Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Numerical Experimentation on Structure Simplification in Bayesian Network

Masakazu Shiba, Atsushi Takahashi, Shingo Aoki and Hiroshi Tsuji
Department of Computer Science and Intelligent Systems
Osaka Prefecture University
Osaka, Japan
{shiba@mis., takahashi@mis., aoki@, tsuji@}
cs.osakafu-u.ac.jp

Shuki Inoue
Department of Energy Use R&D Center
The Kansai Electric Power Co., Inc.
Hyogo, Japan
inoue.shuuki@d4.kepco.co.jp

***Abstract*—In order to use Bayesian network for a complex world, this paper proposes a framework to simplify the network structure. The framework consists of "elimination", "segmentation" and "aggregation". To verify the framework, this paper discusses the prediction problem on consumer's preferences for the future products. The numerical experimentation shows how the framework could be applied and how it resolved the limitation of Bayesian network for the complex world.**

***Keywords*— Bayesian Network, Sensitivity Analysis, Data Mining, Questionnaire Analysis, Decision Support, Future Forecast**

I. INTRODUCTION

As consumer's life values are diversified, requirement analysis for products plays an important role for developing markets. To avoid the waste of product development, it is necessary to analyze consumer's preferences with their attributes and value senses. From this viewpoints, we have reported on causal relationship based on information technique [1][2][3]. One of our studies concerns on Bayesian network based on the data of respondent attributes, preference on value senses and purchase demand level.

As Bayesian network expresses causal relationship by digraph structure and carries out probabilistic inference, it works for finding the difference of some parameters at the same time [4]. For example, the technique was used for the recommendation system [5] and thermal comfort analysis [2]. As the number of parameters increases, Bayesian network requires volumes of data for precise inference. Then, it becomes difficult to collect sufficient volumes of data from the view of time and cost. Therefore, Bayesian network has generally simulated in the limited data and the constructed model was also small parameters [6].

To overcome such tradeoff problem with a lot of parameters for modeling in case of the limited samples, this paper proposes a framework to simplify the network structure. The framework consists of "elimination", "segmentation" and "aggregation". To verify the framework, this paper discusses the prediction problem on consumer's preferences for the future products introduced in [1]. The numerical experimentation shows how the framework could be applied and how it resolves the limitation of Bayesian network.

This paper is organized as follows: In the second chapter, we review the outline and problems of Bayesian network. In the third chapter, we propose a framework to simplify the network structure. In the fourth chapter, we show how the framework could be applied and examine how it resolved the limitation of Bayesian network. We conclude remarks in the fifth chapter.

II. BAYESIAN NETWORK

A. *Overview of Bayesian network*

Bayesian network [4] expresses relationships between nodes by conditional probability and directed acyclic graph. Each node is connected with others by direct link and the link from node X to node Y indicates that node Y has received the direct effect from node X. In other words, there is causal relationship in the direction of the link. The source node in this link is called "parent" and the target node from the parent node is called "child" as shown in Fig. 1 [5].

Bayesian network works for carrying out probabilistic inference by setting "evidence". The "evidence" assumes that an event happens with probability. Once evidence is assigned for a set of node, the probabilities which other events happen are calculated based on the condition that an event happens. Prior probability is a probability before setting evidence and posterior probability is a probability after setting evidence. Probabilistic inference analyzes relationship between prior probability and posterior probability. Flexible evidence assignment enables the sensitivity analysis in decision making as shown in Fig. 1.

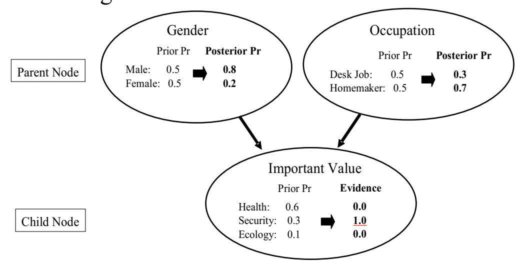


Figure 1. Example model of BN

B. Problem of Bayesian network

In Bayesian network, joint probability is used when the probabilities are calculated. Joint probability is defined as following formula (1):

$$p(X_1, X_2, ..., X_n) = \prod_{i=1}^{n} p(X_i \mid X_1, .., X_{i-1}). \quad (1)$$

As the number of nodes increases, the number of inference error also increases under the condition of the same sample size. In probabilistic inference, as the number of nodes and its parameters increases, the required sample size also increases for validating precise inference [5]. Thus, there is a tradeoff problem for verifying the events that require a lot of parameters for modeling in case of the limited samples.

This paper assumes that we are in the limited samples. Then keeping the relationship among nodes, we simplify the network structure.

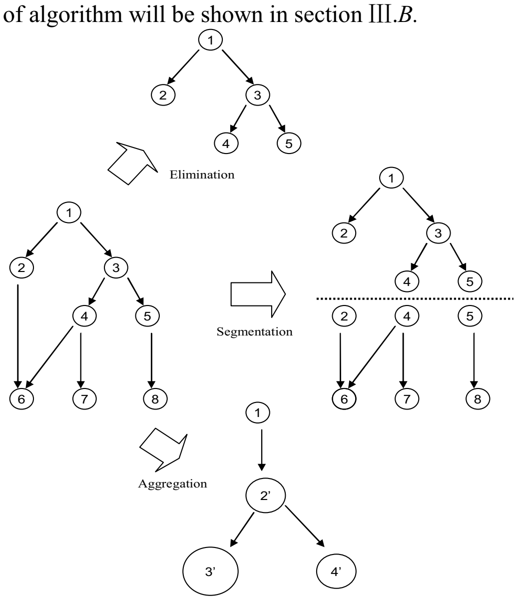
III. PROPOSITION OF THE FRAMEWORK TO SIMPLIFY NETWORK STRUCTURE

A. Overview of framework

The framework for simplifying network structure consists of three ways as follows:

(i) Elimination,
(ii) Segmentation,
(iii) Aggregation.

Fig. 2 illustrates an example of simplifying network structure. In elimination, the edge nodes 6, 7 and 8 which have no outbound link are removed from the network. In segmentation, the intermediate nodes 2, 4, and 5 divide the network into two networks. Aggregation groups nodes as follows: (1) the nodes 2 and 3 into node 2', (2) the nodes 4, 6 and 7 into node 3' and (3) the nodes 5 and 8 into node 4'. The detail of algorithm will be shown in section III.*B*.



Figurre 2. Overview of simplifying network structure

B. Detail of algorithm for structure simplification

(i) Elimination

In elimination, a part of nodes in a network, called target nodes, are removed for simplifying network structure as shown in Fig. 3 where node 6, 7 and 8 are target nodes.

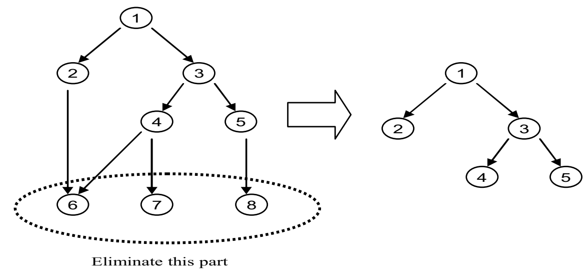


Figure 3. Elimination

[Algorithm of elimination]

Step1 Eliminate a target node.

Step2 Adjust a link. (If a target node lies on the top layer or the bottom layer in the network, a link from a target node is removed. If a target node lies on the middle layer in the network, a link connects the source node for a target node and the terminal node for a target node.)

Step3 Repeat Step1 and 2 until all target nodes are eliminated.

(ii) Segmentation

In segmentation, part of nodes divide the network into some networks for simplifying network structure as shown in Fig. 4.

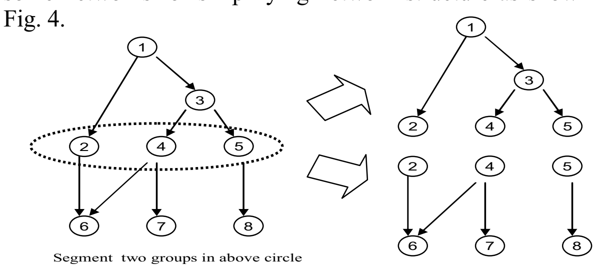


Figure 4. Segmentation

Bayesian network has characteristic called *dependency separation* [5]. Primarily, the nodes in Bayesian network have dependence. However, the nodes which satisfy dependency separation are independent. Dependency separation establishes particular nodes in three network structures. Dependency separation corresponds to stochastic independence as shown in Fig. 5.

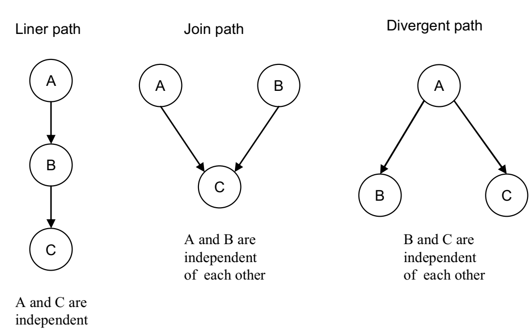


Figure 5. Dependency separation

Only the nodes which satisfy dependency separation can be segmented.

[Algorithm of segmentation]

Step1 Check the target nodes of segmentation on structural viewpoint. (Only the nodes which satisfy topology order* besides parent-child relationship can be segmented)

Step2 Check whether the target nodes belong to middle layer or not.

Step3 Check whether the target nodes satisfy dependency separation.

Step4 Segment in the target nodes. (Segmented nodes remain in each network.)

*Topology order required that the ancestral nodes always lie on the upper layer than the descendant nodes [5].

(iii) Aggregation

Aggregation merges part of nodes for simplifying network structure as shown in Fig. 6.

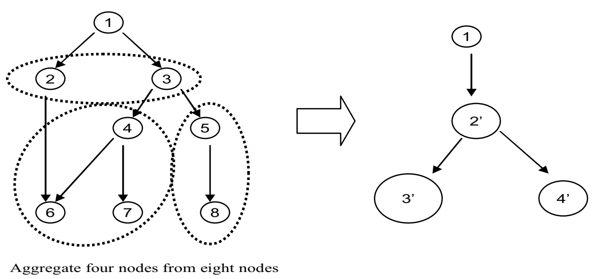


Figure 6. Aggregation

[Algorithm of aggregation]

Step1 Check the target nodes of aggregation on structural viewpoint where the nodes requires the relation of topology order.

Step2 Check the consistency of changes in aggregated nodes.

Step3 Adjust the data which are used in aggregated nodes.

Step4 Aggregate the target nodes.

Step5 Slot the aggregated nodes to the network and adjust links. (Aggregated nodes catch all links which the nodes before aggregation catch and release all links which the nodes before aggregation release.)

IV. NUMERICAL EXPERIMENTATION

A. Questionnaire data [1]

To verify the proposed framework, we have used data collected as "questionnaire on consumer electronics in the future" [7]. The questions consisted of respondent attributes, preference on value senses and purchase demand level. The sample size of this questionnaire is 1,030. There are eight nodes on respondent attributes, two nodes on preference on value senses and five nodes on purchase demand level. Furthermore, each node has some parameters. The detail parameters are listed in Table 1.

B. Modeling

To construct a BN model, the causal relationships between variables must be set beforehand on the basis of the hypothesized structure. We assumed that causal relationships sequence is respondent attributes ⇒value senses ⇒purchase demand levels. That is, the purchase demand levels to the products are influenced from person's value senses, and the value senses assumes the causal relation of being influenced from person's attributes. Because the order of the occurrence among the respondent attributes cannot be determined manually, we constructed the order structure automatically using the *K2* algorithm [5]. We used commercial software "*BAYONET*" (Http://www.msi.co.jp/BAYONET) to construct the model as shown in Fig. 7.

We operate the structure of the constructed model to carry out numerical experimentation.

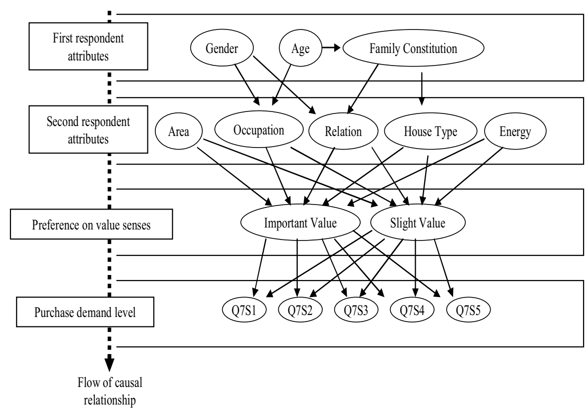


Figure 7. Constructed BN model (Original BN)

C. Result

(i) Elimination

Let us compare two networks: the constructed BN (original BN) and the BN which eliminates the first respondent attributes to the constructed BN as shown in Fig. 8.

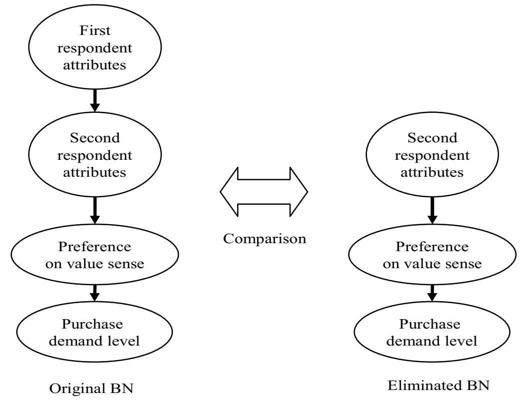


Figure 8. Experimentation on elimination

The elimination process is as follows:

[Eliminate Gender, Age and Family Constitution.]

Step1 Eliminate Gender node.

Step2 Remove the links which Gender node releases.

Step3 Eliminate Age node.

Step4 Remove the links which Age node releases.

Step5 Eliminate Family Constitution node.

TABLE Ⅰ. PARAMETERS IN EACH NODE

	Node	Parameter							
First respondent attributes	Gender	Male(515)	Female(515)						
	Age	20's(206)	30's(206)	40's(206)	50's(206)	Over 60(206)			
	Family Constitution	Living Alone (164)	Couple, No Children(131)	Couple, Child ;Living Separetely(92)	Couple, Child ;Under 20 old (242)	Couple, Child ;Over 20 old (224)	Couple, Married Child (7)	More than 3 Generations (77)	Others (93)
Second respondent attributes	Occupation	Aggriculture, Student(125)	Self-employed, Disemployed(159)	Operation Work (81)	Desk Job(251)	Executive Officer(70)	Specialist Personnel(87)	Homemaker (257)	
	Area	Kanto(632)	Tyubu(135)	Kinki(263)					
	Relation	Head of Family(502)	Spouse(354)	Child(146)	Others(28)				
	House Type	City, Single House(94)	Suburb, Single House(369)	Country, Single House (76)	City, Apartment(227)	Suburb, Apartment(237)	Country, Apartment(27)		
	Energy	Gas, Electricity ;Private Power Generation(37)	Gas, Electricity (919)	All Electrification ;Private Power Generation(6)	All Electrification (46)	Others(22)			
Preference on value senses	Inportant Value	Comfort(134)	Security(235)	Health(389)	Ecology(85)	Convenience(101)	Housework(86)		
	Slight Value	Comfort(99)	Security(36)	Health(45)	Ecology(378)	Convenience(190)	Housework(282)		
Purchase demand level	Q7S1	Want(557)	Don't Want(473)						
	Q7S2	Want(727)	Don't Want(303)						
	Q7S3	Want(524)	Don't Want(506)						
	Q7S4	Want(474)	Don't Want(556)						
	Q7S5	Want(520)	Don't Want(510)						

Q7S1～Q7S5 is products.

Step6 Remove the links which Family Constitution node releases.

As an example, it shows part of results in a variety of prior probabilities and deviation of the original and the eliminated networks as shown in Table Ⅱ and Ⅲ.

TABLE Ⅱ. A VARIETY OF PRIOR PROBABILITIES (ELIMINATION)

Node	Parameter	Prior Probability		
		Calculate from Raw Data	ORIGINAL	ELIMINATED
Occupation	Agriculture, Student	0.121	0.123	0.122
	Self-employed, Disemployed	0.154	0.154	0.154
	Operation Work	0.079	0.083	0.079
	Desk Job	0.244	0.237	0.243
	Executive Officer	0.068	0.073	0.068
	Specialist Personnel	0.084	0.088	0.085
	Homemaker	0.250	0.243	0.249
House Type	City, Single House	0.091	0.095	0.092
	Suburb, Single House	0.358	0.350	0.357
	Country, Single House	0.074	0.078	0.074
	City, Apartment	0.220	0.218	0.220
	Suburb, Apartment	0.230	0.227	0.230
	Country, Apartment	0.026	0.032	0.027

TABLE III. DEVIATION OF ORIGINAL BN AND ELIMINATED BN

Node	Parameter	Impotant Value: Comfort		Slight Value: Convenience	
		Deviation		Deviation	
		ORIGINAL	ELIMINATED	ORIGINAL	ELIMINATED
Relation	Head of Family	0.026	0.022	-0.031	-0.016
	Spouse	-0.032	-0.024	0.036	0.021
	Child	0.003	0.000	-0.003	-0.003
	Others	0.003	0.001	-0.003	-0.002
Q7S1	Want	0.015	0.015	0.026	0.026
Q7S2		0.027	0.027	0.037	0.036
Q7S3		-0.040	-0.040	0.015	0.015
Q7S4		0.011	0.011	-0.022	-0.022
Q7S5		0.005	0.005	0.032	0.032

※Set the evidence to "Important Value:~" and "Slight Value~".

※※Deviation is the value which subtract prior probability from posterior probability.

Comparing the original BN with the eliminated BN, it is observed that there is the difference on prior probability and deviation between prior probability and posterior probability. The reason why there is the difference on deviation comes from the fact that there is the difference on prior probability. So let us pay attention to prior probability. The prior probabilities in the eliminated BN are close to prior probabilities calculated from raw data. Thus elimination decreases the error. Therefore, we found that the eliminated BN analyzed more accurately than the original BN.

(ii) Segmentation

In the experimentation, we segment the constructed BN in preference on value senses into the two segmented BN as shown in Fig. 9.

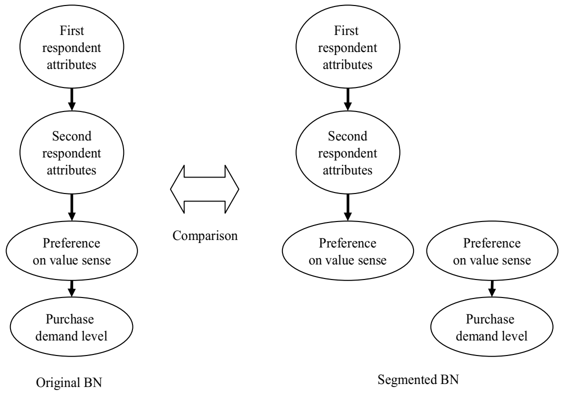


Figure 9. Experimentation on segmentation

The segmentation process is as follows:

[Segment in Important Value and Slight Value.]

Step1 Confirm that the nodes of Important Value and Slight Value satisfy topology order and exchangeability.

Step2 Confirm that the nodes of Important Value and Slight Value belong to middle layer.

Step3 Confirm that the nodes of Important Value and Slight Value satisfy the condition of divergent path in dependency separation.

Step4 Segment the original BN in the nodes of Important Value and Slight Value.

As an example, it shows part of results in a variety of prior probabilities and deviation of the original and the segmented networks as shown in Table Ⅳ and Ⅴ.

TABLE VI. A VARIETY OF PRIOR PROBABILITIES (SEGMENTATION)

Node	Parameter	Prior Probability		
		Calculate from Raw Data	ORIGINAL	SEGMENTED
Impotant Value	Comfort	0.130	0.158	0.130
	Security	0.228	0.185	0.228
	Health	0.378	0.227	0.376
	Ecology	0.083	0.145	0.083
	Convenience	0.098	0.143	0.098
	Housework	0.083	0.143	0.084
Q7S1	Want	0.541	0.529	0.537
Q7S2		0.706	0.663	0.674
Q7S3		0.509	0.506	0.511
Q7S4		0.460	0.473	0.463
Q7S5		0.505	0.498	0.501

TABLE V. DEVIATION OF ORIGINAL BN AND SEGMENTED BN

Node	Parameter	Important Value: Security		Slight Value: Housework	
		Deviation		Deviation	
		ORIGINAL	SEGMENTED	ORIGINAL	SEGMENTED
Relation	Head of Family	−0.003	−0.003	0.034	0.034
	Spouse	0.002	0.002	−0.016	−0.016
	Child	0.004	0.004	−0.013	−0.013
	Others	−0.003	−0.003	−0.005	−0.005
Q7S1	Want	0.002	0.013	−0.027	−0.014
Q7S2		0.051	0.070	−0.023	−0.010
Q7S3		0.028	0.033	−0.011	0.009
Q7S4		−0.009	−0.010	−0.039	−0.045
Q7S5		0.027	0.015	−0.035	−0.022

※Set the evidence to "Important Value:~" and "Slight Value~".

※※Deviation is the value which subtract prior probability from posterior probability.

In the original BN and the segmented BN, there is also the difference on prior probability and deviation (the value which subtracts prior probability from posterior probability). The reason why there is the difference on deviation comes from the fact that there is the difference on prior probability. So we pay attention to prior probability. In prior probability of the original BN and the segmented BN, prior probabilities of the segmented BN are close to the prior probabilities calculated from the raw data. Thus segmentation decreases the error. Therefore, we found that the segmented BN analyzes more accurately than the original BN.

(iii) Aggregation

Let us compare the original BN with the BN that the nodes of Important Value and Slight Value in preference on value senses are aggregated as shown in Fig. 10.

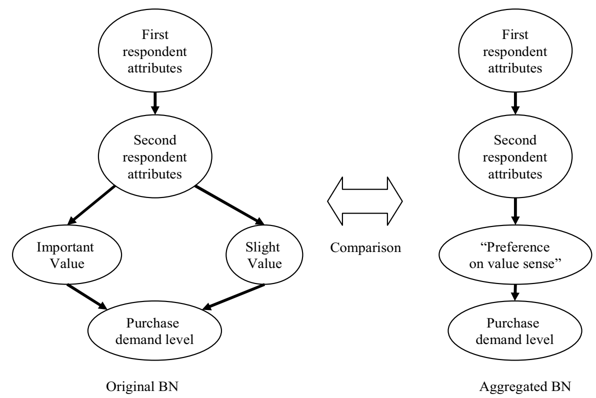


Figure 10. Experimentation on aggregation

The aggregation process is as follows:

[Aggregate Important Value and Slight Value]

Step1 Confirm that the nodes of Important Value and Slight Value satisfy topology order and exchangeability.

Step2 Confirm the consistency between nodes.

Step3 Use the data of Important Value mainly and a part of the data of Slight Value(Security).

Step4 Aggregate the nodes of Important Value and Slight Value.

Step5 Slot the node which aggregate the nodes of Important Value and Slight Value into the position between the second respondent attributes and purchase demand level. Catch all links from the second respondent attributes and release all links to products.

We aggregate the nodes of Important Value and Slight Value according to their process as shown in Fig. 11.

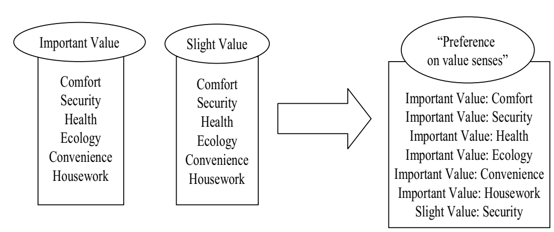


Figure 11. Result of aggregation

As an example, part of numerical results in deviation of the original and the aggregated networks (successful examples and unsuccessful examples) is shown in Table Ⅵ and Table Ⅶ.

Such as "Important Value: Health", it observes that the deviations of aggregated BN are bigger than the deviations of the original BN. However, such as "Slight Value: Ecology", some results of the aggregated BN are the opposite to results of the original BN. Therefore, we found that though there is lost information, the aggregated BN is more proper to analysis in the limited samples than the original BN.

TABLE Ⅵ. DEVIATION OF ORIGINAL BN AND AGGREGATED BN(SUCCESSFUL EXAMPLES)

Node	Parameter	Impotant Value: Comfort		Important Value: Health	
		Deviation		Deviation	
		ORGINAL	AGGREGATED	ORGINAL	AGGREGATED
Relation	Head of Family	0.026	0.031	0.016	0.020
	Spouse	−0.032	−0.035	0.010	0.012
	Child	0.003	0.001	−0.018	−0.022
	Others	0.003	0.002	−0.008	−0.010
Q7S1	Want	0.015	−0.007	0.017	0.005
Q7S2		0.027	0.006	−0.043	−0.055
Q7S3		−0.040	−0.038	0.013	0.035
Q7S4		0.011	0.015	−0.023	−0.025
Q7S5		0.005	0.006	0.000	0.013

※Set the evidence to "Important Value:~".

※※Deviation is the value which subtract prior probability from posterior probability.

TABLE Ⅶ. DEVIATION OF ORIGINAL BN AND AGGREGATED BN (UNSUCCESSFUL EXAMPLES)

Node	Parameter	Slight Value: Comfort		Slight Value: Ecology	
		Deviation		Deviation	
		ORGINAL	AGGREGATED	ORGINAL	AGGREGATED
Relation	Head of Family	−0.021	0.020	0.030	0.006
	Spouse	0.005	0.012	−0.005	−0.018
	Child	0.010	−0.022	−0.016	0.009
	Others	0.006	−0.010	−0.009	0.003
Q7S1	Want	0.034	0.005	0.018	−0.038
Q7S2		−0.026	−0.055	0.033	−0.045
Q7S3		−0.005	0.035	−0.004	−0.058
Q7S4		0.013	−0.025	0.021	0.012
Q7S5		−0.021	0.013	0.016	−0.061

※Set the evidence to "Slight Value~".

※※Deviation is the value which subtract prior probability from posterior probability.

V. CONCLUSION

To overcome the tradeoff problem for verifying the events that require a lot of parameters for modeling in case of the limited samples, this paper has proposed a framework for simplifying the network structure: elimination, segmentation and aggregation. On elimination and segmentation, the numerical experimentation has found that the eliminated BN and the segmented BN analyzed causal relationship for future product preference more accurately than the original BN. On aggregation, the experimentation has also viewed that the aggregated BN is more proper to analysis in the limited samples than the original BN. Thus, the paper has shown that the proposed framework enables to analyze in case of restriction of many parameters and the limited samples.

ACKNOWLEDGMENTS

The authors would like to express sincere thanks to Mr. Eiji Mimura and Mr. Kazunari Asari of the Kansai Electric Power Co, Inc. for their useful advises and encouragement.

REFERENCES

[1] H. Tsuji, M. Kojima, A. Takahashi, M. Nakano, S. Aoki, S. Inoue, K. Asari, E. Mimura: Preference Mining on Future Home Energy Consumption, IEEE International Conference on Systems, Man & Cybernetics (IEEE/SMC 2008), pp3697-3701, October 2008.

[2] S. Aoki, E. Mukaii, H. Tsuji, S, Inoue and E. Mimura: Bayesian Networks for Thermal Comfort Analysis, IEEE International Conference on Systems, Man & Cybernetics (IEEE/SMC 2007), pp.1919-1923, October 2007.

[3] A. Koretsune, S. Kubo, M. Nakano, H. Tsuji, Y. Jinno and E. Mimura:Causal Analysis for Thermal Comfort Votes,IEEE International Conference on Systems, Man & Cybernetics (IEEE/SMC 2006), pp1382-1387, October 2006.

[4] R. E. Neapolitan: Learning Bayesian Network, Artificial Intelligence, Prentice Hall, 2004.

[5] K. Shigemasu, M. Ueno, Y. Motomura: Outline of Bayesian Network (in Japanese), Baifukan Co.,Ltd, 2006.

[6] E. Horvitz, P. Koch, R. Sarin, J. Apacible, M. Subramani: Bay- -esphone: Precomputation of Context-Sensitive Policies for Inquiry and Action in Mobile Devices, L. Ardissono, P. Brna, and A. Mitrovic(Eds.): UM2005, LNAI3538, Springer-Verlag Berlin Heidelberg, pp.251-260, 2005.

[7] A. Takahashi, S. Aoki, H. Tsuji and S. Inoue; Bayesian Network for Future Home Energy Consumption, in Lecture Notes on Artificial Intelligence (LNAI 5243: Ed by A. Dengel et al.(Eds)), Springer-Verlag Berlin Heidelberg, pp.372-379 , 2008.