Liu, Xuqing; **Liu, Xinsheng**

Structure learning of Bayesian networks by continuous particle swarm optimization algorithms. (English) Zbl 07192618

J. Stat. Comput. Simulation 88, No. 8, 1528-1556 (2018).

Summary: In this paper, the problem of learning Bayesian network (BN) structures is studied by virtue of particle swarm optimization (PSO) algorithms. After analysing the optimal flying behaviours of some classic PSO algorithms, we put forward a new PSO-based method of learning BN structures. In this method, we treat the position of a particle as an imaginary likelihood that represents to what extent the associated edges exist, treat the velocity as the corresponding increment or decrement of likelihood that represents how the position changes in the process of flying, and treat the BN structures outputted as appendants of positions. The resulting algorithm and its improved version with expert knowledge integrated are illustrated to be efficient in collecting the randomly searched information from all particles. The numerical study based on two bechmarking BNs shows the superiority of our algorithms in the sense of precision, speed, and accuracy.

MSC:

62 Statistics

Keywords:

Bayesian network; structure learning; (conditional) mutual information; (continuous) particle swarm optimization

Software:

BNT

Full Text: DOI

References:

[1] Pearl J. Probabilistic reasoning in intelligent systems: networks of plausible inference. San Francisco (CA): Morgan Kaufmann; 1988. [Google Scholar] · Zbl 0746.68089

[2] Neapolitan RE. Learning Bayesian networks. Upper Saddle River (NJ): Prentice Hall; 2004. [Google Scholar]

[3] Zhang L, Guo H. Introduction to Bayesian networks. Beijing: Science Press; 2006. [Google Scholar]

[4] Chickering DM.Learning Bayesian networks is NP-complete. In: Fisher D, Lenz HJ, editors. Learning from data. New York: Springer-Verlag; 1996. p. 121-130. [Crossref], [Google Scholar]

[5] Robinson RW.Counting unlabeled acyclic digraphs. In: Little CHC, editor. Combinatorial mathematics V. Lecture Notes in Statistics, vol. 112. New York: Springer; 1977. p. 28-43. [Crossref], [Google Scholar] · Zbl 0376.05031

[6] Cooper GF, Herskovits E.A Bayesian method for the induction of probabilistic networks from data. Mach Learn. 1992;9(4):309-347. [Crossref], [Web of Science ®], [Google Scholar] · Zbl 0766.68109

[7] Alcobé JR. Incremental hill-climbing search applied to Bayesian network structure learning. First International Workshop on Knowledge Discovery in Data Streams; Pisa; 2004. [Google Scholar]

[8] Cheng J, Greiner R, Kelly J, et al. Learning Bayesian networks from data: an information-theory based approach. Artif Intell. 2002;137(1-2):43-90. doi: 10.1016/S0004-3702(02)00191-1[Crossref], [Web of Science ®], [Google Scholar] · Zbl 0995.68114

[9] Aliferis CF, Statnikov A, Tsamardinos I, et al. Local causal and Markov blanket induction for causal discovery and feature selection for classification part I: algorithms and empirical evaluation. J Mach Learn Res. 2010;11:171-234. [Web of Science ®], [Google Scholar] · Zbl 1242.68197

[10] Brown LE, Tsamardinos I, Aliferis CF. A novel algorithm for scalable and accurate Bayesian network learning. Eleventh World Congress on Medical Informatics; San Francisco; 2004. p. 711-715. [Google Scholar]

[11] Tsamardinos I, Brown LE, Aliferis CF.The max-min hill-climbing Bayesian network structure learning algorithm. Mach Learn. 2006;65(1):31-78. doi: 10.1007/s10994-006-6889-7[Crossref], [Web of Science ®], [Google Scholar]

[12] Larrañaga P, Poza M, Yurramendi Y, et al. Structure learning of Bayesian networks by genetic algorithms: performance analysis of control parameters. IEEE Trans Pattern Anal Mach Intell. 1996;18(9):912-926. doi: 10.1109/34.537345[Crossref],

[Web of Science ®], [Google Scholar]

[13] Zhang L, Zhang J.Structure learning of BN based on improved genetic algorithm. Comput Syst Appl. 2011;20(9):68-72. [Google Scholar]

[14] Janžura M, Nielsen J.A simulated annealing-based method for learning Bayesian networks from statistical data. Int J Intell Syst. 2006;21(3):335-348. doi: 10.1002/int.20138[Crossref], [Web of Science ®], [Google Scholar] · Zbl 1101.62015

[15] Wang C, Wang Y.Constrained simulated annealing method for Bayesian network structure learning. J Henan Norm Univ Nat Sci. 2013;41(2):6-9. [Google Scholar] · Zbl 1289.68098

[16] Du T, Zhang SS, Wang Z. Efficient learning Bayesian networks using pso. Berlin: Springer Berlin Heidelberg; 2005. p. 151-156. doi: 10.1007/11596448_22[Crossref], [Google Scholar]

[17] Du T, Zhang S, Wang Z.Learning Bayesian networks from data by particle swarm optimization. J Shanghai Jiaotong Univ Sci. 2006;E-11(4):423-429. [Google Scholar] · Zbl 1113.68435

[18] Sahin F, Devasia A.Distributed particle swarm optimization for structural Bayesian network learning. In: Chan FTS, Tiwari MK, editors. Swarm intelligence: focus on ant and particle swarm optimization. Vienna: I-Tech Education and Publishing; 2007. p. 505-532. [Crossref], [Google Scholar]

[19] Wang T, Yang J.A heuristic method for learning Bayesian networks using discrete particle swarm optimization. Knowl Inf Syst. 2010;24(2):269-281. doi: 10.1007/s10115-009-0239-6[Crossref], [Web of Science ®], [Google Scholar]

[20] Wang C, Lü J.Hybrid particle swarm algorithm for learning Bayesian network structure. Sci Technol Rev. 2013;31(22):50-55. [Google Scholar]

[21] Santos FP, Maciel CD. A PSO approach for learning transition structures of higher-order dynamic Bayesian networks. Biosignals and Biorobotics Conference (2014): Biosignals and Robotics for Better and Safer Living (BRC), 5th ISSNIP-IEEE; Salvador; 2014. p. 1-6. [Google Scholar]

[22] Schlüter F.A survey on independence-based Markov networks learning. Artif Intell Rev. 2012. Available from: http://link.springer.com/article/10.10 012-9346-y[Web of Science ®], [Google Scholar]

[23] Zhang HY, Wang LW, Chen YX.Research progress of probabilistic graphical models: a survey. J Softw. 2013;24(11):2476-2497. doi: 10.3724/SP.J.1001.2013.04486[Crossref], [Google Scholar] · Zbl 1299.68091

[24] Liu J, Li H, Luo X.Learning technique of probabilistic graphical models: a review. Acta Automat Sin. 2014;40(6):1025-1044. [Google Scholar] · Zbl 1313.68108

[25] Kennedy J, Eberhart RC. Particle swarm optimization. Proceedings of 1995 IEEE International Conference on Neural Networks; Vol. 4; Perth; 1995. p. 1942-1948. [Google Scholar]

[26] Kennedy J, Eberhart RC. A discrete binary version of the particle swarm algorithm. IEEE International Conference on Systems, Man, and Cybernetics; Vol. 5; Orlando; 1997. p. 4104-4108. [Google Scholar]

[27] Di RH, Gao XG.Bayesian network structure learning based on restricted particle swarm optimization. Syst Eng Electronics. 2011;33(11):2423-2427. [Google Scholar] · Zbl 1265.68133

[28] Heckerman D, Geiger D, Chickering DM.Learning Bayesian networks: the combination of knowledge and statistical data. Mach Learn. 1995;20(3):197-243. [Crossref], [Web of Science ®], [Google Scholar] · Zbl 0831.68096

[29] Schwarz G.Estimating the dimension of a model. Ann Statist. 1978;6(2):461-464. doi: 10.1214/aos/1176344136[Crossref], [Web of Science ®], [Google Scholar] · Zbl 0379.62005

[30] Akaike H.Information theory and an extension of the maximum likelihood principle. In: Parzen E, Tanabe K, Kitagawa G, editors. Selected papers of Hirotugu Akaike Springer Series in Statistics (Perspectives in Statistics). New York: Springer; 1998. p. 199-213. [Crossref], [Google Scholar]

[31] Draper D.Assessment and propagation of model uncertainty (with discussion). J R Stat Soc Ser B. 1995;57(1):45-97. [Google Scholar] · Zbl 0812.62001

[32] Rahman MS, King ML.Improved model selection criterion. Comm Stat Simul Comput. 1999;28(1):51-71. doi: 10.1080/03610919908813535[Taylor \& Francis Online], [Web of Science ®], [Google Scholar] · Zbl 1100.62585

[33] Lanterman AD.Schwarz, Wallace, and Rissanen: Intertwining themes in theories of model selection. Int Stat Rev. 2001;69(2):185-212. doi: 10.1111/j.1751-5823.2001.tb00456.x[Crossref], [Web of Science ®], [Google Scholar] · Zbl 1211.62007

[34] Rahman MS.Model selection criteria in residual sum of squares form. Int Rev Bus Res Pap. 2010;6(4):501-511. [Google Scholar]

[35] Heng XC, Qin Z, Wang XH, et al. Research on learning Bayesian networks by particle swarm optimization. Inform Technol J. 2006;5(3):540-545. doi: 10.3923/itj.2006.540.545[Crossref], [Google Scholar]

[36] Li XL, Wang SC, He XD.Learning Bayesian networks structures based on memory binary particle swarm optimization. In: Wang TD, Li X, Wang X, editors. Simulated evolution and learning. Lecture Notes in Computer Science, vol 4247. Berlin: Springer; 2006. p. 568-574. [Crossref], [Google Scholar]

[37] Du ZH, Wang YW.A novel structure learning method for constructing gene regulatory network. J Comput Appl. 2009;29(6):1539-1543. [Google Scholar]

[38] Aouay S, Jamoussi S, BenAyed Y. Particle swarm optimization based method for Bayesian network structure learning. Fifth International Conference on Modeling, Simulation and Applied Optimization (ICMSAO), 2013; Hammamet; 2013. p. 1-6. [Google Scholar]

[39] de Campos LM.A scoring function for learning Bayesian networks based on mutual information and conditional independence tests. J Mach Learn Res. 2006;7:2149-2187. [Web of Science ®], [Google Scholar] · Zbl 1222.62036

[40] Martínez-Rodríguez AM, May JH, Vargas LG.An optimization-based approach for the design of Bayesian networks. Math Comput Model. 2008;48(7-8):1265-1278. doi: 10.1016/j.mcm.2008.01.007[Crossref], [Google Scholar] · Zbl 1187.90086

[41] Lee CP, Leu Y, Yang WN.Constructing gene regulatory networks from microarray data using GA/PSO with DTW. Appl Soft Comput. 2012;12(3):1115-1124. doi: 10.1016/j.asoc.2011.11.013[Crossref], [Web of Science ®], [Google Scholar]

[42] Shi Y, Eberhart R A modified particle swarm optimizer. Proceedings of the IEEE International Conference on Evolutionary Computation, Anchorage, AK, USA; IEEE; 1998. p. 69-73 [Google Scholar]

[43] Lepar V, Shenoy PP A comparison of Lauritzen-Spiegelhalter, Hugin, and Shenoy-Shafer architectures for computing marginals of probability distributions. Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence. San Francisco, CA: Morgan Kaufmann; 1998. p. 328-337 [Google Scholar]

[44] Beinlich IA, Suermondt HJ, Chavez RM, et al. The ALARM monitoring system: a case study with two probabilistic inference techniques for belief networks. In: Second European conference on artificial intelligence in medicine. London: Springer-Verlag; 1989. p. 247-256. [Google Scholar]

[45] Chickering DM, Meek C.On the incompatibility of faithfulness and monotone DAG faithfulness. Artif Intell. 2006;170(8):653-666. doi: 10.1016/j.artint.2006.03.001[Crossref], [Web of Science ®], [Google Scholar] · Zbl 1131.68082

[46] Poli R, Kennedy J, Blackwell T.Particle swarm optimization – an review. Swarm Intell. 2007;1(1):33-57. doi: 10.1007/s11721-007-0002-0[Crossref], [Google Scholar]

[47] Bollobás B. Random graphs. 2nd ed. Cambridge: Cambridge University Press; 2001. [Crossref], [Google Scholar] · Zbl 0979.05003

[48] Murphy KP. Bayes net toolbox for Matlab. Version: FullBNT-1.0.7; 2007. [Google Scholar]

[49] Brown G, Pocock A, Zhao MJ, et al. Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. J Mach Learn Res. 2012;13:27-66. (Version: MIToolbox-2.0). [Web of Science ®], [Google Scholar] · Zbl 1283.68283

[50] Pellet JP, Elisseeff A.Using Markov blankets for causal structure learning. J Mach Learn Res. 2008;9:1295-1342. [Web of Science ®], [Google Scholar] · Zbl 1225.68205