

Реализация принятия решения при выборе автомобиля для города с помощью байесовских сетей доверия в системе Netica

Бокач Никита Александрович

*Приамурский государственный университет имени Шолом-Алейхема
Студент*

Научный руководитель:

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 автомобиля для езды по городу, учитывая параметры заказчика. Исследование проводится в системе создания байесовских сетей доверия Netica. Описан алгоритм создания байесовской сети доверия.

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

Implementation of decision-making when choosing a car for a city using Bayesian trust networks in the Netica system

Bokach Nikita Aleksandrovich

*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

The paper considers the problem of choosing a car for driving around the city, taking into account the parameters of the customer. The research is carried out in the system for creating Bayesian trust networks Netica. An algorithm for creating a Bayesian trust network is described.

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

Сегодня люди часто задаются вопросом, какой выбрать себе автомобиль, чтобы он был комфортным, и имел необходимые для пользователя параметры. 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.

Данная модель дает возможность определить связи между переменными и вычислить их вероятностные значения, учитывая, что события условно независимы. Что дает нам возможность выстроить сеть без построения больших таблиц вероятностей, вследствие чего уменьшается количество и объем вычислений.

Психологические эксперименты показали, что люди часто неверно оценивают вероятность события, на основе полученного опыта (апостериорная вероятность), поскольку игнорируют саму вероятность предположения (априорная вероятность). Поэтому правильный результат по формуле Байеса может сильно отличаться от интуитивно ожидаемого. Теорема Байеса названа в честь её автора Томаса Байеса (1702–1761) – английского математика и священника, который первым предложил использование теоремы для корректировки убеждений, основываясь на 11 обновлённых данных. Его работа «An Essay towards solving a Problem in the Doctrine of Chances» впервые опубликована в 1763 году, через 2 года после смерти автора. До того, как посмертная работа Байеса была принята и прочитана в Королевском обществе, она была значительно отредактирована и обновлена Ричардом Прайсом. Однако эти идеи не предавались публичной огласке до тех пор, пока не были вновь открыты и развиты Лапласом, впервые опубликовавшим современную формулировку теоремы в своей книге 1812 года «Аналитическая теория вероятностей». В статье А.В.Сироткин описал алгоритмы построение 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. Коршевнок 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].

Основной целью данного исследования является реализация принятия решений при выборе автомобиля с помощью байесовских сетей доверия в системе Netica.

Для начала рассмотрим интерфейс выбранной программы. 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 решения поставленной задачи необходима таблица, содержащая основные параметры автомобилей, это будут наши входные данные. Netica предоставляет возможность добавления входных данных, посредством загрузки файла Microsoft Excel. Однако важно учесть, что заголовок столбца не должен начинаться с символов: /, ,, .

Рассмотрим пример задачи. Для выбора автомобиля, клиенту предоставляется огромный выбор технической составляющей каждой модели автомобиля. Такие как: Объем двигателя, мощность двигателя (л.с), коробка переключения передач, расход топлива. Цель задачи определить вероятность покупки той или иной модели автомобиля, учитывая заявленные требования клиента.

Рассмотрим интерфейс программы Netica:

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

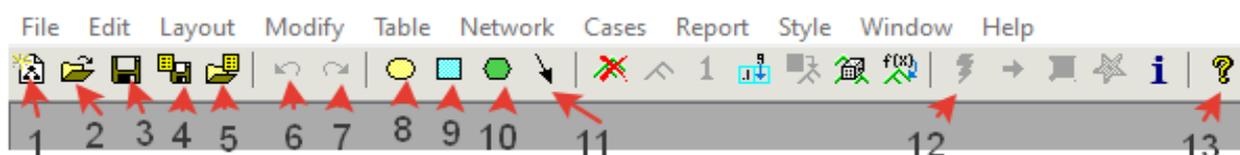


Рис. 1. Панель программы Netica

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

Составим таблицу с входными данными для программы Netica. Таблица заполняется данными для будущей модели системы. Для получения данные были посещены интернет площадки по продаже и покупке автомобилей. Выбраны варианты как с автоматической так и с механической коробкой переключения передач. В таблице первые строки служат для названия узлов в программе, строки 2-15 являются вариантами вероятности выбора того или иного узла.

	A	B	C	D	E	
1	Marka-model-god	Obem dvigatelya	moshnost	cena	rashod-na-100km	kpp
2	Nissan Sunny-2001	1,5	105	195000		6,6 auto
3	Toyota Platz-2000	1,3	88	170000		6 mechanica
4	Nissan Primera-2001	2,00	150	240000		7,7 auto
5	Renault Logan-2013	1,4	75	200000		6,9 mechanica
6	Ford Focus-2006	2,00	145	238000		7,5 mechanica
7	Nissan Cedric-2000	2,00	125	250000		7 auto
8	Toyota Corolla Fielder-2012	1,50	103	685000		7 auto
9	Toyota Premio-2002	1,50	109	455000		6,5 auto
10	Nissan X-Trail-2008	2,00	141	740000		9 auto
11	Nissan R'nessa-1998	2,00	140	300000		12 auto
12	Volkswagen Passat-1990	1,80	112	80000		11 mechanica
13	Mazda Mazda6-2009	1,80	120	650000		11 mechanica
14	Ford Focus-2007	1,80	125	269000		8 mechanica
15	Nissan Primera-2006	1,80	109	350000		9 mechanica

Рис. 2. Результат выборки

После заполнения таблицы данными, переходим в программу Netica. Создаем новую сеть: File – New – Network

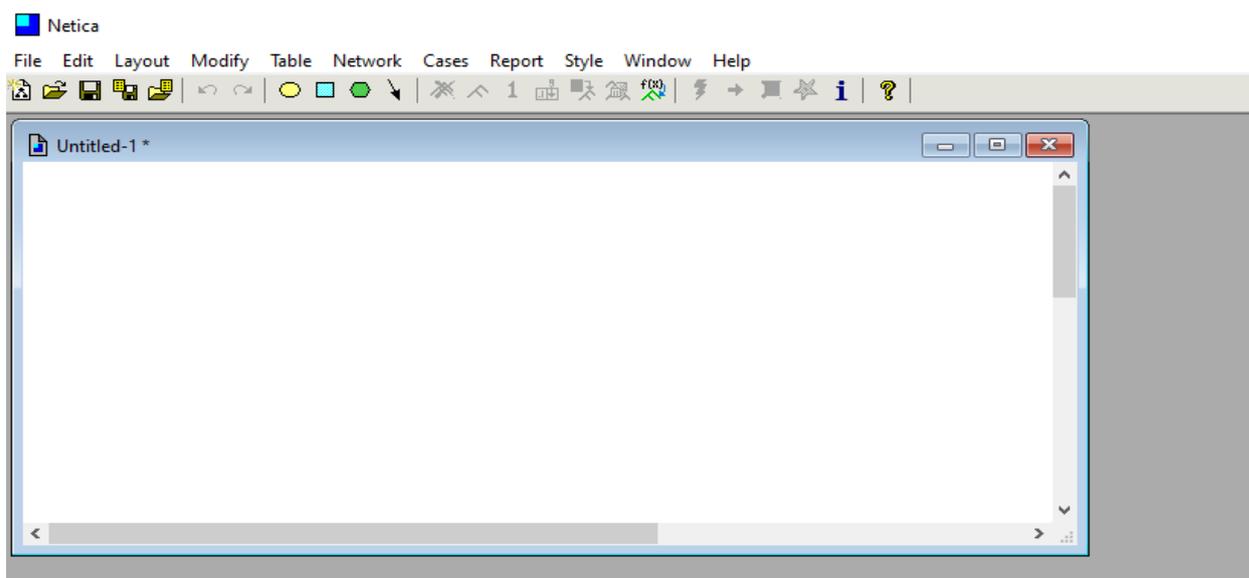


Рис. 3. Создание новой сети

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.

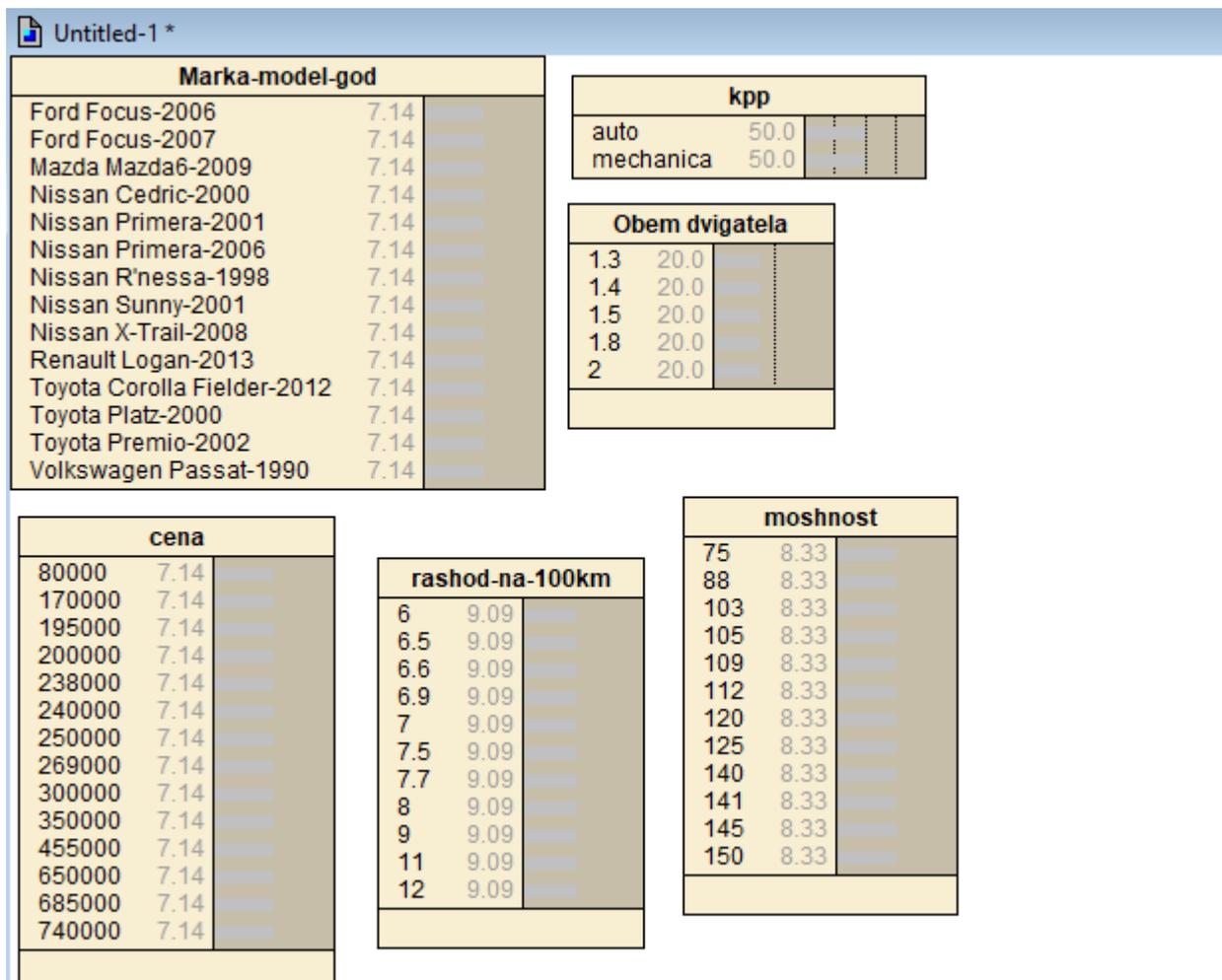


Рис. 4. Результат обучения сети

Чтобы определить систематизацию данных необходимо задать связи между узлами. Для этого выполним операцию Tree Augmented Naive Bayes. Но перед этим необходимо выбрать целевой узел. Мы выбрали Pashod-na-100km. Так как автомобиль на данный момент должен быть экономичный. После чего выполнили команду Cases - Learn - Learn TAN Structure, в диалоговом окне была выбрана наша таблица Excel. Результат можно увидеть на Рис.5.

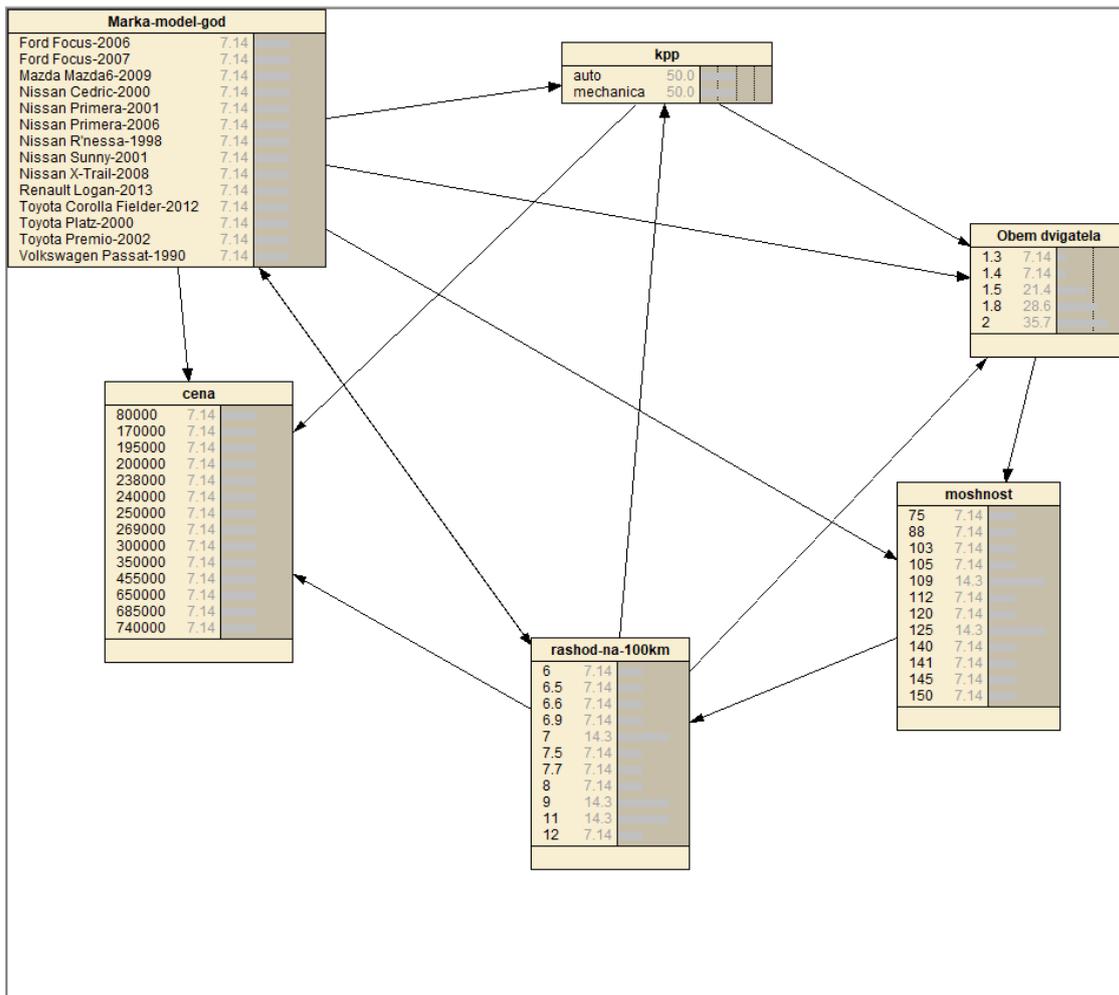


Рис.5 Добавление связей с помощью Tree Augmented Naive Bayes

Рассматривая построенные связи можно заметить, что они не совсем точны и не совсем корректно описывают нашу систему. Например, не совсем корректно будет утверждение, что коробка переключения передач (kpp) зависит от расхода топлива (rashod-na-100km). Определим связи с точки зрения логических рассуждений, удалив некорректные и добавив верные суждения. Результат можно увидеть на рис.6.

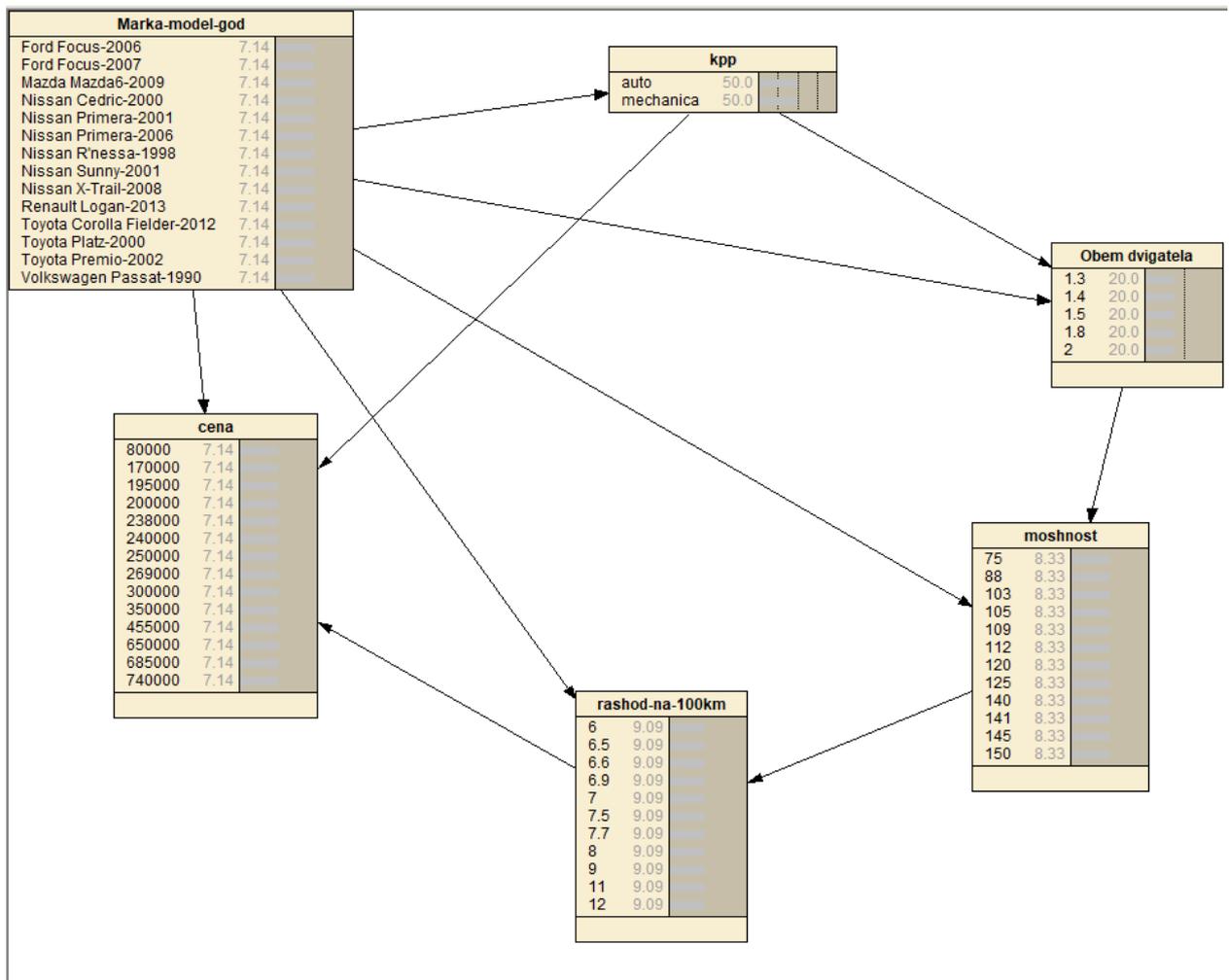


Рис. 6. Удаление противоречивых связей и оптимальный порядок узлов

Далее для получения результатов необходимо выполнить обучение, полученное сети, используя EM-алгоритм. Воспользуемся клавишей в панели быстрого доступа Cases - Learn - Learn Using EM, выбираем файл MC Excel и ждем "открыть". Результат на рис. 7.

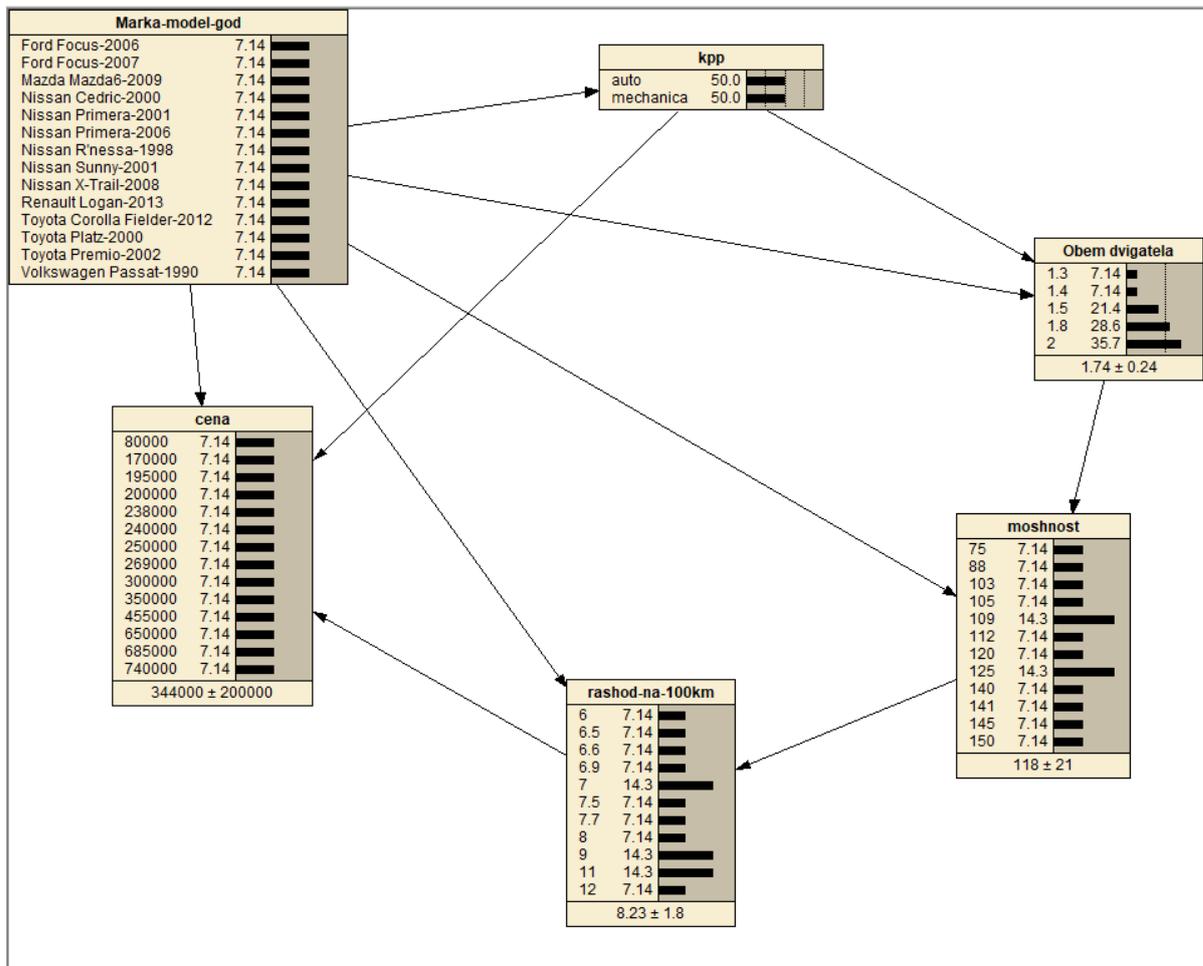


Рис. 7. Результат EM-алгоритма

Применим полученную сеть для поставленной задачи. Рассмотрим различные варианты поведения сети, в процесса изменения различных факторов.

Так при выборе объема двигателя в 2 литра и автоматической коробкой переключения передач, система предоставляет нам несколько моделей машин.

Результат на рис. 8.

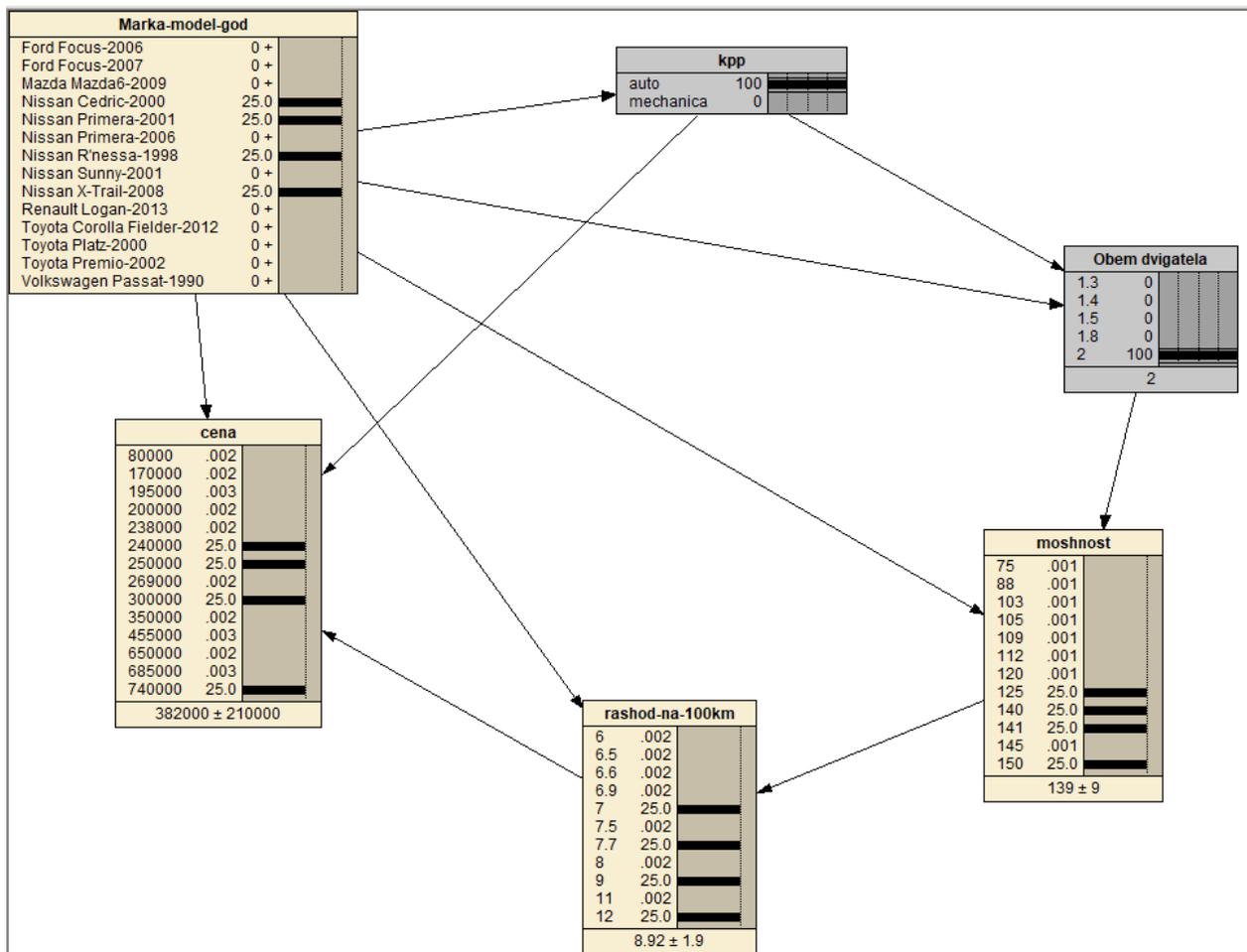


Рис. 8. Результат выбора по предпочтениям

Добавим предпочтения по таким параметрам как: Расход топлива (reshod-na-100km). Таким образом система показывает, что с вероятностью 100% клиенту следует выбрать следующую модель авто: Nissan Cedric 2000г. Результат на рис. 9.

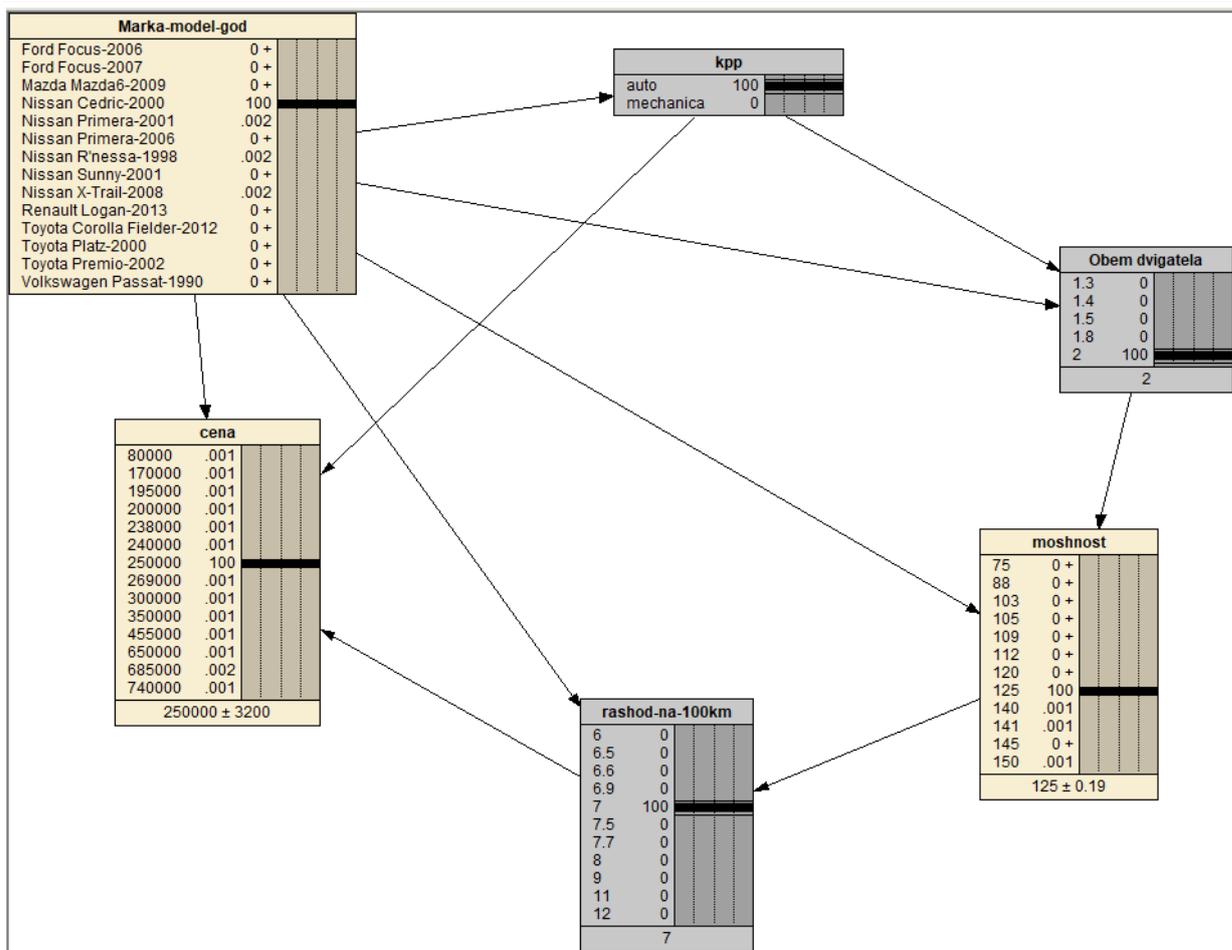


Рис. 9. Предпочтения по коробке переключения передач, объему двигателя и расходу топлива

Предположим, что клиент хочет выбрать максимально бюджетный автомобиль с наименьшим расходом топлива. Для этого он выбирает самое меньшее значение расхода топлива (reshod-na-100km). Система выдает результат, что подходящей моделью авто является: Toyota platz 2000г с двигателем объемом 1.3 литра и расходом бл на 100км. Результат выборки можно увидеть на рис. 10.

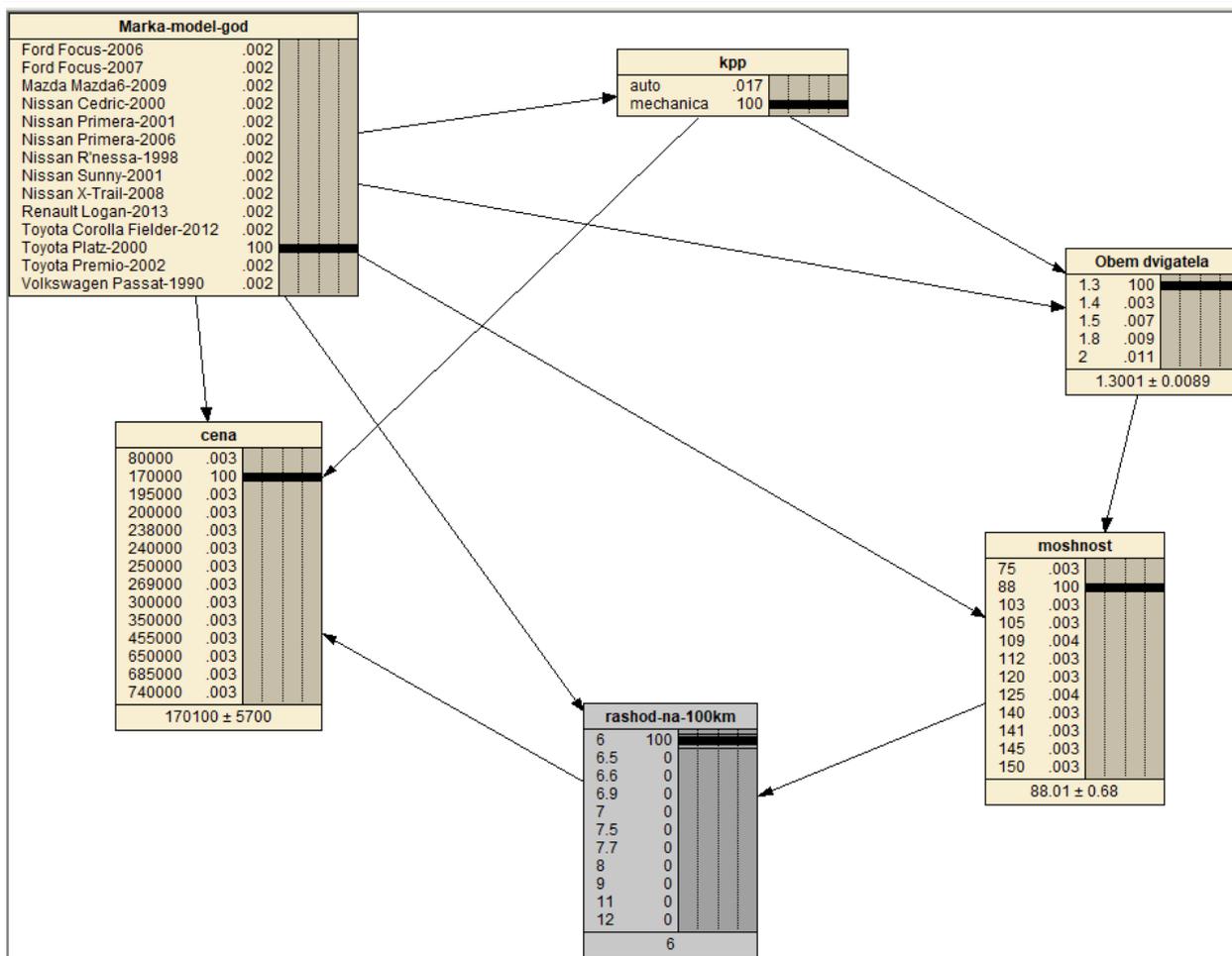


Рис.10 Выбор предпочтений клиента по наименьшему расходу топлива.

Данная система позволяет определить вероятность покупки определенной модели авто из существующего списка по определенным параметрам для клиента.

Таким образом, в ходе нашего исследования была разработана байесовская система доверия, которая помогает упростить выбор покупателям в подборе автомобиля, путем отбора характеристик по запросу клиента. Так мы доказали что байесовская сеть может решить широкий спектр различных проблем в поиске решений.

Поставленная цель была достигнута.

Библиографический список

1. Мальчевская Е.А., Бирилло А.И., Харитонов Н.А. Развитие матрично-векторного подхода в алгоритмах локального априорного вывода в алгебраических байесовских сетях // Труды СПИИРАН . 2017.
2. Звягин Л.С. Применение байесовского подхода в измерениях аналитических данных как фактор формирования процессов системного экономического развития // Молодой ученый. 2017.
3. Баркалов С.А., Азаранова Т.В., Полухин П.В. Управление процессом тестирования Веб-приложений методом фаззинга на основе динамических

- байесовских сетей // Вестник Южно-Уральского государственного университета. 2017.
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. Тулупьев А.Л. Преобразование ациклических байесовских сетей доверия в алгебраические байесовские сети // Известия высших учебных заведений. Приборостроение. 2009. Т. 52. № 3. С. 21-23.
 8. Мотиенко А.И., Макеева С.М. Басова О.О Проактивное управление робототехническими системами спасения пострадавших // Труды СПИИРАН. 2016. №3.