Synthesis and Performance Analysis of Network Topology using Graph Theory

U N Bassey[1*] K N Bassey[2] E N Bassey[3]
1 Department of Mathematics, University of Ibadan, Ibadan, Nigeria
2 IT/Telecom Infrastructure Projects, Shell Nig. P. O. Box 1259, Warri, Nigeria
3 Department of Chemical Engineering, Akwa Ibom State University, Ikot Akpaden. P. O. Box 3864, Uyo,Nigeria
[*] E-mail of the corresponding author: un.bassey@ui.edu.ng

Abstract
In this work, the peculiarities of network topology have been explored to evolve techniques for the solution of practical problems which manifest in the form of graphs. They also give insight into the scope and possible areas for improvement of existing networks and as well the cost implication of incorporating efficiciency factors into new designs. The shortest route algorithm was applied in defining the topology that maximizes reliability during resource transmission. The algorithm was implemented using TORA Software on an Excel platform.
Key words: topology, graphs, vertices, nodes, edges, telecommunication network

1. Introduction

1.1 *Network Topology as a Graph*

Large Scale Systems encountered in the industries constitute a network of related physical units and their interconnections. The structure of such arrangements is referred to as Network Topology. The practical network problems that generally conform to a graph structure include the following (Beasley 2004):

- Communication Networks- telecommunication systems, the internet among others
- Oil /Gas and Water Pipelines
- Road and Rail Networks etc.

1.2 *Challenges*

As the world economy is becoming very competitive there is a great drive to streamline and integrate information flows between diverse businesses to achieve excellence and improved profitability. This has therefore brought a tremendous pressure on the information and communication infrastructure leading to occasional congestion, interferences, and sometimes outright failure. Again since at the moment, the Nigerian economy derives the largest portion of its revenue from the oil and gas sector, the related infrastructure for the operation of this sector is of paramount importance with respect to costs efficiencies and systems sustainability.

It may be noted that these systems are often complex by virtue of the large number of units involved and their interactions. The complexity usually associated with practical systems may be illustrated as in Figure 1 which shows an **IT** Integration Network in an oil and gas environment where infrastructure like VSAT, Routers, Switches, Workstations etc are linked and integrated (Bassey 2001). So the challenge the Engineer is constantly facing is the task of searching for innovative configurations to ensure that the systems work together as desired, perform optimally and meet the requirements of the customers at minimum costs. Consequently, the problems which must be overcome include creating models which are amenable to computer manipulation, evolving algorithms for their implementation, and evaluating the information flows for efficiency (Bassey 2000; Bassey *et al*. 2005). The relationship between network topology and the associated information flows if explored, offers useful principles for tackling these problems.

1.3 *Concepts of Graph*

The synthesis and optimisation of practical networks can be greatly enhanced by simulating the network topologies using graph models. Information flows represent the values of the variables in the models and are the function of the network topology.

A graph is a set of vertices, $V(v_1, v_2, \ldots v_n)$ -, nodes *(N),* points, elements, units subsystems, events etc. and a set of lines joining the nodes called edges, $E(e_1, e_2 \ldots e_n)$ – communication links, flows, streams etc. (Koryachko et al, 1970). Graphs can be directed (digraphs) or undirected.

1.4 *Useful Properties of Graphs*

The network structure affects the information flows within the system (i.e. network performance, throughput, costs, etc). It is therefore necessary to keep the link as simple as possible, and also make connections, which present the least operating costs. Various properties of graphs are helpful in the study of practical systems, and of particular importance to a variety of networks are the properties such as Isomorphism, Subgraphs and Tree Structures.

1.4.1 *ISOMORPHISM*

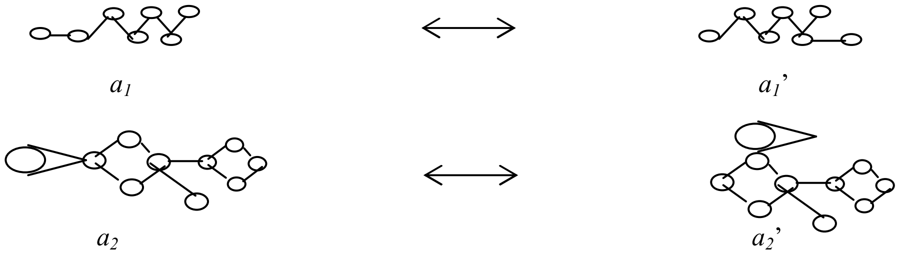


Figure 2, Isomorphic Graphs, $a_1 \equiv a_1$' *and* $a_2 \equiv a_2$'

Two graphs $G = (V, E)$ and $G' = (V', E')$ are Isomorphic when there is a one to one correspondence: ie $V \to V'$, $E \to E'$ such that if $(v_i , v_j) \in V$, $(v_{i'}, v_{j'}) \in V'$ then the edge $e = (e_i, e_j) \in E$, $e' = (e_i', e_j') \in E'$. Figure 2, illustrates Isomorphic Graphs.

Isomorphism forms the basis for representing physical objects (or systems) such as telephone, process, data communication networks, etc by graph models.

1.4.2 Subgraphs

Subgraphs (Kafarov *et al*. 1974) consist of subsets of vertices and subsets of edges both belonging to a complete graph G. Thus for a graph, $G = (V, E)$, $G' = (V', E')$ is the subgraph if : $V' \subseteq U$ and $E' \subseteq E$ and the edges link the vertices V. Figure 3 shows the subgraphs G_1 and G_2 of the complete graph G.

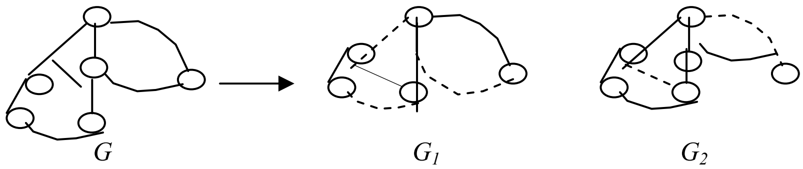


Figure 3. Graph, G and its Subgraphs G_1 and G_2

To break a graph into its component parts is to split it at its cut vertex since at the cut vertex, there is minimum interaction between the component parts. Subgraphs are particularly useful in establishing the concept of decomposing systems into subsystems; links and nodes into segments, etc where each subsystem or segment can be treated as an independent structure. Examples abound in the industry, for instance in the telecommunication system, cut vertex may be the Port Interface Modules (PIM) where station transmission /signal cards are interfaced. In process systems, cut vertex may be the boundaries between the individual units or operations.

1.4.3 *Cyclic Graph and Tree Formation*

Another property of importance is that every cyclic graph contains at least one loop, and that a graph without loops is called a tree. The tree that links all the nodes

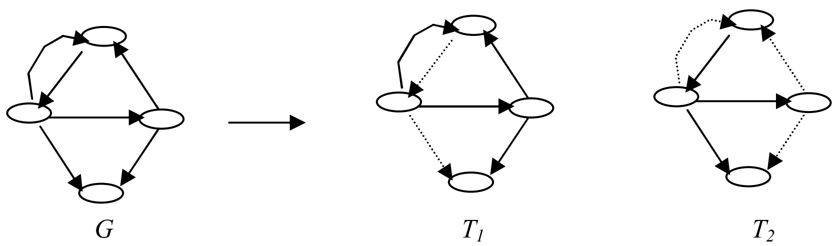


Figure 4. Cyclic Graph G_1 and its Spanning Trees T_1 and T_2.

without allowing loops is called a spanning tree, (Linnhoff et al., 1979). Figure 4 shows a tree structure ***G*** and its spanning trees T_1 and T_2 formed by breaking some edges. The spanning tree forms the basis for the design of minimum cost networks expressed mathematically as follows:

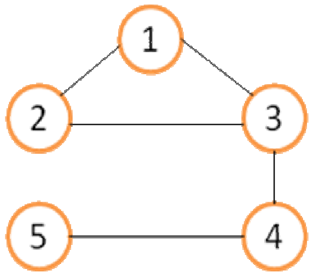
$$E = V\text{-}1. \qquad (1)$$

where: E , V – number of edges and vertices respectively in the network.

1.4.3 *Graph Matrix*

The Boolean Matrix is another interesting property of a graph by which information embodied in the graph can easily be decoded and converted into a machine language (Scheid 1983; Aspen 2000).

Of particular interest in this work, is the Incidence Matrix which is shown in relation to the graph structure, $G = (V, E)$ in Figure 5.



G

$$A = \begin{vmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{vmatrix}$$

Figure 5. Graph and its Incidence Matrix.

The Incidence Matrix, of a graph, G with V vertices may be defined as $|V|$ by $|V|$ matrix , $A = a_{ij}$, whose entries are chosen such that the entry, $a_{ij} = 1$ if ith vertex is joined by an edge to jth vertex in G, and 0 if otherwise, (Nozari, et al. 1981). Other criteria (Ahuja 1982; Koryachko 1970; Kafarov 1974) for generating the graph matrix may also apply. The Adjacency Matrix is a veritable tool in the development of simple solution approaches to network problems.

2. Application of Graph Theory to Network Problems

The graph concepts allow for the configuration of units into a network and the reconfiguration of the existing network into minimum cost or more reliable structure (Ahuja 1982; Steward 1987; Lucey 2002). It may also permit the decomposition of networks into subsystems for better and easier analysis. In practical terms the structures so evolved are known as network topologies. Thus network topologies may change with time via addition or subtraction of nodes or links to allow for determination of suitable information routes or implementation of flow control. Furthermore changes may occur on account of temporary failure of some nodes or due to link congestion (Ahuja 1982). Various algorithms have therefore been developed to serve as tools for appropriately defining network topologies depending on the specific problem objectives. These include minimal spanning tree, shortest - route algorithm, maximum flow algorithm, minimum – cost capacitated network algorithm. Application of these methods has been reported widely and may be found in the works of Wayne (1999), Taha (2008) and Beasley (2004) for transportation, foreign exchange/multinational tax planning networks, Nozari et al. (1981) for computerized facility layout etc.

2.1 *Telecommunication Network Case Study*

A typical telecommunication system (Bassey 2001), used for linking some oil and gas operational centres is illustrated in Figure 6. It consists of a set of nodes which are the equipment such as radios, duplexers, multiplexers, PABX, switches, bridges etc and edges which are links used for voice or data communication. The designed link capacities are as given on the left hand side corners of the nodes where the nodes 1, 2 and 3 are terminal concentrators. Figure 3, emphasizes the operational topology of the network for the connections at node 1 in which possible loops 1,4,5,1; 1,8,9,1 and 1,2,3,1 are indicated. Such loops are capable of causing information recycle resulting in inefficient operation of the network. Figure 8 further shows the architecture of the base case network when all the nodes are interconnected and the distances between the nodes are as indicated on the edges

The objective of this paper, was to carry out performance analysis of the base case telecommunication network and there from synthesize topology for the network which shall be the least prone to failure.

2.2 *Performing the Analysis*

It is always the desire of System designers /analysts to thrive to select the most effective architecture or design constraints which drive a system to maximum performance. Network performance is concerned with measuring or estimating the parameters that represent

the network behavior which in most cases is the benchmark of what a system is capable of doing. The parameters of general interest are the network delays, throughput, link utilization etc. These parameters can consequently be measured in terms of transmission efficiency / reliability and operational cost of the network topology (Leino *et al*. 2006; Ekpenyong *et al*. 2009).

2.3 *Estimating Network Parameters*

2.3.1 *Network Delay*

This is estimated as the delay prior to start of transmission of information as well as the delay incurred in processing same through the link and is given by the expression (Ahuja 1982) in equation (1):

$$T = \sum_{i=1}^{M} \lambda_i \gamma^{1} [1/(uL_i - \lambda_i)] \quad (2)$$

where:

T - network delay, s

λ_i - information traffic in i[th] link, bits/s

γ- total information traffic in the network (throughput), bits/s

L_i-Link capacity, bits/s

$1/u$-average size of information, bits

M- total number of links

2.3.2 *Network Throughput*

Throughput may be defined as the magnitude of information in bits per second received correctly and it is expressed as shown in equation (2):

$$\gamma = \sum_{j=1}^{N} \sum_{k=1}^{N} \gamma_{jk} \quad (3)$$

where:

γ - network throughput, bits/s

γ_{jk} - average traffic from node j to node k, bits/s

N - total number of nodes from j to k

2.3.3 *Link Utilization*

The Link Utilisation p, can be estimated by using the expression in equation (3):

$$p = \lambda_i / uL_i \quad (4)$$

where all the notations remain as earlier defined.

2.3.4 *Network Cost*

Assuming the distance matrix D = d_{jk}, of dimension NxN for all the nodes of the assigned region and the cost $C_{(x)}$, of communicating between the nodes j and k is proportional to the distance between the nodes, d_{jk}. Then introducing Boolean variables we have as follows:

$$x_j = \begin{cases} 1, \text{ if the link } d_{jk}, \text{is within the network topology} \\ 0, \text{ if otherwise} \end{cases} \quad (5)$$

$$C_{(x)} = \sum_{j=1}^{N} \sum_{k=1}^{N} d_{jk} x_{jk} \quad (6)$$

Under the constraints:

$$\sum_{j=1}^{N} x_{jk} = 1, \quad \sum_{k=1}^{N} x_{jk} = 1, \quad x_{jk} = 0 \text{ or} 1 \quad (7)$$

$$j, k = 1,2,3 \ldots N$$

3. Synthesis of Network Topology

From the analysis above it may be observed that synthesis of the network to evolve the topology that will satisfy the needed objective shall invariably translate into estimating the variables, $x_{jk} \in \{0,1\}$ using a chosen criterion that will minimize the goal function in equation (5) (Koryachko et al., 1970).

Assuming the associated probabilities for routes with successful transmission are $P_1, P_2, \ldots P_k$, then probability of successful transmission through the network is the product $\prod P_i = P_1 . P_2 . \ldots . P_k$. By using the shortest route algorithm this problem can be solved by converting elements of $log \prod P_i$ into $log\ P_1 + log\ P_2 + \ldots + log\ P_k = \sum log\ P_k$. (Taha, 2008).

For the most reliable topology, the objective is therefore to select the routes which maximize the probabilities of successful transmission through the network. However since log $P_k \leq 0$, then the maximization of $log\ P_k$ is replaced by the minimization of $- log\ P_k$, which values are positive.

Accordingly, Table 1 gives the probability matrix for the network shown in Figure 8. These were obtained from estimates of the operation statistics for the telecommunication
network in Figure 6. The figures in brackets in Table 1 are the – *log* P_k values for the links.

4. Analysis of Results.

The shortest route algorithm due to Dijkstra was employed in solving the problem using the data in Table 1. The solution output from TORA software (Taha 2005) for transmission from node 1 to node 10 is as shown in Figure 9. From the result, it was found that the most reliable network topology is defined by the links 1-2, 1-3, 1-5-6, 1-7, 1-4-8, 1-9 and 1-10 with the maximum probability of 84.1 percent.

5. Conclusion

The properties of graphs useful for the analysis and synthesis of practical networks have been enunciated. Examples of algorithms based on graph structures capable of providing cost effective solutions to the synthesis of networks as well as permitting incorporation of efficiency factors to drive network performance, have been given . The study of an existing Telecommunication network where the probabilities for successful transmission through the links varied from 0.93 to 1, has shown that the most reliable topology that can be synthesized for the network using Dijkstra's shortest route algorithm gives a maximum performance probability of 0.841 for a destination from nodes 1 to 10.

References

Ahuja, V. (1982). *Design and Analysis of Computer Communication Networks*. Mc Graw - Hill Coy, New York
Aspen – Products – Integration (2000). http:/www.aspenntech.com/index.asp? menuchoice. (December 4.)
Bassey, K. N. (2001). An overview of **IT** Infrastructure in SPDC: Identification of Bottlenecks and Possible Solutions. Technical Report -SSIN-TSSW, SPDC Warri.Nigeria.
Bassey, E.N. (2000). Of Process Network Topology and Information Flow: The Theoretical Framework. Technical Presentation. September Meeting of the Nigerian Society of Chemical Engineers, Warri.
Bassey, E. N., Bassey, O. B., & Bassey, U. N. (2005). The Design and Analysis of Process Networks using Spanning Tree Technology. *Global Journal of Mechanical Engineering,* 6, (1), 59-68
Beasley, J. E. (2004). "OR-Notes". *mscmga.ms.ic.ac.uk/jeb/or/graph.html.*
Kafarov, V. V. et al. (1974). *Principles of Mathematical Modelling of Chemical Process Systems.* Khimia, Moscow
Koryachko, V. et al. (1970) *CAD Basics*. Mir Publisher, Moscow
Leino, J., Penttinen, A., & Virtamo, J. (2006). "Flow Level Performance Analysis of Wireless DataNetworks: A Case Study", Proceedings of the 2006 IEEE International Conference on Communications (Reprint 6 pages)
Linnhoff, B. et al. (1979). "Understanding Heat Exchanger Network". Computers and Chemical Engineering, 3, 295-302
Lucey, T. (2002). Quantitative Technique. (6th Ed.). Thompson Learning High Holborn House, 50-51 Bedford Row, London WC1R 4LR, 374-413
Nozari, A. and Enscore (Jr), E. E. (1981). "Computerised Facility Layout with Graph Theory", Computer & Industrial Engineering. 5, (3), 183-193
Scheid, F. (1983). Theory and Problems of Computers and Programming. Schaum's Outline Series, McGraw – Hill Book Company, Singapore
Steward D. V. (1987). Software Engineering with Systems Analysis and Design. Brooks Cole Pub. Coy. California.
Taha H.A. (2005). Operations Research: An Introduction. (7th ed.). Prentice-Hall of India Private Limited, New Delhi.
Taha H.A. (2008). Operations Research: An Introduction. (8th Ed.). PHI Learning Private Limited. New Delhi.
Wayne, K. D. (1999). "Generalized Maximum Flow Algorithms". PhD Dissertation, Cornell University, USA.
Winston, W. L. (1994). *Operations Research: Application and Algorithnms*. Wordsworth Publishing Coy, Belmont, California, USA.

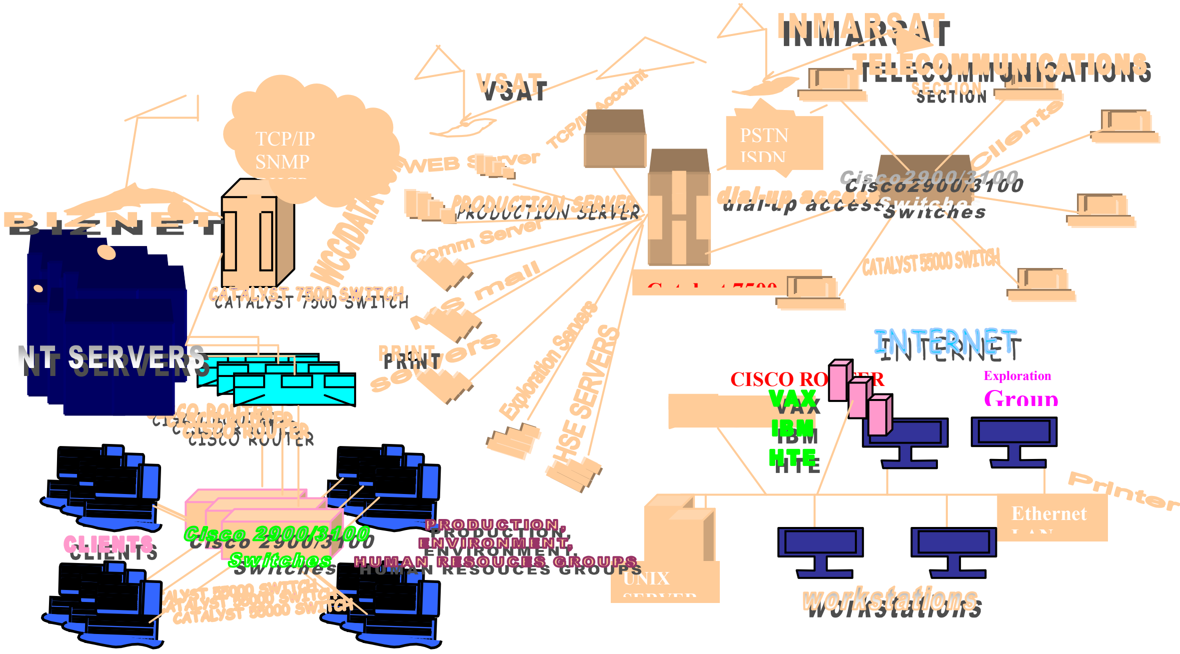


FIGURE 1 : A TYPICAL IT INTEGRATION IN AN OIL AND GAS BUSINESS

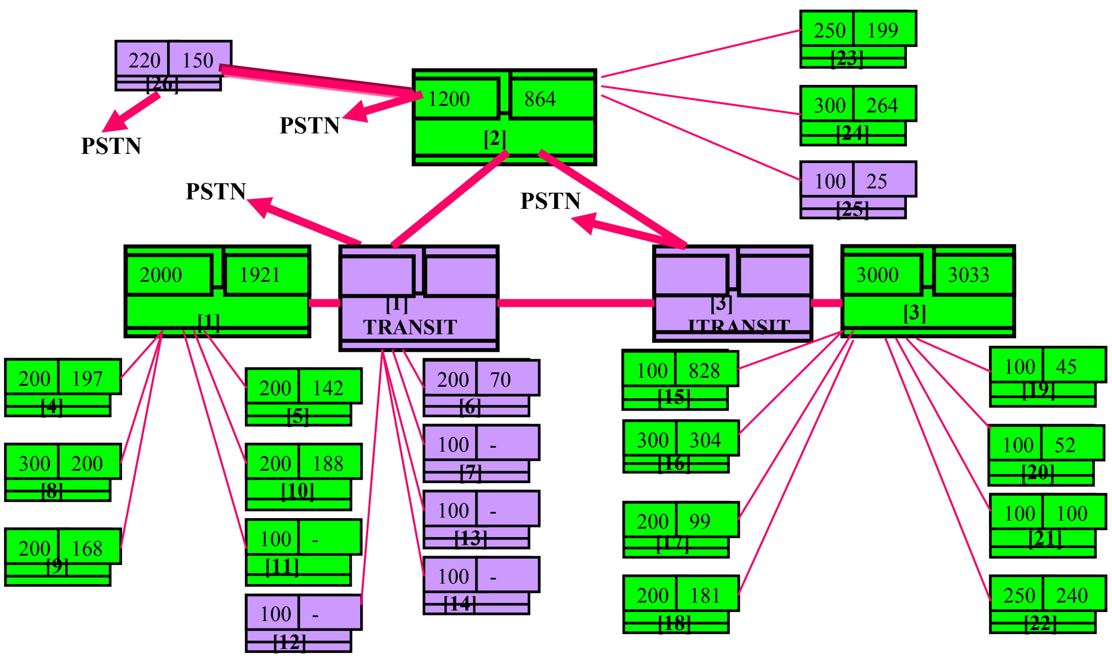


Figure 6. Typical Practical Telecommunication Network

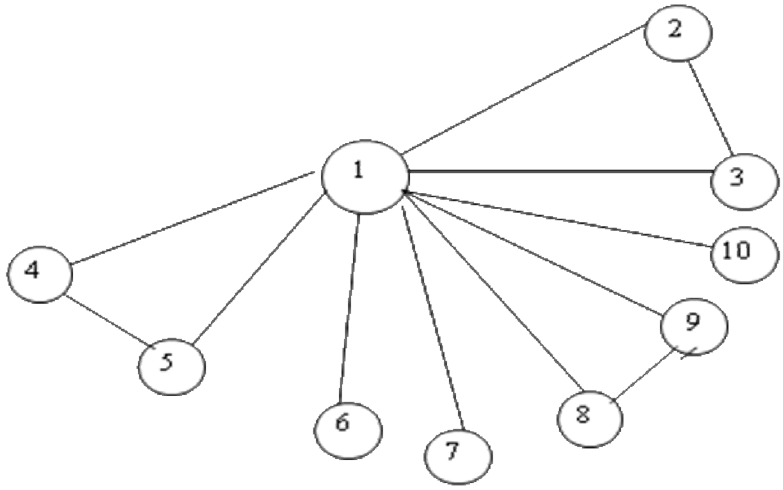


Figure 7. Structural Configuration of the Base Case Network.

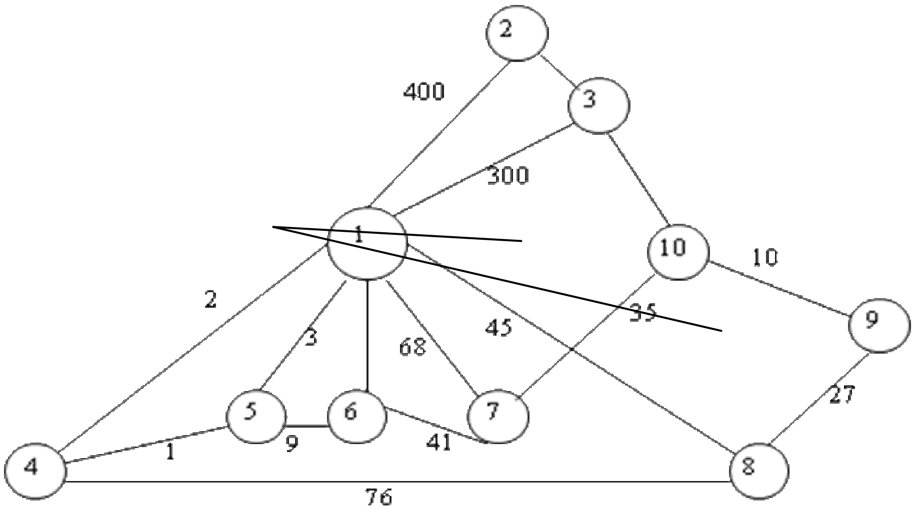


Figure 8. Network Architecture for Base Case with all nodes interconnected

Table1. Probability Matrix for the Telecommunication Network.

From / To	2	3	4	5	6	7	8	9	10
1	**1 [0]**	**1 [0]**	**0.99 [0.0044]**	**0.99 [0.0044]**	**0.90 [0.0458]**	**0.99 [0.0044]**	**0.93 [0.0315]**	**0.97 [0.0132]**	**0.99 [0.0044]**
2		**0.99 [0.0044]**							
3									**0.99 [0.0044]**
4				**0.95 [0.0223]**			**0.95 [0.0223]**		
5					**0.95 [0.0223]**				
6						**0.95 [0.0223]**			
7									**0.99 [0.0044]**
8								**0.95 [0.0223]**	
9									**0.95 [0.0223]**

Figure 9. 'Tora' Solution of the most reliable Network Topology