Structure Evolution of Dynamic Bayesian Network for Traffic Accident Detection

Ju-Won Hwang, Young-Seol Lee, and Sung-Bae Cho
Dept. of Computer Science
Yonsei University
Seoul, Korea
juwon@sclab.yonsei.ac.kr, tiras@sclab.yonsei.ac.kr, sbcho@cs.yonsei.ac.kr

Abstract—Recently, Bayesian network has been widely used to cope with the uncertainty of real world in the field of artificial intelligence. Dynamic Bayesian network, a kind of Bayesian network, can solve problems in dynamic environments. However, as node and state values of node in Bayesian network grow, it is very difficult to define structure and parameter of Bayesian network. This paper proposes a method which generates and evolves structure of dynamic Bayesian network to deal with uncertainty and dynamic properties in real world using genetic algorithm. Effectiveness of the generated structure of dynamic Bayesian network is evaluated in terms of evolution process and the accuracy in a domain of the traffic accident detection.

Keywords—structure of dynamic Bayesian network; Bayesian network, evolution

I. INTRODUCTION

Every year, traffic congestion and traffic accidents have been rapidly increasing in proportion to increasing number of vehicles. Although the roadway design and signal system have been improved to relieve traffic congestion, traffic casualties and property damage do not decrease. The traffic accident is a serious issue of society because vehicle is a primary means of transportation. In our previous study, we developed a real-time traffic accident detection system (RTADS) that helps us to cope with accidents and discover the causes of traffic accident by detecting the accident [1]. We gathered video data recorded at several intersections and used them to detect accidents at different intersections which have different traffic flow and intersection design. However, because the data gathered from intersections have incompleteness, uncertainty and complicated causal dependency between them, we adopted probability-based networks which calculate based on the probability for correct accident detection.

The RTADS consists of data collection unit, preprocessing unit, accident inference unit, and accident alarm unit. In data collection unit, system gathers video from a camera located on the intersections. Videos are sent to preprocessing unit in a server. The preprocessing unit extracts basic information of objects and features of the objects such as coordinate, moving vector, velocity and direction. The accident inference unit infers types and probability of the accident using above features and relationships between features with Bayesian network. Finally, accident alarm unit reports the result of inference to operators so that they can rapidly handle the accidents. An overview of the system is presented in Fig. 1.

Bayesian network, a probabilistic approach, has been used to deal with the uncertainty of problems in various fields. Particularly, Bayesian network was successfully used to work out the uncertainty for robot in dynamic and uncertain environments [2].

Bayesian network is represented as directed acyclic graph (DAG) and comprised of nodes and arcs which mean environment variables and the dependence between nodes, respectively [3]. It is the strong point of Bayesian network to reflect the knowledge of expert. However, this approach which reflects knowledge of the expert is expensive. On the other hand, we can easily model by learning data, because we can collect large amounts of data at low cost in proportion to increasing number of data. The necessity of learning of Bayesian network is increased by these properties. As the DAG structure can be generated in geometrical progression in spite of low number of nodes, it is a NP-hard problem to find the optimal DAG structure [4]. Therefore, many studies have been working for learning of Bayesian network.
Sechin et al. [5] suggested the structure learning approach of Bayesian network from data using the semantic genetic algorithm. Lee et al. [6] investigated the structure learning of Bayesian network by using the double genetic algorithm that uses the individual order and connectivity between individuals. In many cases of traffic accidents, traffic environment at the different area have different characteristics such as road type, traffic jam, and road shape. Even if a probabilistic model for a specific area is designed, it cannot guarantee that the model works well at another area. In order to treat variables in the diverse environment, it is helpful to find an optimistic structure of the probability model using GA (genetic algorithm). General Bayesian network training method is based on ML (Maximum Likelihood) method. It sometimes fails to get optimal solution. We tried to overcome the limit for training structure of dynamic Bayesian network using GA.

In this paper, we aim to find the optimal structure and the candidate structure of dynamic Bayesian network using genetic algorithm. We propose representation methods of a gene and the three steps method for fitness evaluation of populations of each created generation by genetic algorithm. The experiment is conducted to evaluate each population and the accuracy of inference result on a set of real traffic data.

This paper is organized as follows. In section II, we will deal with the background knowledge to understand our proposed approach. In Section III, we will explain individual representation methods, restrictions, genetic operators, fitness function, and selection mechanism of structure evolution of dynamic Bayesian network using genetic algorithm. In Section IV, the performance of individuals is evaluated. Section V wraps up the paper with a conclusion and suggestions for future work.

II. RELATED WORKS

A. Bayesian Network

The algorithm which is to solve the uncertainty of environment in many studies has been issuing. The probabilistic approach is the most successful method, though there are many methods to deal with uncertainty [2]. Especially, there have been many studies using Bayesian network in artificial intelligence. Bayesian network is a favorable method to model the causal relationship between cause and effect, because Bayesian network can effectively constitute probabilistic model for efficient inference and learning. It is not impossible but inefficient to represent the causal relationship between all the information of the specific domain in real world. Therefore, to use the structure of Bayesian network which is defined the conditional probability in direct link is more efficient than all causal relationships between cause and effect [7]. Bayesian network has been widely used because that method is intuitive graphics model and has efficiency of inference and learning algorithms. Bayesian network has nodes, arcs and conditional probabilistic table that represent probabilistic relationships between significant variables and is represented as the directed acyclic graph (DAG) [8]. In a probabilistic model, an input is considered as an observation or evidence. In Bayesian network, an input is denoted as a node in the network. Therefore, all inputs are included as nodes in the Bayesian network. For example, if we want to input a fact that weather is rainy, state value of 'weather node' in a Bayesian network is set by 'Rainy' state.

B. Dynamic Bayesian Network

Dynamic Bayesian network is a kind of Bayesian network for dynamic processing about data of time series. Generally, some natural phenomena such as flipping a coin does not change current probabilistic distribution depending on previous result. On the while, some phenomena like natural language processing, and earthquakes have causal dependencies between current state at time t and previous state at time t-1. Standard Bayesian network does not consider such temporal causality. Dynamic Bayesian network is used to represent the temporal relationship among nodes. This dynamic Bayesian network is useful for the tasks such as monitoring, diagnosis, and prediction. An example of structure of dynamic Bayesian network is presented in Fig. 2.

![Fig. 2. A structure of dynamic Bayesian network](image)

C. Genetic Algorithm

The individuals which have fitness greater than other individuals survive in biological evolution of nature. New populations are generated by crossover and mutation. Genetic algorithm is used to simulate evolution process of nature in computer, thereby solving a complex problem in real world. Genetic algorithm has a wide range of applications and has been used adaptive explorations and optimization problem because simple structure and the typical method are features of the genetic algorithm. There are many kinds of genetic algorithm: evolution strategies, evolutionary programming and genetic programming. The following pseudo code is a simple genetic algorithm (SGA).

```plaintext
Procedure SGA()
    initialize(Population);
    evaluate(Population);
    while (not terminal condition satisfied) do
        MatingPool = reproduce(Population);
        MutationPool = crossover(MatingPool);
        Population = mutation(MutationPool);
        evaluate(Population);
    end while
end procedure
```
There are many studies of the structure evolution using the genetic algorithm. Larranaga et al. [3] analyzed the performance of evolved network that are ASIA and ALARM using some parameters of the genetic algorithm. Wong et al. [9] proposed the new learning method of Bayesian network based on the MDL and the Evolutionary Programming (EP). Li et al. [10] presented the search method for the optimal structure using the repair operator and Kullback-Leibler distance between Bayesian network structure and data. In this paper, we consider the weight of arcs which are connected from previous time \((t-1)\) to current time \((t)\) and the whole structure of dynamic Bayesian network.

III. STRUCTURE EVOLUTION OF DYNAMIC BAYESIAN NETWORK USING GENETIC ALGORITHM

The execution sequence of evolution of dynamic Bayesian network is presented in Fig. 3. In this paper, the proposed method initializes each population that is dynamic Bayesian network. There are score metrics which calculate score of Bayesian network, but we design three steps of fitness evaluation to consider characteristics of dynamic Bayesian network.

\[
C_g = \begin{cases} 
1 & \text{if arc is connected from } i \text{ to } j \text{ and } i \text{ is a parent of } j, \\
0 & \text{otherwise.} 
\end{cases} 
\]  

(3)

The representation of a population is defined as follows:

\[
C_{11}, C_{12}, \ldots, C_{1n}, C_{21}, C_{22}, \ldots, C_{2n}, \ldots, C_{n1}, C_{n2}, \ldots, C_{nn} 
\]  

(2)

In this paper, the representation of dynamic Bayesian network structure as follows and Fig. 4 is the example of representation of Bayesian network:

\[
\begin{align*}
C_{11} & C_{12} \ldots C_{1n} \\
C_{21} & C_{22} \ldots C_{2n} \\
& \vdots \\
C_{n1} & C_{n2} \ldots C_{nn}
\end{align*}
\]

Fig. 4. An example of representation of Bayesian network

If connection matrix is used, circle may be occurred, but that is against the DAG. In Fig. 5, \(C_{44}\) and \(C_{55}\) are represented as 1. That is violated in DAG because circles are occurred by connected arc to itself.

Fig. 5. A violated case 1 in DAG

In Fig. 6, a graph has values of \(C_{21}, C_{31},\) and \(C_{32}\) which correspond to \(C_{12}, C_{13},\) and \(C_{23}\). That case is against the DAG of Bayesian network. In this paper, we substitute 0 for 1 for the cases 1 and 2.

Fig. 6. A violated case 2 in DAG

A. Population Representation Method

In genetic algorithm, the most important thing is to reflect characteristics of problem into a chromosome. The connection matrix is used to represent a population of dynamic Bayesian network. In the previous method, Bayesian network is represented by connection matrix \(C = (C_{ij})\) of \(n \times n\) where

\[
C_g = \begin{cases} 
1 & \text{if } j \text{ is parent of } i, \\
0 & \text{otherwise.} 
\end{cases} 
\]  

(1)

B. Restrictions

The order hypothesis and the repair operator have been used to apply genetic algorithm to Bayesian network. These methods such as the order hypothesis and the repair operator are proposed to avoid violated cases in DAG [11]. The order hypothesis means that if \(X_i\) is a parent of \(X_j\), order of \(X_i\) have to be earlier than \(X_j\). In this case, the gene size of each population
is reduced to $2^{\binom{n}{2}}$ because if the number of nodes is $n$, the gene size is $\binom{n}{2}$. However, if we do not assume the order between nodes, a wrong structure of DAG can be generated. Also, the solution space may be increased as follows:

$$f(n) = \sum_{i=1}^{n} (-1)^{i+1} \binom{n}{i} 2^{i(n-i)} f(n-i), \ (f(0) = 1, \ f(1) = 1) \quad (4)$$

In this paper, the order hypothesis and the repair operator are used to apply genetic algorithm into dynamic Bayesian network.

We restrict the connection from current nodes ($t$) to previous nodes ($t-1$) because arcs are connected from previous nodes ($t-1$) to current nodes ($t$) in general cases. In other words, the previous nodes are parents of the current nodes and the order of previous nodes have to be earlier than that of current nodes. Also, we restrict the connection between different names from previous nodes to current nodes. Lastly, we divide node by three kinds such as observation node, middle node, and result node to represent the causal relationship between cause and effect during evolution process. When we want to make a probabilistic model using Bayesian network, we first must define input and output variables. For example, input variables in traffic accident recognition will be road type, motion vector of a car, etc. Output variables may be occurrence of traffic accidents or accident types. Observation node corresponds to input variables for a model. Result nodes denote output variables. Middle node is sometimes a virtual node which is between observation and result. It is used for effective abstraction or efficient calculation.

C. Genetic Operators

There are crossover, mutation, inversion, substitution, duplication, supplement, and removal as operator.

We can reproduce new population by exchanging part of chromosome between two populations. This paper uses acquired result from random-number generator. If acquired result is less than predefined crossover rate, crossover is occurred by the random-number as a negative number that is less than predefined maximum value by entering gene size (Fig. 7).

Fig. 7. An example of crossover

Mutation is a genetic operator that alters one or more gene values in a chromosome from its initial state and is used to maintain genetic diversity to next generation. If acquired result from random-number generator is less than predefined mutation rate, mutation is occurred by the random-number as a negative number that is less than predefined maximum value by entering gene size (Fig. 7).

Fig. 8. An example of mutation

Fig. 9. An example of a violated structure in DAG

D. Fitness Function

Fitness function evaluates the performance of population for the given problem. Fitness function, therefore, evaluates the suitability of dynamic Bayesian network structure. The reasons of optimal model of Bayesian network are as follows [2].

- The incorrect structure may involve the wrong assumption about causality and domain structure.
- If unnecessary arcs are added to network, the number of probability parameters and complexity metric of model are increased, while the efficiency of model is decreased (Fig. 10).
If necessary arcs are removed from network, the number of probability parameters is decreased, while we cannot obtain the correct result.

In this paper, we devise a method that can evaluate the characteristics of dynamic Bayesian network. In the previous study [11], the authors developed a method to extract optimal Bayesian network structure. However, it cannot be applied to dynamic Bayesian network because it does not consider temporal dependencies and restrictions. Dynamic Bayesian network structure can be divided into the structure ($S_1$) of previous time ($t-1$), the structure ($S_2$) of current time ($t$), and the whole dynamic Bayesian network ($S_3$) (Fig. 11). The fitness function evaluates using three-steps ($S_1$, $S_2$, and $S_3$).

$$F = (N + (P_1 - P_2)) \times K = (N + D) \times K$$  \hspace{1cm} (5)

$S_1$ is evaluated using $N$, $D$, the number of arcs of each Bayesian network ($K_1$, $K_2$), and the number of connection of nodes of the same name from the previous time ($t-1$) to the current time ($t$).

$$F = (N + (P_1 - P_2)) \times K_1 \times K_2 \times K_3 = (N + D) \times K_1 \times K_2 \times K_3$$  \hspace{1cm} (6)

**E. Selection Mechanism**

Selection mechanism induces to make well-adjusted populations survive and to make maladjusted populations die out. Populations are selected based on the fitness. This paper uses the ranking selection which decides the probability of descendant to next generation by ranking the fitness score. The ranking selection is performed as follows.

$$p_{j,t} = \frac{\lambda - \text{rank}(g(I'_j)) + 1}{\hat{\lambda}(\hat{\lambda} + 1)/2}$$  \hspace{1cm} (7)

$I'_j$ means the $j$th in generation ($t$), $\text{rank}(g(I'_j))$ is ordering of all populations according to the fitness. The selection rate ($P_{j,t}$) of each population ($I'_j$) depends on the size of population ($\lambda$).

Also, we use the elitist preserving selection which preserves the best population to the next generation to prevent loss of best population by the crossover and the mutation. The elitist preserving selection is used in combination with other selection operators.

**IV. EXPERIMENTAL RESULTS**

The experiments are performed on the evolution process of dynamic Bayesian network and the precision ratio of evolved dynamic Bayesian network for a domain of the traffic accident detection. Video scripts were directly gathered at an intersection and used to evaluate the performance of the real-time traffic accident detection system. Because directly acquired videos include a small case of accidents for experiments, we can acquire relatively higher detection rate (DR), correct detection rate (CDR) and lower false alarm rate (FAR). Total 70 video scripts are used in the experiments, which contain 33 accidents, 10 situations similar to accidents, and 27 normal situations.

**A. Evolution of Dynamic Bayesian Network**

We compare the fitness of all generation with the following parameters.

- Number of generations ($G$): 150
- Number of populations ($\lambda$): 100
- Crossover rate ($P_c$): 0.3
- Mutation rate ($P_m$): 0.01
- Selection rate: 0.5
- Elitist preserving selection rate: 0.01

We can check slippery situations of the fitness from 200,000 to 250,000 in Fig. 12. The maximum fitness is found in the 61st generation but that is a slippery situation and
collected in the 71st generation. The five dynamic Bayesian networks correspond to initial, 10, 20, 50, and 100 generations, respectively.

Fig. 13 shows the evolution process of dynamic Bayesian network. The network structures of initial and 10 generations are insufficient but dynamic Bayesian network structures become efficient as the more population evolved.

**B. Test of Evolved Dynamic Bayesian Network**

In this test, the evolved dynamic Bayesian networks are evaluated using detection ratio, precision ratio, and recall ratio in a domain of the traffic accident detection. We use the virtual data and EM algorithm to set the conditional probability tables of nodes. Fig. 14 shows the detection rate using the following parameters.

- Number of generations ($G$): 100
- Number of populations ($\lambda$): 50
- Crossover rate ($P_c$): 0.3
- Mutation rate ($P_m$): 0.01
- Selection rate: 0.5
- Elitist preserving selection rate: 0.01

![Detection Rate](image)

We can detect the traffic accidents more than 65% in the 1st generation. In this case, arcs have the linkages between nodes which have higher casual relationship. Therefore, to compare detection rate, precision ratio, and recall ratio between initial population and evolved population may be useless. The evolved structure has the complexity of model and parameters. So, the detection ratio can be decreased by occurring during modeling process. Therefore, the good detection ratio of the evolved structure is the major outcome (Table 1). Table 1 shows the precision ratio and recall ratio more than 80% when the threshold is set as 80.

<table>
<thead>
<tr>
<th>Threshold</th>
<th>Precision ratio</th>
<th>Recall ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>91.61</td>
<td>98.01</td>
</tr>
<tr>
<td>60</td>
<td>90.65</td>
<td>96.04</td>
</tr>
<tr>
<td>70</td>
<td>86.35</td>
<td>90.84</td>
</tr>
<tr>
<td>80</td>
<td>83.79</td>
<td>87.38</td>
</tr>
<tr>
<td>85</td>
<td>69.38</td>
<td>86.97</td>
</tr>
</tbody>
</table>

*(T,T) : Experimental results = true / Actual output = true
(T,F) : Experimental results = true / Actual output = false
(F,T) : Experimental results = false / Actual output = true
(F,F) : Experimental results = false / Actual output = false

*Precision ratio : $(T,T) / (T,T) + (T,F)$
*Recall ratio : $(T,T) / (T,T) + (F,T)$
V. CONCLUSIONS

In this paper, the representation method of individual and three steps method of fitness function are proposed to model the relationship of time series between nodes of dynamic Bayesian network. We can confirm that the method produces the stable structure of dynamic Bayesian network after evolution and the higher detection, precision, and recall ratios are obtained from the evolved structures. For future works, we are planning to cover an issues about computation time. A good probabilistic model has to show computational efficiency as well as the accuracy. In addition to accuracy measure, it may be useful to consider computation time for fitness.

ACKNOWLEDGMENT

This work was supported by the IT R&D program of MKE/KEIT (10033807, Development of context awareness based on self learning for multiple sensors cooperation).

REFERENCES