



Estimation of state changes in system descriptions for dynamic Bayesian networks by using a genetic procedure and particle filters



Jianjun Lu ^{a,*}, Shozo Tokinaga ^{b,1}

^a College of Economics and Management, China Agricultural University, Beijing 100083, China

^b Graduate School of Economics, Kyushu University, Fukuoka 812-8581, Japan

ARTICLE INFO

Article history:

Accepted 21 February 2014

Available online 22 March 2014

Keywords:

Dynamic Bayesian network

Genetic procedure

Particle filter

State estimation

Acrylic graph

ABSTRACT

This paper deals with the estimation of state changes in system descriptions for dynamic Bayesian networks (DBNs) by using a genetic procedure and particle filters (PFs). We extend the DBN scheme to more general cases with unknown Directed Acyclic Graph (DAG) and state changes. First, we summarize the basic model of DBN where the DAG can be changed and the state transition occurs. In the genetic procedure to estimate DAG changes, we utilize the mutation operation (called Evolutionary Programming: EP) to the DAG to maintain consistency. By defining the possible DAG structure and state changes as particles, we formalize the optimization as the PF procedure. The weight of a particle representing the DAG and state transition is defined as the capability to approximate the probability distribution function obtained from a table of cases. We apply the estimation scheme of the paper to an artificially generated DBN, in which the state of the variables and the changed structure of the DAG are already known, to prove the applicability of the method, and discuss its applicability to debt rating.

© 2014 Elsevier B.V. All rights reserved.

1. Introduction

A Bayesian network (BN) model is a multivariate statistical model that uses Directed Acyclic Graph (DAG) to represent statistical dependencies among variables and has been applied in various fields (Cooper and Herskovits, 1992; Etzeberria et al., 1997; Lam, 1998; Man Leung et al., 2002; Pearl, 1988; Wong et al., 1999). In recent years, in DAGs, dynamic Bayesian networks (Dynamic BN:DBNs) that allow variables in the DAG of a Bayesian network to be time dependent have been proposed and applied (Fearnhead, 2006; Friedman et al., 1998; Nielsen and Nielsen, 2008; Punsakaya et al., 2002; Wang et al., 2008, 2011). Additionally, the application of particle filters (PFs) has been proposed for the estimation of the state from observed data in a DBN (Andrieu et al., 2004; Arulampalam et al., 2002; Doucet et al., 2001; Gustafsson et al., 2002; Tokinaga and Tan, 2010; Wang et al., 2011). However, in the conventional method of DBNs, the allowable range of a DAG is limited to the graph shape that is known in advance, and the state probabilities belong to the pattern already known. Therefore, this conventional DBN method cannot be applied to the unknown DAG shape and state changes (Wang et al., 2011). In this paper, we deal with the estimation of state changes in system descriptions for dynamic Bayesian networks by using a genetic procedure and particle filters (PFs) (Bordley and Kadane, 1999; Tokinaga and Ikeda, 2011).

First, we organize the relation equation between the basic DAG model and the change of the DAG shape (Fearnhead, 2006; Friedman et al., 1998; Nielsen and Nielsen, 2008; Punsakaya et al., 2002; Wang et al., 2008, 2011). We describe the method for estimating the DAG shape change and variable state transition by using a genetic approach (Alvarez-Diaz and Miguez, 2008; Chen and Duan, 2011; Chi and Tang, 2007; Tokinaga and Ikeda, 2011; Wong et al., 1999). In this case, in the estimation of DAG shape change, we do not use the entire process included in the crossover procedure of Genetic Programming (GP), but use methods based on mutations, such as changing the direction of a branch of the DAG, in order to maintain consistency in the process (called Evolutionary Programming: EP in Wong et al., 1999) (Alvarez-Diaz and Miguez, 2008; Antoci et al., 2012; Chen and Duan, 2011; Chi and Tang, 2007; Tokinaga and Ikeda, 2011; Wong et al., 1999). In addition, we introduce function $f_k(z)$ to the dynamics description, which represents the state transition, and we apply GP to estimate the description shape (Ikeda and Tokinaga, 2007a,b; Lu et al., 2006, 2007; Tokinaga and Kishikawa, 2010).

Secondly, with respect to the issue of state estimation, we focus on the change of the joint distribution of the state variables between time $t + 1$ and time t , and propose a state estimation method by using PFs, which are applied to state estimation in a nonlinear state equation. In the estimation of the state change by PFs, the particles to represent the DBN structure and state transition are given in multiples, so the weight of a particle representing the DAG and state transition is defined as the capability to approximate the probability distribution function obtained from a table of cases. We apply the estimation scheme of the paper to the artificially generated DBN, in which the state of the

* Corresponding author. Tel.: +86 10 6273 6564; fax: +86 10 6273 7830.

E-mail addresses: ljun@cau.edu.cn (J. Lu), tokinaga@econ.kyushu-u.ac.jp (S. Tokinaga).

¹ Tel./fax: +81 92 642 2466.

variables and the changed structure of the DAG are already known, in order to prove the applicability of the method, and discuss its applicability to real data.

In the following, we describe the problem formulation and the estimation of the DBN description change by EP GP in Section 2. In Section 3, we explain the estimation of the state change by PFs and in Section 4, we show the application and the results. Finally, the paper concludes in Section 5.

2. Issue formulation and DBN description change estimation by EP GP

2.1. DBN and DAG

We suppose that there are observational data on the vector of random variables $X = (X_1, X_2, \dots, X_n)$. In the following, we call this database a “case table” (Cooper and Herskovits, 1992). The purpose of BNs is to express the relationship between the random variable X_1, X_2, \dots, X_n by a stochastic graph model; in this case, the probability distribution obtained from the relationship representation and case table that has been observed should be consistent. Nodes in this graph correspond to a variable; the DAG is a relationship between the bindings of these nodes, which represent the dependencies between the variables. X_k is the parent variable that precedes the variable positions, i.e., the DAG is described by the conditional probability using variable X_j corresponding to the node that is coupled in the direction of flow to the node corresponding to the variable X_k . This conditional probability is defined as $P(X_k|Pa(X_k))$; here, $Pa(X_k)$ denotes the set of variables of the parent variable X_k . The joint distribution of the random variable X_1, X_2, \dots, X_n by this can be expressed as follows (Cooper and Herskovits, 1992; Etxeberria et al., 1997; Lam, 1998; Man Leung et al., 2002; Pearl, 1988; Wong et al., 1999):

$$P(X) = \prod_{k=1}^n P(X_k|Pa(X_k)), X = (X_1, X_2, \dots, X_n) \quad (1)$$

Such an optimal DAG estimating problem is a joint distribution obtained when assuming a certain DAG. Various methods, including genetic methods, have been proposed to derive a DAG that is equal to the probability distribution obtained from a given case table (Cooper and Herskovits, 1992; Etxeberria et al., 1997; Lam, 1998; Man Leung et al., 2002; Wong et al., 1999). In this paper, we focus on the issue of estimating a current change from the shape of the DAG and therefore, we will skip the details of how to determine the optimal DAG starting from the initial state.

In the following, we extend the BN to model DBN, which introduces the time change of the DAG and the time change of the state variable. We assume this time change to be a transition based on the Markov process, and we assume a Bayesian estimation equation in time t to include the state variable at time $t - 1$ and shape change of the DAG. By introducing time t , a vector of random variables is displayed as $X(t) = (X_1(t), X_2(t), \dots, X_n(t))$. With respect to the transition of the state variables, it is assumed that for the description using the dynamics as follows, we introduce a nonlinear function $f_k(\cdot)$ here:

$$X_k(t + 1) = f_k(X_k(t)) + v_k(t) \quad (2)$$

Here, $v_k(t)$ is a function of the approximation errors, which are assumed to be normal random, i.e., the mean is 0 and the variance is σ_k^2 .

The reason we introduce such dynamics, is that in many state estimation problems the observed value of the variable is not a true value; there are many cases in which it receives some conversion (deviation) (Korolkiewicz and Elliott, 2008; Tokinaga and Tan, 2010). In addition, generally, it is necessary to consider the statistical distribution if the size of the case table is not guaranteed to be large enough.

After the introduction of such a transition, Eq. (1) will be $P(X(t)) = \prod_{k=1}^n P(X_k(t)|Pa(X_k(t)))$, in this case $Pa(X_k(t))$. At the same time variable of time t , the node of the link flowing into the node corresponding

to $X_k(t)$, will contain some of the variables of time $t - 1$. In the DBN, when representing $Pa(X_k(t))$ at time t by graph $G(t)$, we assume that this graph shape may be different at $G(t - 1)$.

In general, estimating the state variables $X(t)$ from observed data is a problem of state estimation in a dynamic system; however, considering the application field, we are targeting the problem of estimating the function $f_k(z)$ used in the state transition. Thus, it becomes a problem of estimating the structure of graph $G(t)$, which represents the DBN. Therefore, target (state) $s(t)$ can be estimated as follows:

$$s(t) = [f(z), G(t)], f(z) = (f_1(z), f_2(z), \dots, f_m(z)) \quad (3)$$

Observation data used in the estimation are values of $X(t)$ observed up to the time t .

2.2. DAG change estimation by EP

As already mentioned, this paper does not deal with the problem of estimating the BN in order to express it directly from the observed data. We estimate the change of the DAG by using a method based on the mutation that has been proposed as one way of estimating these DAGs in the BN (Lam, 1998; Wong et al., 1999). GP is a typical genetic method, which can be applied to general function approximations (Ikeda and Tokinaga, 2007a,b; Lu et al., 2006, 2007; Tokinaga and Kishikawa, 2010). That is, in generating offspring when applying GP, there are several operations that can be used, such as crossover operation and mutation operation; we use mutation operation only.

We use the mutation method in the EP procedure as described in Wong et al. (1999). In addition to the GP method, we use data that represent the current structure of the individual DAG (such as a connection matrix, for example). EP is a method of genetic processing of a set of individuals at a certain time, leaving a large fitness from one individual to the next process. The definition of fitness will be described later, but basically it is: if assuming a structure DAG that corresponds to an individual, when calculating the probability of simultaneous distribution of formula (1), by how much to approximate the probability that is determined from the case table.

The EP approach is summarized as the following four steps.

- Step M1 Adding branch
Insert a branch between the two nodes N_i and N_j arbitrarily selected in the current DAG. If it already exists, then this operation is not performed.
- Step M2 Change in branch direction
Converting the direction of the branch that exists between two nodes arbitrarily selected in the current DAG.
- Step M3 Transformation of parent variables
Switch the node that flows into the arbitrarily selected node N_i in the original branch to another node.
- Step M4 Removing branches and bypass
Once the middle stage of the EP method has been applied, redundant branches may have been formed between the nodes. In particular, if there are three nodes A, B, and C, in addition to the link through the A-B-C, a direct branch extending from node A to node C is formed. Extending the branch directly from node A to node C will have a detrimental influence when estimating the structure, so in this case this branch is removed.
Because the Step M4 procedure is used only to organize the data, we do not add a mutation procedure.

2.3. EP application efficiency using stochastic approximation

It is also possible to search randomly to estimate the graph change to t from time $t - 1$ (changes to $G(t)$ from $G(t - 1)$) by using EP, but in this paper, we improve the efficiency by using an approximate

Download English Version:

<https://daneshyari.com/en/article/5054240>

Download Persian Version:

<https://daneshyari.com/article/5054240>

[Daneshyari.com](https://daneshyari.com)