

Bibliography

- [1] Abbe, E. 2018. Community Detection and Stochastic Block Models: Recent Developments. *Journal of Machine Learning Research*, **18**(177), 1–86. (Cited on p. 356.)
- [2] Acebrón, J. A., Bonilla, L. L., Pérez Vicente, C. J., Ritort, F., and Spigler, R. 2005. The Kuramoto Model: A Simple Paradigm for Synchronization Phenomena. *Reviews of Modern Physics*, **77**(1), 137–185. (Cited on p. 235.)
- [3] Adamic, L. A., and Adar, E. 2003. Friends and Neighbors on the Web. *Social Networks*, **25**(3), 211–230. (Cited on p. 131.)
- [4] Aggarwal, C. C., Zhao, Y., and Yu, P. S. 2011. Outlier Detection in Graph Streams. Pages 399–409 of: *2011 IEEE 27th International Conference on Data Engineering*. Hannover, Germany: IEEE. (Cited on p. 469.)
- [5] Ahn, Y.-Y., Bagrow, J. P., and Lehmann, S. 2010. Link Communities Reveal Multiscale Complexity in Networks. *Nature*, **466**(7307), 761–764. (Cited on pp. 184, 201.)
- [6] Ahn, Y.-Y., Ahnert, S. E., Bagrow, J. P., and Barabási, A.-L. 2011. Flavor Network and the Principles of Food Pairing. *Scientific Reports*, **1**(1), 196. (Cited on pp. 9, 10, 22, 24, 26.)
- [7] Airoldi, E. M., Blei, D., Fienberg, S., and Xing, E. 2008. Mixed Membership Stochastic Blockmodels. In: *Advances in Neural Information Processing Systems*, vol. 21. Curran Associates, Inc. (Cited on p. 358.)
- [8] Al Hasan, M., and Dave, V. S. 2018. Triangle Counting in Large Networks: A Review. *WIREs Data Mining and Knowledge Discovery*, **8**(2), e1226. (Cited on p. 469.)
- [9] Albert, R., Jeong, H., and Barabási, A.-L. 2000. Error and Attack Tolerance of Complex Networks. *Nature*, **406**(6794), 378–382. (Cited on p. 167.)
- [10] Allen, A. 2007. *Vaccine: The Controversial Story of Medicine's Greatest Life-saver*. New York, NY: W.W. Norton. (Cited on p. 67.)
- [11] Almeida-Neto, M., Guimarães, P., Guimarães Jr, P. R., Loyola, R. D., and Ulrich, W. 2008. A Consistent Metric for Nestedness Analysis in Ecological Systems: Reconciling Concept and Measurement. *Oikos*, **117**(8), 1227–1239. (Cited on p. 202.)
- [12] Alon, U., and Yahav, E. 2021. On the Bottleneck of Graph Neural Networks and

- Its Practical Implications. In: *International Conference on Learning Representations*. Virtual Event, Austria: OpenReview.net. (Cited on p. 444.)
- [13] Alstott, J., Bullmore, E., and Plenz, D. 2014. Powerlaw: A Python Package for Analysis of Heavy-Tailed Distributions. *PLoS ONE*, **9**(1), e85777. (Cited on pp. 156, 163.)
- [14] Amstrup, S. C., McDonald, T. L., and Manly, B. F. J. (eds.). 2005. *Handbook of Capture-Recapture Analysis*. Princeton, NJ: Princeton University Press. (Cited on pp. 391, 395.)
- [15] Ashburner, M., Ball, C. A., Blake, J. A. et al. 2000. Gene Ontology: Tool for the Unification of Biology. *Nature Genetics*, **25**(1), 25–29. (Cited on p. 116.)
- [16] Ashmore, R., Calinescu, R., and Paterson, C. 2021. Assuring the Machine Learning Lifecycle: Desiderata, Methods, and Challenges. *ACM Computing Surveys*, **54**(5), 111:1–111:39. (Cited on p. 31.)
- [17] Atmar, W., and Patterson, B. D. 1993. The Measure of Order and Disorder in the Distribution of Species in Fragmented Habitat. *Oecologia*, **96**(3), 373–382. (Cited on p. 202.)
- [18] Babai, L. 2016. Graph Isomorphism in Quasipolynomial Time [Extended Abstract]. Pages 684–697 of: *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*. STOC '16. New York, NY: Association for Computing Machinery. (Cited on p. 443.)
- [19] Babai, L. 2019. Canonical Form for Graphs in Quasipolynomial Time: Preliminary Report. Pages 1237–1246 of: *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2019. New York, NY: Association for Computing Machinery. (Cited on p. 443.)
- [20] Backstrom, L., Boldi, P., Rosa, M., Ugander, J., and Vigna, S. 2012. Four Degrees of Separation. Pages 33–42 of: *Proceedings of the 4th Annual ACM Web Science Conference*. Evanston, IL: Association for Computing Machinery. (Cited on pp. 341, 458.)
- [21] Baek, J., Kang, M., and Hwang, S. J. 2022. Accurate Learning of Graph Representations with Graph Multiset Pooling. In: *International Conference on Learning Representations*. Virtual Event, Austria: OpenReview.net. (Cited on p. 442.)
- [22] Bagrow, J., and Ahn, Y.-Y. 2022. Network Cards: Concise, Readable Summaries of Network Data. *Applied Network Science*, **7**(1), 1–17. (Cited on pp. 225, 232, 474.)
- [23] Bagrow, J. P. 2008. Evaluating Local Community Methods in Networks. *Journal of Statistical Mechanics: Theory and Experiment*, **2008**(05), P05001. (Cited on p. 184.)
- [24] Bagrow, J. P. 2012. Communities and Bottlenecks: Trees and Treelike Networks Have High Modularity. *Physical Review E*, **85**(6), 066118. (Cited on p. 179.)
- [25] Bagrow, J. P. 2020. Democratizing AI: Non-Expert Design of Prediction Tasks. *PeerJ Computer Science*, **6**(Sept.), e296. (Cited on p. 374.)
- [26] Bagrow, J. P., and Boltt, E. M. 2005. Local Method for Detecting Communities. *Physical Review E*, **72**(4), 046108. (Cited on pp. 184, 185, 201.)

- [27] Bagrow, J. P., and Boltt, E. M. 2019. An Information-Theoretic, All-Scales Approach to Comparing Networks. *Applied Network Science*, **4**(1), 45. (Cited on pp. 24, 216, 228.)
- [28] Bagrow, J. P., Lehmann, S., and Ahn, Y.-Y. 2015. Robustness and Modular Structure in Networks. *Network Science*, **3**(4), 509–525. (Cited on pp. 386, 395.)
- [29] Bagrow, J. P., Danforth, C. M., and Mitchell, L. 2017. Which Friends Are More Popular than You?: Contact Strength and the Friendship Paradox in Social Networks. Pages 103–108 of: *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2017*. Sydney, Australia: Association for Computing Machinery. (Cited on p. 323.)
- [30] Bagrow, J. P., Liu, X., and Mitchell, L. 2019. Information Flow Reveals Prediction Limits in Online Social Activity. *Nature Human Behaviour*, **3**(2), 122–128. (Cited on p. 37.)
- [31] Baker, R. J. 2010. *CMOS: Circuit Design, Layout, and Simulation*. 3rd edn. IEEE Press Series on Microelectronic Systems. Piscataway, NJ; Hoboken, NJ: IEEE. (Cited on p. 26.)
- [32] Balcilar, M., Renton, G., Héroux, P., Gaüzère, B., Adam, S., and Honeine, P. 2022. Analyzing the Expressive Power of Graph Neural Networks in a Spectral Perspective. In: *International Conference on Learning Representations*. Virtual Event, Austria: OpenReview.net. (Cited on p. 443.)
- [33] Ball, P. 2006. *Critical Mass: How One Thing Leads to Another*. New York, NY: Farrar, Straus and Giroux. (Cited on p. 5.)
- [34] Banerjee, O., El Ghaoui, L., and d’Aspremont, A. 2008. Model Selection Through Sparse Maximum Likelihood Estimation for Multivariate Gaussian or Binary Data. *The Journal of Machine Learning Research*, **9**(June), 485–516. (Cited on p. 371.)
- [35] Barabási, A.-L. 2016. *Network Science*. Cambridge: Cambridge University Press. (Cited on pp. 16, 472.)
- [36] Barabási, A.-L., and Albert, R. 1999. Emergence of Scaling in Random Networks. *Science*, **286**(5439), 509–512. (Cited on pp. 344, 346, 349.)
- [37] Barabási, A.-L., and Oltvai, Z. N. 2004. Network Biology: Understanding the Cell’s Functional Organization. *Nature Reviews Genetics*, **5**(2), 101–113. (Cited on p. 25.)
- [38] Barber, M. J., and Clark, J. W. 2009. Detecting Network Communities by Propagating Labels under Constraints. *Physical Review E*, **80**(2), 026129. (Cited on p. 469.)
- [39] Barnes, E. R. 1982. An Algorithm for Partitioning the Nodes of a Graph. *SIAM Journal on Algebraic Discrete Methods*, **3**(4), 541–550. (Cited on p. 427.)
- [40] Bartomeus, I., Vilà, M., and Santamaría, L. 2008. Contrasting Effects of Invasive Plants in Plant–Pollinator Networks. *Oecologia*, **155**(4), 761–770. (Cited on p. 23.)
- [41] Bascompte, J. 2007. Networks in Ecology. *Basic and Applied Ecology*, **8**(6), 485–490. (Cited on p. 25.)

- [42] Bassler, P. J., Mattiello, J., and LeBihan, D. 1994. Estimation of the Effective Self-Diffusion Tensor from the NMR Spin Echo. *Journal of Magnetic Resonance, Series B*, **103**(3), 247–254. (Cited on p. 8.)
- [43] Bassler, P. J., Mattiello, J., and LeBihan, D. 1994. MR Diffusion Tensor Spectroscopy and Imaging. *Biophysical Journal*, **66**(1), 259–267. (Cited on p. 8.)
- [44] Bassett, D. S., and Sporns, O. 2017. Network Neuroscience. *Nature Neuroscience*, **20**(3), 353–364. (Cited on pp. 25, 82.)
- [45] Batagelj, V., and Brandes, U. 2005. Efficient Generation of Large Random Networks. *Physical Review E*, **71**(3), 036113. (Cited on p. 469.)
- [46] Batson, J., Spielman, D. A., Srivastava, N., and Teng, S.-H. 2013. Spectral Sparsification of Graphs: Theory and Algorithms. *Communications of the ACM*, **56**(8), 87–94. (Cited on p. 134.)
- [47] Battiston, S., Caldarelli, G., May, R. M., Roukny, T., and Stiglitz, J. E. 2016. The Price of Complexity in Financial Networks. *Proceedings of the National Academy of Sciences*, **113**(36), 10031–10036. (Cited on p. 22.)
- [48] Belkin, M., and Niyogi, P. 2003. Laplacian Eigenmaps for Dimensionality Reduction and Data Representation. *Neural Computation*, **15**(6), 1373–1396. (Cited on p. 430.)
- [49] Belkin, M., Hsu, D., Ma, S., and Mandal, S. 2019. Reconciling Modern Machine-Learning Practice and the Classical Bias–Variance Trade-Off. *Proceedings of the National Academy of Sciences*, **116**(32), 15849–15854. (Cited on p. 277.)
- [50] Bender, E. M., Gebru, T., McMillan-Major, A., and Shmitchell, S. 2021. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? Pages 610–623 of: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. Virtual Event, Canada: Association for Computing Machinery. (Cited on p. 29.)
- [51] Bengio, Y., Ducharme, R., Vincent, P., and Janvin, C. 2003. A Neural Probabilistic Language Model. *The Journal of Machine Learning Research*, **3**(null), 1137–1155. (Cited on p. 445.)
- [52] Bentley, J., Knuth, D., and McIlroy, D. 1986. Programming Pearls: A Literate Program. *Communications of the ACM*, **29**(6), 471–483. (Cited on p. 313.)
- [53] Bentley, J. L. 1975. Multidimensional Binary Search Trees Used for Associative Searching. *Communications of the ACM*, **18**(9), 509–517. (Cited on p. 426.)
- [54] Bharat, K., Broder, A., Henzinger, M., Kumar, P., and Venkatasubramanian, S. 1998. The Connectivity Server: Fast Access to Linkage Information on the Web. *Computer Networks and ISDN Systems*, **30**(1-7), 469–477. (Cited on p. 469.)
- [55] Bhattacharya, A., Friedland, S., and Peled, U. N. 2008. On the First Eigenvalue of Bipartite Graphs. *The Electronic Journal of Combinatorics [electronic only]*, **15**(1). (Cited on p. 197.)
- [56] Bishop, C. M. 2006. *Pattern Recognition and Machine Learning*. Berlin, Heidelberg: Springer. (Cited on pp. 277, 472.)
- [57] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. 2008. Fast

- Unfolding of Communities in Large Networks. *Journal of Statistical Mechanics: Theory and Experiment*, **2008**(10), P10008. (Cited on pp. 179, 358.)
- [58] Bloom, B. H. 1970. Space/Time Trade-Offs in Hash Coding with Allowable Errors. *Communications of the ACM*, **13**(7), 422–426. (Cited on p. 455.)
- [59] Boguñá, M., Pastor-Satorras, R., and Vespignani, A. 2004. Cut-Offs and Finite Size Effects in Scale-Free Networks. *The European Physical Journal B*, **38**(2), 205–209. (Cited on pp. 163, 173.)
- [60] Boguñá, M., Bonamassa, I., De Domenico, M., Havlin, S., Krioukov, D., and Serrano, M. Á. 2021. Network Geometry. *Nature Reviews Physics*, **3**(2), 114–135. (Cited on p. 438.)
- [61] Böhning, D., van der Heijden, P. G. M., and Bunge, J. (eds.). 2018. *Capture-Recapture Methods for the Social and Medical Sciences*. Boca Raton, FL: CRC Press/Taylor & Francis Group. (Cited on pp. 391, 395.)
- [62] Boldi, P., and Vigna, S. 2004. The WebGraph Framework I: Compression Techniques. Pages 595–602 of: *Proceedings of the 13th International Conference on World Wide Web*. New York, NY: Association for Computing Machinery. (Cited on pp. 451, 469.)
- [63] Boldi, P., Rosa, M., and Vigna, S. 2011. HyperANF: Approximating the Neighbourhood Function of Very Large Graphs on a Budget. Pages 625–634 of: *Proceedings of the 20th International Conference on World Wide Web*. WWW '11. New York, NY: Association for Computing Machinery. (Cited on pp. 456, 457, 469.)
- [64] Boldi, P., Rosa, M., Santini, M., and Vigna, S. 2011. Layered Label Propagation: A Multiresolution Coordinate-Free Ordering for Compressing Social Networks. Pages 587–596 of: *Proceedings of the 20th International Conference on World Wide Web*. WWW '11. New York, NY: Association for Computing Machinery. (Cited on pp. 452, 469.)
- [65] Bordes, A., Usunier, N., Garcia-Duran, A., Weston, J., and Yakhnenko, O. 2013. Translating Embeddings for Modeling Multi-relational Data. In: *Advances in Neural Information Processing Systems*, vol. 26. Curran Associates, Inc. (Cited on p. 437.)
- [66] Borgatti, S. P., and Everett, M. G. 2000. Models of Core/Periphery Structures. *Social Networks*, **21**(4), 375–395. (Cited on pp. 194, 195, 201.)
- [67] Borgatti, S. P., Carley, K. M., and Krackhardt, D. 2006. On the Robustness of Centrality Measures under Conditions of Imperfect Data. *Social Networks*, **28**(2), 124–136. (Cited on pp. 377, 395.)
- [68] Borwein, J., and Bailey, D. 2013. The Reinhart-Rogoff Error – or How Not to Excel at Economics. *The Conversation*, Apr. (Cited on p. 36.)
- [69] Bostrom, N. 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford: Oxford University Press. (Cited on p. 277.)
- [70] Boswell, D., and Foucher, T. 2011. *The Art of Readable Code*. Theory in Practice. Sebastopol, CA: O'Reilly Media. (Cited on pp. 293, 300.)

- [71] Bouveyron, C., Celeux, G., Murphy, T. B., and Raftery, A. E. 2019. *Model-Based Clustering and Classification for Data Science: With Applications in R*. Cambridge: Cambridge University Press. (Cited on p. 428.)
- [72] boyd, d., Golder, S., and Lotan, G. 2010. Tweet, Tweet, Retweet: Conversational Aspects of Retweeting on Twitter. Pages 1–10 of: *2010 43rd Hawaii International Conference on System Sciences*. Honolulu, HI: IEEE. (Cited on p. 84.)
- [73] Breiger, R. L., Boorman, S. A., and Arabie, P. 1975. An Algorithm for Clustering Relational Data with Applications to Social Network Analysis and Comparison with Multidimensional Scaling. *Journal of Mathematical Psychology*, **12**(3), 328–383. (Cited on p. 374.)
- [74] Brenner, S. 2009. In the Beginning Was the Worm *Genetics*, **182**(2), 413–415. (Cited on p. 82.)
- [75] Brin, S., and Page, L. 1998. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Computer Networks and ISDN Systems*, **30**(1), 107–117. (Cited on p. 189.)
- [76] Broder, A. 1997. On the Resemblance and Containment of Documents. Pages 21–29 of: *Proceedings. Compression and Complexity of SEQUENCES 1997 (Cat. No.97TB100171)*. Salerno, Italy: IEEE. (Cited on p. 455.)
- [77] Brody, S., Alon, U., and Yahav, E. 2022. How Attentive Are Graph Attention Networks? In: *International Conference on Learning Representations*. Virtual Event, Austria: OpenReview.net. (Cited on pp. 273, 442.)
- [78] Broido, A. D., and Clauset, A. 2019. Scale-Free Networks Are Rare. *Nature Communications*, **10**(1), 1017. (Cited on p. 346.)
- [79] Brown, T., Mann, B., Ryder, N. et al. 2020. Language Models Are Few-Shot Learners. Pages 1877–1901 of: *Advances in Neural Information Processing Systems*, vol. 33. Vancouver, BC: Curran Associates, Inc. (Cited on p. 261.)
- [80] Bryan, K., and Leise, T. 2006. The \$25,000,000,000 Eigenvector: The Linear Algebra behind Google. *SIAM Review*, **48**(3), 569–581. (Cited on p. 427.)
- [81] Buehrer, G., and Chellapilla, K. 2008. A Scalable Pattern Mining Approach to Web Graph Compression with Communities. Pages 95–106 of: *Proceedings of the 2008 International Conference on Web Search and Data Mining. WSDM '08*. New York, NY: Association for Computing Machinery. (Cited on p. 452.)
- [82] Buluç, A., Meyerhenke, H., Safro, I., Sanders, P., and Schulz, C. 2016. Recent Advances in Graph Partitioning. Pages 117–158 of: Kliemann, L., and Sanders, P. (eds.), *Algorithm Engineering*, vol. 9220. Cham, Switzerland: Springer. (Cited on pp. 409, 427.)
- [83] Buolamwini, J., and Gebru, T. 2018. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. Pages 77–91 of: *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. PMLR. (Cited on p. 28.)
- [84] Burnham, K. P., and Anderson, D. R. 2004. Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, **33**(2), 261–304. (Cited on p. 162.)

- [85] Burt, R. S. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press. (Cited on p. 15.)
- [86] Busetto, L., Wick, W., and Gumbinger, C. 2020. How to Use and Assess Qualitative Research Methods. *Neurological Research and Practice*, **2**(1), 14. (Cited on p. 226.)
- [87] Butts, C. T. 2003. Network Inference, Error, and Informant (in)Accuracy: A Bayesian Approach. *Social Networks*, **25**(2), 103–140. (Cited on p. 134.)
- [88] Butts, C. T. 2009. Revisiting the Foundations of Network Analysis. *Science*, **325**(5939), 414–416. (Cited on p. 86.)
- [89] Callaway, D. S., Hopcroft, J. E., Kleinberg, J. M., Newman, M. E. J., and Strogatz, S. H. 2001. Are Randomly Grown Graphs Really Random? *Physical Review E*, **64**(4), 041902. (Cited on p. 201.)
- [90] Callaway, E. 2022. ‘The Entire Protein Universe’: AI Predicts Shape of Nearly Every Known Protein. *Nature*, **608**(7921), 15–16. (Cited on p. 17.)
- [91] Cameron, A. C., and Trivedi, P. K. 2013. *Regression Analysis of Count Data*. 2nd edn. Econometric Society Monographs. Cambridge; New York, NY: Cambridge University Press. (Cited on pp. 159, 163.)
- [92] Cantwell, G. T., Kirkley, A., and Newman, M. E. J. 2021. The Friendship Paradox in Real and Model Networks. *Journal of Complex Networks*, **9**(2), cnab011. (Cited on pp. 317, 324, 325.)
- [93] Cao, J., Packer, J. S., Ramani, V. et al. 2017. Comprehensive Single-Cell Transcriptional Profiling of a Multicellular Organism. *Science*, **357**(6352), 661–667. (Cited on p. 375.)
- [94] Cao, M., Shu, N., Cao, Q., Wang, Y., and He, Y. 2014. Imaging Functional and Structural Brain Connectomics in Attention-Deficit/Hyperactivity Disorder. *Molecular Neurobiology*, **50**(3), 1111–1123. (Cited on p. 85.)
- [95] Caretta Cartozo, C., and De Los Rios, P. 2009. Extended Navigability of Small World Networks: Exact Results and New Insights. *Physical Review Letters*, **102**(23), 238703. (Cited on p. 344.)
- [96] Carmi, S., Carter, S., Sun, J., and ben-Avraham, D. 2009. Asymptotic Behavior of the Kleinberg Model. *Physical Review Letters*, **102**(23), 238702. (Cited on p. 344.)
- [97] Casey, B. J., Cannonier, T., Conley, M. I. et al. 2018. The Adolescent Brain Cognitive Development (ABCD) Study: Imaging Acquisition across 21 Sites. *Developmental Cognitive Neuroscience*, **32**(Aug.), 43–54. (Cited on pp. 19, 80.)
- [98] Chami, I., Wolf, A., Juan, D.-C., Sala, F., Ravi, S., and Ré, C. 2020. Low-Dimensional Hyperbolic Knowledge Graph Embeddings. Pages 6901–6914 of: *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*. Online: Association for Computational Linguistics. (Cited on p. 438.)
- [99] Chami, I., Abu-El-Haija, S., Perozzi, B., Ré, C., and Murphy, K. 2022. Machine Learning on Graphs: A Model and Comprehensive Taxonomy. *Journal of Machine Learning Research*, **23**(89), 1–64. (Cited on p. 445.)

- [100] Cheeger, J. 1969. A Lower Bound for the Smallest Eigenvalue of the Laplacian. Pages 195–200 of: Gunning, R. C. (ed.), *Problems in Analysis: A Symposium in Honor of Salomon Bochner (PMS-31)*. Princeton, NJ: Princeton University Press. (Cited on p. 427.)
- [101] Chen, B. L., Hall, D. H., and Chklovskii, D. B. 2006. Wiring Optimization Can Relate Neuronal Structure and Function. *Proceedings of the National Academy of Sciences*, **103**(12), 4723–4728. (Cited on p. 227.)
- [102] Chen, M., Mao, S., and Liu, Y. 2014. Big Data: A Survey. *Mobile Networks and Applications*, **19**(2), 171–209. (Cited on p. 469.)
- [103] Ching, A., Edunov, S., Kabiljo, M., Logothetis, D., and Muthukrishnan, S. 2015. One Trillion Edges: Graph Processing at Facebook-Scale. *Proceedings of the VLDB Endowment*, **8**(12), 1804–1815. (Cited on pp. 448, 453.)
- [104] Christen, P. 2012. *Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection*. Data-Centric Systems and Applications. Berlin; New York: Springer. (Cited on p. 116.)
- [105] Chu, C.-C., and Iu, H. H.-C. 2017. Complex Networks Theory For Modern Smart Grid Applications: A Survey. *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, **7**(2), 177–191. (Cited on p. 26.)
- [106] Chung, F. 2007. Four Proofs for the Cheeger Inequality and Graph Partition Algorithms. Page 378 of: *Proceedings of ICCM*, vol. 2. Providence, RI; Boston, MA: American Mathematical Society & International Press of Boston, for Citeseer. (Cited on p. 427.)
- [107] Chung, F., and Lu, L. 2002. Connected Components in Random Graphs with Given Expected Degree Sequences. *Annals of Combinatorics*, **6**(2), 125–145. (Cited on pp. 467, 468.)
- [108] Chung, F. R. K. 1997. *Spectral Graph Theory*. Providence, RI: American Mathematical Society. (Cited on pp. 406, 413, 414, 427.)
- [109] Clauset, A., Shalizi, C. R., and Newman, M. E. J. 2009. Power-Law Distributions in Empirical Data. *SIAM Review*, **51**(4), 661–703. (Cited on pp. 153, 163.)
- [110] Cleveland, W. S., and McGill, R. 1984. Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods. *Journal of the American Statistical Association*, **79**(387), 531–554. (Cited on p. 214.)
- [111] Cohen, R., Erez, K., ben-Avraham, D., and Havlin, S. 2000. Resilience of the Internet to Random Breakdowns. *Physical Review Letters*, **85**(21), 4626–4628. (Cited on pp. 167, 384, 386, 395.)
- [112] Cohen, R., Erez, K., ben-Avraham, D., and Havlin, S. 2001. Breakdown of the Internet under Intentional Attack. *Physical Review Letters*, **86**(16), 3682–3685. (Cited on p. 167.)
- [113] Colizza, V., Flammini, A., Serrano, M. A., and Vespignani, A. 2006. Detecting Rich-Club Ordering in Complex Networks. *Nature Physics*, **2**(2), 110–115. (Cited on p. 202.)
- [114] Cong, J., and Liu, H. 2014. Approaching Human Language with Complex Networks. *Physics of Life Reviews*, **11**(4), 598–618. (Cited on p. 26.)

- [115] Conover, M. D., Gonçalves, B., Flammini, A., and Menczer, F. 2012. Partisan Asymmetries in Online Political Activity. *EPJ Data Science*, **1**(1), 6. (Cited on p. 84.)
- [116] Conte, D., Foggia, P., Sansone, C., and Vento, M. 2004. Thirty Years of Graph Matching in Pattern Recognition. *International Journal of Pattern Recognition and Artificial Intelligence*, **18**(03), 265–298. (Cited on p. 232.)
- [117] Cormen, T. H., Leiserson, C. E., Rivest, R. L., and Stein, C. 2022. *Introduction to Algorithms*. 4th edn. Cambridge, MA: MIT Press. (Cited on p. 61.)
- [118] Cutler, A. 1991. Nested Faunas and Extinction in Fragmented Habitats. *Conservation Biology*, **5**(4), 496–504. (Cited on p. 202.)
- [119] Dai, H., Li, H., Tian, T. et al. 2018. Adversarial Attack on Graph Structured Data. Pages 1115–1124 of: *Proceedings of the 35th International Conference on Machine Learning*. PMLR. (Cited on p. 444.)
- [120] Dalton, J. 1808. *A New System of Chemical Philosophy. Part I*. Cambridge: Cambridge University Press. (Cited on p. 4.)
- [121] Dantzig, G. B., and Fulkerson, D. R. 1955. *On the Max Flow Min Cut Theorem of Networks*. Santa Monica, CA: RAND corporation. (Cited on pp. 201, 427.)
- [122] Dawid, A. P., and Skene, A. M. 1979. Maximum Likelihood Estimation of Observer Error-Rates Using the EM Algorithm. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **28**(1), 20–28. (Cited on p. 374.)
- [123] de Saussure, F. 1959. *Course in General Linguistics*. New York, NY: Philosophical Library. (Cited on p. 432.)
- [124] Dean, J., and Ghemawat, S. 2008. MapReduce: Simplified Data Processing on Large Clusters. *Communications of the ACM*, **51**(1), 107–113. (Cited on pp. 452, 453.)
- [125] Decelle, A., Krzakala, F., Moore, C., and Zdeborová, L. 2011. Asymptotic Analysis of the Stochastic Block Model for Modular Networks and Its Algorithmic Applications. *Physical Review E*, **84**(6), 066106. (Cited on pp. 356, 357.)
- [126] Decelle, A., Krzakala, F., Moore, C., and Zdeborová, L. 2011. Inference and Phase Transitions in the Detection of Modules in Sparse Networks. *Physical Review Letters*, **107**(6), 065701. (Cited on pp. 356, 357.)
- [127] Dempster, A. P., Laird, N. M., and Rubin, D. B. 1977. Maximum Likelihood from Incomplete Data Via the EM Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, **39**(1), 1–22. (Cited on pp. 134, 360, 362.)
- [128] Dessimoz, C., and Škunca, N. (eds.). 2017. *The Gene Ontology Handbook. Methods in Molecular Biology*, vol. 1446. New York, NY: Springer New York. (Cited on p. 116.)
- [129] Devlin, J., Chang, M.-W., Lee, K., and Toutanova, K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Pages 4171–4186 of: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. Minneapolis, MN: Association for Computational Linguistics. (Cited on p. 261.)

- [130] Devroye, L. 1986. *Non-Uniform Random Variate Generation*. New York, NY: Springer. (Cited on p. 466.)
- [131] Dodds, P. S., Muhamad, R., and Watts, D. J. 2003. An Experimental Study of Search in Global Social Networks. *Science*, **301**(5634), 827–829. (Cited on p. 341.)
- [132] Donath, W. E., and Hoffman, A. J. 1973. Lower Bounds for the Partitioning of Graphs. *IBM Journal of Research and Development*, **17**(5), 420–425. (Cited on pp. 427, 428.)
- [133] Dong, X., Thanou, D., Rabbat, M., and Frossard, P. 2019. Learning Graphs From Data: A Signal Representation Perspective. *IEEE Signal Processing Magazine*, **36**(3), 44–63. (Cited on p. 238.)
- [134] Doolittle, W. F. 2013. Is Junk DNA Bunk? A Critique of ENCODE. *Proceedings of the National Academy of Sciences*, **110**(14), 5294–5300. (Cited on p. 122.)
- [135] Dunbar, R. I. M. 1993. Coevolution of Neocortical Size, Group Size and Language in Humans. *Behavioral and Brain Sciences*, **16**(4), 681–694. (Cited on p. 181.)
- [136] Dunning, T. 2012. *Natural Experiments in the Social Sciences: A Design-Based Approach*. Cambridge: Cambridge University Press. (Cited on p. 474.)
- [137] Durand, M., and Flajolet, P. 2003. Loglog Counting of Large Cardinalities. Pages 605–617 of: Di Battista, G., and Zwick, U. (eds.), *Algorithms—ESA 2003*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. (Cited on pp. 455, 469.)
- [138] Dwork, C., and Roth, A. 2014. The Algorithmic Foundations of Differential Privacy. *Foundations and Trends® in Theoretical Computer Science*, **9**(3–4), 211–407. (Cited on p. 475.)
- [139] Eades, P. 1984. A Heuristic for Graph Drawing. *Congressus Numerantium*, **vol.42**, 149–160. (Cited on p. 221.)
- [140] Easley, D., and Kleinberg, J. 2010. *Networks, Crowds, and Markets: Reasoning about a Highly Connected World*. Cambridge: Cambridge University Press. (Cited on p. 472.)
- [141] Edwards, P. N., Mayernik, M. S., Batcheller, A. L., Bowker, G. C., and Borgman, C. L. 2011. Science Friction: Data, Metadata, and Collaboration. *Social Studies of Science*, **41**(5), 667–690. (Cited on p. 292.)
- [142] Efron, B., and Hastie, T. 2016. *Computer Age Statistical Inference*. Cambridge: Cambridge University Press. (Cited on pp. 62, 277, 472.)
- [143] Ehrlich, B. 2022. *The Brain in Search of Itself: Santiago Ramón y Cajal and the Story of the Neuron*. New York, NY: Farrar, Straus and Giroux. (Cited on p. 16.)
- [144] Einstein, A. 1905. Über Die von Der Molekularkinetischen Theorie Der Wärme Geforderte Bewegung von in Ruhenden Flüssigkeiten Suspendierten Teilchen. *Annalen der Physik*, **322**(8), 549–560. (Cited on p. 327.)
- [145] Ellson, J., Gansner, E., Koutsofios, L., North, S. C., and Woodhull, G. 2002. Graphviz—Open Source Graph Drawing Tools. Pages 483–484 of: Mutzel, P.,

- Jünger, M., and Leipert, S. (eds.), *Graph Drawing*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. (Cited on p. 206.)
- [146] Emmons, S. W. 2015. The Beginning of Connectomics: A Commentary on White et al. (1986) 'The Structure of the Nervous System of the Nematode *Caenorhabditis Elegans*'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **370**(1666), 20140309. (Cited on p. 82.)
- [147] Eom, Y.-H., and Jo, H.-H. 2015. Generalized Friendship Paradox in Complex Networks: The Case of Scientific Collaboration. *Scientific Reports*, **4**(1), 4603. (Cited on p. 324.)
- [148] Erdős, P., and Rényi, A. 1959. On Random Graphs I. *Publicationes Mathematicae*, **6**(1), 290–297. (Cited on pp. 163, 349.)
- [149] Erdős, P., and Rényi, A. 1960. On the Evolution of Random Graphs. *Publ. Math. Inst. Hung. Acad. Sci.*, **5**(1), 17–60. (Cited on pp. 163, 349.)
- [150] Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Pages 226–231 of: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. KDD'96, vol. 96. Portland, OR: AAAI Press. (Cited on p. 428.)
- [151] Eswaran, D., Faloutsos, C., Guha, S., and Mishra, N. 2018. SpotLight: Detecting Anomalies in Streaming Graphs. Pages 1378–1386 of: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. KDD '18. New York, NY: Association for Computing Machinery. (Cited on p. 469.)
- [152] Euzenat, J., and Shvaiko, P. 2013. *Ontology Matching*. 2nd edn. Berlin, Heidelberg: Springer. (Cited on p. 116.)
- [153] Evans, T. S., and Lambiotte, R. 2009. Line Graphs, Link Partitions, and Overlapping Communities. *Physical Review E*, **80**(1), 016105. (Cited on pp. 184, 201.)
- [154] Evtushenko, A., and Kleinberg, J. 2023. Node-Based Generalized Friendship Paradox Fails. *Scientific Reports*, **13**(1), 2074. (Cited on p. 324.)
- [155] Fan, J., Han, F., and Liu, H. 2014. Challenges of Big Data Analysis. *National Science Review*, **1**(2), 293–314. (Cited on p. 469.)
- [156] Feld, S. L. 1991. Why Your Friends Have More Friends Than You Do. *American Journal of Sociology*, **96**(6), 1464–1477. (Cited on p. 324.)
- [157] Ferguson, N. 2019. *The Square and the Tower: Networks and Power, from the Freemasons to Facebook*. New York, NY: Penguin. (Cited on p. 16.)
- [158] Fiedler, M. 1973. Algebraic Connectivity of Graphs. *Czechoslovak Mathematical Journal*, **23**(2), 298–305. (Cited on pp. 412, 427, 428.)
- [159] Firth, J. R. 1957. *Studies in Linguistic Analysis*. Publications of the Philological Society. Oxford: Blackwell. (Cited on pp. 272, 432.)
- [160] Flajolet, P., and Nigel Martin, G. 1985. Probabilistic Counting Algorithms for Data Base Applications. *Journal of Computer and System Sciences*, **31**(2), 182–209. (Cited on pp. 455, 469.)

- [161] Flajolet, P., Fusy, É., Gandouet, O., and Meunier, F. 2007. HyperLogLog: The Analysis of a near-Optimal Cardinality Estimation Algorithm. Pages 137–156 of: *Discrete Mathematics and Theoretical Computer Science*, vol. AH. Nancy, France: Discrete Mathematics and Theoretical Computer Science. (Cited on pp. 455, 456, 469.)
- [162] Ford, L. R., and Fulkerson, D. R. 1956. Maximal Flow Through a Network. *Canadian Journal of Mathematics*, **8**, 399–404. (Cited on p. 201.)
- [163] Ford, L. R., and Fulkerson, D. R. 1962. *Flows in Networks*. Princeton, NJ: Princeton University Press. (Cited on p. 201.)
- [164] Fortunato, S. 2010. Community Detection in Graphs. *Physics Reports*, **486**(3), 75–174. (Cited on p. 201.)
- [165] Fortunato, S., and Barthélemy, M. 2007. Resolution Limit in Community Detection. *Proceedings of the National Academy of Sciences*, **104**(1), 36–41. (Cited on p. 201.)
- [166] Fortunato, S., and Newman, M. E. J. 2022. 20 Years of Network Community Detection. *Nature Physics*, **18**(8), 848–850. (Cited on p. 201.)
- [167] Fosdick, B. K., Larremore, D. B., Nishimura, J., and Ugander, J. 2018. Configuring Random Graph Models with Fixed Degree Sequences. *SIAM Review*, **60**(2), 315–355. (Cited on p. 163.)
- [168] Frank, O., and Strauss, D. 1986. Markov Graphs. *Journal of the American Statistical Association*, **81**(395), 832–842. (Cited on pp. 163, 367.)
- [169] Frantz, T. L., Cataldo, M., and Carley, K. M. 2009. Robustness of Centrality Measures under Uncertainty: Examining the Role of Network Topology. *Computational and Mathematical Organization Theory*, **15**(4), 303–328. (Cited on pp. 377, 395.)
- [170] Friedman, J., Hastie, T., and Tibshirani, R. 2008. Sparse Inverse Covariance Estimation with the Graphical Lasso. *Biostatistics*, **9**(3), 432–441. (Cited on pp. 373, 374.)
- [171] Fruchterman, T. M. J., and Reingold, E. M. 1991. Graph Drawing by Force-Directed Placement. *Software: Practice and Experience*, **21**(11), 1129–1164. (Cited on p. 221.)
- [172] Gao, X., Xiao, B., Tao, D., and Li, X. 2010. A Survey of Graph Edit Distance. *Pattern Analysis and Applications*, **13**(1), 113–129. (Cited on p. 232.)
- [173] Garcia, D. 2017. Leaking Privacy and Shadow Profiles in Online Social Networks. *Science Advances*, **3**(8), e1701172. (Cited on p. 37.)
- [174] Garey, M., Johnson, D., and Stockmeyer, L. 1976. Some Simplified NP-complete Graph Problems. *Theoretical Computer Science*, **1**(3), 237–267. (Cited on p. 410.)
- [175] Gauvin, L., Génois, M., Karsai, M. et al. 2022. Randomized Reference Models for Temporal Networks. *SIAM Review*, **64**(4), 763–830. (Cited on p. 243.)
- [176] Gawande, A. 2010. *The Checklist Manifesto: How to Get Things Right*. New York, NY: Metropolitan Books. (Cited on p. 285.)

- [177] Gebru, T., Morgenstern, J., Vecchione, B. et al. 2020. Datasheets for Datasets. *arXiv:1803.09010 [cs]*, Mar. (Cited on pp. 31, 474.)
- [178] Gelman, A., and Hill, J. 2007. *Data Analysis Using Regression and Multi-level/Hierarchical Models*. Analytical Methods for Social Research. Cambridge; New York: Cambridge University Press. (Cited on pp. 116, 134.)
- [179] Gelman, A., Carlin, Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. 2014. *Bayesian Data Analysis*. 3rd edn. Chapman & Hall/CRC Texts in Statistical Science. Boca Raton, FL: CRC Press. (Cited on pp. 62, 375, 472.)
- [180] Gertner, J. 2013. *The Idea Factory: Bell Labs and the Great Age of American Innovation*. London: Penguin Books. (Cited on p. 313.)
- [181] Ghosh, S., Das, N., Gonçalves, T., Quaresma, P., and Kundu, M. 2018. The Journey of Graph Kernels through Two Decades. *Computer Science Review*, **27**(Feb.), 88–111. (Cited on p. 232.)
- [182] Gilbert, E. N. 1959. Random Graphs. *The Annals of Mathematical Statistics*, **30**(4), 1141–1144. (Cited on pp. 163, 349.)
- [183] Girvan, M., and Newman, M. E. J. 2002. Community Structure in Social and Biological Networks. *Proceedings of the National Academy of Sciences*, **99**(12), 7821–7826. (Cited on pp. 185, 201.)
- [184] Golub, G. H., and Van Loan, C. F. 2013. *Matrix Computations*. 4th edn. Johns Hopkins Studies in the Mathematical Sciences. Baltimore, MD: The Johns Hopkins University Press. (Cited on pp. 400, 427.)
- [185] Gonzalez, J. E., Low, Y., Gu, H., Bickson, D., and Guestrin, C. 2012. Power-Graph: Distributed Graph-Parallel Computation on Natural Graphs. Pages 17–30 of: *Proceedings of the 10th USENIX Conference on Operating Systems Design and Implementation*. OSDI'12. Hollywood, CA: USENIX Association. (Cited on p. 453.)
- [186] Goodfellow, I., Bengio, Y., and Courville, A. 2016. *Deep Learning*. Cambridge, MA: MIT Press. (Cited on pp. 277, 472.)
- [187] Gower, J. C. 1975. Generalized Procrustes Analysis. *Psychometrika*, **40**(1), 33–51. (Cited on p. 371.)
- [188] Goyal, P., and Ferrara, E. 2018. Graph Embedding Techniques, Applications, and Performance: A Survey. *Knowledge-Based Systems*, **151**(July), 78–94. (Cited on p. 445.)
- [189] Granovetter, M. 1976. Network Sampling: Some First Steps. *American Journal of Sociology*, **81**(6), 1287–1303. (Cited on p. 395.)
- [190] Granovetter, M. S. 1973. The Strength of Weak Ties. *American Journal of Sociology*, **78**(6), 1360–1380. (Cited on pp. 21, 343.)
- [191] Gray, J., Liu, D. T., Nieto-Santisteban, M., Szalay, A., DeWitt, D. J., and Heber, G. 2005. Scientific Data Management in the Coming Decade. *ACM SIGMOD Record*, **34**(4), 34–41. (Cited on p. 292.)
- [192] Grohe, M., and Schweitzer, P. 2020. The Graph Isomorphism Problem. *Communications of the ACM*, **63**(11), 128–134. (Cited on p. 445.)

- [193] Grover, A., and Leskovec, J. 2016. Node2vec: Scalable Feature Learning for Networks. Pages 855–864 of: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. New York, NY: Association for Computing Machinery. (Cited on p. 435.)
- [194] Groves, R. M. 2011. Three Eras of Survey Research. *Public Opinion Quarterly*, **75**(5), 861–871. (Cited on pp. 65, 68, 71.)
- [195] Guimerà, R., and Sales-Pardo, M. 2009. Missing and Spurious Interactions and the Reconstruction of Complex Networks. *Proceedings of the National Academy of Sciences*, **106**(52), 22073–22078. (Cited on p. 134.)
- [196] Guimerà, R., Sales-Pardo, M., and Amaral, L. A. N. 2004. Modularity from Fluctuations in Random Graphs and Complex Networks. *Physical Review E*, **70**(2), 025101. (Cited on p. 179.)
- [197] Gutmann, M., and Hyvärinen, A. 2010. Noise-Contrastive Estimation: A New Estimation Principle for Unnormalized Statistical Models. In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. Proceedings of Machine Learning Research. Sardinia, Italy: PMLR. (Cited on p. 435.)
- [198] Hagberg, A., and Lemons, N. 2015. Fast Generation of Sparse Random Kernel Graphs. *PLOS ONE*, **10**(9), e0135177. (Cited on p. 469.)
- [199] Hall, K. M. 1970. An R-Dimensional Quadratic Placement Algorithm. *Management Science*, **17**(3), 219–229. (Cited on p. 221.)
- [200] Hamilton, W., Ying, Z., and Leskovec, J. 2017. Inductive Representation Learning on Large Graphs. In: Guyon, I., Luxburg, U. V., Bengio, S. et al. (eds.), *Advances in Neural Information Processing Systems*, vol. 30. Long Beach, CA: Curran Associates, Inc. (Cited on pp. 273, 442.)
- [201] Hanel, R., Corominas-Murtra, B., Liu, B., and Thurner, S. 2017. Fitting Power-Laws in Empirical Data with Estimators That Work for All Exponents. *PLOS ONE*, **12**(2), e0170920. (Cited on p. 163.)
- [202] Harris, Z. S. 1954. Distributional Structure. *WORD*, **10**(2-3), 146–162. (Cited on pp. 272, 432.)
- [203] Hartle, H., Klein, B., McCabe, S. et al. 2020. Network Comparison and the Within-Ensemble Graph Distance. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, **476**(2243), 20190744. (Cited on p. 232.)
- [204] Hashimoto, K.-i. 1989. Zeta Functions of Finite Graphs and Representations of P-Adic Groups. Pages 211–280 of: *Automorphic Forms and Geometry of Arithmetic Varieties*. Advanced Studies in Pure Mathematics, vol. 15. Cambridge, MA: Academic Press. (Cited on pp. 357, 403.)
- [205] Hastie, T., Tibshirani, R., and Friedman, J. H. 2009. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd edn. Springer Series in Statistics. New York, NY: Springer. (Cited on pp. 277, 428, 472.)
- [206] Hastie, T., Tibshirani, R., and Wainwright, M. 2015. *Statistical Learning with*

- Sparsity: The Lasso and Generalizations*. New York, NY: Taylor & Francis. (Cited on pp. 277, 375.)
- [207] Head, M. L., Holman, L., Lanfear, R., Kahn, A. T., and Jennions, M. D. 2015. The Extent and Consequences of P-Hacking in Science. *PLOS Biology*, **13**(3), e1002106. (Cited on pp. 35, 71.)
- [208] Healy, K. 2018. *Data Visualization: A Practical Introduction*. Princeton, NJ: Princeton University Press. (Cited on p. 473.)
- [209] Heinrich, J. 2001. *Drug Safety: Most Drugs Withdrawn in Recent Years Had Greater Health Risks for Women*. Tech. rept. GAO-01-286R. US Government Printing Office, Washington, DC. (Cited on p. 28.)
- [210] Hellman, M. 2002. An Overview of Public Key Cryptography. *IEEE Communications Magazine*, **40**(5), 42–49. (Cited on p. 313.)
- [211] Hennig, C., Meila, M., Murtagh, F., and Rocci, R. (eds.). 2016. *Handbook of Cluster Analysis*. Chapman & Hall/CRC Handbooks of Modern Statistical Methods. Boca Raton, FL; London; New York: CRC Press, Taylor & Francis Group. (Cited on p. 428.)
- [212] Heule, S., Nunkesser, M., and Hall, A. 2013. HyperLogLog in Practice: Algorithmic Engineering of a State of the Art Cardinality Estimation Algorithm. Pages 683–692 of: *Proceedings of the 16th International Conference on Extending Database Technology*. EDBT '13. New York, NY: Association for Computing Machinery. (Cited on p. 469.)
- [213] Hočevár, T., and Demšar, J. 2014. A Combinatorial Approach to Graphlet Counting. *Bioinformatics*, **30**(4), 559–565. (Cited on p. 175.)
- [214] Hodas, N., Kooti, F., and Lerman, K. 2013. Friendship Paradox Redux: Your Friends Are More Interesting Than You. *Proceedings of the International AAAI Conference on Web and Social Media*, **7**(1), 225–233. (Cited on p. 324.)
- [215] Hofer, M., Hellmann, S., Dojchinovski, M., and Frey, J. 2020. The New DBpedia Release Cycle: Increasing Agility and Efficiency in Knowledge Extraction Workflows. Pages 1–18 of: Blomqvist, E., Groth, P., de Boer, V. et al. (eds.), *Semantic Systems. In the Era of Knowledge Graphs*. Lecture Notes in Computer Science. Cham, Switzerland: Springer. (Cited on p. 449.)
- [216] Hoff, P. D., Raftery, A. E., and Handcock, M. S. 2002. Latent Space Approaches to Social Network Analysis. *Journal of the American Statistical Association*, **97**(460), 1090–1098. (Cited on pp. 369, 370, 371, 374.)
- [217] Hofman, J. M., Watts, D. J., Athey, S. et al. 2021. Integrating Explanation and Prediction in Computational Social Science. *Nature*, **595**(7866), 181–188. (Cited on p. 474.)
- [218] Holland, P. W., and Leinhardt, S. 1981. An Exponential Family of Probability Distributions for Directed Graphs. *Journal of the American Statistical Association*, **76**(373), 33–50. (Cited on p. 163.)
- [219] Holland, P. W., Laskey, K. B., and Leinhardt, S. 1983. Stochastic Blockmodels: First Steps. *Social Networks*, **5**(2), 109–137. (Cited on pp. 163, 374.)

- [220] Holme, P. 2019. Rare and Everywhere: Perspectives on Scale-Free Networks. *Nature Communications*, **10**(1), 1016. (Cited on p. 346.)
- [221] Holme, P., and Saramäki, J. 2012. Temporal Networks. *Physics Reports*, **519**(3), 97–125. (Cited on pp. 249, 472.)
- [222] Holme, P., and Saramäki, J. (eds.). 2019. *Temporal Network Theory*. Computational Social Science. Cham, Switzerland: Springer Nature. (Cited on p. 249.)
- [223] Holten, D., and Van Wijk, J. J. 2009. Force-Directed Edge Bundling for Graph Visualization. *Computer Graphics Forum*, **28**(3), 983–990. (Cited on p. 213.)
- [224] Hosmer, D. W., Lemeshow, S., and Sturdivant, R. X. 2013. *Applied Logistic Regression*. Hoboken, NJ: John Wiley & Sons. (Cited on p. 256.)
- [225] Howe, J., et al. 2006. The Rise of Crowdsourcing. *Wired magazine*, **14**(6), 1–4. (Cited on p. 374.)
- [226] Humphreys, I. R., Pei, J., Baek, M. et al. 2021. Computed Structures of Core Eukaryotic Protein Complexes. *Science*, **374**(6573), eabm4805. (Cited on p. 17.)
- [227] Hyndman, R. J., and Athanasopoulos, G. 2021. *Forecasting: Principles and Practice*. 3rd edn. Melbourne, Australia: Otexts, Online Open-Access Textbooks. (Cited on p. 249.)
- [228] Israel, M., and Hay, I. 2006. *Research Ethics for Social Scientists*. Thousand Oaks, CA: SAGE. (Cited on p. 37.)
- [229] Jackson, M. O. 2019. The Friendship Paradox and Systematic Biases in Perceptions and Social Norms. *Journal of Political Economy*, **127**(2), 777–818. (Cited on p. 325.)
- [230] James, G., Witten, D., Hastie, T., and Tibshirani, R. 2021. *An Introduction to Statistical Learning: With Applications in R*. 2nd edn. Springer Texts in Statistics. New York, NY: Springer. (Cited on pp. 277, 472.)
- [231] Jha, M., Seshadhri, C., and Pinar, A. 2015. A Space-Efficient Streaming Algorithm for Estimating Transitivity and Triangle Counts Using the Birthday Paradox. *ACM Transactions on Knowledge Discovery from Data*, **9**(3), 15:1–15:21. (Cited on p. 454.)
- [232] Kanare, H. M. 1985. *Writing the Laboratory Notebook*. Washington, DC: American Chemical Society. (Cited on pp. 283, 287.)
- [233] Kanungo, T., Mount, D., Netanyahu, N., Piatko, C., Silverman, R., and Wu, A. 2002. An Efficient K-Means Clustering Algorithm: Analysis and Implementation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**(7), 881–892. (Cited on p. 426.)
- [234] Karger, D., Oh, S., and Shah, D. 2011. Iterative Learning for Reliable Crowdsourcing Systems. In: *Advances in Neural Information Processing Systems*, vol. 24. Granada, Spain: Curran Associates, Inc. (Cited on p. 374.)
- [235] Karinthy, F. 1929. Chain-Links. *Everything is different*, 21–26. (Cited on p. 15.)
- [236] Karrer, B., and Newman, M. E. J. 2011. Stochastic Blockmodels and Community Structure in Networks. *Physical Review E*, **83**(1), 016107. (Cited on p. 355.)
- [237] Katz, L. 1953. A New Status Index Derived from Sociometric Analysis. *Psychometrika*, **18**(1), 39–43. (Cited on pp. 189, 408, 427.)

- [238] Kaufman, L., and Rousseeuw, P. J. 2009. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York, NY: John Wiley & Sons. (Cited on pp. 232, 428.)
- [239] Kernighan, B. W., and Lin, S. 1970. An Efficient Heuristic Procedure for Partitioning Graphs. *The Bell System Technical Journal*, **49**(2), 291–307. (Cited on pp. 201, 427.)
- [240] Kernighan, B. W. 2020. *UNIX: A History and a Memoir*. Seattle, WA: Kindle Direct Publishing. (Cited on p. 313.)
- [241] Kessler, M. D., Yerges-Armstrong, L., Taub, M. A. et al. 2016. Challenges and Disparities in the Application of Personalized Genomic Medicine to Populations with African Ancestry. *Nature Communications*, **7**(1), 12521. (Cited on p. 28.)
- [242] Kipf, T. N., and Welling, M. 2017. Semi-Supervised Classification with Graph Convolutional Networks. In: *International Conference on Learning Representations*. Toulon, France: OpenReview.net. (Cited on pp. 273, 442.)
- [243] Kirchgässner, G., Wolters, J., and Hassler, U. 2013. *Introduction to Modern Time Series Analysis*. 2nd edn. Springer Texts in Business and Economics. Berlin, Heidelberg: Springer. (Cited on p. 249.)
- [244] Kleinberg, J. M. 1999. Authoritative Sources in a Hyperlinked Environment. *Journal of the ACM*, **46**(5), 604–632. (Cited on p. 189.)
- [245] Kleinberg, J. M. 2000. Navigation in a Small World. *Nature*, **406**(6798), 845–845. (Cited on p. 344.)
- [246] Kleinberg, J. M., Kumar, R., Raghavan, P., Rajagopalan, S., and Tomkins, A. S. 1999. The Web as a Graph: Measurements, Models, and Methods. Pages 1–17 of: Goos, G., Hartmanis, J., van Leeuwen, J. et al. (eds.), *Computing and Combinatorics*, vol. 1627. Berlin, Heidelberg: Springer. (Cited on pp. 346, 349.)
- [247] Knuth, D. E. 1984. Literate Programming. *The Computer Journal*, **27**(2), 97–111. (Cited on p. 301.)
- [248] Kojaku, S., and Masuda, N. 2017. Finding Multiple Core-Periphery Pairs in Networks. *Physical Review E*, **96**(5), 052313. (Cited on p. 195.)
- [249] Kojaku, S., Yoon, J., Constantino, I., and Ahn, Y.-Y. 2021. Residual2Vec: Debiassing Graph Embedding with Random Graphs. Pages 24150–24163 of: *Advances in Neural Information Processing Systems*, vol. 34. Online: Curran Associates, Inc. (Cited on pp. 435, 445.)
- [250] Koponen, J., and Hildén, J. 2019. *Data Visualization Handbook*. Aalto, Finland: Aalto University. (Cited on p. 473.)
- [251] Kosinski, M., Stillwell, D., and Graepel, T. 2013. Private Traits and Attributes Are Predictable from Digital Records of Human Behavior. *Proceedings of the National Academy of Sciences*, **110**(15), 5802–5805. (Cited on p. 32.)
- [252] Köster, J., and Rahmann, S. 2012. Snakemake—a Scalable Bioinformatics Workflow Engine. *Bioinformatics*, **28**(19), 2520–2522. (Cited on p. 302.)
- [253] Krapivsky, P. L., and Redner, S. 2001. Organization of Growing Random Networks. *Physical Review E*, **63**(6), 066123. (Cited on pp. 345, 346, 348, 349.)

- [254] Krioukov, D., Papadopoulos, F., Kitsak, M., Vahdat, A., and Boguñá, M. 2010. Hyperbolic Geometry of Complex Networks. *Physical Review E*, **82**(3), 036106. (Cited on p. 438.)
- [255] Kruja, E., Marks, J., Blair, A., and Waters, R. 2002. A Short Note on the History of Graph Drawing. Pages 272–286 of: Mutzel, P., Jünger, M., and Leipert, S. (eds.), *Graph Drawing*. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. (Cited on p. 221.)
- [256] Kruschke, J. K. 2015. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. 2nd edn. Boston, MA: Academic Press. (Cited on pp. 62, 472.)
- [257] Kruskal, J. B., and Wish, M. 1978. *Multidimensional Scaling*. Sage University Papers Quantitative Applications in the Social Sciences, no. 11. Newbury Park, CA: SAGE. (Cited on p. 370.)
- [258] Kryshchavych, A., Schwede, T., Topf, M., Fidelis, K., and Moul, J. 2021. Critical Assessment of Methods of Protein Structure Prediction (CASP)—Round XIV. *Proteins: Structure, Function, and Bioinformatics*, **89**(12), 1607–1617. (Cited on p. 17.)
- [259] Krzakala, F., Moore, C., Mossel, E. et al. 2013. Spectral Redemption in Clustering Sparse Networks. *Proceedings of the National Academy of Sciences*, **110**(52), 20935–20940. (Cited on pp. 356, 357, 421.)
- [260] Kumar, S., Morstatter, F., and Liu, H. 2014. *Twitter Data Analytics*. Springer-Briefs in Computer Science. New York, NY: Springer. (Cited on p. 71.)
- [261] Landauer, T. K., Foltz, P. W., and Laham, D. 1998. An Introduction to Latent Semantic Analysis. *Discourse Processes*, **25**(2-3), 259–284. (Cited on pp. 438, 445.)
- [262] Lang, T. A., and Altman, D. G. 2015. Basic Statistical Reporting for Articles Published in Biomedical Journals: The “Statistical Analyses and Methods in the Published Literature” or the SAMPL Guidelines. *International Journal of Nursing Studies*, **52**(1), 5–9. (Cited on p. 163.)
- [263] Latapy, M., Viard, T., and Magnien, C. 2018. Stream Graphs and Link Streams for the Modeling of Interactions over Time. *Social Network Analysis and Mining*, **8**(1), 61. (Cited on p. 240.)
- [264] Lazer, D., Pentland, A., Adamic, L. et al. 2009. Computational Social Science. *Science*, **323**(5915), 721–723. (Cited on p. 25.)
- [265] Lazer, D. M. J., Pentland, A., Watts, D. J. et al. 2020. Computational Social Science: Obstacles and Opportunities. *Science*, **369**(6507), 1060–1062. (Cited on p. 474.)
- [266] Lee, C., and Wilkinson, D. J. 2019. A Review of Stochastic Block Models and Extensions for Graph Clustering. *Applied Network Science*, **4**(1), 1–50. (Cited on pp. 353, 374.)
- [267] Lee, J. R., Gharan, S. O., and Trevisan, L. 2014. Multiway Spectral Partitioning and Higher-Order Cheeger Inequalities. *Journal of the ACM*, **61**(6), 1–30. (Cited on pp. 417, 427.)

- [268] Lehmann, S. 2019. Fundamental Structures in Temporal Communication Networks. Pages 25–48 of: Holme, P., and Saramäki, J. (eds.), *Temporal Network Theory*. Computational Social Sciences. Cham, Switzerland: Springer. (Cited on p. 473.)
- [269] Levy, O., and Goldberg, Y. 2014. Neural Word Embedding as Implicit Matrix Factorization. In: *Proceedings of the 27th International Conference on Neural Information Processing Systems*, vol. 2. Montreal, Canada: Curran Associates, Inc. (Cited on pp. 272, 438, 439, 440, 445.)
- [270] Levy, O., Goldberg, Y., and Dagan, I. 2015. Improving Distributional Similarity with Lessons Learned from Word Embeddings. *Transactions of the Association for Computational Linguistics*, **3**(Dec.), 211–225. (Cited on p. 272.)
- [271] Lewandowsky, S., Mann, M. E., Bauld, L., Hastings, G., and Loftus, a. E. F. 2013. The Subterranean War on Science. *APS Observer*, **26**. (Cited on p. 284.)
- [272] Lewis, S. 2019. The Racial Bias Built into Photography. *The New York Times*, Apr. (Cited on p. 29.)
- [273] Liben-Nowell, D., and Kleinberg, J. 2007. The Link-Prediction Problem for Social Networks. *Journal of the American Society for Information Science and Technology*, **58**(7), 1019–1031. (Cited on p. 134.)
- [274] Liggett, T. M. 1999. *Stochastic Interacting Systems: Contact, Voter, and Exclusion Processes*. Grundlehren Der Mathematischen Wissenschaften, no. 324. Berlin; New York: Springer. (Cited on p. 235.)
- [275] Liggett, T. M. 2005. *Interacting Particle Systems*. Classics in Mathematics. Berlin; New York: Springer. (Cited on p. 235.)
- [276] Lin, J. W.-B., Aizenman, H., Espinel, E. M. C., Gunnerson, K. N., and Liu, J. 2022. *An Introduction to Python Programming for Scientists and Engineers*. Cambridge; New York, NY: Cambridge University Press. (Cited on p. 61.)
- [277] Lin, T., Wang, Y., Liu, X., and Qiu, X. 2022. A Survey of Transformers. *AI Open*, **3**(Jan.), 111–132. (Cited on p. 433.)
- [278] Liu, Y., Safavi, T., Dighe, A., and Koutra, D. 2018. Graph Summarization Methods and Applications: A Survey. *ACM Computing Surveys*, **51**(3), 62:1–62:34. (Cited on pp. 232, 452, 469.)
- [279] Louis, A., Raghavendra, P., Tetali, P., and Vempala, S. 2012. Many Sparse Cuts via Higher Eigenvalues. Pages 1131–1140 of: *Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing*. New York, NY: Association for Computing Machinery. (Cited on pp. 417, 427.)
- [280] Low, Y., Gonzalez, J., Kyrola, A., Bickson, D., Guestrin, C., and Hellerstein, J. 2010. GraphLab: A New Framework for Parallel Machine Learning. Pages 340–349 of: *Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence*. UAI'10. Arlington, VA: AUAI Press. (Cited on p. 453.)
- [281] Low, Y., Bickson, D., Gonzalez, J., Guestrin, C., Kyrola, A., and Hellerstein, J. M. 2012. Distributed GraphLab: A Framework for Machine Learning and Data Mining in the Cloud. *Proceedings of the VLDB Endowment*, **5**(8), 716–727. (Cited on p. 453.)

- [282] Lü, L., and Zhou, T. 2011. Link Prediction in Complex Networks: A Survey. *Physica A: Statistical Mechanics and its Applications*, **390**(6), 1150–1170. (Cited on p. 134.)
- [283] Luck, K., Kim, D.-K., Lambourne, L. et al. 2020. A Reference Map of the Human Binary Protein Interactome. *Nature*, **580**(7803), 402–408. (Cited on pp. 24, 84, 130, 389, 390.)
- [284] Lusher, D., Koskinen, J., and Robins, G. (eds.). 2013. *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications*. Structural Analysis in the Social Sciences, no. 32. Cambridge: Cambridge University Press. (Cited on p. 163.)
- [285] Macleod, M., Collings, A. M., Graf, C. et al. 2021. The MDAR (Materials Design Analysis Reporting) Framework for Transparent Reporting in the Life Sciences. *Proceedings of the National Academy of Sciences*, **118**(17), e2103238118. (Cited on p. 163.)
- [286] Madden, S. 2012. From Databases to Big Data. *IEEE Internet Computing*, **16**(3), 4–6. (Cited on p. 469.)
- [287] Makalowski, W. 2003. Not Junk After All. *Science*, **300**(5623), 1246–1247. (Cited on p. 122.)
- [288] Makin, T. R., and Orban de Xivry, J.-J. 2019. Ten Common Statistical Mistakes to Watch out for When Writing or Reviewing a Manuscript. *eLife*, **8**(Oct.), e48175. (Cited on p. 163.)
- [289] Malewicz, G., Austern, M. H., Bik, A. J. et al. 2010. Pregel: A System for Large-Scale Graph Processing. Pages 135–146 of: *Proceedings of the 2010 ACM SIGMOD International Conference on Management of Data*. SIGMOD '10. New York, NY: Association for Computing Machinery. (Cited on p. 453.)
- [290] Manzoor, E., Milajerdi, S. M., and Akoglu, L. 2016. Fast Memory-efficient Anomaly Detection in Streaming Heterogeneous Graphs. Pages 1035–1044 of: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. New York, NY: Association for Computing Machinery. (Cited on p. 469.)
- [291] Marai, G. E., Pinaud, B., Bühler, K., Lex, A., and Morris, J. H. 2019. Ten Simple Rules to Create Biological Network Figures for Communication. *PLOS Computational Biology*, **15**(9), e1007244. (Cited on p. 221.)
- [292] Mariani, M. S., Ren, Z.-M., Bascompte, J., and Tessone, C. J. 2019. Nestedness in Complex Networks: Observation, Emergence, and Implications. *Physics Reports*, **813**(June), 1–90. (Cited on pp. 197, 198, 202.)
- [293] Maron, H., Ben-Hamu, H., Shamir, N., and Lipman, Y. 2019. Invariant and Equivariant Graph Networks. In: *International Conference on Learning Representations*. New Orleans, LA: OpenReview.net. (Cited on p. 443.)
- [294] Martin, C., and Niemeyer, P. 2019. Influence of Measurement Errors on Networks: Estimating the Robustness of Centrality Measures. *Network Science*, **7**(2), 180–195. (Cited on pp. 377, 395.)

- [295] Martin, T., Ball, B., and Newman, M. E. J. 2016. Structural Inference for Uncertain Networks. *Physical Review E*, **93**(1), 012306. (Cited on p. 134.)
- [296] Marx, V. 2013. The Big Challenges of Big Data. *Nature*, **498**(7453), 255–260. (Cited on p. 469.)
- [297] Masuda, N., and Lambiotte, R. 2020. *A Guide to Temporal Networks*. 2nd edn. Series on Complexity Science, vol. 06. London: World Scientific. (Cited on pp. 238, 240, 241, 242, 249, 472.)
- [298] Masuda, N., Porter, M. A., and Lambiotte, R. 2017. Random Walks and Diffusion on Networks. *Physics Reports*, **716–717**(Nov.), 1–58. (Cited on p. 428.)
- [299] Matejka, J., and Fitzmaurice, G. 2017. Same Stats, Different Graphs: Generating Datasets with Varied Appearance and Identical Statistics through Simulated Annealing. Pages 1290–1294 of: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. Denver, CO: Association for Computing Machinery. (Cited on p. 203.)
- [300] McGregor, A. 2014. Graph Stream Algorithms: A Survey. *ACM SIGMOD Record*, **43**(1), 9–20. (Cited on pp. 453, 454, 469.)
- [301] McGregor, A., Vorotnikova, S., and Vu, H. T. 2016. Better Algorithms for Counting Triangles in Data Streams. Pages 401–411 of: *Proceedings of the 35th ACM SIGMOD-SIGACT-SIGAI Symposium on Principles of Database Systems*. San Francisco, CA: Association for Computing Machinery. (Cited on pp. 454, 469.)
- [302] McInnes, L., Healy, J., and Melville, J. 2018. *UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction*. arXiv:1802.03426. (Cited on p. 259.)
- [303] Mead, C., and Conway, L. 1980. *Introduction to VLSI Systems*. Reading, MA: Addison-Wesley. (Cited on p. 26.)
- [304] Meinshausen, N., and Bühlmann, P. 2006. High-Dimensional Graphs and Variable Selection with the Lasso. *The Annals of Statistics*, **34**(3). (Cited on p. 371.)
- [305] Menczer, F., Fortunato, S., and Davis, C. A. 2020. *A First Course in Network Science*. Cambridge: Cambridge University Press. (Cited on pp. 16, 472.)
- [306] Meusel, R., Vigna, S., Lehmborg, O., and Bizer, C. 2014. Graph Structure in the Web — Revisited: A Trick of the Heavy Tail. Pages 427–432 of: *Proceedings of the 23rd International Conference on World Wide Web*. WWW '14 Companion. New York, NY: Association for Computing Machinery. (Cited on p. 448.)
- [307] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., and Dean, J. 2013. Distributed Representations of Words and Phrases and Their Compositionality. In: *Advances in Neural Information Processing Systems*, vol. 26. Lake Tahoe, NV: Curran Associates, Inc. (Cited on pp. 272, 431, 434, 445.)
- [308] Milgram, S. 1963. Behavioral Study of Obedience. *The Journal of Abnormal and Social Psychology*, **67**, 371–378. (Cited on p. 341.)
- [309] Milgram, S. 1967. The Small-World Problem. *Psychology Today*, **1**, 61. (Cited on pp. 15, 192, 458.)

- [310] Miller, J. C., and Hagberg, A. 2011. Efficient Generation of Networks with Given Expected Degrees. Pages 115–126 of: Frieze, A., Horn, P., and Pralat, P. (eds.), *Algorithms and Models for the Web Graph*, vol. 6732. Berlin, Heidelberg: Springer. (Cited on pp. 468, 469.)
- [311] Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. 2002. Network Motifs: Simple Building Blocks of Complex Networks. *Science*, **298**(5594), 824–827. (Cited on p. 201.)
- [312] Mitchell, T. M. 1997. *Machine Learning*. New York, NY: McGraw-Hill. (Cited on pp. 254, 472.)
- [313] Molloy, M., and Reed, B. 1995. A Critical Point for Random Graphs With a Given Degree Sequence. *Random Structures & Algorithms*, **6**(2-3), 161–180. (Cited on pp. 163, 333.)
- [314] Molloy, M., and Reed, B. 1998. The Size of the Giant Component of a Random Graph with a Given Degree Sequence. *Combinatorics, Probability and Computing*, **7**(3), 295–305. (Cited on pp. 163, 333.)
- [315] Molnar, C. 2020. *Interpretable Machine Learning*. Morrisville, NC: Lulu Press, Inc. (Cited on p. 275.)
- [316] Moore, P. B., Hendrickson, W. A., Henderson, R., and Brunger, A. T. 2022. The Protein-Folding Problem: Not Yet Solved. *Science*, **375**(6580), 507–507. (Cited on p. 17.)
- [317] Moreno, J. L. 1934. *Who Shall Survive?: A New Approach to the Problem of Human Interrelations*. Washington, DC: Nervous and Mental Disease Publishing Co. (Cited on pp. 7, 16, 25, 78, 221.)
- [318] Moreno, J. L. 1953. *Who Shall Survive? Foundations of Sociometry, Group Psychotherapy and Socio-Drama*. 2nd edn. Oxford: Beacon House. (Cited on p. 7.)
- [319] Moreno, R. 2021. *Words of the Daughter: A Memoir*. Morrisville, NC: Lulu Press, Inc. (Cited on p. 16.)
- [320] Morris, C., Ritzert, M., Fey, M. et al. 2019. Weisfeiler and Leman Go Neural: Higher-Order Graph Neural Networks. *Proceedings of the AAAI Conference on Artificial Intelligence*, **33**(01), 4602–4609. (Cited on p. 443.)
- [321] Morris, R. 1978. Counting Large Numbers of Events in Small Registers. *Communications of the ACM*, **21**(10), 840–842. (Cited on p. 469.)
- [322] Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., and Onnela, J.-P. 2010. Community Structure in Time-Dependent, Multiscale, and Multiplex Networks. *Science*, **328**(5980), 876–878. (Cited on pp. 181, 201.)
- [323] Munafò, M. R., Nosek, B. A., Bishop, D. V. M. et al. 2017. A Manifesto for Reproducible Science. *Nature Human Behaviour*, **1**(1), 0021. (Cited on p. 163.)
- [324] Murphy, K. P. 2022. *Probabilistic Machine Learning: An Introduction*. Cambridge, MA: MIT Press. (Cited on p. 473.)
- [325] Murray, D., Yoon, J., Kojaku, S. et al. 2021 (June). *Unsupervised Embedding of Trajectories Captures the Latent Structure of Mobility*. arXiv:2012.02785. (Cited on p. 435.)

- [326] Nadakuditi, R. R., and Newman, M. E. J. 2012. Graph Spectra and the Detectability of Community Structure in Networks. *Physical Review Letters*, **108**(18), 188701. (Cited on p. 357.)
- [327] Narayanan, A., Shi, E., and Rubinstein, B. I. P. 2011. Link Prediction by De-Anonymization: How We Won the Kaggle Social Network Challenge. Pages 1825–1834 of: *The 2011 International Joint Conference on Neural Networks*. San Jose, CA: IEEE. (Cited on p. 33.)
- [328] National Academies of Sciences, Engineering, and Medicine. 2019. *Reproducibility and Replicability in Science*. Washington, DC: National Academies Press. (Cited on p. 474.)
- [329] Nemeth, R. J., and Smith, D. A. 1985. International Trade and World-System Structure: A Multiple Network Analysis. *Review (Fernand Braudel Center)*, **8**(4), 517–560. (Cited on p. 194.)
- [330] Newman, E. A., Araque, A., Dubinsky, J. M., Swanson, L. W., King, L. S., and Himmel, E. (eds.). 2017. *The Beautiful Brain: The Drawings of Santiago Ramón y Cajal*. New York, NY: Abrams. (Cited on p. 16.)
- [331] Newman, M. E. J. 2002. Assortative Mixing in Networks. *Physical Review Letters*, **89**(20), 208701. (Cited on pp. 111, 201.)
- [332] Newman, M. E. J. 2003. Mixing Patterns in Networks. *Physical Review E*, **67**(2), 026126. (Cited on pp. 111, 116, 163, 173, 201.)
- [333] Newman, M. E. J. 2006. Finding Community Structure in Networks Using the Eigenvectors of Matrices. *Physical Review E*, **74**(3), 036104. (Cited on pp. 421, 427.)
- [334] Newman, M. E. J. 2006. Modularity and Community Structure in Networks. *Proceedings of the National Academy of Sciences*, **103**(23), 8577–8582. (Cited on pp. 201, 417, 418, 427.)
- [335] Newman, M. E. J. 2012. Communities, Modules and Large-Scale Structure in Networks. *Nature Physics*, **8**(1), 25–31. (Cited on p. 201.)
- [336] Newman, M. E. J. 2013. Spectral Methods for Community Detection and Graph Partitioning. *Physical Review E*, **88**(4), 042822. (Cited on p. 427.)
- [337] Newman, M. E. J. 2018. Network Structure from Rich but Noisy Data. *Nature Physics*, **14**(6), 542–545. (Cited on pp. 134, 365, 374.)
- [338] Newman, M. E. J., and Girvan, M. 2004. Finding and Evaluating Community Structure in Networks. *Physical Review E*, **69**(2), 026113. (Cited on p. 201.)
- [339] Newman, M. E. J., and Park, J. 2003. Why Social Networks Are Different from Other Types of Networks. *Physical Review E*, **68**(3), 036122. (Cited on pp. 380, 395.)
- [340] Newman, M. E. J., Strogatz, S. H., and Watts, D. J. 2001. Random Graphs with Arbitrary Degree Distributions and Their Applications. *Physical Review E*, **64**(2), 026118. (Cited on p. 340.)
- [341] Newman, M. E. J., Watts, D. J., and Strogatz, S. H. 2002. Random Graph Models of Social Networks. *Proceedings of the National Academy of Sciences*, **99**(suppl. 1), 2566–2572. (Cited on p. 380.)

- [342] Newman, M. 2018. *Networks*. Oxford: Oxford University Press. (Cited on pp. 16, 172, 472.)
- [343] Newman, M. 2005. Power Laws, Pareto Distributions and Zipf's Law. *Contemporary Physics*, **46**(5), 323–351. (Cited on p. 163.)
- [344] Ng, A., Jordan, M., and Weiss, Y. 2001. On Spectral Clustering: Analysis and an Algorithm. *Advances in Neural Information Processing Systems*, **14**. (Cited on p. 428.)
- [345] Nica, B. 2018. *A Brief Introduction to Spectral Graph Theory*. EMS Textbooks in Mathematics. Zürich, Switzerland: European Mathematical Society. (Cited on p. 427.)
- [346] Nickel, M., Tresp, V., and Kriegel, H.-P. 2011. A Three-Way Model for Collective Learning on Multi-Relational Data. Pages 809–816 of: *Proceedings of the 28th International Conference on International Conference on Machine Learning*. ICML'11. Madison, WI: Omnipress. (Cited on p. 437.)
- [347] Nickel, M., and Kiela, D. 2017. Poincaré Embeddings for Learning Hierarchical Representations. In: *Advances in Neural Information Processing Systems*, vol. 30. Long Beach, CA: Curran Associates, Inc. (Cited on p. 438.)
- [348] Nilsson, N. J. 2009. *The Quest for Artificial Intelligence*. Cambridge: Cambridge University Press. (Cited on p. 276.)
- [349] Nolte, D. D. 2019. *Introduction to Modern Dynamics: Chaos, Networks, Space and Time*. 2nd edn. Oxford; New York, NY: Oxford University Press. (Cited on p. 428.)
- [350] Nowicki, K., and Snijders, T. A. B. 2001. Estimation and Prediction for Stochastic Blockstructures. *Journal of the American Statistical Association*, **96**(455), 1077–1087. (Cited on p. 374.)
- [351] Obermeyer, Z., Powers, B., Vogeli, C., and Mullainathan, S. 2019. Dissecting Racial Bias in an Algorithm Used to Manage the Health of Populations. *Science*, **366**(6464), 447–453. (Cited on pp. 30, 31.)
- [352] O'Neil, C. 2017. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. New York, NY: Crown. (Cited on pp. 37, 277.)
- [353] Ozella, L., Paolotti, D., Lichand, G. et al. 2021. Using Wearable Proximity Sensors to Characterize Social Contact Patterns in a Village of Rural Malawi. *EPJ Data Science*, **10**(1), 46. (Cited on pp. 25, 37, 246.)
- [354] Palla, G., Derényi, I., Farkas, I., and Vicsek, T. 2005. Uncovering the Overlapping Community Structure of Complex Networks in Nature and Society. *Nature*, **435**(7043), 814–818. (Cited on pp. 181, 201.)
- [355] Palmer, C. R., Gibbons, P. B., and Faloutsos, C. 2002. ANF: A Fast and Scalable Tool for Data Mining in Massive Graphs. Pages 81–90 of: *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '02. New York, NY: Association for Computing Machinery. (Cited on pp. 456, 457, 469.)

- [356] Panko, R. R. 1998. What We Know About Spreadsheet Errors. *Journal of Organizational and End User Computing*, **10**(2), 15–21. (Cited on p. 36.)
- [357] Pantel, P. 2005. Inducing Ontological Co-Occurrence Vectors. Pages 125–132 of: *Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics*. ACL '05. Ann Arbor, MI: Association for Computational Linguistics. (Cited on pp. 432, 445.)
- [358] Park, J., Wood, I. B., Jing, E. et al. 2019. Global Labor Flow Network Reveals the Hierarchical Organization and Dynamics of Geo-Industrial Clusters. *Nature Communications*, **10**(1), 3449. (Cited on p. 21.)
- [359] Park, J., and Newman, M. E. J. 2004. Solution of the Two-Star Model of a Network. *Physical Review E*, **70**(6), 066146. (Cited on p. 369.)
- [360] Park, J., and Newman, M. E. J. 2005. A Network-Based Ranking System for US College Football. *Journal of Statistical Mechanics: Theory and Experiment*, **2005**(10), P10014–P10014. (Cited on p. 216.)
- [361] Pascual, M., and Dunne, J. A. 2005. *Ecological Networks: Linking Structure to Dynamics in Food Webs*. Oxford: Oxford University Press. (Cited on p. 25.)
- [362] Pastor-Satorras, R., and Vespignani, A. 2001. Epidemic Spreading in Scale-Free Networks. *Physical Review Letters*, **86**(14), 3200–3203. (Cited on p. 333.)
- [363] Patterson, B. D., and Atmar, W. 1986. Nested Subsets and the Structure of Insular Mammalian Faunas and Archipelagos. *Biological Journal of the Linnean Society*, **28**(1–2), 65–82. (Cited on p. 202.)
- [364] Peel, L., Larremore, D. B., and Clauset, A. 2017. The Ground Truth about Metadata and Community Detection in Networks. *Science Advances*, **3**(5), e1602548. (Cited on p. 107.)
- [365] Peixoto, T. P. 2013. Parsimonious Module Inference in Large Networks. *Physical Review Letters*, **110**(14), 148701. (Cited on p. 354.)
- [366] Peixoto, T. P. 2014. Hierarchical Block Structures and High-Resolution Model Selection in Large Networks. *Physical Review X*, **4**(1), 011047. (Cited on p. 358.)
- [367] Pelleg, D., and Moore, A. W. 2000. X-Means: Extending K-means with Efficient Estimation of the Number of Clusters. Pages 727–734 of: *Proceedings of the Seventeenth International Conference on Machine Learning*. ICML '00. San Francisco, CA: Morgan Kaufmann. (Cited on p. 426.)
- [368] Pereira, D. A., and Williams, J. A. 2007. Origin and Evolution of High Throughput Screening. *British Journal of Pharmacology*, **152**(1), 53–61. (Cited on p. 82.)
- [369] Perozzi, B., Al-Rfou, R., and Skiena, S. 2014. DeepWalk: Online Learning of Social Representations. Pages 701–710 of: *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. New York, NY: Association for Computing Machinery. (Cited on pp. 435, 445.)
- [370] Perry, B. L., Pescosolido, B. A., and Borgatti, S. P. 2018. *Egocentric Network Analysis: Foundations, Methods, and Models*. Cambridge: Cambridge University Press. (Cited on p. 201.)

- [371] Pfeiffer, J., and Neville, J. 2021. Methods to Determine Node Centrality and Clustering in Graphs with Uncertain Structure. *Proceedings of the International AAAI Conference on Web and Social Media*, **5**(1), 590–593. (Cited on pp. 387, 395.)
- [372] Phan, T. Q., and Airolidi, E. M. 2015. A Natural Experiment of Social Network Formation and Dynamics. *Proceedings of the National Academy of Sciences*, **112**(21), 6595–6600. (Cited on p. 474.)
- [373] Pilgrim, C., and Hills, T. T. 2021. Bias in Zipf's Law Estimators. *Scientific Reports*, **11**(1), 17309. (Cited on p. 163.)
- [374] Pimentel, J. F., Murta, L., Braganholo, V., and Freire, J. 2019. A Large-Scale Study About Quality and Reproducibility of Jupyter Notebooks. Pages 507–517 of: *2019 IEEE/ACM 16th International Conference on Mining Software Repositories (MSR)*. Montreal, Canada: IEEE. (Cited on p. 302.)
- [375] Porter, M., and Gleeson, J. 2016. *Dynamical Systems on Networks: A Tutorial*. Frontiers in Applied Dynamical Systems: Reviews and Tutorials, no. 4. Cham, Switzerland: Springer. (Cited on pp. 235, 428.)
- [376] Porter, M. A., Onnela, J.-P., Mucha, P. J., et al. 2009. Communities in Networks. *Notices of the AMS*, **56**(9), 1082–1097. (Cited on p. 201.)
- [377] Pothén, A., Simon, H. D., and Liou, K.-P. 1990. Partitioning Sparse Matrices with Eigenvectors of Graphs. *SIAM Journal on Matrix Analysis and Applications*, **11**(3), 430–452. (Cited on pp. 412, 427.)
- [378] Price, D. D. S. 1976. A General Theory of Bibliometric and Other Cumulative Advantage Processes. *Journal of the American Society for Information Science*, **27**(5), 292–306. (Cited on p. 349.)
- [379] Price, D. J. d. S. 1965. Networks of Scientific Papers: The Pattern of Bibliographic References Indicates the Nature of the Scientific Research Front. *Science*, **149**(3683), 510–515. (Cited on p. 349.)
- [380] Proulx, S. R., Promislow, D. E. L., and Phillips, P. C. 2005. Network Thinking in Ecology and Evolution. *Trends in Ecology & Evolution*, **20**(6), 345–353. (Cited on p. 25.)
- [381] Pržulj, N. 2007. Biological Network Comparison Using Graphlet Degree Distribution. *Bioinformatics*, **23**(2), e177–e183. (Cited on p. 228.)
- [382] Qiu, J., Dong, Y., Ma, H., Li, J., Wang, K., and Tang, J. 2018. Network Embedding as Matrix Factorization: Unifying DeepWalk, LINE, PTE, and Node2vec. Pages 459–467 of: *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*. WSDM '18. New York, NY: Association for Computing Machinery. (Cited on pp. 440, 441, 445.)
- [383] Radford, A., Narasimhan, K., Salimans, T., and Sutskever, I. 2018. *Improving Language Understanding by Generative Pre-Training*. (Cited on p. 261.)
- [384] Raghavan, U. N., Albert, R., and Kumara, S. 2007. Near Linear Time Algorithm to Detect Community Structures in Large-Scale Networks. *Physical Review E*, **76**(3), 036106. (Cited on pp. 458, 459, 469.)

- [385] Ramani, A. S., Eikmeier, N., and Gleich, D. F. 2019. Coin-Flipping, Ball-Dropping, and Grass-Hopping for Generating Random Graphs from Matrices of Edge Probabilities. *SIAM Review*, **61**(3), 549–595. (Cited on p. 469.)
- [386] Ramón y Cajal, S. 1899. *Comparative Study of the Sensory Areas of the Human Cortex*. Worcester, MA: Clark University. (Cited on p. 5.)
- [387] Ramón y Cajal, S. 2004. *Advice for a Young Investigator*. Cambridge, MA: MIT Press. (Cited on p. 16.)
- [388] Ramón y Cajal, S., and Cowan, W. M. 1989. *Recollections of My Life*. Cambridge, MA: MIT Press. (Cited on p. 16.)
- [389] Ranshous, S., Shen, S., Koutra, D., Harenberg, S., Faloutsos, C., and Samatova, N. F. 2015. Anomaly Detection in Dynamic Networks: A Survey. *WIREs Computational Statistics*, **7**(3), 223–247. (Cited on p. 469.)
- [390] Reed, B. J., and Segal, D. R. 2006. Social Network Analysis and Counterinsurgency Operations: The Capture of Saddam Hussein. *Sociological Focus*, **39**(4), 251–264. (Cited on p. 33.)
- [391] Reichardt, J., and Bornholdt, S. 2006. Statistical Mechanics of Community Detection. *Physical Review E*, **74**(1), 016110. (Cited on p. 181.)
- [392] Reichardt, J., and Bornholdt, S. 2006. When Are Networks Truly Modular? *Physica D: Nonlinear Phenomena*, **224**(1-2), 20–26. (Cited on p. 179.)
- [393] Reinhart, C. M., and Rogoff, K. S. 2010. Growth in a Time of Debt. *American Economic Review*, **100**(2), 573–578. (Cited on p. 36.)
- [394] Resnik, D. B. 1998. *The Ethics of Science: An Introduction*. London: Routledge. (Cited on p. 37.)
- [395] Ribeiro, P., Paredes, P., Silva, M. E. P., Aparicio, D., and Silva, F. 2022. A Survey on Subgraph Counting: Concepts, Algorithms, and Applications to Network Motifs and Graphlets. *ACM Computing Surveys*, **54**(2), 1–36. (Cited on p. 201.)
- [396] Rider, P. R. 1955. Truncated Binomial and Negative Binomial Distributions. *Journal of the American Statistical Association*, **50**(271), 877–883. (Cited on p. 393.)
- [397] Robins, G., and Alexander, M. 2004. Small Worlds Among Interlocking Directors: Network Structure and Distance in Bipartite Graphs. *Computational & Mathematical Organization Theory*, **10**(1), 69–94. (Cited on p. 194.)
- [398] Robins, G., Pattison, P., Kalish, Y., and Lusher, D. 2007. An Introduction to Exponential Random Graph (P^*) Models for Social Networks. *Social Networks*, **29**(2), 173–191. (Cited on pp. 163, 374.)
- [399] Robinson, I., Webber, J., and Eifrem, E. 2015. *Graph Databases: New Opportunities for Connected Data*. Sebastopol, CA: O'Reilly Media. (Cited on p. 469.)
- [400] Rombach, M. P., Porter, M. A., Fowler, J. H., and Mucha, P. J. 2014. Core-Periphery Structure in Networks. *SIAM Journal on Applied Mathematics*, **74**(1), 167–190. (Cited on pp. 195, 202.)
- [401] Rota, G.-C. 1964. The Number of Partitions of a Set. *The American Mathematical Monthly*, **71**(5), 498. (Cited on p. 178.)

- [402] Roughan, M., and Tuke, J. 2015. The Hitchhikers Guide to Sharing Graph Data. Pages 435–442 of: *2015 3rd International Conference on Future Internet of Things and Cloud*. Rome, Italy: IEEE. (Cited on p. 98.)
- [403] Roughan, M., Tuke, S. J., and Maennel, O. 2008. Bigfoot, Sasquatch, the Yeti and Other Missing Links: What We Don't Know about the as Graph. Pages 325–330 of: *Proceedings of the 8th ACM SIGCOMM Conference on Internet Measurement*. Vouliagmeni, Greece: Association for Computing Machinery. (Cited on pp. 392, 395.)
- [404] Rousseeuw, P. J. 1987. Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *Journal of Computational and Applied Mathematics*, **20**(Nov.), 53–65. (Cited on p. 426.)
- [405] Rozenblatt-Rosen, O., Regev, A., Oberdoerffer, P. et al. 2020. The Human Tumor Atlas Network: Charting Tumor Transitions across Space and Time at Single-Cell Resolution. *Cell*, **181**(2), 236–249. (Cited on p. 473.)
- [406] Rubin, D. B. 1976. Inference and Missing Data. *Biometrika*, **63**(3), 581–592. (Cited on p. 134.)
- [407] Rubin, D. B. 1996. Multiple Imputation after 18+ Years. *Journal of the American Statistical Association*, **91**(434), 473–489. (Cited on pp. 116, 134.)
- [408] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. 1986. Learning Representations by Back-Propagating Errors. *Nature*, **323**(6088), 533–536. (Cited on p. 258.)
- [409] Russell, M. A. 2013. *Mining the Social Web: Data Mining Facebook, Twitter, LinkedIn, Google+, GitHub, and More*. Sebastopol, CA: O'Reilly Media. (Cited on p. 71.)
- [410] Russell, S. J., and Norvig, P. 2021. *Artificial Intelligence: A Modern Approach*. New York, NY: Pearson. (Cited on p. 473.)
- [411] Russell, S. J., Russell, S., and Norvig, P. 2020. *Artificial Intelligence: A Modern Approach*. New York, NY: Pearson. (Cited on p. 277.)
- [412] Salganik, M. J. 2018. *Bit by Bit: Social Research in the Digital Age*. Princeton, NJ: Princeton University Press. (Cited on pp. 25, 65, 68, 71, 82.)
- [413] Sambourg, L., and Thierry-Mieg, N. 2010. New Insights into Protein-Protein Interaction Data Lead to Increased Estimates of the *S. Cerevisiae* Interactome Size. *BMC Bioinformatics*, **11**(1), 605. (Cited on p. 395.)
- [414] Sarigol, E., Garcia, D., and Schweitzer, F. 2014. Online Privacy as a Collective Phenomenon. Pages 95–106 of: *Proceedings of the Second ACM Conference on Online Social Networks*. COSN '14. New York, NY: Association for Computing Machinery. (Cited on p. 37.)
- [415] Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., and Monfardini, G. 2009. The Graph Neural Network Model. *IEEE Transactions on Neural Networks*, **20**(1), 61–80. (Cited on p. 445.)
- [416] Schaub, M. T., Delvenne, J.-C., Rosvall, M., and Lambiotte, R. 2017. The Many Facets of Community Detection in Complex Networks. *Applied Network Science*, **2**(1), 1–13. (Cited on p. 201.)

- [417] Schimel, J. 2012. *Writing Science: How to Write Papers That Get Cited and Proposals That Get Funded*. New York, NY: Oxford University Press. (Cited on p. 473.)
- [418] Schreier, A. A., Wilson, K., and Resnik, D. 2006. Academic Research Record-Keeping: Best Practices for Individuals, Group Leaders, and Institutions. *Academic Medicine*, **81**(1), 42–47. (Cited on pp. 287, 292.)
- [419] Schwabish, J. 2021. *Better Data Visualizations: A Guide for Scholars, Researchers, and Wonks*. New York, NY: Columbia University Press. (Cited on p. 473.)
- [420] Scopatz, A., and Huff, K. D. 2015. *Effective Computation in Physics: Field Guide to Research with Python*. Sebastopol, CA: O'Reilly Media. (Cited on pp. 102, 312, 313.)
- [421] Seal, H. L. 1952. The Maximum Likelihood Fitting of the Discrete Pareto Law. *Journal of the Institute of Actuaries*, **78**(1), 115–121. (Cited on p. 163.)
- [422] Seebauer, E. G., and Barry, R. L. 2000. *Fundamentals of Ethics for Scientists and Engineers*. Oxford: Oxford University Press. (Cited on p. 37.)
- [423] Serafino, M., Cimini, G., Maritan, A. et al. 2021. True Scale-Free Networks Hidden by Finite Size Effects. *Proceedings of the National Academy of Sciences*, **118**(2), e2013825118. (Cited on p. 346.)
- [424] Serrano, M. Á., Boguñá, M., and Vespignani, A. 2009. Extracting the Multiscale Backbone of Complex Weighted Networks. *Proceedings of the National Academy of Sciences*, **106**(16), 6483–6488. (Cited on pp. 126, 134, 375.)
- [425] Sheppard, C. 2017. *Tree-Based Machine Learning Algorithms: Decision Trees, Random Forests, and Boosting*. Seattle, WA: CreateSpace. (Cited on p. 275.)
- [426] Sherman, R. M., Forman, J., Antonescu, V. et al. 2019. Assembly of a Pan-Genome from Deep Sequencing of 910 Humans of African Descent. *Nature Genetics*, **51**(1), 30–35. (Cited on p. 29.)
- [427] Shervashidze, N., Vishwanathan, S. V. N., Petri, T., Mehlhorn, K., and Borgwardt, K. 2009. Efficient Graphlet Kernels for Large Graph Comparison. Pages 488–495 of: *Proceedings of the Twelfth International Conference on Artificial Intelligence and Statistics*. Proceedings of Machine Learning Research, vol. 5. Clearwater Beach, FL: PMLR. (Cited on p. 228.)
- [428] Shi, J., and Malik, J. 2000. Normalized Cuts and Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **22**(8), 888–905. (Cited on pp. 412, 413, 428.)
- [429] Shneiderman, B. 2020. Bridging the Gap Between Ethics and Practice: Guidelines for Reliable, Safe, and Trustworthy Human-centered AI Systems. *ACM Transactions on Interactive Intelligent Systems*, **10**(4), 26:1–26:31. (Cited on p. 31.)
- [430] Shumway, R. H. 2017. *Time Series Analysis and Its Applications: With R Examples*. New York, NY: Springer Science+Business Media. (Cited on p. 249.)
- [431] Simkin, M., and Roychowdhury, V. 2011. Re-Inventing Willis. *Physics Reports*, May, S0370157310003339. (Cited on p. 349.)

- [432] Simon, H. A. 1955. On a Class of Skew Distribution Functions. *Biometrika*, **42**(3–4), 425–440. (Cited on pp. 344, 349.)
- [433] Sinclair, A., and Jerrum, M. 1989. Approximate Counting, Uniform Generation and Rapidly Mixing Markov Chains. *Information and Computation*, **82**(1), 93–133. (Cited on p. 427.)
- [434] Sirugo, G., Williams, S. M., and Tishkoff, S. A. 2019. The Missing Diversity in Human Genetic Studies. *Cell*, **177**(1), 26–31. (Cited on p. 28.)
- [435] Smith, R. C. 2013. *Uncertainty Quantification: Theory, Implementation, and Applications*. Computational Science and Engineering Series. Philadelphia, PA: Society for Industrial and Applied Mathematics. (Cited on p. 395.)
- [436] Snijders, T. A. B., Pattison, P. E., Robins, G. L., and Handcock, M. S. 2006. New Specifications for Exponential Random Graph Models. *Sociological Methodology*, **36**(1), 99–153. (Cited on p. 369.)
- [437] Snijders, T. A., and Nowicki, K. 1997. Estimation and Prediction for Stochastic Blockmodels for Graphs with Latent Block Structure. *Journal of Classification*, **14**(1), 75–100. (Cited on pp. 353, 374.)
- [438] Sood, V., and Redner, S. 2005. Voter Model on Heterogeneous Graphs. *Physical Review Letters*, **94**(17), 178701. (Cited on p. 235.)
- [439] Soranzo, N., Bianconi, G., and Altafini, C. 2007. Comparing Association Network Algorithms for Reverse Engineering of Large-Scale Gene Regulatory Networks: Synthetic versus Real Data. *Bioinformatics*, **23**(13), 1640–1647. (Cited on p. 134.)
- [440] Soundarajan, S., Eliassi-Rad, T., and Gallagher, B. 2014. A Guide to Selecting a Network Similarity Method. Pages 1037–1045 of: *Proceedings of the 2014 SIAM International Conference on Data Mining (SDM)*. Philadelphia, PA: Society for Industrial and Applied Mathematics. (Cited on p. 232.)
- [441] Spielman, D. A. 2007. Spectral Graph Theory and Its Applications. Pages 29–38 of: *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS'07)*. Providence, Rhode Island: IEEE. (Cited on p. 427.)
- [442] Sporns, O. 2016. *Networks of the Brain*. Cambridge, MA: MIT Press. (Cited on pp. 25, 82.)
- [443] Staiano, J., Lepri, B., Aharony, N., Pianesi, F., Sebe, N., and Pentland, A. 2012. Friends Don't Lie: Inferring Personality Traits from Social Network Structure. Pages 321–330 of: *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. UbiComp '12. New York, NY: Association for Computing Machinery. (Cited on p. 37.)
- [444] Stewart, J. 2017. *Python for Scientists*. 2nd edn. Cambridge; New York, NY: Cambridge University Press. (Cited on p. 61.)
- [445] Stopczynski, A., Sekara, V., Sapiezynski, P. et al. 2014. Measuring Large-Scale Social Networks with High Resolution. *PLOS ONE*, **9**(4), e95978. (Cited on p. 473.)
- [446] Strang, G. 2019. *Linear Algebra and Learning from Data*. Wellesley, MA: Wellesley-Cambridge Press. (Cited on pp. 62, 277, 427.)

- [447] Strang, G. 2023. *Introduction to Linear Algebra*. 6th edn. Wellesley, MA: Wellesley-Cambridge. (Cited on pp. 61, 427.)
- [448] Stumpf, M. P. H., Thorne, T., de Silva, E. et al. 2008. Estimating the Size of the Human Interactome. *Proceedings of the National Academy of Sciences*, **105**(19), 6959–6964. (Cited on pp. 390, 395.)
- [449] Sugimoto, C. R., Ahn, Y.-Y., Smith, E., Macaluso, B., and Larivière, V. 2019. Factors Affecting Sex-Related Reporting in Medical Research: A Cross-Disciplinary Bibliometric Analysis. *The Lancet*, **393**(10171), 550–559. (Cited on p. 28.)
- [450] Sullivan, D. 2020. *A Reintroduction to Our Knowledge Graph and Knowledge Panels*. (Cited on p. 449.)
- [451] Sullivan, T. J. 2015. *Introduction to Uncertainty Quantification*. Texts in Applied Mathematics, no. 63. Cham, Switzerland: Springer. (Cited on p. 395.)
- [452] Sun, J., Bagrow, J. P., Bollt, E. M., and Skufca, J. D. 2009. Dynamic Computation of Network Statistics via Updating Schema. *Physical Review E*, **79**(3), 036116. (Cited on pp. 459, 465, 466, 469.)
- [453] Tang, J., Qu, M., Wang, M., Zhang, M., Yan, J., and Mei, Q. 2015. LINE: Large-scale Information Network Embedding. Pages 1067–1077 of: *Proceedings of the 24th International Conference on World Wide Web*. WWW '15. Geneva, Switzerland: International World Wide Web Conferences Steering Committee. (Cited on pp. 437, 440.)
- [454] Taylor, D., Myers, S. A., Clauset, A., Porter, M. A., and Mucha, P. J. 2017. Eigenvector-Based Centrality Measures for Temporal Networks. *Multiscale Modeling & Simulation*, **15**(1), 537–574. (Cited on p. 241.)
- [455] Taylor, J. R. 1997. *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. 2nd edn. Sausalito, CA: University Science Books. (Cited on pp. 134, 392.)
- [456] The Gene Ontology Consortium, Carbon, S., Douglass, E. et al. 2021. The Gene Ontology Resource: Enriching a Gold Mine. *Nucleic Acids Research*, **49**(D1), D325–D334. (Cited on p. 116.)
- [457] Thiebes, S., Lins, S., and Sunyaev, A. 2021. Trustworthy Artificial Intelligence. *Electronic Markets*, **31**(2), 447–464. (Cited on p. 31.)
- [458] Thorndike, R. L. 1953. Who Belongs in the Family? *Psychometrika*, **18**(4), 267–276. (Cited on p. 426.)
- [459] Tibshirani, R., Walther, G., and Hastie, T. 2001. Estimating the Number of Clusters in a Data Set via the Gap Statistic. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **63**(2), 411–423. (Cited on p. 426.)
- [460] Ting, D. 2016. Towards Optimal Cardinality Estimation of Unions and Intersections with Sketches. Pages 1195–1204 of: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. New York, NY: Association for Computing Machinery. (Cited on p. 456.)
- [461] Traag, V. A., Waltman, L., and van Eck, N. J. 2019. From Louvain to Leiden:

- Guaranteeing Well-Connected Communities. *Scientific Reports*, **9**(1), 5233. (Cited on p. 179.)
- [462] Tracy, S. J. 2010. Qualitative Quality: Eight “Big-Tent” Criteria for Excellent Qualitative Research. *Qualitative Inquiry*, **16**(10), 837–851. (Cited on p. 226.)
- [463] Trouillon, T., Welbl, J., Riedel, S., Gaussier, É., and Bouchard, G. 2016. Complex Embeddings for Simple Link Prediction. Pages 2071–2080 of: *Proceedings of the 33rd International Conference on International Conference on Machine Learning*. ICML’16, vol. 48. New York, NY: [JMLR.org](http://jmlr.org). (Cited on p. 437.)
- [464] Tufte, E. R. 2001. *The Visual Display of Quantitative Information*. 2nd edn. Cheshire, CT: Graphics Press. (Cited on p. 473.)
- [465] Tukey, J. W. 1977. *Exploratory Data Analysis*. Addison-Wesley Series in Behavioral Science. Reading, MA: Addison-Wesley. (Cited on pp. 71, 162.)
- [466] Tutte, W. T. 1963. How to Draw a Graph. *Proceedings of the London Mathematical Society*, **3**(1), 743–767. (Cited on p. 221.)
- [467] Ugander, J., Backstrom, L., Marlow, C., and Kleinberg, J. 2012. Structural Diversity in Social Contagion. *Proceedings of the National Academy of Sciences*, **109**(16), 5962–5966. (Cited on p. 201.)
- [468] Ulrich, W., Almeida-Neto, M., and Gotelli, N. J. 2009. A Consumer’s Guide to Nestedness Analysis. *Oikos*, **118**(1), 3–17. (Cited on p. 202.)
- [469] van Buuren, S. 2018. *Flexible Imputation of Missing Data*. 2nd edn. Chapman and Hall/CRC Interdisciplinary Statistics Series. Boca Raton, FL: CRC Press, Taylor and Francis Group. (Cited on pp. 116, 134.)
- [470] van den Heuvel, M. P., and Sporns, O. 2011. Rich-Club Organization of the Human Connectome. *Journal of Neuroscience*, **31**(44), 15775–15786. (Cited on p. 202.)
- [471] van der Maaten, L., and Hinton, G. 2008. Visualizing Data Using T-SNE. *Journal of Machine Learning Research*, **9**(86), 2579–2605. (Cited on p. 259.)
- [472] van Mieghem, P. 2010. *Graph Spectra for Complex Networks*. Cambridge: Cambridge University Press. (Cited on pp. 406, 427.)
- [473] Vanderplas, J. T. 2016. *Python Data Science Handbook: Essential Tools for Working with Data*. Sebastopol, CA: O’Reilly Media. (Cited on p. 102.)
- [474] Vaswani, A., Shazeer, N., Parmar, N. et al. 2017. Attention Is All You Need. In: *Advances in Neural Information Processing Systems*, vol. 30. Long Beach, CA: Curran Associates, Inc. (Cited on p. 273.)
- [475] Veličković, P., Cucurull, G., Casanova, A., Romero, A., Liò, P., and Bengio, Y. 2018. Graph Attention Networks. In: *International Conference on Learning Representations*. Vancouver, BC: OpenReview.net. (Cited on pp. 273, 442.)
- [476] Vitter, J. S. 1985. Random Sampling with a Reservoir. *ACM Transactions on Mathematical Software*, **11**(1), 37–57. (Cited on p. 453.)
- [477] Voitalov, I., van der Hoorn, P., van der Hofstad, R., and Krioukov, D. 2019. Scale-Free Networks Well Done. *Physical Review Research*, **1**(3), 033034. (Cited on p. 346.)

- [478] von Luxburg, U. 2007. A Tutorial on Spectral Clustering. *Statistics and Computing*, **17**(4), 395–416. (Cited on pp. 205, 232, 423, 425, 428.)
- [479] Vreeman, R. C., and Carroll, A. E. 2008. Festive Medical Myths. *BMJ*, **337**, a2769. (Cited on p. 86.)
- [480] Wagner, C., Strohmaier, M., Olteanu, A., Kıcıman, E., Contractor, N., and Eliassi-Rad, T. 2021. Measuring Algorithmically Infused Societies. *Nature*, **595**(7866), 197–204. (Cited on p. 473.)
- [481] Walpole, R. E., Myers, R. H., Myers, S. L., and Ye, K. (eds.). 2017. *Probability & Statistics for Engineers & Scientists*. 9th edn. Boston, MA: Pearson. (Cited on p. 62.)
- [482] Wang, C. C., Prather, K. A., Sznitman, J. et al. 2021. Airborne Transmission of Respiratory Viruses. *Science*, **373**(6558), eabd9149. (Cited on p. 86.)
- [483] Wang, Y. X. R., and Bickel, P. J. 2017. Likelihood-Based Model Selection for Stochastic Block Models. *The Annals of Statistics*, **45**(2), 500–528. (Cited on p. 354.)
- [484] Wasserman, L. A. 2004. *All of Statistics: A Concise Course in Statistical Inference*. New York, NY: Springer Science & Business Media. (Cited on pp. 62, 375, 472.)
- [485] Wasserman, S., and Faust, K. 1994. *Social Network Analysis: Methods and Applications*. Structural Analysis in the Social Sciences. Cambridge: Cambridge University Press. (Cited on pp. 25, 163, 201, 380, 472.)
- [486] Watts, D. J., and Strogatz, S. H. 1998. Collective Dynamics of ‘Small-World’ Networks. *Nature*, **393**(6684), 440–442. (Cited on pp. 15, 201, 343, 346, 349.)
- [487] Weisfeiler, B., and Leman, A. 1968. The Reduction of a Graph to Canonical Form and the Algebra Which Appears Therein. *Nauchno-Tekhnicheskaya Informatsia, Seriya*, **2**(9), 12–16. (Cited on p. 443.)
- [488] Wernicke, S. 2006. Efficient Detection of Network Motifs. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, **3**(4), 347–359. (Cited on p. 175.)
- [489] White, H. C., Boorman, S. A., and Breiger, R. L. 1976. Social Structure from Multiple Networks. I. Blockmodels of Roles and Positions. *American Journal of Sociology*, **81**(4), 730–780. (Cited on p. 374.)
- [490] Wickham, H. 2014. Tidy Data. *Journal of Statistical Software*, **59**(Sept.), 1–23. (Cited on pp. 109, 116, 217.)
- [491] Wickham, H., and Grolemund, G. 2016. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. Sebastopol, CA: O’Reilly Media. (Cited on pp. 61, 102.)
- [492] Wilf, H. S. 2006. *Generatingfunctionology*. Wellesley, MA: A K Peters, Ltd. (Cited on p. 349.)
- [493] Wilson, G., Aruliah, D. A., Brown, C. T. et al. 2014. Best Practices for Scientific Computing. *PLOS Biology*, **12**(1), e1001745. (Cited on p. 300.)
- [494] Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L., and Teal, T. K.

2017. Good Enough Practices in Scientific Computing. *PLOS Computational Biology*, **13**(6), e1005510. (Cited on p. 300.)
- [495] Wong, D. M. 2013. *The Wall Street Journal Guide to Information Graphics: The Do's And Don'ts Of Presenting Data Facts And Figures*. Washington, DC: National Geographic Books. (Cited on p. 473.)
- [496] Xin, R. S., Gonzalez, J. E., Franklin, M. J., and Stoica, I. 2013. GraphX: A Resilient Distributed Graph System on Spark. Pages 1–6 of: *First International Workshop on Graph Data Management Experiences and Systems*. GRADES '13. New York, NY: Association for Computing Machinery. (Cited on p. 453.)
- [497] Xu, K., Hu, W., Leskovec, J., and Jegelka, S. 2019. How Powerful Are Graph Neural Networks? In: *International Conference on Learning Representations*. New Orleans, LA: OpenReview.net. (Cited on p. 443.)
- [498] Xu, M. 2021. Understanding Graph Embedding Methods and Their Applications. *SIAM Review*, **63**(4), 825–853. (Cited on p. 445.)
- [499] Xu, R., and Wunsch, D. 2005. Survey of Clustering Algorithms. *IEEE Transactions on Neural Networks*, **16**(3), 645–678. (Cited on p. 232.)
- [500] Yanai, I., and Lercher, M. 2020. A Hypothesis Is a Liability. *Genome Biology*, **21**(1), 231. (Cited on p. 70.)
- [501] Yang, B., Yih, W.-t., He, X., Gao, J., and Deng, L. 2015. *Embedding Entities and Relations for Learning and Inference in Knowledge Bases*. arXiv:1412.6575. (Cited on p. 437.)
- [502] Young, J.-G., Desrosiers, P., Hébert-Dufresne, L., Laurence, E., and Dubé, L. J. 2017. Finite-Size Analysis of the Detectability Limit of the Stochastic Block Model. *Physical Review E*, **95**(6), 062304. (Cited on p. 357.)
- [503] Young, J.-G., Cantwell, G. T., and Newman, M. E. J. 2021. Bayesian Inference of Network Structure from Unreliable Data. *Journal of Complex Networks*, **8**(6), cnaa046. (Cited on p. 394.)
- [504] Yule, G. U. 1924. II.—A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F. R. S. *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, **213**(402–410), 21–87. (Cited on pp. 344, 349.)
- [505] Zachary, W. W. 1977. An Information Flow Model for Conflict and Fission in Small Groups. *Journal of Anthropological Research*, **33**(4), 452–473. (Cited on pp. 23, 82, 355, 418.)
- [506] Zhang, C., Bengio, S., Hardt, M., Recht, B., and Vinyals, O. 2021. Understanding Deep Learning (Still) Requires Rethinking Generalization. *Communications of the ACM*, **64**(3), 107–115. (Cited on p. 277.)
- [507] Zhang, H., Goel, A., Govindan, R., Mason, K., and Van Roy, B. 2004. Making Eigenvector-Based Reputation Systems Robust to Collusion. Pages 92–104 of: Leonardi, S. (ed.), *Algorithms and Models for the Web-Graph*, vol. 3243. Berlin, Heidelberg: Springer. (Cited on p. 189.)
- [508] Zhou, J., Cui, G., Hu, S. et al. 2020. Graph Neural Networks: A Review of Methods and Applications. *AI Open*, **1**(Jan.), 57–81. (Cited on p. 445.)

- [509] Zhou, S., and Mondragon, R. 2004. The Rich-Club Phenomenon in the Internet Topology. *IEEE Communications Letters*, **8**(3), 180–182. (Cited on p. 202.)
- [510] Zhou, Z.-H. 2021. *Machine Learning*. Singapore: Springer Singapore. (Cited on p. 472.)
- [511] Ziemann, M., Eren, Y., and El-Osta, A. 2016. Gene Name Errors Are Widespread in the Scientific Literature. *Genome Biology*, **17**(1), 177. (Cited on p. 36.)
- [512] Zuboff, S. 2020. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York, NY: PublicAffairs. (Cited on pp. 32, 34.)
- [513] Zuckerman, E. W., and Jost, J. T. 2001. What Makes You Think You're so Popular? Self-Evaluation Maintenance and the Subjective Side of the "Friendship Paradox". *Social Psychology Quarterly*, **64**(3), 207–223. (Cited on p. 324.)
- [514] Zügner, D., Akbarnejad, A., and Günnemann, S. 2018. Adversarial Attacks on Neural Networks for Graph Data. Pages 2847–2856 of: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. KDD '18. New York, NY: Association for Computing Machinery. (Cited on p. 444.)

