

Bayesian network variable elimination method optimal elimination order construction

Dongping Ren^{*}, *Jianxi Guo*, and *Xiaoli Hao*

Naval Logistics Academy, Tianjin 300450, China

Abstract. Variable Elimination (VE) is the most basic one of many Bayesian network inference algorithms. The speed and complexity of reasoning mainly depend on the order of elimination. Finding the optimal elimination order is a Nondeterministic Polynomial Hard (NP-Hard) problem, which is often solved by heuristic search in practice. In order to improve the speed of reasoning of the variable elimination method, the minimum, maximum potential, minimum missing edge and minimum added complexity search methods are studied. The Asian network is taken as an example to analyze and calculate the complexity and elimination of the above search method. Meta-order, through MATLAB R2018a, the above different search methods were constructed and reasoned separately. Finally, the performance of the four search methods was compared by inference time analysis. The experimental results show that the minimum increase complexity search method is better than other search methods, and the average time consuming is at least 0.012s, which can speed up the reasoning process of Bayesian network.

1 Introduction

The Bayesian network is derived from the uncertain knowledge representation model proposed by Pearl^[1]. Because of its rigorous theoretical foundation of probability theory, a graph model clearly shows the interdependence between random events, which has become a research hotspot in the fields of artificial intelligence, pattern recognition and fault diagnosis^[2]. There are two main methods of Bayesian network reasoning: exact reasoning and approximate reasoning. Exact reasoning mainly includes variable elimination method, group tree propagation algorithm, multi-tree propagation and graph reduction algorithm. Among them, the variable elimination method is the most basic reasoning algorithm, and it is easier to understand. Its advantage lies in its versatility and simplicity. It can solve the reasoning of complex Bayesian networks, especially complex fault diagnosis Bayesian network.

The complexity of the variable elimination method mainly depends on the order of its elimination. Finding the optimal elimination order is an NP-Hard problem^[3], at present, approximate algorithms are commonly used to solve the problem, mainly including minimum degree search^[4], maximum potential search^[2, 5], and minimum deficiency search^[2], Xiang

* Corresponding author: 1419863531@qq.com

Guang-jun through the analysis of the variable elimination method, puts forward an algorithm for the parallel operation of Minimum Deficiency Search and variable elimination^[9], but the algorithm did not find the optimal elimination order, and there are still certain defects; Gao Wen-yu proposed a new search algorithm—minimum increased complexity search is presented^[6], but it is only explained through simulation experiments, and is not used in examples. The specific implementation process of its search method is not given, and it is not compared with the above search method. In response to the above problems, this paper uses the Bayesian network classic model Asian network to verify the pros and cons of the above search algorithm in the variable elimination method.

2 Bayesian network

The Bayesian network is a directed acyclic graph, in which the nodes represent random variables in the event, and the directed edges represent the interdependence between the nodes. Each node has a probability distribution, and the probability distribution of the root node belongs to the marginal distribution, which can also be called a prior probability; the non-root node is a conditional probability.

In probability theory, the joint probability can be expressed in the form of a conditional probability chain^[7], its form is:

$$P(Y_1, Y_2, \dots, Y_k) = \prod_{i=1}^k P(Y_i | Y_{i-1}, \dots, Y_1) \tag{1}$$

Formula (1) can also be called a chain rule, where Y is a random variable, $P(Y_1, Y_2, \dots, Y_k)$ is the joint probability of Y_1, Y_2, \dots, Y_k , $P(Y_i | Y_{i-1}, \dots, Y_1)$ is the conditional probability of Y_i, Y_{i-1}, \dots, Y_1 .

Applying the assumption of conditional independence of Bayesian network to chain rule formula (1), we can get:

$$P(Y_1, Y_2, \dots, Y_k) = \prod_{i=1}^k P(Y_i | Parent(Y_i)) \tag{2}$$

where $P(Y_i | Parent(Y_i))$ is the conditional probability of Bayesian network node Y_i , and $Parent(Y_i)$ represents the directly connected parent node of Y_i .

For the following discussion, a classic model of Bayesian network-Asian network^[8] is introduced here. Smoking may cause lung cancer or bronchitis, and travel to Asia may cause tuberculosis. Any of these three diseases may cause breathing difficulties. If you have tuberculosis or lung cancer, the X-ray chest X-ray result may be positive. Combining these causal relationships, you can get the classic Asian network model, as shown in Figure 1, where V indicates that you have been in Asia, and T indicates that you have been infected with tuberculosis. S means smoking, L means lung cancer, B means bronchitis, E means tuberculosis or lung cancer, X means taking X-rays, and D means breathing difficulties. Using formula (2) to express the joint probability distribution of each node in Figure 1, the following formula can be obtained:

$$P(V, T, S, L, B, E, X, D) = P(V)P(S)P(T|V)P(L|S)P(B|S)P(D|B, E)P(X|E)P(E|T, L) \tag{3}$$

Table 1 shows the node probability of the Asian network, where $P(i)$ represents the prior probability of node i , $P(i|j)$ represents the conditional probability of node i when j is true, and $P(i|\bar{j})$ represents the conditional probability of node i when j is false.

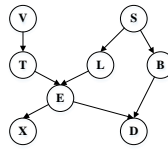


Fig. 1. Asian network model.

Table 1. Probability of asian network nodes.

Symbolic representation	Probability	Symbolic representation	Probability
$P(V)$	0.01	$P(T V)$	0.05
$P(T \bar{V})$	0.01	$P(S)$	0.50
$P(L S)$	0.10	$P(L \bar{S})$	0.01
$P(B S)$	0.60	$P(B \bar{S})$	0.30
$P(E L, T)$	1	$P(E L, \bar{T})$	1
$P(E \bar{L}, T)$	1	$P(E \bar{L}, \bar{T})$	0
$P(X E)$	0.98	$P(X \bar{E})$	0.05
$P(D E, B)$	0.90	$P(D E, \bar{B})$	0.70
$P(D \bar{E}, B)$	0.80	$P(D \bar{E}, \bar{B})$	0.10

3 Variable elimination method

Suppose F is a function of variable set $X = \{X_1, X_2, \dots, X_n\}$, and $f = \{f_1, f_2, \dots, f_m\}$ is also a set of functions, where the variables involved in f are all a subset of the variable set X . If $F = \prod_{i=1}^m f_i$, then f is a decomposition of F , and f_1, f_2, \dots, f_m is called the decomposition factor of F . Elimination methods can be divided into two types. One is to start from $F(X_1, X_2, \dots, X_n)$ and complete the elimination of X_1 through $G(X_2, \dots, X_n) = \sum_{X_1} F(X_1, X_2, \dots, X_n)$; the other is to start from $f = \{f_1, f_2, \dots, f_m\}$, use formula $\sum_{X_1} \prod_{i=1}^k f_i$ to eliminate all related functions $\{f_1, f_2, \dots, f_k\}$ with X_1 , and return $\sum_{X_1} \prod_{i=1}^k f_i$ to f , achieve The elimination of X_1 .

Through the above description, if the first elimination method is adopted, the computational complexity increases exponentially with respect to the number n of variables. If the second elimination method is adopted, the computational complexity only depends on the number of variables related to the decomposition factor f_1, f_2, \dots, f_k of X_1 , and the computational complexity will be greatly reduced.

Suppose X is used to represent the set of all variable nodes in a Bayesian network N , and f is the set of probability distributions of all nodes in N . According to the definition of Bayesian network, f can be regarded as a decomposition of the joint probability distribution $P(X)$ represented by N . Suppose the evidence set $E = \{E_1, E_2, \dots, E_m\}$ is observed, and their value is denoted as e . In the factor of the set f , set all the variables

representing the evidence as the actual observation value, and then another set of functions can be obtained, which is denoted as f' . This step is called setting evidence, f' is a decomposition of the function $P(Y, E=e)$, where $Y = X \cap \bar{E}$.

Suppose a subset of Y is Q , and eliminate all the variables in Y but not in Q from f' to get another set of functions, denoted as f'' . f'' is a decomposition of $P(Q, E=e)$. Multiply all the factors in f'' to get $P(Q, E=e)$. According to the definition of conditional probability, formula (4) can be obtained:

$$P(Q|E=e) = \frac{P(Q, E=e)}{P(E=e)} \tag{4}$$

where $P(Q|E=e) = \sum_Q P(Q, E=e)$. The above process is an elimination process to reduce calculation variables, referred to as VE algorithm^[2, 9, 10].

The VE algorithm needs to provide five inputs: (1) Bayesian network N , which is also a decomposition of the joint distribution $P(X)$. (2) The set E of evidence variables. (3) The value e of the evidence variable set E . (4) Query the set Q of variables. (5) The order ρ of all variables not in $Q \cup E$ is the order of elimination. For the acquisition of the joint distribution $P(X)$, formula (2) is used, just write the joint distribution according to the network topology diagram; the evidence set variable set E and the value e of E are all observed by personnel; the query variable Q is the desired event; the acquisition of the elimination order ρ needs to be acquired through the search method, and different search methods acquire different elimination orders. The above is the input amount of the variable elimination method. The code implementation of VE inference algorithm is shown in Table 2.

4 Analysis of elimination complexity

It can be seen from the implementation process of the variable elimination method that the most time-consuming and memory-consuming step is the call to the function $Elim(f, Z)$. The calculation amount of this function far exceeds the calculation of other steps, so the calculation complexity of the entire variable elimination method is mainly determined by this function.

The function $Elim(f, Z)$ means to pick out all functions $\{f_1, \dots, f_k\}$ containing Z from f , multiply them to obtain the intermediate function G , and then eliminate Z from G . Let X_1, \dots, X_k be all the variables in G except Z . If the function is expressed as a multi-dimensional table, the number of functions that G needs to store is $|Z| \prod_{i=1}^k |X_i|$. In summary, $|Z| \prod_{i=1}^k |X_i|$ is a measure of the complexity of the function G , and therefore also a measure of the complexity of $Elim(f, Z)$, which is called the elimination cost of variable Z , which is recorded as $\delta(Z)$ ^[2, 6].

$$\delta(Z) = |Z| \prod_{i=1}^k |X_i| \tag{5}$$

Table 2. Implementation process of VE algorithm.

Function	Input	Output	Process
$VE(N, E, e, Q, \rho)$	N, E, e, Q, ρ	$P(q, E=e)$	<ol style="list-style-type: none"> $f \leftarrow$ The set of all probability distributions in N; in the factor of f, set the evidence variable E to the observed value e While $(\rho \neq \phi)$, Let Z be the variable whose ρ is rearranged in front, and delete Z from ρ $f \leftarrow Elim(f, Z)$ End while all factors in f are multiplied, a function $h(Q)$ with Q is obtained Return $h(Q)/\sum_Q h(Q)$
$Elim(f, Z)$	F, Z	Another set of functions	<ol style="list-style-type: none"> Delete all functions involving Z from f, and let these functions be $\{f_1, \dots, f_k\}$ $G \leftarrow \prod_{i=1}^k f_i$ $h \leftarrow \sum_Z g$ Put h back into f Return f

When using variable elimination method for reasoning, multiple variables need to be eliminated, and set them as Z_1, Z_2, \dots, Z_n in turn. Here, the total elimination cost $\delta = \sum_{i=1}^n \delta(Z_i)$ is used to measure the total complexity of the variable elimination method. Take the Asian network as an example, the function set is $f = \{P(V), P(S), P(T|V), P(L|S), P(B|S), P(D|B, E), P(X|E), P(E|T, L)\}$. Set the evidence node as $X=1, E=1$, calculate the credibility of V , that is, find the probability of $P(V|X=1, E=1)$, the elimination order is $\rho = \langle D, B, L, S, T \rangle$, and all variables in the network take binary values. When eliminating D , $G = P(D|B, E)$ needs to be calculated, involving 3 variables: D, B, E , so the elimination cost of D is $\delta(D) = 2 \times 2 \times 2 = 8$. By analogy, the elimination costs of the remaining four nodes B, L, S , and T are 4, 8, 2, and 4 respectively. Then the total cost of elimination in this order of elimination is 26.

The structure graph is to consider a set of functions $F = \{f_1, \dots, f_m\}$. The structure graph of F can be defined as starting from an empty graph. For each variable in F , add a node to the graph accordingly. For any two variables X and Y , if they appear in the same factor f_i at the same time, an edge is added between their corresponding nodes. The graph model constructed according to this rule is a structural graph. Figure 2 shows the structure of the Asian network.

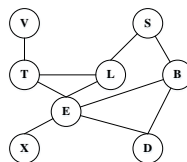


Fig. 2. Asia network structure diagram.

Eliminating the cost of a variable Z from a set of functions f can be calculated from the structure diagram of the function. As mentioned earlier, eliminate the cost $\delta(Z) = |Z| \prod_{i=1}^l |X_i|$ of Z from f , where X_1, \dots, X_l are variables other than Z , Z are all variables involved in the function $G = \prod_{i=1}^k f_i$, and f_1, \dots, f_k are all factors involved in Z . According to the definition of the structure diagram above, in the structure diagram of f , $\{X_1, \dots, X_l\}$ is just the set of all nodes adjacent to Z , denoted as $nb(Z)$, so:

$$f(Z) = |Z| \prod_{X \in nb(Z)} |X| \tag{6}$$

In the Asian network, eliminating the variable T from f requires calculating $G = P(T|V)P(E|T,L)$, involving 4 nodes $\{T, V, E, L\}$, so $\delta(T) = |Z| \prod_{i=1}^l |X_i| = |T| \cdot |V| \cdot |E| \cdot |L| = 2 \times 2 \times 2 \times 2 = 16$. In the structure diagram, the node adjacent to T uses $nb(T) = \{V, L, E\}$, so $\delta(T)$ can also be calculated according to formula (5), namely $\delta(T) = |Z| \prod_{X \in nb(Z)} |X| = 16$.

5 Elimination order construction

5.1 Minimal degree method

In the structure diagram, the minimum degree search method is to put the node variable with the smallest degree at the end of the elimination sequence queue, and delete the node and the directed edge connecting the node in the network. If there are multiple nodes with the same degree, choose one of them and repeat the above process until all the nodes in the structure diagram are eliminated. This elimination method is called the minimum degree method. Taking the Asian network model as an example, set the evidence node as $X=1$ and $E=1$, where 1 means that the event is true. Find the posterior probabilities of V, T, S, L, B , and D respectively. That is to find the probability of $P(V|X=1, E=1)$, $P(T|X=1, E=1)$, $P(S|X=1, E=1)$, $P(L|X=1, E=1)$, $P(B|X=1, E=1)$, $P(D|X=1, E=1)$. The minimum degree search method is used to find the elimination order, and the total elimination order is $\rho = \langle E, D, B, L, S, T, X, V \rangle$. The order of elimination when finding node V is $\rho = \langle D, B, L, S, T \rangle$, and the joint distribution can be obtained from the formula (3) in the second section:

$$F = \{P(V), P(S), P(T|V), P(L|S), P(B|S), P(X=1|E=1), P(D|B, E=1), P(E=1|L, T)\} \tag{7}$$

Eliminate element D as follows:

$$F = \{P(V), P(S), P(T|V), P(L|S), P(B|S), P(X=1|E=1), P(E=1|L, T), \varphi_1(B)\} \tag{8}$$

among them $\varphi_1(B) = \sum_D P(D|B, E=1)$

Eliminate element B as follows:

$$F = \{P(V), P(S), P(T|V), P(L|S), P(X=1|E=1), P(E=1|L, T), \varphi_2(S)\} \tag{9}$$

among them $\varphi_2(S) = \sum_B P(B|S) \varphi_1$

Eliminate element L as follows:

$$F = \{P(V), P(S), P(T|V), P(X = 1|E = 1), \varphi_3(S, T), \varphi_2(S)\} \quad (10)$$

among them $\varphi_3(S, T) = \sum_L P(E = 1|L, T)P(L|S)$

Eliminate element S as follows:

$$F = \{P(V), P(T|V), P(X = 1|E = 1), \varphi_4(T)\} \quad (11)$$

among them $\varphi_4(T) = \sum_S P(S)\varphi_2(S)\varphi_3(S, T)$

Eliminate element T as follows:

$$F = \{P(V), P(X = 1|E = 1), \varphi_5(V)\} \quad (12)$$

among them $\varphi_5(V) = \sum_T P(T|V)\varphi_4$

Calculate the posterior probability distribution of V :
 $h(V = 1) = P(V = 1)P(X = 1|E = 1)\varphi_5 = 0.00100205$, $h(V = 0) = P(V = 0)P(X = 1|E = 1)\varphi_5 = 0.09920295$

Calculate the maximum hypothesis test of V : $P(V = 1|X = 1, E = 1) = \frac{h(V = 1)}{\sum_V h(V)} = 0.01$

The above process is the inference calculation process of the variable elimination method. The Asian network is constructed in MATLAB R2018a, and $P(V = 1|X = 1, E = 1) = 0.01$, which takes 0.0750s (the subsequent inference calculation process of different search methods are the same as above, and will not be repeated). Similarly, the elimination order when seeking T is $\rho = \langle D, B, L, S, V \rangle$; the elimination order when seeking S is $\rho = \langle D, B, L, T, V \rangle$; the elimination order when seeking L is $\rho = \langle D, B, S, T, V \rangle$; the elimination order when seeking B is $\rho = \langle D, L, S, T, V \rangle$; the elimination order when seeking D is $\rho = \langle B, L, S, T, V \rangle$. The inference time of each node can be obtained as 0.0239s, 0.0127s, 0.0164s, 0.0146s, 0.0264s. The average inference time is 0.0282s.

5.2 Maximum potential search

The maximum potential search is to number all nodes in an undirected graph according to the following rules: In an undirected graph with n nodes, in the number i step, select the neighboring nodes with the most numbered nodes, and these nodes are not numbered, And number it as $n-i+1$. If there are multiple such nodes, choose one of them. When all the nodes in the undirected graph are numbered, the numbers are sorted in ascending order. The order is the elimination order of the corresponding nodes. This is the general idea of the maximum potential search method. Use the maximum potential search method to find the elimination order of the Asian network: the first step is to number the B node as 8, the second step can be the number of the node S or D , here the D node is numbered as 7, and the third step is for the node S as 6. The fourth step is to number the node L as 5, the fifth step is to number the E node as 4, the sixth step can be the number of the node T or X , here you choose to number T as 3, and the seventh step is to number the node V as 2, Step 8 number the last node X as 1. So far, all nodes have been numbered, and their elimination order is $\rho = \langle X, V, T, E, L, S, D, B \rangle$. Excluding the evidence nodes X and E , the elimination order of the corresponding nodes can be obtained. The elimination order when seeking V is $\rho = \langle T, L, S, D, B \rangle$; the elimination order when seeking T is $\rho = \langle V, L, S, D, B \rangle$; the elimination order when seeking S is $\rho = \langle V, T, L, D, B \rangle$;

when seeking L The order of elimination is $\rho = \langle V, T, S, D, B \rangle$; the order of elimination when seeking B is $\rho = \langle V, T, L, S, D \rangle$; the order of elimination when seeking D is $\rho = \langle V, T, L, S, B \rangle$. The inference time of each node can be obtained as 0.0258s, 0.0126s, 0.0118s, 0.0126s, 0.0118s, 0.0146s. The average inference time is 0.0149s.

5.3 Minimal missing edge search

The number of missing edges of a node Z in the undirected graph is the number of edges that need to be added when using the function $Elim(f, Z)$ to eliminate Z . The minimum missing edge search calculates the number of missing edges of the node while deleting the node from the undirected graph. Each step needs to calculate the number of missing edges of the node. When eliminating nodes, select the node with the smallest number of missing edges to delete. When there are multiple nodes with the same number of missing edges at the same time, choose one of them, and loop until all nodes are equal. been deleted. Still taking the Asian network model as an example, at the initial stage, the number of missing edges of nodes $\{V, S, T, L, B, E, X, D\}$ are $\{0, 1, 2, 2, 2, 8, 0, 0\}$, choose one, choose V here. After removing V , the number of missing edges of nodes $\{S, T, L, B, E, X, D\}$ are $\{1, 0, 2, 2, 8, 0, 0\}$, choose one of them, and choose X here. After eliminating X , the number of missing edges of the node $\{S, T, L, B, E, D\}$ are $\{1, 0, 2, 2, 6, 0\}$, choose one of them, here we choose T . After eliminating T , the number of missing edges of node $\{S, L, B, E, D\}$ is $\{1, 1, 3, 3, 0\}$, and D is eliminated. After eliminating D , the number of missing edges of nodes $\{S, L, B, E\}$ are $\{1, 1, 1, 1\}$ respectively. Since nodes S, L, B , and E form a ring, the number of missing edges is the same and can be eliminated for any one, choose to eliminate E here, and the remaining 3 nodes can be eliminated arbitrarily. The order of selection here is L, S, B . In summary, the elimination order of the least missing edge search method is $\rho = \langle V, X, T, D, E, L, S, B \rangle$. Excluding the evidence nodes X and E , the elimination order of other nodes can be obtained. The elimination order when seeking V is $\rho = \langle T, D, L, S, B \rangle$; the elimination order when seeking T is $\rho = \langle V, D, L, S, B \rangle$; the elimination order when seeking S is $\rho = \langle V, T, D, L, B \rangle$; when seeking L The order of elimination is $\rho = \langle V, T, D, S, B \rangle$; the order of elimination when seeking B is $\rho = \langle V, T, D, L, S \rangle$; the order of elimination when seeking D is $\rho = \langle V, T, L, S, B \rangle$. The inference time of each node can be obtained as 0.0150s, 0.0122s, 0.0119s, 0.0119s, 0.0121s, 0.0125s. The average inference time is 0.0126s.

5.4 Minimal increase in complexity search

The main idea is to delete a node when eliminating elements in the structure graph, and all edges connected to the node will be deleted. In order to reduce the complexity of the graph, it is often hoped to delete as many as possible. The number of edges removed is represented by e_d ; when the node is deleted, the deleted graph needs to be formed into a complete graph, so some edges will be added between some nodes to form a complete graph, but the added edges will increase the complexity of the graph, so we hope to add as few edges as possible. Record the number of edges added as e_a , and use e_a / e_d to get a new measurement index, that is, the minimum added complexity, which is represented by IC. When selecting the elimination order, first calculate the IC value of all nodes, select the smallest IC value for elimination, if there are multiple identical values, choose one of them, and then calculate the IC values of the remaining nodes, and repeat until all The node is eliminated, the elimination order can be obtained. The minimum increased complexity search method is a dynamic search process. Starting from the entire network graph, it constantly loops judgments to find

the optimal solution, improves the overall running time of the reasoning, and avoids partial elimination and overall comparison. The lack of slowness. This algorithm is suitable for the elimination of multiple query variables during fault diagnosis, and can improve the overall running time. Taking the Asian network as an example, using the minimum increase complexity search method, Table 3 is the IC value of the initial state node.

Table 3. Initial state node IC value.

node	e_a	e_d	e_a/e_d
V	0	1	0
S	1	2	0.5
T	2	3	0.67
L	2	3	0.67
B	2	3	0.67
E	9	5	1.8
X	0	1	0
D	0	2	0

It can be seen from the table that the IC values of nodes *V*, *X*, *D* are all 0, and one of them can be selected for elimination. Here, node *V* is selected, and after node *V* is eliminated, the IC values of other nodes can be obtained, as shown in Table 4.

Table 4. Node IC value after eliminating node *V*.

node	e_a	e_d	e_a/e_d
S	1	2	0.5
T	0	2	0
L	2	3	0.67
B	2	3	0.67
E	9	5	1.8
X	0	1	0
D	0	2	0

It can be seen from the table that the IC values of nodes *T*, *X*, and *D* are all 0, and one of them can be selected for elimination. Here, node *D* is selected, and after node *D* is eliminated, the IC values of the remaining nodes can be obtained, as shown in Table 5.

Table 5. Node IC value after eliminating node *D*.

node	e_a	e_d	e_a/e_d
S	1	2	0.5
T	0	2	0
L	2	3	0.67
B	1	2	0.5
E	5	4	1.25
X	0	1	0

It can be seen from the table that the IC values of nodes *T* and *X* are both 0, and one of them can be selected for elimination. Here, node *X* is selected, and after node *X* is eliminated, the IC values of the remaining nodes can be obtained, as shown in Table 6.

Table 6. Node IC value after eliminating node X.

node	e_a	e_d	e_a/e_d
S	1	2	0.5
T	0	2	0
L	2	3	0.67
B	1	2	0.5
E	2	3	0.67

It can be seen from the table that the IC values of node *T* are all 0, and the IC values of the remaining nodes can be obtained after node *T* is eliminated, as shown in Table 7.

Table 7. Node IC value after eliminating node T.

node	e_a	e_d	e_a/e_d
S	1	2	0.5
L	1	2	0.5
B	1	2	0.5
E	1	2	0.5

It can be seen from the table that the IC values of nodes *S*, *L*, *B*, and *E* are all 0.5, and one of them can be selected for elimination. Here, node *S* is selected, and the remaining nodes *L*, *B*, and *E* after eliminating node *S* form a triangle graph. All nodes are equivalent and can be combined arbitrarily. Finally, the elimination order $\rho = \langle V, D, X, T, S, L, B, E \rangle$ is obtained. The inference time of each node can be obtained as 0.0119s, 0.0130s, 0.0124s, 0.0116s, 0.0115s, 0.0116s. The average inference time is 0.012s. Figure 3 is a comparison of the average time consumption of the four inference algorithms. It can be clearly seen that the minimum increase complexity search algorithm runs better than the other three search methods.

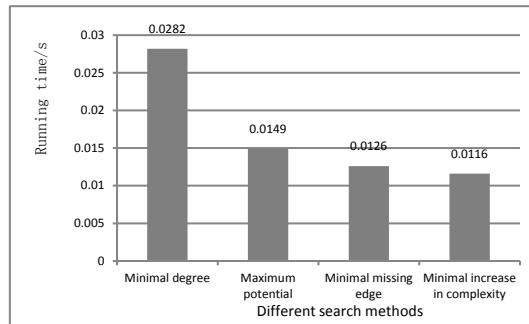


Fig. 3. Comparison of different inference algorithms.

6 Conclusion

Bayesian network has a wide range of application prospects, which can be used in artificial intelligence, fault diagnosis and other fields. The speed of calculation directly determines whether the method can be applied in the above-mentioned fields. Therefore, this paper takes the Asian network model as an example, and discusses the realization process of the variable elimination method in the reasoning example. The Asian network is constructed and inferred through MATLAB R2018a, and the four elimination sequence structures of the variable elimination method are compared through experiments. Method, experiments show that the

minimum increase complexity search algorithm is better than other search algorithms, which can effectively reduce the time of inference operation.

References

1. Pearl J. Fusion, propagation, and structuring in belief networks [J]. *Artificial Intelligence*, 1986(3):241-288.
2. Zhang Lianwen. *Introduction to Bayesian Nets* [M]. Science Press, 2006.
3. Cooper G F. The computational complexity of probabilistic inference using bayesian belief networks [J]. *Artificial Intelligence*, 1990, 42 (2):393-405.
4. Amestoy P R, Davis T A, Duff I S. Algorithm 837: AMD, an approximate minimum degree ordering algorithm [J]. *ACM Transactions on Mathematical Software*, 2004, 30(3):381-388.
5. Berry A, Blair J R S, Heggernes P, et al. Maximum cardinality search for computing minimal triangulations of graphs [J]. *Algorithmica (New York)*, 2004, 39 (4):287-298.
6. Gao Wenyu, Zhang Li. Approximate construction algorithm for Bayesian net optimal elimination order [J]. *Journal of Computer Applications*, 2011, 31 (08):2072-2074.
7. Li Jianchuan. *Bayesian network fault diagnosis and maintenance decision-making method and application research* [D]. National University of Defense Technology of the Chinese People's Liberation Army, 2002.
8. Zhang N L. Hierarchical latent class models for cluster analysis [Z]. 2002230-237.
9. Xiang Guangjun, Kong Bing, Ou Jiaqin. Research on Parallelization of Bayesian Network VE Inference Algorithm [J]. *Journal of Yunnan University (Natural Science Edition)*, 2010, 32 (04):392-395.
10. Zhang N L, Poole D. Exploiting Causal Independence in Bayesian Network Inference [J]. *CoRR*, 1996.