Bang Ye Wu and Kun-Mao Chao

Spanning Trees and Optimization Problems

CRC PRESS
Boca Raton London New York Washington, D.C.

Preface

The research on spanning trees has been one of the most important areas in algorithm design. People who are interested in algorithms will find this book informative and inspiring. The new results are still accumulating, and we try to make clear the whole picture of the current status and future developments.

This book is written for graduate or advanced undergraduate students in computer science, electrical engineering, industrial engineering, and mathematics. It is also a good reference for professionals.

Our motivations for writing this book:

1. To the best of our knowledge, there is no book totally dedicated to the topics of spanning trees.

2. Our recent progress in spanning trees reveals a new line of investigation.

3. Designing approximation algorithms for spanning tree problems has become an exciting and important field in theoretical computer science.

4. Besides numerous network design applications, spanning trees have also been playing important roles in newly established research areas, such as biological sequence alignments, and evolutionary tree construction.

This book is a general and rigorous text on algorithms for spanning trees. It covers the full spectrum of spanning tree algorithms from classical computer science to modern applications. The selected topics in this book make it an excellent handbook on algorithms for spanning trees. At the end of every chapter, we report related work and recent progress.

We first explain general properties of spanning trees. We then focus on three categories of spanning trees, namely, minimum spanning trees, shortest-paths trees, and optimum routing cost spanning trees. We also show how to balance the tree costs. Besides the theoretical description of the methods, many examples are used to illustrate the ideas behind them. Moreover, we demonstrate some applications of these spanning trees. We explore in details some other interesting spanning trees, including maximum leaf spanning trees and minimum diameter spanning trees. In addition, Steiner trees and evolutionary trees are also discussed. We close this book by summarizing other important problems related to spanning trees.

Writing a book is not as easy as we thought at the very beginning of this project. We have tried our best to make it consistent and correct. However, it's a mission impossible for imperfect authors to produce a perfect book.

Should you find any mathematical, historical, or typographical errors, please let us know.

We are extremely grateful to Richard Chia-Tung Lee, Webb Miller, and Chuan Yi Tang, who always make the subject of algorithms exciting and beautiful in their superb lectures. Their guidance and suggestions throughout this study were indispensable.

We thank Vineet Bafna, Giuseppe Lancia, and R. Ravi for their collaborations on the minimum routing cost spanning tree problem. We thank Tao Jiang and Howard Karloff for suggesting the merger of two different works at the early stage of our investigation. We are also thankful to Piotr Berman, Xiaoqiu Huang, Yuh-Dauh Lyuu, Anna Östlin, Pavel Pevzner, and the anonymous reviewers for their valuable comments.

It has been a pleasure working with CRC Press in the development of this book. We are very proud to have this book included in the CRC series on Discrete Mathematics and Its Applications, edited by Kenneth H. Rosen. Ken also provided critical reviews and invaluable information for which we are grateful. We thank Sunil Nair for his final approval of our proposal. Richard O'Hanley was the first to approach us about the possibility of publishing a book at CRC Press. Robert B. Stern then handled the proposal review and contract arrangements efficiently. Bob also proposed many constructive suggestions throughout the project. Jamie B. Sigal helped us with both production and permissions issues, and his gentle reminders kept us moving at a good pace. William R. Palmer III resolved the questions arising in prepress. Nishith Arora revised the LaTeX style files in a timely manner. Julie Spadaro set the production schedule in a perfect way, and kindly copyedited our manuscript for us to review and correct.

Finally, we thank our families for their love, patience, and encouragement. We thank our wives, Mei-Ling Cheng and Pei-Ju Tsai, and our sons, Ming-Hsuan Wu and Leo Liang Chao, for tolerating our absentmindedness during the writing of this book. We promise to work less than 168 hours a week by not taking on a new grand project immediately.

Bang Ye Wu
`bangye@mail.stu.edu.tw`
`http://www.personal.stu.edu.tw/bangye`

Kun-Mao Chao
`kmchao@csie.ntu.edu.tw`
`http://www.csie.ntu.edu.tw/~kmchao`

December 2003

About the Authors

Bang Ye Wu was born in Kaohsiung, Taiwan, in 1964. He earned the B.S. degree in electrical engineering from Chung Cheng Institute of Technology, Taiwan, in 1986, and the M.S. and the Ph.D. degrees in computer science from National Tsing-Hua University, Taiwan, in 1991 and 1999 respectively. He is currently an assistant professor and the head of the Department of Computer Science and Information Engineering, Shu-Te University. Before joining the faculty of Shu-Te University, he worked in the Chung-Shan Institute of Science and Technology, Taiwan, as a research assistant (1986-1989), an assistant research fellow (1991-1995), and an associate research fellow (1999-2000). His current research interests include algorithms and bioinformatics.

Kun-Mao Chao was born in Tou-Liu, Taiwan, in 1963. He earned the B.S. and M.S. degrees in computer engineering from National Chiao-Tung University, Taiwan, in 1985 and 1987, respectively, and the Ph.D. degree in computer science from The Pennsylvania State University, University Park, in 1993. He is currently a professor of the Department of Computer Science and Information Engineering, National Taiwan University. From 1987 to 1989, he served in the ROC Air Force Headquarters as a system engineer. From 1993 to 1994, he worked as a postdoctoral fellow at Penn State's Center for Computational Biology. In 1994, he was a visiting research scientist at the National Center for Biotechnology Information, National Institutes of Health, Bethesda, Maryland. Before joining the faculty of National Taiwan University, he taught in the Department of Computer Science and Information Management, Providence University, from 1994 to 1999, and the Department of Life Science, National Yang-Ming University, from 1999 to 2002. His current research interests include algorithms and bioinformatics. Dr. Chao is a member of Phi Tau Phi and Phi Kappa Phi.

Contents

References

[1] R.K. Ahuja, T.L. Magnanti, and J.B. Orlin. *Network Flows – Theory, Algorithms, and Applications*. Prentice-Hall, 1993.

[2] I. Althöfer, G. Das, D. Dobkin, D. Joseph, and J. Soares. On sparse spanners of weighted graphs. *Discrete Comput. Geom.*, 9:81–100, 1993.

[3] S. Arora. Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems. *J. ACM*, 45(5):753–782, 1998.

[4] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi. *Complexity and Approximation – Combinatorial Optimization Problems and Their Approximability Properties*. Springer-Verlag, 1999.

[5] B. Awerbuch, A. Baratz, and D. Peleg. Cost-sensitive analysis of communication protocols. In *Proceedings of the 9th Symposium on Principles of Distributed Computing*, pages 177–187, 1990.

[6] V. Bafna, E.L. Lawler, and P. Pevzner. Approximation algorithms for multiple sequence alignment. In *Proceedings of the 5th Combinatorial Pattern Matching conference*, LNCS 807, pages 43–53. Springer-Verlag, 1994.

[7] H.J. Bandelt. Recognition of tree metrics. *SIAM J. Discrete Math*, 3(1):1–6, 1990.

[8] Y. Bartal. Probabilistic approximation of metric spaces and its algorithmic applications. In *Proceedings of the 37th Annual IEEE Symposium on Foundations of Computer Science*, pages 184–193, 1996.

[9] Y. Bartal. On approximating arbitrary metrics by tree metrics. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 161–168, 1998.

[10] R. Bellman. On a routing problem. *Quar. Appl. Math.*, 16:87–90, 1958.

[11] P. Berman and V. Ramaiyer. Improved approximations for the Steiner tree problem. *J. Algorithms*, 17(3):381–408, 1994.

[12] M. Bern and P. Plassmann. The Steiner problem with edge lengths 1 and 2. *Inf. Process. Lett.*, 32(4):171–176, 1989.

[13] K. Bharath-Kumar and J.M. Jaffe. Routing to multiple destinations in computer networks. *IEEE Trans. Commun.*, 31(3):343–351, 1983.

[14] O. Borůvka. O jistém problému minimálním (about a certain minimal problem). *Práca Moravské Přirodovědecké Společnosti*, 3:37–58, 1926. (In Czech.).

[15] P. Buneman. A note on metric properties of trees. *J. Comb. Theory B*, 17:48–50, 1974.

[16] L. Cai. NP-completeness of minimum spanner problems. *Discrete Appl. Math.*, 48:187–194, 1994.

[17] A. Cayley. A theorem on trees. *Quart. J. Math.*, 23:376–378, 1889.

[18] B. Chazelle. A minimum spanning tree algorithm with inverse-Ackermann type complexity. *J. ACM*, 47:1028–1047, 2000.

[19] B. Chazelle. The soft heap: an approximate priority queue with optimal error rate. *J. ACM*, 47:1012–1027, 2000.

[20] D. Cheriton and R. E. Tarjan. Finding minimum spanning trees. *SIAM J. Comput.*, 5:724–742, 1976.

[21] B.V. Cherkassky, A.V. Goldberg, and T. Radzik. Shortest paths algorithms: theory and experimental evaluation. In *Proceedings of the Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 516–525, 1994.

[22] L.P. Chew. There are planar graphs almost as good as the complete graph. *J. Comput. Syst. Sci.*, 39(2):205–219, 1989.

[23] H.S. Connamacher and A. Proskurowski. The complexity of minimizing certain cost metrics for k-source spanning trees. *Discrete Appl. Math.*, 131:113–127, 2003.

[24] S.A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the 3rd Annual ACM Symposium on Theory of Computing*, pages 151–158, 1971.

[25] T.H. Cormen, C.E. Leiserson, and R.L. Rivest. *Introduction to Algorithms*. The MIT Press, 1994.

[26] J. Culberson and P. Rudnicki. A fast algorithm for constructing trees from distance matrices. *Inf. Process. Lett.*, 30:215–220, 1989.

[27] E.V. Denardo and B.L. Fox. Shortest-route methods: 1. reaching, pruning, and buckets. *Oper. Res.*, 27:161–186, 1979.

[28] E.W. Dijkstra. A note on two problems in connection with graphs. *Numer. Math.*, 1:269–271, 1959.

[29] R. Dionne and M. Florian. Exact and approximate algorithms for optimal network design. *Networks*, 9(1):37–60, 1979.

[30] B. Dixon, M. Rauch, and R. Tarjan. Verification and sensitivity analysis of minimum spanning trees in linear time. *SIAM J. Comput.*, 21:1184–1192, 1992.

[31] D-.Z Du and F.K. Hwang. A proof of the Gilbert-Pollak conjecture on the Steiner ratio. *Algorithmica*, 7:121–135, 1992.

[32] M. Farach, S. Kannan, and T. Warnow. A robust model for finding optimal evolutionary trees. *Algorithmica*, 13:155–179, 1995.

[33] S. Fekete, S. Khuller, M. Klemmstein, B. Raghavachari, and N.E. Young. A network-flow technique for finding low-weight bounded-degree spanning trees. *J. Algorithms*, 24(2):310–324, 1997.

[34] D. Feng and R. Doolittle. Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.*, 25:351–360, 1987.

[35] M. Fischetti, G. Lancia, and P. Serafini. Exact algorithms for minimum routing cost trees. *Networks*, 39:161–173, 2002.

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

[37] M. Fredman and D. E. Willard. Trans-dichotomous algorithms for minimum spanning trees and shortest paths. *J. Comput. Syst. Sci.*, 48:424–436, 1994.

[38] M.L. Fredman and R.E. Tarjan. Fibonacci heaps and their uses in improved network optimization algorithms. *J. ACM*, 34:596–615, 1987.

[39] M. Fürer and B. Raghavachari. Approximating the minimum-degree Steiner tree to within one of optimal. *J. Algorithms*, 17(3):409–423, 1994.

[40] H. N. Gabow, Z. Galil, T. Spencer, and R. E. Tarjan. Efficient algorithms for finding minimum spanning trees in undirected and directed graphs. *Combinatorica*, 6:109–122, 1986.

[41] G. Galbiati, F. Maffioli, and A. Morzenti. A short note on the approximability of the Maximum Leaves Spanning Tree Problem. *Inf. Process. Lett.*, 52(1):45–49, 1994.

[42] G. Gallo and S. Pallottino. Shortest paths algorithms. *Ann. Oper. Res.*, 13:3–79.

[43] M.R. Garey and D.S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W.H. Freeman and Company, San Francisco, 1979.

[44] N. Garg. A 3-approximation for the minimum tree spanning k vertices. In *Proceedings of the 37th Annual IEEE Symposium on Foundations of Computer Science*, pages 302–309, Burlington, Vermont, 1996.

[45] E.N. Gilbert and H.O. Pollak. Steiner minimal trees. *SIAM J. Appl. Math.*, 16(1):1–29, 1968.

[46] A.V. Goldberg. A simple shortest path algorithm with linear average time. In *Proceedings of the 9th Annual European Symposium Algorithms*, pages 230–241, 2001.

[47] R. L. Graham and P. Hell. On the history of the minimum spanning tree problem. *Ann. Hist. Comput.*, 7:43–57, 1985.

[48] D. Gusfield. Efficient methods for multiple sequence alignment with guaranteed error bounds. *Bull. Math. Biol.*, 55:141–154, 1993.

[49] D. Gusfield. *Algorithms on Strings, Trees, and Sequences – Computer Science and Computational Biology.* Cambridge University Press, 1997.

[50] T. Hagerup. Improved shortest paths in the word RAM. In *Proceedings of the 27th International Colloquium on Automata, Languages and Programming*, pages 61–72, 2000.

[51] P. Hansen and M. Zheng. Shortest shortest path trees of a network. *Discrete Appl. Math.*, 65:275–284, 1996.

[52] R. Hassin and A. Tamir. On the minimum diameter spanning tree problem. *Inf. Process. Lett.*, 53:109–111, 1995.

[53] J. Hein. An optimal algorithm to reconstruct trees from additive distance data. *Bull. Math. Biol.*, 51:597–603, 1989.

[54] J.-M. Ho, D.T. Lee, C.-H. Chang, and C.K. Wong. Minimum diameter spanning trees and related problems. *SIAM J. Comput.*, 20:987–997, 1991.

[55] T.C. Hu. Optimum communication spanning trees. *SIAM J. Comput.*, 3:188–195, 1974.

[56] F.K. Hwang. On Steiner minimal trees with rectilinear distance. *SIAM J. Appl. Math.*, 30:104–114, 1976.

[57] F.K. Hwang and D.S. Richards. Steiner tree problems. *Networks*, 22:55–89, 1992.

[58] O.H. Ibarra and C.E. Kim. Fast approximation algorithms for the knapsack and sum of subset problems. *J. ACM*, 22:463–468, 1975.

[59] V. Jarník. O jistém problému minimálním (about a certain minimal problem). *ráca Moravské Přirodovědecké Společnosti*, 6:57–63, 1930.

[60] D.S. Johnson, J.K. Lenstra, and A.H.G. Rinnooy Kan. The complexity of the network design problem. *Networks*, 8:279–285, 1978.

[61] D. R. Karger, P. N. Klein, and R. E. Tarjan. A randomized linear-time algorithm to find minimum spanning trees. *J. ACM*, 42:321–328, 1995.

[62] O. Kariv and S.L. Hakimi. An algorithmic approach to network location problems I: The p-centers. *SIAM J. Appl. Math.*, 37:513–537, 1979.

[63] R.M. Karp. Reducibility among combinatorial problems. In R.E. Miller and J.W. Thatcher, editors, *Complexity of Computer Computations*, pages 85–103. Plenum Press, New York, 1975.

[64] S. Khuller, B. Raghavachari, and N. Young. Balancing minimum spanning trees and shortest-path trees. *Algorithmica*, 14:305–321, 1995.

[65] S. Khuller, B. Raghavachari, and N. Young. Low degree spanning trees of small weight. *SIAM J. Comput.*, 25:355–368, 1996.

[66] V. King. A simpler minimum spanning tree verification algorithm. *Algorithmica*, 18:263–270, 1997.

[67] J. Komlós. Linear verification for spanning trees. *Combinatorica*, 5:57–65, 1985.

[68] L. Kou, G. Markowsky, and L. Berman. A fast algorithm for Steiner trees. *Acta Inform.*, 15(2):141–145, 1981.

[69] J. B. Kruskal. On the shortest spanning subtree of a graph and the travelling salesman problem. *Proc. Amer. Math. Soc.*, 7:48–50, 1956.

[70] E.L. Lawler. Fast approximation algorithms for knapsack problems. *Math. Oper. Res.*, 4(4):339–356, 1979.

[71] C. Levcopoulos and A. Lingas. There are planar graphs almost as good as the complete graphs and as cheap as minimum spanning trees. *Algorithmica*, 8(3):251–256, 1992.

[72] H.-I Lu and R. Ravi. The power of local optimization: Approximation for maximum-leaf spanning trees. In *Proceedings of the 13th Allerton Conference on Communication, Control and Computing*, pages 533–542, 1992.

[73] H.-I Lu and R. Ravi. Approximating maximum leaf spanning trees in almost linear time. *J. Algorithms*, 29(1):132–141, 1998.

[74] U. Meyer. Single-source shortest paths on arbitrary directed graphs inlinear average time. In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 797–806, 2001.

[75] D. Peleg and J.D. Ullman. An optimal synchronizer for the hypercube. In *Proceedings of the 6th Symposium on Principles of Distributed Computing*, pages 77–85, 1987.

[76] S. Pettie and V. Ramachandran. An optimal minimum spanning tree algorithm. *J. ACM*, 49:16–34, 2002.

[77] P. Pevzner. Multiple alignment, communication cost, and graph matching. *SIAM J. Appl. Math.*, 52:1763–1779, 1992.

[78] J. Plesnik. The complexity of designing a network with minimum diameter. *Networks*, 11:77–85, 1981.

[79] R. C. Prim. Shortest connection networks and some generalizations. *Bell. Syst. Tech. J.*, 36:1389–1401, 1957.

[80] H. Prüfer. Never beweis eines satzes über permutationen. *Arch. Math. Phys. Sci.*, 27:742–744, 1918.

[81] R. Raman. Priority queues: small, monotone and trans-dichotomous. In *Proceedings of the 4th Annual European Symposium Algorithms*, pages 121–137, 1996.

[82] R. Raman. Recent results on single-source shortest paths problem. *SIGACT News*, 28:81–87, 1997.

[83] D. Sankoff and J. B. Kruskal, editors. *Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison.* Addison Wesley, 1983.

[84] H. Takahashi and A. Mastsuyama. An approximate solution for the Steiner problem in graphs. *Math. Jap.*, 24:573–577, 1980.

[85] R. E. Tarjan. Efficiency of a good but not linear set-union algorithm. *J. ACM*, 22:215–225, 1975.

[86] R. E. Tarjan. Applications of path compressions on balanced trees. *J. ACM*, 26:690–715, 1979.

[87] M. Thorup. Undirected single-source shortest paths with positive integer weights in linear time. *J. ACM*, 46:362–394, 1999.

[88] M. Thorup. On RAM priority queues. *SIAM J. Comput.*, 30:86–109, 2000.

[89] P.M. Vaidya. A sparse graph almost as good as the complete graph on points in k dimensions. *Discrete Comput. Geom.*, 6:369–381, 1991.

[90] M. Waterman, T. Smith, M. Singh, and W. Beyer. Additive evolutionary trees. *J. Theor. Biol.*, 64:199–213, 1977.

[91] M.S. Waterman. *Introduction to Computational Biology.* Chapman & Hall, CRC Press, 1995.

[92] R. Wong. Worst-case analysis of network design problem heuristics. *SIAM J. Algebra. Discr.*, 1:51–63, 1980.

[93] B.Y. Wu. A polynomial time approximation scheme for the two-source minimum routing cost spanning trees. *J. Algorithms*, 44:359–378, 2002.

[94] B.Y. Wu. Approximation algorithms for optimal p-source communication spanning trees. Unpublished manuscript, 2003.

[95] B.Y. Wu, K.-M. Chao, and C.Y. Tang. Approximation and exact algorithms for constructing minimum ultrametric trees from distance matrices. *J. Comb. Optim.*, 3:199–211, 1999.

[96] B.Y. Wu, K.-M. Chao, and C.Y. Tang. Approximation algorithms for some optimum communication spanning tree problems. *Discrete Appl. Math.*, 102:245–266, 2000.

[97] B.Y. Wu, K.-M. Chao, and C.Y. Tang. Approximation algorithms for the shortest total path length spanning tree problem. *Discrete Appl. Math.*, 105:273–289, 2000.

[98] B.Y. Wu, K.-M. Chao, and C.Y. Tang. A polynomial time approximation scheme for optimal product-requirement communication spanning trees. *J. Algorithms*, 36:182–204, 2000.

[99] B.Y. Wu, K.-M. Chao, and C.Y. Tang. Light graphs with small routing cost. *Networks*, 39:130–138, 2002.

[100] B.Y. Wu, G. Lancia, V. Bafna, K.-M. Chao, R. Ravi, and C.Y. Tang. A polynomial time approximation scheme for minimum routing cost spanning trees. *SIAM J. Comput.*, 29:761–778, 2000.

[101] B.Y. Wu and C.Y. Tang. An $O(n)$ algorithm for finding an optimal position with relative distances in an evolutionary tree. *Inf. Process. Lett.*, 63:263–269, 1997.

[102] Y.F. Wu, P. Widmayer, and C.K. Wong. A faster approximation algorithm for the Steiner problem in graphs. *Acta Inform.*, 23(2):223–229, 1986.

[103] Y. Xu, V. Olman, and D. Xu. Clustering gene expression data using a graph-theoretic approach: An application of minimum spanning trees. *Bioinformatics*, 18:536–545, 2002.

[104] A. Yao. An $O(|E| \log\log |V|)$ algorithm for finding minimum spanning trees. *Inf. Process. Lett.*, 4:21–23, 1975.

[105] A. Zelikovsky. An 11/6-approximation algorithm for the network Steiner problem. *Algorithmica*, 9:463–470, 1993.

Index