

## Реализация принятие решений при выборе хостинга с помощью байесовских сетей доверия в системе Netica

*Халиманенков Андрей Сергеевич*

*Приамурский государственный университет имени Шолом-Алейхема  
Студент*

*Баженов Руслан Иванович*

*Приамурский государственный университет имени Шолом-Алейхема  
к.п.н., доцент, зав. кафедрой информационных систем, математики и  
правовой информатики*

### Аннотация

В работе рассматривается задача по выбору оптимального варианта услуг хостинга, соответствующего целям заказчика. Исследование производится в системе создания байесовских сетей доверия Netica. Описан алгоритм создания байесовской сети доверия.

**Ключевые слова:** Байесовские сети доверия, система Netica, оптимальный выбор.

### Realization decision-making at the choice of the hosting by means of Bayesian networks of trust in the Netica

*Khalimanenkov Andrey Sergeevich*

*Sholom-Aleichem Priamursky State University  
Student*

*Bazhenov Ruslan Ivanovich*

*Sholom-Aleichem Priamursky State University  
Candidate of pedagogical sciences, associate professor, Head of the Department  
of Information Systems, Mathematics and Legal Informatics*

### Abstract

In work the task of the choice of an optimal variant of the choice of the hosting services the customer's purpose is considered. The study is the creation of Bayesian networks Netica. Describes in detail the algorithm for generating the Bayesian belief networks.

**Keywords:** Bayesian belief networks, system Netica, optimum choice.

В жизни современного человека мы часто сталкиваемся с различными ситуациями, выбираем средства и пути, которые удовлетворят нас исходом в той или иной ситуации. Хотя у нас и не всегда есть полное представление о ситуации, но чаще всего мы находим оптимальный путь для решения.

В подобных случаях нам помогает вычислительная модель под названием байесовские сети доверия, которая позволяет существенно сократить время для принятия верного решения, исключив возможность ошибки и повторения ранее проведённых этапов анализа ситуации для правильного выбора.

Эта модель позволяет задать связи между переменными и разделить вычисление их значений, чтобы остановить рассуждения на определенном вопросе, считая остальные события условно независимыми. Тем самым, используя данный подход, мы избегаем построения огромных таблиц вероятностей и в целом уменьшаем объем подсчетов.

Областями для использования байесовских сетей считаются экспертные системы, которые работают с вероятностями. Байесовские сети имеют применение в следующих областях:

- медицина и здравоохранение;
- космическая и военная;
- компьютерные технологии и программное обеспечение;
- изображение и видео.

В настоящее время байесовские сети имеют огромную популярность в сфере научно-исследовательской деятельности. Исследователь А.Л.Тулупьев в своей работе рассматривает преобразование ациклических байесовских сетей доверия в алгебраические байесовские сети. Преобразование осуществляется путем последовательного расчета тензоров совместной вероятности на основе тензоров условной вероятности, хранящихся в узлах байесовской сети доверия. Процесс завершается переходом от тензоров совместной вероятности к оценкам вероятности на идеалах конъюнктов [1]. В работе Е.Д. Бычкова, Б.К. Сагиновой, Н.Н. Нарутта рассматривается применение аппарата нечетких байесовских сетей доверия для оценки состояния качества телекоммуникационных услуг [2]. В статье А.В.Сироткин описал алгоритмы построения алгебраической байесовской сети, семантически эквивалентной многосвязной байесовской сети доверия [3]. Г.Г. Маклакова оценила качество услуг телекоммуникационной сети дистанционного обучения на основе байесовских сетей доверия [4]. Д.А. Яковлев и И.С. Синева изучили процесс детектирование веб-роботов, используя в работе байесовские сети [5]. В работе А.Н. Размахниной и Р.И. Баженова был реализован процесс принятия решений с помощью байесовских сетей доверия в системе Hugin [6]. А.Н. Терентьев, П.И. Бидюк и Л.А. Коршевнюк предложили новый, более простой и точный алгоритм вероятностного вывода в байесовских сетях на основе обучающих данных [7]. Так же в сегменте иноязычных научных исследований байесовские сети активно изучаются и применяются для решения задач. Н. Gan, Y. Zhang, Q. Song использовали байесовскую сеть доверия для положительного немеченого изучения с неопределенностью в решении [8]. Авторами T.D. Phan, J.C.R. Smart, S.J.Capon, W.L. Hadwen, O.Sahin был проведен систематический обзор в сфере управления водными ресурсами, используя байесовские сети [9]. F. Zare, H.K. Zare, M.S. Fallahnezhad реализовали

систему оценки проектов по разработке программного обеспечения на основе оптимальных Байесовских сетей [10]. Построили байесовскую сеть для исследования реакции хозяйственных мероприятий на промысел F.M. Underwood, G. Parkes, J. H. Swasey [11]. L. Mkrtchyan, L. Podofillini, V.N. Dang продемонстрировали методы построения таблиц условных вероятностей в байесовской сети доверия [12].

Целью данного исследования является реализация принятия решений при выборе компьютера с помощью байесовских сетей доверия в системе Netica.

Перед началом выполнения поставленной задачи рассмотрим программу Netica. Netica — достаточно удобная в работе программа для работы с графовыми вероятностными моделями. Она имеет интуитивно понятный интерфейс пользователя для ввода данных. Для выполнения многих операций достаточно воспользоваться несколькими щелчками мыши, что облегчает работу пользователя в системе. Netica является продуктом компании Norsys Software Corp. Программа имеет коммерческую и бесплатную версию. Для использования бесплатной версии, необходимо скачать приложение на официальном сайте <http://www.norsys.com/>, оставить диалоговое окно для ввода пароля пустым и нажать 'Limited Mode'. Бесплатная демо-версия является полнофункциональной, но ограничивает созданную модель по размеру.

Для реализации поставленной в данном исследовании цели рассмотрим задачу.

Покупатели, зайдя на большое количество интернет-страниц, увидели огромный выбор по предоставлению услуг хостинга. На выбор были предложены различные сервисы, отличающиеся по количеству доступных сайтов и доменов, баз данных и систем управления ими; объему оперативной и постоянной памяти; наличием таких параметров как CMS, сертификата безопасности SSL; поддержкой различных языков программирования, и соответственно цене. Целью задачи является определение вероятностей приобретения услуги того или иного хостинга для каждого покупателя, преследующего различные цели.

Рассмотрим интерфейс программы Netica:

На рисунке 1 указано расположение кнопок, выполняемые ими функции указаны в таблице.

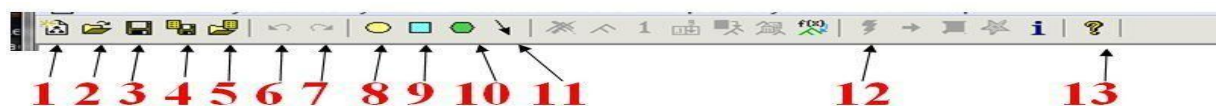


Рис. 1. Панель программы Netica

Таблица 1. Функции кнопок Netica

№	Наименование	Функция
1	New Net	Создать новую сеть
2	Open	Открыть файл

3	Save	Сохранить
4	Save Case	Сохранить как ...
5	Get Case From File	Добавить таблицу данных
6	Undo	Предыдущее действие
7	Redo	Сделать вновь отмененное действие
8	Add Nature Node	Добавить узел
9	Add Decision Node	Добавить узел с таблицей вероятностей
10	Add Utility Node	Добавить вспомогательный узел
11	Add Link	Добавить связь
12	Compile	Компилировать сеть
13	Help Topics	Помощь

Для работы в программе Netica в самом начале необходима таблица входных данных. Её создание возможно в MS Excel. Важно учесть, что идентификатор переменной не может начинаться с цифры и содержать символы такие как: /, ,,.

Таблица заполняется данными для будущей модели системы. Для получения данных провели выборку из лидирующих представителей хостинг-услуг и их предложений. Были выписаны цены и предлагаемые возможности. В таблице столбцы первой строки будут служить названиями для узлов в программе, строки 2-7 каждого столбца являются вариантами вероятности выбора того или иного узла.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Source	Number of sites	Number of DB	Price	Prog. lang.	E-mail accounts	Memory, Gb	domain	SSL	RAM	CMS	DBMS
2	<a href="http://www.fozzy.com">www.fozzy.com</a>	1	1	129	3	1000	10		0 yes	512	4	2
3	<a href="https://sherlockhost.ru/">https://sherlockhost.ru/</a>	15	1000	199	3	1000	1000		1 yes	1536	1	1
4	<a href="https://timeweb.com/ru/services/hosting/">https://timeweb.com/ru/services/hosting/</a>	1	1	160	6	1000	5		0 yes	1024	3	2
5	<a href="https://www.reg.ru/hosting/">https://www.reg.ru/hosting/</a>	7	1000	129	3	1000	12		4 yes	1024	3	1
6	<a href="https://www.nic.ru/catalog/hosting/shared/">https://www.nic.ru/catalog/hosting/shared/</a>	10	0	129	0	10	4		10 no	64	0	0
7	<a href="https://www.hostland.ru/hosting">https://www.hostland.ru/hosting</a>	3	1000	109	6	1000	5		0 yes	512	0	2
8												

Рис. 2. Результат выборки хостингов

Заполнив таблицу результатами опроса, переходим непосредственно в программу Netica. Создаем новую сеть: File - New - Network

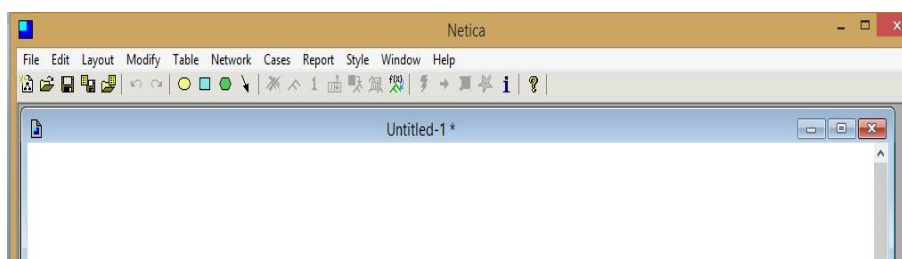


Рис. 3. Создание новой сети

Далее необходимо обучить сеть согласно данным, созданных ранее MS Excel. Для этого используем команду Cases - Learn - Add Case File Nodes,

в появившемся окне находим наш файл MS Excel и нажимаем "Открыть". Результат данной операции показан на рис. 4.

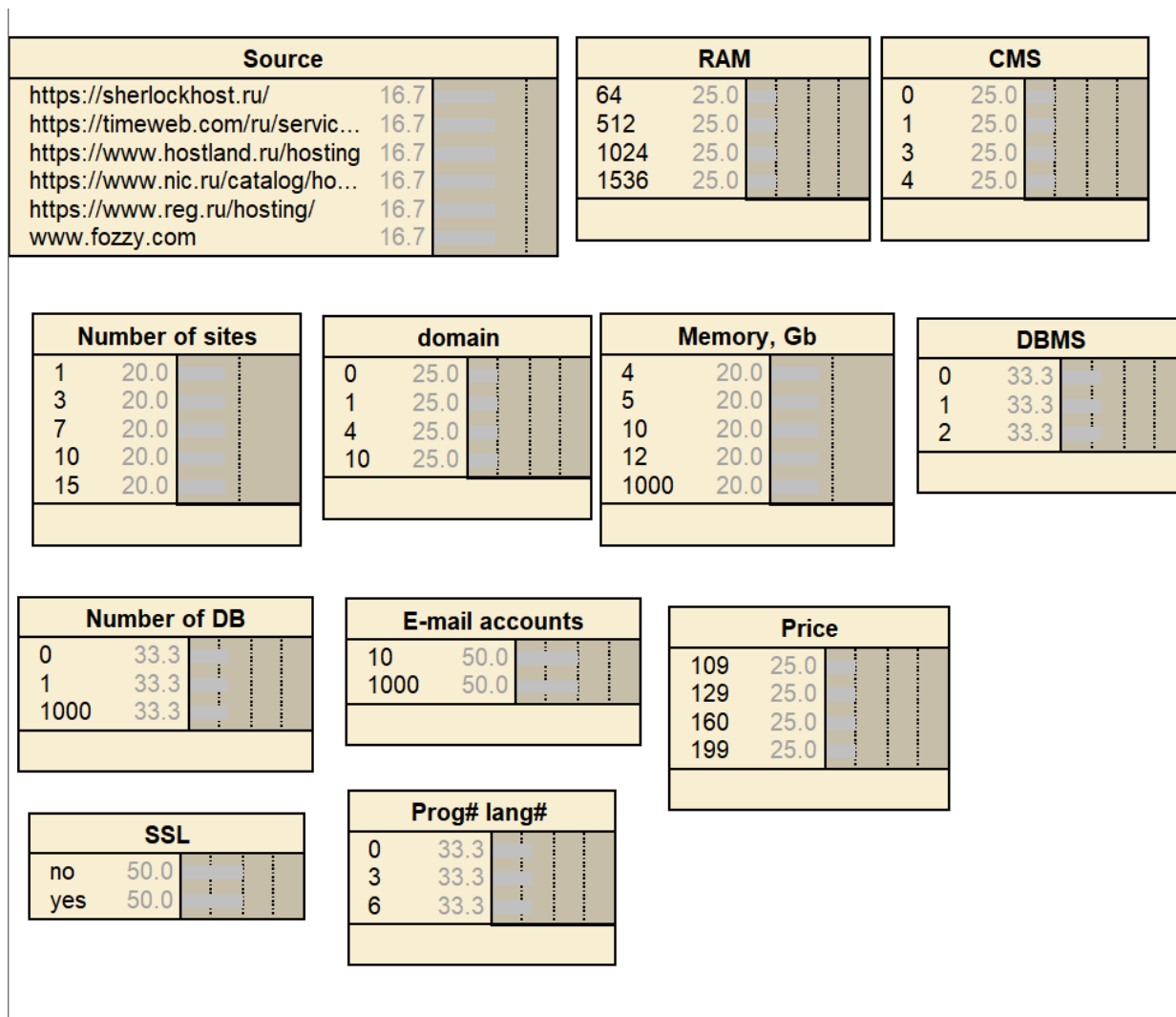


Рис. 4. Результат обучения сети

После обучения сети необходимо задать связи полученным узлам сети. Для этого выбираем целевой узел (В нашей случае целевым узлом является Price), и щелчком левой кнопки мыши делаем его активным. Для получения связей обучаем сеть с помощью Tree Augmented Naive Bayes: используем кнопки Cases - Learn - Learn TAN Structure, в появившемся окне выбираем, используемый ранее файл MS Excel и жмем "открыть". Результат видим на рис. 5.

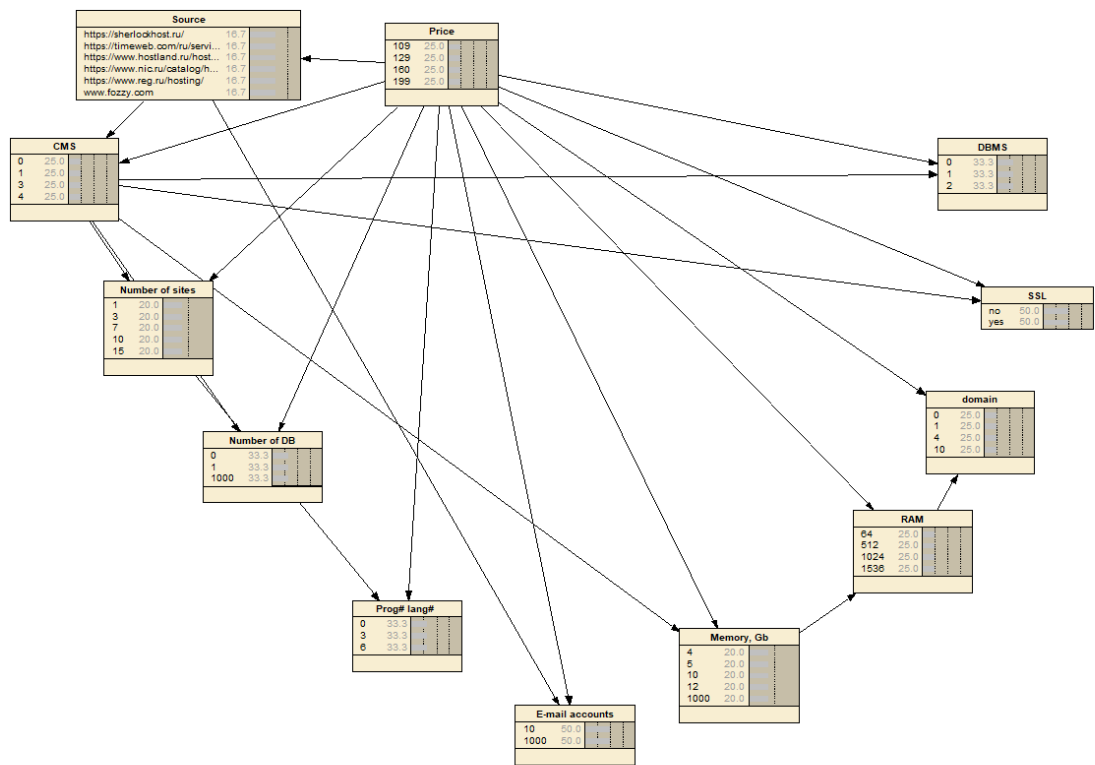


Рис. 5. Добавление связей между узлами с помощью Tree Augmented Naive Bayes

Рассматривая построенные связи, можно заметить, что некоторые из них являются противоречивыми. К примеру, Number of DB (количество баз данных) не может зависеть от количества сайтов (Number of sites), так же как и E-mail accounts (электронные почтовые ящики), не могут зависеть от Price (цены). Далее проверим все связи с точки зрения логического рассуждения и выбрав лишние, удалим стрелки с помощью клавиши "delete" на клавиатуре. Так же расставим узлы в более оптимальном порядке. Результат видим на рис. 6.

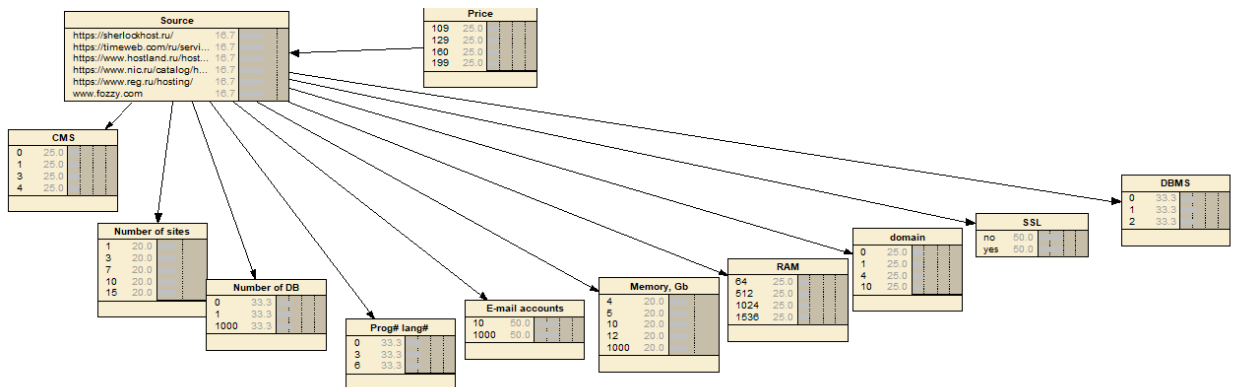


Рис. 6. Удаление противоречивых связей и оптимальный порядок узлов

Далее для получения результатов необходимо выполнить обучение полученной сети, используя EM-алгоритм.

Воспользуемся клавишей в панели быстрого доступа Cases - Learn - Learn Using EM, выбираем файл MS Excel и жмем "открыть". Результат на рис. 7.

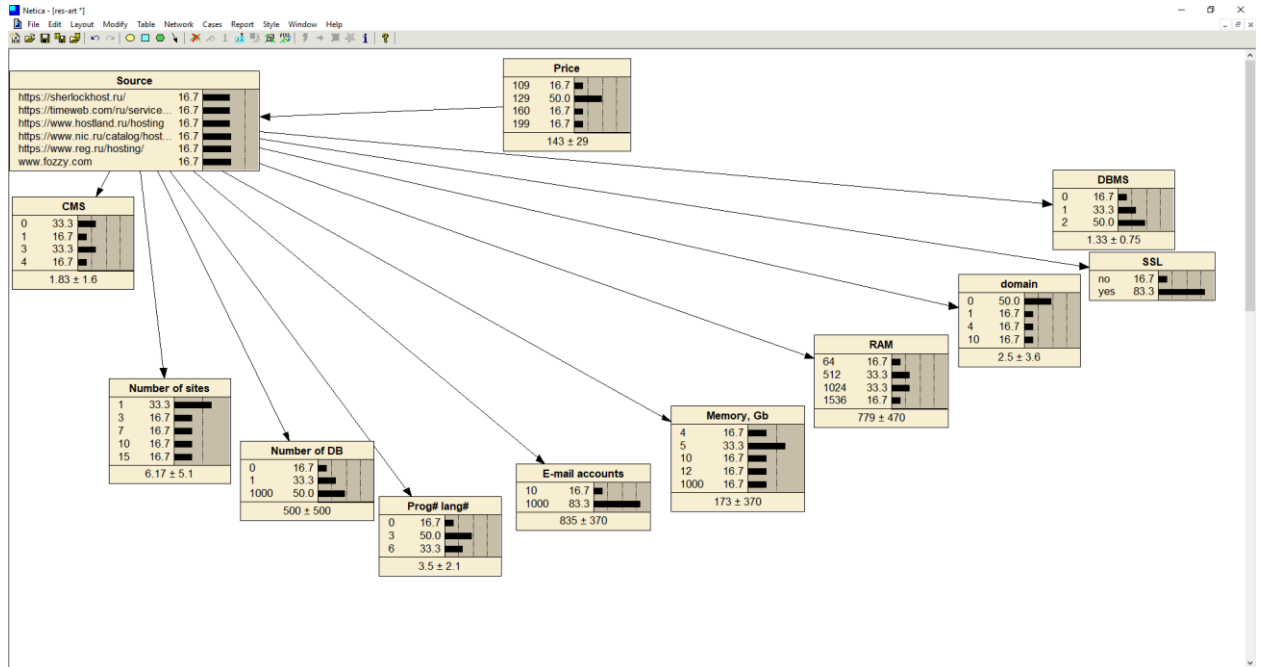


Рис. 7. Результат EMалгоритма

Далее применим полученную сеть для поставленной нами задачи. Для начала определим различные варианты поведения сети, в процессе изменения различных факторов.

При выборе хостинга за 129 руб./мес. и наличия SSL-сертификата, система говорит, что с вероятностью 50% будет выбран сервис www.reg.ru/hosting/ и с той же вероятностью будет выбран www.fozzy.com. Результат на рис. 8.

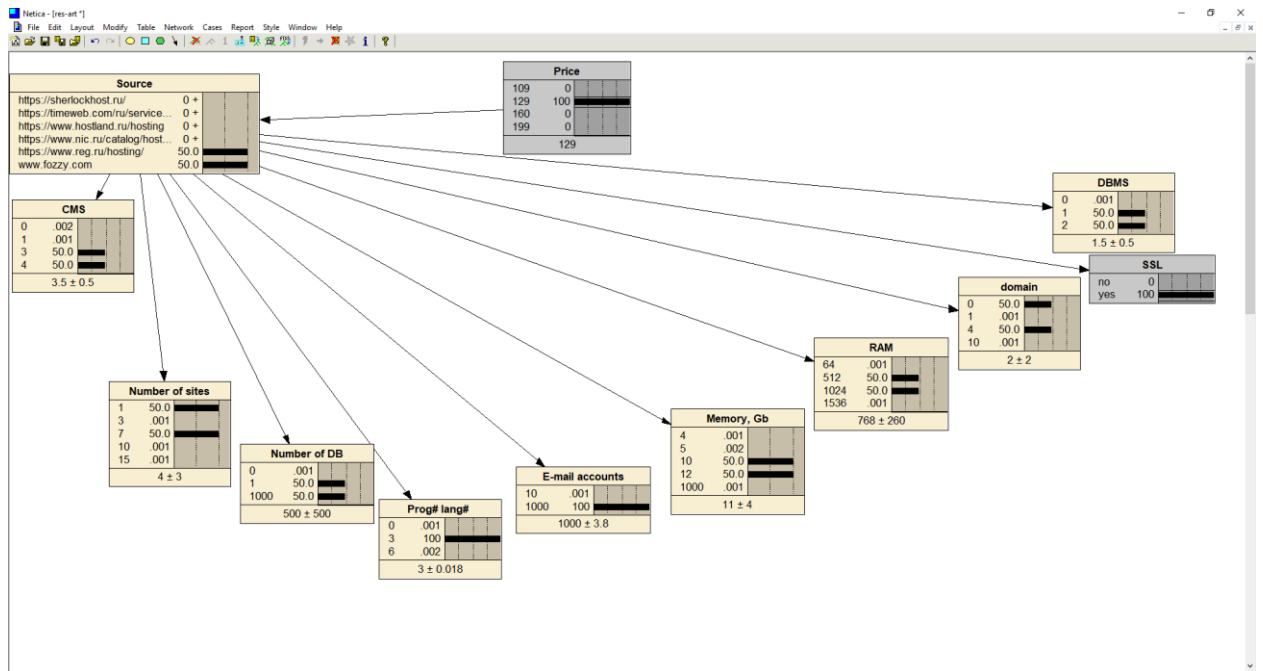


Рис. 8. Результат выбора по предпочтениям

Добавим предпочтения по таким факторам как: объём жесткого диска (12 Гб), оперативной памяти (1 Гб), и количество доступных баз данных (1000). Результат на рис. 9.

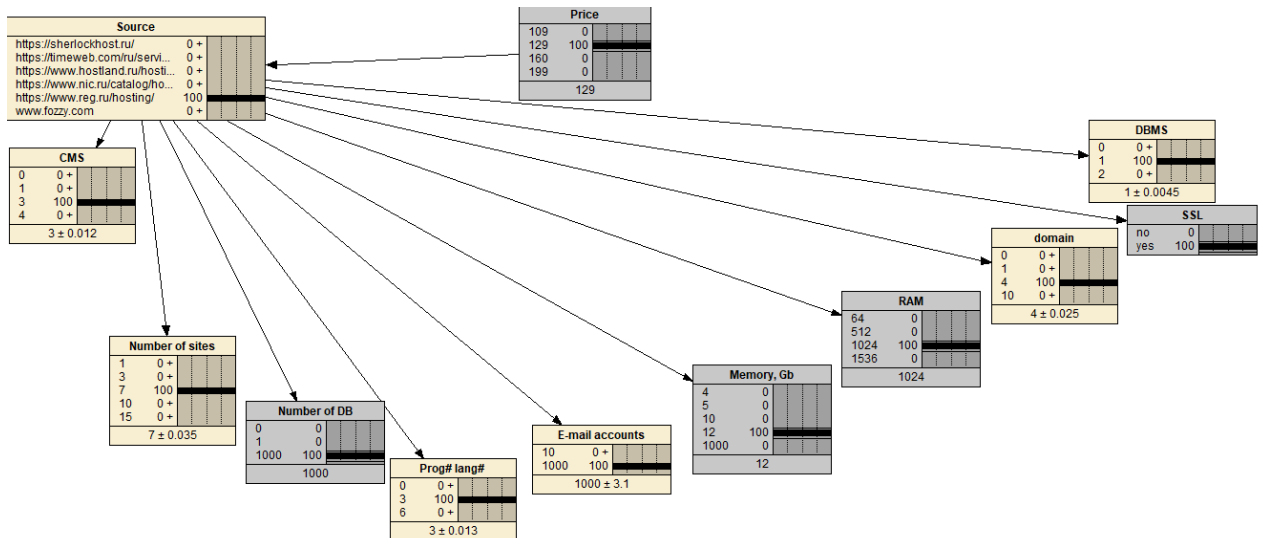


Рис. 9. Предпочтения по цене, объёму памяти и сертификату безопасности

Предположим ситуация, когда заказчику нужно создать сервис по рассылке электронных писем. В данном случае, важно большое количество E-mail аккаунтов и наличия сертификата-безопасности, для снижения рисков перехвата конфиденциальных сообщений и данных. Результат на рис. 10.



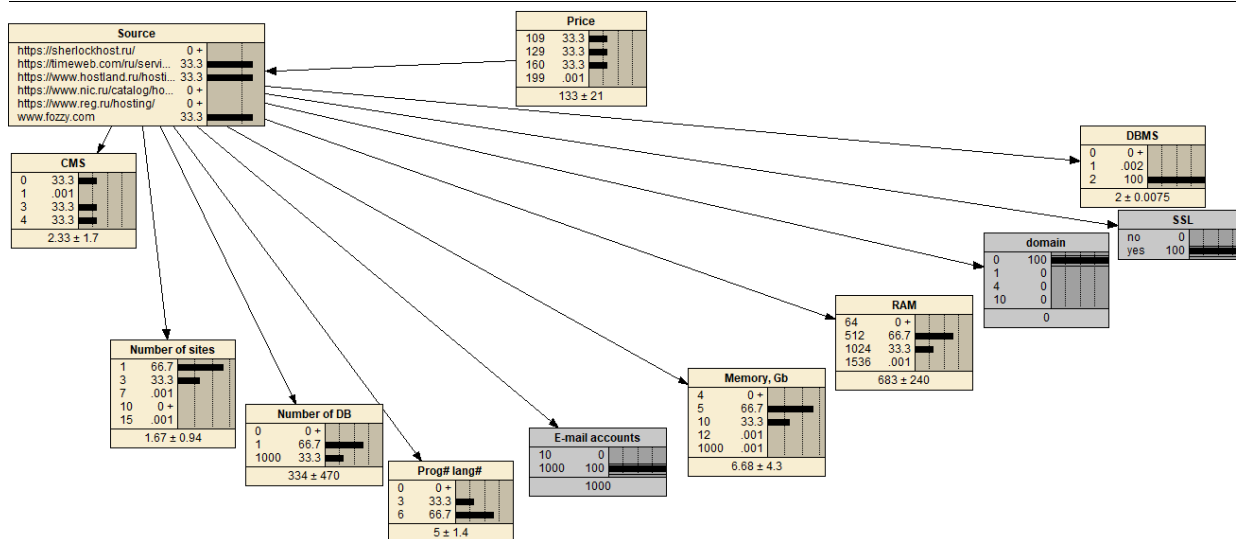


Рис. 10. Выбор варианта для почтовой службы с огромным количеством почтовых ящиков

Для конечного выбора следует выбрать больший объём памяти, т.к. для рассылки сообщений с вложенными файлами нужен достаточный размер жёсткого диска сервера. Результат на рис. 11.

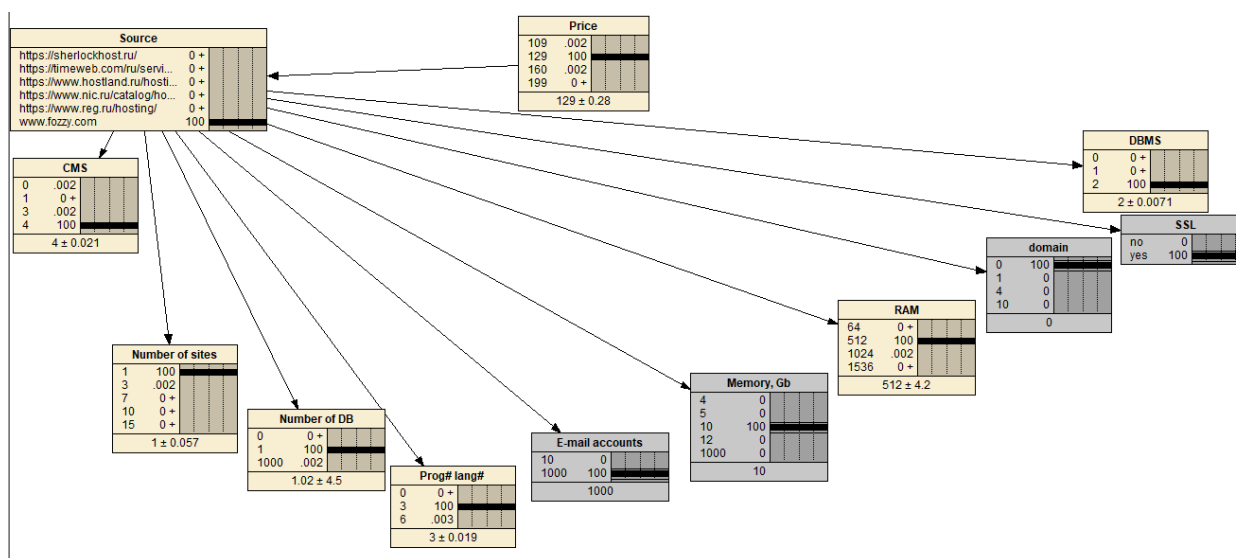


Рис. 11. Выбор наилучшего хостинга для почтовых рассылок

Данная сеть позволяет определить вероятность выбора того или иного сервиса по хостингу любому заказчику, достаточно выбрать приоритетные характеристики.

Таким образом, в ходе исследования была реализована полностью работающая сеть, позволяющая осуществить подбор хостинга для каждого покупателя, опираясь на его предпочтения. Мы показали, что байесовские сети - это модель способная помочь в поиске сложных решений и определении вероятности определенных событий. Поставленные в работе цели были осуществлены.

**Библиографический список**

1. Тулупьев А.Л. Преобразование ациклических байесовских сетей доверия в алгебраические байесовские сети // Известия высших учебных заведений. Приборостроение. 2009. Т. 52. № 3. С. 21-23.
2. Бычков Е.Д., Сагинова Б.К., Нарутта Н.Н. Нечеткая байесовская сеть в модели качества обслуживания услуг сети связи. // Россия молодая: передовые технологии – в промышленность! 2013. № 1. С. 202-205 .
3. Сироткин А.В. Алгоритмы построение алгебраической байесовской сети, семантически эквивалентной многосвязной байесовской сети доверия // Международная конференция по мягким вычислениям и измерениям. 2009. Т. 1. С. 131-134.
4. Маклакова Г.Г. Система оценки качества услуг телекоммуникационной сети дистанционного обучения на основе байесовских сетей доверия // Новые компьютерные технологии. 2008. Т. 6. № 1(6). С. 72-73.
5. Яковлев Д.А., Синева И.С. Детектирование веб-роботов с помощью байесовских сетей. // Технологии информационного общества. 2017. С. 292-293.
6. Размахнина А.Н., Баженов Р.И. Исследование процесса принятия решений с помощью байесовских сетей доверия в системе Hugin // Постулат. 2017. № 8. С. 1-8.
7. Терентьев А.Н., Бидюк П.И., Коршевнюк Л.А. Алгоритм вероятностного вывода в байесовских сетях // Системні дослідження та інформаційні технології. 2009. № 2. С. 107-111.
8. Gan H., Zhang Y., Song Q. Bayesian belief network for positive unlabeled learning with uncertainty // Pattern Recognition Letters. 2017. Т. 90. С. 28-35.
9. Phan T. D., Smart J. C.R., Capon S. J., Hadwen W. L., O. Sahin Applications of Bayesian belief networks in water resource management: A systematic review // Environmental Modelling & Software. 2016. Т. 85. С. 98-111.
10. Zare F., Zare H. K., Fallahnezhad M.S. Software effort estimation based on the optimal Bayesian belief network // Applied Soft Computing. 2016. Т. 49. С. 968-980.
11. Underwood F.M., Parkes G., Swasey J. H. Building Bayesian Belief Networks to investigate how fishery performance responds to management interventions // Fisheries Research. 2016. Т. 182. С. 28-38.
12. Mkrtchyan L., Podofillini L., Dang V.N. Methods for building Conditional Probability Tables of Bayesian Belief Networks from limited judgment: An evaluation for Human Reliability Application // Reliability Engineering & System Safety. 2016. Т. 151. С. 93-112.