

A.S. Shayakhmetova¹, M. Satymbekov¹

¹Институт информационных и вычислительных технологий,
г.Алматы, Казахстан

ANALYSIS OF THE MODERN SOFTWARE FOR BAYESIAN NETWORKS

Abstract. This article provides an analytical review of the most common modern software for Bayesian networks. The history of development of Bayesian networks was considered. The relevance of the application of Bayesian networks in various areas of research has been featured. The most common software for Bayesian networks were considered. The comparative analysis of their characteristics was carried out and tools of development were considered. There were identified the features of the functioning of Bayesian networks in the financial sphere for risk research, in development of various expert systems, artificial intelligence systems, etc. The advantages of Bayesian networks (take into account cause-effect relationships, uncertainties, the ability to work with incomplete and inaccurate information and take into account random factors) are shown in creation modern intellectual technologies. The advantages of application of Bayesian networks and the prospects of their using for development of complex applications have been analyzed.

Keywords: Bayesian networks, analytical review, software.

• • •

Аннотация. В данной статье проведен аналитический обзор наиболее распространенных современных программных продуктов по байесовским сетям. Рассмотрена история развития байесовских сетей. Отмечена актуальность применения байесовского подхода в различных областях исследования. Рассмотрены наиболее распространённые программные продукты по байесовским сетям. Проведен сравнительный анализ их характеристик и рассмотрены инструменты разработки. Выделены особенности функционирования приложений по байесовским сетям в финансовой сфере для исследования рисков, при разработке различных экспертных систем, систем искусственно-

го интеллекта и др. Показаны достоинства байесовского подхода (учитывают причинно-следственные связи, неопределенности, возможность работать с неполной и неточной информацией и учитывают случайные факторы) при создании современных интеллектуальных технологий. Проанализированы особенности применения байесовских сетей и перспективы их использования для разработки сложных приложений.

Ключевые слова: байесовские сети, аналитический обзор, программный продукт.

• • •

Түйіндеме. Зерттеудің бұл мақалада байестік желі бойынша кең таралған заманауи программалық өнімдерге аналитикалық шолу жасалды. Байестік желілердің даму тарихы келтірілген. Зерттеудің түрлі салаларында байестік тәсілдердің қолданылуының өзектілігі қарастырылған. Байестік желілер бойынша ең кең таралған программалық өнімдер келтірілген. Олардың сипаттамаларына салыстырмалы талдау жүргізіліп, оларды құру құралдары қарастырылған. Әртүрлі сараптау жүйелері мен жасанды интеллект жүйелерін құруда, қаржы саласындағы тәуекелдерді зерттеуде байестік желі қосымшаларының ерекшелігі анықталды. Заманауи интеллектуалды технологияларды құруда байестік тәсілдердің (себеп-салдар байланысы, белгісіздік, толық емес және анық емес ақпаратпен жұмыс істеу мүмкіндігі, кездейсоқ факторларды ескеру) ерекшеліктері көрсетілді. Байестік желілерді қолданудың ерекшеліктері мен күрделі қосымшаларды құрудағы болашағына талдау жүргізілді.

Түйінді сөздер: байестік желілер, аналитикалық талдау, программалық өнім.

Introduction. Nowadays the interest in application of Bayesian networks (BS) in various fields of research has been grown. The application of BS in medicine, economics, sociology, psychology, ecology, information and communication technologies, bioinformatics and many other areas of science is actual problem [1]. Many applications of Bayesian networks in medicine have been developed. The possibility of using Bayesian network for differential diagnosis of arterial hypertension of various etiologies in a patient's primary treatment to a general practitioner was presented in work [2].

Bayesian networks are successfully used in the development of various expert systems and artificial intelligence systems [3]. The structure of knowledge of students is can show the structure of the. It is possible to show the structure of knowledge of students using the structure of Bayesian network. The publication [4] considers the use of Bayesian network in

the construction of model of students to assess the level of competence formation. The developed algorithms for construction competency models based on BS were described in detail. In article [5], the realization of Bayesian model of logical inference was implemented in the development of an expert system for assessing the compliance of the level of training of graduates of technical specialties with the requirements of modern employers in the field of IT services.

One of advantages of the application of Bayesian networks is possible to work with incomplete, inaccurate and inconsistent information. The paper [6] describes the application of Bayesian network of trust to assess economic risk and support decision making in a case uncertainty.

Bayesian networks are one of the most suitable models for dealing with incomplete, inaccurate and contradictory information. The study [7] describes the use of Bayesian networks to solve the problem of finding the place and type of failure of complex technical system. A method and algorithm for the optimal search for a place and type of failure of a complex technical system have been developed, which allows to reduce the time and other types of resources for conducting inspections significantly. Experimental researches confirm the effectiveness of this method.

Bayesian theory is named after Thomas Bayes and was published in 1763 [8]. Bayesian networks are widely used graphic models with a variety of applications applied in the field of knowledge detection, classification, forecasting and control [9]. The networks consist of two main parts [2]: a graphical structure, which defines a set of dependencies and independencies in a set of random variables and a set of probability distributions determining the strength of the dependency relationships that are encoded in the graphical structure.

There are various software for work with Bayesian networks of trust [10]. One of the features of using the apparatus of Bayesian networks of trust is visual visualization of the interrelations between the elements.

Overview of Software for Bayesian Networks. There is a large number of software for implementation of Bayesian network. The most famous software are GeNIe & SMILE [11], RISO [12], Bansys [13], Sam lam [14], Agena Risk [15], Bayesia [16], Hugin Expert [17], Netica [18], BNet [19] and others.

The most of the listed software are based on the programming language C++, Python and R. Each software has its own advantages. Most

of them are in open source and are free. Software in closed source do not allow to change the source code and are mostly paid. The most well-known software and their characteristics are presented in table 1.

Table 1 - The main software products for Bayesian networks

Name	Authors	Graph. interf. user	Free or Commercial	Which graphics are supported
GeNIe & SMILE	Decision systems laboratory	Yes	Free	Oriented
RISO	Dodier (U. Colorado)	Yes	Free	Oriented
Bansys	U. Helsinki	No	Free	Oriented
Sam lam	Darwiche(Ucla)	Yes	Free	Oriented
Agena Risk	Norman Fen-ton, Martin Neil	Yes	Free	Oriented
Bayesia	Bayesia ltd	Yes	Free	Chain diagrams
Hugin Expert	Hugin	Yes	Commercial	Chain diagrams
Netica	Norsys[9]	Yes	Commercial	Oriented
BNet	Murphy (U.C.Berkeley)[10]	Yes	Free	Oriented

GeNIe & SMILE [11]. The software product is a graphical interface for the SMILE library (Structural Modeling, Inference and Learning Engine), which provides the possibility of building Bayesian networks. In this software, the amount of designed network depends on the performance of the computer. The resulting models can be integrated into other systems and on most platforms.

RISO [12] - supports distributed trust networks that operate on different hosts and are integrated into one large trust network. The means of communication is Internet, and the Java Remote Method Invocation (RMI) is protocol.

Bansys [13]. Software for data mining with Bayesian networks. The system implements several algorithms for studying the topology of Bayesian networks. It has several modules (module of learning topology, module of learning parameter (probability), output module, correction

modulation module, classification module) that allow the user to perform data mining using Bayesian networks.

Sam lam [14]. A complex modeling and analysis tool based on Bayesian networks, developed in java is Sam lam. This software product includes two main components:

- A graphical interface that allows users to develop Bayesian network models and save them in different formats;

- The reasoning mechanism supports a variety of tasks, such as classical output, parameter estimation, time-based trade-offs and sensitivity analysis.

AgenaRisk uses the latest developments in the field of artificial intelligence and visualization to solve complex risk problems [20]. This product includes intelligent analytics and scaling the level of monitoring and risk assessment. It is ideal for planning risk scenarios. AgenaRisk, is used in many industries: banking, defense, aerospace, energy, telecommunications and technology [21].

Bayesia [16]. Bayesia software focuses on all aspects of decision support with Bayesian networks and includes BayesiaLab, BEST and BRICKS. These applications consist of individual decision support and large-scale policy analysis and risk assessment of industrial systems.

Hugin Expert is an advanced intelligent analytical solution from HUGIN EXPERT [17]. Hugin is a software implementation of decision-making system based on Bayesian networks of trust. It has two versions of Pro and Explorer. It is operates in the Windows operating system. It has a version for Unix. This system has a developed interface. The system allows to create knowledge bases and facts. The paper [22] presents a methodology for calculation project risks based on Bayesian networks using the Hugin software package. This package uses two main operating modes:

- mode of editing and building the cause-effect network;
- a mode of calculation of probabilistic estimations for decision-making on all events in the cause-effect network.

Bayesian networks provide a convenient apparatus for research risks, draw up their models and quantify them.

Netica [18]. The most widely used software for the development of Bayesian networks in the world is Netica. This software is simple, reliable and highly efficient. It is possible easily find patterns in data, create diagrams, code knowledge and create probabilistic expert systems using this product [23]. A study of the migration factors of rural youth based on Bayesian networks of trust is conducted in [24].

This research is conducted on the results of students' questionnaires. Based on the results of the questionnaire, a Bayesian network of trust was

developed. The network structure is trained on the basis of an assessment of the statistical significance between the factors. The final network includes 11 factors, the software part is implemented in the package NETICA, which has an intuitive and pleasant user interface. Application of Bayesian network technologies will allow to bring the results to a qualitatively new level. Netica is used to manage uncertainty in business, engineering, medicine or the environment. It is a tool of choice for many leading global companies and government agencies.

BNet [25]. Belief Networks is a powerful modeling tool that covers a large number of models for creating and learning BS. This software product implements most of the logical inference mechanisms and has open source code. Allows you to develop your own models and algorithms.

The conclusion. Thus, as a result of the review, the most known software products on Bayesian networks were considered and a comparative analysis of their characteristics was carried out. Examples of practical applications developed with the help of existing software in various fields of industry, science and education are given. There are several advantages of using the Bayesian network in applied research areas:

- allows to describe sufficiently well-defined cause-effect relationships;
- the ability to work with incomplete and inaccurate information;
- allows to bring the results to a qualitatively new level;
- high performance;
- presentation of results in an intuitive graph form.

In this direction, the work has started only. In the future, there will be practical applications with specific examples.

Source of research funding. The work has been performed under grant project of «Development and software implementation of a package for solving applied problems in Bayesian networks» (2018 - 2020).

References

1. *Toropova A.V., Suvorova A.V., Tulup'ev A.L.* Diagnostika soglasovannosti v modeli dlya otsenivaniya intensivnosti sotsial'no-znachimogo povedeniya // *Nechetkiye sistemy i myagkiye vychisleniya.* - 2015. - Vol. 10. - No 1. - P. 93-107.
2. *Dimitrova L.K., Golubeva O.A.* Primeneniye bayyesovskoy seti v differentsial'noy diagnostike arterial'noy gipertenzii // *Materialy III Mezhdunar. nauch. konf. Tekhnicheskoye nauki v Rossii i za rubezhom.* — M.: Buki-Vedi, 2014. — P. 4-14. — URL, [Electronic resource]:

<https://moluch.ru/conf/tech/archive/90/6030/> (date of the application: 14.03.2018).

3. *Maslennikov Ye.D., Sulimov V.B.* Predskazaniya na osnove bayesovskikh setey doveriya: algoritm i programmaya realizatsiya // Vychislitel'nyye metody i programmirovaniye. – 2010. – Vol. 11. – P. 94-107.

4. *Khlopotov M.V.* Primeneniye bayesovskoy seti pri postroyenii modeley obuchayushchikhsya dlya otsenki urovnya sformirovannosti kompetentsiy // Internet-zhurnal Naukovedeniye, 2014. – No 5 (24). – P. 1-28.

5. *Naseykina L.i., Sokolova I.M.* Razrabotka ekspertnoy sistemy otsenki professional'noy kompetentnosti budushchikh IT-spetsialistov // Materialy XXXVII Mezhdunar. nauch.-prakt. konf. «Ot teorii k praktike», Novosibirsk: SibAK, 2014. – No 8(33). – P. 148-154.

6. *Musina V.F.* Bayesovskiye seti doveriya kak veroyatnostnaya graficheskaya model' dlya otsenki ekonomicheskikh riskov // Trudy SPIIRAN. – 2013. – No 2(25). – P. 235-254.

7. *Fefelov A. A.* Ispol'zovaniye bayesovskikh setey dlya resheniya zadachi poiska mesta i tipa otkaza slozhnoy tekhnicheskoy sistemy // Informatsionno-izmeritel'nyye sistemy, 2013. – No 2. – P. 87-93.

8. Bayes theorem. [Electronic resource]: <http://m.diary.ru/~BratstvoStali/p207345776.htm>. (date of the application: 5.03.2018).

9. *Peter van Beek, Hella-Franziska Hoffmann* Machine Learning of Bayesian Networks Using Constraint Programming // International Conference on Principles and Practice of Constraint Programming, 2015. – P. 429-445.

10. *Murphy K.* A brief introduction to Graphical models and Bayesian Networks [Electronic resource]: site The University of British Columbia – Electron. text. data – 2010. [Electronic resource]: <http://www.cs.ubc.ca/~murphyk/Bayes/bnintro.html> (date of the application: 10.01.2018).

11. GeNIe & SMILE, [Electronic resource]: <http://genie.sis.pitt.edu>. (date of the application: 8.02.2018).

12. *Zvyagin L.S.* Primeneniye bayesovskogo podkhoda v izmereniyakh analiticheskikh dannykh kak faktor formirovaniya protsessov sistemnogo ekonomicheskogo razvitiya // Molodoy uchenyy, 2017. – №22. – P. 256-261.

13. *Murphy K.* Software Packages for Graphical Models URL: [Electronic resource]: <http://www.cs.ubc.ca/~murphyk/Software/bnsoft.html> (date of the application: 20.05.2016).

14. *Borisov V.V, Zakharov A.S.* Approximate reasoning based on temporal fuzzy Bayesian belief networks //Software & systems, 2016. – No 2. – P. 27-33.

15. AgenaRisk Bayesian network tool: [Electronic resource]: <http://www.agenarisk.com> (date of the application: 27.02.2018).

16. Bayesia, [Electronic resource]: <http://www.bayesia.com> (date of the application: 1.03.2018).

17. Hugin, [Electronic resource]:<http://www.hugin.com> (date of the application: 5.03.2018).

18. Netica, [Electronic resource]: <http://www.norsys.com> (date of the application: 14.03.2018).

19. BNet, [Electronic resource]: <http://www.cra.com/bnet> (date of the application: 23.03.2018).

20. AgenaRisk, [Electronic resource]:URL: <http://www.agenarisk.com/> (date of the application: 25.03.2018).

21. *Fenton N., Neil M., Marsh W., Hearty P., Marquez D., Krause P., Mishra R.* Predicting software defects in varying development lifecycles using Bayesian nets // Information and Software Technology, 2007. – Vol.1 (49). – P. 32-43.

22. *Fomenko A.O.* Metodika rascheta proyektnykh riskov na osnove primeneniya bayyesovskikh setey // Visnyk NTUU «KPI» Informatyka, upravlinnya ta obchyslyvalna tekhnika, 2011. – No 54. – P. 127-129.

23. Netica URL, [Electronic resource]: <http://www.norsys.com/> (date of the application: 7.02.2018).

24. *Kuznetsova A. A., Pon'kina Ye. V., Bednarikova Z., Bovarova M.* Issledovaniye faktorov migratsii sel'skoy molodezhi na osnove bayyesovskikh setey doveriya // Materialy Vserossiyskoy konferentsii po matematike. – Barnaul, 2016. – P. 133–137.

25. *Murphy K.* Bayes Net Toolbox for Matlab, [Electronic resource]: <http://bnt.googlecode.com/svn/trunk/docs/Talks/> (23.01.2018).

Щаяхметова А.С., PhD, e-mail: asemshayakhmetova@mail.ru

Сатымбеков М., докторант, e-mail: m_satymbekov@mail.ru