NEWS
OF THENATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
PHYSICO-MATHEMATICAL SERIES
ISSN 1991-346X **https://doi.org/10.32014/2020.2518-1726.73**
Volume 4, Number 332 (2020), 119 – 126

O. Mamyrbayev[1,2], N. Litvinenko[1], A. Shayakhmetova [1,2]

[1] Institute of Information and Computational Technologies, Almaty, Kazakhstan;
[2] Al-Farabi Kazakh National University, Almaty, Kazakhstan.
E-mail: morkenj@mail.ru, n.litvinenko@inbox.ru, asemshayakhmetova@mail.ru

EVIDENCES PROPAGATIONS IN BAYESIAN NETWORKS

Abstract. This paper is devoted to some problems of the distribution of several evidences in Bayesian networks. Currently, there are many different algorithms for calculations in Bayesian networks. Unfortunately, the description of most algorithms is either absent or only the idea of algorithms is described. Not only algorithms but also ideas for constructing these algorithms are quite complex. Many questions arise in the process of considering these algorithms remain unanswered. Some of them can be understood by testing the appropriate software, but many questions remain unanswered.

We use the idea of dividing the set of network vertices into sets by analogy using the concept of “Generation”. The concept of "Generation" is convenient to use in the absence of evidence. The presence of evidence requires a rather complicated adjustment of this concept. However, as a result, the propagation of evidence becomes more visible, and the corresponding algorithms are greatly simplified.

The presence of several evidences in some cases leads to contradictions, which solutions should be provided for by the algorithms of the Bayesian network nodes calculations. The modified concept of “Generation” allows one to find more visual and adequate approaches to resolving contradictions.

Keywords: Bayesian networks, oriented graphs, generation, propagating.

Introduction. Since the beginning of 21st century Bayesian Networks is the most popular tool of artificial intelligence in different researches. Models that use a Bayesian networks are usually insensitive to wrong, incomplete, and redundant data. Bayesian networks allow the use of heterogeneous data in various studies.

Bayesian networks, as a tool for studying models with uncertainties, is considered by many authors. Pearl J. was the first one who considered more completely the Bayesian networks tool in his works [1] and [2]. The Bayesian network theory is described quite well in [3], [4], [5].

The use of Bayesian networks in practice is practically impossible without the use of computer technology and related software. Currently, there are many programs for work with Bayesian networks. For example, BayesiaLab ([6], [7], [8]), AgenaRisk ([9], [10], [11], [12], [13], [14]), Hugin Expert. Unfortunately, we could not find a description of the algorithms for calculations in Bayesian networks which were used in these works.

Main definitions. Bayesian network theory is based on some sections of probability theory and graph theory. The definitions and concepts of graph theory used in BN theory can be found in [15], [16], [17]. The necessary concepts in probability theory can be found in [18], [19], [20]. The basic principles of BN theory can be found in [3], [4], [5].

For brevity, we will not give them here.

Use of the concept «Generation». We can distinguish two types of generations – generations of descendants and generations of ancestors. For a generation of descendants, this is a set of vertices which have parents only from earlier generations (or do not have parents at all), and have children only in later generations (or do not have children at all).

For a generation of ancestors, this concept is similar - it is a set of vertices which have children only in later generations (or do not have children at all) and have parents only from earlier generations (or do not have parents at all).

The only difference is that the construction of generations of descendants begins with nodes that do not have parents, and the construction of generations of ancestors begins with nodes that do not have descendants.

Generations of descendants are constructed, starting from vertices which do not have parents.

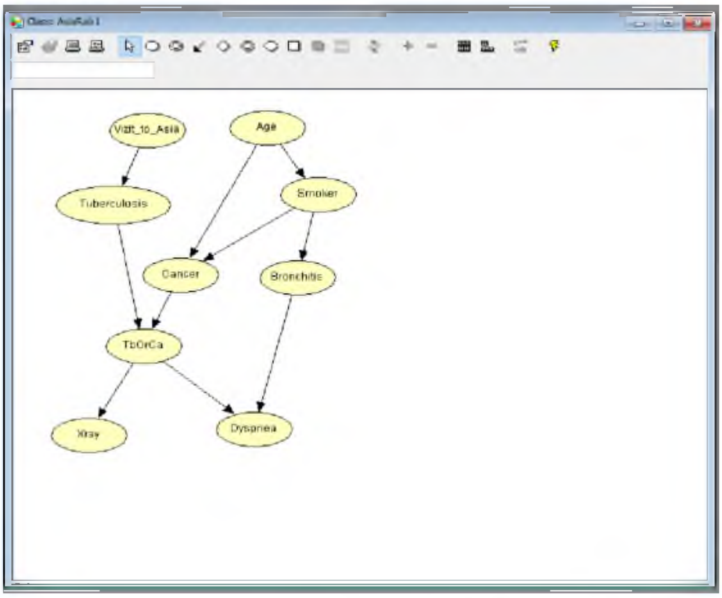

Figure 1 - Partitioning into generations of descendants

Definition. Generations of descendants are defined as follows:

- Nodes without parents belong to the 0 generation of descendants.
- Nodes with only 0 generation of parents belong to 1 generation of descendants.
- Nodes with only 0 and 1 generation of parents belong to 2 generation of descendants.
- ………….
- Nodes with 0, 1, 2, … K generation of parents belong to K+1 generation of descendants.
- ………….

The example of partitioning into generations of descendants is shown in figure 1.

Generations of descendants for a given Bayesian network:

- Vertices Age and Visit_to_Asia will be assigned to the 0 generation.
- Vertices Smoker and Tuberculosis will be assigned to the 1 generation.
- Vertices Cancer and Bronchitis will be assigned to the 2 generation.
- The only vertex TbOrCa will be assigned to the 3 generation.
- Vertices XRay и Dyspnea will be assigned to the 4 generation.

Generations of ancestors are constructed starting from vertices which do not have children.

Definition. Generations of ancestors are defined as follows:

- Nodes with no children belong to the 0 generation of ancestors.
- Nodes with only 0 generation of children belong to 1 generation of ancestors.
- Nodes with only 0 and 1 generation of children belong to the 2 generation of ancestors.
- ………….
- Nodes with only 0, 1, 2, … K generation of children belong to the K+1 generation of ancestors.
- ………….

The example of partitioning into generations of ancestors is shown in figure 2.

Generations of ancestors for a given Bayesian network:

- Vertices XRay and Dyspnea will be assigned to the 0 generation.
- Vertices TbOrCa and Bronchitis will be assigned to the 1 generation.

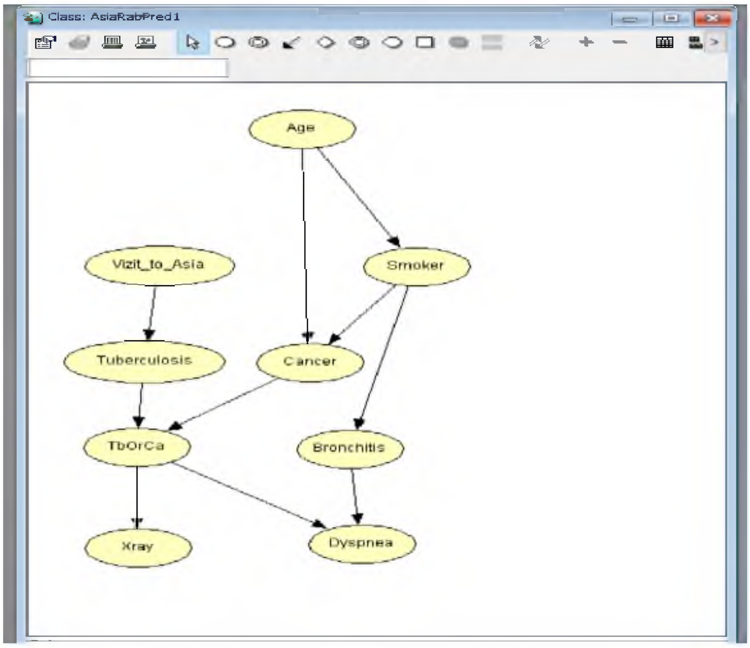

Figure 2 - Partitioning into generations of ancestors

- Vertices Cancer and Tuberculosis will be assigned to the 2 generation.
- Vertices Smoker and Visit_to_Asia will be assigned to the 3 generation.
- The only vertex Age will be assigned to the 4 generation.

In most cases, the partitioning of the set of vertices into generation of descendants and generation of ancestors is significantly different. However, it is not difficult to come up with a graph (Bayesian network) in which the partitioning into generations of descendants and generations of ancestors is completely identical.

Propagation of the several evidences with the use of "Generation" concept

Initialization of a Bayesian network, i.e. calculation of node values from the values of unconditional variables (nodes without parents) and conditional probability tables is a fairly simple task. Nevertheless, we briefly describe the initialization algorithm using the concept of "Generation".

1. We separate the vertices of the BN into generations of descendants, as described above. Recall that the 0 generation consists of vertices without parents. For nodes of 0 generation calculations are not needed.
2. We will calculate the values for all nodes of the 1 generation, using the values of the nodes of the 0 generation and the formula for total probability.
3. We will calculate the values of the nodes of the 2 generation, using the values of the nodes of the 0 generation and the calculated values of the nodes of the 1 generation, as well as the formula for the total probability.
4. Similarly, we will calculate the values of the nodes of the third generation.
5. … etc.

At some point, some nodes in the Bayesian network get evidence. It is required to recalculate the values of the network nodes considering the obtained evidence. One of the difficult issues is determining the order of calculation of Bayesian network nodes. Obviously, the nodes that obtained the evidence do not need calculations. Also, nodes that do not have parents do not need calculations, if among the descendants of these nodes there are no nodes that have obtained evidence.

It would be interesting to collect in one set all the nodes for which we do not need to make calculations. This set will be called the zero level of the Bayesian network nodes.

In the next set we will collect all the nodes for the calculation of the values of which there is enough information from the nodes of the zero level. This set will be called the first level of the Bayesian network nodes.

In the next set we will collect all the nodes, for the calculation of the values of which there is enough information from the nodes of the zero and first levels. This set will be called the second level of the Bayesian network nodes.

In the next set we will collect all the nodes, for the calculation of the values of which there is enough information from the nodes of the zero, first and second levels. This set will be called the third level of the Bayesian network nodes.Etc.

Let us consider in more detail the construction of various levels of Bayesian network nodes. The nodes that do not require calculations (nodes of the zero level):

1. Nodes that obtained evidences.
2. All descendants and ancestors of the nodes that obtained evidences must be recalculated. It is also necessary to recalculate all nodes that are descendants of all the ancestors of the nodes that obtained certificates. These nodes do not belong to 0 level. We remove these nodes from the Bayesian network. We also remove the nodes that obtained evidence. Then we consider the remaining nodes of the Bayesian network. These nodes together with arcs form a certain subnetwork of the studied network. From this subnet, we select nodes that do not have parents. These nodes are also included in the 0 level.

To determine the first, second and other levels, we separate the set of Bayesian network nodes into two subsets (two subnets): BN1 and BN2. BN1 subnet includes nodes that obtained evidences, these nodes' descendants, ancestors, and also the descendants of their ancestors. We assign the remaining nodes to the BN2 subnet. In the BN2 subnet we define "Generations" as described earlier. If the node belongs to some generation, then we will assume that this node belongs to the same level. It is easy to understand that for the calculation of a node of a generation, information on other nodes of a given generation, older generations, as well as on the nodes of the BN1 subnet, is not required. Only data on the nodes of the previous levels (maybe not all) are needed.

Some nodes (BN2 subnet) are already distributed across levels. Then we divide the set of nodes of subnetwork BN1 by levels. First, we replace all the BN1 nodes that are the ancestors of the nodes that obtained evidences with descendant nodes. To do this, we change the direction of the corresponding arcs. Let's denote the new subnet by BN1A. This subnet differs from the BN1 subnet only in the direction of some arcs. The BN1A subnet will no longer have the ancestors of the nodes obtained evidences - all nodes will be descendants of the nodes obtained evidences. We divide the BN1A subnet into generations, as described above. Then we assign a level number equal to a generation number to each node.

The nodes of the original Bayesian network are distributed at the same levels as the same nodes in the auxiliary networks BN1A and BN2.

Thus, all nodes of the Bayesian network will be separated by levels (zero, first, second, etc.). Nodes belonging to level zero do not need calculations. For the calculation of nodes of the first level we need only data of some nodes of the zero level. For the calculation of nodes of the second level, we need only the data of some nodes of the zero and first levels. Etc. For calculations we will mainly use the law of total probability and Bayes formula.

Below we will provide 2 examples of distribution by levels of the training Bayesian network. The network in both examples is the same, but the nodes that obtained evidences are different. Each level in the figure is located on a separate line. The first line contains nodes of the zero level.

Example 1.

Figure 3a shows some Bayesian network. We have obtained 3 evidences for this network for nodes C3, C4 and C6. In what order should we make calculations? To determine the order of calculations, we separate the set of BN nodes into several levels. We refer to the zero level those nodes in which calculations are not necessary. Obviously, these will be the nodes that obtained evidences (nodes C3, C4 and C6), as well as the nodes that do not have parents, and the nodes among whose descendants there are no nodes that obtained evidences (nodes C8 and C10). In figure 3b, the zero level is in the top line.

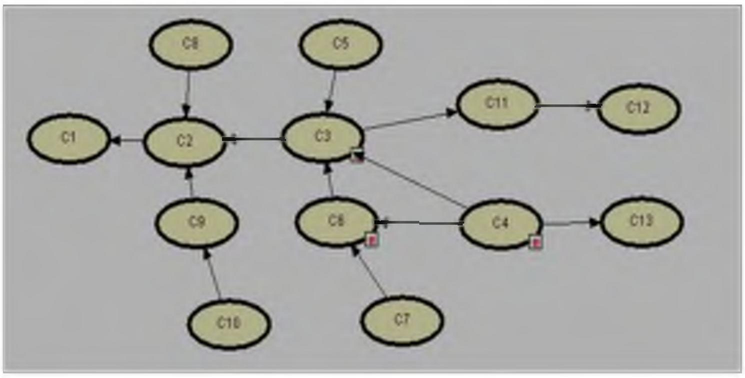

Figure 3a - Example 1

The first level includes nodes which calculation is required the only information of zero level. These will be nodes C5, C11, C13, C7 and C9. In figure 3b, these nodes are in the second line.

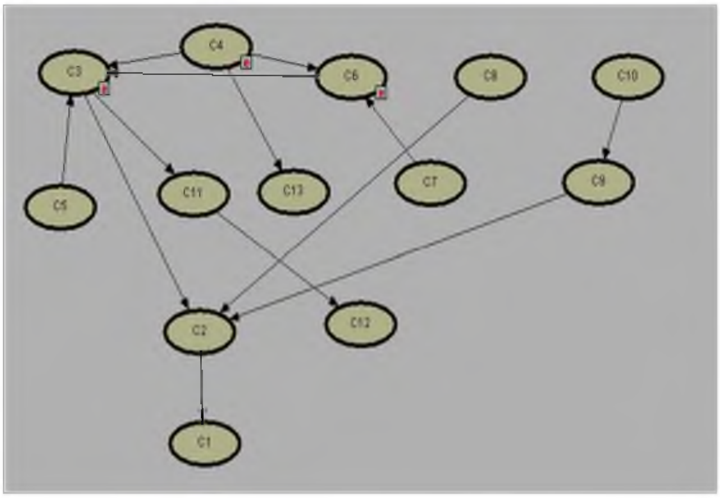

Figure 3b - Example 1

The second level includes nodes which calculation is required information of the zero and first levels. These will be nodes C2 and C12. In figure 3b, these nodes are in the third line.

The third level includes nodes which calculation is required information of the zero, first and second levels. These will be the only node C1. In figure 3b, these nodes are in the fourth line.

The order of calculations:

1. Calculation of nodes of the first level using data from some nodes of the zero level.
2. Calculation of nodes of the second level using data from some nodes of the zero and first levels.
3. Calculation of nodes of the third level using data from some nodes of the zero, first and second levels.

Example 2.

We consider the same Bayesian network as in the figure 3a. However, we obtain 3 evidences for the nodes C8, C10 and C13. To determine the order of calculations, we partition the set of BN nodes into several levels. We refer to the zero level those nodes in which calculations are not necessary. Obviously, it will be nodes, obtained evidences (nodes C8, C10 and C13), as well as the nodes that do not have parents, and the nodes among whose descendants there are no nodes that obtained evidences (nodes C5 and C7). In figure 4a, the zero level is in the top line.

The first level includes nodes which calculation is required the only information of zero level. These will be nodes C9 and C4. In figure 4a, these nodes are in the second line.

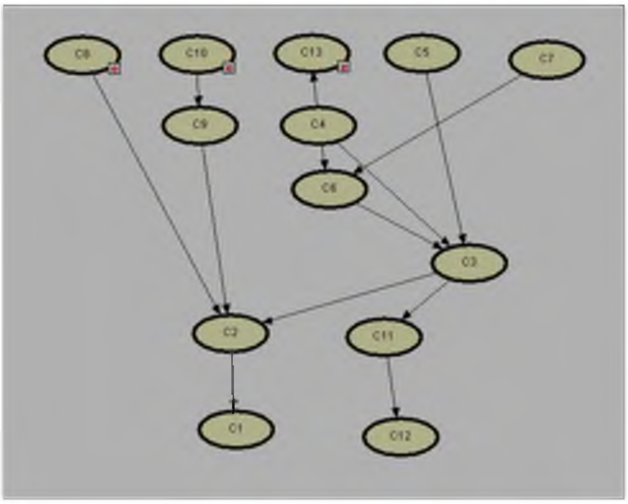

Figure 4a - Example 2

The second level includes nodes which calculation is required information of the zero and first levels. These will be the only node C6. In Figure 4a, these nodes are in the third line.

The third level includes nodes which calculation is required information of the zero, first and second levels. These will be the only node C3. In figure 4a, these nodes are in the fourth line.

The fourth level includes nodes which calculation is required information of the zero, first, second and third levels. These will be nodes C2 and C11. In Figure 4a, these nodes are in the fifth line.

The fifth level includes nodes which calculation is required information of the zero, first, second, third and fourth levels. These will be nodes C1 and C12. In Figure 4a, these nodes are in the sixth line.

The order of calculations:

Steps 1-3 we repeat as in the Example 1

4. Calculation of nodes of the fourth level using data from some nodes of the zero, first, second and third levels.
5. Calculation of nodes of the fifth level using data from some nodes of the zero, first, second, third and fourth levels.

Conclusion. This paper shows one of the options for determining the propagation order of Bayesian network nodes in the process of obtaining evidence with use the idea of partitioning of Bayesian network nodes into independent sets (generations, levels). Nodes, which belong to zero level do not need calculations. For the calculation of nodes of the next level, we use only the data of some nodes of the previously calculated levels.

This method allows to develop simpler algorithms for calculations in networks, some nodes of which have obtained evidences.

Acknowledgement

The work was performed under the grant of the Committee of Science of the Ministry of Education and Science of the Republic of Kazakhstan (2018-2020), on the theme “Development and software implementation of an application package for solving applied problems on Bayesian networks.

Ө. Мамырбаев[1,2], Н. Литвиненко[1], Ә. Шаяхметова [1,2]

[1]Ақпараттық және есептеуіш технологиялар институты, Алматы, Қазақстан;
[2]Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан

БАЙЕС ЖЕЛІЛЕРІНДЕ КУӘЛІКТЕРДІ НАСИХАТТАУ

Аннотация. Байес желісі көптеген айнымалыларды, сондай-ақ осы айнымалылардың арасындағы әртүрлі ықтималдық тәуелділіктерді сипаттайтын графикалық ықтималдық моделі. Байес желілерінің жалпы математикалық аппаратын американдық ғалым, Тьюринг сыйлығының лауреаты Pearl J құрды.

Байес желісі сұраулардың әртүрлі типтеріне жауап алуға мүмкіндік береді:

- Куәләндірудің ықтималдық бағасы
- Априорлық маргиналдық ықтималдықтарды бағалау.
- Апостериорлық маргиналдық ықтималдықтарды есептеу.
- Апостериорлық максимум есебі.
- Ықтимал болатын оқиғаны түсіндіруді зерттеу.

Байес желілеріндегі есептеулер өте күрделі және көлемді. 10 тораптан тұратын байес желілеріндегі есептеулер, әдетте, есептеу техникасын пайдалануды, алгоритмдерді дайындауды, бағдарламалық кодта алгоритмдерді іске асыруды талап етеді. Байес желілерінде есептеулерді жүргізу кезінде әртүрлі алгоритмдер қолданылады, оларды келесідей жіктеуге болады:

- Толық аралықтар немесе өрескел күш әдісі. Толық таңдаудың кемшіліктеріне, әдетте, тапсырманы шешуге кететін уақыттың жеткілікті үлкен шығындарын жатқызуға болады.
- Кластерлеудің түрлі идеяларын пайдаланатын алгоритмдер. Бұл алгоритмдер жиі өрескел күш әдісімен салыстырғанда жақсы уақыт ұтысын береді. Әдетте, осы алгоритмдерге салынған идеялар жеткілікті ашық және түсінікті.
- Түйіндер арасындағы ақпаратты беру(пропагациялау) идеяларын пайдаланатын алгоритмдер. Бұл алгоритмдер графтар теориясында жақсы білімді талап етеді.
- Әртүрлі сұрыптауды құруда қолданатын алгоритмдер
- Monte Carlo әдісі идеяларын пайдаланатын алгоритмдер

Мақала Байес желілерінде бірнеше дәлелдемелерді таратудың кейбір мәселелеріне арналған. Қазіргі уақытта Байес желілерінде есептеулердің көптеген алгоритмдері бар. Өкінішке орай, алгоритмдердің көпшілігінің сипаттамасы қарастырылмаған немесе тек алгоритм идеялары сипатталған. Алгоритмдер ғана емес, сонымен қатар бұл алгоритмдерді құру идеялары да өте күрделі. Осы алгоритмдерді қарастыру кезінде оқырманға қойылған көптеген сұрақтар жауапсыз қалады. Байес желілерінде алгоритмдерді құру кезінде дәлелдерді тарату мәселесі маңызды болып табылады.

Желілік шындардың жиынтығын жиынтықтарға бөлу идеясы аналогия арқылы «Ұрпақ» ұғымын қолданады. «Ұрпақ» ұғымы дәлел болмаған кезде қолдануға ыңғайлы. Дәлелдердің болуы осы тұжырымдаманы едәуір күрделі түзетуді талап етеді. Алайда, нәтижесінде дәлелдердің таралуы айқынырақ болады, сәйкесінше алгоритмдер айтарлықтай жеңілдетілген.

Кейбір жағдайларда бірнеше дәлелдердің болуы қарама-қайшылықтарға әкеледі, олардың шешімі Байес желілеріндегі есептеу алгоритмдерімен қамтамасыз етілуі керек. «Ұрпақ» тұжырымдамасы туындайтын қайшылықтарды шешуге көрнекі және барабар тәсілдерді табуға мүмкіндік береді.

Түйін сөздер: Байестік желілер, бағытталған графтар, генерация, тарату.

О. Мамырбаев[1,2], Н. Литвиненко[1], А. Шаяхметова [1,2]

[1]Институт информационных и вычислительных технологии, Алматы, Казахстан;
[2] Казахский национальный университет имени аль-Фараби, Алматы, Казахстан

ПРОПАГАЦИЯ СВИДЕТЕЛЬСТВ В БАЙЕСОВСКИХ СЕТЯХ

Аннотация. Байесовская сеть представляет собой графовую вероятностную модель, описывающую множество переменных, а также различные вероятностные зависимости между этими переменными. Общий математический аппарат байесовских сетей разработан американским учёным Pearl J, лауреатом премии Тьюринга.

Байесовская сеть позволяет получить ответы на различные типы запросов:

- Вероятностная оценка свидетельств.
- Оценка априорных маргинальных вероятностей.
- Расчет апостериорных маргинальных вероятностей.
- Расчет апостериорного максимума.
- Исследование наиболее вероятного объяснения события.

Расчеты в байесовских сетях достаточно сложны и объемны. Расчеты в байесовских сетях, содержащих более 10 узлов, как правило уже требуют использования вычислительной техники, разработки алгоритмов, реализации алгоритмов в программном коде. При проведении расчетов в байесовских сетях используются различные алгоритмы, которые можно классифицировать, например, следующим образом:

- Полный перебор или метод грубой силы. К недостаткам полного перебора можно отнести, как правило, достаточно большие затраты времени на решение задачи.

- Алгоритмы, использующие различные идеи кластеризации. Данные алгоритмы часто дают хороший выигрыш времени по сравнению с методом грубой силы. Обычно идеи, заложенные в данных алгоритмах достаточно прозрачны и понятны.
- Алгоритмы, использующие идеи передачи(пропагации) информации между узлами. Данные алгоритмы требуют достаточно хороших знаний в теории графов.
- Алгоритмы, использующие построение различных выборок.
- Алгоритмы, использующие идеи метода Monte Carlo.

Статья посвящена некоторым проблемам распространения нескольких свидетельств в байесовских сетях. В настоящее время существует множество различных алгоритмов для расчетов в байесовских сетях. К сожалению описание большинства алгоритмов либо отсутствует, либо описывается лишь идея алгоритмов. Не только алгоритмы, но также и идеи построения данных алгоритмов достаточно сложны. Многие вопросы, возникающие у читателя при рассмотрении данных алгоритмов, остаются без ответа. Что-то можно понять, тестируя соответствующее программное обеспечение, но многие вопросы остаются без ответа.

Используется идея разбиения множества вершин сети на множества по аналогии с использованием понятия «Поколение». Понятие «Поколение» удобно использовать при отсутствии свидетельств. Наличие свидетельств требует достаточно сложной корректировки данного понятия. Однако в результате пропагация свидетельств становится более наглядной, а соответствующие алгоритмы значительно упрощаются.

Наличие нескольких свидетельств в некоторых случаях приводит к противоречиям, решение которых должно быть предусмотрено алгоритмами расчетов узлов байесовской сети. Модифицированное понятие «Поколение» позволяет находить более наглядные и адекватные подходы к разрешению возникающих противоречий.

Ключевые слова: Байесовские сети, ориентированные графы, генерация, распространение.

The information about the authors:

Mamyrbayev O., PhD, deputy general Director, Institute of Information and Computational Technologies, Almaty, Kazakhstan, Al-Farabi Kazakh National University, 050040 Almaty, Kazakhstan. https://orcid.org/0000-0001-8318-3794;

Litvinenko N., junior researcher, Institute of Information and Computational Technologies, Almaty, Kazakhstan, n.litvinenko@inbox.ru, https://orcid.org/0000-0002-0576-8305;

Shayakhmetova A., PhD, senior researcher, Al-Farabi Kazakh National University, 050040 Almaty, Kazakhstan, asemshayakhmetova@mail.ru, https://orcid.org/0000-0002-4072-3671

REFERENCES

[1] Pearl J., How to Do with Probabilities what People Say You Can't // Artificial Intelligence Applications / Ed. Weisbin C.R. IEEE, North Holland. 1985. P. 6–12.

[2] Pearl J., Probabilistic Reasoning in Intelligent Systems. San Francisco: Morgan Kaufmann Publishers. 1988. 552 p.

[3] Jensen F.V., Nielsen T.D. Bayesian Networks and Decision Graphs. Springer. 2007. 447 p.

[4] Barber D., Bayesian Reasoning and Machine Learning. 2017. 686 p. http://web4.cs.ucl.ac.uk/staff/D.Barber/textbook/020217.pdf

[5] Neapolitan R.E., Learning Bayesian Networks. 704 p. http://www.cs.technion.ac.il/~dang/books/Learning%20Bayesian%20Networks(Neapolitan,%20Richard).pdf

[6] Conrady S., Jouffe L., Bayesian Networks and BayesiaLab, A Practical Introduction for Researchers. https://www.researchgate.net/publication/282362899_Bayesian_Networks_ BayesiaLab_-_A_Practical_Introduction_for_Researchers

[7] BayesiaLab User Guide. https://library.bayesia.com/display/BlabC/BayesiaLab+User+ Guide

[8] Litvinenko N., Litvinenko A., Mamyrbayev O., Shayakhmetova A., Work with Bayesian Networks in BAYESIALAB. Almaty: IPIC, 2018. 311 p. (in Rus). ISBN 978-601-332-206-3.

[9] Fenton N. and Neil M., Risk Assessment and Decision Analysis with Bayesian Networks. Queen Mary, University of London and Agena Ltd. CRC Press. ISBN: 9781439809105, ISBN 10: 1439809100

[10] AgenaRisk 7.0 User Manual. https://www.agenarisk.com

[11] Advanced Modelling with AgenaRisk. https://www.agenarisk.com

[12] Agena's Bayesian Network Technology. https://www.agenarisk.com

[13] Basic Modelling with AgenaRisk. https://www.agenarisk.com

[14] N. Litvinenko, A. Litvinenko, O. Mamyrbayev, A. Shayakhmetova, AgenaRisk. Work with Bayesian Networks. Almaty: IPIC, 2019. 236 p. (in Rus). ISBN 978-601-332-335-0.

[15] D.V. Karpov, Graph theory. https://logic.pdmi.ras.ru/~dvk/graphs_ dk.pdf (in Rus)

[16] O. Ore. Graph theory. M.: Science. 1980. 336 p. (in Rus)

[17] Ph. Kharari, Graph theory. M.: Mir. - 1973. 300 p. (in Rus)

[18] Gmurman V.E., Probability Theory and Mathematical Statistics: Manual, M., 2003. 479 p. (in Rus)

[19] Kolmogorov A.N., Basic Concepts of Probability Theory: Manual. M: Science, 1974. 412 p. (in Rus)

[20] Mamyrbayev, O. Zh., Shayakhmetova, A. S., Seisenbekova, P. B. The methodology of creat-ing an intellectual environment of increasing the competence of students based on a bayesian ap-proach // News of the National academy of sciences of the republic of Kazakhstan. Series physico-mathematical. 2019. Vol. 4(326). P. 50-58.