

TPR values. The resulting ROC *plot* depicts relative trade-offs between true positive predictions and false positive prediction across different parameter values; see Fig. 5.23 for an illustration. An alternative plot is the *recall precision plot* obtained by plotting TPR versus PPV values.

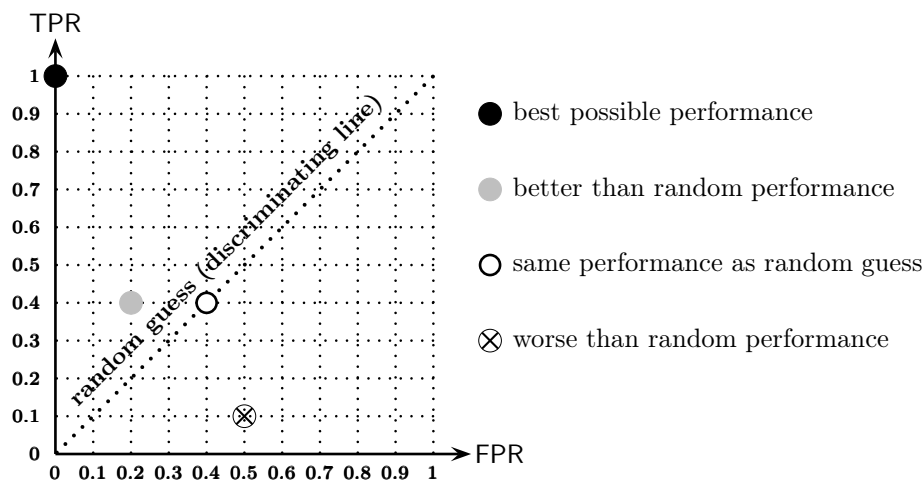


Figure 5.23 Two dimensional ROC space obtained by plotting FPR versus TPR values.

Examples of gold standard networks We give two examples of gold standard networks that can be used for benchmark testing of reverse engineering methods. Further discussion on generation of gold standard networks can be found later in Section 6.3.3.1.

(i) Gene regulatory networks with external perturbations can be generated from the differential equation models using the software package in [61].

(ii) Time courses can be generated from the Boolean model of network of segment polarity genes involved in pattern formation in the *Drosophila melanogaster* embryo. This model was proposed by Albert and Othmer [7]; the network for each cell has 15 nodes.

REFERENCES

1. B. Alberts. Molecular biology of the cell, New York: Garland Publishers, 1994.
2. R. Albert and A.-L. Barabási. Statistical mechanics of complex networks, *Reviews of Modern Physics*, 74 (1), 47-97, 2002.
3. R. Albert, B. DasGupta, A. Gitter, G. Gürsoy, R. Hegde, P. Pal, G. S. Sivanathan and E. Sontag. A New Computationally Efficient Measure of Topological Redundancy of Biological and Social Networks, *Physical Review E*, 84 (3), 036117, 2011.
4. R. Albert, B. DasGupta, R. Dondi and E. Sontag. Inferring (Biological) Signal Transduction Networks via Transitive Reductions of Directed Graphs, *Algorithmica*, 51 (2), 129-159, 2008.

5. R. Albert, B. DasGupta, R. Dondi, S. Kachalo, E. Sontag, A. Zelikovsky and K. Westbrooks. A Novel Method for Signal Transduction Network Inference from Indirect Experimental Evidence, *Journal of Computational Biology*, 14 (7), 927-949, 2007.
6. R. Albert, B. DasGupta and E. Sontag. Inference of signal transduction networks from double causal evidence, in *Methods in Molecular Biology: Topics in Computational Biology*, D. Fenyo (ed.), 673, Chapter 16, Springer, 2010.
7. R. Albert and H. Othmer. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*, *Journal of Theoretical Biology*, 223, 1-18, 2003.
8. M. Andrec, B. N. Kholodenko, R. M. Levy, and E. D. Sontag. Inference of Signaling and Gene Regulatory Networks by Steady-State Perturbation Experiments: Structure and Accuracy, *Journal of Theoretical Biology*, 232, 427-441, 2005.
9. D. Angeli and E.D. Sontag. Monotone control systems, *IEEE Transactions on Automatic Control*, 48, 1684-1698, 2003.
10. J. Aracena, M. González, A. Zuñiga, M. Méndez and V. Cambiazo. Regulatory network for cell shape changes during *Drosophila* ventral furrow formation, *Journal of Theoretical Biology*, 239 (1), 49-62, 2007.
11. M.J. Beal and F. Falciani. A Bayesian approach to reconstructing genetic regulatory networks with hidden factors, *Bioinformatics*, 21 (3), 349-356, 2005.
12. N. Beckage, L. Smith and T. Hills. Semantic network connectivity is related to vocabulary growth rate in children, *Annual Meeting of The Cognitive Science Society*, 2769-2774, 2010.
13. P. Berman, B. DasGupta and M. Karpinski. Approximating Transitive Reduction Problems for Directed Networks, *Algorithms and Data Structures Symposium*, F. Dehne, M. Gavrilova, J.-R. Sack and C. D. Tóth (eds.), LNCS 5664, 74-85, 2009.
14. P. Berman, B. DasGupta and E. Sontag. Randomized Approximation Algorithms for Set Multicover Problems with Applications to Reverse Engineering of Protein and Gene Networks, *Discrete Applied Mathematics*, 155 (6-7), 733-749, 2007.
15. P. Berman, B. DasGupta and E. Sontag. Computational Complexities of Combinatorial Problems With Applications to Reverse Engineering of Biological Networks, in *Advances in Computational Intelligence: Theory and Applications*, F.-Y. Wang and D. Liu (editors), *Series in Intelligent Control and Intelligent Automation Volume 5*, World Scientific publishers, 303-316, 2007.
16. P. Berman, B. DasGupta and E. Sontag. Algorithmic Issues in Reverse Engineering of Protein and Gene Networks via the Modular Response Analysis Method, *Annals of the New York Academy of Sciences*, 1115, 132-141, 2007.
17. J. A. Bondy and U. S. R. Murty. *Graph Theory*, Springer, 2008.
18. T. Chen, V. Filkov and S. Skiena. Identifying Gene Regulatory Networks from Experimental Data, *Third Annual International Conference on Computational Molecular Biology*, 94-103, 1999.
19. D.-Z. Cheng, H. Qi, and Z. Li. *Analysis and control of boolean networks: A semi-tensor product approach*, London: Springer, 2011.
20. T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein. *Introduction to Algorithms*, The MIT Press, 2001.
21. E. J. Crampin, S. Schnell, and P. E. McSharry. Mathematical and computational techniques to deduce complex biochemical reaction mechanisms, *Progress in Biophysics & Molecular Biology*, 86, 77-112, 2004.

22. B. DasGupta, G. A. Enciso, E. Sontag and Y. Zhang. Algorithmic and complexity results for decompositions of biological networks into monotone subsystems, *Biosystems*, 90 (1), 161-178, 2007.
23. B. DasGupta, P. Vera-Licona and E. Sontag. Reverse Engineering of Molecular Networks from a Common Combinatorial Approach, in *Algorithms in Computational Molecular Biology: Techniques, Approaches and Applications*, M. Elloumi and A. Zomaya (editors), 941-955, Chapter 40, John Wiley & Sons, Inc., 2011.
24. L. Dall'Asta, I. Alvarez-Hamelina, A. Barrata, A. Vázquezb and A. Vespignania. Exploring networks with traceroute-like probes: Theory and simulations, *Theoretical Computer Science*, 355, 6-24, 2006.
25. E. H. Davidson. *The regulatory genome*, Academic Press, 2006.
26. N. Dojer, A. Gambin, A. Mizera, B. Wilczynski and J. Tiuryn. Applying dynamic Bayesian networks to perturbed gene expression data, *BMC Bioinformatics*, 7 (1), 249, 2006.
27. G. Enciso and E. Sontag. On the stability of a model of testosterone dynamics, *Journal of Mathematical Biology*, 49, 627-634, 2004.
28. U. Feige. A threshold for approximating set cover, *Journal of the ACM*, 45, 634-652, 1998.
29. S. Fields. High-throughput two-hybrid analysis: the promise and the peril, *FEBS Journal*, 272 (21), 5391-5399, 2005.
30. N. Friedman, M. Linial, I. Nachman and D. Pe'er. Using Bayesian Networks to Analyze Expression Data, *Journal of Computational Biology*, 7 (3-4), 600-620, 2000.
31. C. Friedman, P. Kra, H. Yu, M. Krauthammer and A. Rzhetsky. GENIES: a natural-language processing system for the extraction of molecular pathways from journal articles, *Bioinformatics*, 17 (Suppl 1), S74-82, 2001.
32. A. de la Fuente, N. Bing, I. Hoeschele and P. Mendes. Discovery of meaningful associations in genomic data using partial correlation coefficients, *Bioinformatics*, 20, 3565-3574, 2004.
33. L. Giot, J. S. Bader, C. Brouwer, A. Chaudhuri, B. Kuang, Y. Li, Y. L. Hao, C. E. Ooi, B. Godwin, E. Vitols, G. Vijayadamodar, P. Pochart, H. Machineni, M. Welsh, Y. Kong, B. Zerhusen, R. Malcolm, Z. Varrone, A. Collis, M. Minto, S. Burgess, L. McDaniel, E. Stimpson, F. Spriggs, J. Williams, K. Neurath, N. Ioime, M. Agee, E. Voss, K. Furtak, R. Renzulli, N. Aanensen, S. Carrolla, E. Bickelhaupt, Y. Lazovatsky, A. DaSilva, J. Zhong, C. A. Stanyon, R. L. Finley, K. P. White, M. Braverman, T. Jarvie, S. Gold, M. Leach, J. Knight, R. A. Shinkets, M. P. McKenna, J. Chant and J. M. Rothberg. A protein interaction map of *Drosophila melanogaster*, *Science*, 302 (5651), 1727-1736, 2003.
34. M. Girvan and M. E. J. Newman. Community structure in social and biological networks, *Proceedings of the National Academy of Sciences USA*, 99, 7821-7826, 2002.
35. R. Guimerà, M. Sales-Pardo and L. A. N. Amaral. Classes of complex networks defined by role-to-role connectivity profiles, *Nature Physics*, 3, 63-69, 2007.
36. J. Hann and M. Kamber. *Data Mining: Concepts and Techniques*, Morgan Kaufman Publishers, 2000.
37. S. Huang. Gene expression profiling, genetic networks and cellular states: an integrating concept for tumorigenesis and drug discovery, *Journal of molecular medicine*, 77 (6), 469-480, 1999.

38. T. E. Ideker, V. Thorsson and R. M. Karp. Discovery of regulatory interactions through perturbation: inference and experimental design, *Pacific Symposium on Biocomputing*, 5, 305-316, 2000.
39. A. S. Jarrah, R. Laubenbacher, B. Stigler and M. Stillman. Reverse-engineering polynomial dynamical systems, *Advances in Applied Mathematics*, 39 (4) 477-489, 2007.
40. L. J. Jensen, J. Saric and P. Bork. Literature mining for the biologist: from information retrieval to biological discovery, *Nature Review Genetics*, 7 (2), 119-129, 2006.
41. H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai and A.-L. Barabási. The large-scale organization of metabolic networks, *Nature*, 407, 651-654, 2000.
42. D. S. Johnson. Approximation Algorithms for Combinatorial Problems, *Journal of Computer and Systems Sciences*, 9, 256-278, 1974.
43. S. Kachalo, R. Zhang, E. Sontag, R. Albert and B. DasGupta. NET-SYNTHESIS: A software for synthesis, inference and simplification of signal transduction networks, *Bioinformatics*, 24 (2), 293-295, 2008.
44. R. Kannan, P. Tetali and S. Vempala. Markov-chain algorithms for generating bipartite graphs and tournaments, *Random Structures and Algorithms*, 14, 293-308, 1999.
45. N. Karmarkar. A New Polynomial-time Algorithm for Linear Programming, *Combinatorica*, 4, 373-395, 1984.
46. S. A. Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets, *Journal of Theoretical Biology*, 22, 437-467, 1969.
47. B. N. Kholodenko, A. Kiyatkin, F. Bruggeman, E.D. Sontag, H. Westerhoff, and J. Hoek. Untangling the Wires: A Novel Strategy to Trace Functional Interactions in Signaling and Gene Networks, *Proceedings of the National Academy of Sciences USA*, 99, 12841-12846, 2002.
48. B. N. Kholodenko and E.D. Sontag. Determination of Functional Network Structure from Local Parameter Dependence Data, *arXiv physics/0205003*, May 2002.
49. B. Kolb and I. Q. Whishaw. *Fundamentals of Human Neuropsychology*, Freeman, New York, 1996.
50. S. Khuller, B. Raghavachari and N. Young. On strongly connected digraphs with bounded cycle length, *Discrete Applied Mathematics*, 69 (3), 281-289, 1996.
51. S. Khuller, B. Raghavachari and N. Young. Approximating the minimum equivalent digraph, *SIAM Journal of Computing*, 24 (4), 859-872, 1995.
52. B. Krupa. On the Number of Experiments Required to Find the Causal Structure of Complex Systems, *Journal of Theoretical Biology*, 219 (2), 257-267, 2002.
53. R. J. Larsen and M. L. Marx. *An Introduction to Mathematical Statistics and Its Applications*, Third Edition, 2000.
54. E. A. Leicht and M. E. J. Newman. Community Structure in Directed Networks, *Physical Review Letters*, 100, 118703, 2008.
55. S. Li, C. M. Armstrong, N. Bertin, H. Ge, S. Milstein, M. Boxem, P.-O. Vidalain, J.-D. J. Han, A. Chesneau, T. Hao, D. S. Goldberg, N. Li, M. Martinez, J.-F. Rual, P. Lamesch, L. Xu, M. Tewari, S. L. Wong, L. V. Zhang, G. F. Berriz, L. Jacotot, P. Vaglio, J. Reboul, T. Hirozane-Kishikawa, Q. Li, H. W. Gabel, A. Elewa, B. Baumgartner, D. J. Rose, H. Yu, S. Bosak, R. Sequerra, A. Fraser, S. E. Mango, W. M. Saxton, S. Strome, S. van den Heuvel, F. Piano, J. Vandenhaute, C. Sardet,

- M. Gerstein, L. Doucette-Stamm, K. C. Gunsalus, J. W. Harper, M. E. Cusick, F. P. Roth, D. E. Hill, and M. Vidal. A map of the interactome network of the metazoan *C. elegans*, *Science*, 303, 540-543, 2004.
56. T. I. Lee, N. J. Rinaldi, F. Robert, D. T. Odom, Z. Bar-Joseph, G. K. Gerber, N. M. Hannett, C. T. Harbison, C. M. Thompson, I. Simon, J. Zeitlinger, E. G. Jennings, H. L. Murray, D. B. Gordon, B. Ren, J. J. Wyrick, J.-B. Tagne, T. L. Volkert, E. Fraenkel, D. K. Gifford, and R. A. Young. Transcriptional regulatory networks in *Saccharomyces cerevisiae*, *Science*, 298 (5594), 799-804, 2002.
 57. S. Li, S. M. Assmann and R. Albert. Predicting Essential Components of Signal Transduction Networks: A Dynamic Model of Guard Cell Abscisic Acid Signaling, *PLoS Biology*, 4(10), e312, 2006.
 58. A. Ma'ayan, S. L. Jenkins, S. Neves, A. Hasseldine, E. Grace, B. Dubin-Thaler, N. J. Eungdamrong, G. Weng, P. T. Ram, J. J. Rice, A. Kershenbaum, G. A. Stolovitzky, R. D. Blitzer and R. Iyengar. Formation of Regulatory Patterns During Signal Propagation in a Mammalian Cellular Network, *Science*, 309 (5737), 1078-1083, 2005.
 59. E. M. Marcotte, I. Xenarios and D. Eisenberg. Mining literature for protein-protein interactions, *Bioinformatics*, 17 (4), 359-363, 2001.
 60. N. Megiddo. Linear programming in linear time when the dimension is fixed, *Journal of ACM*, 31, 114-127, 1984.
 61. P. Mendes. Biochemistry by numbers: simulation of biochemical pathways with Gepasi 3, *Trends in Biochemical Sciences*, 22, 361-363, 1997.
 62. J. D. Murray. *Mathematical Biology, I: An introduction*, New York, Springer, 2002.
 63. N. Nariai, Y. Tamada, S. Imoto and S. Miyano. Estimating gene regulatory networks and protein-protein interactions of *Saccharomyces cerevisiae* from multiple genome-wide data, *Bioinformatics*, 21 (suppl 2), ii206-ii212, 2005.
 64. M. E. J. Newman. The structure and function of complex networks, *SIAM Review*, 45, 167-256, 2003.
 65. M. E. J. Newman. Detecting community structure in networks, *European Physics Journal B*, 38, 321-330, 2004.
 66. M. E. J. Newman and G. T. Barkema. *Monte Carlo Methods in Statistical Physics*, Oxford University Press, 1999.
 67. M. E. J. Newman and M. Girvan. Finding and evaluating community structure in networks, *Physical Review E*, 69, 026113, 2004.
 68. M. E. J. Newman, S. H. Strogatz and D. J. Watts. Random graphs with arbitrary degree distributions and their applications, *Physical Review E*, 64 (2), 026118-026134, 2001.
 69. J. A. Papin and B. O. Palsson. Topological analysis of mass-balanced signaling networks: a framework to obtain network properties including crosstalk, *Journal of Theoretical Biology*, 227 (2), 283-297, 2004.
 70. I. Pournara and L. Wernisch. Reconstruction of gene networks using Bayesian learning and manipulation experiments, *Bioinformatics*, 20 (17), 2934-2942, 2004.
 71. E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai and A.-L. Barabási. Hierarchical Organization of Modularity in Metabolic Networks, *Science*, 297 (5586), 1551-1555, 2002.
 72. J.J. Rice, Y. Tu and G. Stolovitzky. Reconstructing biological networks using conditional correlation analysis, *Bioinformatics*, 21 (6), 765-773, 2005.

73. S. D. M. Santos, P. J. Verveer, and P. I. H. Bastiaens. Growth factor-induced MAPK network topology shapes Erk response determining PC-12 cell fate, *Nature Cell Biology*, 9, 324-330, 2007.
74. S. S. Shen-Orr, R. Milo, S. Mangan and U. Alon. Network motifs in the transcriptional regulation network of *Escherichia coli*, *Nature Genetics* 31, 64-68, 2002.
75. H. L. Smith. *Monotone Dynamical Systems*, Providence, R.I., AMS 1995.
76. E. D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, Springer, New York, 1998.
77. E. D. Sontag. Molecular systems biology and control, *European Journal of Control*, 11 (4-5), 396-435, 2005.
78. E. D. Sontag, A. Kiyatkin and B. N. Kholodenko. Inferring Dynamic Architecture of Cellular Networks Using Time Series of Gene Expression, Protein and Metabolite Data, *Bioinformatics*, 20, 1877-1886, 2004.
79. J. Stark, R. Callard and M. Hubank. From the Top Down: Towards a Predictive Biology of Signalling Networks, *Trends in Biotechnology*, 21, 290-293, 2003.
80. G. Tononi, O. Sporns and G. M. Edelman. A measure for brain complexity: relating functional segregation and integration in the nervous system, *Proceedings of the National Academy of Sciences USA*, 91 (11), 5033-5037, 1994.
81. G. Tononi, O. Sporns and G. M. Edelman. A complexity measure for selective matching of signals by the brain, *Proceedings of the National Academy of Sciences USA*, 93, 3422-3427, 1996.
82. G. Tononi, O. Sporns and G. M. Edelman. Measures of degeneracy and redundancy in biological networks, *Proceedings of the National Academy of Sciences USA*, 96, 3257-3262, 1999.
83. L. Trevisan. Non-approximability results for optimization problems on bounded-degree instance, *Thirty third ACM Symposium on Theory of Computing*, 453-461, 2001.
84. V. Vazirani. *Approximation Algorithms*, Springer-Verlag, 2001.
85. A. Wagner. Estimating Coarse Gene Network Structure from Large-Scale Gene Perturbation Data, *Genome Research*, 12 (2), 309-315, 2002.
86. J. Yu, V. Smith, P. Wang, A. Hartemink and E Jarvis. Advances to Bayesian network inference for generating causal networks from observational biological data, *Bioinformatics*, 20, 3594-3603, 2004.
87. R. Zhang, M. V. Shah, J. Yang, S. B. Nyland, X. Liu, J. K. Yun, R. Albert and T. P. Loughran, Jr. Network Model of Survival Signaling in LGL Leukemia, *Proceedings of the National Academy of Sciences USA*, 105 (42), 16308-16313, 2008.
88. M. Zou and S. D. Conzen. A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data, *Bioinformatics*, 21 (1), 71-79, 2005.

EXERCISES

5.1 Suppose that your data suggests that Protein A inhibits the transcription of the RNA that codes Protein B, whereas Protein B in turn enhances the production