Clustering and Blockmodeling

Vladimir Batagelj
University of Ljubljana, Slovenia

June 28, 2002

References

[1] Aarts, E. & Lenstra, J.K. (Eds.)(1997). *Local Search in Combinatorial Optimization*. Wiley, Chichester.

[2] Anderberg, M.R. (1973), Cluster analysis for applications. Probability and mathematical statistics, vol.19. Academic Press, New York.

[3] Batagelj, V. (1981), Note on ultrametric hierarchical clustering algorithms. *Psychometrika*, **46**(3), 351-352.

[4] Batagelj, V. (1984), Notes on the dynamic clusters method, in *Proceedings of the 4th Conference on Applied Mathematics 1984*, Split, University of Split, 1985, 139-146.

[5] Batagelj, V. (1985), Algorithmic aspects of clustering problem, Proceedings of 7th International Symposium "Computer at the University", Cavtat 1985, SRCE, Zagreb, p. 502 1 - 502 15.

[6] Batagelj, V. (1986), On adding clustering algorithms. COMPSTAT 86, short communications, Rome.

[7] Batagelj, V. (1988), Generalized Ward and related clustering problems. In: Bock, H.H. (Ed.): *Classification and related methods of data analysis*, North-Holland, Amsterdam, p. 67–74.

[8] Batagelj, V. (1989), *Similarity measures between structured objects*, in A. Graovac (Ed.), Proceedings MATH/CHEM/COMP 1988, Dubrovnik, Yugoslavia 20-25 June 1988, Studies in Physical and Theoretical Chemistry. Vol 63, pp. 25-40, Amsterdam: Elsevier.

[9] Batagelj, V., Pisanski, T., and Simões-Pereira, J.M.S. (1990), An algorithm for tree-realizability of distance matrices. *International Journal of Computer Mathematics*, **34**, pp. 171-176.

[10] Batagelj, V., Ferligoj, A., and Doreian, P. (1992), Direct and Indirect Methods for Structural Equivalence. *Social Networks* **14**, 63-90.

[11] Batagelj, V., Doreian, P., and Ferligoj, A. (1992), An Optimizational Approach to Regular Equivalence. *Social Networks* **14**, 121-135.

[12] Batagelj, V., Korenjak-Černe, S., and Klavžar, S. (1994) Dynamic Programming and Convex Clustering. *Algorithmica* **11**, 93-103.

[13] Batagelj, V. (1994), Semirings for Social Networks Analysis. *Journal of Mathematical Socilogy*, **19**1, 53-68.

[14] Batagelj, V., and Bren, M. (1995), Comparing Resemblance Measures, *Journal of Classification,* 12(1), 73-90.

[15] Batagelj, V. (1997), Notes on blockmodeling. *Social Networks* **19**, 143-155.

[16] Batagelj, V., Mrvar, A. and Zaveršnik, M. (1999), Partitioning approach to visualization of large graphs, In Kratochvíl, J. (ed), *Lecture notes in computer science*, 1731, Springer, Berlin, 90–97.

[17] Batagelj, V., and Mrvar, A. (2001), A subquadratic triad census algorithm for large sparse networks with small maximum degree. *Social networks*, **23**(3), 237-243.

[18] Batagelj, V. and Zaveršnik, M. (2002) Generalized cores, submitted.

[19] Batagelj, V. and Zaveršnik, M. (2002) Triangular connectivity and its generalizations, in preparation.

[20] Benzécri, J.P. (1976), *L'analyse des donnees, Tome 1., 2.*, Paris: Dunod.

[21] Bock, H.H. and Diday, E., (Eds.) (2000), *Analysis of Symbolic Data.* Exploratory methods for extracting statistical information from complex data. Springer, Heidelberg.

[22] Borgatti, S.P., and Everett, M.G. (1989), The Class of all Regular Equivalences: Algebraic Structure and Computation, Social Networks, **11**, 65-88.

[23] Borgatti, S.P., and Everett, M.G. (1992). Notions of positions in social network analysis. In P. V. Marsden (Ed.), *Sociological Methodology* (pp. 1–35). San Francisco: Jossey-bass.

[24] Boyd, J.P. (1991). *Social Semigroups. A Unified Theory of Scaling and Blockmodelling as Applied to Social Networks*. Fairfax: George Mason University Press.

[25] Brigham, R.C., and Dutton, R.D. (1985). A compilation of relations between graph invariants. *Networks*, *15*, 73–107.

[26] Brucker, P. (1978), On the complexity of clustering problems, in *Optimization and Operations Research*, Proceedings, Bonn 1977, in Lecture Notes in Economics and Mathematical Systems, Vol. 157, eds. R.Henn, B. Korte and W. Oettli, Berlin: Springer-Verlag.

[27] Bruynooghe, M. (1977), Méthodes nouvelles en classification automatique des données taxinomiques nombreuses. Statistique et Analyse des Données, **3**, 24-42.

[28] Carré, B. (1979). *Graphs and Networks*. Oxford: Clarendon.

[29] E. Chávez, G. Navarro, R. Baeza-Yates, and J. Marroqu'in (1999). Searching in Metric Spaces. Technical Report TR/DCC-99-3, Dept. of Computer Science, Univ. of Chile.
`http://citeseer.nj.nec.com/avez99searching.html`

[30] De Nooy, W., Mrvar A., Batagelj, V. (2002), Exploratory Social Network Analysis with Pajek. Manuscript of a textbook.

[31] Degenne, A., and Forseé, M. (1999), *Introducing social networks*, London: Sage.

[32] Dickerson M.T., and Eppstein D. (1996), Algorithms for proximity problems in higher dimensions. Comp. Geom. Theory & Applications **5**, 277-291.
`http://www.ics.uci.edu/~eppstein/pubs/a-dickerson.html`

[33] Diday, E. et al. (1979), *Optimisation en classification automatique, Tome 1., 2.*, Rocquencourt: INRIA.

[34] Dieudonne, J. (1960), Foundations of modern analysis. Academic Press, New York.

[35] Doreian, P., Batagelj, V., Ferligoj, A. (2000), Symmetric-acyclic decompositions of networks. *J. classif.*, **17**(1), 3-28.

[36] Doreian, P., Batagelj, V., Ferligoj, A. (2001), Positional analyses of sociometric data. Chapter for the book: Models and methods in social network analysis. New York: Cambridge University Press (forthcoming).

[37] Doreian, P., Batagelj, V., Ferligoj, A. (2002), Generalized blockmodeling. Manuscript of a book.

[38] Doreian, P., Fujimoto K. (2001), Structures of Supreme Court Voting. University of Pittsburgh, manuscript, version November 3, 2001.

[39] Doreian, P., and Mrvar, A. (1996). A partitioning approach to structural balance. *Social Networks*, *18*, 149–168.

[40] Esfahanian A-H., On the Evolution of Graph Connectivity Algorithms. `http://www.cse.msu.edu/~esfahani/book_chapter/Graph_connectivity_chapter.pdf`

[41] Everett, M.G. (1983), EBLOC: A graph theoretic blocking algorithm for social networks. Social Networks **5**, 323-346.

[42] Everett, M.E and Borgatti, S.P. (1996), Exact Colorations of Graphs and Digraphs. *Social Networks*, *18*, 319-331.

[43] Everitt, B. (1974), Cluster Analysis. Heinemann Educational Books LTD., London.

[44] Fuast, K. (1988), Comparison of Methods for Positional Analysis: Structural and General Equivalences, *Social Networks*, *10*, 313-341.

[45] Ferligoj A., Batagelj V. (1982), Clustering with relational constraint. *Psychometrika*, **47**(4), 413-426.

[46] Ferligoj A., Batagelj V. (1983), Some types of clustering with relational constraints. *Psychometrika*, **48**(4), 541-552.

[47] Ferligoj, A., and Batagelj, V. (1992), Direct Multicriteria Clustering Algorithms. *Journal of Classification*, **9**(1), 43-61.

[48] Fisher, W.D. (1958), On grouping for maximum homogeneity. Journal of the American Statistical Association, **53**, 789-798.

[49] Fukunaga, K., Narendra, P.M. (1975), A branch and bound algorithm for computing k-nearest neighbors. *IEEE Transactions on Computers*, **C-24**, 750-753.

[50] Ganti V. etal. (1998), Clustering large datasets in arbitrary metric spaces. http://www.cs.wisc.edu/ vganti/pubs.html

[51] Garey, M.R. and Johnson, D.S. (1979), *Computer and intractability*, San Francisco: Freeman.

[52] Godehardt, E. (1988), *Graphs as Structural Models*, Brounschweig: Vieweg.

[53] Gordon, A.D.(1981), Classification. Monographs on Applied Probability and Statistics. Chapman and Hall, London.

[54] Gordon, A.D. (1996). A survey of constrained classification. *Computational Statistics & Data Analysis*, *21*, 17–29.

[55] Gower, J.C.(1971), A general coefficient of similarity and some of its properties. Biometrics **27**, 857-871.

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

[57] Harel, D., and Koren, J. (2001), On Clustering Using Random Walks, LNCS 2245, Berlin: Springer-Verlag, 18-41.

[58] Hartigan, J.A. (1975), *Clustering Algorithms*, New York: John-Wiley.

[59] Hubálek, Z. (1982), "Coefficients of association and similarity, based on binary (presence-absence) data: an evaluation", *Biological Review 57*, 669-689.

[60] Hummon, N.P. & Doreian, P. (1989). Connectivity in a citation network: The development of dna theory. *Social Networks*, **11**, 39–63.

[61] Jambu M., Lebeaux M-O. (1983), Cluster analysis and data analysis, North-Holland, Amsterdam.

[62] Janowitz, M.F. (1978). An order theoretic model for cluster analysis. *SIAM J. Appl. Math.*, **34**, 55–72.

[63] Jardin N., Sibson R.(1971), Mathematical taxonomy. John Wiley, London.

[64] Jensen R.E. (1969), A dynamic programming algorihm for cluster analysis. Operational research, **7**, 1034-1057.

[65] Kamvar, S.D., and Klein, D., and Manning, C.D. (2002), Interpreting and Extending Classical Agglomerative Clustering Algorithms using a Model-Based Approach. ICML 2002.
`http://nlp.stanford.edu/~danklein/project-clustering.shtml`

[66] Kashyap R.L., Oommen B.J. (1983), A common basis for similarity measures involving two strings. International Journal of Computer Mathematics 13, 17-40.

[67] Korenjak-Černe S. and Batagelj V. (1998), Clustering large datasets of mixed units, in Rizzi, A., Vichi, M., Bock, H.-H. (Eds.): *Advances in Data Science and Classification*. Springer, Berlin, p. 43–48.

[68] Levenshtein V.I. (1966), Binary codes capable of correcting deletions, insertions, and reversals. Soviet Physics - Doklady, 10, 707-710.

[69] Lorrain, F. and H.C. White (1971): Structural equivalence of individuals in social networks. *Journal of Mathematical Sociology*, 1:49–80.

[70] Lorrain, F. (1975): Réseaux sociaux et classifications sociales. Paris: Hermann.

[71] Marriott, F.H.C. (1982), Optimization methods of cluster analysis, *Biometrika*, **69**, 417-421.

[72] Matula D.W.(1977) Graph theoretic techniques for cluster analysis algorithms. Van Ryzin J. (ed.): Classification and clustering. Academic Press, New York, 95-127.

[73] Mirkin, B.G. (1980), *Analiz kačestvennyh priznakov i struktur*, Moskva: Statistika.

[74] Mirkin, B.G., and Rodin C.N. (1977), *Grafy i geny*, Moskva: Nauka.

[75] Moody, J. (2001), Peer influence groups: identifying dense clusters in large networks. *Social Networks* **23**, 261-283.

[76] Moreno, J.L. (1953), *Who shall survive?*, Beacon N.Y.: Beacon House.

[77] Murtagh, F. (1985), Multidimensional Clustering Algorithms, *Compstat lectures*, **4**, Vienna: Physica-Verlag.

[78] Murtagh, F. (1999), Clustering in Massive Data Sets, Proceedings of the Beilstein Workshop 2000: Chemical Data Analysis in the Large, May 22nd - 26th 2000, Bozen, Italy, p. 28-51
http://www.beilstein-institut.de/bozen2000/proceedings/contents.html

[79] Olson, C.F. (1995), Parallel Algorithm for Hierarchical Clustering, Parallel computing, **21**, 1313-1325.
http://faculty.washington.edu/cfolson/papers.html

[80] Pattison, P.E. (1988), Network Models; Some Comments on Papers in this Special Issue, *Social Networks*, *10*, 383–411.

[81] Pattison, P.E. (1993). *Algebraic Models for Social Networks*. Cambridge, England: Cambridge University Press.

[82] Roberts, F.S. (1976), Discrete mathematical models with applications to social, biological, and environmental problems. Englewood Cliffs, N.J.: Prentice Hall.

[83] Schreider, Ju.A. (1975), *Equality, resemblance, and order*, Moskva: Mir.

[84] Scott, J. (2000), Social Network Analysis: A Handbook, 2nd edition. London: Sage Publications.

[85] Seidman, S.B., (1983), Network Structure And Minimum Degree. *Social Networks*, **5**, 269-287.

[86] Shamos, M.I. (1976), Geometry and statistics: Problems at the interface, in J. Traub (Ed.): *Algorithms and Complexity* (New directions and recent results). New York: Academic Press, 251-288.

[87] Sneath P.H.A., Sokal R.R. (1973), Numerical taxonomy. W.H. Freeman and Company, San Francisco.

[88] Snyder, D., Kick, E. (1979), Structural position in the world system and economic growth 1955-70: A multiple network analysis of transnational interactions. *American Journal of Sociology* **84**, 1096-1126.

[89] Späth H. (1977), Cluster Analyse Algorithmen zur Objekt-Klassifizierung und Datenreduction. R. Oldenbourg, München.

[90] Trinajstić, N. (1983), Chemical graph theory, Vol. 2. CRC Press, Boca Raton.

[91] Van Cutsem, B., ed. (1994), *Classification and Dissimilarity Analysis,* Lecture Notes in Statistics, New York: Springer Verlag.

[92] Ward, J.H. (1963), Hierarchical grouping to optimize an objective function, *American Statistical Association Journal*, **58**, 236-244.

[93] Wasserman, S. & Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

[94] White, D.R. and K.P. Reitz (1983), Graph and semigroup homomorphisms on networks of relations. *Social Networks*, 5:193–234.

[95] Index of clustering papers on the WWW:
`http://www.ece.nwu.edu/~harsha/Clustering/clus.html`