90	0.5
M	M			Generation 1	1.5	95	1
M	L			Generation 4	1	95	0.5
L	H			Generation 1	1	90	0.33
L	M			Generation 1	1	95	0.5
L	L			Generation 1	1	98	1
		H	H	Generation 2	1	90	0.33
		H	M	Generation 1	1	90	0.5
		H	L	Generation 1	1	90	1

In this instance, the tool has discovered all feasible solutions by generation 4, as shown in Table 12.

5.8.10 Performance Evaluation

Hitherto, a number of experiments have been conducted to verify the MOEA scheme is operating appropriately. These experiments actually provide solutions for a rural community taking into account the end-to-end speed, Minimum Residual Battery Life (RBLm) and Transmit Power Setting Impactor (TPSI). The tool currently takes as inputs the location of the relay nodes placed between two fixed points and a transmission range profile. The range profile is such that the further apart are two adjacent nodes that represent a point-to-point link, the lower will be the transmission speed. Beyond a specific range, we assume no communication is possible.

5.8.10.1 New Topographical Test

We now turn our attention to a more complex rural topology as shown in Figure 63. In this instance there are seven relay nodes and the possible downstream FSO links are indicated accordingly. As FSO technology is used, it is essential to have a line of sight (LOS) between nodes. The access point, S, is a typical base station where a broadband Internet point of presence is situated. The seven relay nodes provide several paths through which broadband services can be delivered to the rural community, located at destination node, D.

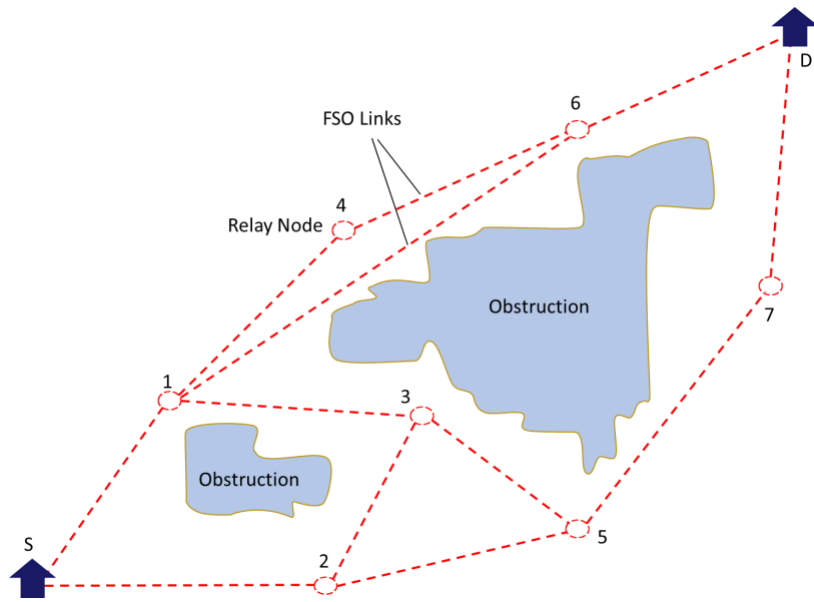


Figure 63: A Complex Rural Topology to Provide Broadband

All nodes transmit data at a particular link speed based on the distance between two points and the transmit power. This link speed is determined by a given range profile which, in this instance is shown in Figure 65. The final E2E speed of a path is determined by the slowest link speed along it.

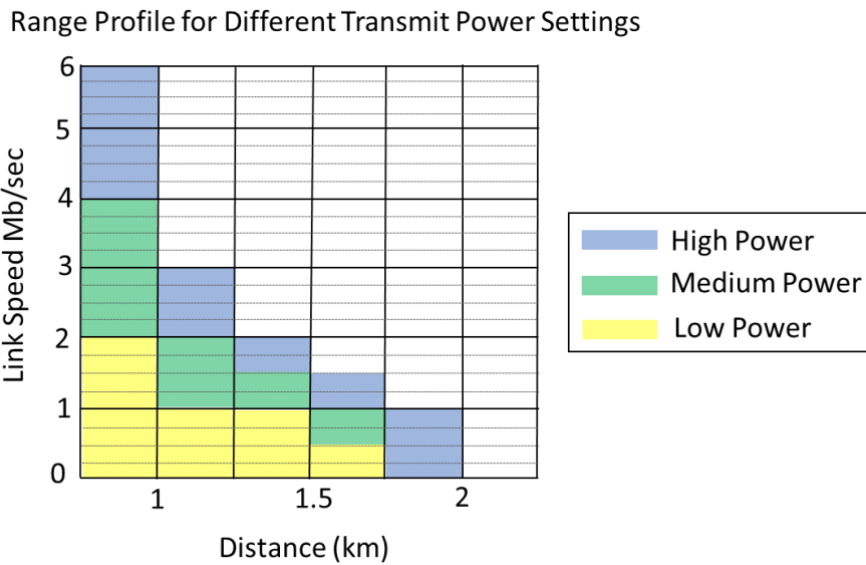


Figure 64: Performance Evaluation Range Profile

The tool is set with the following parameter values:

- Number of generations = 100
- New chromosome injected (per iteration) = 5
- Population size = 40
- Initial residual battery life (all relay nodes) = 100
- Topology grid size = 45 x 45; each grid square represents an area of 100m x 100m

The topology is shown in Figure 65 where the blue-coloured coordinates represent obstructions. Here, the seven relay nodes provide multiple path alternatives along with different power setting combinations. According to our objectives and fitness values, we show the path consisting of the green nodes is selected by the tool as the best transmission path in terms of overall speed. The red marked nodes represent points along alternate paths that are not selected in this instance. However, depending on the RBL of the nodes, this situation can change in order to increase the lifetime of the network.

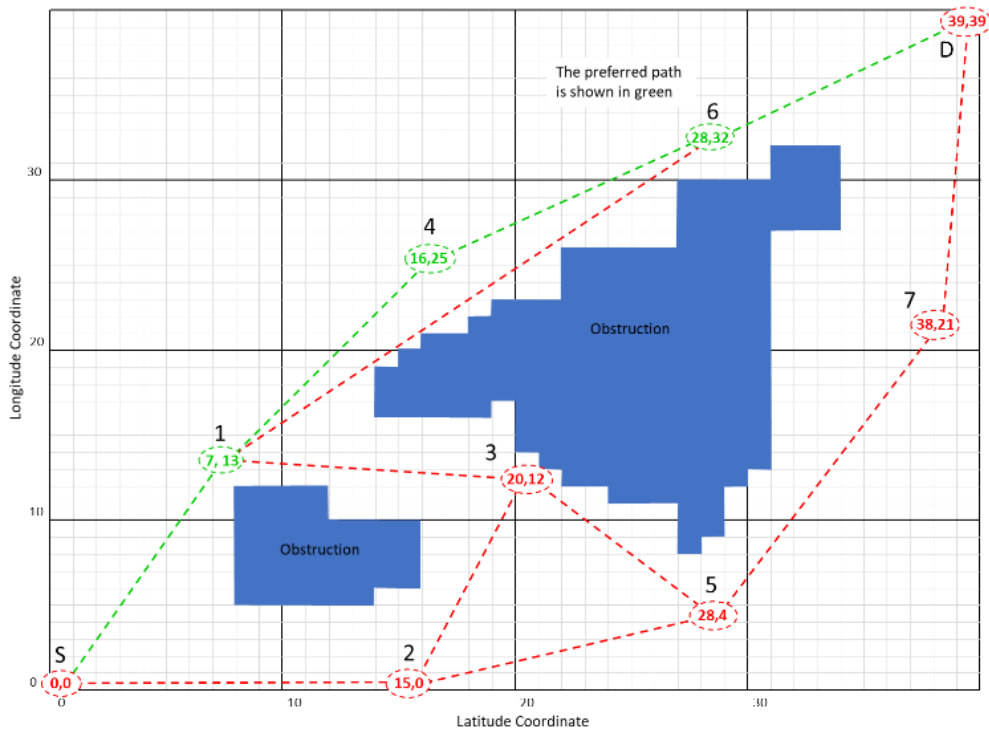


Figure 65: Final Output after 100 Generations

Table 24 shows the temporary population of the final generation and provides an indication of how the tool has decided the most appropriate transmission path. This table shows how the solutions evolve over generations to create a temporary population and provide diverse solutions. The temporary population is shown to understand how Pareto Front (PF) built where a number of diverse paths exist, as shown in Figure 65. Previously, the tool was considering all solutions in PS move to secondary population. According to Figure 51, only two paths exist, and therefore, the PS was designed with only eight solutions and thus, the secondary population was displayed.

Table 24: Tabular Output of the Temporary Population

Chromosome No	Source	Node 1	Node 2	Node 3	Node 4			E2E Speed	RBL	TPS
1	(0,0)(H)(6.0)	(15,0)(M)(4.5)	(20,12)(H)(6.0)	(28,4)(H)(4.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.17
2	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(L)(3.0)	(28,4)(H)(4.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.2
3	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(M)(3.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.17
4	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(L)(2.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.2
5	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(H)(4.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.14
6	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(H)(4.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.14
7	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(M)(3.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.17
8	(0,0)(H)(6.0)	(15,0)(M)(4.5)	(20,12)(H)(6.0)	(28,4)(H)(4.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.17
9	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(M)(3.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.17
10	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(L)(2.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.2
11	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(H)(4.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.14
12	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(M)(4.5)	(28,4)(H)(4.0)	(38,21)(M)(3.0)	(39,39)		3	90	0.2
13	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(M)(4.5)	(28,4)(H)(4.0)	(38,21)(M)(3.0)	(39,39)		3	90	0.2
14	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(H)(4.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.14
15	(0,0)(H)(6.0)	(15,0)(M)(4.5)	(20,12)(H)(6.0)	(28,4)(M)(3.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.2
16	(0,0)(H)(6.0)	(15,0)(L)(3.0)	(28,4)(M)(3.0)	(38,21)(L)(2.0)	(39,39)			2	95	0.5
17	(0,0)(H)(6.0)	(15,0)(M)(4.5)	(28,4)(H)(4.0)	(38,21)(L)(2.0)	(39,39)			2	90	0.25
18	(0,0)(H)(6.0)	(7,13)(L)(3.0)	(16,25)(L)(3.0)	(28,32)(M)(4.5)	(39,39)			3	95	0.5
19	(0,0)(H)(6.0)	(15,0)(L)(3.0)	(28,4)(L)(2.0)	(38,21)(H)(4.0)	(39,39)			2	90	0.33
20	(0,0)(H)(6.0)	(15,0)(M)(4.5)	(20,12)(H)(6.0)	(28,4)(M)(3.0)	(38,21)(H)(4.0)	(39,39)		3	90	0.2
21	(0,0)(H)(6.0)	(7,13)(M)(4.5)	(16,25)(M)(4.5)	(28,32)(M)(4.5)	(39,39)			4.5	95	1
22	(0,0)(H)(6.0)	(7,13)(H)(6.0)	(16,25)(H)(6.0)	(28,32)(H)(6.0)	(39,39)			6	90	1
23	(0,0)(H)(6.0)	(7,13)(L)(3.0)	(16,25)(L)(3.0)	(28,32)(L)(3.0)	(39,39)			3	98	1
24	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(H)(4.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.14
25	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(L)(2.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.2
26	(0,0)(H)(6.0)	(15,0)(M)(4.5)	(20,12)(M)(4.5)	(28,4)(H)(4.0)	(38,21)(M)(3.0)	(39,39)		3	90	0.25
27	(0,0)(H)(6.0)	(15,0)(L)(3.0)	(20,12)(H)(6.0)	(28,4)(L)(2.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.33
28	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(M)(4.5)	(28,4)(H)(4.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.17
29	(0,0)(H)(6.0)	(15,0)(L)(3.0)	(20,12)(H)(6.0)	(28,4)(M)(3.0)	(38,21)(L)(2.0)	(39,39)		2	90	0.25
30	(0,0)(H)(6.0)	(7,13)(L)(3.0)	(16,25)(L)(3.0)	(28,32)(M)(4.5)	(39,39)			3	95	0.5
31	(0,0)(H)(6.0)	(15,0)(L)(3.0)	(20,12)(L)(3.0)	(28,4)(M)(3.0)	(38,21)(L)(2.0)	(39,39)		2	95	0.5
32	(0,0)(H)(6.0)	(7,13)(M)(4.5)	(16,25)(H)(6.0)	(28,32)(H)(6.0)	(39,39)			4.5	90	0.33
33	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(28,4)(H)(4.0)	(38,21)(H)(4.0)	(39,39)			4	90	0.5
34	(0,0)(H)(6.0)	(15,0)(M)(4.5)	(20,12)(L)(3.0)	(28,4)(H)(4.0)	(38,21)(M)(3.0)	(39,39)		3	90	0.33
35	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(H)(4.0)	(38,21)(H)(4.0)	(39,39)		4	90	0.33
36	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(M)(4.5)	(28,4)(H)(4.0)	(38,21)(M)(3.0)	(39,39)		3	90	0.2
37	(0,0)(H)(6.0)	(7,13)(M)(4.5)	(16,25)(M)(4.5)	(28,32)(L)(3.0)	(39,39)			3	95	0.33
38	(0,0)(H)(6.0)	(15,0)(L)(3.0)	(20,12)(L)(3.0)	(28,4)(M)(3.0)	(38,21)(M)(3.0)	(39,39)		3	95	1
39	(0,0)(H)(6.0)	(15,0)(H)(6.0)	(20,12)(H)(6.0)	(28,4)(H)(4.0)	(38,21)(M)(3.0)	(39,39)		3	90	0.17
40	(0,0)(H)(6.0)	(7,13)(M)(4.5)	(16,25)(H)(6.0)	(28,32)(M)(4.5)	(39,39)			4.5	90	0.5

From Table 24, it is hard to determine which chromosomes are better between the 40 chromosomes. Therefore, we need to plot the objective values of the chromosomes to find the non-dominated solutions. The results of the final generation are plotted in Figure 66.

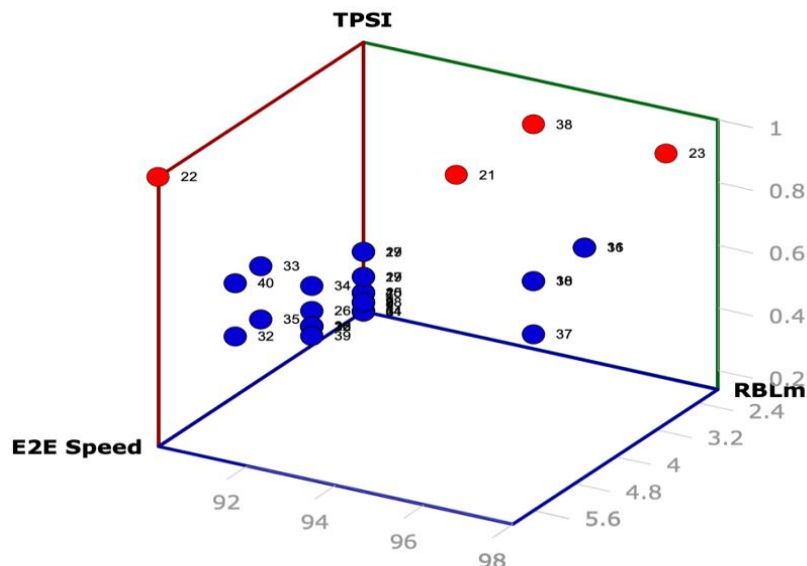


Figure 66: Pareto Set of the Final Generation

The red dots are the non-dominated solutions, which represent the best solutions for this topology based on the parameter preferences. For example, chromosome 21 provides a good compromise between E2E speed and RBLm, whereas chromosome 22 provides the best E2E speed and TPSI but the worst RBLm. Conversely, chromosome 23 has a poor E2E speed but offers the best RBLm and TPSI values, which would extend the lifetime of the network if it was relying on battery power. Notably, chromosome 38 has a relatively poor E2E speed fair RBL but a good TPSI value. It forms part of the Pareto Front (PF) as it is strictly no worse than any of the other solutions. In MOEA, there is typically no single answer that is the best; instead, we are presented with a selection of answers which are the best, based on specific preferences. Finally, it should be noted that there is a good distribution of solutions. This is at least in part because of two features we employ. Firstly, we use random injection of fresh viable chromosomes at the end of each generation to preserve population diversity. Secondly, we use a crowding-distance filter when determining membership of the secondary population.

5.8.10.2 Simulation Time

The tool is meant to be operated on a periodic basis as part of the control plane functionality. Each time it runs, the current environmental conditions and the battery life of the relay nodes are taken into account. The chosen path and transmit power setting at each node are recalculated to ensure the network continues to operate at or close to optimality. Thus, one factor that is important is the run-time of the MOEA to ensure decisions can be made in a timely manner.

In 100 generations situation, the execution time of the simulation was 1 minute, 13 seconds to run 100 generations of the MOEA on a 2.8 GHz Intel Core™ i7 based laptop with 16 GB of SDRAM. Increasing the number of generations to 1000 led to a runtime of 9 minutes, 53 seconds, although there was no discernible benefit in the quality of the solutions found.

We analyse the results after 100, 200, 500 and 1000 generations to see whether any difference can be found while increasing the generations. This experiment will allow us to build confidence in the tool and also help confirm when convergence is reached.

5.8.10.2.1 100 Generations

The first experiment is with 100 generations. Typically, we run all experiments with 100 generations and thus, this output is our base output. This output will be compared with all other outputs with a different number of generations. Hence, the 100-generation output is shown in

Table 25.

Table 25: Final Output with 100 Generations

	Coordinate	Power setting	Leftover
Node 0	(0,0)	()	100
Node 1	(9,19)	(L)	98
Node 2	(27,32)	(L)	98
Node 3	(20,1)	()	100
Node 4	(35,23)	()	100
Node 5	(44,44)	()	100

5.8.10.2.2 200 Generations

This experiment is to re-run the TE tool for 200 generations. After running 200 generations, the final output is shown in Table 26. After comparing with

Table 25, we see no difference between the output of 100 generations and 200 generations. Therefore, it is correct to say that, in this instance, there is no perceived benefit from running the tool for 200 generations.

Table 26: Final Output with 200 Generations

	Coordinate	Power setting	Leftover
Node 0	(0,0)	()	100
Node 1	(9,19)	(L)	98
Node 2	(27,32)	(L)	98
Node 3	(20,1)	()	100
Node 4	(35,23)	()	100
Node 5	(44,44)	()	100

5.8.10.2.3 500 Generations

This experiment is to run the TE tool with 500 generations. After running 500 generations, the final output is shown in Table 27. After comparing with

Table 25 and Table 26, the output is identical.

Table 27: Final Output with 500 Generations

	Coordinate	Power setting	Leftover
Node 0	(0,0)	()	100
Node 1	(9,19)	(L)	98
Node 2	(27,32)	(L)	98
Node 3	(20,1)	()	100
Node 4	(35,23)	()	100
Node 5	(44,44)	()	100

5.8.10.2.4 1000 Generations

This time we increase the number of generations to 1000. This is very likely to be enough to see whether any further changes in the output would appear. After running for 1000 generations, the final output is shown in Table 28. After comparing

Table 25, Table 26 and Table 27, the output is identical except for different power settings. However, the output/path chosen to transmit data is not changed. Therefore, it is correct to say that running the tool for many generations has yielded no discernible benefit.

Table 28: Final Output with 1000 Generations

	Coordinate	Power setting	Leftover
Node 0	(0,0)	()	100
Node 1	(9,19)	(M)	95
Node 2	(27,32)	(M)	95
Node 3	(20,1)	()	100
Node 4	(35,23)	()	100
Node 5	(44,44)	()	100

5.8.10.3 Results Analysis

One of the experiments that need to be addressed in performance evaluation is whether it can identify the best possible outcome for the rural community to provide the best End-to-End speed. Therefore, Table 29 has been created with all possible paths that can transmit data to the rural community.

Table 29: All Possible Paths in the given Topology

No	Possible Path
1	S - 1 - 4 - 6 - D
2	S - 1 - 6 - D
3	S - 1 - 3 - 5 - 7 - D
4	S - 1 - 3 - 2 - 5 - 7 - D
5	S - 1 - 5 - 7 - D
6	S - 2 - 5 - 7 - D
7	S - 2 - 3 - 5 - 7 - D
8	S - 2 - 3 - 1 - 4 - 6 - D

In order to confirm the appropriate functioning of the tool and to better appreciate the choices available to the network operator, we consider all the possible paths along with their E2E speed on the highest and lowest relay node transmit power settings as shown in Table 30 and Table 31, respectively.

Referring to Table 30 we can confirm the tool was able to find the best possible E2E speed by the final generation. As expected, in this instance some or all of the relay nodes would be operating at the maximum TPS setting. The most appropriate solution would be Option 1, yielding 6 Mbps although this consequently has a high energy drain.

Table 30: All Paths with High TPSI and the corresponding E2E Speed

No	Possible Path	E2E Speed	Energy Consumption
1	S - 1 - 4 - 6 - D	6	40
2	S - 1 - 6 - D	1	30
3	S - 1 - 3 - 5 - 7 - D	3	50
4	S - 1 - 3 - 2 - 5 - 7 - D	1	60
5	S - 1 - 5 - 7 - D	1	40
6	S - 2 - 5 - 7 - D	3	40
7	S - 2 - 3 - 5 - 7 - D	3	50
8	S - 2 - 3 - 1 - 4 - 6 - D	6	60

Observing Table 30, it seems paths 1 and 7 have the highest E2E speed. Regarding energy consumption, path 1 has lower energy consumption compared to path 7. It will be interesting to see whether we find at least one path in the final generation. Therefore, we investigated in Table 24 and confirmed that that path appears as Chromosome 22. This ensures that our tool is behaving as expected.

When configuring all the nodes for low TPSI, some links (due to the range) are unable to function and so the E2E speed drops to zero as seen in Table 31. For those paths remaining, Option 1 provides the best E2E speed with the low lowest energy drain. One would therefore expect this path to be chosen unless factors such as poor weather were to interfere with the S-1 link.

Table 31: All Paths with Low TPSI and the Corresponding E2E Speed

No	Possible Path	E2E Speed	Energy Consumption
1	S - 1 - 4 - 6 - D	3	40
2	S - 1 - 6 - D	0	30
3	S - 1 - 3 - 5 - 7 - D	1	50
4	S - 1 - 3 - 2 - 5 - 7 - D	0	60
5	S - 1 - 5 - 7 - D	0	40
6	S - 2 - 5 - 7 - D	1	40
7	S - 2 - 3 - 5 - 7 - D	1	50
8	S - 2 - 3 - 1 - 4 - 6 - D	3	60

Considering a low TPSI illustrates an exciting feature of our tool. According to the range profile in Figure 65, there is zero E2E speed, path 2 path 4 do not exist under low TPS. For high TPSI, path 1 path 7 have the highest E2E speed, and path 1 has more survival chance to a subsequent generation because path 1 has better energy consumption with reasonable E2E speed. So, it will be interesting to observe whether it appears the final generation. Therefore, we examine Table 24 and it is confirmed that it appears in the final generation as chromosome 23. So, this ensures that this tool is behaving as expected.

5.8.10.4 Impact of Node Count

This experiment considers the impact of the execution time on the complexity of the topology. In this instance we grow the size of the network topology in terms of node count whilst maintaining the same Average Node Degree (AND). The AND that has been selected for this experiment is 3, except for the case when the number of nodes is 2. This is because it is impossible to have AND of 3. Also, the number of nodes is excluding the Source and Destination. The experiment runs for 100 generations. After running this experiment, a graph of the execution time is shown in Figure 67.

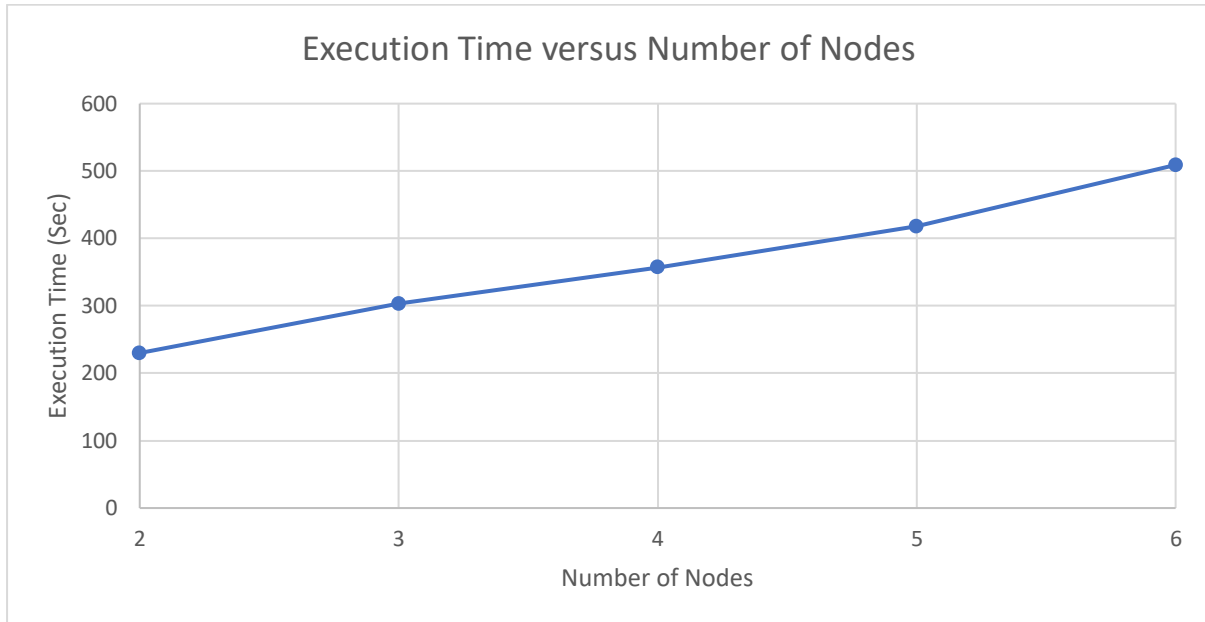


Figure 67: Execution Time versus the Number of Nodes

As expected, as the complexity of the topology increases, in terms of the number of nodes, the execution time increases. The experiment shows that the execution time of the TE tool is roughly linearly increasing while increasing the number of nodes. The performance measured in seconds, and this performance has been calculated by publishing every generation before achieving the final output. This means computation calculation might longer than just printing only final generation output. Despite the increasing number of nodes, the performance was not extremely degraded as diverse path selection is limited to an extent by maintaining the same AND. Again, bidirectional settings of nodes provide additional performance time and worsen performance.

5.8.10.5 Impact of Average Node Degree (AND)

This experiment is performed to discover the extent to which the performance of the tool is affected by the complexity of the topology under consideration. In this instance the number of relay nodes is set to 4, not including the Source and Destination. We then examine the performance of the tool when the AND is varied between 1 and 4. After running this experiment, a graph of the execution time is shown in Figure 68.

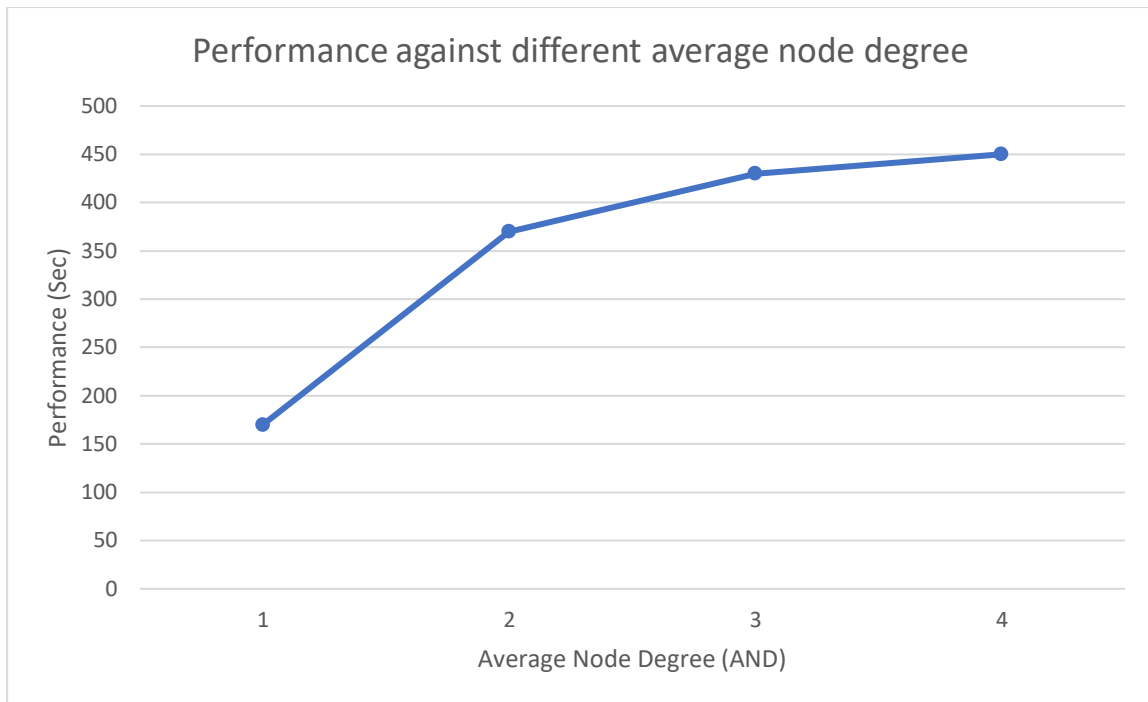


Figure 68: Execution Time versus Average Node Degree

This experiment shows that the AND has a noticeable impact on performance because the TE tool runs longer while increasing node degree. The degree of a node is the number of edges connected to the node. In other words, the number of possible paths increases as AND increases. Running TE tool with different average node degrees provides insight into how our tool provides the output by considering all possible paths. In this experiment, the number of nodes is 4, and the number of generations is 100. Figure 68 shows that the TE tool started taking a longer time in AND of 2 than 1. As AND started increasing after more than three, performance fell linearly to find the final output despite having a similar number of nodes and generations. The reason behind this is TE tool is considering and processing all possible paths and takes longer when average node degree is higher.

5.8.11 Impact of Weather

Weather is the biggest drawback in rural broadband. Therefore, the tool must be capable of taking into account changing conditions.

5.8.11.1 Design Process

Weather is a significant drawback with rural broadband. In some cases, the weather is so severe that the optical link between two nodes becomes inaccessible. Therefore, we need some alternative arrangement to tackle this problem. Our approach is to rerun the TE tool periodically, and at each epoch we select the most appropriate path and power setting for the relay nodes. This selection is based on a suitable balance between TPSI, RBLm and E2E speed given the operator's preferences. Thus, if the weather has a specific impact on one or more links, as illustrated in Figure 69, the algorithm will determine whether an increase of power setting or an alternative path selection, can offer superior service for the next epoch.

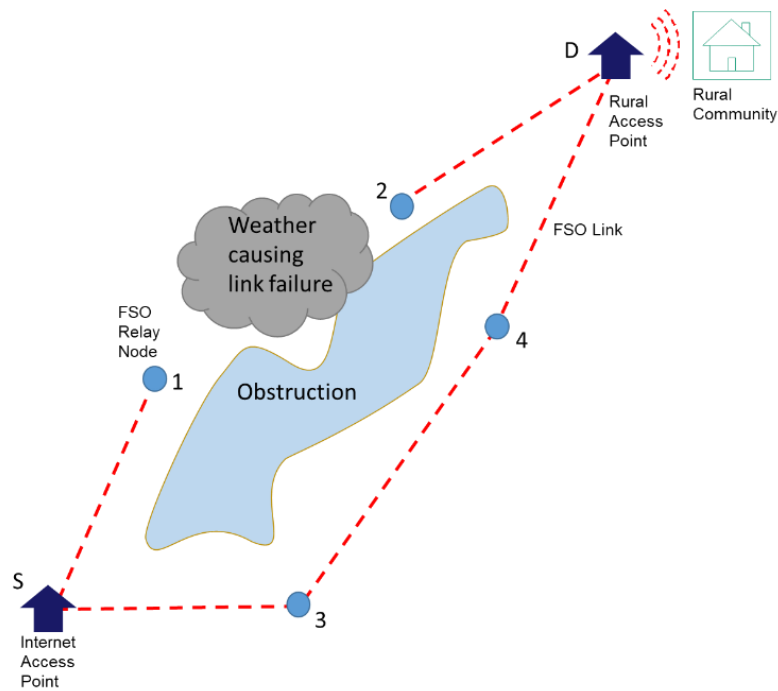


Figure 69: Rural Broadband Impacted by Weather

In some cases, the weather is so severe that the optical link between two nodes fails. Therefore, we need some alternative arrangement to tackle this problem. The current tool has been designed under a range profile to allocate End-to-End (E2E) speed. It also introduces our third objective, Power Setting, with three different power setting. The power setting determines the link speed, but this now takes into account the weather conditions. This power setting also dictates what impact will be on residual battery life. The tool operates as follows:

1. Before allocating any power setting, the first tool checks the weather where each relay node is placed and assigns it one of three weather status markers:
 - a. Clear Weather
 - b. Light Weather
 - c. Bad Weather

Once the tool assigns the weather, it uses a range profile associated with the weather status marker and the power setting to determine the link speed. This means the new range profile will have three different range profiles, shown in Table 32. This allows us to assign an E2E speed for the selected path. Different weather conditions can be applied to different parts of the network over time.

In order to assess the impact on the link performance in different weather conditions, we refine the range profile as shown in Table 32. As the severity of the weather worsens, the link speed is degraded accordingly. This is offset to some extent by increasing the transmit power setting.

Table 32: Range Profile considering the Impact of Weather

Distance (km)	Weather State								
	Clear Weather			Light Weather			Bad Weather		
	High TPS	Medium TPS	Low TPS	High TPS	Medium TPS	Low TPS	High TPS	Medium TPS	Low TPS
0 – 0.5	20	15	10	10	7.5	5	5	3.75	2.5
0.6 – 1	12	9	6	6	4.5	3	3	2.25	1.5
1.1 – 1.5	6	4.5	3	3	2.25	1.5	1.5	1.125	0
1.6 – 2	3	2.25	1.5	1.5	1.125	0	0	0	0
2.1 – 2.5	2	1.5	1	1	0	0	0	0	0
2.6 – 3	1	0	0	0	0	0	0	0	0
> 3	0	0	0	0	0	0	0	0	0

The operation of the TE tool is the same as before in terms of determining the path and transmit power setting of each link; however, the link-speed is derated in accordance with. Non-viable chromosomes are eliminated.

Figure 69 shows a suitable topographical representation for this experiment where the weather has impacted node 1 at a particular time. Therefore, a weather impactor can eliminate network outage due to weather by switching the transmitting path to keep the network alive.

5.8.12 Results

We start by performing several simple weather experiments where the conditions are static. In these initial experiments the scenario is configured thus:

- a) Number of generations = 100
- b) Topology grid size = 45 x 45; each grid square represents an area of 100m x 100m
- c) Range = as per Table 5
- d) New chromosomes injected per iteration = 5
- e) Population size = 20
- f) The potential of poor weather is assigned to only one node in any individual epoch.
- g) The RBL of each node is initialised to 100 units.

The tool is run with three different weather conditions to see the impact of cloud and precipitation on the output. In each instance a single relay node is selected to experience the specific weather effect. All the remaining nodes are assumed to be in clear weather conditions at all times. The chosen topology is as shown in Figure 51 and Figure 69. In each weather scenario after running the tool, the output we obtained from the different weather conditions are as follows:

5.8.12.1 Clear Weather

Clear weather is assumed to be normal conditions. The clear weather range profile in Table 32 provides the link speed based on a node's Transmit Power Setting (TPS) and distance between nodes. This data is used to calculate the best possible path in order to transmit data. The final output is presented in Table 33 showing the path and associated parameter settings. The final output is obtained from the secondary population of the final generation.

Table 33: Final Secondary Population under Clear Weather Conditions

Secondary Population							
Chromosome No	Source	Relay Node i	Relay Node j	Destination	E2E Speed	RBLm	TPSI
1	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(H)(2.0)	(44,44)	1	90	0.33
2	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(L)(1.0)	(44,44)	1	98	1
3	(0,0)(H)(2.0)	(9,19)(M)(1.5)	(27,32)(M)(1.5)	(44,44)	1.5	95	1
4	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(L)(1.0)	(44,44)	1	90	1
5	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(M)(1.5)	(44,44)	1	95	0.5
6	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(M)(1.5)	(44,44)	1.5	90	0.5
7	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(H)(2.0)	(44,44)	2	90	1
8	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(M)(1.5)	(44,44)	1	90	0.5

Figure 70 illustrates the results provided in Table 33 of the viable solutions. We have four non-dominated solutions. These are chromosomes: 2, 3, 4 and 7; however, Chromosome 4 has a high Transmit Power Setting Impactor (TPSI) that is unwanted. Chromosome 2 has a good Minimum Residual Battery Life (RBLm) which means the node will last longer although it does not have the best E2E speed. Chromosome 3 has a reasonable E2E speed and RBLm with excellent TPSI, providing a better compromise between all the objectives. Finally, Chromosome 7 displays the best E2E speed.

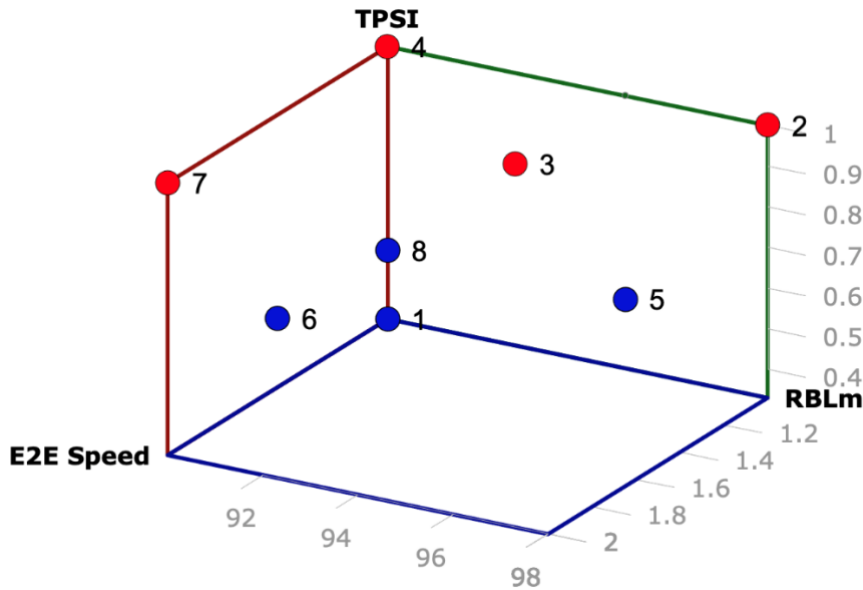


Figure 70: Clear Weather Final Generation Secondary Population

5.8.12.2 Light Weather

We use the term “light weather” to represent situations where low cloud or drizzle has a marginal impact on the FSO link performance due to increased scatter. Light weather reduces link speed for a given TPS and in certain cases, nodes can become entirely inaccessible.

To evaluate the TE tool under light weather conditions we refer back to Table 32 to calculate the derated link speed for Node 4 at location (35,23), assuming all the remaining nodes continue to experience clear conditions. The secondary population of the final generation in this light weather scenario is shown in Table 34.

Table 34: Final Secondary Population under Light Weather Conditions

Secondary Population							
Chromosome No	Source	Relay Node i	Relay Node j	Destination	E2E Speed	RBLm	TPS
1	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(H)(2.0)	(44,44)	1	90	0.33
2	(0,0)(H)(2.0)	(9,19)(M)(1.5)	(27,32)(L)(1.0)	(44,44)	1	95	0.5
3	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(M)(1.5)	(44,44)	1.5	90	0.5
4	(0,0)(H)(2.0)	(9,19)(L)(1.0)	(27,32)(L)(1.0)	(44,44)	1	98	1
5	(0,0)(H)(2.0)	(9,19)(M)(1.5)	(27,32)(M)(1.5)	(44,44)	1.5	95	1
6	(0,0)(H)(2.0)	(9,19)(H)(2.0)	(27,32)(H)(2.0)	(44,44)	2	90	1

Using the results from Table 34, we illustrate the possible solutions in Figure 71. We note that the weather at Node 4 precludes its use even at a high TPS and so the only viable solutions are along the path S-1-2-D. The TE tool thus limits itself to viable solutions avoiding this node and, consequentially there are fewer options than are represented in Figure 71, for the clear weather case.

Figure 71 shows that we have six chromosomes plotted including three non-dominated solutions. These are chromosomes: 4, 5 and 6. Chromosome 4 provides very good Minimum Residual Battery Life (RBLm) which means the node will last longer with a very good (low) TPSI, but it also yields the lowest E2E speed. For this scenario, Chromosome 4 can possibly be chosen though it does not have the best E2E speed. Chromosome 5 has a decent/average E2E speed and RBLm with excellent TPSI. Chromosome 6 also has a good E2E speed and TPSI, but the RBLm is a drawback here. Chromosome 5 thus shows a better compromise between all objectives.

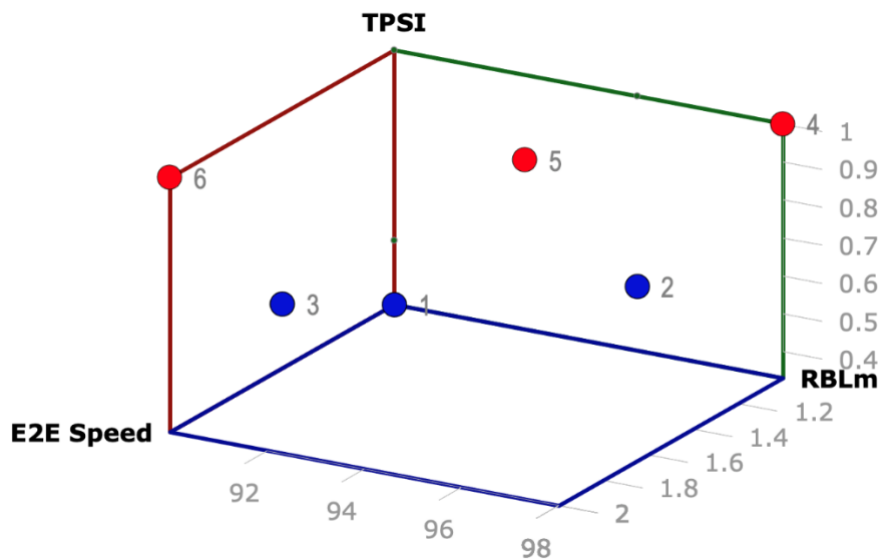


Figure 71: Light Weather Final Generation Secondary Population

5.8.12.3 Bad Weather

In a rural context, we consider bad weather to be severe, harsh, foggy, rainy, or snowy conditions. This is fairly common in a rural community since many are situated in hilly terrain. In cases whether the bad weather is relatively localised, our TE tool can explore the possibility of altering transmit power settings and/or selecting an alternate path to the destination. A

typical example is shown in Figure 69 where the path S-1-2-D is no longer viable due to the conditions.

For this scenario we again use Table 32 to calculate link speeds under the particular conditions where one of the nodes, i.e., Node 1 (27,32), experiences bad weather along its link to Node 2. Apart from the localised “bad” conditions at Node 1, the tool is configured as for the “clear” and “light” weather setup and run. The secondary population of the final generation for this bad weather case is shown in Table 35.

Table 35: Final Secondary Population under Bad Weather Conditions

Secondary Population							
Chromosome No	Source	Node 1	Node 2	Destination	E2E Speed	RBL	TPS
1	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(H)(2.0)	(44,44)	1	90	0.33
2	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(M)(1.5)	(44,44)	1	90	0.5
3	(0,0)(H)(2.0)	(20,1)(H)(1.0)	(35,23)(L)(1.0)	(44,44)	1	90	1

It is notable that the path created via Node 1 is completely inaccessible and the number of viable solutions is severely diminished. The data in Table 35 is plotted in Figure 72.

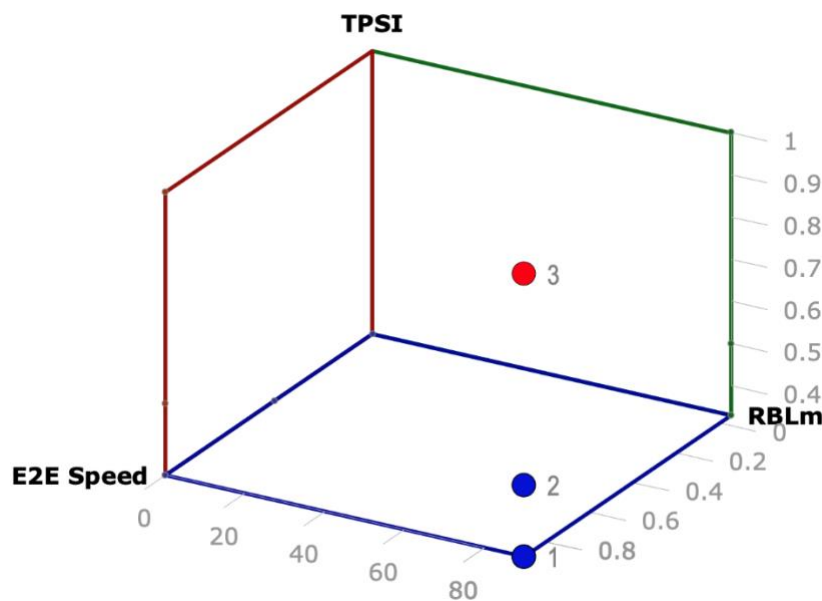


Figure 72: Bad Weather Final Generation Secondary Population

Only three viable chromosomes are plotted, and a single non-dominated solution exists. Chromosome 3 has the same RBLm and E2E speed as the others, but its TPSI is the lowest, and thus would be selected as the path solution to the rural community.

5.8.13 Real-life Scenario

A performance experiment has been set up to provide further clarification in order to find real-life scenarios. This experiment will show how the system evolves with different epochs. One of the real-life scenarios where the tool has to run every 2 hours interval for 12 hours of a specific day. This experiment acts as a real-life scenario where a typical rural network is running under various weather conditions. These weather conditions are influencing on Rural

Broadband network. It is essential to test our tool under a real-life situation where the tool has to work with an intelligent system to divert traffic whenever any effect arises in the overall transmitting network.

There are a number of behaviours of the tool is outlined while running this experiment, and those are below:

- The condition of the weather can be changing over time.
- The RBLm of the current transmitting path will be changed in each simulation.
- The E2E speed and TPSI will also be changing over each simulation.

In order to cover those performances for this experiment, some specific configurations are essential to place on the tool. Therefore, the new tool conditions are the following:

- The tool is able to create a system where the condition of weather can be inputted.
- Each simulation provides a final output which is a current transmitting path for rural broadband. Thus, the RBL of the nodes used for transmitting will be changed according to their TPS. Again, the RBL of transmitting nodes in the final generation of each simulation becomes the starting point for the next simulation.
- Based on the weather effect, the new E2E speed is calculated using Table 32.
- Finally, a new TPS assigns for each node in each simulation.

Again, the tool runs its simulations under a specific configuration which are below:

- Number of generations = 100
- Topology grid size = 45 x 45; each grid square represents an area of 100m x 100m
- Maximum range = 3 km
- New chromosome injected per iteration = 5
- Population size = 20
- Number of epochs = 6; the tool is run every 2 hours of simulated time for 12 hours.

As mentioned above, each simulation runs under different weather condition, and the weather conditions have been selected in this experiment shown in Table 36.

Table 36: Weather Conditions for Successive Epochs

Epoch number	Weather impacted Node	Weather
1	Node 1	Clear
2	Node 3	Light
3	Node 2	Light
4	Node 4	Clear
5	Node 3	Clear
6	Node 3	Light

Please note that the weather condition is not pre-selected, so it has been assigned randomly while running at the start of the simulation. Also, the weather is currently impacting individual nodes, and thus, in some cases, that node becomes inaccessible. It makes that transmitting path inaccessible as well.

5.8.13.1 1st Simulation

Figure 6 shows that the 1st simulation has been set as clear weather and runs the best possible path to provide broadband service to the rural community. The nodes are battery-

powered except for source and destination; the rest of the nodes got full charge as 100 unit. The final output of this simulation is as follows:

Table 37: Final Output of the 1st Simulation

	Coordinate	TPS	RBLm
Node 0	(0,0)	()	100
Node 1	(9,19)	(L)	98
Node 2	(27,32)	(L)	98
Node 3	(20,1)	()	100
Node 4	(35,23)	()	100
Node 5	(44,44)	()	100

Table 37 shows that the tool has chosen a path that consisted of node 1 and node 2. The transmitting path has an E2E Speed of 1 Mbps. This is not the best E2E Speed but considering the best TPSI as one and very good RBLm make this path one of the best solutions. The nodes are energy efficient with low TPS and thus saves perfect battery life.

5.8.13.2 2nd Simulation

According to Figure 69, the 2nd simulation has been set as light weather, which means the network is experiencing a light weather effect. As the weather is getting slightly flawed, E2E Speed and new E2E Speed have been calculated using Table 32. The nodes are battery-powered except source and destination, node 1 & 2 got RBL of 98 units as used in the first simulation, but node 3 & 4 still got 100 units of RBL. The final output of this simulation is as follows:

Table 38: Final Output of the 2nd Simulation

	Coordinate	TPS	RBLm
Node 0	(0,0)	()	100
Node 1	(9,19)	(H)	88
Node 2	(27,32)	(H)	88
Node 3	(20,1)	()	100
Node 4	(35,23)	()	100
Node 5	(44,44)	()	100

Table 38 shows that the simulation has chosen a similar path as the first simulation. Nevertheless, this simulation decided to set higher power to provide broadband service through node 1 & 2. This action provides the E2E Speed of 2 Mbps which is the best E2E Speed possible in this topology. This path might not be energy efficient, but considering the light weather effect, this path has the best TPSI though RBL is not good. As node 3 & 4 is unchanged, node 1 & 2 got 88 unit after simulation. It seems like RBL is in the worst. Despite the worst RBL, the path got the best E2E Speed and TPSI. Thus, the solution seems valid for the rural

broadband problem. Moreover, the light weather is impacting node 3. This weather has put node 3 in limited access, and thus, any transmitting path associated with node 3 will also be impacted.

5.8.13.3 3rd Simulation

For the next 2 hours, the tool has decided on light weather again. According to Table 36, the 3rd simulation has been set as light weather, which means node 2 is experiencing a light weather effect. As the weather is getting slightly flawed, this affects E2E Speed and calculated using Table 32. As The nodes are battery-powered except source and destination, node 1 & 2 got starting RBL of 88 unit as it was ended in the second simulation, but node 3 & 4 still got 100 unit of starting RBL. After running the tool, the final output of this simulation is as follows:

Table 39: Final Output of the 3rd Simulation

	Coordinate	TPS	RBLm
Node 0	(0,0)	()	100
Node 1	(9,19)	()	88
Node 2	(27,32)	()	88
Node 3	(20,1)	(H)	90
Node 4	(35,23)	(L)	98
Node 5	(44,44)	()	100

Table 39 shows that the tool has chosen a path with node 3 & 4 as it has switched the path. As light weather impacts node 2 and lower starting RBL in node 1 & 2 lead to switching the path and choosing that path. The transmitting path has 1 Mbps E2E Speed which has decent bitrate. As the topology designed, node 3 has to be high TPS, and therefore, the path got the best TPSI value, and the best RBL compares to the path consist of node 1 & 2. So, the chosen path seems valid for the rural broadband problem.

5.8.13.4 4th Simulation

For the next 2 hours, the weather has changed to clear weather. According to Table 36, the 4th simulation has been set as clear weather, which means all nodes have no impact on the objectives. As the weather is getting better, this has no effect on E2E Speed and calculated using Table 32. As The nodes are battery-powered except source and destination, node 1 & 2 got starting RBL of 88 unit; node 3 got 90 unit and node 4 got 98 unit as it was ended in the third simulation. The final output of this simulation is as follows:

Table 40: Final Output of the 4th Simulation

	Coordinate	TPS	RBLm
Node 0	(0,0)	()	100
Node 1	(9,19)	(M)	83
Node 2	(27,32)	(M)	83
Node 3	(20,1)	()	90
Node 4	(35,23)	()	98
Node 5	(44,44)	()	100

Table 40 shows that the tool has chosen a path with node 1 & 2 as it has switched the path. As in clear weather and better E2E Speed, this leads to switch the path and choose that path. The transmitting path has 1.5 Mbps E2E Speed which has better E2E Speed than the alternative path. Also, it has the best TPSI value. Despite having worse RBL, the path has better E2E Speed and TPSI. So, the chosen path seems valid for the rural broadband problem.

5.8.13.5 5th Simulation

In the fifth simulation, the weather has stayed to clear weather for the next 2 hours. According to Table 36, the 5th simulation has been set as clear weather, which means all nodes have no impact on the objectives. As the weather stays better, this does not affect E2E Speed and calculated using Table 32. The nodes are battery-powered except source and destination, node 1 & 2 got starting RBL of 83-unit, node 3 & 4 stayed the same as it was ended in the fourth simulation. The final output of this simulation is as follows:

Table 41: Final Output of the 5th Simulation

	Coordinate	TPS	RBLm
Node 0	(0,0)	()	100
Node 1	(9,19)	(L)	81
Node 2	(27,32)	(L)	81
Node 3	(20,1)	()	90
Node 4	(35,23)	()	98
Node 5	(44,44)	()	100

Table 41 shows that the tool has stayed the same as the path chosen for the last interval. The transmitting path has 1 Mbps E2E Speed with decent E2E Speed and similar to an alternative path as in clear weather. Also, it has the best TPSI value. Despite having worse RBL, the path has decent E2E Speed and TPSI. So, the chosen path seems valid for the rural broadband problem.

5.8.13.6 6th Simulation

For the next 2 hours, the weather has changed to light weather as the weather has to change. According to Table 36, the 6th simulation has been set as light weather, which means

all nodes have no impact on the objectives. As the weather becomes better, this does not affect E2E Speed and calculated using Table 32. The nodes are battery-powered except source and destination, node 1 & 2 got starting RBL of 81 unit, node 3 & 4 stayed the same as it was ended in the fifth simulation. The final output of this simulation is as follows:

Table 42: Final Output of the 6th Simulation

	Coordinate	TPS	RBLm
Node 0	(0,0)	()	100
Node 1	(9,19)	()	81
Node 2	(27,32)	()	81
Node 3	(20,1)	(H)	80
Node 4	(35,23)	(L)	96
Node 5	(44,44)	()	100

Table 42 shows that the tool has chosen a path with node 3 & 4 as the tool has switched the path. Despite light weather impacting node 3, it only survives with high TPS, and lower starting RBL leads to switching the path and choosing that path. The transmitting path has 1 Mbps E2E Speed which has decent bitrate. As mentioned, node 3 has to be high TPS, and therefore, the path got the best TPSI value, and the best RBL compares to the path consist of node 1 & 2. So, the chosen path seems valid for the rural broadband problem.

5.8.14 Performance over Multiple Epochs

The true benefit of traffic engineering is the ability to modify the flow of traffic based on changing current conditions in order to maintain a suitable level of service. However, it is useful to quantify the benefit that TE can give to the network. We thus conduct an experiment to assess the efficacy of our approach in an evolving situation where the weather conditions change over time. In this experiment, the aim is to maintain the highest E2E speed regardless of what is happening to the other objectives. Other experiments could give more consideration to factors such as residual battery life to preserve network operations for as long as possible or consider a balanced compromise between these factors; however, we leave these for further study.

We represent time in epochs or snapshots in time when the tool assesses the current situation and decides the most appropriate course of action in response. In this case, our epochs are chosen to be every 2-hours, and we run an experiment across a 12-hour interval. During this time, we introduce changing weather conditions and observe the E2E speed provided by our scheme relative to a system where the configuration is static, being configured in the first epoch and thenceforth remaining unchanged.

With our dynamic scheme, several factors are taken into account at each epoch. These are the revised link speeds due to the weather conditions at each relay node, based on Table 32, as well as the RBLm and TPSI values. It should be noted that the RBLm of the nodes in a given epoch influences the path selection in the next epoch as there will be a reluctance to use nodes

whose residual battery levels are low even if they offer greater E2E speeds. Again, the tool is configured as follows:

- Number of generations = 100
- Topology grid size = 45 x 45; each grid square represents an area of 100m x 100m
- Maximum range = 3 km
- New chromosome injected per iteration = 5
- Population size = 20
- Number of epochs = 6; the tool is run every 2 hours of simulated time for 12 hours.

The weather conditions at each epoch are pre-selected before running the simulation. The parameters of the tool have been selected discretely and used the value in order to make the process much easier to understand the calculations. Also, the weather currently impacts individual nodes, and thus, in some cases, that node becomes inaccessible. This makes any transmission path using this node no longer viable.

We run our tool with the topology shown in Figure 63. This experiment assumes the localized poor weather conditions move as time progresses across epochs. At each epoch, the current conditions are used to determine the most appropriate path and power setting of each node. This can be based on the achievable E2E speed and the RBLm of the various nodes. The balance between these is an operator decision that can be factored into the TE algorithm. In this particular case, we give preference to the E2E speed.

According to Section 5.8.11 and at Epoch 3, the bad weather impacts Node 6, and this bad weather then travels across the terrain such that it reaches Node 5 at Epoch 4. After this, it moves away from the area. We ran the experiment for six epochs with our TE mechanism enabled and disabled. In the disabled case, the path configuration chosen in Epoch 1 remains fixed for the remainder of the experiment.

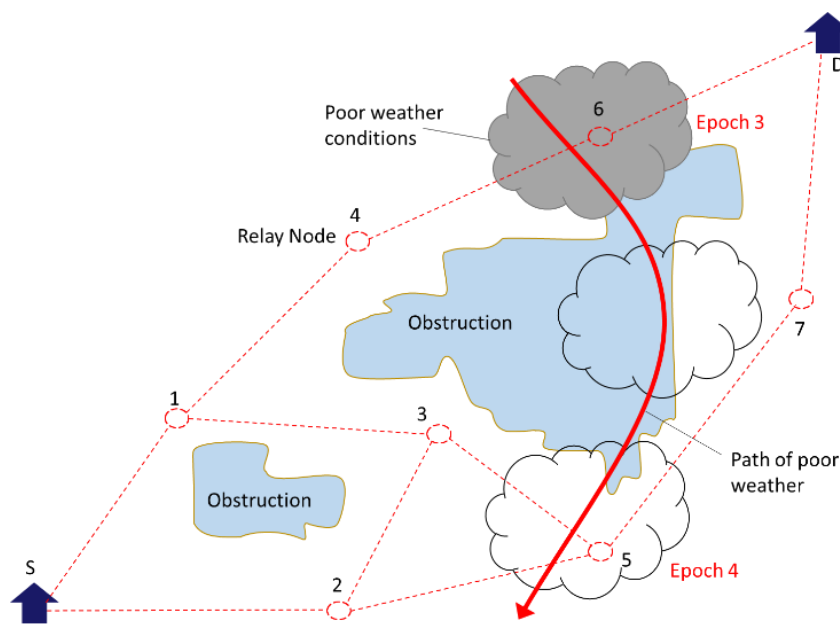


Figure 73: Weather Conditions across Different Epochs

Conversely, when the TE mechanism is enabled, the configuration is recalculated at each epoch based on the current conditions. This means that alternative configurations can be selected if they are superior to the current one. The results are presented in Table 43.

Table 43: E2E Speed Across Multiple Epochs

Epoch Number	Weather	E2E Speed (Mbps)	
		TE Disabled	TE Active
1	Clear	2	2
2	Clear	2	2
3	Bad (at Node 6)	0	2
4	Bad (at Node 5)	0	1.12
5	Clear	2	2
6	Clear	2	2

Table 43 shows that when the TE mechanism is enabled, the system is able to cope with dynamic weather conditions and manages to maintain a transmission path between the Internet point-of-presence (S) and the rural community (D). A visualization of this situation is provided in Figure 74.

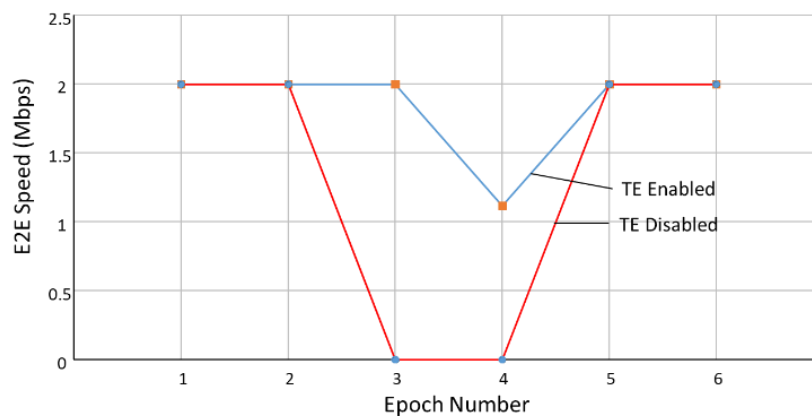


Figure 74: Comparative Performance with TE Mechanism Enabled/Disabled under Dynamic Weather Conditions

From Figure 74, the red line represents the transmission performance without the TE process in place, and the blue one shows with the TE mechanism enabled. The network might have lost the connection on Epoch 3 and 4 without TE. While active TE mechanism allows to keep network alive shown in blue line. We observe that the E2E speed in Epoch 3 and 4 have improved when the TE mechanism is active as the system has adapted to the changing conditions. TE has improved significantly to the network availability and performance as well.

5.9 Discussion

The primary aim of our research was to develop a tool that can provide a reliable FSO transmission system accounting for the changing weather conditions whilst providing connectivity to a rural area. The transmission speed of any free-space optical communication channel is impacted by weather. For example, [128] provides an analysis of the attenuation of an FSO link associated with different weather conditions. Given the dynamic nature of the

weather, it is thus highly desirable to have a TE mechanism that can reconfigure the network to maintain satisfactory operation. In Section 5.8.1 we show that the TE tool is able to find a best possible route to transmit data after considering multiple constraints.

Although FSO has many desirable features, its biggest drawback is its performance in response to various atmospheric effects. This research has created a tool for FSO relay node placement [127] in order to provide rural Internet access, taking into account possible compromises between E2E speed and path diversity. The network-planning tool considers link speed based on different FSO range profiles, along with the E2E Speed and path diversity to identify superior path alternatives. The TE tool supports a maintenance of a data flow which can be reconfigured in order to keep QoS, in terms of E2E speed as high as possible, or to achieve a suitable balanced compromise taking into account the residual battery life of relay nodes. Furthermore, as the tool is typically capable of finding (near)optimal solutions within several minutes, it can be run periodically in order to allow the network to be reconfigured according to changing weather conditions, to ensure availability where possible.

The execution time of the tool on a 2.8 GHz Intel Core™ i7 based laptop is the order of tens of seconds for a typical scenario with several relay nodes between the source and destination access points. This is fast enough to be responsive to changing weather conditions. However, for more complex scenarios with many paths and power settings a reasonable solution is to further restrict the number of generations in the search process. This is acceptable as we observe that good solutions are typically identified quickly.

There are number of researchers addressed different types of network issues. An SDN controller also can be used for TE to improve network utilization and reduce packet loss and delay when the entire network consists of OpenFlow and traditional networks together [129]. Dekker and Colbert [130] investigate the robustness of network topologies using graph-theoretic concepts by assessing robustness according to the traffic levels induced in the network by node failures. However, our TE tool has designed to eliminate the rural broadband drawback such as weather factors. The output from Section 5.2 and 5.8.1 shows that this research has provided robust optimal solutions to rural problem.

One of the key factors in rural broadband is weather. Our tool has considered the weather factor and provided solutions for this problem. However, further research can be conducted as explained on Chapter 6. Figure 74 shows that our tool is able to keep network alive despite weather impacting the relay nodes.

Chapter 6. Conclusions and Future Work

Broadband connections have become indispensable for accessing healthcare, advancing education, improving agriculture, and growing small businesses. Urban areas have become accustomed to ongoing capital investment to expand capacity in regions with broadband coverage. The time has come to extend this coverage to the rural areas that currently lack it entirely. The primary motivation for this research is to explore and facilitate the rollout of backbone infrastructure for rural areas. Most of these rural populations reside in economically backward and geographically isolated communities. This research has developed a tool that can reach these rural communities and provide them with high-speed Internet services by deploying a network planning tool for relay node deployment. This network planning tool will ensure the provision of high-speed Internet infrastructure to these rural communities.

Initially, the research has created a tool for FSO-based relay node placement in order to provide rural Internet access taking into account possible compromises between end-to-end speed and path diversity. The network-planning tool can accommodate obstacles and different FSO range profiles. Although FSO has many desirable features, its biggest drawback is its performance in response to various atmospheric effects. Therefore, our research has addressed the issue of using traffic engineering mechanisms that can offset these effects in order to maintain availability. This introduces further trade-offs in terms of transmission power, end-to-end speed, energy drain and the use of diverse paths.

In terms of network planning tools, they can cover a wide range of platform systems and applications-oriented to the industry. The market for commercial planning tools is aimed at providing a complete set of solutions to design and analyse networks [131] [132] [133]. Another network planning tool that is capable of meeting industrial requirements is introduced in [83]. These are designed for commercial use and are not open source. In comparison, the tools presented in this thesis are based on optimisation techniques that are described in detail and provide considerable scope for enhancement.

This network planning tool is currently confined to the placement of FSO nodes given various performance constraints and a specific topology. It can create a topological diagram and place nodes such as to avoid obstructions whilst ensuring suitable compromises between diversity, and hence resilience, and communication speed. Parameters such as the number of nodes used, their power consumption and anticipated weather conditions have been used to specify suitable node locations. However, this work could be readily extended to include other technologies such as different forms of radio wireless, allowing for hybrid network architectures to be created.

One possible future work is integrating the network planning tool the Geographical Information System (GIS) to import real-world data. This information could provide an insight into what technologies might be best suited to specific geographical areas taking into account the exact location of hills, their elevation etc. More work could also be done on including CapEx and OpEx data into the tool, so the financial impact of different solutions could be explored.

A TE has been designed, implemented and evaluated to provide energy-efficient configurations coupled with the best E2E speed whilst preserving battery life. The energy consumption of nodes could become one of the internet's dominant energy factors as more reliance is placed on combinations of renewable energy sources with battery storage. This research takes a fresh look at the traffic engineering mechanism for rural broadband from energy awareness whilst achieving good E2E speed. Our TE tool has improved the reliability of the network and enhanced the survivability of the network against failures, typically resulting from adverse weather. The TE has improved the network availability in a significant way and using the right technique, and the availability improves to keep the network alive as long as there is an alternative path.

Both in terms of the network planning tool and traffic engineering, optimization is achieved firstly through genetic algorithms. These limit the tools to finding one solution at a time, where multiple factors have to be combined into a weighted fitness function. As such, the result is a restricted view of what is possible. We, therefore, devote much research to the design and implementation of multi-objective genetic mechanisms, where conflicting objectives can be considered and a set of solutions presented that show a range of (near) optimal compromises. This allows the operator to make more informed choices. We also take into account temporal changes in the system due to weather effects or other factors that can impair the operation of relay nodes and provide a means of identifying suitable alternatives.

MOEA has been used to design and implement a network planning tool that considered two significant objectives to allow this tool to provide high-speed broadband access for rural people. Diverse relay node placement and end-to-end speed are the objectives that ensure a workable tool is created. However, the use of an MOEA allows other objectives to be taken into consideration. In [134] a tool called IncrEase is described that upgrades the current mobile tower. Despite attempting to support rural broadband, it does not consider aspects such as weather, terrain, or bitrate. Our MOEA obtains a set of Pareto optimal non-dominated solutions where none of the objective function values can be improved without degrading other objective function values. It is relatively straightforward to add objectives such as financial cost. Our research also could provide the ability to consider a selection of heterogeneous network technologies. This would allow the rural network to comprise both line-of-sight and wireless communication channels, which may permit a more effective solution to be found considering all trade-offs. Different types of FSO and even hybrid technologies could be supported, such as Li-Fi [3], Hybrid FSO/RF [135] etc.

Our research has been extended to design a traffic engineering system to operate over the deployed infrastructure. This tool is designed to maintain services over the network despite changes in operating conditions. Software-Defined Networking (SDN) is an emerging network architecture where network control is decoupled from forwarding and is directly programmable [136]. Although SDN is not used in our work at this time, our TE mechanism could be integrated into an SDN enabled network, allowing for automated reconfiguration of the system based on changing conditions and operator preferences.

In Figure 75, an SDN-enabled reference network is illustrated. A new framework would show how SDN would be used to monitor and manage a rural Internet infrastructure with our

TE mechanism at the heart of the management system. Weather, link BER and RBLm could be sent to the control node from the various relay locations and the network then configured to provide the consumer bitrate most efficiently in terms of the path and power settings. Furthermore, if an outage arises, the system could quickly provide an alternative pathway, should one exist and so restore the service.

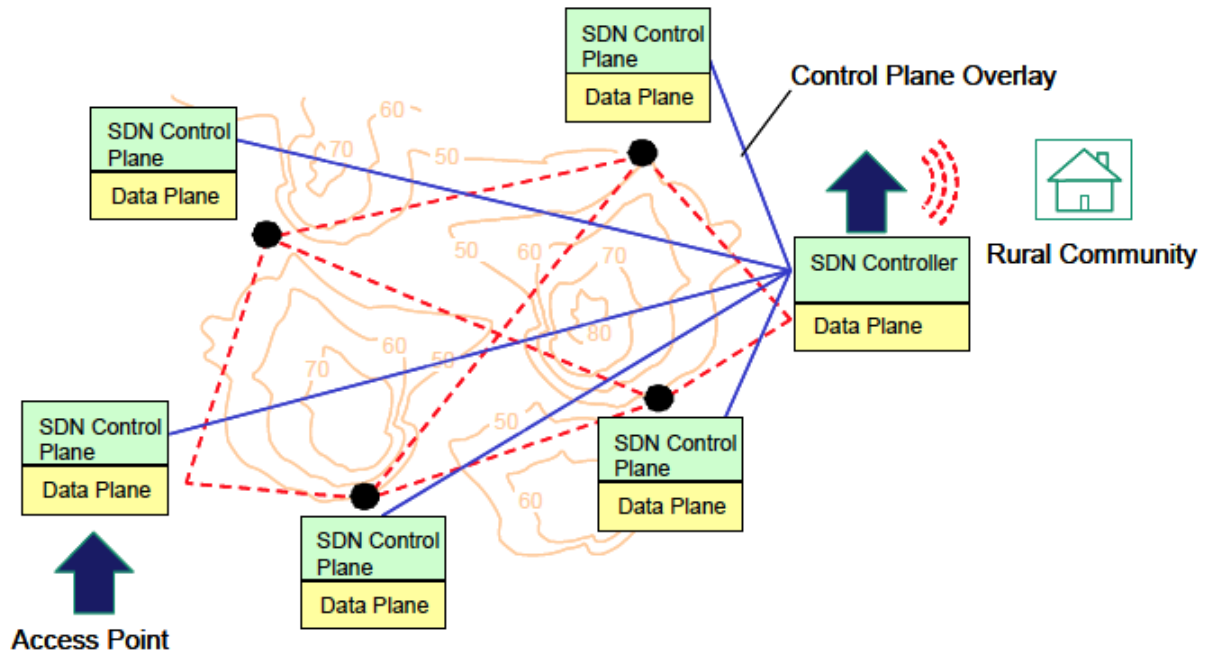


Figure 75: SDN Controlled Network Infrastructure

Another element of the research is the transmission of control messages between nodes. The current TE tool has an assumption of transmitting network management information. We assume that control messages are transmitted between relay nodes. This research can be extended to build a system where the network management information will communicate efficiently and with the right technology.

The current tool concentrates on the impact of weather, and this has been categorised into three different conditions: Clear, Light and Bad. We do not focus on specific atmospheric effects but use a set of range profiles to derate the link speed based on the magnitude of the impairment. However, a more precise system could be considered, taking into account Signal to Noise Ratio (SNR), Bit Error Rate (BER) and Pointing Error (PE) [137] [138] [139] etc., on optical networks. Indeed, a number of studies [140] [141] [142] [143] [144] have examined the atmospheric effect on Free-space Optical communication. Our research does not model atmospheric effects, but our tool could utilise additional information to represent these phenomena and respond more precisely. Furthermore, the line of sight (LOS) modelling in the tool represents the “world” as a two-dimensional grid with the ability to include obstacles. This representation could be enhanced to consider elevation, too.

The TE tool could be extended with some performance estimate-based techniques on the effective transmission bitrate. These techniques include adjustable Forward Error Correction (FEC) [145], adjustable power and multipath error amelioration. These techniques could be used to address specific optical communication problems in order to achieve higher

transmission bitrates. Our tool could be extended to consider these techniques according to the problem description.

Another aspect worth pursuing is the introduction of hysteresis, such that changes to the path are only implemented when there is sufficient justification. This would ensure route flapping is suppressed. Also, the MOEA-based TE tool currently presents a Pareto set of possible solutions. It is left to the operator to select the most suitable one. Going forward, it would be possible to integrate fuzzy logic decision-making into selecting the preferred solution (chromosome) based on the changing conditions. This would be a relatively straightforward means of taking into account various preferences and aspects of the situation in order to make decisions without the need for human involvement.

Finally, this research has given a good solution and possibility of providing high-speed broadband to a rural community. A defined and tested technology, FSO has chosen in this research and the viability of FSO is high regards as this technology is low cost and optics omitted transmission, where physical digging isn't necessary and suitable for remote rural areas. Both NP and TE tools have been designed to eliminate any technology drawback. The NP tool is designed to determine the deployment of relay nodes appropriately to provide broadband access while considering a set of rural topographical constraints. Therefore, the TE tool has considered redirecting the traffic flows across the access infrastructure in order to meet varying needs. This tool also considered and provided a solution for one of the major rural drawbacks, weather. TE tool has been able to find a solution for weather situations which is very common in rural areas. Lastly, both tools together can provide a fair and good broadband solution for rural communities where they can enjoy high-speed broadband.

Chapter 7. References

- [1] J. L. Riding, J. C. Ellershaw, A. V. Tran, L. J. Guan and T. Smith, "Economics of Broadband Access Technologies for Rural areas," IEEE, 2009.
- [2] F. Darbari, M. Brew, S. Weiss and R. W. Stewart, "Practical Aspects of Broadband Access for Rural Communities Using a cost and power efficient Multi-Hop/Relay Network," in IEEE Globecom 2010 workshop on Broadband Wireless Access, 2010.
- [3] P. Kuppusamy, S. Muthuraj and S. Gopinath, "Survey and Challenges of Li-Fi with Comparison of Wi-Fi," in IEEE WiSPNET, 2016.
- [4] G. Soni and N. Singh, "Analysis of Future Fiber-Wireless (FiWi) Networks- EPON & WiMAX," International Journal of New Technology and Research (IJNTR), vol. 1, no. 1, pp. 19-22, 2015.
- [5] A. Haider, R. Rahman, O. Faruk, F. Alam and R. M. Huq, "Towards an IEEE 802.22 (WRAN) based wireless broadband for rural Bangladesh - Antenna design and coverage planning," in Electrical, Computer and Communication Engineering (ECCE), Cox's Bazar, 2017.
- [6] J. Wang, T. Shake, P. Deutsch, A. Coyle and B.-N. Cheng, "Topology management algorithms for large-scale aerial high capacity directional networks," in Military Communications Conference, USA, 2016.
- [7] K. Werbach, "The Development of Fixed Broadband Networks," 08 January 2015. [Online]. Available: <http://dx.doi.org/10.1787/5jz2m5mlb1q2-en>. [Accessed 06 September 2017].
- [8] J. Fang, J. Wu and W. Tian, "On the relationship between Internet and Traditional Communication Industry," Communication and Information Engineering, Changshu University, Jiangsu, China, 2018.
- [9] P. Tsatsou, "Why Internet use? A quantitative examination of the role of everyday life and Internet policy and regulation," Technology in Society, vol. 33, pp. 1-2, 2011.
- [10] J. M. Pedersen and M. T. Riaz, "Bringing fiber to the home to rural areas in Denmark," in Applied Sciences in Biomedical and Communication, 2009.
- [11] Ofcom, "UK fixed-line broadband performance, November 2015: Research Report," 24 March 2016. [Online]. Available: https://www.ofcom.org.uk/__data/assets/pdf_file/0017/71540/fixed-bb-speeds-nov15-report.pdf. [Accessed 8 October 2018].
- [12] H. Samimi and M. Uysal, "End-to-End performance of mixed RF/FSO transmission systems," IEEE Journal of Optical Communication, vol. 5, no. 11, pp. 1139-1144, Nov 2013.

- [13] N. I. Miridakis, M. Matthaïou and G. K. Karagiannidis, "Multiuser relaying over mixed RF/FSO links," *IEEE transaction Communication*, vol. 62, no. 5, pp. 1634-1645, May 2014.
- [14] S. J. Habib and P. N. Marimuthu, "Optimized capacity planning and performance measurement through OPNET Modeler," in *Computer Applications and Industrial Electronics*, Malaysia, 2010.
- [15] A. Erkollar, W. Krug and H. C. Mayr, "Planning in virtual enterprises using the NETSIM approach," in *Technology and Innovation Management*, Portland, 1999.
- [16] L. V. De Souza, J. C. W. A. Costa and C. R. L. Frances, "Integrated Solutions for Broadband Access in Brazilian Amazon Rural Areas," in *Globecom Workshops*, Anaheim, CA, USA, 2012.
- [17] B4RN, "Broadband for the Rural North Ltd," December 2011. [Online]. Available: <http://b4rn.org.uk/>. [Accessed 12 July 2017].
- [18] "Gigaclear," December 2010. [Online]. Available: <http://www.gigaclear.com/gigaclear-story/>. [Accessed 12 July 2017].
- [19] LN Communications Ltd, "Boundless Networks," September 2015. [Online]. Available: <http://www.boundlessnetworks.co.uk/about-us/>. [Accessed 12 July 2017].
- [20] "Village Networks," 2003. [Online]. Available: <http://www.villagenetworks.co.uk/>. [Accessed 12 July 2017].
- [21] M. Zhang and R. S. Wolff, "Using Wi-Fi for Cost-Effective Broadband Wireless Access in Rural and Remote Areas," in *Wireless Communications and Networking Conference*, Georgia, 2004.
- [22] E. Lawler, "Combinatorial Optimization in Networks and Matroids," Holt, Rinehart and Winston, New York, 1976.
- [23] F. Araujo, B. Ribeiro and L. Rodrigues, "A neural network for shortest path computation," *IEEE Transaction Neural Networks*, vol. 12, pp. 1067-1073, 2001.
- [24] C. Ahn and R. Ramakrishna, "A genetic algorithm for shortest path routing problem and the sizing of populations," *IEEE transaction of Evolutionary computation*, vol. 12, pp. 1067-1073, 2002.
- [25] A. W. Mohemmed, N. Saahoo and T. Geok, "Solving Shortest path problem using particle swarm optimization," *Applied Soft Computing*, vol. 8, pp. 1643-1653, 2008.
- [26] A. Osyczka, *Multicriterion optimization in engineering*, Ellis Horwood Series, 1984.
- [27] K. Deb, *Multi-objective optimization using evolutionary algorithms*, John Wiley & Sons, 2001.
- [28] D. E. Goldberg, *Genetic Algorithm in Search, Optimization & Machine Learning*, Addison-Wesley, 2001.

- [29] S. Leue and W. Wei, "Integer Linear Programming-Based Property Checking for Asynchronous Reactive Systems," *IEEE Transactions on Software Engineering*, vol. 39, no. 2, pp. 216-236, 2013.
- [30] M. Lukasiewicz, M. Glab, C. Haubelt, J. Teich, R. Regler and B. Lang, "Concurrent Topology and Routing Optimization in Automotive Network Integration," in *Design Automation Conference*, Anaheim, 2008.
- [31] P. Singh, R. Kumar and P. P. Chakrabarti, "Multiobjective EA Approach for Improved Quality of Solutions for Spanning Tree Problem," in *International Conference on Evolutionary Multi-Criterion Optimization*, Guanajuato, Mexico, 2005.
- [32] J. Uerpmann, D. Ziegenbein and R. Ernst, "Dynamic response time optimization for SDF graphs," in *International Conference on Computer Aided Design*, San Jose, USA, 2000.
- [33] N. Banerjee and R. Kumar, "Multiobjective network design for realistic traffic models," in *Genetic and Evolutionary Computation Conference*, London, UK, 2007.
- [34] A. Konak and A. Smith, "Capacitated Network Design Considering Survivability: An Evolutionary Approach," *Engineering Optimization*, vol. 36, no. 2, pp. 189-205, 2004.
- [35] H. Perros and S. Blake, "Bandwidth Provisioning in ADSL Access Networks," *International Journal of Network Management*, vol. 19, no. 5, pp. 427-444, 2009.
- [36] O. Osterbo, "Capacity Dimensioning for Real Time Services in Access networks," Telenor Group, 2003.
- [37] D. Nowak, P. Perry and J. Murphy, "Bandwidth allocation for service agreement aware Ethernet passive optical networks," in *IEEE Globecom 04*, Dallas, 2004.
- [38] M. Zukerman, T. Neame and R. Addie, "Internet traffic modeling and future technology implications," in *IEEE INFOCOM*, San Francisco, 2003.
- [39] A. Muradova, "Calculation in the NGN networks of indexes of reliability of tracts of transmission of packet information," *The Advanced Science Journal*, vol. 2, pp. 24-28, 2014.
- [40] E. Chromy and M. Kavacky, "Asynchronous Networks and Erlang Formulas," *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 2, no. 2, pp. 85-89, 2010.
- [41] L. Vidmar, B. Peternel, M. Stular and A. Kos, "Broadband Access Network Investment Optimization in Rural Areas," in *15th IEEE Mediterranean Electrotechnical Conference*, Valletta, 2010.
- [42] T. Rokkas, D. Katsianis, D. Varoutas and T. Sphicopoulos, "Free Space Optical Technology as an alternative Last Mile Solution: A Techno-Economic Analysis," in *Telecommunication Techno-Economics CTTE*, Helsinki, 2007.

- [43] N. Cvijetic, M. Cvijetic, M. Huang, E. Ip, Y. Huang and T. Wang, "Terabit Optical Access Networks based on WDM-OFDMA-PON," *Journal of Lightwave Technology*, vol. 30, no. 4, pp. 493-503, 2012.
- [44] S. Sadchikova, M. Abdujapparova and D. Normatova, "Algorithm of optimal technology selection of broadband access network," *Telecommunication Computing Electronics and Control*, vol. 17, no. 5, pp. 2244-2249, 2019.
- [45] Ofcom, "Connected Nations 2020: UK report," 17 December 2020. [Online]. Available: https://www.ofcom.org.uk/__data/assets/pdf_file/0024/209373/connected-nations-2020.pdf. [Accessed 12 February 2021].
- [46] C. Hope, "Better Broadband," 27 March 2017. [Online]. Available: <http://www.telegraph.co.uk/news/2017/03/26/welsh-village-broadband-slow-locals-get-4am-send-emails/>. [Accessed 10 July 2017].
- [47] P. Pavon-Marino and J. L. Izquierdo-Zaragoza, "Net2Plan: An open source network planning tool for bridging the gap between academia and industry," *IEEE Network*, vol. 29, no. 5, pp. 90-96, 2015.
- [48] G. Diviney, "An introduction to short-range wireless data communications," in *Proceedings of the Embedded Systems Conference*, San Francisco, 2003.
- [49] A. Kretsis, P. Kokkinos, K. Christodouloupoulos and E. Varvarigos, "Mantis: Optical Network Planning and Operation Tool," in *15th International Conference on Transparent Optical Networks (ICTON)*, Cartagena, Spain, 2013.
- [50] K. Rusek, J. Suarez-Varela, P. Almasan, P. Barlet-Ros and A. Cabellos-Aparicio, "RouteNet: Leveraging Graph Neural Networks for Network Modeling and Optimization in SDN," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 10, pp. 2260-2270, 2020.
- [51] A. Banerjee, "Towards Technologically and Competitively Neutral Fiber to the Home (FTTH) infrastructure," in *Proceedings of the TPRC conference*, 2003.
- [52] B. Whitacre, R. Gallardo and S. Stover, "Broadband's contribution to economic growth in rural areas:," *Telecommunications Policy*, vol. 38, no. 11, pp. 1011-1023, 2014.
- [53] C. Z. Qiang, C. Rossotto and K. Kimura, "Economic Impacts of Broadband. In *Information and Communications for Development*," World Bank, Washington D.C., 2009.
- [54] Federal Communications Commission, "2018 Broadband Deployment Report," 2018. [Online]. Available: <https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2018-broadband-deployment-report>. [Accessed 31 January 2021].
- [55] C. I. Canfield, O. Egbue, J. Hale and S. Long, "Opportunities and Challenges for Rural Broadband Infrastructure Investment," in *2019 International Annual Conference of the American Society for Engineering Management*, Philadelphia, PA, 2019.

- [56] L. Galloway, "Can broadband access rescue the rural economy?," *Journal of Small Business and Enterprise Development*, vol. 4, no. 14, pp. 641-653, 2007.
- [57] K. L. Alvin Yau, C. J. Perera, H. T. D. Chieng, K. H. Kwong and A. Ting, "Exploring sustainability of the deployment of wireless broadband in Malaysia rural areas: Comprehensive insights," 2011 IEEE Colloquium on Humanities, Science and Engineering, CHUSER 2011, pp. 148-152, 2011.
- [58] G. Taylor, "Remote Rural Broadband Systems in Canada," *Telecommunications Policy*, vol. 42, no. 9, pp. 744-756, 2018.
- [59] G. Ford and T. Koutsky, "Broadband and economic development: A municipal case study from Florida," *Review of Urban and Regional Development Studies*, vol. 17, no. 3, pp. 216-229, 2005.
- [60] H. Hudson, "Bridging the broadband divide: Strategies for rural and developing regions," *Americas*, vol. 35, pp. 1504-1505, 2003.
- [61] T. Turk, B. J. Blazic and P. Trkman, "Factors and sustainable strategies fostering the adoption of broadband communications in an enlarged European Union," *Technological Forecasting and Social Change*, vol. 75, no. 7, pp. 933-951, 2008.
- [62] D. Brake, "A Policymaker's Guide to Rural Broadband Infrastructure," April 2017. [Online]. Available: <http://www2.itif.org/2017-rural-broadband-infrastructure.pdf>. [Accessed 12 February 2021].
- [63] OFCOM, "Ofcom Connected Nations Report, 2018," 18 December 2018. [Online]. Available: <https://www.ofcom.org.uk/research-and-data/multi-sector-research/infrastructure-research/connected-nations-2018>. [Accessed 12 February 2021].
- [64] S. Verbrugge, K. Casier, J. V. Ooteghem and B. Lannoo, "Practical steps in techno-economic evaluation of network deployment planning part 1: methodology overview," in *International Telecommunications Network Strategy and Planning Symposium*, Budapest, Hungary, 2008.
- [65] C. J. Bouras, *Trends in Telecommunications Technologies*, intechweb.org, 2010.
- [66] ZigBee Alliance, "ZigBee specification," 27 June 2013. [Online]. Available: <http://www.zigbee.org/Specifications/ZigBee/FAQ.aspx>. [Accessed 5 September 2017].
- [67] Bluetooth SIG, "Bluetooth qualification and membership fees," Bluetooth SIG, 05 September 2017. [Online]. Available: <https://www.bluetooth.com/develop-with-bluetooth/qualification-listing/qualification-listing-fees>. [Accessed 05 September 2017].
- [68] D. Kravets, "U.N. report declares internet access a human right," 6 March 2011. [Online]. Available: <https://www.wired.com/2011/06/internet-a-human-right/>. [Accessed 29 August 2017].

- [69] M. Jackson, "Ofcom 2016 Report – Average UK Home Broadband Speeds Reach 28.9Mbps," 25 March 2016. [Online]. Available: <http://www.ispreview.co.uk/index.php/2016/03/ofcom-2016-report-average-uk-home-broadband-speeds-reach-28-9mbps.html>. [Accessed 29 August 2017].
- [70] Countryside Alliance, "Options for those struggling with poor broadband in rural areas," 21 April 2015. [Online]. Available: <http://www.countryside-alliance.org/options-for-those-struggling-with-poor-broadband-in-rural-areas/>. [Accessed 7 August 2017].
- [71] R. Flickenger, S. Okay, E. Pietrosemoli, M. Zennaro and C. Fonda, "Very long distance wi-fi networks," second ACM SIGCOMM workshop on Networked systems for developing regions, pp. 1-6, 18 August 2008.
- [72] J. C. Juarez, A. Dwivedi, A. R. Hammons, S. D. Jones, V. Weerackody and R. A. Nichols, "Free-Space Optical Communications for Next-generation Military Networks," IEEE Communications Magazine, vol. 44, no. 11, 2006.
- [73] D.-G. A. Pedro and H. F. Dean, "Initial Population for Genetic Algorithms: A Metric Approach," in International Conference on Genetic and Evolutionary Methods, Las Vegas, USA, 2007.
- [74] Ofcom, "Connected Nations 2019 - UK report," 18 March 2020. [Online]. Available: https://www.ofcom.org.uk/__data/assets/pdf_file/0023/186413/Connected-Nations-2019-UK-final.pdf. [Accessed 19 January 2021].
- [75] D. Kohn, "Providing global broadband internet access using low-earth-orbit satellites," Computer Networks and ISDN Systems, vol. 29, no. 15, pp. 1763-1768, 1997.
- [76] I. d. Portillo, B. G. Cameron and E. F. Crawley, "A technical comparison of three low earth orbit satellite constellation systems to provide global broadband," Acta Astronautica, vol. 159, pp. 123-135, 2019.
- [77] J. Porter, "Amazon will launch thousands of satellites to provide internet around the world," 4 April 2019. [Online]. Available: <https://www.theverge.com/2019/4/4/18295310/amazon-project-kuiper-satellite-internet-low-earth-orbit-facebook-spacex-starlink>. [Accessed 21 February 2021].
- [78] X. Wang, S. Mao and M. Gong, "A survey of LTE wi-fi coexistence in unlicensed bands," ACM Mobile Computer Communication, vol. 20, no. 3, 2016.
- [79] Microsoft, "A Rural Broadband Strategy," 10 July 2017. [Online]. Available: https://msblob.blob.core.windows.net/ncmedia/2017/07/Rural-Broadband-Strategy-Microsoft-Whitepaper-FINAL-7-10-17.pdf?utm_source=Direct. [Accessed 20 July 2017].
- [80] E. Leitgeb, M. Gebhart, U. Birnbacher, S. Sheikh Muhammad and C. Chlestil, "Applications of Free Space Optics for Broadband Access," in Optical Networks and Technologies, Pisa, Italy, 2005.

- [81] D. Rodewald, "MRV introduces industry's first 10G ethernet wireless point-to-point system," MRV Communications Inc., Chatsworth, CA, USA, 2008.
- [82] M. A. Khalighi, "Survey on Free Space Optical Communication: A Communication Theory Perspective," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2231-2258, 2014.
- [83] "Net2Plan - The open-source network planner," [Online]. Available: <http://www.net2plan.com/>. [Accessed 24 July 2017].
- [84] R. Sivaraj and T. Ravichandran, "A Review of selection methods in Genetic Algorithm," *International journal of engineering science and technology*, vol. 3, no. 5, 2011.
- [85] Y. John and R. Langari, *Fuzzy Logic: Intelligence, Control, and Information*, Prentice-Hall Inc, 1999, pp. 469-470.
- [86] S. Rajasekaran and G. A. Vijayalakshmi Pai, *NEURAL NETWORKS, FUZZY SYSTEMS AND EVOLUTIONARY ALGORITHMS : SYNTHESIS AND APPLICATIONS*, PHI Learning Pvt. Ltd., 2017.
- [87] S. Forrest, "Genetic algorithms for multiobjective optimization: Formulation discussion and generalization," *Proceedings of the Fifth International Conference on Genetic Algorithms*, pp. 416-423, 1993.
- [88] N. Srinivas and K. Deb, "Multiobjective function optimization using nondominated sorting genetic algorithms," *Evolutionary Computing*, vol. 2, no. 3, pp. 221-248, 1995.
- [89] A. E. Eiben, T. Back, M. Schoenauer and H. P. Schwefel, "Multiobjective optimization using evolutionary algorithms: A comparative case study," in *Parallel Problem Solving From Nature V*, Berlin, Germany, Springer-Verlag, 1998, pp. 292-301.
- [90] R. Steuer, *Multiple Criteria Optimization: Theory, Computation and Application*, John Wiley & Sons, Inc, 1986.
- [91] W. Yuanzhi, "Constraint cellular ant algorithm for the multi-objective vehicle routing problem," *Journal of Software*, vol. 8.6, pp. 1339-1345, 2013.
- [92] G. Vassil, H. Toshev and C. Korsemov, "Survey of evolutionary algorithms used in multiobjective optimization," *Problems of engineering cybernetics and robotics*, vol. 60, pp. 42-54, 2009.
- [93] F. Kursawe, "Evolution strategies for vector optimization," in *Proc. 10th International conference Multiple Criteria Decision Making*, Taiwan, 1992.
- [94] H. Ishibuchi and T. Murata, "Multi-objective genetic local search algorithm," *IEEE International Conference Evolutionary Computation*, pp. 119-124, May 20-22, 1996.
- [95] D. S. Todd and P. Sen, "A Multiple criteria genetic algorithm for containership loading," in *International Conference Genetic Algorithms*, San Francisco, CA, 1997.

- [96] A. G. Cunha, P. Oliviera and J. Covas, "Use of genetic algorithms in multicriteria optimization to solve industrial problems," in International Conference Genetic Algorithms, San Francisco, CA, 1997.
- [97] S. Yilmaz, K. Ivanov, S. Levine and M. Mahgerefteh, "Application of genetic algorithms to optimize burnable poison placement in pressurized water reactors," Annual Nuclear Energy, pp. 446-456, 2006.
- [98] S. Potti and C. Chinnasamy, "Strength Pareto Evolutionary Algorithm based Multi-Objective Optimization for shortest Path routing problem in computer networks," Journal of Computer Science, vol. 7, no. 1, pp. 17-26, 2011.
- [99] E. Zitzler, M. Laumanns and L. Thiele, "SPEA2: Improving the strength Pareto evolutionary algorithm," Evolutionary Methods for Design, Optimization and Control with Applications to Industrial Problems, pp. 95-100, 2002.
- [100] N. Makariye, "Towards shortest path computation using Dijkstra algorithm," in 2017 International Conference on IoT and Application, Nagapattinam, India, 2017.
- [101] N. Jasika, N. Alispahic, A. Elma, K. Ilvana, L. Elma and N. Nosovic, "Dijkstra's shortest path algorithm serial and parallel execution performance analysis," in 2012 Proceedings of the 35th International Convention MIPRO, Opatija, Croatia, 2012.
- [102] K. Deb, A. Pratap, S. Agarwal and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," Evolutionary Computation, vol. 6, no. 2, pp. 182-197, 2002.
- [103] Y. Fu, H. Wang and G. Jiang, "An improved NSGA-II to solve multi-objective optimization problem," in The 26th Chinese Control and Decision Conference, Changsha, China, 2014.
- [104] B. Luo, J. Zheng, J. Xie and J. Wu, "Dynamic Crowding Distance? A new Diversity Maintenance Strategy for MOEAs," in Fourth International conference on Natural Computation, Jinan, China, 2008.
- [105] K. Kirkpatrick, "Software-defined networking," Communication of the ACM, vol. 56, no. 9, pp. 16-19, 2013.
- [106] F. de Oliveira Silva, J. de Souza Pereira, P. Rosa and S. Kofujo, "Enabling Future Internet Architecture Research and Experimentation by Using Software Defined Networking," in 2012 European Workshop on Software Defined Networking, Darmstadt, Germany, 2012.
- [107] G. Goth, "Software-defined networking could shake up more than packets," IEEE Internet Computing, vol. 15, no. 4, pp. 6-9, 2011.
- [108] W. Xia, Y. Wen, D. Niyato and H. Xie, "A Survey on Software-Defined Networking," IEEE Communications Surveys & Tutorials, vol. 17, no. 1, pp. 27-51, 2015.
- [109] R. Masoudi and A. Ghaffari, "Software defined networks: A survey," Journal of Network and Computer Applications, vol. 67, pp. 1-25, 2016.

- [110] F. A. Lopes, M. Santos, R. Fidalgo and S. Fernandes, "A Software Engineering Perspective on SDN Programmability," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1255-1272, 2016.
- [111] K. Bakshi, "Considerations for Software Defined Networking (SDN): Approaches and use cases," in *2013 IEEE Aerospace Conference*, USA, 2013.
- [112] D. B. Fogel, "An Introduction to Simulated Evolutionary Optimization," *IEEE Transactions on Neural Networks*, vol. 5, no. 1, pp. 3-14, 1994.
- [113] M. Srinivas and L. M. Patnaik, "Adaptive probabilities of crossover and mutation in genetic algorithms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 24, no. 4, pp. 656-667, 1994.
- [114] R. Belea, S. Caraman and V. Palade, "Diagnosing the Population State in a Genetic Algorithm Using Hamming Distance," *the Proceedings of Knowledge-based and Intelligent Engineering Systems KES 2004*, pp. 246-255, 2004.
- [115] E. Zitzler and L. Thiele, "An evolutionary algorithm for multiobjective optimization: the strength pareto approach," *Swiss Federal Institute of Technology (ETH) Zurich*, Zurich, 1998.
- [116] K. Grunert, R. H. Meyer, M. Knofel and R. Zhao, "From strategy to implementation - Tool based planning of optical networks," in *Asia Communications and Photonics Conference and Exhibition (ACP)*, Shanghai, China, 2011.
- [117] A. F. Santos, G. S. Couto, R. H. C. Manicoba, K. D. R. Assis and R. C. Almeida, "SimRSA: An education tool for network planning in Spectrum-sliced Elastic Optical Path Networks," in *16th International Conference on Transparent Optical Networks (ICTON)*, Graz, Austria, 2014.
- [118] N. Priyanka, G. Achala, V. Seshasai and G. Sadashivappa, "Bandwidth planning tool for optical networks," in *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, Bangalore, India, 2016.
- [119] M. A. Cavalcante, H. A. Pereira and R. C. Almeida Jr., "SimEON: an open-source elastic optical network simulator for academic and industrial purposes," *Photonic Network Communications*, vol. 34, pp. 193-201, 2017.
- [120] A. Kretsis, P. Soumplis, K. Christodoulopoulos and E. Varvarigos, "DEPLOY Tool: Incrementally Planning Marginless Elastic Optical Networks," in *20th International Conference on Transparent Optical Networks (ICTON)*, Bucharest, Romania, 2018.
- [121] X. Chen, "Energy Efficient Wired Networking," September 2014. [Online]. Available: <https://qmro.qmul.ac.uk/jspui/handle/123456789/7966>. [Accessed 24 August 2018].
- [122] D. Goldberg, *Genetic algorithms in search, optimization and machine learning*, Boston, MA, USA: Addison-Wesley Longman Publishing Co. Inc., 1989.

- [123] E. Zitzler and L. Thiele, "Multiobjective evolutionary algorithms: a comparative case study and the strength Pareto approach," *IEEE Transactions on Evolutionary Computation*, vol. 3, no. 4, pp. 257-271, Nov 1999.
- [124] M. L. Mauldin, "Maintaining diversity in Genetic Search," in *National Conference on Artificial Intelligence*, 1984.
- [125] D. Gupta and S. Ghafir, "An overview of methods maintaining diversity in Genetic Algorithms," *International Journal of Emerging Technology and Advanced Engineering*, vol. 2, no. 5, pp. 56-60, May 2012.
- [126] R. Kumar, P. Parida and M. Gupta, "Topological design of communication networks using multiobjective genetic optimization," in *Congress on Evolutionary Computation*, Honolulu, USA, 2002.
- [127] N. Islam and C. Phillips, "Cost-effective Network Planning for Rural Communities," in *International Conference on Computing, Electronics & Communications Engineering (iCCECE)*, London, 2019.
- [128] S. A. Zabidi, W. A. Khateeb, M. R. Islam and A. W. Naji, "The effect of weather on free space optics communication (FSO) under tropical weather conditions and a proposed setup for measurement," in *International Conference on Computer and Communication Engineering*, Kuala Lumpur, Malaysia, 2010.
- [129] A. Agarwal, M. Kodialam and T. Lakshman, "Traffic engineering in software defined networks," *INFOCOM*, pp. 2211-2219, 2013.
- [130] H. Dekker and B. D. Colbert, "Network Robustness and Graph Topology," *Australasian Computer Science Conference*, vol. 26, pp. 359-368, 2004.
- [131] J. Moysen, L. Giupponi and J. Mangues-Bafalluy, "Machine Learning enabled network planning tool," in *International Symposium on Personal, Indoor and Mobile Radio Communications*, 2016.
- [132] Riverbed, "Riverbed opnet netone," [Online]. Available: <https://support.riverbed.com/content/support/software/steelcentralnpm/net-planner.html>. [Accessed 18 August 2017].
- [133] CelPlan, "CelTrace Wireless Global Solutions," [Online]. Available: <http://www.celplan.com/products/indoor/celtrace.asp>. [Accessed 18 August 2017].
- [134] G. Bernardi, M. K. Marina, F. Talamona and D. Rykovanov, "IncrEase: A Tool for Incremental Planning of Fixed Broadband Wireless Access Networks".
- [135] S. Enayati and H. Saeedi, "Deployment of Hybrid FSO/RF Links in Backhaul of Relay-based Rural area cellular networks: advantages and performance analysis," *IEEE communications letters*, vol. 20, no. 9, pp. 1824-1827, 2016.
- [136] P. Alto, "Software-defined networking: The new norm for networks," April 2012. [Online]. Available: <https://www.opennetworking.org/sdn-resources/sdn-definition>. [Accessed 13 August 2017].

- [137] I. Lee, Z. Ghassemlooy, W. P. Ng and M. A. Khalighi, "Reducing pointing errors in free-space optical communication links over turbulences with a partially coherent Gaussian beam," in *EEE International Conference on Communications Workshops (ICC)*, Kuala Lumpur, 2016.
- [138] R. Yang, Q. Liu and Z. Wu, "Effects of atmospheric turbulence scintillation on the error performance of partially coherent laser communication," in *3rd European Conference on Antennas and Propagation*, Berlin, 2009.
- [139] A. K. Rahman, S. A. Aljjunid, M. S. Anuar and H. A. Fadhil, "Improve the bit error rate performance of free space optic due to atmospheric turbulence using new dual diffuser modulation," in *IEEE 5th International Conference on Photonics (ICP)*, Kuala Lumpur, 2014.
- [140] L. Nandita, "Performance Analysis of Free-Space Optical Wireless Channels Using Finite-SNR DMT," in *International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, Chennai, 2017.
- [141] R. Narasimhan, "Finite-SNR diversity multiplexing tradeoff for correlated rayleigh and rician MIMO channels," *IEEE transactions on Information Theory*, vol. 52, no. 9, pp. 3965-3979, 2006.
- [142] E. Stauffer, O. Oyman, R. Narasimhan and A. Paulraj, "Finite-SNR diversity-multiplexing tradeoffs in fading relay channels," *IEEE Journal on Selected Areas in Communications*, vol. 25, no. 2, 2007.
- [143] N. Lavanis and D. Jalihal, "Finite-SNR diversity multiplexing tradeoff of simo diversity combining schemes," *IEEE Vehicular Technology Conference*, pp. 1-5, 2009.
- [144] X. Zhu and J. M. Kahn, "Free-space optical communication through atmospheric turbulence channels," *IEEE Transactions on communications*, vol. 50, no. 8, pp. 1293-1300, 2002.
- [145] Z. A. Almaalie, Z. Ghassemlooy, A. A. S. Al-Rubaie, I. E. Lee and H. Le-Minh, "Forward error correction with physical layer network coding in two-way relay free space optical links," in *Computer Science and Electronic Engineering Conference (CEEC)*, Colchester, 2016.
- [146] GSM Association, "The state of mobile internet connectivity 2021," GSMA, September 2021. [Online]. Available: <https://www.gsma.com/r/wp-content/uploads/2021/09/The-State-of-Mobile-Internet-Connectivity-Report-2021.pdf>. [Accessed 12 December 2021].
- [147] C. Butler, "Disaggregating the Drivers of Mobile Technology Adoption: The Threat of Unobservable Gender Biases," in *The 48th Research Conference on Communication, Information and Internet Policy*, Texas, USA, 2020.