Evaluating the uncertainty of a Bayesian network query response by using joint probability distribution

Yang SHAO, Toshinori MIYOSHI, Yasutaka HASEGAWA, Hideyuki BAN
Research and Development Group, Hitachi, Ltd.
1-280 Higashi-Koigakubo, Kokubunji-shi, Tokyo 185-8601, Japan
Email: {yang.shao.kn,toshinori.miyoshi.pd,yasutaka.hasegawa.mp,hideyuki.ban.et}@hitachi.com

***Abstract*—Bayesian network is a powerful tool to represent patterns inside past data. It can be used to predict future by calculating the posterior probability of future events. Machine learning techniques that can construct a Bayesian network from past data automatically are well developed in recent years. If we consider past data as a sampling set from an original probabilistic distribution, the "learning" process is actually trying to reproduce the original probabilistic distribution from the sampling set. Therefore, the finiteness of size of sampling set will bring uncertainties to the reproduced parameters of constructed Bayesian network. When the constructed Bayesian network is used to predict future, the uncertainties of reproduced parameters will be transferred to the uncertainty of query response. Here, the query response is the posterior probability that we are interested in. Evaluating the uncertainty of query response is critical to some strict industrial applications. Previous researches have proposed a method to evaluate the uncertainty. The consequence is shown as a variance of the query response. However, the conventional method need to work together with the bucket elimination, an exact inference method. Therefore, the conventional method can not deal with large Bayesian networks that used in real applications because of its calculation cost. We proposed a new approach to calculate the uncertainty of query responses by using joint probability distribution in this research. The proposed method can work with any inference method. Therefore, it can give an approximate evaluation even when the Bayesian network is large by using an approximate inference method. To investigate the accuracy of our proposed method, six well used public Bayesian networks are used as test cases. By comparing the approximate results with the exact results, an average error of -13.60% is got.**

Keywords*—*Bayesian network, machine learning, effective sample size, variance of query response, approximate inference

I. Introduction

Be accompanied by the increasing of sensor data, recognizing patterns from past data automatically and using the patterns to predict future data becomes more and more realizable in a lot of fields. A good prediction of future data can effectively help us to optimize current products and services. Bayesian network [1]–[4] is a graphical representation of the joint probability distribution for a set of variables. It is a compact form to keep probabilistic information among variables and a powerful tool to represent patterns inside past data [5], [6]. As an attractive research field [7], Bayesian network has several features that make it useful in many real-life data analysis problems [8], such as medical field [9], weather forecast [10], financial risk management [11], fault diagnosis [12], environmental modelling [13], ecological modelling [14] and classification [15]. It provides a natural way to handle missing data, allows combination of data with domain knowledge, provides an understandable way to show complex causal relationships among variables for human being and provides a method for avoiding over fitting of data [5].

Two main techniques are required by using Bayesian network as a prediction tool. First, once a Bayesian network successfully constructed, the probabilities of future events with evidence can be calculated. This calculation process is called "inference". Previous researches have proposed a lot of inference methods. They have different computational complexities and features [16], [17]. Some of them are called exact inference methods, such as bucket elimination [18], generalizing variable elimination [19] and junction tree. Exact inference methods require relatively more time to calculate, but can give the exact probabilities of future events. Some other methods are called approximate inference methods, such as loopy belief propagation [20]–[24], differential method [25] and Gibbs sampling. Approximate inference methods require relatively less time to calculate, but can only give approximate probabilities of future events. Second, to fully exploit the advantages of big data, a lot of machine learning techniques for Bayesian network construction are proposed [7], [26]–[28]. A Bayesian network can be learned from past data automatically by using these machine learning techniques. This process is called "learning". Some of learning methods can even allow a combination of data with domain knowledge [29] or a learning of hidden variables [30], [31].

With "inference" and "learning" techniques, the process of using Bayesian network to predict future data from past data can be described as follow. First, it is supposed that future data will have the same probabilistic distribution with the past data. And then, the probabilistic distribution can be learned from past data automatically by using machine learning techniques and can be represented as a Bayesian network. At last, the probabilistic distribution implicated in the Bayesian network that learned from past data can be used to predict future data. Actually, if we consider records of past data and future data as sampling sets from one original probabilistic distribution, the prediction process is actually trying to reproduce the original probabilistic distribution from one sampling set first and use the reproduced probabilistic distribution to predict another sampling set then. Therefore, we may have two kinds of error in this process of using past data to predict future data. First one comes from the reproducing process of the probabilistic distribution. And second one comes from the assumption that future data will have the same probabilistic distribution with the past data. To strict industrial applications, such as medical diagnosis, gene expression analysis, business decision support and automatic drive, both errors should be evaluated.

However, previous researches on Bayesian network seldom distinguished two kinds of errors that mentioned before. They are regularly evaluated together and represented by one number, prediction error [8]. This situation brings difficulty to further improvements. Because that the improvement strategies of two kinds of errors are essentially different. Reproducing error can be reduced by sampling more data from the original probabilistic distribution or improving the machine learning algorithms. But the error comes from the difference of past and future probabilistic distribution might require completely new information to reduce.

We will focus on evaluating the uncertainty of a query response, which is the error of prediction result caused by the reproducing process of Bayesian network in this research. The consequence will be shown as a variance of the query response. A conventional method that can calculate the variance of query response has been proposed in previous researches [32]–[35]. However, this conventional method can only be used with an exact inference method called bucket elimination [18]. Exact inference methods can not deal with large Bayesian networks because of their high calculation costs. Therefore, to deal with large Bayesian networks used in industrial applications, a new method that can be used with approximate inference methods is required.

We propose a new approach that can be used with any inference method to evaluate the variance of query response by using joint probability distribution in this research. We also use six well used public Bayesian networks as test cases to investigate the accuracy of our approach when it used with loopy belief propagation, an approximate inference method. In chapter 2, we show an introduction of conventional method and our proposed method. In chapter 3, we show test cases with public Bayesian networks, and discuss the results.

II. Methods

A Bayesian network commonly returns a posterior probability $P(H=h|\mathbf{E}=\mathbf{e})$ as the response to a query. Here, H is the queried variable, $\mathbf{E}$ is the set of evidence variables, h and $\mathbf{e}$ are assignments of H and $\mathbf{E}$. The posterior probability $P(H=h|\mathbf{E}=\mathbf{e})$ is called query response in this research. Equation (1) proposed by Allen et al. [33] is used to evaluate the variance of the query response in this research.

$$\begin{aligned}\sigma^2_{P(H=h|\mathbf{E}=\mathbf{e})} &= \sum_{C\in\Omega}\sum_{\mathbf{f}\in\mathbf{F}}\frac{1}{1+m_{C|\mathbf{F}=\mathbf{f}}}\\ &\left\{\sum_{c\in C}\left[\frac{\partial P(H=h|\mathbf{E}=\mathbf{e})}{\partial P(C=c|\mathbf{F}=\mathbf{f})}\right]^2 P(C=c|\mathbf{F}=\mathbf{f})\right.\\ &\left.-\left[\sum_{c\in C}\frac{\partial P(H=h|\mathbf{E}=\mathbf{e})}{\partial P(C=c|\mathbf{F}=\mathbf{f})}P(C=c|\mathbf{F}=\mathbf{f})\right]^2\right\}\end{aligned} \tag{1}$$

Here, $\sigma^2_{P(H=h|\mathbf{E}=\mathbf{e})}$ is the evaluation of the variance of the query response $P(H=h|\mathbf{E}=\mathbf{e})$, C is a variable of the Bayesian network, $\mathbf{F}$ is the set of parent variables of C, c and $\mathbf{f}$ are assignments of C and $\mathbf{F}$, Ω is the set of all variables in the Bayesian network. The conditional probabilities $P(C=c|\mathbf{F}=\mathbf{f})$ are called parameters in this research. Because parameters are supposed to be learned from a dataset with limited size, each parameter has a value of $m_{C|\mathbf{F}=\mathbf{f}}$ called Effective Sample Size to measure its uncertainty. As mentioned before, the variance of the query response comes from the uncertainties of the parameters in the Bayesian network. And the uncertainties of the parameters come from the process of reproducing the original conditional probability distributions from a sampling dataset (past data) with limited size. Therefore, the variance of the query response is a function of the parameters and the Effective Sample Sizes associated with parameters.

To calculate the variance of query response by using equation (1), an important amount in equation (1), the partial derivative of the query response with respect to the parameters is required. The amount is noted as below.

$$\frac{\partial P(H=h|\mathbf{E}=\mathbf{e})}{\partial P(C=c|\mathbf{F}=\mathbf{f})} \tag{2}$$

A. Conventional Method

Allen et al. [33] proposed a method to calculate the amount in equation (2) by using intermediate results that generated by using bucket elimination inference method to calculate the query response.The essential of the conventional method can be explained as follow.

First, from the factorization definition of Bayesian network, we have equation (3).

$$P(H=h|\mathbf{E}=\mathbf{e}) = \sum_{C\in\Omega/\{H\}}\prod_{C\in\Omega}P_{\mathbf{E}}(C|\mathbf{F}) \tag{3}$$

Here, $P_{\mathbf{E}}(C|\mathbf{F})$ mean conditional probabilities (parameters) $P(C=c|\mathbf{F}=\mathbf{f})$ that consistent with $\mathbf{E}=\mathbf{e}$ and $H=h$. Then, for a specific instantiation of a specific parameter ($C_X = c_x, \mathbf{F_X}=\mathbf{f_x}$), the partial derivative of query response with respect to the parameter can be calculated by equation (4).

$$\frac{\partial P(H=h|\mathbf{E}=\mathbf{e})}{\partial P(C_X=c_x|\mathbf{F_X}=\mathbf{f_x})} = \sum_{C\in\Omega/\{H\}}\prod_{C\in\Omega/\{C_X\}}P_{\mathbf{E}}(C|\mathbf{F}) \tag{4}$$

At last, conventional method indicated that a reversed process of bucket elimination can be used to calculate the equation (4).

Therefore, same as bucket elimination method, the computational complexity of the conventional method is $O(n\exp(w))$. Here, n is the number of variables in the Bayesian network, w is the treewidth of the Bayesian network.

B. Proposed Method

In this research, we propose another approach to calculate the partial derivative of query response with respect to the parameters by using a joint probability distribution instead of by using intermediate results of bucket elimination inference method. The new equation is shown as follow.

$$\frac{\partial P(H=h|\mathbf{E}=\mathbf{e})}{\partial P(C=c|\mathbf{F}=\mathbf{f})} = \frac{P(C=c,\mathbf{F}=\mathbf{f},H=h|\mathbf{E}=\mathbf{e})}{P(C=c|\mathbf{F}=\mathbf{f})} \tag{5}$$

The proof of equation (5) is shown as follow. First, from the definition of the marginal probability we know that

$$\begin{aligned} & P(H=h|\mathbf{E}=\mathbf{e}) \\ = & \sum_{c\in C,\mathbf{f}\in\mathbf{F}} P(C=c,\mathbf{F}=\mathbf{f},H=h|\mathbf{E}=\mathbf{e}) \end{aligned} \tag{6}$$

Then, from Bayes' theorem, we know that

$$\begin{aligned} & P(C=c,\mathbf{F}=\mathbf{f},H=h|\mathbf{E}=\mathbf{e}) \\ = & P(C=c,\mathbf{F}=\mathbf{f}|\mathbf{E}=\mathbf{e}) \\ & \times P(H=h|C=c,\mathbf{F}=\mathbf{f},\mathbf{E}=\mathbf{e}) \\ = & P(C=c|\mathbf{F}=\mathbf{f},\mathbf{E}=\mathbf{e}) \\ & \times P(\mathbf{F}=\mathbf{f}|\mathbf{E}=\mathbf{e}) \\ & \times P(H=h|C=c,\mathbf{F}=\mathbf{f},\mathbf{E}=\mathbf{e}) \\ = & P(C=c|\mathbf{F}=\mathbf{f}) \\ & \times P(\mathbf{F}=\mathbf{f}) \\ & \times P(\mathbf{E}=\mathbf{e}|C=c,\mathbf{F}=\mathbf{f}) \\ & \div P(\mathbf{F}=\mathbf{f},\mathbf{E}=\mathbf{e}) \\ & \times P(\mathbf{F}=\mathbf{f}|\mathbf{E}=\mathbf{e}) \\ & \times P(H=h|C=c,\mathbf{F}=\mathbf{f},\mathbf{E}=\mathbf{e}) \end{aligned} \tag{7}$$

Therefore, we have

$$\begin{aligned} & P(H=h|\mathbf{E}=\mathbf{e}) \\ = & \sum_{c\in C,\mathbf{f}\in\mathbf{F}} [P(C=c|\mathbf{F}=\mathbf{f}) \\ & \times P(\mathbf{F}=\mathbf{f}) \\ & \times P(\mathbf{E}=\mathbf{e}|C=c,\mathbf{F}=\mathbf{f}) \\ & \div P(\mathbf{F}=\mathbf{f},\mathbf{E}=\mathbf{e}) \\ & \times P(\mathbf{F}=\mathbf{f}|\mathbf{E}=\mathbf{e}) \\ & \times P(H=h|C=c,\mathbf{F}=\mathbf{f},\mathbf{E}=\mathbf{e})] \end{aligned} \tag{8}$$

Here, we found that, for a specific instantiation $(C=c,\mathbf{F}=\mathbf{f})$,

$$\begin{aligned} & [P(C=c|\mathbf{F}=\mathbf{f}) \\ & \times P(\mathbf{F}=\mathbf{f}) \\ & \times P(\mathbf{E}=\mathbf{e}|C=c,\mathbf{F}=\mathbf{f}) \\ & \div P(\mathbf{F}=\mathbf{f},\mathbf{E}=\mathbf{e}) \\ & \times P(\mathbf{F}=\mathbf{f}|\mathbf{E}=\mathbf{e}) \\ & \times P(H=h|C=c,\mathbf{F}=\mathbf{f},\mathbf{E}=\mathbf{e})] \end{aligned} \tag{9}$$

is the only item in equation (8) that depends on the parameter $P(C=c|\mathbf{F}=\mathbf{f})$. And there is no part of the amount in equation (9) being a function of $P(C=c|\mathbf{F}=\mathbf{f})$ except itself. Therefore,

$$\begin{aligned} & \frac{\partial P(H=h|\mathbf{E}=\mathbf{e})}{\partial P(C=c|\mathbf{F}=\mathbf{f})} \\ = & P(\mathbf{F}=\mathbf{f}) \\ & \times P(\mathbf{E}=\mathbf{e}|C=c,\mathbf{F}=\mathbf{f}) \\ & \div P(\mathbf{F}=\mathbf{f},\mathbf{E}=\mathbf{e}) \\ & \times P(\mathbf{F}=\mathbf{f}|\mathbf{E}=\mathbf{e}) \\ & \times P(H=h|C=c,\mathbf{F}=\mathbf{f},\mathbf{E}=\mathbf{e})] \end{aligned} \tag{10}$$

Using equation (7) again, we have

$$\begin{aligned} & \frac{\partial P(H=h|\mathbf{E}=\mathbf{e})}{\partial P(C=c|\mathbf{F}=\mathbf{f})} \\ = & \frac{P(C=c,\mathbf{F}=\mathbf{f},H=h|\mathbf{E}=\mathbf{e})}{P(C=c|\mathbf{F}=\mathbf{f})} \end{aligned} \tag{11}$$

And this is exactly equation (5).

In this new equation, exact answer of joint probability distribution $P(C=c,\mathbf{F}=\mathbf{f},H=h|\mathbf{E}=\mathbf{e})$ will spend time to calculate. However, previous researches have proposed a lot of approximate inference methods that can calculate the joint probability distribution in a Bayesian network quickly, such as loopy belief propagation method and Gibbs sampling method. The accuracy and the computational complexity of using proposed method to calculate the variance of query response depend on which inference method is used to calculate the joint probability distribution. If an exact inference method such as variable elimination method is used, the variance of query response will be same as the result calculated by the conventional method, but the computational complexity will be $O(n^2\exp(w))$, worse than the conventional method. Here, n is the number of variables in the Bayesian network, w is the treewidth of the Bayesian network. However, if an approximate inference method such as loopy belief propagation is used, the computational complexity will be reduced to $O(n^2)$. When Bayesian network becomes complex, the $\exp(w)$ part will increase rapidly and the proposed method will be faster than the conventional method because of the absence of the $\exp(w)$ part.

III. Experiments

To evaluate the accuracy of our proposed method when used with approximate inference method, six public Bayesian networks as test cases are used. All of them are well used in previous researches. Their query responses and corresponding variances are calculated by both exact and approximate methods. The difference of two results are compared to show the accuracy of the proposed method when used with approximate inference method. For each case, the query responses and corresponding variances of all variables that being their state "1" without any evidences are calculated. The average error between exact results and approximate results for all variables are investigated. Synchronous loopy belief propagation is used as the approximate inference method to calculate the joint probability distribution in equation (5). The loop numbers are set to n. The Effective Sample Size of the learning dataset (past data) is set to be $200n$. Here, n is the number of variables in Bayesian networks.

A. Results

Table I to Table VI show the results of test cases respectively. We calculated query responses by using exact inference method as reference values to show impressions of the amount of the variances of the query responses. But it should be mentioned that the point in this research is the variances of the query responses, not the query responses themselves. We calculate the variances of the query responses by using both exact and approximate inference methods and calculate the error rates between two results. Table VII shows the overall results.

B. Discussion

There are three facts that can be observed from the tables of results. First, for each Bayesian network, the error rates of variances are large for some variables but quite small for other variables. Generally, the differences between large error rates and small error rates are remarkable. For example, in the Bayesian network earthquake, variable Alarm and Merycalls have large error rates more than 20%, but the error rates of other variables are smaller than 10%, two of them are even smaller than 1%. No variable in Bayesian network cancer has large error rate. Error rates of all variables are smaller than 1% except Xray, its error is 1.7%. Besides, although the largest error rate of Bayesian network asia is about 100%, more than half of error rates are smaller than 5%. This fact shows that our proposed method can work well regularly except some special conditions in each Bayesian network. Second, even including the variables that have large error rates, the average error rates of variance are no more than 25% for all test cases. One of them is even smaller than 1%. And if we get rid of the variables that have large error rates over than 10%, we find that the average error rates of variance are no more than 4% for all test cases. This fact means that our proposed method can work well in most of Bayesian networks and work excellent in some of them. The average error of all test cases is -13.60%. Third, by comparing the error rates of variance with the query responses, we found that there is little relevance between them. This fact means that the most important factor that decides which variables have large error rates and which do not is not their probability distribution.

Based on facts mentioned before, we show some point of views on our proposed method. First, the error of variance comes from the approximate inference method that we used to calculate the joint probability distribution in equation (5), does not come from equation (5) itself. Because we have proved equation (5) theoretically, and the only part that have an approximate factor in all process is the calculation of joint probability distribution. This point of view is proved by using exact inference method instead of the approximate inference method to calculate the joint probability distribution in equation (5). When we use exact inference method, all of the results of variance are same as the results calculated by the conventional method. Second, we think that the reason of some variables have large error rates is that the approximate inference method we used in this research, synchronous loopy belief propagation method. Synchronous loopy belief propagation method sometimes will have some misconvergence variables depending on the structure of Bayesian network. When this situation happened, the joint probability distribution would be not correct, and the partial derivative would have errors. Third, to solve this problem and improve the results of our proposed method, we need to make efforts on improving the approximate inference method, and target on reducing the rate of misconvergence variables.

IV. Conclusion

A. Conclusion

In this research, first, we proposed a new equation, equation (5), to calculate the partial derivative of query response with respect to the parameters of a Bayesian network. It calculates the partial derivative by using a joint probability distribution instead of by using intermediate results of an exact inference method, bucket elimination inference method. Once we can calculate the partial derivative of query response with respect to the parameters of a Bayesian network, the variance of the prediction result can be calculated. This means that we can evaluate the influence of the reproducing error of Bayesian network models to its query responses quantitatively. And the prediction result can be shown as an expected value with a confidence interval of certain confidence level. For strict business, such as financial prediction and healthcare service, a believable prediction result is critical. Showing a confidence interval together with a prediction result can make the result more believable.

Compared with the conventional method, our proposed method can use an approximate inference method to get the variances of query responses. Therefore, to complex Bayesian network, only proposed method might be calculated in a practical time. The computational complexity of the variance of query response is reduced from $O(n\exp(w))$ to $O(n^2)$ by using proposed method.

To evaluate the accuracy of proposed method, we use six well used public Bayesian networks as test cases, and calculate their variances of query responses by both exact and approximate methods. The comparison of the results shows that our proposed method has an average error of -13.60%. Detailed analysis of the experiment results shows that the main reason of error is the synchronous loopy belief propagation, the approximate inference method we used in this research, not the proposed approach itself. Depending on the structure of Bayesian network, synchronous loopy belief propagation method sometimes will have some misconvergence variables. This is the problem we need to solve in the future research.

B. Future work

Loopy belief propagation is a well used approximate inference method for Bayesian network. However, the convergence speed and the convergence rate of this method strongly depend on the way of belief propagation. From previous researches [24], actually we know that propagating belief synchronous is not the best way. Therefore, we shall try some asynchronous ways to propagate the beliefs in the future research. We hope the improvement of propagation way can increase the convergence rate of loopy belief propagation method and make our proposed approach more accurate. Besides belief propagation methods, from previous researches, we also know that some sampling inference methods can give approximate inference results with error boundaries. These methods will also be useful to work with our proposed method to get an evaluation of the variance of query response with error boundary.

TABLE I. RESULT OF "EARTHQUAKE"

Variable	Query response (prior)	Variance (conventional method)	Variance (proposed method)	Error rate of variance
Burglary	1.00%	0.3145%	0.3145%	0.00%
Earthquake	2.00%	0.4425%	0.4425%	0.00%
Alarm	1.61%	0.3946%	0.2313%	-41.38%
Johncalls	6.37%	0.7701%	0.6932%	-9.99%
Marycalls	2.11%	0.4509%	0.3594%	-20.29%

TABLE II. RESULT OF "CANCER"

Variable	Query response (prior)	Variance (conventional method)	Variance (proposed method)	Error rate of variance
Pollution	10.00%	0.9482%	0.9482%	0.00%
Smoker	30.00%	1.4484%	1.4484%	0.00%
Cancer	1.16%	0.3376%	0.3345%	-0.92%
Xray	20.81%	1.2827%	1.2607%	-1.72%
Dyspnoea	30.41%	1.4532%	1.4484%	-0.33%

TABLE III. RESULT OF "ASIA"

Variable	Query response (prior)	Variance (conventional method)	Variance (proposed method)	Error rate of variance
Asia	1.00%	0.3145%	0.3145%	0.00%
Tub	1.04%	0.3200%	0.3197%	-0.09%
Smoke	50.00%	1.5803%	1.5803%	0.00%
Lung	5.50%	0.7202%	0.7060%	-1.97%
Bronc	45.00%	1.5717%	1.4985%	-4.66%
Either	6.48%	0.7742%	0.0000%	-100.00%
Xray	11.03%	0.9873%	0.6755%	-31.58%
Dysp	43.60%	1.5542%	1.1174%	-28.10%

REFERENCES

[1] J.Pearl, Probabilistic Reasoning in Intelligent Systems, Morgan Kaufmann, 1988.

[2] D.Heckerman, A Tutorial on Learning With Bayesian Networks, Technical Report MSR-TR-95-06, 1996.

[3] F.V.Jensen, An Introduction to Bayesian Networks, Springer Verlag, 1996.

[4] N.Friedman, M.Goldszmidt, D.Heckerman, S.Russell, Challenge: What is the Impact of Bayesian Networks on Learning?, Proceedings of the 15th International Joint Conference on Artificial Intelligence, pp.10-15, 1997.

[5] A.Darwiche, Modeling and Reasoning with Bayesian Networks, Cambridge, 2009.

[6] D.Koller, N.Friedman, Probabilistic Graphical Models: Principles and Techniques, the MIT Press, 2009.

[7] R.Daly, Q.Shen, and S.Aitken, Learning Bayesian networks: approaches and issues, The Knowledge Engineering Review, Vol.26:2, pp.99-157, 2011.

[8] D.Heckerman, A.Mamdani, M.Wellman, Real-world applications of Bayesian networks, Communications of the ACM, Vol.38, No.3, pp.25-26, 1995.

[9] E.H.Herskovits, C.F.Cooper, Algorithms for Bayesian belief-network precomputation, Methods of Information in Medicine, pp.362370, 1991.

TABLE IV. RESULT OF "CHILD"

Variable	Query response (prior)	Variance (conventional method)	Variance (proposed method)	Error rate of variance
BirthAsphyxia	10.00%	0.2371%	0.2371%	0.00%
HypDistrib	90.18%	0.2048%	0.2321%	13.32%
HypoxiaInO2	11.00%	0.1527%	0.1426%	-6.62%
CO2	71.86%	0.2354%	0.2254%	-4.26%
ChestXray	21.71%	0.1542%	0.1031%	-33.13%
Grunting	22.36%	0.2952%	0.2790%	-5.47%
LVHreport	28.67%	0.2872%	0.1925%	-32.98%
LowerBodyO2	37.14%	0.2541%	0.2431%	-4.32%
RUQO2	35.28%	0.2500%	0.2415%	-3.40%
CO2Report	74.35%	0.2890%	0.2371%	-17.97%
XrayReport	24.76%	0.1483%	0.1023%	-31.06%
Disease	4.76%	0.0676%	0.0544%	-19.54%
GruntingReport	25.65%	0.3297%	0.2566%	-22.15%
Age	64.90%	0.2414%	0.2350%	-2.65%
LVH	27.84%	0.2508%	0.2341%	-6.64%
DuctFlow	52.67%	0.1975%	0.1682%	-14.84%
CardiacMixing	5.46%	0.0844%	0.0775%	-8.16%
LungParench	71.56%	0.2148%	0.2064%	-3.89%
LungFlow	19.81%	0.2015%	0.1985%	-1.48%
Sick	31.64%	0.3524%	0.3505%	-0.55%

TABLE V. RESULT OF "INSURANCE"

Variable	Query response (prior)	Variance (conventional method)	Variance (proposed method)	Error rate of variance
GoodStudent	4.38%	0.1260%	0.1183%	-6.10%
Age	20.00%	0.1814%	0.1814%	-0.00%
SocioEcon	42.00%	0.1683%	0.1673%	-0.58%
RiskAversion	1.50%	0.0413%	0.0419%	1.32%
VehicleYear	34.80%	0.2879%	0.2768%	-3.85%
ThisCarDam	73.36%	0.1728%	0.0360%	-79.14%
RuggedAuto	49.24%	0.1817%	0.1673%	-7.93%
Accident	71.59%	0.1798%	0.1080%	-39.96%
MakeModel	14.10%	0.0949%	0.0945%	-0.47%
DrivQuality	37.14%	0.2166%	0.0834%	-61.50%
Mileage	10.00%	0.1021%	0.1021%	0.00%
Antilock	20.08%	0.2057%	0.2040%	-0.84%
DrivingSkill	32.49%	0.2133%	0.7488%	250.99%
SeniorTrain	11.70%	0.1606%	0.1107%	-31.07%
ThisCarCost	82.34%	0.1417%	0.0411%	-71.02%
Theft	0.12%	0.0238%	0.0347%	46.16%
CarValue	42.58%	0.1977%	0.0956%	-51.66%
HomeBase	26.18%	0.1503%	0.0925%	-38.47%
AntiTheft	41.89%	0.2412%	0.1864%	-22.70%
PropCost	56.29%	0.1743%	0.1185%	-32.01%
OtherCarCost	86.12%	0.1260%	0.0679%	-46.13%
OtherCar	69.65%	0.2960%	0.2902%	-1.96%
MedCost	92.81%	0.0850%	0.0699%	-17.79%
Cushioning	32.86%	0.1672%	0.1276%	-23.65%
Airbag	43.25%	0.3003%	0.2255%	-24.89%
ILiCost	96.88%	0.0600%	0.0600%	0.09%
DrivHist	57.68%	0.2168%	0.1482%	-31.64%

[10] B.Abramson, J.Brown, W.Edwards, A.Murphy, R.L.Winkler, Hailfinder: A Bayesian system for forecasting severe weather, International Journal of Forecasting, Vol.12, No.1, pp.5771, 1996.

[11] C.Shenoy, P.P.Shenoy, Bayesian network models of portfolio risk and return, Computational Finance, Vol.1999, pp.87, 2000.

[12] S.Verron, T.Tiplica, A.Kobi, Bayesian networks and mutual information for fault diagnosis of industrial systems, Workshop on Advanced Control and Diagnosis, 2006.

[13] L.Uusitalo, Advantages and challenges of Bayesian networks in environmental modelling, Ecological Modelling, Vol.203, pp.312-318, 2007.

[14] B.G.Marcot, Metrics for evaluating performance and uncertainty of Bayesian network models, Ecological Modelling, Vol.230, pp.50-62, 2012.

[15] A.Antonucci, M.E.G.V.Cattaneo, G.Corani, Likelihood-based robust classification with Bayesian networks, Advances in Computational Intelligence, pp.491-500, 2012.

[16] G.Cooper, The computational complexity of probabilistic inference using Bayesian belief networks, Artificial Intelligence Vol.42, pp.393-405, 1990.

[17] C.Butz, S.Hua, J.Chen, H.Yao, A simple graphical approach for understanding probabilistic inference in Bayesian networks, Information Sciences, Vol.179, No.6, pp.699716, 2009.

[18] R.Dechter, Bucket elimination: A unifying framework for probabilistic inference, Learning and Inference in Graphical Models, 1998.

[19] F.G.Cozman, Generalizing Variable Elimination in Bayesian Networks, Workshop on Probabilistic Reasoning in Artificial Intelligence, 2000.

[20] K.P.Murphy, Y.Weiss, and M.I.Jordan, Loopy belief propagation for approximate inference: An empirical study, In Proceeding of Uncertainty in Artificial Intelligence, pp.467-475, 1999.

[21] A.L.Madsen, F.V.Jensen, LAZY propagation: A junction tree inference algorithm based on lazy evaluation, Artificial Intelligence, Vol.113, pp.203-245, 1999.

[22] A.Choi, A.Darwiche, An edge deletion semantics for belief propagation and its practical impact on approximation quality, Proceedings of the National Conference on Artificial Intelligence, Vol.21, No.2, pp.1107, 2006.

[23] A.Choi, A.Darwiche, Focusing Generalizations of Belief Propagation on Targeted Queries, Proceedings of the 23rd national conference on Artificial intelligence, pp.1024-1030, 2008.

[24] G.Elidan, I.McGraw, D.Koller, Residual belief propagation: Informed scheduling for asynchronous message passing, Proceedings of the 22nd Conference on Uncertainty in Artificial Intelligence, 2006.

[25] A.Darwiche, A Differential Approach to Inference in Bayesian Networks, Journal of the ACM, Vol.50, No.3, pp.280-305, 2003.

[26] S.Acid, L.M.Campos, J.M.Fernandez-Luna, S.Rodriguez, J.M.Rodriguez, J.L.Salcedo, A comparison of learning algorithms for Bayesian networks: a case study based on data from an emergency medical service, Artifical Intelligence in Medicine, Vol.30, pp.215-232, 2004.

[27] L.D.Fu, A comparison of state-of-the-art algorithms for learning Bayesian network structure from continuous data, Master Thesis, 2005.

[28] D.Margaritis, Learning Bayesian Network Model Structure from Data, PhD Thesis, 2003.

[29] D.Heckerman, D.Geiger, D.M.Chickering, Learning Bayesian Networks: The Combination of Knowledge and Statistical Data, Machine Learning Vol.20, pp.197-243, 1995.

[30] J.Binder, D.Koller, S.J.Russell, K.Kanazawa, Adaptive probabilistic networks with hidden variables, Machine Learning, Vol.29 2-3, pp.213-244, 1997.

[31] G.Elidan, LEARNING HIDDEN VARIABLES IN PROBABILISTIC GRAPHICAL MODELS, PhD Thesis, 2004.

[32] T.V.Allen, R.Greiner, P.M.Hooper, Bayesian error-bars for belief net inference, Proceedings of the 17th conference on Uncertainty in artificial intelligence, 2001.

[33] T.V.Allen, A.Singh, R.Greiner, P.M.Hooper, Quantifying the uncertainty of a belief net response: Bayesian error-bars for belief net inference, Artificial Intelligence Vol.172, pp.483513, 2008.

[34] P.M.Hooper, Exact distribution theory for belief net responses, Bayesian Analysis Vol.3, pp.615-624, 2008.

[35] P.Hooper, Y.Abbasi-Yadkori, R.Greiner, B.Hoehn, Improved Mean and Variance Approximations for Belief Net Responses via Network Doubling, Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence, pp.1821, 2009.

TABLE VI. RESULT OF "ALARM"

Variable	Query response (prior)	Variance (conventional method)	Variance (proposed method)	Error rate of variance
HISTORY	5.45%	0.1319%	0.0685%	-48.07%
CVP	11.43%	0.1477%	0.0706%	-52.18%
PCWP	11.43%	0.1477%	0.0706%	-52.18%
HYPOVOLEMIA	20.00%	0.2325%	0.2325%	0.00%
LVEDVOLUME	8.86%	0.1411%	0.0767%	-45.61%
LVFAILURE	5.00%	0.1267%	0.1267%	0.00%
STROKEVOLUME	18.08%	0.1830%	0.1134%	-38.03%
ERRLOWOUTPUT	5.00%	0.1267%	0.1267%	0.00%
HRBP	17.60%	0.1533%	0.0495%	-67.69%
HREKG	16.94%	0.1500%	0.0521%	-65.25%
ERRCAUTER	10.00%	0.1744%	0.1744%	-0.00%
HRSAT	16.94%	0.1500%	0.0521%	-65.25%
INSUFFANESTH	10.00%	0.1744%	0.1744%	0.00%
ANAPHYLAXIS	1.00%	0.0578%	0.0578%	0.00%
TPR	30.68%	0.1811%	0.1768%	-2.39%
EXPCO2	4.32%	0.0665%	0.0679%	2.16%
KINKEDTUBE	4.00%	0.1139%	0.1139%	0.00%
MINVOL	71.52%	0.1562%	0.0478%	-69.39%
FIO2	5.00%	0.1267%	0.1267%	0.00%
PVSAT	79.98%	0.1357%	0.0468%	-65.48%
SAO2	79.64%	0.1375%	0.0606%	-55.91%
PAP	4.96%	0.0841%	0.0841%	-0.04%
PULMEMBOLUS	1.00%	0.0578%	0.0578%	0.00%
SHUNT	89.69%	0.1550%	0.1271%	-17.98%
INTUBATION	92.00%	0.1051%	0.1051%	0.00%
PRESS	2.72%	0.0464%	0.0323%	-30.41%
DISCONNECT	10.00%	0.1744%	0.1744%	-0.00%
MINVOLSET	5.00%	0.0844%	0.0844%	-0.00%
VENTMACH	5.00%	0.0633%	0.0633%	-0.00%
VENTTUBE	19.22%	0.1701%	0.0338%	-80.13%
VENTLUNG	74.26%	0.1446%	0.0829%	-42.69%
VENTALV	69.58%	0.1600%	0.0445%	-72.19%
ARTCO2	18.02%	0.1321%	0.0670%	-49.31%
CATECHOL	10.01%	0.1307%	0.2796%	113.93%
HR	1.40%	0.0456%	0.0457%	0.14%
CO	17.23%	0.1685%	0.0750%	-55.48%
BP	39.00%	0.1868%	0.0902%	-51.71%

TABLE VII. RESULTS OF TEST CASES

Name of Bayesian network	Average error rate of variance	Average error rate of variance for converged variables
earthquake	-14.33%	-3.33%
cancer	-0.59%	-0.59%
asia	-20.80%	-1.34%
child	-10.29%	-3.95%
insurance	-10.92%	-1.85%
alarm	-24.63%	-1.01%