COMMUNICATION, TRANSMISSION, AND TRANSPORTATION NETWORKS

HOWARD FRANK · IVAN T. FRISCH

Network Analysis Corporation, formerly, University of California, Berkeley

ADDISON-WESLEY PUBLISHING COMPANY

Reading, Massachusetts · Menlo Park, California · London · Don Mills, Ontario

This book is in the
ADDISON-WESLEY SERIES IN ELECTRICAL ENGINEERING

Consulting Editors
David K. Cheng, Leonard A. Gould, Fred K. Manasse

 Library of Congress Catalog Card No. 78-119666.

PREFACE

Modern society is dominated by a complex of networks for the transmission of information, the transportation of people, and the distribution of goods and energy. This complex includes such diverse systems as the telephone network, gas and oil pipelines, airline networks, and networks of computers serving as data banks and remote processing units. The enormous cost of these networks demands that existing ones be rationally used and new ones intelligently planned.

The purpose of this book is to develop a unified treatment of the fundamental theory of networks. We present the underlying problems and properties common to many classes, and techniques for their solution. We have intended the treatment to be rigorous and include many problems concerning new mathematical developments relevant to the study of physical networks. In all cases, we present algorithms which are computationally efficient. These algorithms usually require the use of a computer for implementation. In a few cases, where a rigorous theory is not available, we give heuristic methods whose merits have been substantiated by successful, documented computer programs.

The material is organized with respect to classes of physical problems rather than the techniques used to solve them. For example, a unified treatment of network vulnerability is presented even though significantly different mathematical techniques are required to treat various aspects of the problem. In spite of this problem orientation, many new theoretical results appear throughout the book. Problems or techniques which do not seem to be presently applicable to the study of physical systems are omitted or mentioned only in problems at the ends of chapters.

All basic concepts are defined and many exercises are included to illustrate them. Hence no prior background in networks is required. However, in almost all cases, the material is developed to the level of present research. Indeed, we believe that much of the material in the book is new or novel. As far as we know, this is the first book to treat the theory of probabilistic graphs in depth. The material we present on the graph theoretic formulation of vulnerability has never appeared in any book; nor has the general solution of the multiterminal synthesis problem. Numerous other algorithms are presented in book form for the first time. Many of these are based on our own research and that of our graduate students.

The book contains parallel treatments of deterministic and probabilistic networks. A knowledge of the fundamental concepts of probability theory is required for the latter. The entire book can be covered in a one-year sequence of courses for students with no previous background in graph theory or networks. Several possible one-semester or quarter courses can also be taught from it. A course treating only deterministic problems could be based on Chapters 1, 2, 3, 5, Sections 1 through 7, 9, 10, and 13 of Chapter 6, and Chapter 7. A course emphasizing probabilistic concepts could be built on Chapters 1, 2, Sections 1, 2, 3, and 6 of Chapter 3, Chapter 4, Sections 1, 2, 3, 5 through 8, and 10 through 12 of Chapter 6, and Chapters 8 and 9. A course combining both deterministic and probabilistic treatments would consist of Chapters 1, 2, Sections 1 through 8 of Chapter 3, Sections 1 through 6 of Chapter 4, Sections 1 through 6 of Chapter 5, and either Sections 1 through 10 of Chapter 6 or Chapter 7 and either Sections 2, 3, and 7 of Chapter 8 or Chapter 9. We have taught courses using a number of these options.

Each of the above sequences requires approximately forty hours of lectures. For students who have already had a course in network flows the following option might be desirable: Chapter 4, Sections 1 through 6 of Chapter 5, Sections 5 through 13 of Chapter 6, Chapter 7, and Chapter 8. A shorter course on vulnerability can be taught using Chapters 1, 2, Sections 1, 2, 3, and 6 of Chapter 3, and Chapters 7, 8, and 9. For students with a background in network flow theory, Chapters 7 and 8 can form the basis of a twenty-hour course on vulnerability. We have taught several such variations at Berkeley.

ACKNOWLEDGMENTS

We are indebted to many people for their generous help. We both were initiated into the study of flow networks as doctoral students—Howard Frank, while working with S. L. Hakimi at Northwestern University, and Ivan T. Frisch, while working with W. H. Kim at Columbia University. At Berkeley these studies were supported in part by an Army Research Office Contract on "Information Processing" and in part by grants from the Joint Services Electronics Program and the National Science Foundation. In this regard we are grateful to J. E. Norman of the Army Research Office, Durham, North Carolina, for his continued interest and encouragement. Some material was developed while one of us (I. T. F.) was on a leave of absence at the Bell Telephone Laboratories, Whippany, New Jersey, as a Ford Foundation Resident in Engineering Practice. The original motivation for some of the work discussed in Chapter 7 is due to D. Gillette, Executive Director, Transmission Systems, at the Bell Laboratories in Holmdel, and to A. R. Eckler and F. D. Benedict, who headed groups at the Bell Laboratories in Whippany, working on the vulnerability problem. Some of the results presented in the book were developed while one of us (H. F.) was on a leave of absence at the Office of Emergency Preparedness (O.E.P.), Executive Office of the President, Washington, D.C. The enthusiastic support of R. Truppner, Director of the O.E.P.'s National Resource Analysis Center (N.R.A.C.), and R. H. Kupperman, Chief of the N.R.A.C.'s Systems Evaluation Division, enabled us to apply many of our results to problems of national importance. Much of the work at the O.E.P. was strongly influenced by the far-sighted recommendations of David Rosenbaum, presently with the MITRE Corporation, Maclean, Virginia.

We are grateful to L. A. Zadeh and E. S. Kuh, who encouraged us to develop our network courses at Berkeley, and to L. Farbar, Director of Engineering Extension at the University of California at Berkeley, for his aid in presenting a two-week extension course. Helen Barry was most helpful in administering this course. In addition, portions of the book were presented in the UCLA extension courses on Large Scale Systems, organized by C. T. Leondes, and Queueing Systems, organized by L. Kleinrock, and in short courses sponsored by the Network Analysis Corporation.

We have incorporated many suggestions from our former graduate students: J. Ayoub, S. Chaubey, W. S. Chou, M. El-Bardai, M. Malek-Zaverei, S. Sankaran, D. K. Sen, N. P. Shein, and especially P. Jabedar-Maralani. B. Rothfarb made numerous valuable suggestions about several versions of the manuscript and was one of the reviewers of the final manuscript. A number of our other colleagues have read portions of the manuscript and offered criticisms. Among them are F. T. Boesch, of the Bell Telephone Laboratories, S. L. Hakimi, of Northwestern University, T. C. Hu, of the University of Wisconsin, W. S. Jewell, of the University of California, Berkeley, L. Kleinrock, of the University of California, Los Angeles, D. J. Kleitman, of the Massachusetts Institute of Technology, E. Lawler, of the University of California, Berkeley, W. Mayeda, of the University of Illinois, and K. Steiglitz, of Princeton University.

Jane Frank and Vivian Frisch typed many pages of original notes; helped with all sorts of clerical details; strove mightily to suppress laughter while we argued bitterly over whether ρ should represent connectivity or number of components; and acted as witnesses while the order of our names on the title page was decided by the flip of a quarter.

Glen Cove, New York H. F.
September 1970 I. T. F.

NOTES TO THE READER

1. Sections marked with an asterisk contain advanced topics not essential to the continuity of the main development and may be omitted on first reading.
2. Unless otherwise indicated, a reference to a section, figure, formula, etc., is to an item in the same chapter as the reference.
3. References in the text are indicated by the first two letters of the author's surname followed by an indexing number. The references are listed alphabetically at the end of each chapter.
4. Many of the problems at the ends of chapters are based on results from published papers. In such cases, a reference to the appropriate paper is given in the problem.

CONTENTS

CHAPTER 1

GRAPHS AND PHYSICAL MODELS

1.1 GRAPHS AS STRUCTURAL MODELS

Many systems involve the communication, transmission, transportation, flow, or movement of commodities. The commodity under consideration may be a tangible item, such as railway cars, automobiles, oil drums or water, or an intangible item, such as information, disease, "friendship," or heredity. Thus, a highway system, a telephone network, an interconnection of warehouses and retail outlets, a power grid, or an airline network all involve the flow of commodities through a network. Often these networks can be modeled by a mathematical entity called a *graph*.

A graph may be considered to be a collection of points called vertices connected by lines called branches. The modeling of some physical systems by graphs is quite natural, while for others the relationship between the graph theoretic model and the original system is extremely subtle. In the former group are communication or transportation systems. The branches of the graph can represent roads, telegraph wires, railroad tracks, airline routes, water pipes—in general, channels through which the commodities are transmitted. The vertices of the graph can represent communities, highway intersections, telegraph stations, railroad yards, airline terminals, water reservoirs and outlets—in general, points where flow originates, is relayed, or terminates.

Two physical networks may be structurally similar, but have significantly different characteristics. For example, the interconnections in both an electrical network and a telephone system can be specified by a graph. However, the branches of the electrical network model are characterized by parameters such as resistance, inductance, or capacitance whereas the branches of the telephone network model are characterized by parameters such as the number of wires per trunk, the maximum transmission rate, and the cost per unit length. We must account for these parameters as part of the model, to obtain meaningful results.

With each branch and vertex of a graph, we can associate a number of parameters that represent the natural limitations and capabilities of the branches and vertices. For example, a power system might be modeled by a graph in which the branches represent power lines and the vertices represent power generation stations, substations, and customers. The important parameters of the system are incorporated into the model as numbers, or *weights*, on the branches and vertices of the graph. These weights may be either fixed or random. Thus, for the power system, a typical vertex

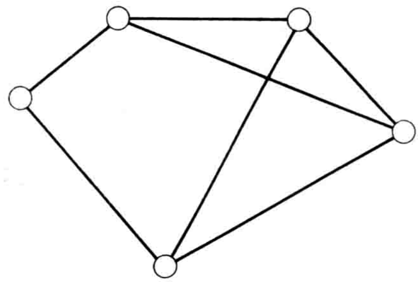

Fig. 1.1.1 Pictorial representation of a graph.

representing a power generator might have the following weights: maximum power output, number of generators at the station, reliability of each generator, and cost per kilowatt-hour. A typical branch might have three weights corresponding to the maximum power-handling capacity, reliability, and cost.

The purpose of the branch and vertex weights is to include nonstructural information into the graph theoretic model of a system. The following examples further illustrate this point.

Example 1.1.1 A traffic network. Let each vertex of a graph represent a city. Two vertices are connected by a branch if there is a highway between the corresponding cities. A number is associated with each branch to indicate the length of the corresponding highway. A second weight represents the maximum number of cars that can be accommodated per unit length per unit time, and a third branch weight could be the speed limit.

Example 1.1.2 An airline system. Let each vertex of a graph represent an airline terminal. Two vertices are connected by a branch if there is a direct air link between the terminals. Each vertex of the graph has a weight indicating the number of airplanes that the terminal can handle in a given interval of time. This vertex weight could be a fixed number if the traffic-handling capability of the terminal is assumed to be constant, it could be time-varying if the traffic-handling capability fluctuates, or it could be a random variable if it depends on unpredictable elements such as weather.

Example 1.1.3 A telegraph system. Let each vertex of a graph represent a telegraph station. Two vertices are connected if there is a telegraph wire between the corresponding telegraph stations. That is, there will be a branch between two vertices if the stations can communicate directly without any intermediate relay station. The number of telegraph operators at each station is limited, so there is a maximum number of messages which can be simultaneously transmitted and received. This can be included in the model by an appropriate vertex weight. The maximum number of messages on a branch is determined by the number of telegraph wires. We weight each branch by the maximum number of simultaneous messages that can be handled. Another possible consideration is the time delay required to send a message through the network. At each station there will be a "waiting" time for an available operator or channel, as well as the time required to transmit the message. The total time delay at a given station will usually be a random variable and can be represented by an additional vertex weight.

Example 1.1.4 An economic model. Suppose we are given a system of factories, warehouses, and outlets connected by a set of highways, railroads, and waterways. This system can be structurally modeled by a graph with the branches representing transportation channels and the vertices representing factories, warehouses, and outlets. In the graph, the factories are source vertices, the outlets are terminal vertices, and the warehouses are relay vertices. We could further distinguish between the vertices. For example, some source vertices might produce one type of commodity while others produce a different type. Among the many possible vertex weights are the rate of production of the ith commodity, the production cost per unit of the ith commodity, and the time required to produce a unit of the ith commodity. For the relay vertices, a single weight representing the storage space might suffice. A terminal vertex could be weighted with numbers which indicate the types of commodity which are sold at that vertex, the price of each commodity, the amount of local storage available, and the demand (usually a random variable) for each commodity. Typical branch weights could be maximum volume per unit time, cost of transportation per unit of ith commodity, and transmission time per unit of ith commodity.

1.2 TYPICAL PROBLEMS

The utility of graphs as models depends on the nature of the physical problem to be solved. The type of problem for which a graph is most obviously useful is that of connectivity. Given a system and its graph, we might be interested in determining whether or not a particular commodity can be transported from one given location to another. We are interested in finding a "path" between two given vertices over which the commodity can be sent. A more general, but similar, problem is to establish whether or not a given commodity can be sent from *any* point to *any other* point. Here, we must determine whether there is at least one path from any vertex to any other vertex.

The connectivity problems mentioned above are structural problems. The existence of a path between a pair of vertices implies that some amount of flow can be transmitted between these two vertices. There is no information about the quantity of flow which can be sent. To include this information, we must consider weighted graphs. Suppose we weight each branch and vertex by a number which indicates the maximum amount of flow that it can accommodate. These weights represent the capacities of the channel, sources, terminals, and relay points in the original system. We might then ask: What is the maximum amount of flow which can be sent between a given pair of vertices? In a power system, this number might correspond to the maximum power that a particular generating station could supply to a particular user; in a telegraph system, it might represent the maximum attainable rate of information transmission between two telegraph stations.

The problem of finding the maximum amount of a given quantity which can be transmitted between two points is known as the Maximum Flow Problem. A generalization of this problem is to find the maximum amount of several commodities which can be *simultaneously* sent between several pairs of points. This is known as the Multicommodity Maximum Flow Problem. Both of these are analysis problems. Given a system and its graph theoretic model, we can attempt to analyze the graph and find the maximum flow rates. We can also formulate an analogous synthesis problem. Suppose we are given a set of stations and maximum flow-rate requirements. We would like to design a system which satisfies these requirements. Furthermore, since there may be many such systems, we would like to select a system which is in some sense "optimal." One possible optimality criterion could be minimum cost. This synthesis problem has many variations. For example, we may assume that the behavior of the network to be designed can be precisely predicted. In this case, flow-rate requirements can be exactly satisfied. In other cases, there may be random elements in the design or behavior of the system and so the meaning of the phrase "satisfy flow rate requirement" must be interpreted probabilistically.

The preceding connectivity and maximum-flow problems are closely related to a group of problems which may be termed problems of "vulnerability" and "reliability." Here, we are given a system operating in a "hostile" environment. This hostility may be the result of natural disturbances, equipment failure, or enemy attack. The effect of the hostility is to disrupt communications. Given an existing system, we must analyze it to determine the system degradation which could occur. Given a set of performance criteria, we must design a system which minimizes the possible system degradation. Again, both the analysis and synthesis problems can be posed in either deterministic or probabilistic terms.

It is usually possible to route a given commodity over many different paths, with one routing possibly better than another. If a poor route is selected, it may block an additional flow which might otherwise have been established. Among the problems we must therefore consider are finding the shortest, the least expensive, or the most reliable route for a given commodity.

In many physical systems network traffic is a function of time. All available routes between a pair of stations may be occupied and so it is impossible to send additional flow between these stations. Thus, a subscriber at one of these stations must wait until a channel is available. The expected value of the time he must wait is an important parameter of the system. Typical analysis problems are to find the average waiting time and to investigate the effect of network structure and various routing procedures on this waiting time.

The examples given above should suffice to point out the wide range and applicability of the model and the nature of the problems that can be posed. To solve these problems, we require some concepts and results from the theory of graphs. These are the subject of the next chapter.

1.3 RELATED READING

The material relevant to the study of networks is scattered throughout the journals of a number of different disciplines, because, historically, problems modeled by graphs have arisen in varied and seemingly unrelated situations. Furthermore, solutions to these problems have been achieved by the use of several different branches of mathematics.

As we have seen, one motivation for the study of networks is the investigation of traffic in communication systems. A. K. Erlang, whose work is summarized by Brockmeyer *et al.* [BR1], was one of the earliest and most significant innovators in this area. Since his work, a tremendous body of literature has developed. References to this literature can be found in the recent books by Benes [BE2] and Kleinrock [KL1]. A good summary of the physical bases of these problems is given by Rubin and Haller [RU1]. The results of extensive computer simulation are given by Weber [WE1].

The second major impetus for the development of the theory of networks was the formulation of mathematical models for economic and distribution systems, both of which are included under the generic title of operations research. Probably the earliest link between problems in this area and communication networks was Hitchcock's solution to the "transportation problem" [HI1]. A number of the major results and references in this area are given by Ford and Fulkerson [FO1], who have been among the most original and prolific contributors to the development of the theory of flows in networks.

Another stimulus to the study of networks has been in the study of steady-state flow of information through a communication system. This approach seems to have been first adopted by Elias, Feinstein and Shannon [EL1], although Mayeda [MA1] was the first to formulate and solve significantly different problems arising primarily from the new viewpoint. This new stimulus came at a time when many electrical engineers were prepared to apply their knowledge of graph theory acquired through the study of electrical networks. Hence, a new body of literature was developed by this group. Some of the early work in this area is given by Kim and Chien [KI1].

The primary mathematical disciplines relevant to the study of networks in this book are graph theory, combinatorics, probability theory, and statistics. Secondary use is made of mathematical programming and queuing theory. All the necessary results from the theory of graphs are derived in this book and hence no outside references are required. However, for those readers wishing to delve more deeply into this subject, a number of excellent books and bibliographies are now available [BE3, BE4, BU1, HA3, HA4, KI1, KO1, OR1, DE1, SE1, TU1, TU2, ZY1].

The book is also self-contained with regard to theorems from combinatorics. Indeed, much of the material on flows in networks actually comprises a significant new branch of combinatorics. We do not emphasize this viewpoint. However, as an example, Ford and Fulkerson [FO1] use the theory of flows to solve many purely combinatorial problems such as the assignment of entries to matrices of zeros and ones to satisfy various constraints. The classical results of combinatorics can be found in several references [BE1, RI1, RY1].

Probability theory and statistics are the only essential disciplines in which we must assume some background on the part of the reader. However, except for a few sections, only elementary knowledge is required. Furthermore, the book is organized so that the material on random networks appears in separate chapters or sections parallel to corresponding sections on deterministic networks. Hence, it is possible to read those sections which do not require probabilistic theory independently of the others. The required material on probability and statistics can be found in references [CR1, FI1, LE1, MO1]. Reference [CR1] is a classic work in probability theory and statistics, which contains all of the necessary foundations in probability theory. Reference [KO1] is an elementary text on statistics while references [FI1] and [LE1] are intermediate and advanced level texts, respectively. Fisz's book [FI1] may be considered as the primary reference for both probability theory and statistics. Queuing theory is required in only one chapter of the book and only the most elementary concepts are needed. For a more extensive treatment and a guide to the literature in queueing theory as applied to networks one can refer to the work of Kleinrock [KL1].

Mathematical programming is useful in the study of communication networks, first because it is an efficient computational tool and second because it enables one to place many algorithms and results in a general framework. We regard mathematical programming as a secondary discipline for this book, since we can derive most of our results without resorting to it. Furthermore, where flow techniques are applicable, they are usually more efficient than general programming methods. The few theorems we do need are given in the next chapter without their proofs. The proofs of the theorems and the general theory of mathematical programming are expounded in a number of excellent books among which are those of Dantzig [DA1], Hadley [HA1, HA2], Berge and Ghouila-Houri [BE4], and Simonard [SI1]. The relationship between programming techniques and flow problems is developed in the books by Kaufmann [KAI] and Hu [HU1]. Further generalizations of graph theory and programming techniques using matroids appear in a book being written by E. Lawler.

REFERENCES

BE1 E. F. Beckenbach (ed.), *Applied Combinatorial Mathematics*, Wiley, New York, 1964.

BE2 V. R. Benes, *Mathematical Theory of Connecting Networks and Telephone Traffic*, Academic Press, New York, 1965.

BE3 C. Berge, *The Theory of Graphs and Its Applications*, Wiley, New York, 1962.

BE4 C. Berge and A. Ghouila-Houri, *Programming, Games and Transportation Networks*, Wiley, New York, 1962.

BR1 E. Brockmeyer, H. L. Helstrom, and A. Jensen, *The Life and Works of A. K. Erlang*, Danish Academy of Technical Sciences, No. 2, Copenhagen, 1948.

BU1 R. G. Busacker and T. L. Saaty, *Finite Graphs and Networks: An Introduction with Applications*, McGraw-Hill, New York, 1965.

CR1 H. Cramer, *Mathematical Models of Statistics*, Princeton University Press, Princeton, 1963.

DA1 G. B. Dantzig, *Introduction to Linear Programming*, Princeton University Press, Princeton, 1963.

DE1 N. Deo, *An Extensive English Language Bibliography on Graph Theory and Its Applications*, National Aeronautics and Space Administration, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, Cal., 1969.

EL1 P. Elias, A. Feinstein, and C. E. Shannon, "A Note on the Maximum Flow through a Network," *IRE Trans. Inform. Theory* **IT-2,** 117–119 (1956).

FI1 M. Fisz, *Probability Theory and Mathematical Statistics*, Wiley, New York, 1963.

FO1 L. R. Ford, Jr., and D. R. Fulkerson, *Flows in Networks*, Princeton University Press, Princeton, 1962.

HA1 G. Hadley, *Linear Programming*, Addison-Wesley, Reading, Mass., 1962.

HA2 G. Hadley, *Nonlinear and Dynamic Programming*, Addison-Wesley, Reading, Mass., 1964.

HA3 F. Harary, *Graph Theory*, Addison-Wesley, Reading, Mass., 1969.

HA4 F. Harary, R. Z. Norman, and D. Cartwright, *Structural Models: An Introduction to the Theory of Directed Graphs*, Wiley, New York, 1965.

HI1 F. L. Hitchcock, "The Distribution of a Product from Several Sources to Numerous Localities," *J. Math. Phys.* **20,** 224–230 (1941).

HU1 T. C. Hu, *Integer Programming and Network Flows*, Addison-Wesley, Reading, Mass., 1969.

KA1 A. Kaufmann, *Graphs, Dynamic Programming, and Finite Games*, Academic Press, New York, 1967.

KI1 W. H. Kim and R. T. Chien, *Topological Analysis and Synthesis of Communication Networks*, Columbia University Press, New York, 1962.

KL1 L. Kleinrock, *Communication Nets: Stochastic Message Flow and Delay*, McGraw-Hill, New York, 1964.

KO1 D. König, *Theorie der Endlichen und Unendlichen Graphen*, Chelsea, New York, 1950.

LE1 E. L. Lehman, *Testing Statistical Hypotheses*, Wiley, New York, 1959.

MA1 W. Mayeda, "Terminal and Branch Capacity Matrices of a Communication Net," *IRE Trans. Circuit Theory* **CT-7.** 261–269 (1960).

MO1 A. M. Mood and F. A. Graybill, *Introduction to the Theory of Statistics*, McGraw-Hill, New York, 1963.

OR1 O. Ore, *Theory of Graphs*, American Mathematical Society, Vol. XXXVIII, Colloquium Publications, Providence, R.I., 1962.

RI1 J. Riordan, *An Introduction to Combinatorial Analysis*, Wiley, New York, 1958.

RU1 M. Rubin and C. E. Haller, *Communication Switching Systems*, Reinhold, New York, 1966.

RY1 H. J. Ryser, *Combinatorial Mathematics*, Carus Mathematical Monograph 14, Mathematical Association of America, Wiley, New York, 1963.

SE1 S. Seshu and M. B. Reed, *Linear Graphs and Electrical Networks*, Addison-Wesley, Reading, Mass., 1961.

SI1 M. Simonard, *Linear Programming*, Prentice-Hall, Englewood Cliffs, N.J., 1966.

TU1 J. Turner, *Key Word Indexed Bibliography on Graph Theory*, Stanford Research Institute Report, SRI Project 145591-W.O.A14, February 1964.

TU2 W. Tutte, *Connectivity in Graphs*, University of Toronto Press, Toronto, 1968.

WE1 J. H. Weber, "A Simulation Study of Routing and Control in Communication Networks," *Bell System Tech. J.* **43**. 2639–2676 (1964).

ZY1 A. A. Zykov, "Bibliography on Graph Theory," in *Theory of Graphs and Its Applications: Proceedings of the Symposium Held in Smolenice in June 1963*, Academic Press, New York, 1964.