Автор: к.т.н., м.н.с. Терентьев А. Н.
каф. ММСА НТУУ “КПИ” УНК ИПСА
Киев–2011

Описание алгоритма эвристического метода построения байесовской сети на основе использования ЗВИ и функции ОМД / MDL

Сокращение ОМД означает Описание Минимальной Длиной, это прямой перевод названия с английского языка, MDL – Minimum Description Length.

Метод ОМД состоит из нескольких этапов. Сначала он выполняет вычисление значения взаимной информации между всеми вершинами, после чего выполняется целенаправленный поиск, который использует принцип ОМД в качестве оценочной функции, которая используется на каждой итерации построения сети Байеса.

Значение обоюдной информации

На первом этапе вычисляется значение обоюдной информации, которое было предложено исследователями Шоу и Лью в 1968 году. Данная мера была предложена для оценки степени зависимости между двумя переменными X_i и X_j. Для вычисления используется формула:

$$MI(X_i, X_j) = P(X_i, X_j) \cdot \log\left(\frac{P(X_i, X_j)}{P(X_i) \cdot P(X_j)}\right) = \sum_{x_i, x_j} P(x_i, x_j) \cdot \log\left(\frac{P(x_i, x_j)}{P(x_i) \cdot P(x_j)}\right)$$

где запись $\sum\limits_{x_i, x_j}$ означает перебор по всем состояниям переменных X_i и X_j.

MI – это сокращение от английского Mutual Information.

ЗВИ – это русскоязычное сокращение, которое означает Значение Взаимной Информации.

По своей сути ЗВИ это аналог корреляции, но по своему содержанию – это количество информации содержащейся в переменной X_i о переменной X_j.

ЗВИ принимает неотрицательные значения, то есть $MI(X_i, X_j) \geq 0$.

В случае, если X_i и X_j полностью независимы одна от другой, то $MI(X_i, X_j) = 0$, так как $P(X_i, X_j) = P(X_i) \cdot P(X_j)$ и тогда

$$\log\left(\frac{P(X_i, X_j))}{P(X_i) \cdot P(X_j)}\right) = \log\left(\frac{P(X_i) \cdot P(X_j)}{P(X_i) \cdot P(X_j)}\right) = \log(1) = 0$$

В случае, если баейсовская сеть (БС) состоит из N вершин, то для вычисления $MI(X_i, X_j)$ всех пара сочетаний X_i и X_j нужно выполнить $\frac{N \cdot (N-1)}{2}$ вычислений, так как $MI(x^i, x^j) = MI(x^j, x^i)$.

В качестве примера вычислим ЗВИ между X_1 и X_2 для которых задан набор экспериментально полученных значений из 10 наблюдений в табл. 1.

Таблица 1. Набор значений из 10 наблюдений

n	X_1	X_2
1	0	1
2	1	0
3	0	1
4	1	0
5	0	1
6	0	1
7	1	0
8	1	0
9	0	1
10	1	1

Для вычисления ЗВИ между X_1 и X_2 по формуле

$$MI(X_i, X_j) = \sum_{x_i, x_j} P(x_i, x_j) \cdot \log\left(\frac{P(x_i, x_j)}{P(x_i) \cdot P(x_j)}\right)$$

потребуется использование дополнительных значений вероятностей, которые приведены в таблицах 2, 3 и 4.

Таблица 2. Таблица значений вероятностей переменной X_1

$P(X_1)$	
$P(X_1 = 0)$	$P(X_1 = 1)$
0,5	0,5

Таблица 3. Таблица значений вероятностей переменной X_2

$P(X_2)$	
$P(X_2 = 0)$	$P(X_2 = 1)$
0,4	0,6

Таблица 4. Таблица значений совместной вероятности между переменными X_1 и X_2

$P(X_1, X_2)$	
$P(X_1=0, X_2=0)$	0
$P(X_1=0, X_2=1)$	0,5
$P(X_1=1, X_2=0)$	0,4
$P(X_1=1, X_2=1)$	0,1

$$MI(X_1, X_2) = \sum_{x_i, x_j} P\left(X_i, X_j\right) \cdot \log\left(\frac{P(X_i, X_j)}{P(X_i) \cdot P(X_j)}\right) =$$

$$= P\left(X_1 = 0, X_2 = 0\right) \cdot \log\left(\frac{P(X_1 = 0, X_2 = 0)}{P(X_1 = 0) \cdot P(X_2 = 0)}\right) +$$

$$+ P\left(X_1 = 0, X_2 = 1\right) \cdot \log\left(\frac{P(X_1 = 0, X_2 = 1)}{P(X_1 = 0) \cdot P(X_2 = 1)}\right) +$$

$$+ P\left(X_1 = 1, X_2 = 0\right) \cdot \log\left(\frac{P(X_1 = 1, X_2 = 0)}{P(X_1 = 1) \cdot P(X_2 = 0)}\right) +$$

$$+ P\left(X_1 = 1, X_2 = 1\right) \cdot \log\left(\frac{P(X_1 = 1, X_2 = 1)}{P(X_1 = 1) \cdot P(X_2 = 1)}\right) =$$

$$= 0 + 0{,}5 \cdot \log\left(\frac{0{,}5}{0{,}5 \cdot 0{,}6}\right) + 0{,}4 \cdot \log\left(\frac{0{,}4}{0{,}5 \cdot 0{,}4}\right) + 0{,}1 \cdot \log\left(\frac{0{,}1}{0{,}5 \cdot 0{,}6}\right) =$$

$$= 0 + 0{,}2554 + 0{,}2773 - 0{,}1099 = 0{,}4228 \quad .$$

Функция описания минимальной длиной

Согласно теории кодирования разработанной Шенноном, при известном распределении $P(X)$ случайной величины X длина оптимального кода для передачи конкретного значения x по каналу связи стремиться к $L(x) = -\log P(x)$. Энтропия источника $S(P) = -\sum_x P(x) \cdot \log P(x)$ является минимальной ожидаемой величиной закодированного сообщения. Любой другой код, основанный на неверном представлении про источник сообщения, приведет к большей длине сообщения. Другими словами, чем лучше построена модель источника, тем компактнее кодируются данные.

При обучении в качестве источника данных выступает некоторая истинная функция распределения $P(D|h_0)$, где $D=\{d_1,...,d_N\}$ – набор данных, а h – гипотеза про вероятностное происхождение данных, $L(D|h) = -\log P(D|h)$ – эмпирический риск аддитивный по количеству наблюдений и пропорциональный эмпирической ошибке. Отличие между $P(D|h_0)$ и распределение модели $P(D|h)$ по мере Кулбака-Леблера определяется как:

$$\left|P(D|h) - P(D|h_0)\right| = \sum_D P(D|h_0) \cdot \log \frac{P(D|h_0)}{P(D|h)} =$$
$$= \sum_D P(D|h_0) \cdot \left|L(D|h) - L(D|h_0)\right| \geq 0 \quad ,$$

то есть она представляет собой разницу между ожидаемой длиной закодированных данных на основе гипотезы про минимальную возможную длину. Эта разница всегда положительна и равна нулю, только при полном совпадении обоих распределений. Другими словами, гипотеза лучше только в том случае, когда средняя длина закодированных данных как можно меньше.

Принцип ОМД в своей упрощенной и наиболее общей форме утверждает, что среди множества моделей нужно отобрать ту, которая позволит описать данные наиболее сжато (кратко) без потери информации.

В общем виде задача ОМД выглядит следующим образом. Сначала задается множество начальных данных $D=\{d_1,...,d_n\}$, $d_i = \{x_i^{(1)} x_i^{(2)} ... x_i^{(N)}\}$ (нижний индекс – номер наблюдения, а верхний – номер переменной), n – количество наблюдений, каждое наблюдение состоит из N ($N \geq 2$) переменных $X_1, X_2, ..., X_N$, каждая j-я переменная ($j = 1,...,N$) имеет $A^{(j)} = \{0,1,...,\alpha^{(j)} - 1\}$ ($\alpha^{(j)} \geq 2$) состояний, каждая структура БС представляется N множествами предков ($\Pi^{(1)},...,\Pi^{(N)}$), то есть для каждой вершины $j = 1,...,N$, $\Pi^{(j)}$ – это множество отцовских вершин, такое что $\Pi^{(j)} \subseteq \{X_1,...,X_N\} \setminus \{X_N\}$ (вершина не может быть предком самой себе, то есть петли в графе отсутствуют). Тогда ОМД структуры $g \in G$ при заданной последовательности из n наблюдений $x^n = d_1 d_2 ... d_n$ вычисляется по формуле:

$$L(g, x^n) = H(g, x^n) + \frac{k(g)}{2} \cdot \log(n)$$

где $k(g)$ – количество независимых условных вероятностей в структуре БС g, а $H(g, x^n)$ – это эмпирическая энтропия.

$$H(g, x^n) = \sum_{j \in J} H(j, g, x^n)$$

$$k(g) = \sum_{j \in J} k(j, g)$$

где ОМД j-й вершины вычисляется по формуле:

$$L(j, g, x^n) = H(j, g, x^n) + \frac{k(j, g)}{2} \cdot \log(n)$$

где $k(j, g)$ – количество независимых условных вероятностей j-й вершины:

$$k(j, g) = (\alpha^{(j)} - 1) \cdot \prod_{k \in \phi(j)} \alpha^k$$

где $\phi(j) \subseteq \{1,.., j-1, j+1,..., N\}$, это такое множество, для которого $\Pi^{(j)} = \{X_k : k \in \phi^{(j)}\}$.

Эмпирическая энтропия j-й вершины вычисляется по формуле:

$$H(j, g, x^n) = \sum_{s \in S(j,g)} \sum_{q \in A^{(j)}} -n[q, s, j, g] \cdot \log \frac{n[q, s, j, g]}{n[s, j, g]}$$

где

$$n(s, j, g) = \sum_{i=1}^{n} I(\pi_i^{(j)} = s)$$

$$n[q, s, j, g] = \sum_{i=1}^{n} I(X_i = q, \pi_i^{(j)} = s)$$

где $\pi^{(j)} = \Pi^{(j)}$, а $X_k = x_k, \forall k \in \phi^{(j)}$, функция $I(E) = 1$, если предикат $E = true$, в противном случае $I(E) = 0$.

Простой алгоритм построения топологии БС с использованием ОМД выглядит следующим образом: по циклу выполняется перебор всех возможных не цикличных структур сети. В g^* сохраняется оптимальная структура сети. Оптимальной структурой будет та, у которой будет наименьшее значение функции $L(g,x^n)$.

Алгоритм построения БС с использованием ОМД

1. $g^* \leftarrow g_0 (\in G)$;

2. для $\forall g \in G - \{g_0\}$ если $L(g,x^n) < L(g^*,x^n)$ то $g^* \leftarrow g$;

3. в качестве решения на выход подается g^*.

Пример №1 использования функции описания минимальной длиной

На рис. 1 приведена оптимальная структура БС, которая соответствует экспериментально собранным данным из табл. 5.

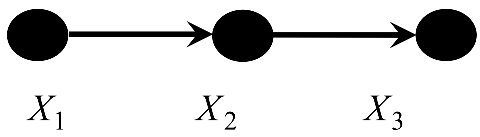


Рис. 1 Оптимальная структура БС, которая соответствует данным из табл. 5

В случае полного перебора всех возможных топологий (структур) сети нужно будет рассмотреть 25 структур. После того как все структуры будут рассмотрены в качестве оптимальной будет выдана структура приведенная на рис. 1.

Таблица 5 Обучающий набор данных из 10 записей

n	X_1	X_2	X_3
1	0	1	1
2	1	0	0
3	0	1	1
4	1	0	0
5	0	1	1
6	0	1	1
7	1	0	1
8	1	0	0
9	0	1	1
10	1	1	1

Вычисление значения ОМД этой структуры выполняется следующим образом.

Вершина X_1.

У вершины X_1 отсутствуют предки, то есть $\Pi^{(1)} = \{\}$. Эмпирическая энтропия вычисляется как

$$H(j=1,g) = -5\cdot\log\left(\frac{5}{10}\right) - 5\cdot\log\left(\frac{5}{10}\right) = 6{,}9315,$$

а количество независимых условных вероятностей

$$k(j=1,g) = 2-1 = 1.$$

То есть длина описания вершины X_1 равняется

$$L(1,g) = 6{,}9315 + \frac{1}{2}\cdot\log(10) = 8{,}0828.$$

При вычислении можно использовать логарифм с любой базой, в рассматриваемом примере используется с базой $e = 2{,}7183$, то есть натуральный логарифм.

Таблица 6. Таблица значений параметров вершины X_1

X_1	$n[q,s,j,g]$	$n[s,j,g]$
0	5	10
1	5	

Вершина X_2**.**

У вершины X_2 одна родительская вершина X_1, то есть $\Pi^{(2)} = \{X_1\}$. Эмпирическая энтропия вычисляется как:

$$H(j=2,g) = \left(-0\cdot\log\left(\frac{0}{5}\right) - 5\cdot\log\left(\frac{5}{5}\right)\right) + \left(-4\cdot\log\left(\frac{4}{5}\right) - 1\cdot\log\left(\frac{1}{5}\right)\right) = 2{,}502,$$

а количество независимых условных вероятностей $k(j=2,g) = (2-1)\cdot 2 = 2$. То есть длина описания вершины равняется:

$$L(2,g) = 2{,}502 + \frac{2}{2}\cdot\log(10) = 4{,}8046.$$

Таблица 7. Таблица значений параметров вершины X_2

X_1	X_2	$n[q,s,j,g]$	$n[s,j,g]$
0	0	0	5
0	1	5	
1	0	4	5
1	1	1	

Вершина X_3**.**

У вершины X_3, как и у X_2, одна родительская вершина X_2, то есть $\Pi^{(3)} = \{X_2\}$. Эмпирическая энтропия вычисляется как:

$$H(j=3,g) = \left(-3\cdot\log\left(\frac{3}{4}\right) - 1\cdot\log\left(\frac{1}{4}\right)\right) + \left(-0\cdot\log\left(\frac{0}{6}\right) - 6\cdot\log\left(\frac{6}{6}\right)\right) = 2.2493,$$

а количество независимых условных вероятностей $k(j=3,g) = (2-1)\cdot 2 = 2$. Тогда длина описания вершины X_3 равняется

$$L(3,g) = 2.2493 + \frac{2}{2} \cdot \log(10) = 4{,}5519$$

Таблица 8. Таблица значений параметров вершины X_3

X_2	X_3	$n[q,s,j,g]$	$n[s,j,g]$
0	0	3	4
0	1	1	
1	0	0	6
1	1	6	

Длина описания всей структуры БС.

Длина описания структуры БС представленной на рис. 1 равняется:

$$L(g,x^n) = \sum_{j=1}^{3} L(j,g,x^n) = 8{,}0828 + 4{,}8046 + 4{,}5519 = 17{,}4393.$$

Пример №2 использования функции описания минимальной длиной

Рассмотрим более сложный пример, когда у дочерней вершины более чем один предок, на рис. 2 приведена такая структура БС, пусть ей соответствует набор данных из табл. 5.

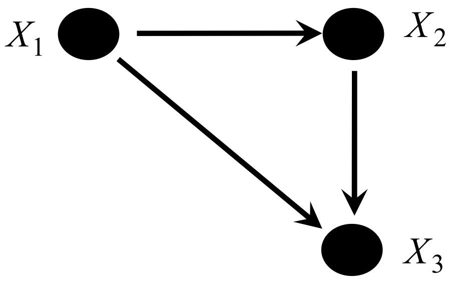


Рис. 2 Структура БС, которая соответствует данным из табл. 5

Значения длин описания для X_1 и X_2 вычисляются также как в примере №1 приведенном выше.

$$L(1,g) = 6{,}9315 + \frac{1}{2} \cdot \log(10) = 8{,}0828$$

$$L(2,g) = 2{,}502 + \frac{2}{2} \cdot \log(10) = 4{,}8046\,,$$

а вот вычисление значения длины описания для X_3 отличается.

Вершина X_3.

У вершины X_3 две родительские вершины X_1 и X_2, то есть $\Pi^{(3)} = \{X_1,\, X_2\}$. Эмпирическая энтропия вычисляется как сумма двух энтропий в зависимости от каждого предка:

$$H(3,g) = H(j=3, g | \Pi = \{X_1\}) + H(j=3, g | \Pi = \{X_2\})$$

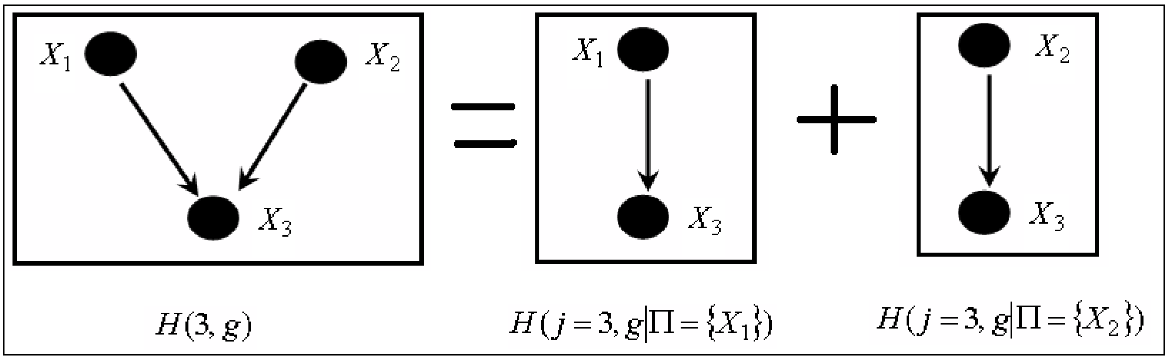


Рис. 4 Вычисление энтропии третьей вершины

$$H(j=3, g | \Pi = \{X_1\}) =$$
$$= \left(-0 \cdot \log\left(\frac{0}{5}\right) - 5 \cdot \log\left(\frac{5}{5}\right)\right) + \left(-3 \cdot \log\left(\frac{3}{5}\right) - 2 \cdot \log\left(\frac{2}{5}\right)\right) = 3.3651$$

Таблица 9. Таблица значений параметров вершины X_3 в зависимости от предка X_1

X_1	X_3	$n[q,s,j,g]$	$n[s,j,g]$
0	0	0	5
0	1	5	
1	0	3	5
1	1	2	

$$H(j=3,g|\Pi=\{X_2\}=$$
$$=\left(-3\cdot\log\left(\frac{3}{4}\right)-1\cdot\log\left(\frac{1}{4}\right)\right)+\left(-0\cdot\log\left(\frac{0}{6}\right)-6\cdot\log\left(\frac{6}{6}\right)\right)=2.2493$$

Таблица 10. Таблица значений параметров вершины X_3 в зависимости от предка X_2

X_2	X_3	$n[q,s,j,g]$	$n[s,j,g]$
0	0	3	4
0	1	1	
1	0	0	6
1	1	6	

$$H(3,g)=H(j=3,g|\Pi=\{X_1\})+H(j=3,g|\Pi=\{X_2\})=$$
$$=3.3651+2.2493=5.6144$$

Количество независимых условных вероятностей $k(j=3,g)=(2-1)\cdot 2\cdot 2=4$.

Длина описания вершины X_3 равняется

$$L(3,g)=5.6144+\frac{4}{2}\cdot\log(10)=10{,}2196$$

Длина описания всей структуры БС.

Длина описания структуры БС представленной на рис. 2 равняется:

$$L(g,x^n)=\sum_{j=1}^{3}L(j,g,x^n)=8{,}0828+4{,}8046+10{,}2196=23{,}107\,.$$

Описание алгоритма эвристического метода построения байесовской сети на основе использования ЗВИ и функции ОМД / MDL

Входные данные.

1. Множество обучающих данных $D=\{d_1,...,d_n\}$, $d_i=\{x_i^{(1)}x_i^{(2)}...x_i^{(N)}\}$ (нижний индекс – номер наблюдения, а верхний – номер переменной),
2. n – количество наблюдений;
3. N – количество вершин (узлов, переменных, признаков).

Первый этап.

Для всех парасочетаний вершин вычисляет ЗВИ $Set\_MI=\left\{MI(X_i,X_j);\ \ \forall i,j\right\}$. После этого элементы множества Set_MI сортируют (упорядочивают / ранжируют) по убыванию значения ЗВИ.

$$Set\_MI=\{MI(X_{m_1},X_{m_2}),MI(X_{m_3},X_{m_4}),MI(X_{m_5},X_{m_6}),...\}\,.$$

Второй этап.

Шаг 1. Из множества ЗВИ Set_MI выбирают первое максимальное значение $MI(X_{m_1},X_{m_2})$. На основе него строится множество моделей G:

{ ($m_1\rightarrow m_2$), ($m_1\leftarrow m_2$), (m_1 не зависит от m_2)}.

Запись вида $m_i\rightarrow m_j$ означает, что вершина X_{m_i} является предком вершины X_{m_j} .

Шаг 2. Выполняется поиск среди всех моделей множества G. В параметре g^* сохраняется оптимальная (лучшая) структура БС. Оптимальной структурой будет та, у которой будет наименьшее значение функции $L(g, x^n)$, где $L(g, x^n)$ – ОМД структуры модели при заданной последовательности из n наблюдений $x^n = d_1 d_2 ... d_n$.

1. $g^* \leftarrow g_0 (\in G)$;

2. для $\forall g \in G - \{g_0\}$ если $L(g, x^n) < L(g^*, x^n)$ то $g^* \leftarrow g$;

3. на выход в качестве решения подается g^*.

Шаг 3.

После того как найдена оптимальная структура (структуры) g^* из G, из множества значений взаимной информации Set_MI выбирается следующее максимальное значение $MI(X_{i\_next}, X_{j\_next})$. По полученному ЗВИ $MI(X_{i\_next}, X_{j\_next})$ и структурой (структурами) g^* строится множество моделей G в виде: { $(g^*;\ i\_next \rightarrow j\_next)$, $(g^*;\ i\_next \leftarrow j\_next)$, $(g^*;\ i\_next$ не зависит от $j\_next)$}. После этого выполняется шаг 2.

Условие завершения алгоритма.

Эвристический метод буде выполнятся до тех пор, пока не будет выполнен анализ всех элементов множества Set_MI. Напомним, что Set_MI содержит $\frac{N \cdot (N-1)}{2}$ элементов.

Выходные данные алгоритма

Оптимальна структура (структуры) g^*.

Пример использования эвристического метода

В качестве примера рассмотрим сеть "Азия" которая состоит из восьми вершин. В табл. 11 приведены значения ЗВИ между всеми восьмью вершинами сети расположенными по убыванию (первый этап алгоритма).

Построение выполнялось на выборке из 7000 обучающих наблюдений. На рис. 5 приведена структура оригинальной БС, на основе которой генерировались значения.

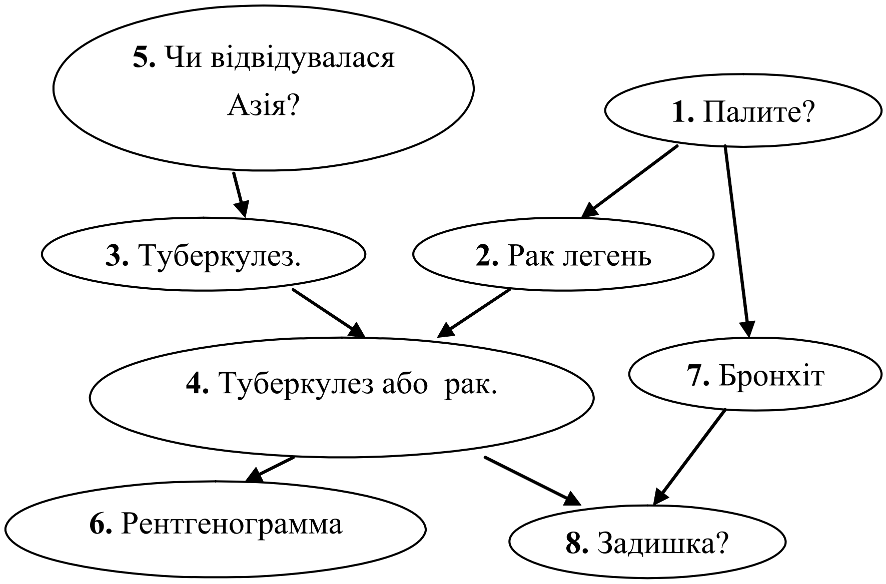


Рис. 5 Оригинальная структура сети "Азия"

Таблица 11 ЗВИ между всеми вершинами БС “Азия”

№	***MI***	***i***	***j***		**№**	***MI***	***i***	***j***
1	0,251	7	8		**15**	0,001227	3	5
2	0,136	2	4		**16**	0,000851	1	6
3	0,125	4	6		**17**	0,000508	2	7
4	0,096	2	6		**18**	0,000381	3	7
5	0,048	1	7		**19**	0,000266	4	5
6	0,036	3	4		**20**	0,000197	1	5
7	0,025	3	6		**21**	0,000128	4	7
8	0,0245	1	8		**22**	0,00012271	2	5
9	0,0132	4	8		**23**	0,00006475	5	6
10	0,0101	2	8		**24**	0,00003950	2	3
11	0,0051	6	8		**25**	0,00003249	5	7
12	0,0031	1	2		**26**	0,00001725	5	8
13	0,0028	3	8		**27**	0,00000303	1	3
14	0,0022	1	4		**28**	0,00000074	6	7

На 1-й итерации по первому значению *MI(7,8)*, отсортированной матрицы *MI*, строится множество моделей из 3 структур, после перебора этого множества выяснилось, что минимальному значению функции ОМД соответствует структуре приведенной на рисунке 6.

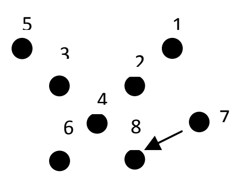
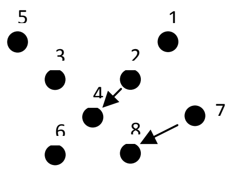
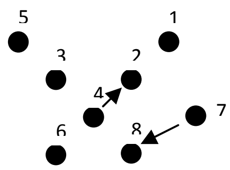


Рис. 6 Структура БС полученной на 1-й итерации

На 2-й итерации по полученной на 1-й итерации структуре (рис. 6) и *MI(2,4)* строится множество моделей из 3 структур, из которых минимальное значение функции ОМД принимают сети изображенные на рис. 7.



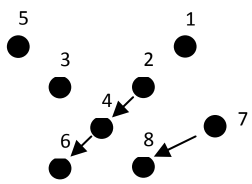


а *L=17820,9586* б

Рис. 7 Структуры БС полученные на 2-й итерации

На 3-й итерации по полученным на 2-й итерации структурам (рис. 7) и *MI(4,6)* строится множество моделей из 6 структур, из которых минимальное значение функции ОМД принимает сеть, изображенная на рис. 8.

На 4-й итерации по полученной на 3-й итерации структуре (рис. 8) и *MI*(2,6) строится множество моделей из 3 структур. Из трех структур минимальное значение функции ОМД принимает та же самая структура, изображенная на рис. 8.



L=16918,4958

Рис. 8 Структуры БС полученные на 3-й и 4-й итерации

На 5-й итерации по полученной на 4-й итерации структуре (рис. 8) и *MI(1,7)* строится множество моделей из 3 структур, среди которых минимальное значение функции ОМД принимают сети, изображенные на рис. 9.

а $L=16570{,}5342$ б

Рис. 9 Структуры БС полученные на 5-й итерации

На 6-й итерации по структурам, полученным на 4-й итерации (рис. 9) и *MI(3,4)* строится множество моделей из 6 структур, среди которых минимальное значение функции ОМД принимают сети изображенные на рис. 10.

На 7-й итерации по структурам, полученным на 6-й итерации и *MI(3,6)* строится множество моделей из 6 структур. В результате получаются те же самые структуры, что и на 6-й итерации (рис. 10).

На 8-й итерации по структурам, полученным на 6-й итерации и *MI(1,8)* строится множество моделей из 6 структур. Результат совпадает с результатом предыдущих двух итераций (рис. 10).

а $L=16312{,}9052$ б

Рис. 10 Структуры БС полученные на 6-й, 7-й и 8-й итерациях

На 9-й итерации по структурам, полученным на 8-й итерации (рис. 3.8) и *MI(4,8)* строится множество моделей з 6 структур, среди которых минимальное значение функции ОМД принимают сети изображенные на рис. 11.

На 10-й итерации по структурам, полученным на 8-й итерации и *MI(2,8)* строится множество моделей из 6 структур. В результате получены те же структуры, что и на 9-й итерации (рис. 11).

На 11-й итерации по структурам, полученным на 9-й итерации и *MI(6,8)* строится множество моделей из 6 структур. Результат совпадает с результатом предыдущих двух итераций (рис. 11).

а *L*= 16153,4686 б

Рис. 11 Структуры БС полученные на 9-й, 10-й та 11-й итерациях

На 12-й итерации по структурам, полученным на 11-й итерации (рис. 11) і *MI(1,2)* строится множество моделей из 6 структур, из которых минимальное значении функции ОМД принимают сети изображенные на рис. 3.12.

На 13-й итерации по структурам, полученным на 12-й итерации и *MI(3,8)* строится множество моделей из 6 структур. В результате получены те же самые структуры, что и на 12-й итерации (рис. 3.12).

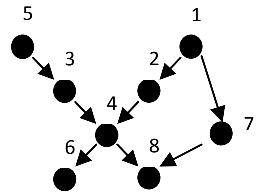
На 14-й итерации по структурам, полученным на 12-й итерации и *MI(1,4)* строится множество моделей из 6 структур. Результат совпадает с результатом предыдущих двух итераций (рис. 3.12).

а *L=16143,5518* б

Рис. 12 Структуры БС полученные на 12-й, 13-й и 14-й итерациях

На 15-й итерации по полученным на 14-й итерации структурам (рис. 3.12) и *MI(3,5)* строится множество моделей из 6 структур, из которых минимальное значение функции ОМД принимает сеть изображенная на рис. 13.

С 16-й по 28-ю итерации никаких изменений в структуре не происходит, то есть результат совпадает с результатом, полученным на 15 итерации (рис. 13).



L=16139,1545

Рис. 3.13 Структура БС полученная с 15 по 28 итерации

Для построения БС "Азия" при простом анализе всех возможных нецикличных структур нужно будет выполнить оценку 783 702 329 343 моделей. Тогда как эвристический метод на 28 итерациях алгоритма выполнил анализ всего лишь 120 структур, причем уже на 15-й итерации, после анализа 81 структуры, метод выдает структуру, которая полностью совпадает с оригинальной (эталонной) сетью "Азия". То есть следующие 13 итераций метода не делают никаких изменений, потому что оптимальная структура уже найдена на 15 итерации.