Information Sciences 348 (2016) 272–289

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

BNC-PSO: structure learning of Bayesian networks by Particle Swarm Optimization

S. Gheisari [a,*], M.R. Meybodi [b]

[a] Department of Computer Engineering, Pardis Branch, Islamic Azad University, Pardis, Iran
[b] Computer Engineering and Information Technology Department, Amirkabir University of Technology, Tehran, Iran

A R T I C L E I N F O

Article history:
Received 15 July 2015
Revised 21 January 2016
Accepted 30 January 2016
Available online 18 February 2016

Keywords:
Bayesian Information Criteria (BIC)
Bayesian Network (BN)
Cross over operation
Mutation operation
Particle Swarm Optimization (PSO)
Structure learning

A B S T R A C T

Structure learning is a very important problem in the field of Bayesian networks (BNs). It is also an active research area for more than 2 decades; therefore, many approaches have been proposed in order to find an optimal structure based on training samples. In this paper, a Particle Swarm Optimization (PSO)-based algorithm is proposed to solve the BN structure learning problem; named BNC-PSO (**B**ayesian **N**etwork **C**onstruction algorithm using **PSO**). Edge inserting/deleting is employed in the algorithm to make the particles have the ability to achieve the optimal solution, while a cycle removing procedure is used to prevent the generation of invalid solutions. Then, the theorem of Markov chain is used to prove the global convergence of our proposed algorithm. Finally, some experiments are designed to evaluate the performance of the proposed PSO-based algorithm. Experimental results indicate that BNC-PSO is worthy of being studied in the field of BNs construction. Meanwhile, it can significantly increase nearly 15% in the scoring metric values, comparing with other optimization-based algorithms.

1. Introduction

Bayesian networks (BNs) are popular within the AI probability and uncertainty community as a method of reasoning under uncertainty [48]. From an informal perspective, a BN is a directed acyclic graph, in which nodes represent random variables, and existence or lack of the arcs represents the dependence relationships between the variables. The relations are further quantified by a set of conditional probability distributions, one for each variable conditioning on its parents. Overall, a BN represents a joint probability distribution over a set of random variables; and provides an efficient device for performing probabilistic inference.

Learning the structure of a BN from data is an important challenge and has been studied extensively during 2 last decades. It is an *NP-hard* problem [9,11,30,54]; so, inferring complete causal models (i.e., causal BNs) is essentially impossible in large-scale data mining applications with thousands of variables [78].

Generally, there are three main approaches for learning BNs from data: scored-based learning, constraint-based learning, and hybrid methods. Algorithms following the first approach evaluate the quality of BNs structures using a scoring function and select the one with the best score [2,13,32]. These methods consider the structure learning problem as a combinatorial optimization problem. However they work well for small datasets, they may fail to find optimal solutions for large datasets. Second group of algorithms typically use statistical tests to identify conditional independence relations from data and build

* Corresponding author. Tel.: +98 9177042533.
E-mail addresses: so_gheisari@pardisiau.ac.ir, so_gheisari@yahoo.com (S. Gheisari), mmeybodi@aut.ac.ir (M.R. Meybodi).

http://dx.doi.org/10.1016/j.ins.2016.01.090
0020-0255/