Bibliography

J. Abello, M. G. C. Resende, and S. Sudarsky. Massive quasi-clique detection. In *LATIN*, pages 598–612, 2002. DOI: 10.1007/3-540-45995-2_51 34, 35

N. Agarwal, H. Liu, L. Tang, and P. S. Yu. Identifying the influential bloggers in a community. In *WSDM '08: Proceedings of the international conference on Web search and web data mining*, pages 207–218, New York, NY, USA, 2008. ACM. ISBN 978-1-59593-927-9. DOI: 10.1145/1341531.1341559 8

A. Anagnostopoulos, R. Kumar, and M. Mahdian. Influence and correlation in social networks. In *KDD '08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 7–15, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-193-4. DOI: 10.1145/1401890.1401897 27, 28, 29

R. Andersen and K. J. Lang. Communities from seed sets. In *WWW '06: Proceedings of the 15th international conference on World Wide Web*, pages 223–232, New York, NY, USA, 2006. ACM. ISBN 1-59593-323-9. DOI: 10.1145/1135777.1135814 9

S. Aral, L. Muchnik, and A. Sundararajan. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proceedings of the National Academy of Sciences*, 106(51):21544–21549, 2009. URL http://www.pnas.org/content/106/51/21544. full. DOI: 10.1073/pnas.0908800106 29

S. Asur, S. Parthasarathy, and D. Ucar. An event-based framework for characterizing the evolutionary behavior of interaction graphs. In *KDD '07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 913–921, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-609-7. DOI: 10.1145/1281192.1281290 9, 76, 78

L. Backstrom, D. Huttenlocher, J. Kleinberg, and X. Lan. Group formation in large social networks: membership, growth, and evolution. In *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 44–54, New York, NY, USA, 2006. ACM. ISBN 1-59593-339-5. DOI: 10.1145/1150402.1150412 9, 76

L. Backstrom, C. Dwork, and J. Kleinberg. Wherefore art thou r3579x?: anonymized social networks, hidden patterns, and structural steganography. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 181–190, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-654-7. DOI: 10.1145/1242572.1242598 11

A.-L. Barabási and R. Albert. Emergence of Scaling in Random Networks. *Science*, 286(5439): 509–512, 1999. URL http://www.sciencemag.org/cgi/content/abstract/286/5439/ 509. DOI: 10.1126/science.286.5439.509 7, 8

L. Becchetti, P. Boldi, C. Castillo, and A. Gionis. Efficient semi-streaming algorithms for local triangle counting in massive graphs. In *KDD '08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 16–24, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-193-4. DOI: 10.1145/1401890.1401898 8

V. Blondel, J. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008:P10008, 2008. DOI: 10.1088/1742-5468/2008/10/P10008 48

I. Borg and P. Groenen. *Modern Multidimensional Scaling: theory and applications*. Springer, 2005. 37

U. Brandes. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2): 163–177, 2001. DOI: 10.1080/0022250X.2001.9990249 16, 18, 46, 48, 97

J. S. Breese, D. Heckerman, and C. Kadie. Empirical analysis of predictive algorithms for collaborative filtering. In *UAI '98: Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, 1998. 10

D. Chakrabarti and C. Faloutsos. Graph mining: Laws, generators, and algorithms. *ACM Comput. Surv.*, 38(1):2, 2006. ISSN 0360-0300. 7, 8

D. Chakrabarti, R. Kumar, and A. Tomkins. Evolutionary clustering. In *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 554–560, New York, NY, USA, 2006. ACM. ISBN 1-59593-339-5. DOI: 10.1145/1150402.1150467 79

W. Chen, Y. Wang, and S. Yang. Efficient influence maximization in social networks. In *KDD '09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 199–208, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-495-9. DOI: 10.1145/1557019.1557047 26

W. Chen, C. Wang, and Y. Wang. Scalable influence maximization for prevalent viral marketing in large-scale social networks. In *KDD '10: Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1029–1038, New York, NY, USA, 2010. ACM. ISBN 978-1-4503-0055-1. DOI: 10.1145/1835804.1835934 26

Y. Chi, X. Song, D. Zhou, K. Hino, and B. L. Tseng. Evolutionary spectral clustering by incorporating temporal smoothness. In *KDD '07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 153–162, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-609-7. DOI: 10.1145/1281192.1281212 79, 80

N. A. Christakis and J. H. Fowler. The spread of obesity in a large social network over 32 years. *The New England Journal of Medicine*, 357(4):370–379, 2007. DOI: 10.1056/NEJMsa066082 29

A. Clauset, M. Newman, and C. Moore. Finding community structure in very large networks. *Arxiv preprint cond-mat/0408187*, 2004. DOI: 10.1103/PhysRevE.70.066111 48, 52

D. Crandall, D. Cosley, D. Huttenlocher, J. Kleinberg, and S. Suri. Feedback effects between similarity and social influence in online communities. In *KDD '08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 160–168, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-193-4. DOI: 10.1145/1401890.1401914 27, 29

P. Desikan and J. Srivastava. I/o efficient computation of first order markov measures for large and evolving graphs. In *Proceedings of the Tenth Workshop on Web Mining and Web Usage Analysis (WebKDD)*, 2008. 8

I. S. Dhillon. Co-clustering documents and words using bipartite spectral graph partitioning. In *KDD '01: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 269–274, New York, NY, USA, 2001. ACM. ISBN 1-58113-391-X. DOI: 10.1145/502512.502550 69

J. Diesner, T. L. Frantz, and K. M. Carley. Communication networks from the enron email corpus "it's always about the people. enron is no different". *Comput. Math. Organ. Theory*, 11(3):201–228, 2005. ISSN 1381-298X. DOI: 10.1007/s10588-005-5377-0 6

P. Domingos and M. Richardson. Mining the network value of customers. In *KDD '01: Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 57–66, New York, NY, USA, 2001. ACM. ISBN 1-58113-391-X. DOI: 10.1145/502512.502525 24

Y. Dourisboure, F. Geraci, and M. Pellegrini. Extraction and classification of dense communities in the web. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 461–470, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-654-7. DOI: 10.1145/1242572.1242635 9

D. Easley and J. Kleinberg. *Networks, Crowds, and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010. 19, 27

K. El-Arini, G. Veda, D. Shahaf, and C. Guestrin. Turning down the noise in the blogosphere. In *KDD '09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 289–298, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-495-9. DOI: 10.1145/1557019.1557056 24

G. W. Flake, S. Lawrence, and C. L. Giles. Efficient identification of web communities. In *KDD '00: Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 150–160, New York, NY, USA, 2000. ACM. ISBN 1-58113-233-6. DOI: 10.1145/347090.347121 9, 52

R. W. Floyd. Algorithm 97: Shortest path. *Commun. ACM*, 5(6):345, 1962. ISSN 0001-0782. DOI: 10.1145/367766.368168 18

S. Fortunato. Community detection in graphs. *Physics Reports*, 486(3-5):75 – 174, 2010. ISSN 0370-1573. URL http://www.sciencedirect.com/science/article/B6TVP-4XPYXF1-1/2/99061fac6435db4343b2374d26e64ac1. DOI: 10.1016/j.physrep.2009.11.002 53

D. Gibson, R. Kumar, and A. Tomkins. Discovering large dense subgraphs in massive graphs. In *VLDB '05: Proceedings of the 31st international conference on Very large data bases*, pages 721–732. VLDB Endowment, 2005. ISBN 1-59593-154-6. 9, 36

E. Gilbert and K. Karahalios. Predicting tie strength with social media. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 211–220, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-246-7. DOI: 10.1145/1518701.1518736 18, 20

G. H. Golub and C. F. Van Loan. *Matrix computations (3rd ed.)*. Johns Hopkins University Press, Baltimore, MD, USA, 1996. ISBN 0-8018-5414-8. 18, 71

L. Goodman. Snowball sampling. *The Annals of Mathematical Statistics*, 32(1):148–170, 1961. DOI: 10.1214/aoms/1177705148 94

A. Goyal, F. Bonchi, and L. V. Lakshmanan. Learning influence probabilities in social networks. In *WSDM '10: Proceedings of the third ACM international conference on Web search and data mining*, pages 241–250, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-889-6. DOI: 10.1145/1718487.1718518 21

M. Granovetter. The Strength of Weak Ties. *American Journal of Sociology*, 78(6):1360, 1973. DOI: 10.1086/225469 18

M. Granovetter. Threshold models of collective behavior. *American journal of sociology*, 83(6):1420, 1978. DOI: 10.1086/226707 22

D. Gruhl, R. Guha, D. Liben-Nowell, and A. Tomkins. Information diffusion through blogspace. In *WWW '04: Proceedings of the 13th international conference on World Wide Web*, pages 491–501, New York, NY, USA, 2004. ACM. ISBN 1-58113-844-X. DOI: 10.1145/988672.988739 23

M. S. Handcock, A. E. Raftery, and J. M. Tantrum. Model-based clustering for social networks. *Journal of The Royal Statistical Society Series A*, 127(2):301–354, 2007. URL http://ideas.repec.org/a/bla/jorssa/v170y2007i2p301-354.html. 37

R. Hanneman and M. Riddle. *Introduction to Social Network Methods*. http://faculty.ucr.edu/ hanneman/, 2005. 36

M. Hechter. *Principles of Group Solidarity*. University of California Press, 1988. 8

S. Hill, F. Provost, and C. Volinsky. Network-based marketing: Identifying likely adopters via consumer networks. *Statistical Science*, 21(2):256–276, 2006. DOI: 10.1214/088342306000000222 29

P. D. Hoff, A. E. Raftery, and M. S. Handcock. Latent space approaches to social network analysis. *Journal of the American Statistical Association*, 97(460):1090–1098, 2002. DOI: 10.1198/016214502388618906 37

J. Hopcroft and R. Tarjan. Algorithm 447: efficient algorithms for graph manipulation. *Commun. ACM*, 16(6):372–378, 1973. ISSN 0001-0782. DOI: 10.1145/362248.362272 33

J. Hopcroft, O. Khan, B. Kulis, and B. Selman. Natural communities in large linked networks. In *KDD '03: Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 541–546, New York, NY, USA, 2003. ACM. ISBN 1-58113-737-0. DOI: 10.1145/956750.956816 36, 78

J. Hopcroft, O. Khan, B. Kulis, and B. Selman. Tracking evolving communities in large linked networks. *Proceedings of the National Academy of Sciences of the United States of America*, 101 (Suppl 1):5249–5253, 2004. URL http://www.pnas.org/content/101/suppl.1/5249. abstract. DOI: 10.1073/pnas.0307750100 78, 79

B. A. Huberman, D. M. Romero, and F. Wu. Social networks that matter: Twitter under the microscope. *First Monday*, 14(1), 2009. 20

A. Java, A. Joshi, and T. Finin. Detecting Commmunities via Simultaneous Clustering of Graphs and Folksonomies. In *Proceedings of the Tenth Workshop on Web Mining and Web Usage Analysis (WebKDD)*. ACM, August 2008. (Held in conjunction with The 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2008)). 9

D. B. Johnson. Efficient algorithms for shortest paths in sparse networks. *J. ACM*, 24(1):1–13, 1977. ISSN 0004-5411. DOI: 10.1145/321992.321993 18

D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM New York, NY, USA, 2003. DOI: 10.1145/956750.956769 8, 21, 23, 24, 25, 26

M.-S. Kim and J. Han. A particle-and-density based evolutionary clustering method for dynamic networks. *Proc. VLDB Endow.*, 2(1):622–633, 2009. ISSN 2150-8097. 82

P. Kolari, T. Finin, and A. Joshi. SVMs for the blogosphere: Blog identification and splog detection. In *AAAI Spring Symposium on Computational Approaches to Analyzing Weblogs*, 2006a. 11

P. Kolari, A. Java, and T. Finin. Characterizing the splogosphere. In *Proceedings of the 3rd Annual Workshop on Weblogging Ecosystem: Aggregation, Analysis and Dynamics, 15th World Wid Web Conference*. Citeseer, 2006b. 11

G. Kossinets and D. J. Watts. Empirical analysis of an evolving social network. *Science*, 311(5757): 88–90, 2006. URL http://www.sciencemag.org/cgi/content/abstract/311/5757/88. DOI: 10.1126/science.1116869 76

G. Kossinets, J. Kleinberg, and D. Watts. The structure of information pathways in a social communication network. In *KDD '08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 435–443, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-193-4. DOI: 10.1145/1401890.1401945 21

R. Kumar, P. Raghavan, S. Rajagopalan, and A. Tomkins. Trawling the web for emerging cyber-communities. *Comput. Netw.*, 31(11-16):1481–1493, 1999. ISSN 1389-1286. DOI: 10.1016/S1389-1286(99)00040-7 32

R. Kumar, J. Novak, P. Raghavan, and A. Tomkins. On the bursty evolution of blogspace. *World Wide Web*, 8(2):159–178, 2005. ISSN 1386-145X. DOI: 10.1007/s11280-004-4872-4 75, 76, 78

R. Kumar, J. Novak, and A. Tomkins. Structure and evolution of online social networks. In *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 611–617, New York, NY, USA, 2006. ACM. ISBN 1-59593-339-5. DOI: 10.1145/1150402.1150476 33, 76

T. La Fond and J. Neville. Randomization tests for distinguishing social influence and homophily effects. In *WWW '10: Proceedings of the 19th international conference on World wide web*, pages 601–610, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-799-8. DOI: 10.1145/1772690.1772752 27, 29

J. Leskovec and C. Faloutsos. Sampling from large graphs. In *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 631–636, New York, NY, USA, 2006. ACM. ISBN 1-59593-339-5. DOI: 10.1145/1150402.1150479 8

J. Leskovec and E. Horvitz. Planetary-scale views on a large instant-messaging network. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pages 915–924, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-085-2. DOI: 10.1145/1367497.1367620 5, 6

J. Leskovec, J. Kleinberg, and C. Faloutsos. Graph evolution: Densification and shrinking diameters. *ACM Trans. Knowl. Discov. Data*, 1(1):2, 2007a. ISSN 1556-4681. DOI: 10.1145/1217299.1217301 76

J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. VanBriesen, and N. Glance. Cost-effective outbreak detection in networks. In *KDD '07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 420–429, New York, NY, USA, 2007b. ACM. ISBN 978-1-59593-609-7. DOI: 10.1145/1281192.1281239 8, 24, 26

J. Leskovec, L. Backstrom, R. Kumar, and A. Tomkins. Microscopic evolution of social networks. In *KDD '08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 462–470, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-193-4. DOI: 10.1145/1401890.1401948 76

D. Liben-Nowell and J. Kleinberg. The link-prediction problem for social networks. *J. Am. Soc. Inf. Sci. Technol.*, 58(7):1019–1031, 2007. ISSN 1532-2882. DOI: 10.1002/asi.20591 10, 11

Y.-R. Lin, Y. Chi, S. Zhu, H. Sundaram, and B. L. Tseng. Analyzing communities and their evolutions in dynamic social networks. *ACM Trans. Knowl. Discov. Data*, 3(2):1–31, 2009. ISSN 1556-4681. DOI: 10.1145/1514888.1514891 82

B. Long, Z. M. Zhang, X. Wú, and P. S. Yu. Spectral clustering for multi-type relational data. In *ICML '06: Proceedings of the 23rd international conference on Machine learning*, pages 585–592, New York, NY, USA, 2006. ACM. ISBN 1-59593-383-2. DOI: 10.1145/1143844.1143918 72

B. Long, P. S. Yu, and Z. M. Zhang. A general model for multiple view unsupervised learning. In *SDM '08: Proceedings of SIAM International Conference on Data Mining*, pages 822–833, 2008. 64

U. v. Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, 2007. ISSN 0960-3174. DOI: 10.1007/s11222-007-9033-z 41, 42

S. A. Macskassy and F. Provost. Classification in networked data: A toolkit and a univariate case study. *J. Mach. Learn. Res.*, 8:935–983, 2007. ISSN 1533-7928. 85, 86

S. C. Madeira and A. L. Oliveira. Biclustering algorithms for biological data analysis: A survey. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 1:24–45, 2004. ISSN 1545-5963. DOI: 10.1109/TCBB.2004.2 69

B. McClosky and I. V. Hicks. Detecting cohesive groups. http://www.caam.rice.edu/ ivhicks/CokplexAlgorithmPaper.pdf, 2009. 34

M. McPherson, L. Smith-Lovin, and J. M. Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27:415–444, 2001. DOI: 10.1146/annurev.soc.27.1.415 27

F. Menczer. Web crawling. In B. Liu, editor, *Web Data Mining*, chapter 8, pages 273–322. Springer, 2006. 94

A. Mislove, M. Marcon, K. P. Gummadi, P. Druschel, and B. Bhattacharjee. Measurement and analysis of online social networks. In _IMC '07: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement_, pages 29–42, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-908-1. DOI: 10.1145/1298306.1298311 6

T. M. Mitchell. Mining Our Reality. _Science_, 326(5960):1644–1645, 2009. URL http://www.sciencemag.org. DOI: 10.1126/science.1174459 11

A. A. Nanavati, S. Gurumurthy, G. Das, D. Chakraborty, K. Dasgupta, S. Mukherjea, and A. Joshi. On the structural properties of massive telecom call graphs: findings and implications. In _CIKM '06: Proceedings of the 15th ACM international conference on Information and knowledge management_, pages 435–444, New York, NY, USA, 2006. ACM. ISBN 1-59593-433-2. DOI: 10.1145/1183614.1183678 6

G. Nemhauser, L. Wolsey, and M. Fisher. An analysis of approximations for maximizing submodular set functions-I. _Mathematical Programming_, 14(1):265–294, 1978. DOI: 10.1007/BF01588971 26

M. Newman. Modularity and community structure in networks. _PNAS_, 103(23):8577–8582, 2006a. DOI: 10.1073/pnas.0601602103 43, 53

M. Newman. Finding community structure in networks using the eigenvectors of matrices. _Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)_, 74(3), 2006b. URL http://dx.doi.org/10.1103/PhysRevE.74.036104. DOI: 10.1103/PhysRevE.74.036104 44, 46, 49

M. Newman and M. Girvan. Finding and evaluating community structure in networks. _Physical Review E_, 69:026113, 2004. URL http://www.citebase.org/abstract?id=oai:arXiv.org:cond-mat/0308217. DOI: 10.1103/PhysRevE.69.026113 9, 46, 97

M. Newman, A.-L. Barabasi, and D. J. Watts, editors. _The Structure and Dynamics of Networks_. 2006. 6

J. Onnela, J. Saramäki, J. Hyvönen, G. Szabó, D. Lazer, K. Kaski, J. Kertész, and A. Barabási. Structure and tie strengths in mobile communication networks. _PNAS_, 104(18):7332–7336, 2007. DOI: 10.1073/pnas.0610245104 19

L. Page, S. Brin, R. Motwani, and T. Winograd. The PageRank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999. URL http://ilpubs.stanford.edu:8090/422/. Previous number = SIDL-WP-1999-0120. 8, 16

G. Palla, I. Derényi, I. Farkas, and T. Vicsek. Uncovering the overlapping community structure of complex networks in nature and society. _Nature_, 435:814–818, 2005. DOI: 10.1038/nature03607 33, 78

G. Palla, A.-L. Barabasi, and T. Vicsek. Quantifying social group evolution. *Nature*, 446(7136): 664–667, April 2007. DOI: 10.1038/nature05670 9, 76, 77, 78

J. Pearl. *Causality: models, reasoning, and inference*. Cambridge University Press, 2000. 28

M. Ramezani, J. Sandvig, R. Bhaumik, R. Burke, and B. Mobasher. Exploring the impact of profile injection attacks in social tagging systems. In *Proceedings of Workshop on Web Mining and Web Usage Analysis*, 2008. 11

M. Richardson and P. Domingos. Mining knowledge-sharing sites for viral marketing. In *KDD '02: Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 61–70, New York, NY, USA, 2002. ACM. ISBN 1-58113-567-X. DOI: 10.1145/775047.775057 8, 24

K. Saito, M. Kimura, K. Ohara, and H. Motoda. Behavioral analyses of information diffusion models by observed data of social network. In *SBP*, pages 149–158, 2010. DOI: 10.1007/978-3-642-12079-4_20 21

P. Sarkar and A. W. Moore. Dynamic social network analysis using latent space models. *SIGKDD Explor. Newsl.*, 7(2):31–40, 2005. ISSN 1931-0145. DOI: 10.1145/1117454.1117459 37, 82

T. C. Schelling. Dynamic models of segregation. *Journal of Mathematical Sociology*, 1:143–186, 1971. DOI: 10.1080/0022250X.1971.9989794 22

P. Sen, G. Namata, M. Bilgic, L. Getoor, B. Galligher, and T. Eliassi-Rad. Collective classification in network data. *AI Magazine*, 29(3):93, 2008. 85

C. Shirky. *Here Comes Everybody: The Power of Organizing without Organizations*. The Penguin Press, 2008. 1

P. Singla and M. Richardson. Yes, there is a correlation: - from social networks to personal behavior on the web. In *WWW '08: Proceeding of the 17th international conference on World Wide Web*, pages 655–664, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-085-2. DOI: 10.1145/1367497.1367586 26

A. Strehl and J. Ghosh. Cluster ensembles — a knowledge reuse framework for combining multiple partitions. *J. Mach. Learn. Res.*, 3:583–617, 2003. ISSN 1533-7928. DOI: 10.1162/153244303321897735 49, 65

J. Sun, C. Faloutsos, S. Papadimitriou, and P. S. Yu. Graphscope: parameter-free mining of large time-evolving graphs. In *KDD '07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 687–696, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-609-7. DOI: 10.1145/1281192.1281266 82, 83

Y. Sun, Y. Yu, and J. Han. Ranking-based clustering of heterogeneous information networks with star network schema. In *KDD '09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 797–806, New York, NY, USA, 2009. ACM. ISBN 978-1-60558-495-9. DOI: 10.1145/1557019.1557107 56

P.-N. Tan, M. Steinbach, and V. Kumar. *Introduction to Data Mining*. Addison Wesley, 2005. 36, 37, 49, 85, 101

L. Tang. *Learning with Large-Scale Social Media Networks*. PhD thesis, Arizona State University, 2010. URL http://www.public.asu.edu/~ltang9/thesis.pdf. 52, 58

L. Tang and H. Liu. Scalable learning of collective behavior based on sparse social dimensions. In *CIKM '09: Proceeding of the 18th ACM conference on Information and knowledge management*, pages 1107–1116, New York, NY, USA, 2009a. ACM. ISBN 978-1-60558-512-3. DOI: 10.1145/1645953.1646094 91

L. Tang and H. Liu. Relational learning via latent social dimensions. In *KDD '09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 817–826, New York, NY, USA, 2009b. ACM. ISBN 978-1-60558-495-9. DOI: 10.1145/1557019.1557109 57, 88

L. Tang and H. Liu. Toward predicting collective behavior via social dimension extraction. *IEEE Intelligent Systems*, 25:19–25, 2010a. ISSN 1541-1672. DOI: 10.1109/MIS.2010.36 89, 90

L. Tang and H. Liu. Graph mining applications to social network analysis. In C. Aggarwal and H. Wang, editors, *Managing and Mining Graph Data*, chapter 16, pages 487–513. Springer, 2010b. DOI: 10.1007/978-1-4419-6045-0_16 31

L. Tang, H. Liu, J. Zhang, and Z. Nazeri. Community evolution in dynamic multi-mode networks. In *KDD '08: Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 677–685, New York, NY, USA, 2008. ACM. ISBN 978-1-60558-193-4. DOI: 10.1145/1401890.1401972 9, 49, 72

L. Tang, X. Wang, and H. Liu. Uncovering groups via heterogeneous interaction analysis. In *ICDM '09: Proceedings of IEEE International Conference on Data Mining*, pages 503–512, 2009. 9, 46, 55, 64, 67

L. Tang, X. Wang, H. Liu, and L. Wang. A multi-resolution approach to learning with overlapping communities. In *Proceedings of Workshop on Social Media Analytics*, 2010. 92

L. Tang, H. Liu, and J. Zhang. Identifying evolving groups in dynamic multi-mode networks. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, forthcoming. 9, 56, 73

C. Tantipathananandh, T. Berger-Wolf, and D. Kempe. A framework for community identification in dynamic social networks. In *KDD '07: Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 717–726, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-609-7. DOI: 10.1145/1281192.1281269 82

M. Thelwall. Bloggers under the london attacks:top information sources and topics. In *WWW:3rd annual workshop on webloging ecosystem: aggreation, analysis and dynamics*, 2006. 2

J. Travers and S. Milgram. An experimental study of the small world problem. *Sociometry*, 32(4): 425–443, 1969. DOI: 10.2307/2786545 5

K. Wakita and T. Tsurumi. Finding community structure in mega-scale social networks: [extended abstract]. In *WWW '07: Proceedings of the 16th international conference on World Wide Web*, pages 1275–1276, New York, NY, USA, 2007. ACM. ISBN 978-1-59593-654-7. DOI: 10.1145/1242572.1242805 48

S. Wasserman and K. Faust. *Social Network Analysis: Methods and Applications*. Cambridge University Press, 1994. 3, 5, 8, 13, 34, 68

D. J. Watts and P. S. Dodds. Influentials, networks, and public opinion formation. *Journal of Consumer Research*, 34(4):441–458, 2007. DOI: 10.1086/518527 26

D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393:440–442, 1998. DOI: 10.1038/30918 7, 8

R. Xiang, J. Neville, and M. Rogati. Modeling relationship strength in online social networks. In *WWW '10: Proceedings of the 19th international conference on World wide web*, pages 981–990, New York, NY, USA, 2010. ACM. ISBN 978-1-60558-799-8. DOI: 10.1145/1772690.1772790 20

T. Yang, Y. Chi, S. Zhu, Y. Gao, and R. Jin. A bayesian approach toward finding communities and their evolutions in dynamic social networks. In *SDM '09: Proceedings of SIAM International Conference on Data Mining*, 2009. 82

H.-J. Zeng, Z. Chen, and W.-Y. Ma. A unified framework for clustering heterogeneous web objects. In *WISE '02: Proceedings of the 3rd International Conference on Web Information Systems Engineering*, pages 161–172, Washington, DC, USA, 2002. IEEE Computer Society. ISBN 0-7695-1766-8. DOI: 10.1109/WISE.2002.1181653 9

X. Zhu, Z. Ghahramani, and J. Lafferty. Semi-supervised learning using gaussian fields and harmonic functions. In *ICML '03: Proceedings of the Twentieth International Conference on Machine Learning*, pages 912–919, 2003. 87