2007 International Conference on Computational Intelligence and Security Workshops

Using Expert's Knowledge to Build Bayesian Networks

Hu Xiao-xuan,
School of Management
Hefei University of Technology,
Hefei, China, 230009
huxiaoxuan@vip.sina.com

Wang Hui
School of Management
Hefei University of Technology,
Hefei , China, 230009
wang_hui_mail@163.com

Wang Shuo
School of Management
Hefei University of Technology,
Hefei, China, 230009
wangshuo@ah.edu.cn

Abstract

Building Bayesian networks is considered to be difficult and time-consuming work. Researchers usually learn networks from data. However, it is of low efficiency because of the huge search space. Since Bayesian networks represent the causal relationships among random variables, domain experts can build Bayesian networks according to their knowledge and experience. In this paper, we develop a method to build Bayesian networks from experts. We use some knowledge elicitation tools to obtain high-quality knowledge, and ensure the validity of networks by combing knowledge from multiple experts. The experimental result indicates that this method can improve the modeling efficiency of Bayesian networks.

1. Introduction

Bayesian networks, also called Probabilistic Networks or Belief Networks, are graphic models representing the probabilistic dependencies among random variables, which offer powerful tools for representing and analyzing problems involving uncertainty. In recent years, Bayesian networks have been successfully used in many fields such as Data Mining [1] and Decision Support Systems [2].

Machine learning algorithm is adopted for automatically learning from data when building the Bayesian networks, such as K2 algorithm [3]. However, the shortcomings of this kind of methods are: (1) if there are a lot of nodes in networks, the search space of learning algorithm is so huge that it often leads to a low learning efficiency. Hence the optimization algorithm can be used to accelerate the learning speed such as genetic algorithm [4] and ant colony algorithm [5], whose effects are satisfying just to a certain degree (2) Learning Bayesian networks from data requires a substantial amount of data; however, we often lack of available database in real-world applications.

Since Bayesian networks express the causal relationships among random variables, domain experts can judge this kind of relationships by their own knowledge and experience. For example, an experienced physician can accurately judge that what symptoms are caused by a certain disease and what corresponding possibility could be. In this paper, we established a process of constructing Bayesian networks fully using expert's knowledge, during which, we gather expert's knowledge through knowledge elicitation tools and combine knowledge from multiple experts so as to achieve rapid modeling as well as ensure the accuracy.

The remainder of this paper is organized as follows. Section 2 briefly introduces the characteristics of Bayesian networks. Section 3 describes how to build the structure of Bayesian networks using expert's knowledge. Section 4 shows how to determine the probability distribution of nodes. In section 5, an example is presented. Finally, we conclude the paper with a summary.

2. Bayesian Networks

Bayesian networks have been a field of widespread interest. A Bayesian network is a directed acyclic graph representing the conditional independence relations among variables. A Bayesian network for a set of variables $X = \{X_1, ..., X_n\}$ consists of (1) a network structure S that encodes a set of conditional independence assertions about variables in X, and (2) a set P of local probability distributions associated with each variable [6]. Together, these components define the joint probability distribution for X.

Let us cite an example mentioned by Russel [7], a Bayesian network used to forecast whether the grass is wet or not. The four binary variables in the network

0-7695-3073-7/07 $25.00 © 2007 IEEE
DOI 10.1109/CIS.Workshops.2007.63

denote Cloudy, Sprinkle, Rain and WetGrass respectively. The network is shown in Figure 1.

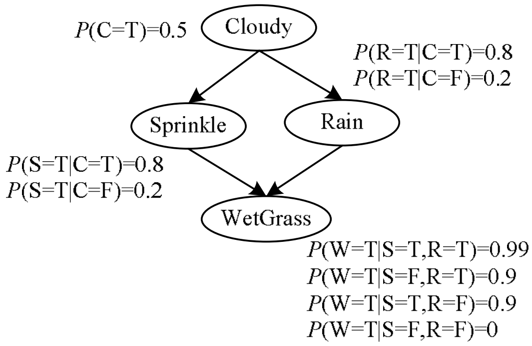

Figure 1. A Bayesian network example

3. Building Bayesian Networks with expert's knowledge

3.1. Building the structure

The building of Bayesian networks usually consists of two phases: (1) determining the structure of the networks, and (2) obtaining the conditional probability tables of each node.

Determining the structure actually focus on determining the causal relationship among variables. Assume there are n random variables, and the causal relationships between each pair of the variables are identified by experts, so it amounts to $n(n-1)/2$ pairs. In order to facilitate the expert's judgment, a table called "causal relationship questionnaire" is adopted, as is shown in Table 1.

Table 1. Causal relationship questionnaire

	variable1	variable2	…	variable n
variable 1	—			
variable 2		—		
⋮			—	
variable n				—

For each pair of random variables X_i, X_j, the content experts fill in the table is one of $\{X_i \to X_j, X_i \leftarrow X_j, X_i \not\leftrightarrow X_j\}$, where $X_i \to X_j$ denotes that X_i is X_j's parent node, while $X_j \to X_i$ interprets that X_j is X_i's parent node, and $X_i \not\leftrightarrow X_j$ represents a non-direct causal relationship between X_i and X_j. Then we can build the structure of Bayesian networks through the questionnaire.

3.2. Specifying Conditional Probability Tables

To specify the conditional probability tables of nodes is the most difficult task in building Bayesian networks, because it is very hard for experts to directly provide numerical probabilities. Renooij [8] found that the verbal description of probability people frequently use has a certain corresponding relation with the actual numerical probability. Therefore, she established a "probability scale" (as shown in Figure 2) and make discrete verbal description of probability continuous in favor of a more correct identification for probability distribution. The experimental result shows that with the adoption of "probability scale", experts can elicit 150 -175 probabilities per hour [9]. What's more, a real-world Bayesian network built by experts for the diagnosis of oesophagus cancer has improved the rate of correct diagnosis to 85% [10].

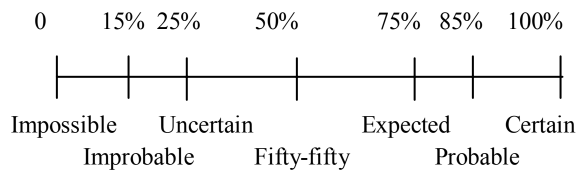

Figure 2. Probability scale

To avoid unprecise experts' knowledge, we select m experts in all from each of whom we attain the probability distribution respectively, then aggregate the probability distribution as the final result. The probability aggregation flow chart is presented in Figure 3.

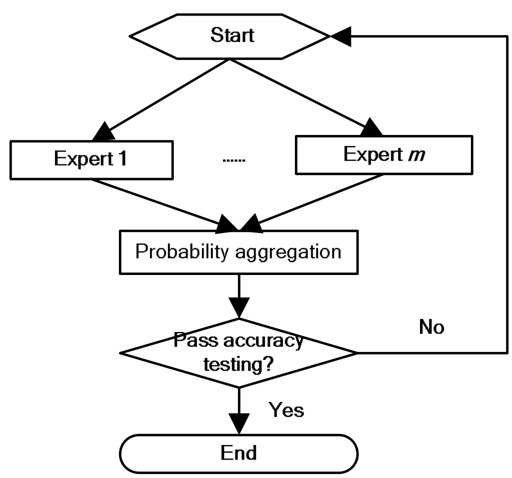

Figure3. Probability aggregation flow chart

The formula of probability aggregation is as follows:

$$P(\bullet) = \sum_{i=1}^{m} \omega_i P_i(\bullet)$$

Where ω_i denotes the weight of expert i, $\sum_{i=1}^{m} \omega_i = 1$, and $P_i(\bullet)$ is the probability given by expert i.

In the process of probability aggregating, considering different degree of expert's authority, we introduce experts' weight when aggregating experts' opinions. We identify the weight of each expert according to the title, basis for judgment, degree of familiarity with the domain, degree of self-confidence about evaluation and the efforts put into the work etc., which is shown in Table 2.

Table 2. Index of determining expert's weight

Expert's weight index	Index analysis
expert's title（a）	professor 、 associate professor 、 assistant professor (scores 10 ， 9 ， 8, respectively)
judgment basis（b）	theoretical analysis, production experience, academic reference works, directly perceived judgment (scores10，9，8，7, respectively)
degree of familiarity with the domain (c)	profession (familiar), relevant profession (more familiar), irrelevant profession (General) （scores 10，9，8, respectively）
efforts put into the work (d)	spending all the time engage in, spend much time and do detail research, spend little time and only do general research (scores 10，9，8, respectively)
degree of self-confidence about evaluation（e）	totally self-confident, much self-confident, general self-confident (scores 10，9 ，8, respectively)

Let H_i be the score for expert *i*, $H_i = a_i \cdot b_i \cdot c_i \cdot d_i \cdot e_i$ .where a_i, b_i, c_i, d_i, e_i represent the expert's title, basis for judgment, degree of familiarity with the domain, the efforts put into the work and the degree of self-confidence about evaluation, respectively. Thus the weight for expert *i* is:

$$\omega_i = H_i / \sum_{i=1}^{m} H_i \text{，} \quad i = 1, 2, \cdots, m$$

To ensure the accuracy, a testing is carried for probability distribution after aggregation. The "and" method is adopted in accuracy testing, which amounts to sum all the probability value of a node and then judge that whether it satisfies the normalization or not. For example, supposing "A" is a binary variable, we firstly inquire of the experts about the value of $P(A = True)$, secondly inquire of the experts about the value of $P(A = False)$, and then add the two values. If the sum equals or approximates 1, that indicates an accurate probability distribution, otherwise modify the probability distribution. Repeat the process until a satisfying result is obtained.

4. Empirical Evaluation

Take demand forecasting problem for example to show how to build Bayesian networks from expert's knowledge. We assume that with a daily sales volume of Hamburg fluctuating between 0 -90, the retailers order from vendors every afternoon, then receive vendors' delivery on next morning and begin the sale of the day. Suppose only the uncertainty of customers but not the uncertainty of vendors is taken into consideration, the best ordering policy can be depicted as: the volume of order equals the volume of sales, through which can we realize the profit maximization.

Noticing that market demand is affected by the following variables: the day, the weather, weather forecast and the season, we establish a table to demonstrate the causal relationship between demand and the above four variables. The contents of the table identified by experts are shown in Table 3.

Table 3. Questionnaire for causal relationships in demand forecast

	Day	Weather	Weather Forecast	Season	Demand
Day	—	$\nleftrightarrow$	$\nleftrightarrow$	$\nleftrightarrow$	$\rightarrow$
Weather	$\nleftrightarrow$	—	$\rightarrow$	$\leftarrow$	$\rightarrow$
Weather Forecast	$\nleftrightarrow$	$\leftarrow$	—	$\nleftrightarrow$	$\nleftrightarrow$
Season	$\nleftrightarrow$	$\rightarrow$	$\nleftrightarrow$	—	$\rightarrow$
Demand	$\leftarrow$	$\leftarrow$	$\nleftrightarrow$	$\leftarrow$	—

The results indicates that the season is the parent node of Weather and Demand, Day is the parent node of Demand, and Weather is the parent node of Demand and Weather Forecast, thus we build a Bayesian network model shown in Figure 4.

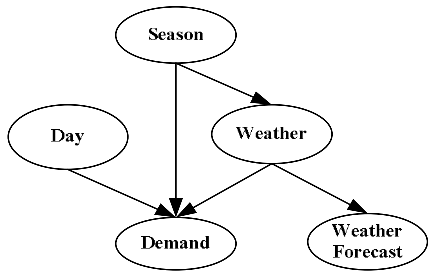

Figure 4. Bayesian network for demand forecast

Let's define the value range for above variables: Season {Jan.-Mar., Apr.-June., July-Sept., Oct.-Dec.}, Day {Saturday, Sunday, others}, Weather {sunny, rainy, cloudy }, Weather Forecast {sunny, rainy, cloudy }, Demand{0-30,31-60,61-90}. Probability scale is used to determine the conditional probabilities of each node in networks which is shown in Table 4 (due to limited space, the listed is only a part).

Table 4. Conditional probability table

Weather / Weather Forecast	sunny	cloudy	rainy
sunny	0.90	0.05	0.02
cloudy	0.07	0.80	0.13
rainy	0.03	0.15	0.85

5. Conclusions

Learning Bayesian networks is a hot research topic in the field of artificial intelligence. Many researchers have studied how to learn Bayesian networks from data. However, in practical application, we often encounter difficulties in data collection. Even with available database, the huge search space also causes a not high enough efficiency of the learning algorithm. Because Bayesian networks structure contain causal semantics, we use the knowledge of domain experts to determine the causal relationship among nodes and the probability distribution for each node. In this process, in order to enable the experts to accurately express knowledge, we use knowledge elicitation tools and choose multiple experts for the fusion of their knowledge so as to ensure the accuracy of the results. Experimental result has proved that this method can improve learning efficiency of Bayesian networks.

References

[1] Heckerman D. Bayesian Networks for Data Mining. Data Mining and Knowledge Discovery, 1997, 1, pp.79-119.

[2] Zhang S Z, Yang N H, Wang X K. Construction and Application of Bayesian Networks in Flood Decision Support System. In: Proceedings of the First International Conference on Machine and Cybermetics, Beijing, 2002, pp.718-722.

[3] Cooper G, Herskovits E.A Bayesian Method for the induction of Probabilistic Networks from data. Machine Learning, 1992, 122(9), pp.309-347.

[4] Larrañaga, P, Poza, M, Yurramendi, Y, Murga, R., Kuijpers, C. Structure learning of Bayesian networks by genetic algorithms: A performance analysis of control parameters [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996(18), pp. 912-926.

[5] de Campos, L. M., Fernández-Luna, J. M., Gámez, J. A., & Puerta, J. M. Ant colony optimization for learning Bayesian networks. International Journal of Approximate Reasoning, 2002, 31, pp.291-311.

[6] Heckerman D. Bayesian Networks for Data Mining. Data Mining and Knowledge Discovery, 1997, 1, pp.79-119.

[7] S. Russell and P. Norvig. Artificial Intelligence: A Modern Approach. Prentice Hall, Englewood Cliffs, NJ, 1995, pp.453-455.

[8] S. Renooij, C. Witteman. Talking probabilities: communicating probabilistic information with words and numbers. International Journal of Approximate Reasoning, 1999, 22, pp.169-194.

[9] L.C. Van der Gaag, S. Renooij, C.L.M. Witteman, etc. How to elicit many probabilities, in: Proceedings of the 15th Conference on Uncertainty in Artificial Intelligence, Morgan Kaufmann, San Francisco, CA, 1999, pp.647-654.

[10] C. Witteman, S. Renooij. Evaluation of a verbal–numerical probability scale. International Journal of Approximate Reasoning. 2003, 33(2), pp.117-131.

Acknowledgement

This research is supported by the National Natural Science Foundation of China (No. 70471046), Research Fund of Hefei University of Technology (No. 061103F).