KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

INSTITUTE OF DISTANCE LEARNING

MINIMUM SPANNING TREE ROUTE FOR MAJOR TOURIST CENTERS IN THE BRONG AHAFO REGION OF GHANA

By:

AKUTSAH, FRANCIS

A THESIS SUBMITTED TO THE DEPARTMENT OF MATHEMATICS IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF DEGREE OF MASTER OF SCIENCE IN INDUSTRIAL MATHEMATICS

JUNE, 2011

CERTIFICATION

I herein certified that this work was carried out solely by AKUTSAH FRANCIS (PG3007609) in the Department of Mathematics, Institute of Distance Learning in partial fulfilment of the requirements for the award of Master of Science in Industrial Mathematics.

AKUTSAH FRANCIS	....	....
(Name of student)	(Signature)	(Date)

DR. F. T. ODURO	....	....
(Supervisor)	(Signature)	(Date)

DECLARATION

I hereby declare that this submission is my own work toward the M Sc. and that to the best of my knowledge it contains no material previously published by another person nor material which has been accepted for the award of any other degree of the University except due acknowledgement as been made in the text.

AKUTSAH FRANCIS ....
Student number: 20103896 Signature Date

Certified by:

DR. F. T. ODURO ....
SUPERVISOR Signature Date

Certified by:

DR. S. K. AMPONSAH ....
Head of Department Signature Date

DEDICATION

To the Glory of God

I dedicate this project to Thomas Akutsah Jr. and all my colleagues for their moral and mutual support.

ACKNOWLEDGEMENT

It is with great pleasure that I take this opportunity to recognize those who have played a major role in bringing this significant work to its full realization. It has been satisfying to see all the pieces come together, often in ways much better than I expected.

It is therefore, my pleasure to express my greatest indebtedness to my supervisor, DR. F. T. ODURO, Department of Mathematics, KNUST, who was always ready to assist me in one way or the other and also helped me in the compilation of this work, may God richly bless him. I also wish to express my profound gratitude to all the lecturers in the Mathematics Department who contributed in one way or the other for the successful completion of this project.

Also, to Mr. Kuvor Wisdom of Ghana Tourist Board (Head Office Accra), Mrs. Matilda Naa (Ghana Tourist Board, Sunyani), Mr. Ernest Osei Bonsu (Ghana Highways Authority Accra) and Mr. Oppong James (Department of Feeder Road, Sunyani) I say may the Lord richly bless you all for your co-operation.

To Mr. Parker George, I say thanks for your guidance in developing the MATLAB Code.

ABSTRACT

The tourism industry is becoming more lucrative in the country. Various measures have been taken by government and private individuals to make the industry more attractive. In this thesis, a mathematical model for finding minimum spanning tree was used to find the minimum spanning tree route for some selected tourist sites in the Brong Ahafo region of Ghana. Prim's Algorithm was used to find the minimum spanning tree. The study reveals that a total minimum distance of three and hundred and sixty kilometres (360Km) will be covered for touring all the eleven selected tourist centers. Also, Wenchi and Buoyem serve as hubs for the tourist industry in the region. Among the recommendations offered was that the Ghana Tourist Board and other Travel and Tour Operators adopt this study as a basis of developing facilities and resources to support the industry in the study area.

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

BCDCP – Base Station Control Dynamic Protocol

Bu – Buoyem

DMSTRP – Dynamic Minimal Spanning Tree Routing Protocol

Du – Duasidan

Fi – Boaben Fiema

Fo – Forikrom

Fu – Fuller falls

GDP Gross Domestic Product

GGP Gross Global Product

GIS – Geographic Information System

GMST Generalised Minimum Spanning Tree

Ha Hani

HCV Hepatitis C Virus

Ki Kintampo

LEACH Low Energy Adaptive Clustering Hierarchy

Ma Bono Manso

MR Magnetic Resonance

MST Minimum Spanning Tree

MSTP Multiple Spanning Tree Protocol

Nk Nkyeraa

NNT Nearest Neighbour Tree

NP Non Polynomial

OECD Organisation for Economic and Cooperative Development

RSTP	Rapid Spanning Tree Protocol
STP	Spanning Tree Protocol
Ta	Tano Boase
UNWTO	United Nations World Tourism Organisation
VLSI	Very Large Scale Integrated
We	Wenchi
WTO	World Tourism Organisation

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.0 OVERVIEW

Most African countries depend greatly on agriculture which also relies on the weather and rainfall. This reduces output, especially during the dry season and sometimes delays in the rainfall coupled with weather changes. Tourism is often developed because it promises to generate employment, enhance community infrastructure and assist in revitalising the flagging economies of rural areas. In Eastern Europe for example, tourism has been identified as a catalyst to stimulate economic growth, increase the viability of underdeveloped regions and improve the standard of living for the local population. In less developed countries in Sub-Sahara Africa, afflicted by debilitating rural poverty, tourism is increasingly seen as one of the few feasible options for economic development. Ghana has also benefited from tourism and is still benefiting from the industry. This has brought about the need to have good modes of transport to the various tourism centers in the country. This work seeks to provide the minimum spanning tree route to the major tourism centers in the Brong Ahafo Region of Ghana.

1.1 BACKGROUND TO THE STUDY

1.1.1 WHAT IS TOURISM?

Tourism has been defined in several ways depending on the use to which the definition is needed, and the writer's intentions.

Mathieson and Wall (1982) define tourism as ‘the temporary movement of people to destinations outside their normal places of work and residence, the activities undertaken during their stay in those destinations and the facilities created to carter for their needs’.

Macintosh and Goelder (1986) also define it as ‘the sum of the phenomena and relationship arising from the interaction s of tourist, business, suppliers, host government and host communities in the process of attracting and hosting these tourist and the other visitors’.

World Tourism Organisation(WTO) defines tourist as a person who ‘travel to and stay in places outside their usual environment for not more than consecutive year for leisure, business and other purposes not related to the exercise of an activity remunerated from within the place visited’.

According to the Longman Dictionary of Contemporary English, ‘the business of providing things for people to do, places for them to stay etc while they are on holiday’ is tourism.

Of late, tourism has become a popular leisure global activity.

SOME TOURIST SITES IN GHANA

There are many tourist attraction sites in Ghana, some of these sites include; Mushroom Rock, Gbele Game Reserve, Hippopotamus Sanctuary, Sankana Slave Cave all in the Upper West Region.   Paga Crocodile Pond, Whistling Rocks, Sirigu Craft Village, Navrongo Basilica (Largest Mud-build Basilica in the world) all in the Upper East Region. Mole National Park, Gambaga Escapement,Salaga Slave Market and Well all in the Northern Region. Kumasi Zoological Gardens, Bobiri Forest Reserve and Butterfly Sanctuary, Bonwire Kente and Craft Village, Lake Bosomtwe, Okonfo Anokye Sword Site, Manhyia Palace, Military Museum all in the Ashanti Region. Bia National Park, Ankasa National Park, Nzulezu (Village on Stilts), Egyambra Crocodile Sanctuary, Fort Metal Cross all in the Western Region.

Kakum National Park, Elmina Castle, Cape Coast Castle, Ajumako Craft Village, Assin Manso Reverential Gardens all in the Central Region. Boti Falls, Aburi Botanical Gardens, Adomi Bridge, Akosombo Hydro Dam all in the Eastern Region. Tafi tome Monkey Sanctuary, Wli Waterfalls, Tagbo Waterfalls, Mount Germini and Afadjato all in the Volta Region. Christiansburg Castle, Kwame Nkrumah Mausoleum, Ostrich Farm all in the Greater Accra Region. Those in the Brong Ahafo Region are; Boabeng Fiema Monkey Sanctuary, Nkyeraa Waterfalls, Tano Sacred Grove, Duasidan Wildlife Sanctuary, Forikrom Boten Shrine and Caves the rest are Buoyam Caves and Bats Colony, Bono Manso Slave Site, Fuller Falls, Kintampo Waterfalls, Dr. K.A Busia Mausoleum and the Hani Archeological Site.

1.1.2 ECONOMIC IMPORTANCE OF TOURISM

Tourism directly or indirectly employs one in every ten (10) employed people on earth, and maintains 3.6 trillion dollars worth of goods and services which is 10.6% the Gross Global Product (Brown 24). Tourism is an important economic sector in Africa within more than half of Sub-Saharan Africa countries (Ashley et al, 2006). Tourism is becoming an increasingly important economic sector for developing countries. According to Pro-poor Tourism in Practice (2004a), the absolute tourism earnings of developing countries grew by 133 percent between 1990 and 2000 and in the least developed countries by 154 percent, as compared with 64 percent for OECD (Organisation for Economic and Cooperative Development) countries and 49 percent for EU countries. According to the World Bank's World Development Indicators (2004), tourism is an important economic activity in 70 percent of developing countries. In 28 of the 49 Sub-Saharan African countries, tourism contributes more than 3 percent of gross domestic product (GDP). The potentials to utilise tourism as a tool for sustained socio-economic development as well as the redistribution of developed-country's wealth is obvious, yet somehow elusive (Harrison,1994). It is estimated that tourism produces 50 to 60% of the total GDP of the Bahamas.

In Ghana, tourism forms an integral part of our Gross Domestic Product(GDP) below is a table showing some figures of tourism to Ghana's GDP.

Source: Ghana Tourist Board: 6/9/2010 **Table: 1.1Tourism in terms of GDP**

	YEAR						
GDP	2002	2003	2004	2005	2006	2007	2008
(%)	4.4	4.7	4.9	5.7	5.8	6.3	6.7

This shows how important the tourism has become in recent times to the economy of Ghana.

1.1.3 SOCIAL IMPACT OF TOURISM

There are various definitions of social development, and most of them converge around the concepts of improving the well-being of a country's citizens, promoting higher standards of living, increasing employment and creating conditions of economic and social progress. Employment is one of the most readily available indicators to begin measuring the social impact of tourism, since job creation generally helps create the opportunities for better standards of living and related conditions of socio-economic progress. In 2006, the tourism economy (direct plus indirect contribution) provided jobs for about 140 million people in the selected subregions and countries of the Asian and Pacific region, representing an average of

8.9 per cent of total employment. Tourism employment in North-East Asia was estimated at 87.7 million jobs, which was 10.1 per cent of the total employment in the subregion. This result can be attributed mainly to China, where 77.6 million people, approximately 1 in every 10 employed persons worked in the tourism economy. In Oceania, the workforce in the tourism economy accounted for 14.5 per cent of total employment in the subregion, which was 1 in every 6.9 jobs (UNWTO, Tourism Highlights 2007 Edition)

1.1.4 WHAT IS A ROUTE?

According to the Longman Dictionary of Contemporary English, route is a way between two places that buses, planes, ships etc regularly travel. There are several modes of transport in the world, air, road, rail and water are the commonest modes of transport. The dominant mode of transport is road transport in the world especially in the third world countries. Tourism routes promise to bring together a variety of activities and attractions under a unified theme and thus stimulate entrepreneurial opportunity through the development of ancillary products and services (Long et al, 1990).

Getz and Page (1997) generally, most routes are initiated with one of the following in mind;

- Diffuse visitors and disperse income from tourism,
- To bring lesser known attractions and features into the tourism business,
- To increase the length of stay and spending by tourist,
- To increase the sustainability of the tourism industry.

Underlying the attractiveness of any route is its perceive ‘distance’ in terms of;

- Geographical distance between the generating regions and the tourism destination
- Travel time needed to cover the geographical distance

- Amount of money needed by a tourist to cover the distance.

Meyer (2004) argues that in the developing world routes generally pass along secondary roads which offer a more relaxed travelling pace and thus opportunities to stop on route. Many of these road networks are comprised of scenic routes, thus increasing the appeal of tourists to choose them as opposed to ‘faster’ highways.

1.2 STATEMENT OF THE PROBLEM

There are several modes of transport in Ghana, road, water, air and rail. The road transportation is the commonest mode of transport in Ghana today. This is also the means by which the various tourist who visit the country use to the various tourist centers. Even though, there are a number of routes that these tourist use any time they want to visit a particular tourist center, one may ask which of these routes is more economical in terms of cost and time consumed without compromising access? Or when tourist comes into the country and he or she is given tourist map, will he or she be in the position to see where to start and end an itinery in order to minimize cost? Or can he or she identify nature of his or her movement at a glance of the map due to the simplicity of the appropriate route? It against this background that the research has been under taken to find the Minimum Spanning Tree Route for Major Tourist Centers using the Brong Ahafo Region of Ghana as a case study**.**

1.3 OBJECTIVES OF THE STUDY

The specific objectives of the study are:

i. To find using Prim's Algorithm, a minimum spanning tree route for the major tourist centers in the Brong Ahafo Region of Ghana based on a road network model.
ii. To determine the implications for tourism of the resulting route.

1.4 METHODOLOGY

Data for the study was collected from the Ghana Tourist Board, Statistical Service Department, Geographic Information System (GIS) and Department of Roads and Highways.

The data involved information about major tourist centers in the Brong Ahafo Region as well as concerning the pertinent road network especially in terms of distances between the nodes in the network.

The use of tourist map, topological maps, road map and also MATLAB and Texas Instrument calculator supported computations and programming in order to get high level of accuracy in the distance calculations and the determination of the minimum spanning tree by means of an implementation of Prim's algorithm.

1.5 SIGNIFICANCE OF THE STUDY

In an attempt to identify routes that will help reduce stress in travelling, high cost of transporting goods and services from one place to another, minimizing time spent in travelling, the study will help motorist to travel with ease by using the shortest path between any two given locations.

It will also help travel and tour operators to maximize profit by using the identified route during their visits. Also, transporting food and other pertinent commodities to and from the centers will be relatively cheaper and less time consuming when the same routes are followed. This will help improve economic activities generally in the country since people can move to and fro at relatively cheaper cost.

1.6 LIMITATIONS OF THE STUDY

Even though the study covers the entire region not every tourist center is captured in the study due to the following setbacks:

1. In sufficient resources and funds to capture every tourist center in the region.
2. Limited time within which the study should be completed.
3. Some of the tourist centers do not have adequate data as at the time the study is been conducted.

1.7 ORGANIZATION OF THE THESIS

The thesis is divided into five chapters. The Chapter 1 is made up of the introduction of the study, this will be made up the background of the study with emphasis on what tourism is, and what a spanning tree is, Statement of the problem and the objectives of the study, methodology to be used in the study and the significance of the study. Also included in this chapter, is limitations of the study and then the organization of the study.

Chapter 2 is made up of review of related literature; emphasising work done by other researchers. Chapter 3 contain methodology, the mathematical model for solving minimum spanning tree problems. Chapter 4 is discussion of results; emphasis will be on results obtained. Chapter 5 captures the Summary of findings, recommendations and conclusion.

CHAPTER 2

LITERATURE REVIEW

2.0 OVERVIEW

Many studies have been conducted on tourism and hospitality industry here in Ghana and various countries where tourism industry is very or becoming lucrative. Some of the studies conducted are on 'impact of tourism on the destination communities' while others also tried to look at 'Effective way of managing the tourism industry' in their respective countries. In this study, emphasis will be on minimum spanning tree route for the major tourist sites in the Brong Ahafo Region of Ghana. The minimum spanning tree problem is one of the most fundamental and intensively studied problems in network optimization with many theoretical and practical applications (Ahuja et al, 1993)

2.1 WIRELESS AND INTERNET NETWORK MODELS

Khan et al (2006) conducted a study into "Distributed Algorithms for Constructing Approximate Minimum Spanning Trees in Wireless Networks". Though there are distributed algorithms for the Minimum Spanning Tree (MST) problem, they realised these algorithms require relatively large number of messages and time, and are fairly involved, making them impractical for resource-constrained networks such as wireless sensor networks. In such networks, a sensor has very limited power, and any algorithm needs to be simple, local, and energy efficient. Motivated by these considerations, they designed and analysed a class of simple and local distributed algorithms called Nearest Neighbour Tree (NNT) algorithms for energy-efficient construction of an approximate MST in wireless networks. Assuming that the nodes are uniformly distributed, they show provable bounds on both the quality of the spanning tree produced and the energy needed to construct them. They show that while NNT produces a close approximation to the MST, it consumes asymptotically less energy than the classical message-optimal distributed MST algorithm due to Gallager, Humblet, and Spira. Further, the NNTs can be maintained dynamically with polylogarithmic rearrangements under node insertions/deletions. They also performed extensive simulations, which show that the bounds are much better in practice. Their results, to the best of their knowledge, demonstrates the first trade-off between the quality of approximation and the energy required for building spanning trees on wireless networks, and motivate similar considerations for other important problems.

Loni (2010) in his work on “Randomly Generated Edge-Disjoint Minimum Spanning Trees based Data Gathering Trees”, Since wireless sensor nodes have limited energy resource that cannot be recharged and are randomly scattered in the observation fields, energy efficiency becomes one of the most important problems. They review basic terminology and protocols that are energy efficient as well as some proposed methods of improvement and performance. They proposed methods of modifying the energy efficient algorithms. Some of the proposed algorithms include: Low-Energy Adaptive Clustering Hierarchy (LEACH), Power-Efficient Gathering in Sensor Information System (PEGASIS) and Base Station Controlled Dynamic Clustering Protocol (BCDCP). They also look at how the algorithms perform on the network and how introducing changes in the network topology and accommodating those changes with the algorithm affect performance.

In a related work by (Guangyan et al, 2006) they proposed another innovative cluster-based routing protocol named Dynamic Minimal Spanning Tree Routing Protocol (DMSTRP), which improves BCDCP by introducing MSTs instead of clubs to connect nodes in clusters. Simulation results show that DMSTRP excels LEACH and BCDCP in terms of both network lifetime and delay when the network size becomes large.

Perez (2009) conducted a study into “improving resource utilization in carrier Ethernet technology”. The study was divided into two classes; the first class relies only on Ethernet control components such as Multiple Spanning Tree Protocol (MSTP) and the Rapid

Spanning Tree Protocol (RSTP). With the MSTP, several spanning trees can be created in the source Ethernet network allowing to route traffic through different paths between a pair of nodes in the network.

The second class relies on improving both Ethernet protocol and forwarding components. The work analyzes and compares label space usage for both architectures to ensure their scalability. They proposed and ILP to calculate optimal performance of this class of approaches and compares them with the label based forwarding technologies to enable to determine, given a specific scenario, which approach to use.

Dippon et al (1999) looked in the “the cost of the local telecommunication network: “A Comparison of Minimum Spanning Trees and the HAI Model”. According to them, under the Telecommunications Act, estimates of local distribution costs may be used to help quantify the subsidy for specified local services whose costs exceed their tariff rates and as a guide for the pricing of unbundled network elements. The most widely-circulated model for estimating these costs, the HAI model, uses a particular procedure to calculate the distribution network and cable length that is required to serve a cluster of customers. They compare the HAI procedure with the minimum spanning tree (MST), which gives the shortest distance for connecting a set of locations. For each cluster in Minnesota they calculated the distribution length with the HAI procedure and the length of the MST. They found that the HAI length is shorter than the MST length in 77% of the main clusters. In low-density areas, the HAI length is less than the MST length for 81% of the main clusters. The too-short cable lengths mean that the HAI model underestimates network costs; this underestimation extends beyond the cost of the cables themselves since many cost components are tied to cable length, such as support structures, maintenance, and associated power and back-up equipment. The use of underestimated costs in determining subsidies and network prices would discourage the provision of services in subsidized areas and encourage inefficient entry that utilizes unbundled network elements.

Farhad et al (2008) Spanning Tree Protocol (STP) is a layer-2 protocol which ensures a loop free topology in Metro Ethernet networks. It is based on Minimum Spanning Tree solution that involves determining the links which can join all the nodes of a network together such that the sum of the costs of the chosen links is minimized. In STP, all customers need to use the same spanning tree and there isn't any traffic engineering mechanism for load balancing. This results in uneven load distribution and bottlenecks, especially close to the root. A solution for this problem is using the multi-criteria Minimum Spanning Tree by considering criterions such as load balance over links and switches. In their previous work, the algorithm was based on computation of the total cost for each possible spanning tree and then selection the best one with minimum total cost. This algorithm is very time consuming, especially when their Metro Ethernet network is large. In this study, they proposed a new approach using Genetic Algorithm. It reduces the computational complexity by selecting the best spanning tree in a stochastic manner.

According to Ahlswede et al (2000) recent research shows that routing alone is not sufficient to achieve the maximum information transmission rate across a communication network.

2.2 IMAGES AND SPANNING TREES

Anthony et al (2007) conducted a study into 'Hierarchical Minimum Spanning Trees for Lossy Image Set Compression'. They propose a hierarchical minimum spanning tree algorithm in which the minimum spanning tree algorithm is first applied to clusters of similar images and then it is applied to the average images of the clusters. It was shown that the new algorithm outperforms the previous image set compression algorithms for image sets which are not very similar, especially at lower bitrates. Furthermore, the computational requirement for a minimum spanning tree is significantly lower than the previous algorithms.

Hero et al (2000) conducted a study into “image registration with minimum spanning tree algorithm”. They propose a novel graph representation method for image registration with R´enyi entropy as the dissimilarity metric between images. The image matching is performed by minimizing the length of the minimum spanning tree (MST) which spans the graph generated from the overlapped images. Their method also takes advantage of the minimum k-point spanning tree (k-MST) approach to robustify the registration against spurious discrepancies in the images. The proposed algorithm is tested in two applications: registering magnetic resonance (MR) images, and registering an electro-optical image with a terrain height map. In both cases the algorithm is shown to be accurate and robust.

Aditee (2004) conducted a study into “affine image registration using minimum spanning tree entropies”. They closely followed Hero's work and have successfully shown how the MST based approach can be used to perform registration over six parameters of the affine transformation. They have computed an information theory based criterion similar to the conventional mutual information using the Ŕenyi entropy as proposed by (Hild et al, 2002) and then minimized it over an affine search space. Also, they combined the use of intensities and features in the same information theory based registration technique.

Mert et al (2008) provide a detailed analysis of the use of minimal spanning graphs as an alignment method for registering multimodal images in their work “Using Spanning Graphs for Efficient Image Registration”. According to them, this yielded an efficient graph theoretic algorithm that, for the first time, jointly estimates both an alignment measure and a viable descent direction with respect to a parameterized class of spatial transformations. They also show how prior information about the inter-image modality relationship from pre-aligned image pairs can be incorporated into the graph-based algorithm. A comparison of the graph theoretic alignment measure is provided with more traditional measures based on plug-in entropy estimators.

This highlighted previously unrecognized similarities between these two registration methods. Their analysis gives additional insight into the tradeoffs the graph-based algorithm is making and how these will manifest themselves in the registration algorithm's performance.

2.3 ALGORITHMS AND REVIEWS

Philipp (2003) Presented an algorithm that computes a spanning tree with stretch $O(OPT^4)$ in time $O(n \log n)$. Besides this, they show a greedy and an evolutionary algorithm and prove that they do not produce a spanning tree with stretch better than O(n). At last they presented two algorithms for which, in the worst example found, the resulting spanning trees have stretch $\Omega(\mathrm{OPT}^2)$ but it remains open how good the approximation factors of these algorithms are.

Narasimhan et al. (2001) gave a practical algorithm that solves the GMST problem. They prove that for uniformly distributed points, in fixed dimensions, an expected $O(n \log n)$ steps suffice to compute the GMST using well separated pair decomposition. Their algorithm, GeoMST2, mimics Kruskal's algorithm (1956) on well separated pairs and eliminates the need to compute bichromatic closest pairs for many well separated pairs.

Smith et al (2010) conducted a study into "Computing Geometric Minimum Spanning Trees Using the Filter-Kruskal Method". They proposed GeoFilter-Kruskal, an algorithm that computes the minimum spanning tree of P using well separated pair decomposition in combination with a simple modification of Kruskal's algorithm. When P is sampled from uniform random distribution, they show that their algorithm runs in $o\ (n\ log^2\ n)$ time with probability at least $1 - \frac{1}{n^c}$ for a given $c > 1$. Although this is theoretically worse compared to known O(n) Sangutharva (2004) or O(n log n) Clarkson et al (1989) algorithms, experiments show that their algorithm works better in practice for most data distributions compared to the current state of the art . Their algorithm is easy to parallelize and to our knowledge, is currently the best practical algorithm on multi-core machines for $d > 2$.

Allison (2009) study "The Geometric Structure of Spanning Trees and Applications to Multiobjective Optimization" They study many different properties of spanning trees, including the graph of tree exchanges. Using this graph, they then study multiobjective optimization with regards to the edge costs. They wrote and implemented a program to enumerate all spanning trees, in order to assess the accuracy of their optimization algorithms and heuristics. They also studied the general case for matroids and wrote a program to estimate the number of bases of matroid polytopes. They considered several fast heuristics that can find the minimum spanning tree for a graph with respect to multiple sets of edge costs, particularly finding the Pareto optima. Although these heuristics could potentially only locate a local minimum, they locate the global minimum in almost every trial, and are extremely efficient.

Erlebach et al (2008) also considered minimum spanning tree problem in a setting where information about the edge weights of the given graph is uncertain. Initially, for each edge e of the graph only a set A_e called an uncertainty area, that contains the actual edge weight w_e is unknown. The algorithm can `update' e to obtain the edge weight $we \in Ae$. The task is to output the edge set of a minimum spanning tree after a minimum number of updates. An algorithm is k-update competitive if it makes at most k times as many updates as the optimum. They presented a 2-update competitive algorithm if all areas *Ae* are open or trivial, which is the best possible among deterministic algorithms. The condition on the areas *Ae* is to exclude degenerate inputs for which no constant update competitive algorithm can exist. Next, they consider a setting where the vertices of the graph correspond to points in Euclidean space and the weight of an edge is equal to the distance of its endpoints. The location of each point is initially given as an uncertainty area, and an update reveals the exact location of the point. they gave a general relation between the edge uncertainty and the vertex uncertainty versions of a problem and use it to derive a 4-update competitive algorithm for the minimum spanning tree problem in the vertex uncertainty model. Again, they show that this is best possible among deterministic algorithms.

Pettie et al (2002) conducted a study into "An Optimal Minimum Spanning Tree Algorithm" they established that the algorithmic complexity of the minimum spanning tree problem is equal to its decision-tree complexity. Specifically, they presented a deterministic algorithm to find a minimum spanning tree of a graph with n vertices and m edges that runs in time $O\big(T*(m,n)\big)$ where $T*$ is the minimum number of edge-weight comparisons needed to determine the solution. According to them, though their time bound is optimal, the exact function describing it is not known at present. The current best bounds known for $T*$ are $T*(m, n) = (m)$ and $T*(m, n) = O(m \, . \, \alpha(m, n))$, where α is a certain natural inverse of Ackermann's function.

Even under the assumption that $T*$ is super linear, they shown that if the input graph is selected from $G_{n,m}$, their algorithm runs in linear time with high probability, regardless of n, m, or the permutation of edge weights.

Jackson et al (2009) study scaling properties of random MSTs using a relation between Kruskal's greedy algorithm for finding the MST, and bond percolation. They solve the random MST problem on the Bethe lattice (BL) with appropriate wired boundary conditions and calculated the fractal dimension D=6 of the connected components. Viewed as a mean-field theory, the result implies that on a lattice in Euclidean space of dimension d, there are of order W^{d-D} large connected components of the random MST inside a window of size W, and that d = d_c = D = 6 is a critical dimension. This differs from the value 8 suggested by Newman and Stein. They also critique the original argument for 8, and provided an improved scaling argument that again yielded d_c = 6. The result implies that the strongly-disordered spin-glass model has many ground states for d > 6, and only of order one below six. The results for MSTs also apply on the Poisson-weighted infinite tree, which is a mean-field approach to the continuum model of MSTs in Euclidean space, and is a limit of the BL. In companion they developed an epsilon = 6-d expansion for the random MST on critical percolation clusters.

Deanne et al (2009) conducted a study on “A Contrasting Look at Network Formation Models and Their Application to Minimum Spanning Tree”. They provided a review of the minimum spanning tree (MST) problem. They also introduced it as a formal optimization problem, which is non-trivial to solve as an integer linear program for large problems. They then review two centralized algorithms, Kruskal’s (1956) and Prim’s (1957), which take advantage of the special network structure in order to more easily solve the MST problem. In contrast to the global algorithms, they review the decentralized algorithm of Gallagher et al (1983) that utilizes “message passing” between nodes to solve for the MST problem.

Brennan (1982) presented a modification to Kruskal's classic minimum spanning tree (MST) algorithm that operated similar in a manner to quicksort; splitting an edge set into “light" and “heavy" subsets. Osipov et al. (2009) further expanded this idea by adding a multi-core friendly filtering step designed to eliminate edges that were obviously not in the MST (Filter-Kruskal). Currently, this algorithm seems to be the most practical algorithm for computing MSTs on multi-core machines.

It is well established that the GMST is a subset of edges in the Delaunay triangulation of a point set (Franco et al 1985) and similarly established that this method is inefficient for any dimension $d > 2$. It was shown by Agarwal et al. (1991) that the GMST problem is related to solving bichromatic closest pairs for some subsets of the input set.

According to (Erickson, 1995) It is known that the GMST problem is harder than bichromatic closest pair problem, and bichromatic closest pair is probably harder than computing the GMST.

The early work in this area was done by Steele (1988) who showed how the length of the minimum spanning tree (MST) of a complete graph from the Euclidean distance between the feature vectors is directly proportional to the integral of an appropriately defined power of the probability density defined on the feature space. Later a group at University of Michigan, Hero et al. (2002) showed how the aforementioned relation between the length of the spanning tree and the integral of the power of the density can be used to estimate the Renyi entropy which in turn can be used to perform registration.

2.4 SPANNING TREE AND HEALTH ISSUES

Spada et al (2004) conducted a study into "Use of the Minimum Spanning Tree Model for Molecular Epidemiological Investigation of a Nosocomial Outbreak of Hepatitis C Virus Infection". The minimum spanning tree (MST) model was applied to identify the history of transmission of hepatitis C virus (HCV) infection in an outbreak involving five children attending a paediatric oncology-haematology outpatient ward between 1992 and 2000. They collected blood samples from all children attending since 1992, all household contacts, and one health care worker positive for antibody to HCV (anti-HCV). HCV RNA detection was performed with these samples and with smears of routinely collected bone marrow samples. For all isolates, they performed sequence analysis and phylogenetic tree analysis of hypervariable region 1 of the E2 gene. The MST model was applied to clinical-epidemiological and molecular data. Sequence analysis and phylogenetic tree analysis revealed a high identity among the isolates. The MST model applied to molecular data, together with the clinical-epidemiological data, allowed them to identify the source of the outbreak and the most probable patient-to-patient chain of transmission. The management of central venous catheters was suspected to be the probable route of transmission. In conclusion, the MST model, supported by an exhaustive clinical-epidemiological investigation, appears to be a useful tool in tracing the history of transmission in outbreaks of HCV infection.

Kayhan Erciyeş (2010) provided a detailed review of basic algorithm techniqueues as applied to bioinformatics problems. According to him, dynamic programming and graph algorithms are of particular concern due to their wide range of applications in bioinformatics. Some of the bioinformatics problems do not have solutions in polynomial time and are called NP-Complete. For these problems, approximation algorithms may be used. They show several examples where approximation algorithms may be used to provide sub-optimal solutions to these problems.

Pozzi et al (2008) considered ‘Dynamical correlations in financial systems’

They discuss and compare the stability and robustness of two methods: the Minimum Spanning Tree (MST) and the Planar Maximally Filtered Graph (PMFG). They constructed such graphs dynamically by considering running windows of the whole dataset. They study their stability and their edges's persistence and came to the conclusion that the Planar Maximally Filtered Graph offers a richer and more significant structure with respect to the Minimum Spanning Tree, showing also a stronger stability in the long run.

2.5 SPANNING TREE AND PROBABILITY

Pegah et al (2011) conducted a study into "Stochastic Minimum Spanning Trees and Related Problems" and came out with the following findings: they investigated the computational complexity of minimum spanning trees and maximum flows in a simple model of stochastic networks, where each node or edge of an undirected master graph can fail with an independent and arbitrary probability. They also showed that computing the expected length of the MST or the value of the max-flow is NP-Hard, but that for the MST it can be approximated within O(log n) factor for metric graphs. The hardness proof for the MST applies even to Euclidean graphs in 3 dimensions. They also show that the tail bounds for the MST cannot be approximated in general to any multiplicative factor unless P = NP. More generally, they also considered the complexity of linear programming under probabilistic constraints, and show it to be NP-Hard. If the linear program has a constant number of variables, then it can be solved exactly in polynomial time.

Farah (2005) looked into "Feature Subset Selection Using Minimum-Cost Spanning Trees". They investigated the use of minimum spanning trees (MST), a graph-theoretic approach, as a criterion function in ranking feature subsets.

The results showed that feature subset selection using minimum cost spanning trees is effective. The MST's approaches result with better subsets than the statistical methods in almost every case. The MST approaches finish in a reasonable amount of time, however the statistical methods are generally faster. Interestingly and un-intiutively their method often did not perform as well as the method by Friedman and Rafsky (1979) on average, therefore further investigation is suggested, in particular a modified K-Nearest Neighbour classifier is proposed. The Karhunen-Loeve transformation generally favoured the Mahalanobis distance classifier, and the K-Nearest Neighbour classifier performed better than the Mahalanobis classifier across the range of data sets experimented with.

Costa et al (2003) in their work on "Manifold Learning with Geodesic Minimal Spanning Trees" they considered the closely related problem of estimating the manifold's intrinsic dimension and the intrinsic entropy of the sample points. They view the sample points as realizations of an unknown multivariate density supported on an unknown smooth manifold. They present a novel geometrical probability approach, called the geodesic-minimal-spanning-tree (GMST), to obtaining asymptotically consistent estimates of the manifold dimension and the R´enyi α-entropy of the sample density on the manifold. The GMST method simply constructs a minimal spanning tree (MST) sequence using a geodesic edge matrix and uses the overall lengths of the MSTs to simultaneously estimate manifold dimension and entropy. They illustrate the GMST approach for dimension and entropy estimation of a human face dataset.

CHAPTER 3

METHODOLOGY

3.0 OVERVIEW

This chapter will concentrate on the methodology that will be used in the study. How data will be collected and analysed in the study. Also, the mathematical theories and instruments that will be used for the study is capture in this chapter.

3.1MATHEMATICAL MODEL

The first step towards solving instances of such large sizes is to find a mathematical formulation of the combinatorial problem such that every solution of the real problem corresponds to a solution of the mathematical model, and vice versa. A graph consists of points and lines connecting pairs of points. The graph associated with the problem is constructed as follows. Each tourist site in the study is declared to be a point. We connect a pair of points by a line segment if there is a direct link between the associated tourist sites. Moreover, with each line segment we associated a length(weight) which corresponds to the distance it takes to travel between the two points that the line connects. Every roundtrip of the tourist site corresponds to some subset of the lines.

3.1.1 GRAPH THEORY

In simple language, graph theory is the study of graphs. Graph has several definitions depending on the researcher or writer. A graph is called *regular* if all its vertices have the same degree.

A graph is a combination of vertices or nodes and edges which connect in some fashion. These graphs are either directed or undirected based on their orientation. If the edges of the graph are represented with ordered pairs of vertices, then the graph *G* is called directed or oriented, otherwise if the pairs are not ordered, it is called undirected or nonoriented graph. If two vertices connected by an edge *ek* = (v_i, v_j) are called end vertices or ends of *ek*. In the directed graph, the vertex v_i is called the *source*, and v_j the *target* vertex of edge *ek*. The elements of the edge set *E* are distinct i.e., more than one edge can join the same vertices. Edges having the same end vertices are called *parallel edges*.

If *ek* = (v_i, v_j), i.e., the end vertices are the same, then *ek* is called a *self-loop*. A graph *G* containing parallel edges and or self-loops is a multigraph. A graph having no parallel edges and self-loops is called a *simple graph*. The number of vertices in *G* is called its *order*, written as |V |; its number of edges is given as |E|. A graph of order zero(0) is called an *empty graph*, and of order one is simply called *trivial graph*. A graph is *finite* or *infinite* based on its order. Two vertices v_i and v_j are neighbours or *adjacent* if they are the end vertices of the same edge *ek* = (v_i, v_j). Two edges e_i and e_j are *adjacent* if they have an end vertex in common, say v_i, i.e., e_i = (v_i, v_j) and e_j = (v_k, v_m). Let *G* = (*V*,*E*) and $G^1 = (V^1, E^1)$ be two graphs. $G^1 = (V^1, E^1)$ is a subgraph of *G* ($G^1 \subseteq G$) if $V^1 \subseteq V$ and $E^1 \subseteq E$, i.e., the graph *G* contains graph G^1

If $G^1 \subseteq G$ and V^1 spans all of *G*, i.e., $V^1 = V$ then G^1 is a spanning subgraph of *G*. Let *G*= (*V*, *E*) be a graph with sets $V = \{v_1, v_2, v_3, \ldots\}$ and $E = \{e_1, e_2, e_3, \ldots\}$.

A walk in a graph *G* is a finite nonempty alternating sequence $v_0, v_2, v_3, \ldots, \ldots\ v_{k-1}$, *ek*, *vk* of vertices and edges in *G* such that $e_i = (v_i, v_{i+1})$ for all $1 \le i \le k$.

A walk is a trail if all its edges are distinct. A trail is closed if its end vertices are the same, otherwise it is opened. A *simple walk* is a walk in which no edge is repeated. The *length* of a walk is its number of edges in the walk. A walk is *closed* when the first

and last vertices, v_0 and v_n are the same. The degree or valency of the vertex V given by d(V) is the number of edges that have V as an endpoint. If *d(V)* = 0, then *V* is called an *isolated* vertex while a vertex of degree 1 is called *pendant*. The edge incident with a pendant vertex is called a *pendant edge.*

A *path* is a walk in which no vertex is repeated. Closed walks are also called *circuits.* A *cycle* of length *n* is a closed walk of length *n*, *n* ≥3, in which the vertices $v_0, v_1, \ldots v_{n-1}$ are all different. A graph that contains no cycles at all is called *acyclic;* a connected acyclic graph is called a *tree* acyclic graphs are called *forests.* If a graph represents a road system, a common weight is the length of the corresponding stretch of road. Weights also often represent costs or durations. The weight of a *path P* is the sum of the weights of the edges in *P.* A cycle that passes through every vertex in a graph is called a *Hamilton cycle* and a graph with such a cycle is called *Hamiltonian.* A *Hamilton path* is a path that contains every vertex. A vertex *V* is called a *cutpoint* in G if G - *V* contains more components than G does; in particular if G is connected, then a cutpoint is a vertex *V* such that (G – *V*) is disconnected. A *bridge* (or *cutedge)* is an edge whose deletion increases the number of components. A minimal collection of edges whose deletion disconnects G is called a *cutset* in G. A cutset partitions the vertex-set *V(G)* into two nonempty components, say *A* and *B,* such that the edges joining vertices in *A* to vertices in *B* are precisely the edges of the cutset. For example, the figure 3.1 below shows the cutpoint and cutedge.

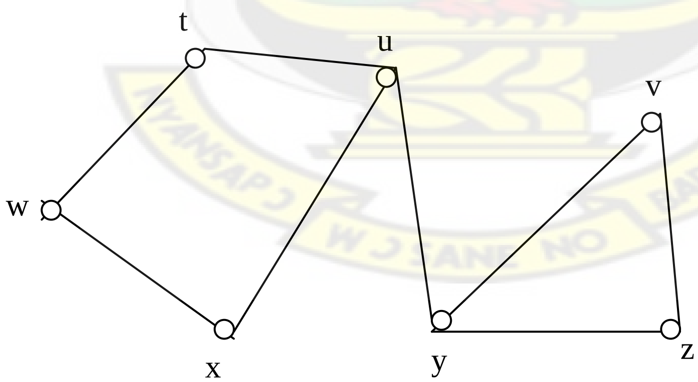


Figure 3.1: A connected graph

From figure 3.1, removal of u or y disconnects the graph, also removal of uy disconnects the graph. Hence u and y are the cut-points and uy is the cut-edge.

Graph theory is applied in almost every day life from communication net work, town or city water supply system, road net work system e.t.c

3.1.2 NETWORKS

A *network* is a specific type of graph, where associated with each arc or node is additional information, such as the cost or capacity of the arc or the demand at a node. Networks are integral to a variety of systems that we rely upon each day. Our transportation system is made up of a variety of networks including road, rail and airline networks. Our electrical system is a network of wires that ensures power reaches homes and businesses. Communications systems, including the Internet, are expanding beyond the typical hard wired lines to include wireless networks. Even individuals' relationships with one another can be viewed as a network of social ties.

Each of these networks plays an important role in society. A transportation network provides a means for goods and people to move from a starting location to a destination. The electrical system continuously balances generation with fluctuating user demand. Communication networks and the Internet provide a massive increase in the amount of easily obtainable information, and they also dramatically decrease the amount of time required to transfer information around the world. The study of social networks is increasingly popular, with sites such as Facebook and Twitter capturing evolving relationships between millions of people. The analysis of networks is even helping to fight terrorism by identifying terrorist networks so that we can determine where it is most effective to disrupt them. The study of networks is actually centuries old. Graph theory dates back to Leonard Euler in 1736 (Biggs, Lloyd and Wilson, 1998), when he proved there was no feasible solution to the Konigsberg Bridge Problem. The development of random graph theory in the 1940s and 1950s generated great interest in the characteristics of graphs and networks (Watts et al, 2006). Most recently, the advent of "network science" during the last decade has witnessed renewed interest in the large-scale properties of graphs (National Research Council, 2005).

Erdős and Rényi (1959) pioneered the exploration of random graphs models, which generated interest in graph and network theory. More recently, the study of network science has focused attention on "small-world networks" and "scale-free networks."

Small-world networks (Watts et al, 1998) are networks that have high local clustering and have path lengths between arbitrarily chosen nodes that are still relatively short. Another area of increasing importance is the use of Hastily Formed Networks (HFNs) in response to humanitarian aid and disaster relief operations, such as a Hurricane Katrina scenario (Denning, 2006). These types of networks require rapid coordination and information between a variety of agencies.

3.1.3 SPANNING TREES

A subgraph T of an undirected graph $G = (V, E)$ is a spanning tree of G if it is a tree and contains every vertex of G, where V is the vertex set and E is the edge set . Let $T \subseteq E$ be the set of edges of a spanning tree of a weighted graph. The weight (or cost) of T is given by

$$\cos t(T) = \sum w(e_i)$$

where $w(e)$ is the weight of edge e.

A graph is called a tree if and only if it is connected and does not contain any circuit. In order words, a tree of the graph G is a connected acyclic sub graph of G. A graph is called a forest if it contains no circuits and the components of the forest are the trees. Vertices of degree one is called a leave and all edges are called branches of the tree. A spanning tree of graph G is a tree of G containing all vertices of G. The edges of a spanning tree are called branches. A forest is simply a set of trees spanning all the vertices of G.

Ahuja et al (1993) A graph, or network, G= (N, A) consists of a set *N* nodes and a set of *A* arcs. The number of nodes is $n = |N|$ and the number of arcs is $m = |A|$. An arc from node *i* to node *j* is denoted as (i, j) where $i, j \in N$. If *G* is a directed graph then $(i, j) \neq (j, i)$, but if *G* is an undirected graph then $(i, j) = (j, i)$. A subgraph of G= (N, A) is a graph G'= (N', A') if N'=N and A'=A. It is a spanning subgraph of G = (N, A) if N'= N. A tree is a connected graph that contains no cycles. A subtree is a connected subgraph of a tree. A spanning tree of *G* is a tree that is a spanning subgraph of *G* and has exactly *n*-1 arcs.

A finite graph can contain finitely many spanning trees. For example, the graph below may have the following spanning trees.

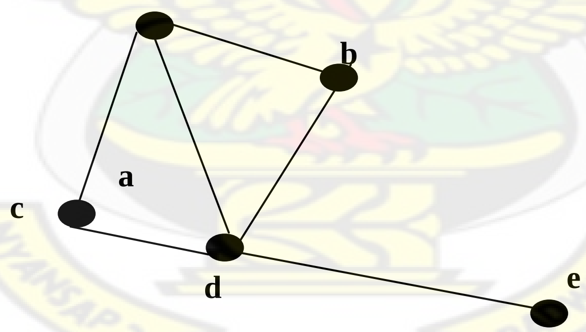


Figure 3.2: undirected graph

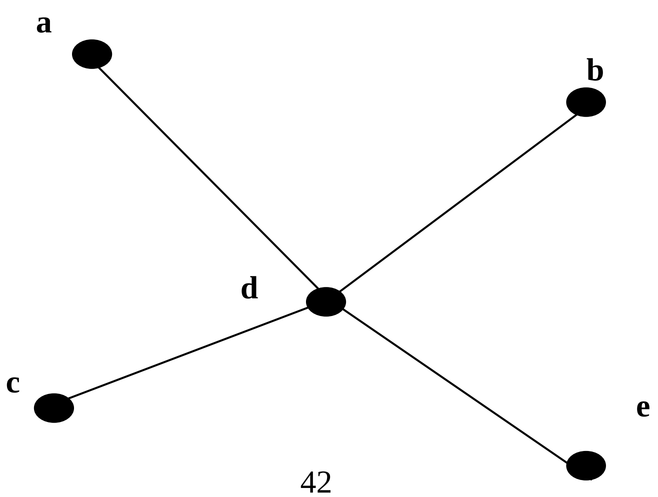


Figure 3.2.1: spanning tree 1

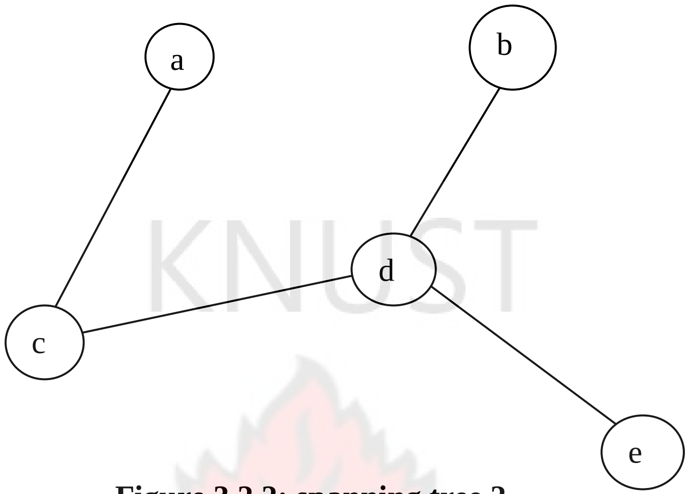


Figure 3.2.2: spanning tree 2

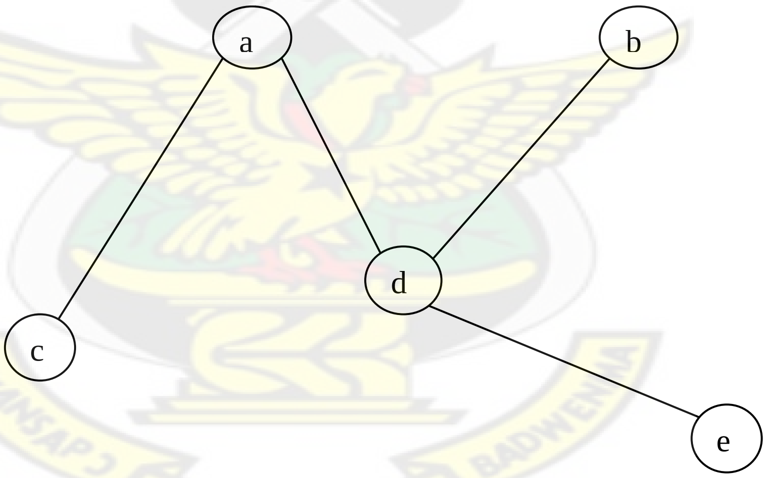


Figure 3.2.3: spanning tree 3

A weighted graph is a graph, in which each edge has a weight (some real number).

The Weight of a Graph is the sum of the weights of all the edges of the graph.

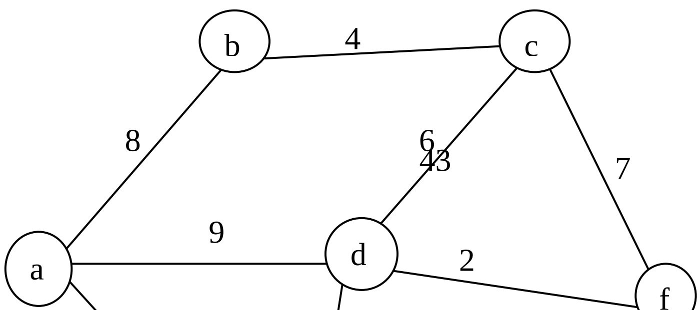


Figure 3.3: weighted Graph

The total weight of figure 3.4 is 5+8+9+4+4+6+2+7 = 45

3.1.4 MINIMUM SPANNING TREES

The problem of finding a minimum spanning tree of a graph is an important building block for many graph algorithms, and has been extensively studied by many scientists and mathematicians. The problem has applications in the design of distributed computer and communication networks, wiring connections, transportation networks among cities, and designing pipe capacities in flow networks

A minimum spanning tree is a mathematical graph theory construct used to connect a set of points at the least possible length of total connecting lines (Biggs, 1994.)

The minimum spanning tree problem is one of the most fundamental and intensively studied problems in network optimization with many theoretical and practical applications. The Minimum Spanning Tree (MST) problem is one of the most typical and well known problems in combinatorial optimization. Borůvka (1926) (cited in Graham and Hell, 1985), used Euclidean MST to find the most economical construction of an electricity network in Moravia.

Let G = (V,E) be a connected graph. A spanning tree is a graph where all the nodes in the graph are connected in some way with the requirement that there are no cycles in the graph. If each edge has a weight or cost connected to it, denoting how much you have to pay in order to use the edge, the total sum of the costs of the edges can vary from one edge to another, in the same graph. The one out of these possibilities which has the lowest sum is called the minimum spanning tree. One of the pioneers in solving the problem of MST was R.C. Prim.

3.1.5 ALGORITHMS FOR SOLVING MST PROBLEM

To find a minimum spanning tree for a given input graph there are several algorithms available, for example, the methods of Kruskal, Prim, Sollin or Borůvka. An algorithm is a systematic logical procedure for solving a problem.

Kruskal and Prim's algorithms for solving Minimum Spanning Tree problem will be considered. The methods have the following optimal conditions: Cut Optimality Conditions and Path Optimality Conditions.

A spanning tree T^* is a minimum spanning tree if and only if it satisfies the following cut optimality conditions: For every tree edge (i, j)$\in T$, $w_{ij} \leq w_{kl}$ holds for every edge (k,l) contained in the cut formed by deleting edge (i,j) from T^*. This is the fundament of the Prim's algorithm. w_{ij} is the weight of the edge from i to j.

A spanning T^* tree is a minimum spanning tree if and only if it satisfies the following path optimality conditions: For every nontree edge (k,l) of G, $w_{ij} \leq w_{kl}$ holds for every edge

(i, j) contained in the path in T^* connecting nodes k and l. This is the basis of Kruskal's algorithm and implies that edges can be added to the current minimum spanning tree in decreasing order of weights except when this operation would result in a cycle.

3.1.5.1 KRUSKAL'S ALGORITHM

Kruskal's algorithm is an algorithm in graph theory that finds a minimum spanning tree for a connected weighted graph. This means it finds a subset of the edges that forms a tree that includes every vertex, where the total weight of all the edges in the tree is minimized. If the graph is not connected, then it finds a minimum spanning forest (a minimum spanning tree for each connected component).

The basic version of Kruskal's algorithm consists of the following four steps:

Step 1. Initialize an empty tree K and sort the edges set E in nondecreasing order of their weights.

Step 2. Take the edge (i, j) with the minimum weight w_{ij} from E.

Step 3. Remove edge (i, j) from E and examine node i and node j:

i. If either node i or node j is not part of K, add edge (i, j) and its nodes to K.

ii. If node i and node j belong to different components of K, then add edge (i, j) and its nodes to K.

Step 4. If |K| = n − c then Stop else go to Step 2.

3.1.5.2 Kruskal's Algorithm Pseudocode

Kruskal (G;w)

1. $A = \emptyset$
2. For each vertex $v \in G.V$
3. Make-Set(v)

4. Put the edges $G.E$ into list sorted by nondecreasing weight
5. For each $(u, v) \in G.E$ taken from the sorted list
6. If find-set(u) ≠ find-set(v)
7. $A = A \cup \{(u, v)\}$
8. $\mathrm{Union}(\mathrm{u}, \mathrm{v})$
9. Return A

Make-Set(v) creates a set containing a single element v.

Find-Set(v) finds a set containing v.

Union(u; v) merges sets containing u and v into a single set.

A is the minimum spanning tree.

Example: Consider the undirected graph below

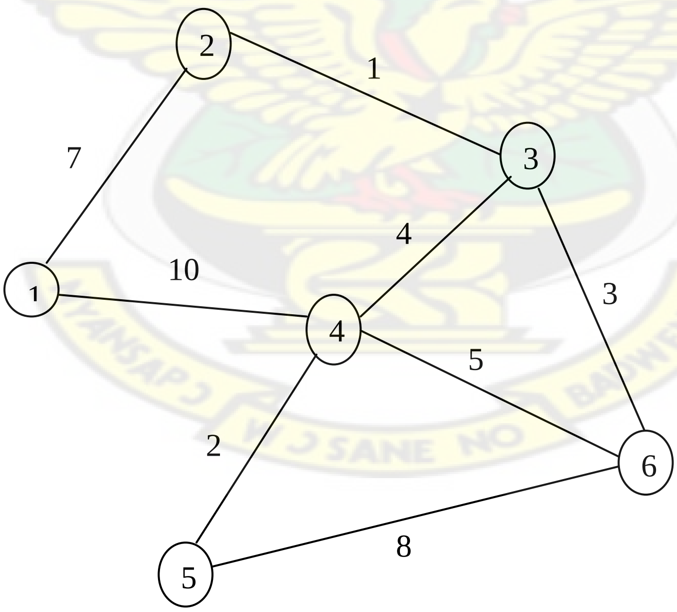


Figure 3.4: undirected weighted graph

The vertex set V= {1, 2, 3, 4, 5, 6} and the edge set

E={(1,2), (1,4), (2,3), (3,4), (3,6), (4,5), (4,6), (5,6)}. Sorting and ordering in nondescending order of their weight yields:

$e_1 = (2,3) = 1, e_2 = (4,5) = 2, e_3 = (3,6) = 3, e_4 = (3,4) = 4, e_5 = (4,6) = 5, e_6 = (1,2) = 7,$

$e_7 = (5,6) = 8, e_8 = (1,4) = 10$

The smallest edge is (2, 3), that is from vertex 2→3

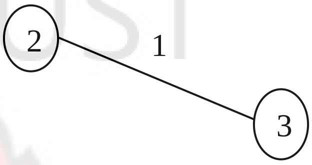


Figure 3.4.1: partial tree 1

The next smallest edge is (4, 5), that is from vertex 4→5, will be added to the partial tree to form.

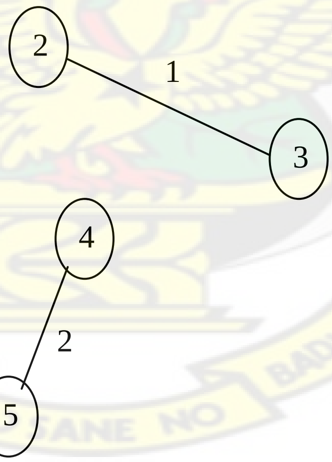


Figure 3.4.2: partial tree 2

The next smallest edge is (3, 6), that is from vertex 3→6 since it does not form a circuit, will be added to the partial tree to form

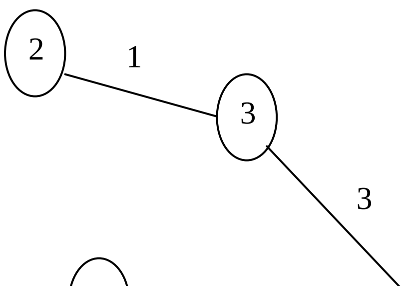


Figure 3.4.3: partial tree 3

The next smallest edge is (3, 4), that is from vertex 3→4, since it does not form a circuit it is added to the partial tree to form

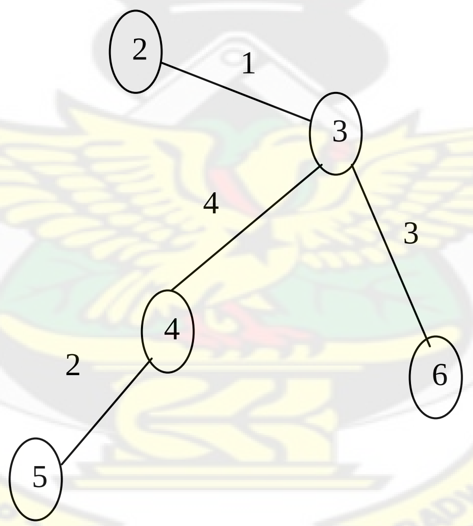


Figure 3.4.4: partial tree 4

The next smallest edge is (4, 6), that is from vertex 4→6, but adding this edge to the partial tree will create a circuit between the vertices 3→4→6→3, therefore this cannot be added to the partial tree.

The next smallest edge is (1, 2), that is from vertex 1→2, adding this to the partial tree does not form any circuit, therefore it will added to form

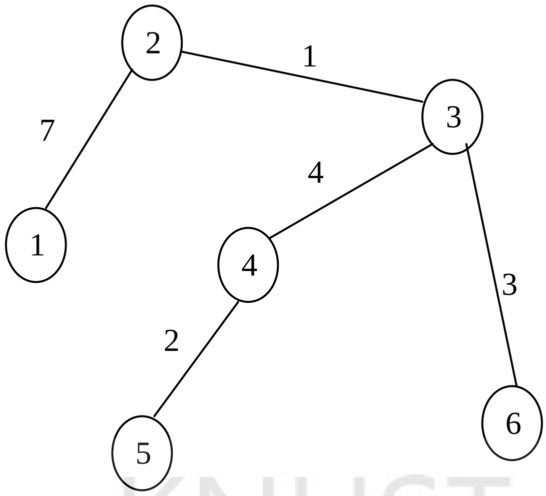


Figure 3.4.5: minimum spanning tree

The next smallest edge is (5, 6), that from vertex 5→6, but this will form circuit between vertices 5→6→3→4→5 therefore cannot be added to the partial tree.

The next smallest edge is (1, 4), that is from vertex 1→4, this also create circuit between the vertices 1→4→3→2→1, this can also not be added to the partial tree. Therefore, the minimum spanning tree for the graph is figure 3.4.5

3.1.5.2 PRIM'S ALGORITHM

Prim's algorithm is one of the best algorithms for solving minimum spanning tree problems. The algorithm represents an $n-node$ network as a square matrix with $n\,rows$ and $n\;columns$. Entry (i, j) of the matrix gives the distance or edge d_{ij} or e_{ij} from node i to node j, which is finite if i is directly linked to j and infinite otherwise.

The algorithm is a greedy algorithm; it starts by selecting an arbitrary vertex as the root of the tree. It then grows the tree by adding a vertex that is closest (has the shortest edge to) the current tree, and adding the shortest edge from any vertex already in the tree to the new vertex. The algorithm terminates once all vertices have been added to the tree. The sum of all added edges is the cost of the minimum spanning tree (MST). The serial computational complexity of the algorithm is $\Theta(N^2)$. Prim's algorithm has the property that the edges in the set A always form a single tree.

If for any undirected graph, $G = (V, E)$, where V is the set of vertices and E is the set of edges. For each $v \in V$ the cost[v] denotes the minimum weight among all edges connecting v to the vertex in the tree T, and the parent[v] denotes parent of v in T. During the algorithm's execution vertices v that are not in T are organised in the minimum-priority queue Q, partition according to cost[v]. Lines 1 to 3 set each cost[v] to infinity usually written as ∞. The parent of each vertex is set to $NULL$ because the construction of the minimum spanning tree is yet to begin. Lines 4 to 6 choose an arbitrary vertex r from V as the root of the tree (starting vertex). The minimum priority queue is set to be all vertices from V. Since r is the starting vertex, cost[r] is set to zero.

During the execution of the while loop from lines 7 to 12, r is the first vertex to be extracted from Q and processed. Line 8 extract a vertex u from Q based on key cost, thus moving u to the vertex set of T. Line 9 considers all vertices adjacent to u. The while loop updates the cost and the parent fields of each vertex v adjacent to u that is not in T. If parent[v] $\neq NULL$ then cost[v] < ∞ and cost[v] is the weight of the edge v to some vertex already in T. Lines 13 and 14 construct the edge set of the minimum spanning tree and return this edge set.

Pseudocode of Prim's algorithm

Given a connected weighted graph $G = (V, E)$ with a weight function w and a minimum spanning tree T can be derived from the code below.

1. for any $v \in V$ do
2. cost[v] $\leftarrow \infty$

3. parent[v] ← $NULL$

4. r ← arbitrary vertex of V

5. Cost[r] ← 0

6. $Q \leftarrow V$

7. While $Q \neq \{\ \}$

8. u ← extractMin(Q)

9. for each $v \in$ adja(u) do

10. if $v \in Q$ and $w(u, v) < cost[v]\ then$

11. parent[v] ← u

12. cost[v] ← $w(u, v)$

13. $T \leftarrow \{(v, parent[v]| \ \ v \in V - \{r\}\}$

14. Return T

3.4.5.3 HOW THE ALGORITHM WORKS

Prim's algorithm works from a starting point and builds up the spanning tree step by step, connecting edges into the existing solution. The algorithm can be stated as follows:

Prim's MST algorithm (from a network)

Step 0: Choose any element r; and set $S = \{r\}$ and $A = \{\ \}$

(r is the root of the spanning tree)

Step 1: Find the lightest edge such that one endpoint is in S and the other is in $V \backslash S$. Add this edge to A and its (other) endpoint to S.

Step 2: If $V \backslash S = \{\ \}$ then stop and output the minimum spanning tree (S, A)

Otherwise go to step 1.

Prim's MST algorithm (from a distance matrix)

Step 0: With the matrix representing the network, choose a starting vertex. Delete the row corresponding to that vertex.

Step 1: Label with '1' the column corresponding to the start vertex and ring the smallest undeleted entry in that column.

Step 2: Delete the row corresponding to the ringed entry.

Step 3: Label (with the next number) the column corresponding to the deleted row.

Step 4: Ring the lowest undeleted entry in all labelled columns.

Step 5: Repeat the last three steps until all rows are deleted. The ringed entries represent the edges in the minimum connector.

When there is a tied in the smallest values, it is broken arbitrary.

EXAMPLE 1

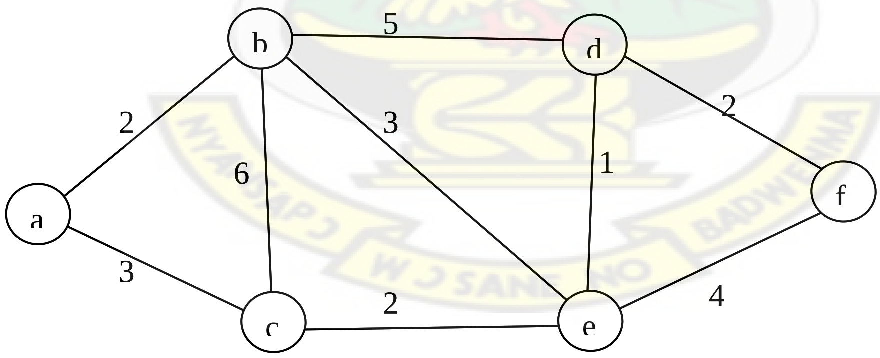


Figure 3.5: Hypothetical network

For example, figure 3.5 a hypothetical network can be put in distance matrix form and solve by Prim's algorithm as follows:

SOLUTION BY MATRIX METHOD

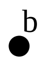
Table 3.1 Node distance matrix 1

	a	b	c	d	e	f
a	∞	2	3	∞	∞	∞
b	2	∞	6	5	3	∞
c	3	6	∞	∞	2	∞
d	∞	5	∞	∞	1	2
e	∞	3	2	1	∞	∞
f	∞	∞	∞	2	4	∞

Choose a starting vertex say **b,** delete row **b,** and look for the smallest entry in column **b.**

Table 3.2 solution matrix 1

		1				
	a	b	c	d	e	f
a	∞	2	3	∞	∞	∞
~~b~~	~~2~~	~~∞~~	~~6~~	~~5~~	~~3~~	~~∞~~
c	3	6	∞	∞	2	∞
d	∞	5	∞	∞	1	2
e	∞	3	2	1	∞	∞
f	∞	∞	∞	2	4	∞



The edge **ba** is the smallest edge joining **b** to the other vertices. Put edge **ba** into the solution. Delete row **a** and look for the smallest entry columns **b** and **a.**

Table 3.3 solution matrix 2

	2	1				
	a	b	c	d	e	f
~~a~~	~~∞~~	~~2~~	~~3~~	~~∞~~	~~∞~~	~~∞~~
c	3	6	∞	∞	2	∞
d	∞	5	∞	∞	1	2
e	∞	3	2	1	∞	∞
f	∞	∞	∞	2	4	∞

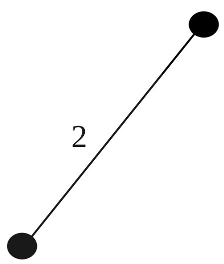


Figure 3.5.1 partial

be is the smallest edge joining **b** and **a** to the other vertices. Put **be** into the solution and delete row **e**. Look for smallest entry in columns **b, a** and **e.**

Table 3.4 solution matrix 3

	2	1			3	
	a	b	c	d	e	f
c	3	6	∞	∞	2	∞
d	∞	5	∞	∞	1	2
~~e~~	~~∞~~	~~3~~	~~2~~	~~1~~	~~∞~~	~~∞~~
f	∞	∞	∞	2	4	∞

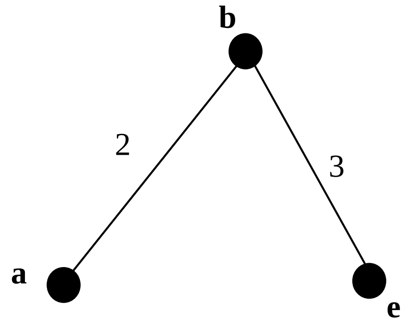


Figure 3.5.2: Partial connection

ed is the smallest edge joining **b, a,** and **e** to the other vertices. Put the edge **ed** into the solution and delete row **d**. Look for the smallest entry in columns **b, a, e** and **d**

Table 3.5 solution matrix 4

	2	1		4	3	
	a	b	c	d	e	f
c	3	6	∞	∞	2	∞
~~d~~	~~∞~~	~~5~~	~~∞~~	~~∞~~	~~1~~	~~2~~
f	∞	∞	∞	2	4	∞

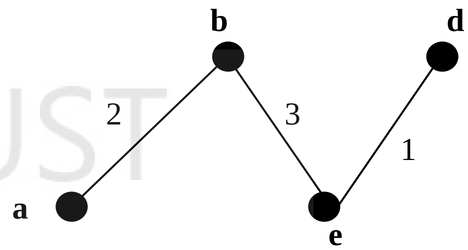


Figure 3.5.3: Partial connection 1

df is the smallest edge joining **b, a, e,** and **d** to the other vertices**.** Put **df** into the solution and delete row **f.** Look for the smallest entry in columns **b, a, e, d** and **f.**

Table 3.6 solution matrix 5

	2	1		4	3	5
	a	b	c	d	e	f
c	3	6	∞	∞	2	∞
~~f~~	~~∞~~	~~∞~~	~~∞~~	~~2~~	~~4~~	~~∞~~

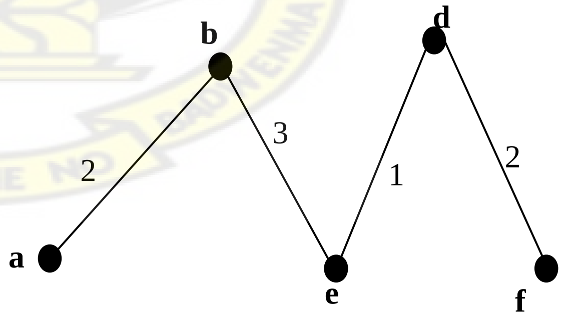


Figure 3.5.4: Partial connected network

ec is the smallest edge joining **b, a, e, d,** and **f** to the other vertices. Put **ec** into the solution.

Table 3.7 solution matrix 6

	2	1	6	4	3	5
	a	b	c	d	e	f
~~c~~	~~3~~	~~6~~	~~∞~~	~~∞~~	~~2~~	~~∞~~

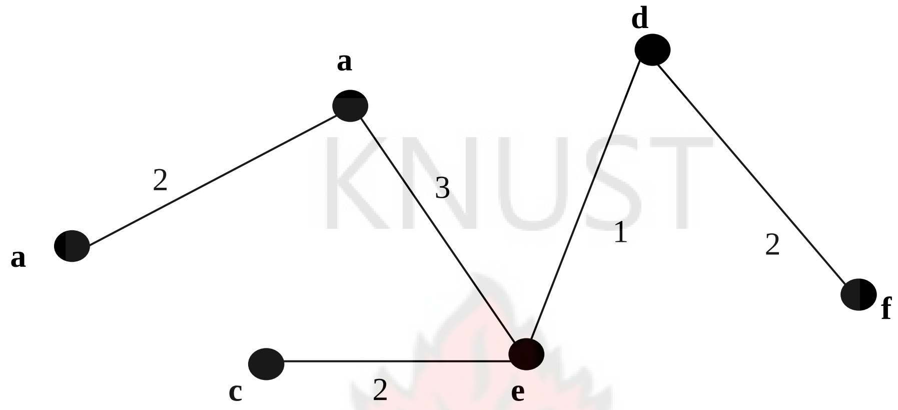


Figure 3.5.5: Minimum spanning tree (MST)

Figure 3.5.5 is the minimum spanning tree for figure 3.5

The minimum length (weight or cost) of the network (figure 3.5) is 10 units, that is the total sum of the edge values (2+3+2+1+2=10)

EXAMPLE 2

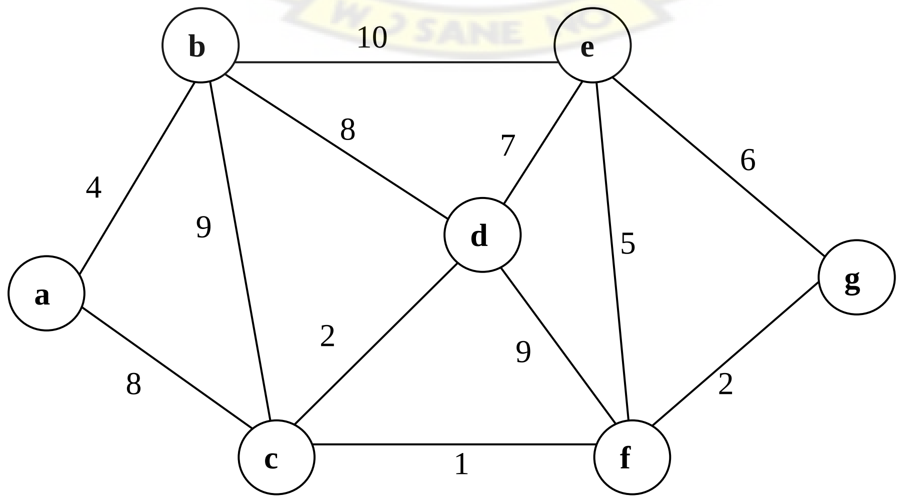


Figure 3.6: Undirected Graph

From figure 3.6, the minimum spanning tree from Prim's algorithm can be obtained from the network method as follows:

SOLUTION BY NETWORK METHOD

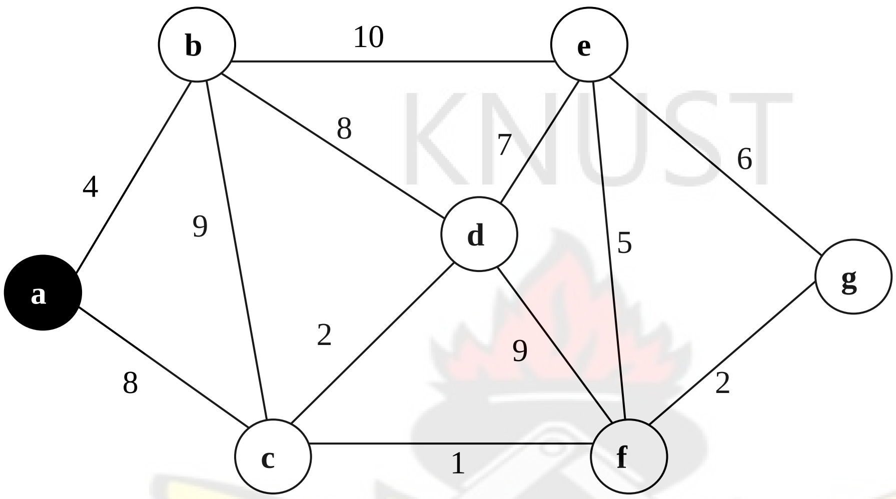


Figure 3.6.1: Undirected Graph 1

Step 0: $S = \{a\}$

$V \backslash S = \{b, c, d, f, g\}$

$A = \{\ \}$

$lightest\ edge = \{a, b\}$

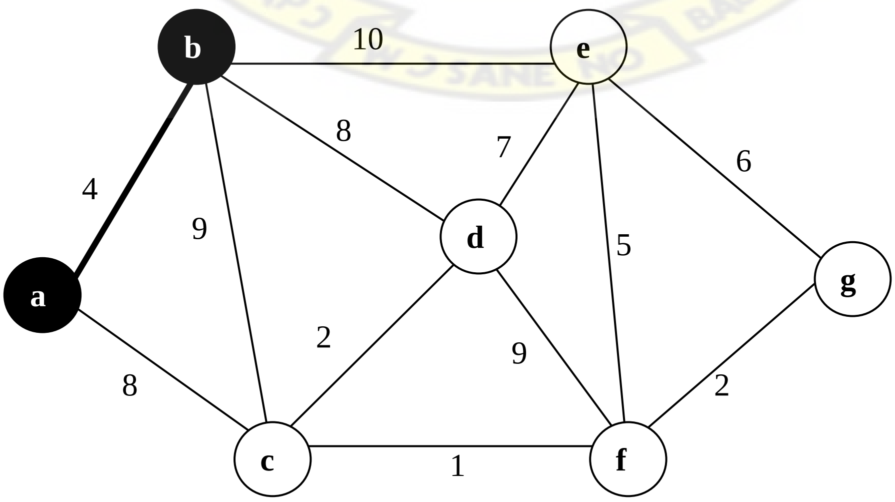


Figure 3.6.2: Undirected Graph 2

Step 1: $S = \{a, b\}$

$V \backslash S = \{c, d, e, f, g\}$

$A = \{\{a, b\}\}$

$lightest\ edge = \{b, d\}, \{a, c\}$

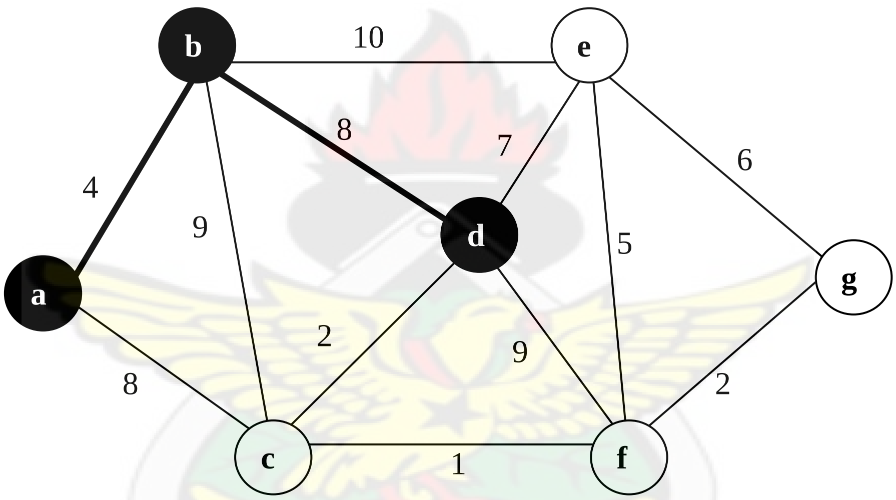


Figure 3.6.3: Undirected Graph 3

Step 1.1: $S = \{a, b, d\}$

$V \backslash S = \{c, e, f, g\}$

$A = \{\{a, b\}, \{b, d\}\}$

$lightest\ edge = \{d, c\}$

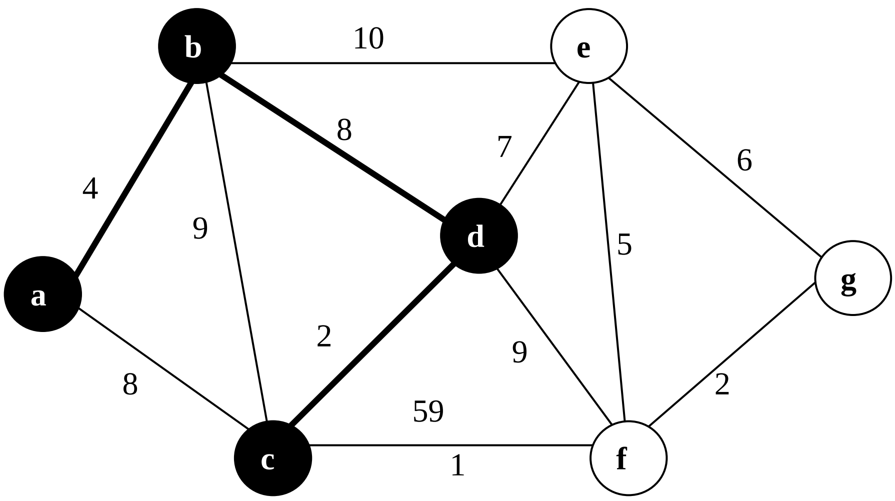


Figure 6.3.4: Undirected Graph 4

Step 1.2: $S = \{a, b, d, c\}$

$V \backslash S = \{e, f, g\}$

$A = \{\{a, b\}, \{b, d\}, \{d, c\}\}$

$lightest\ edge = \{c, f\}$

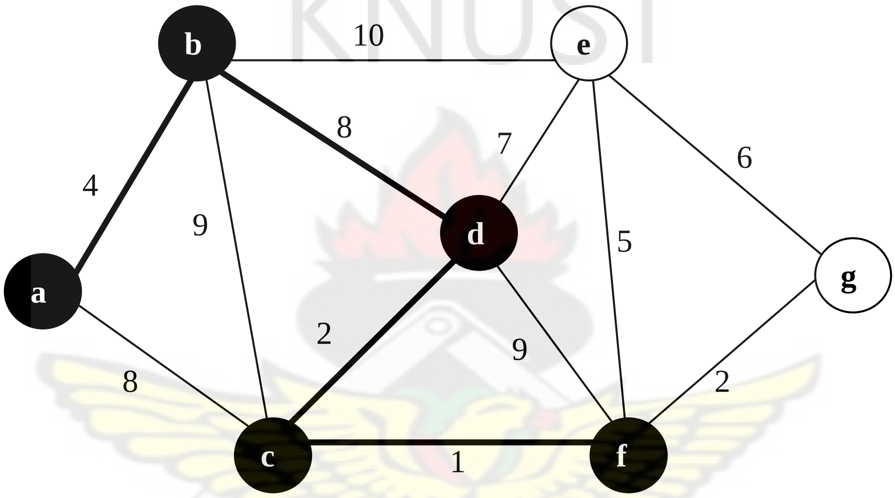


Figure 3.6.5: Undirected Graph 5

Step 1.3: $S = \{a, b, d, c, f\}$

$V \backslash S = \{e, g\}$

$A = \{\{a, b\}, \{b, d\}, \{d, c\}, \{c, f\}\}$

$lightest\ edge = \{f, g\}$

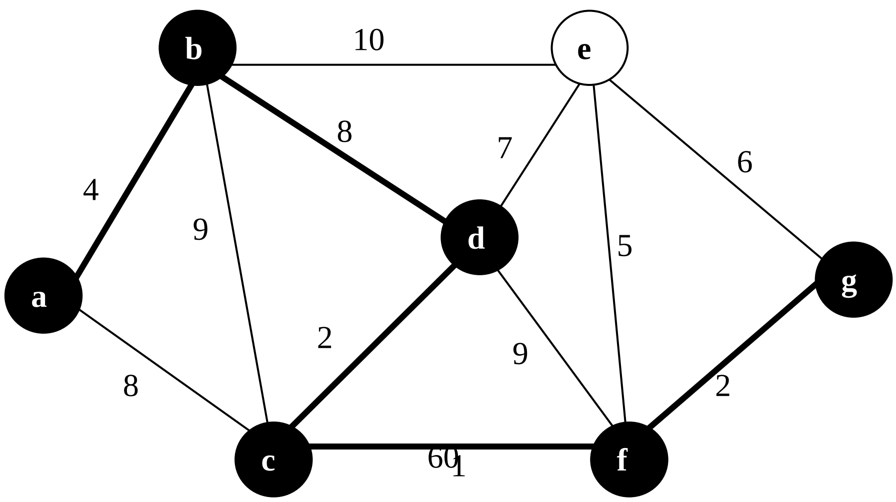


Figure 3.6.6: Undirected Graph 6

Step 1.4: $S = \{a, b, d, c, f, g\}$

$V \backslash S = \{e\}$

$A = \{\{a, b\}, \{b, d\}, \{d, c\}, \{c, f\}, \{f, g\}\}$

$lightest\ edge = \{f, e\}$

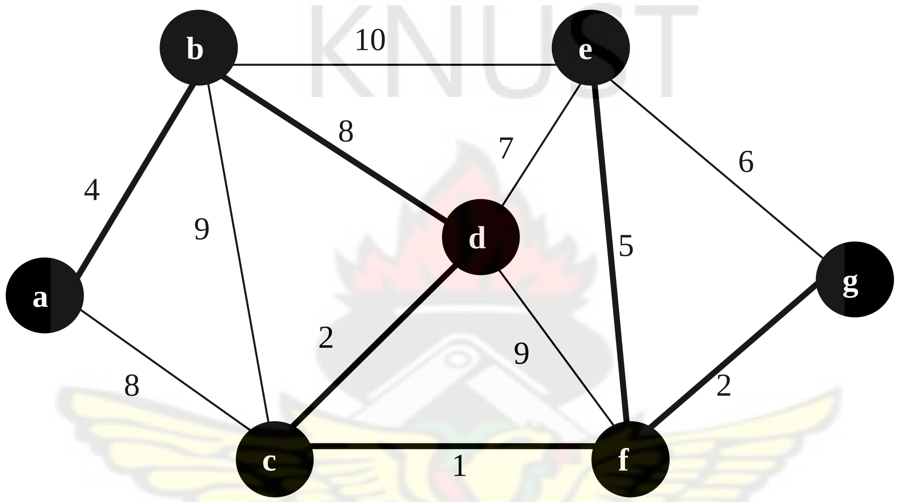


Figure 3.6.7: Undirected Graph 7

Step 2: $S = \{a, b, c, d, e, f, g\}$

$V \backslash S = \{\,\}$

$A = \{\{a, b\}, \{b, d\}, \{d, c\}, \{c, f\}, \{f, g\}, \{f, e\}\}$

Since $V \backslash S = \{\,\}$, it means MST is complete.

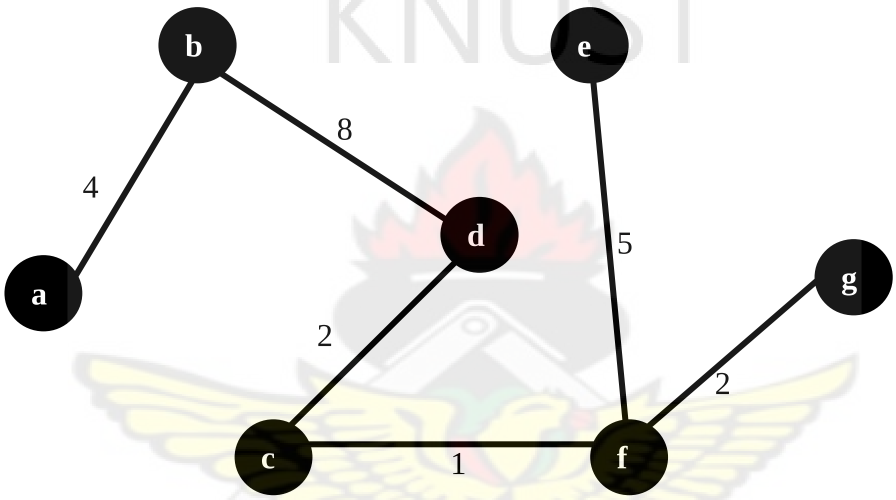


Figure 3.6.8: Minimum Spanning Tree

The minimum total weight of the tree is 4+8+2+1+2+5 = 22

CHAPTER 4

DATA ANALYSIS AND MODELLING

4.0 OVERVIEW

This chapter analyses the secondary data used as input to determine the minimum spanning tree route for the major Tourist centers in the region. The data analysis was done by using the developed MATLAB programme (prims.m) Specifically, Prim's algorithm was used to determine the minimum spanning tree route for the various locations selected for the study.

4.1 DATA COLLECTION

The data used for the study were collected for the Ghana Tourist Board, national and regional headquarters Accra and Sunyani respectively, Ghana Highways Authority (Accra) and the Department of Feeder Roads (Sunyani). The names and location of the various tourist centers in the study were obtained from the Ghana Tourist Board regional headquarters Sunyani.

Feeder Road Map of Brong Ahafo region was obtained from the department of feeder roads (Sunyani) and thread and metre rule was used to calculate the distances along the roads connecting the various tourist sites under consideration in the region. The data in table 4.1 was taken from the Ghana Tourist Board in Sunyani.

Since some of the distances were measured using a thread, there is a possibility of error, therefore the distances might not be exactly the same as on the ground.

Table 4.1 Some Tourist sites in Brong Ahafo and their descriptions

NAME	LOCATION	TYPE OF ATTRACTION	OWNERSHIP
Boaben Fiema Monkey Sanctuary(Fi)	22km north of Nkoranza	Natural Attraction	Community
Nkyeraa Waterfalls	37km from Wenchi	Natural Attraction	Community
Tano Boase Sacred Grove (Ta)	8km from Techiman	Natural Attraction	Community
Duasidan Wildlife Sanctuary (Du)	10km south west of Dormaa Ahenkro	Natural Attraction	Community
Forikrom Boten Shrine and Caves(Fo)	8km off Techiman-Nkoranza road	Natural Attraction	Community
Buoyem Caves Bats Colony (Bu)	11km from Techiman	Natural Attraction	Community
Bono Manso Slave Site (Ma)	12km from Techiman	Historical Attraction	Community
Fuller Falls (Fu)	10km from Kintampo town	Natural Attraction	Community
Kintampo Waterfalls (Ki)	5km from town	Natural Attraction	Community
Dr. K.A. Busia Mauseleum (We)	Wenchi on the Busia road	Historical Attraction	Community

Hani Archaeological Site (Ha)	50km from Wenchi	Historical Attraction	Community

Table 4.1 shows the distance from one tourist site to the nearest town. The remaining distances from one tourist site to the other were calculated using thread to measure the distance along the roads connecting the sites. These distances with the aid of the scale of the regional road map(1:360,000) were converted into kilometres.The data in table 4.1 together with the distances calculated from the regional road map of Brong Ahafo were put together to for the distance node matrix in table 4.2 Where there is a direct link between the two towns, real value is assigned to it otherwise it is infinity.

The distance between any two tourist sites is put in distance matrix form as follows;

Table 4.2: Distance node matrix for the network in figure 4.1

	Du	**Ha**	**We**	**Nk**	**Fu**	**Ki**	**Fi**	**Ma**	**Ta**	**Bu**	**Fo**
Du	∞	109	∞	∞	∞	∞	∞	∞	∞	147	∞
Ha	109	∞	50	71	∞	∞	∞	∞	∞	∞	∞
We	∞	50	∞	37	∞	∞	64	∞	∞	37	∞
Nk	∞	71	37	∞	49	∞	∞	∞	∞	∞	∞
Fu	∞	∞	∞	49	∞	15	∞	∞	∞	∞	∞
Ki	∞	∞	∞	∞	15	∞	53	74	∞	∞	∞
Fi	∞	∞	64	∞	∞	53	∞	∞	∞	∞	28
Ma	∞	∞	∞	∞	∞	74	∞	∞	4	∞	∞
Ta	∞	∞	∞	∞	∞	∞	∞	4	∞	19	12
Bu	147	∞	37	∞	∞	∞	∞	∞	9	∞	19
Fo	∞	∞	∞	∞	∞	∞	28	∞	12	19	∞

The distance node matrix in table 4.2 was used to construct the network in figure 4.1

Figure 4.1 shows network of the various tourist sites in the Brong Ahafo region of Ghana and the distances between them. The distances are in kilometres (km)

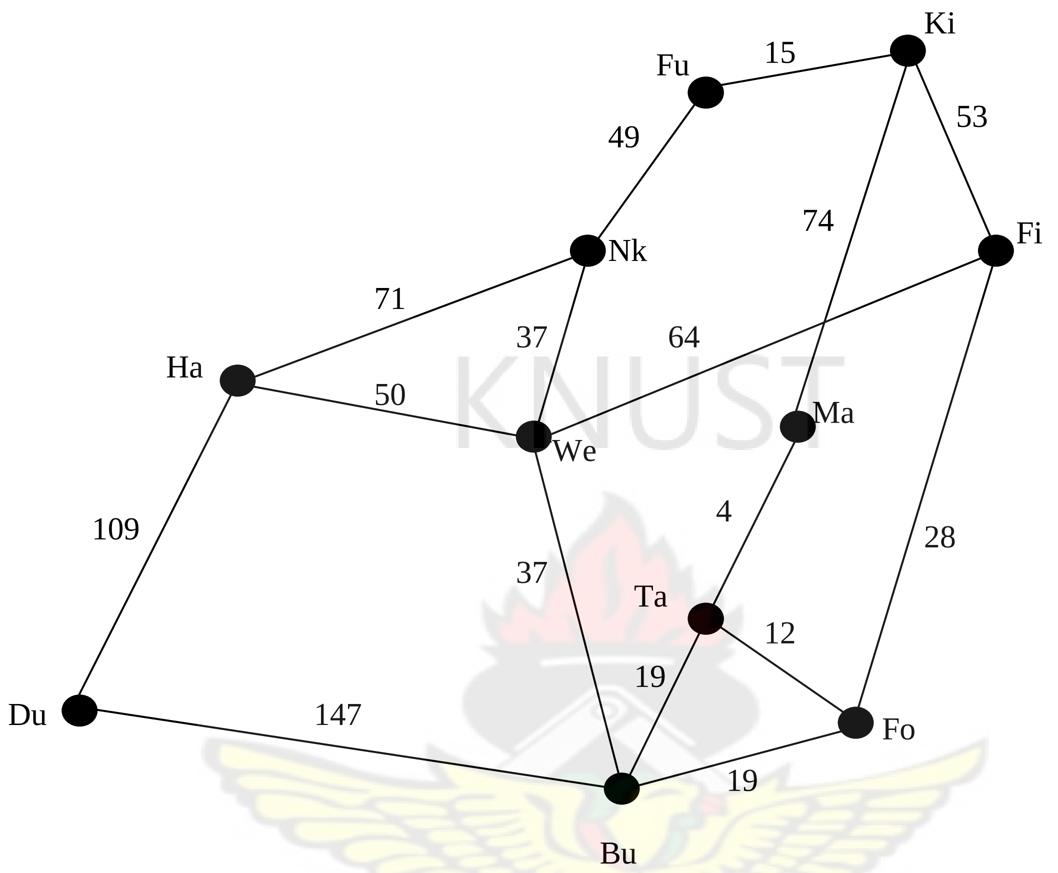


Figure 4.1: Road Network of major tourist sites in Brong Ahafo

4.2 DATA ANALYSIS

Prim's algorithm was used for the analysis of the data. The data in table 4.2 was used as input into the MATLAB program (prims.m) for Prim's algorithm for the analysis. The MATLAB program is in the appendix A. The output from the code starting from **Du** as follows;

Iteration	Starting Node	End Node	Distance (Km)
1.	Du	Ha	109
2.	Ha	We	50
3.	We	Bu	37
4.	Bu	Fo	19
5.	Fo	Ta	12
6.	Ta	Ma	4
7.	Fo	Fi	28
8.	We	Nk	37
9.	Nk	Fu	49
10.	Fu	Ki	15

The output generated from the code was used to construct the minimum spanning tree in figure 4.2

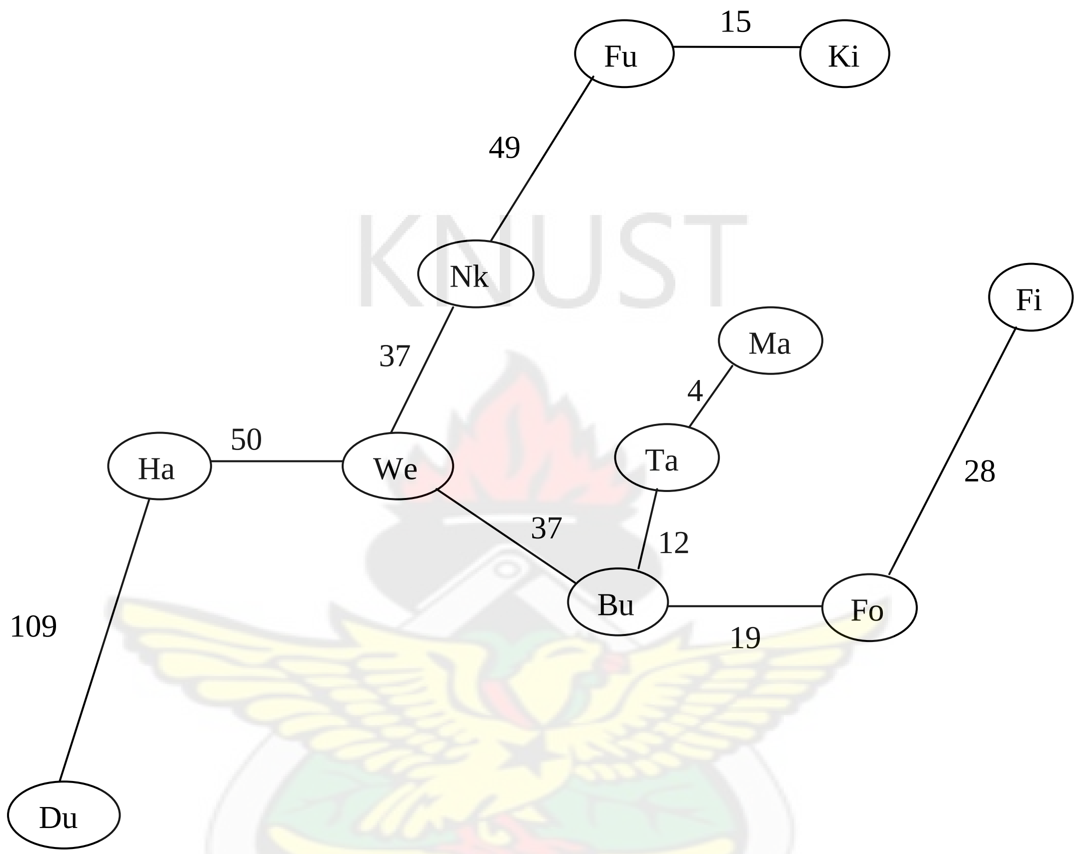


Figure 4.2: Minimum Spanning Tree route for the tourist centers.

The total weight of the of the spanning tree is $360km$ (109+50+37+37+49+15+12+19+28+4)

4.3 FINDINGS

The study has found the minimum spanning tree route for the major tourist centers in the Brong Ahafo region of Ghana. The roads in this spanning tree must thus be given the highest priority in terms of development and maintenance if tourism in the region is to flourish. From figure 4.2, **Du**, **Ki** ,**Ma** and **Fi** are terminal nodes and are the most remote tourist centers from the viewpoint of accessibility. **We** and **Bu** are the two hubs in the network and are the least remote centers from the view point of accessibility. **We** is a major hub of the network and from it one can directly access 3 centers namely **Ha**, **Bu** and **Nk**; and indirectly every other center of the network. **Bu** is the other major hub of the network and from it one can also directly access 3 centers namely **We**, **Ta** and **Fo**; and indirectly every other center of the network.

The two towns **We** and **Bu** can be economically developed and materially resourced to be able to accommodate tourists who plan to spend more than a day in the region. For example, a tourist can stay at **We** and go to **Ha,** afterwards visit **Du** for one trip and return to base and in another next trip, visit **Nk**, **Fu** and finally **Ki** and again return to base. The tourist could then subsequently move base to **Bu** from which he can visit **Ta** and then **Ma** and return to base at **Bu**. In the next trip he could visit **Fo** and then **Fi** and again come back to **Bu**. Such an itinery is both simple convenient and inexpensive. The development of the two hub towns (one of which is already a district capital) should be in terms of providing more hotel facilities, motels, guest houses, garages, internet cafes, restaurants to cater for the expected large number of tourists who would visit and be based at these places. Other tourist paraphernalia and accessories such as tourist maps, historical documents, informative brochures, cultural artefacts, souvenirs, wood carvings etc could also be made available there creating in the process a lot of jobs in the area.

Table 4.3: Distance from one tourist site to another.

FROM	TO	DISTANCE (KM)
Duasidan Wildlife Sanctuary	Hani Archaeological site	109
Hani Archaeological site	Busia Mausoleum-Wenchi	50
Busia Mausoleum-Wenchi	Buoyem Caves& Bats Colony	37
Busia Mausoleum-Wenchi	Nkyeraa waterfalls	37
Nkyeraa waterfalls	Fuller Falls	49
Fuller Falls	Kintampo Waterfalls	15
Buoyem Caves& Bats Colony	Forikrom Boten Shrine & Caves	19
Forikrom Bote Shrine & Caves	Boabeng Fiema Monkey Sanctuary	28
Forikrom Bote Shrine & Caves	Tano Boase Sacred Grove	12
Tano Boase Sacred Grove	Bono Manso Slave Site	4

Table 4.3 gives the shortest distance from one tourist site to the nearest tourist site.

CHAPTER 5

CONCLUSION AND RECOMMENDATION

5.1 SUMMARY

The study sought to find the Minimum Spanning Tree (MST) route for the major tourist sites in the Brong Ahafo region of Ghana using the Prim's Algorithm. The purpose was to find shortest route covering tourist sites in the region affording minimum cost without compromising access and thus connecting all the tourist sites under consideration.

In all, a total of eleven (11) tourist sites were selected for the study, these sites were selected based on how popular the place is in terms of the number of people who visit there. The data used for the study was gathered from the regional offices of the Ghana Tourist Board and the Department of Feeder Roads in Sunyani. Some other information was gathered from the Ghana Highways Authority (Accra) and Ghana Tourist Board (Accra)

Prim's Algorithm was used as the mathematical tool to solve the problem. Also, a MATLAB code for the Prim's algorithm was developed to compute the minimum spanning tree. The analysis shows that a total of approximately three hundred and sixty kilometers (360Km) distance long will be covered by touring all the eleven tourist sites used in the study.

5.2 CONCLUSION

In this study, we presented all the tourist sites selected in the form of a network and Prim's algorithm was applied to the network to construct the minimum spanning tree. The study shows that the minimum total distance connecting the eleven selected tourist sites in the Brong Ahafo region of Ghana is three hundred and sixty kilometres (360km) long.

5.3 RECOMMENDATIONS

Since the tourism industry is becoming more and more lucrative, it is recommended that the routes in the spanning tree be focussed on by the department of roads in terms of maintenance to reduce high cost constructing more roads, since all the communities in the network could be accessed by the roads in the spanning tree. The government can also adopt this by developing the educational and health facilities in these nodal towns rather than trying to build the facilities in every community since all the other communities can easily access the nodal towns in the network.

It is also recommended that further study be conducted in the region to include more other tourist sites that were not included in this study and also try to look at the cost involve in touring these sites. The future study may be conducted to cover the entire country.

Lastly, it is recommended to the Ghana Tourist Board and other travel and tour operators to adopt this study as a basis for developing a data base of distances between the various tourist sites in order to advice tourist who want to visit these sites for them to reduce the number of hours and fuel consumption during their tour.

REFERENCES

1. Addo T. Samuel (2006), Geography Transport and Development: Spatial Trinity, Ghana Uninversity Press.
2. Aditee Kumthekar (2004), Affine Image Registration Using Minimum Spanning Tree Entropies. A thesis presented to the Graduate School, University of Florida.
3. Ahlswede R., Li N., & Yeng R.W (2000), Netwok Information Flow. IEEE Transactions on Information Theory. Vol. 46, pp. 1204-1216.
4. Ahuja R. et al (1993), Network Flows. Prentice-Hall.
5. Alfre O. Hero, Bing Ma, Olivier Michel & John Gorman (2002), Applications of entropic spanning graphs. *IEEE Signal Processing Magazine*, vol. 19, no. 5, pp. 85–95.
6. Alfred O. Hero & Olivier Michel (1998), Robust entropy estimation strategies based on edge weighted random graphs. SPIE'98, San Diego (CA).
7. Bernard Chazelle (2000), A Minimum Spanning Tree Algorithm with Inverse-Ackermann Type Complexity. Journal of the ACM. Vol. 47,No. 6, pp. 1028-1047.
8. Brong Ahafo Road Map, DFR GIS Unit (2007). Scale 1:360000.
9. CTK Network Aviation (2010), Ghana's Tourist Attractions, Ghana Tourist Board.
10. David Joyner, Minh Van Ngugen & Nathann Cohen (2011), Algorithmic Graph Theory. http://code.google.com/p/graph-theory-algorithms-book/
11. Dippon M. Christian & Kennerth E. Train (1999), The Cost Of The Local Telecommunication Network: A Comparison of Minimum Spanning Trees and the HAI Model. National Economic Research Associates. Entropy Estimation Workshop NIPS 2003.
12. Eva Milkova (2008), Combinatorial Optimization: Mutual Relations among Graph Algorithms. WSEAS TRANSACTIONS on MATHEMATICS. Issue 5, Vol. 7 pp. 293-302.

13. Farhad Faghani & Ghasen Mirjalily (2008), Selecting the Best Spanning Tree in Metro Ethernet Network using Genetic Algorithm. IJCSN: International Journal on Computer Science and Network Security. Vol. 8, No.6, pp. 106-111.

14. Gary et al (2009), Chromatic Graph Theory. CRC Press, USA.

15. Gerdts Mathias (2009), Combinatorial Optimisation. MSM 3M02b, www.web.mat.bham.ac.uk/m.gerdts.

16. Getz, D. (2008). Event tourism: definition, evolution, and research. Tourism Management, 29, 403-428.

17. Graham R.L. & Pavol Hell (1995), On the History of Minimum Spanning Tree Problem. Annals of the History of Computing. Vol.7, No.1, pp.43-57.

18. Hall et al (2006), The Geography of Tourism and Recreation: Environment, place and space (3rd Edition). Routledge Taylor & Francis Group. http://www.ebookstore.tandf.co.uk/.

19. http//:www.questionbase.50megs.com.

20. Jeff Erickson (2009), Minimum Spanning Tree. http://creativecommons.org/liceses/by-nc-sa,

21. Jose A. Costa & Alfred O. Hero (2003), Applications of Entropic Graphs to Nonparametric Estimation. Dept. Of Electrical Engineering and Computer Science, University Michigan.

22. Kayhan Erciyes (2010), Design of Algorithms for Bioinformatics. kayhan.erciyes@izmir.edu.tr.

23. Leiper, N. (1990), Tourist attraction system. *Annals of Tourism Research*, 17, 367-384.

24. Liu Yueyang, Ji Hong & Yue Guangxin (2006), An Energy-Efficient PEGASIS Based Enhanced Algorithm in Wireless Sensor Networks. China Communication 2006,Technology Forum, pp. 91-96.

25. Loni Taylor (2010), Randomly Generated Edge-Disjoint Minimum Spanning Trees based Data Gathering Trees. http://www.answers.com/topic/cdma.

26. Luis F. Caro Perez (2009), Improving Resource Utilization in Carrier Ethernet Technologies. A Thesis presented the graduate school, University de Girona, Catalonia, Spain.

27. Maki, D. P. & Thompson, M. (1996), Finite Mathematics (4th Ed.), New York: McGraw-Hill.

28. Maleq Khan, Gopal Pandurangan & Anil V.S. Kumar (2006), Distributed Algorithms for Constructing Approximate Minimum Spanning Tree in Wireless Networks. IEEE Transactions on Parallel Distributed System, pp. 1-14.

29. Mert R. Sabuncu & Peter Ramadge (2008), Using Spanning Graphs for Efficient Image Registration. IEEE Transaction on Image Processing. Vol. 17, No. 5, pp. 788-798.

30. Michael C. Hall & Stephen J. Page (2006), The Geography of Tourism and Recreation, (3rd Edition) Taylor & Francis e-Library.

31. Mohan et al (2001), Estimation of Temperature and Humidity from MST Radar Observations. Annales Geophysicae, Vol. 19, pp. 855–861.

32. Narsingh Deo (1974), Graph Theory with Application to Engineering and Computer Science. Prentice-Hall, Inc., Englewood Cliffs, N.J.

33. Papadimitriou et al (1998), Combinatorial Optimisation-Algorithms and Complexity. Dover Publications.

34. Pearce, D. (1992), Tourist organisations. New York: Wiley.

35. Pergah Kumousi & Subhash Suri (2011), Stochastic Minimum Spanning Trees and Related Problems. Santa Barbara, CA931061, USA.

36. Rama Murthy P. (2007). Operations Research (2nd Ed.), New Delhi: New Age International.

37. Robert C. Prim (1957), Shortest Connection And Some Generalisations. The Bell System Technical Journal, Nov.1957, pp. 1389-1401.

38. Seth Pettie & Vijaya Ramchandran (2002), An Optimal Minimum Spanning Tree Algorithm. Journal of the ACM. Vol. 49, No. 1, pp.16-34.

39. Spada Enea et al (2004), Use of the Minimum Spanning Tree Model for Molecular Epidemiological Investigation of a Nosocomial Outbreak of Hepatitis C Virus Infection. Journal of Clinical Microbiology, Vol.42. No.9, pp. 4230-4236.

40. Taha Hamdy A. (2007), Operations Research; An Introduction (8thEdition), Pearson Prentice Hall.

41. Thomas Erlebach & Michael Hofhan (2008). Computing Minimum Spanning Tree with Uncertainty. Symposium on Theoretical Aspects of Computer Science (Bordeaux), pp. 277-288.

42. Toygar M. Karadeiziz (1996), Degree Constrained Minimum Spanning Tree Problem. CMPE 582 Computer Networks Project Report.

43. UNWTO, Tourism Highlights 2007 Edition.

44. Winston W. L (2004), Operations Research; Applications and Algorithms (4th Edition), Books/Cole-Thomas Learning, Belmont.

45. Wolfram Höpken, Ulrike Gretzel & Rob Law (2009), Information and Communication Technologies in Tourism. Springer Wien New York.

APPENDIX

```
n=input('Enter the number of towns :  ');
Towns=cell(1,n);
for i=1:n;  %This allows entry of the names of the towns
    Towns(1,i)=input('Enter the names of the towns one by one: ');
end
Townsinv=Towns';
D=zeros(n);
for i=1:n;
   for j=1:n;
       if i==j;
           Dis(i,j)=inf;
       elseif i~=j;
      disp(['Distance from ' Townsinv(i,1) 'to' Towns(1,j)]);
   Dis(i,j)=input('Distance =' );
       end
   end
end

m=Dis;m
[m1 n1]=size(m);
x=zeros(n1);%replacement matrix
v=zeros(n1,1);%substitution vector/direction vector
n=v;%path vector
```

```
for i=1:m1;
    v(i,1)=i;
end
f=input('choose a starting point for the search  ');
n(1,:)=v(f,:);v(f,:)=[];
[g h]=min([m(:,n(1,:))]);
[g1 h1]=min(g);
deci=h(1,h1);
x(1,:)=m(f,:);m(f,:)=[];
for i=2:n1-1;
         [t1 t2]=size(g);
      for h2=1:t2;
   if ismember(deci,v)==0;
        g(1,h1)=inf;
        [g1 h1]=min(g);
        deci=h(1,h1);
   elseif ismember(deci,v)==1; break;
   end
      end
      %v==deci;[deci y]=find(ans);
    n(i,:)=v(deci,:);v(deci,:)=[];
    d1=zeros(n1-i+1,i);
    for j=1:i;
         d1(:,j)=m(:,n(j,:));
    end
```

```
    [g h]=min(d1);
    [g1 h1]=min(g);
    deci=h(1,h1);
    x(i,:)=m(deci,:);m(deci,:)=[];
end

n(m1,:)=v(1,:);
x(m1,:)=m(1,:);
disp(['the path is given as follows']);
final=cell(n1,1);
final(:,end)=Townsinv(n(:,end));
final
```