Федеральное государственное автономное
образовательное учреждение
высшего образования
«СИБИРСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Институт космических и информационных технологий
институт

Кафедра информатики
кафедра

УТВЕРЖДАЮ
Заведующий кафедрой
_______ А.И. Рубан
подпись инициалы, фамилия
«16» июня 2016 г.

БАКАЛАВРСКАЯ РАБОТА

27.03.03 « Системный анализ и управление»
код – наименование направления

Алгоритм пропагации свидетельств в байесовской
тема
сети с циклами

Руководитель _______ доцент к.т.н. Даничев А.А.
подпись, дата должность, ученая степень инициалы, фамилия

Выпускник _______ Зыков А.В.
подпись, дата инициалы, фамилия

Красноярск 2016

СОДЕРЖАНИЕ

ВВЕДЕНИЕ ... 3
1 Общие сведения о теории байесовских сетей ... 3
1.1 Байесовские сети ... 4
1.1.1 Условная вероятность ... 4
1.1.2 Байесовская сеть ... 4
1.1.3 Байесовская сеть доверия ... 5
1.2 Формула Байеса и формула полной вероятности ... 6
1.3 Распространение свидетельств в байесовской сети без циклов ... 6
1.4 Циклы в Байесовских сетях доверия ... 7
1.5 Алгоритм пропагации свидетельств в байесовской сети с циклами ... 13
1.6 Пример построения простейшей байесовской сети доверия ... 15
2 Разработанные алгоритмы пропагации свидетельств ... 18
2.1 Измененный алгоритм пропагации ... 18
2.2 Общие формулы для распространения свидетельств ... 23
2.3 Применения формул для распространения свидетельств на конкретном примере ... 25
2.4 Оценка сложности алгоритма ... 27
3 Апробации алгоритмов и их сравнение ... 28
3.1 Апробация алгоритмов на сети без циклов ... 28
3.2 Алгоритм пропагации в байесовских сетях с циклами ... 31
3.2.1 Подсчет вероятностей по алгоритму Николенко-Тулупьева-Сироткина ... 32
3.2.2 Подсчет вероятностей по измененному алгоритму ... 35
3.2.3 Сравнение результатов работы алгоритма для байесовской сети с циклами ... 40
ЗАКЛЮЧЕНИЕ ... 41
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ ... 42

ВВЕДЕНИЕ

Байесовские сети широко применяются в таких областях, как медицина, стратегическое планирование, финансы и экономика, а так же являются особенно полезными при разработке и анализе машинных алгоритмов обучения. Не все алгоритмы распространения свидетельств в байесовских сетях работают для сетей с циклами [1,2,3], поэтому разработка таких алгоритмов вполне актуальна.

Объектом исследования являются Байесовские сети.

Предметом исследования являются алгоритмы пропагации в байесовских сетях с циклами.

Целью данной научной работы является разработка алгоритма распространения свидетельств в байесовской сети с циклами.

Задачи исследования:

- Изучить научно-методическую литература по проблеме исследования;
- Проанализировать имеющиеся алгоритмы распространения свидетельств в байесовских сетях;
- Разработать собственный алгоритм распространения свидетельств в байесовских сетях;
- Сравнить разработанный и уже имеющиеся алгоритмы.

1 Общие сведения о теории байесовских сетей

1.1 Байесовские сети

1.1.1 Условная вероятность

Определение условной вероятности. Представим себе, что мы сидим здесь, в этой комнате, а в это время кто-то за стенкой проводит случайный эксперимент. В этом эксперименте нас интересует событие A. Если мы ничего не знаем о том, как проходит этот эксперимент, то шансы осуществиться событию A выражаются для нас числом $P(A)$. Если же кто-то по секрету сообщил нам, что в эксперименте произошло событие B, то эти шансы вполне могут изменится. Например, если $B \subset A$, то эти шансы возрастут с $P(A)$ до 1, а если $A \cap B = \emptyset$, то упадут с $P(A)$ до 0. Все это приводит к следующему определению [4,5,6,7,8].

Пусть $(\cap, F, P)$— вероятностное пространство, $A, B \in F$ и $P(B) > 0$. Тогда условной вероятностью события A при условии B называется число:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{1.1}$$

Смысл этого определения следующий. Если нам известно, что произошло событие B, то пространство элементарных событий «сужается» до B. В то же время при этом условии событие A, может произойти только вместе с B. Отсюда и получается формула (1.1).

1.1.2 Байесовская сеть

Байесовские сети представляют собой графические модели событий и процессов на основе объединения некоторых результатов теории вероятностей и теории графов [1,2,3]. Несмотря на свое название, эти сети не обязательно подразумевают тесную связь с байесовскими методами. Название связано, прежде всего, с байесовским правилом вероятного вывода. Байесовские сети представляют собой удобный инструмент для описания достаточно сложных процессов и событий с неопределенностями. Они оказались особенно полезными при разработке и анализе машинных алгоритмов обучения. Основной идеей построения графической модели есть понятие модульности, то есть разложение сложной системы на простые элементы. Для объединения отдельных элементов в систему используются результаты теории вероятностей, которые обеспечивают модели состоятельность в целом. Такой граф-теоретический подход к построению графической модели обеспечивает исследователю возможность строить модели процессов с множеством сильно взаимодействующих переменных, а также создавать структуры данных для последующих разработки эффективных алгоритмов их обработки и принятия решений.

1.1.3 Байесовская сеть доверия

Графические модели представляются собой графы, узлы которых соответствуют случайным переменным. Если узлы (переменные) не соединены дугами, то их считают условно независимыми. Ненаправленные графические модели называют также Марковскими случайными полями (МСП). Для МСП независимость можно сформулировать следующим образом: два множества (узлов) A и B являются условно независимыми при наличии в модели третьего множества C, если все пути между узлами множеств A и B разделены узлами множества C.

Байесовская сеть (БС) доверия – это направленный ациклический граф. Байесовская сеть – это пара, в которой первый компонент G, является направленным нециклическим графом, соответствующий случайным переменным. С байесовскими сетями связано более сложное понятие независимости, которое учитывает направленность дуг. Граф записан как набор условий независимости: каждая переменная независима от ее родителей G. Вторая компонента пары, представляет собой набор па- раметров, который определяет сеть. Он содержит параметр $\theta_{x_i|p_a(X_i)} = P(x_i|p_a(X_i))$ для каждого возможного значения x_i из X_i и $p_a(X_i)$ из $P_a(X_i)$. $P_a(X_i)$ обозначает набор родителей X_i в G и $p_a(X_i)$ - родителей. Если смотреть больше чем один граф, тогда мы используем ${P_a}^G(X_i)$. чтобы определить X_i родителей в графе G. Байесовская сеть B определяет распределение вероятности $D = \{x^1, \ldots, x^N\}$ по X, $P_B(X_1, \ldots, X_n) = \prod_{i=1}^{n} P_B(X_i | P_a(X_i))$.

1.2 Формула Байеса и формула полной вероятности

Если событие A может произойти только при выполнении одного из событий $B_1, B_2, \ldots, B_n$, которые образуют полную группу несовместных событий, то вероятность события A вычисляется по формуле (1.2):

$$P(A) = P(B_1) * P(A|B_1) + P(B_2) * P(A|B_2) + \cdots + P(B_n) * P(A|B_n). \quad (1.2)$$

Эта формула называется формулой полной вероятности [5,6,7,8].

Пусть $H_1, H_2, \ldots$ — полная группа событий, и A — некоторое событие, вероятность которого положительна. Тогда условная вероятность того, что имело место событие H_k, если в результате эксперимента наблюдалось событие A, может быть вычислена по формуле (1.3):

$$P(H_k|A) = \frac{P(H_k) * P(A|H_k)}{\sum_{i=1}^{k} P(H_i) * P(A|H_i)} \quad (1.3)$$

Называемой формулой Байеса [5,6,7,8].

1.3 Распространение свидетельств в байесовской сети без циклов

Свидетельства - утверждения вида "событие в узле x произошло".

Рассмотрим случай с 3 вершинами , на рисунке 1.1 мы видим байесовскую сеть , которая представляет собой граф без циклов с сходящийся связью , в данном случае вершины x_1 и x_2 родительские, а вершина x_3 дочерняя.

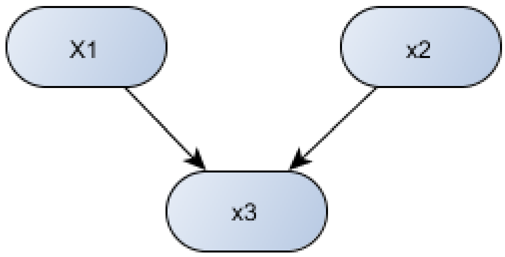

Рисунок 1.1 – Граф с 3 вершинами

И так, как будет происходить распространение свидетельств в данном случае, если распространение свидетельств идет от родительских вершин тоесть от x_1 или x_2, в дочернею вершину x_3, то вероятность события вершины x_3 будет высчитываться с помощью формулы полной вероятности (1.2), в случае же если распространение свидетельства идет от дочерней вершине к родительской будет использована формула Байеса (1.3) [1].

1.4 Циклы в Байесовских сетях доверия

Байесовская сеть доверия — это ациклический направленный граф, вершины которого представляют собой переменные, а связи между ними указывают на непосредственную зависимость между этими переменными. Основная идея байесовской сети доверия заключается в том, что слишком большое для прямого анализа множество утверждений (переменных) структурируется при помощи графа, в котором связаны друг с другом вершины, отвечающие зависимым друг от друга утверждениям, а получающаяся структура поддается достаточно простому анализу.

В байесовской сети доверия связь (говоря языком теории вероятностей и вероятностной логики, зависимость) на самом деле имеется отнюдь не только между узлами, непосредственно связанными ребром. Определение БСД включает в себя понятие так называемой d-разделимости, которая говорит, когда узлы БСД условно независимы.

Для того чтобы определить d-разделимость, сначала определяют три способа связи между узлами. Последовательная связь — это прямая связь между узлами сети представленная на рисунке 1.2. В этом случае 1 влияет на 2, а 2, в свою очередь, влияет на 3, и узлы 1 и 3 получаются связанными. Однако, если в 2 поступило свидетельство, связь между 1 и 3 нарушается: если мы твердо знаем, что промежуточное событие произошло, связи между его причиной и следствием уже нет (ее обеспечивал тот факт, что вероятность следствия зависит от вероятности причины, но если следствие уже произошло, этот механизм больше не работает).

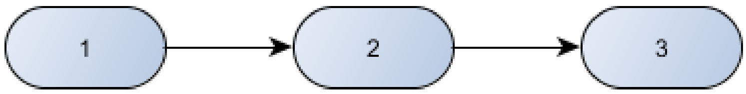

Рисунок 1.2 – Последовательная связь в БСД

В случае расходящейся связи у одного узла сети имеется несколько потомков, данная связь представлена на рисунке 1.3. Информация об одном из потомков может повлиять на вероятность другого потомка одного и того же узла. Это связано с тем, что не только информация о произошедшей причине повышает вероятность (правильнее, конечно, было бы сказать «оценку вероятности», а еще правильнее и более нейтрально — «степень нашей уверенности в», но для краткости мы будем говорить просто «вероятность») следствия, но и случившееся следствие повышает вероятность причины. Но если общий предок уже получил означивание, то связь нарушается. Действительно, если мы уже твердо знаем, что причина двух следствий имеет место, эти следствия становятся независимыми.

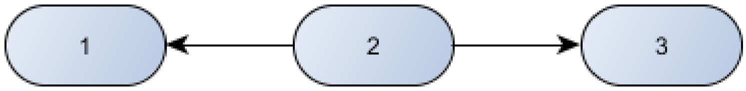

Рисунок 1.3 – Расходящаяся связь в БСД

Сходящаяся связь возникает, когда два узла имеют общего потомка, данная связь представлена на рисунке 1.4. Эта ситуация — самая интересная. Казалось бы, связи между узлами 1 и 3 нет: если произошло 1, то это повлияет на вероятность события 2, но вероятность 3 от этого измениться не должна. Однако ситуация меняется, если свидетельство 2 уже получено. Действительно, после получения свидетельства 2 вероятности обеих его причин повышаются: раз уж что-то произошло, значит, какая-то причина у этого была — неизвестно, какая именно, но обе причины становятся более вероятными. Но если теперь мы вдруг узнаем, что одна из причин произошла, это должно понизить вероятность другой причины: ведь следствие уже объяснено, и его влияние на вторую причину должно уменьшиться. Итак, в случае сходящейся связи зависимость должна появляться, только если общий потомок (или любой из его потомков — там ситуация будет аналогичной) уже получил какое-то означивание.

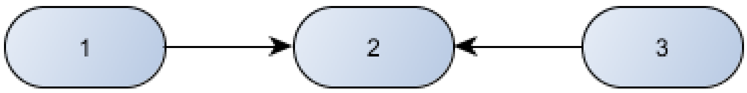

Рисунок 1.4 – Сходящаяся связь в БСД

Теперь можно дать формальное определение d-разделимости.

Определение 1. Два узла направленного графа x и y называются разделенными, если для всякого пути из x в y (путь здесь не учитывает направление ребер) существует такой промежуточный узел (не совпадающий ни с x, ни с y), что либо связь в пути в этом узле последовательная или расходящаяся, и узел z получил означивание, либо связь сходящаяся, и ни узел z, ни какой-либо из его потомков означивания не получил. В противном случае узлы называются d-связанными.

Простой цикл в байесовских сетях доверия сводится к линейному рекурсивному уравнению и нахождению маргинальной вероятности [1].

В данном случае представленном на рисунке 1.5, система линейных уравнений принимает следующий вид.

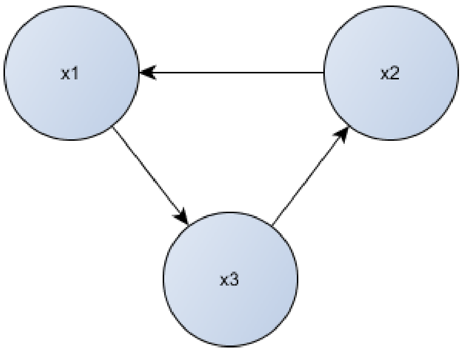

Рисунок 1.5 – Цикл из 3 вершин

$$\begin{cases} p(x_1) = p(x_1 \mid x_3)p(x_3) + p(x_1 \mid \overline{x}_3)(1 - p(x_3)), \\ p(x_2) = p(x_2 \mid x_1)p(x_1) + p(x_2 \mid \overline{x}_1)(1 - p(x_1)), \\ p(x_3) = p(x_3 \mid x_2)p(x_2) + p(x_3 \mid \overline{x}_2)(1 - p(x_2)). \end{cases}$$

Так как в данном случае любая пара переменных является парой соседей в цикле, все вероятности конъюнкций длины два заданы однозначно. Единственным неизвестным является $p(x1x2x3)$. Решив систему, получим маргинальные вероятности вершин цикла и попарные совокупные вероятности. Примем теперь за неизвестную величину $p(x1x2x3)$ и запишем систему ограничений, вытекающую из простейших свойств вероятности:

$$
\begin{cases}
p(x_1x_2x_3) \ge 0, \\
p(x_1x_2) - p(x_1x_2x_3) \ge 0, \\
p(x_1x_3) - p(x_1x_2x_3) \ge 0, \\
p(x_2x_3) - p(x_1x_2x_3) \ge 0, \\
p(x_1) - p(x_1x_2) - p(x_1x_3) + p(x_1x_2x_3) \ge 0, \\
p(x_2) - p(x_1x_2) - p(x_2x_3) + p(x_1x_2x_3) \ge 0, \\
p(x_3) - p(x_1x_3) - p(x_2x_3) + p(x_1x_2x_3) \ge 0, \\
1 - p(x_1) - p(x_2) - p(x_3) + p(x_1x_2) + p(x_2x_3) + p(x_1x_3) - p(x_1x_2x_3) \ge 0.
\end{cases}
$$

Для ее решения достаточно пересечь вышеуказанные интервалы. В явном виде ее решение запишется следующим образом:

$$
\max\left\{\begin{matrix} 0, \\ p(x_1x_2) + p(x_1x_3) - p(x_1), \\ p(x_1x_2) + p(x_2x_3) - p(x_2), \\ p(x_1x_3) + p(x_2x_3) - p(x_3) \end{matrix}\right\} \le p(x_1x_2x_3) \le \min\left\{\begin{matrix} p(x_1x_2), \\ p(x_1x_3), \\ p(x_2x_3), \\ 1 - p(x_1) - p(x_2) - p(x_3) + \\ + p(x_1x_2) + p(x_1x_3) + p(x_2x_3) \end{matrix}\right\}.
$$

Затем можно выразить через $p(x1x2x3)$ и все остальные вероятности, формирующие совокупное распределение:

$$
\begin{aligned}
&p(\bar{x}_1x_2x_3) = p(x_2x_3) - p(x_1x_2x_3), \\
&p(x_1\bar{x}_2x_3) = p(x_1x_3) - p(x_1x_2x_3), \\
&p(x_1x_2\bar{x}_3) = p(x_1x_2) - p(x_1x_2x_3), \\
&p(\bar{x}_1\bar{x}_2x_3) = p(x_3) - p(x_2x_3) - p(x_1x_3) + p(x_1x_2x_3), \\
&p(\bar{x}_1x_2\bar{x}_3) = p(x_2) - p(x_1x_2) - p(x_2x_3) + p(x_1x_2x_3), \\
&p(x_1\bar{x}_2\bar{x}_3) = p(x_1) - p(x_1x_2) - p(x_1x_3) + p(x_1x_2x_3), \\
&p(\bar{x}_1\bar{x}_2\bar{x}_3) = 1 - p(x_1) - p(x_2) - p(x_3) + p(x_1x_2) + p(x_1x_3) + p(x_2x_3) - p(x_1x_2x_3).
\end{aligned}
$$

Как наиболее эффективно решать вышеописанную систему уравнений? Мы опишем один из простейших, но вместе с тем достаточно удобных и эффективных подходов к вычислению $p(x_n)$, а значения остальных оценок вероятностей в таком случае можно будет просто, обойдя цикл, подсчитать по известным значениям условных вероятностей.

Итак, мы решаем систему $\boldsymbol{Ap} = \boldsymbol{B}$. Мы хотим подсчитать $p(x_n)$ по формуле Крамера, разделив определитель матрицы $\boldsymbol{A}'$, которая получается из $\boldsymbol{A}$

заменой последнего столбца на столбец $\boldsymbol{B}$, на определитель матрицы $\boldsymbol{A}$.

Определитель det $\boldsymbol{A}$ вычисляется: $det\,A = 1 + (-1)^n\, \varepsilon_{1n}\varepsilon_{21}\varepsilon_{32} \dots \varepsilon_{n,n-1}$.

Вот так выглядит матрица $\boldsymbol{A}$:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & p(x_1 \mid \overline{x}_n) \\ \xi_{21} & 1 & 0 & \cdots & 0 & p(x_2 \mid \overline{x}_1) \\ 0 & \xi_{32} & 1 & \cdots & 0 & p(x_3 \mid \overline{x}_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \xi_{n,n-1} & p(x_n \mid \overline{x}_{n-1}) \end{pmatrix}.$$

Предлагаемый нами способ вычислять ее определитель — это рекуррентно, один за другим, вычислять аналогичные определители меньшего порядка, раскладывая определители по последней строке. Вот как это выглядит (мы обозначаем через Δ_i определитель подматрицы порядка i (при этом положим $\Delta_0 = 0$):

$$\Delta_0 = 0,$$

$$\Delta_1 = -\xi_{1n}\Delta_0 + p(x_1 \mid \overline{x}_n),$$

$$\Delta_2 = \begin{vmatrix} 1 & p(x_1 \mid \overline{x}_n) \\ \xi_{21} & p(x_2 \mid \overline{x}_1) \end{vmatrix} = -\xi_{21}\Delta_1 + p(x_2 \mid \overline{x}_1),$$

$$\Delta_3 = \begin{vmatrix} 1 & 0 & p(x_1 \mid \overline{x}_n) \\ \xi_{21} & 1 & p(x_2 \mid \overline{x}_1) \\ 0 & \xi_{32} & p(x_3 \mid \overline{x}_2) \end{vmatrix} = -\xi_{32}\Delta_2 + p(x_3 \mid \overline{x}_2),$$

$$\Delta_4 = \begin{vmatrix} 1 & 0 & 0 & p(x_1 \mid \overline{x}_n) \\ \xi_{21} & 1 & 0 & p(x_2 \mid \overline{x}_1) \\ 0 & \xi_{32} & 1 & p(x_3 \mid \overline{x}_2) \\ 0 & 0 & \xi_{43} & p(x_4 \mid \overline{x}_3) \end{vmatrix} = -\xi_{43}\Delta_3 + p(x_4 \mid \overline{x}_3),$$

$$\dots,$$

$$\Delta_i = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & p(x_1 \mid \overline{x}_n) \\ \xi_{21} & 1 & 0 & \cdots & 0 & p(x_2 \mid \overline{x}_1) \\ 0 & \xi_{32} & 1 & \cdots & 0 & p(x_3 \mid \overline{x}_2) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & p(x_{i-1} \mid \overline{x}_{i-2}) \\ 0 & 0 & 0 & \cdots & \xi_{i,i-1} & p(x_i \mid \overline{x}_{i-1}) \end{vmatrix} = -\xi_{i,i-1}\Delta_{i-1} + p(x_i \mid \overline{x}_{i-1}).$$

Эти формулы представляют собой простую рекурсию. Кроме того, они удачны с вычислительной точки зрения — для вычисления $p(x_n)$ потребуется произвести лишь одно деление $p(x_n) = \frac{\Delta_n}{\det \boldsymbol{A}}$, а определитель будет вычислен за линейное количество сложений и умножений; такие параметры позволяют рассчитывать на быстроту и хорошую устойчивость метода.

После получения $p(x_n)$ остальные вероятности вычислить уже совсем несложно:

$$p(x_i) = p(x_i \mid x_{i-1})p(x_{i-1}) + p(x_i \mid \overline{x}_{i-1})(1 - p(x_{i-1})),\ p(x_i x_{i-1}) = p(x_i \mid x_{i-1})p(x_{i-1}).$$

1.5 Алгоритм пропагации свидетельств в байесовской сети с циклами

И так что мы должны сделать наш алгоритм можно разбить на четыре шага:

- Морализовать базовый граф БСД;
- Триангулировать моральный граф;
- Построить дерево смежности;
- Выполнить собственно пропагацию свидетельств.

Рассмотрим 4 шаг на примере представленном на рисунке 1.6.

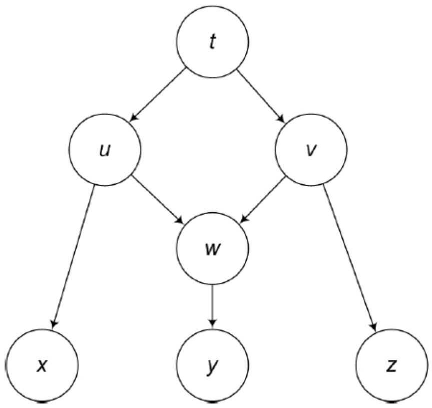

Рисунок 1.6 – Пример графа

И так предположим что мы хотим найти Z. В данном случае такая клика одна- {V,Z} представленная на рисунке 1.7.

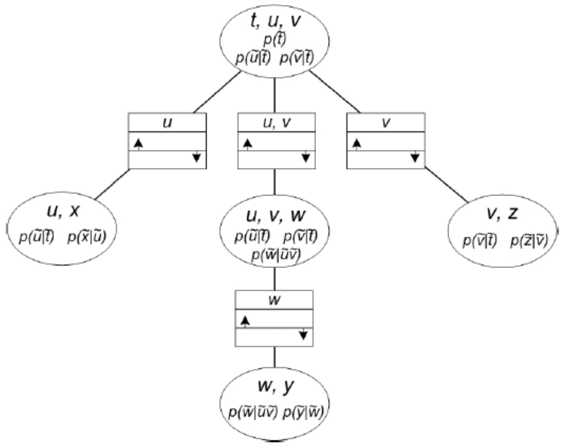

Рисунок 1.7 – Дерево сочленений

Эта клика становится временным корнем дерева, а остальные вершины, начиная с листьев, посылают к временному корню сообщения, в которых записан результат суммирования по переменным, не содержащимся в ближайшем сепараторе.

Каждая вершина посылает сообщение "наверх" тогда, когда она получила все сообщения "снизу".

Весь алгоритм можно представить как последовательную передачу сообщений между узлами тогда, когда выполняется условие: Пусть в моральном графе есть клика С, которой соответствует набор вероятностей ψ_C, а у нее - соседний сепаратор S. Пусть остальные соседи $C - S_1, \ldots, S_K$ - уже получили сообщения ψ_i от своих соседей, т.е. во входящих почтовых ящиках $S_1, \ldots, S_K$ уже лежат готовые распределения вероятностей. Назовем такую ситуацию условием передачи сообщения.

На этапе сбора сведений условие выполняется снизу вверх, и узлы передают сообщения на верх.

Потом, когда мы добрались до корня и пересчитали там, мы начинаем двигаться обратно; условие не меняется, но теперь оно работает для всех исходящих сепараторов.

Алгоритм заканчивает работу, когда пустых почтовых ящиков больше нет [1].

1.6 Пример построения простейшей байесовской сети доверия

В этом примере рассматриваемом кандидатом технических наук доцентом Хабаровым Сергеем Петровичем в его курсе лекций "Интеллектуальные информационные системы" мы будем рассматривать небольшую яблочную плантацию "яблочного Джека". Однажды Джек обнаружил, что его прекрасное яблочное дерево лишилось листвы. Теперь он хочет выяснить, почему это случилось.

Он знает, что листва часто опадает, если:

- дерево засыхает в результате недостатка влаги;
- или дерево болеет.

Данная ситуация может быть смоделирована байесовской сетью доверия представленной на рисунке 1.8, содержащей 3 вершины: "Болеет", "Засохло" и "Облетело".

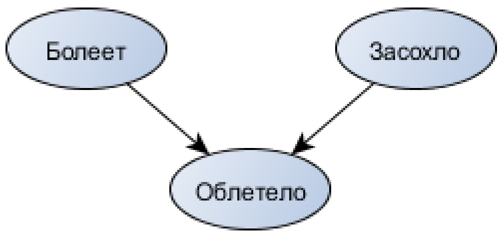

Рисунок 1.8 – Пример байесовской сети доверия с тремя событиями

В данном простейшем случае рассмотрим ситуацию, при которой каждая вершина может принимать всего лишь два возможных состояний и, как следствие находится в одном из них, данная ситуация показана в таблице 1.1.

Таблица 1.1 – Таблица состояний байесовской сети

Вершина(событие)БСД	Состояние 1	Состояние 2
болеет	болеет	нет
засохло	засохло	нет
облетело	да	нет

Вершина "Болеет" говорит о том, что дерево заболело, если оно находится в состоянии "болеет", в противном случае она находится в состоянии "нет". Аналогично для других двух вершин. Рассматриваемая байесовская сеть доверия, моделирует тот факт, что имеется причинно-следственная зависимость от события "Болеет" к событию "Облетело" и от события "Засохло" к событию "Облетело". Это отображено стрелками на байесовской сети доверия представленной на рисунке 1.8.

Когда есть причинно-следственная зависимость от вершины A к другой вершине B, то мы ожидаем, что когда A находится в некотором определённом состоянии, это оказывает влияние на состояние B. Следует быть внимательным, когда моделируется зависимость в байесовских сетях доверия. Иногда совсем не очевидно, какое направление должна иметь стрелка.

Например, в рассматриваемом примере, мы говорим, что имеется зависимость от "Болеет" к "Облетело", так как когда дерево болеет, это может вызывать опадание его листвы. Опадание листвы является следствием болезни, а не болезнь - следствием опадания листвы.

На приведенном выше рисунке дано графическое представление байесовской сети доверия. Однако, это только качественное представление байесовской сети доверия. Перед тем, как назвать это полностью байесовской сетью доверия необходимо определить количественное представление, то есть множество таблиц условных вероятностей представленных в таблице 1.2 и 1.3.

Таблица 1.2 – Таблица априорных вероятностей

Априорная вероятность (Болеет)		Априорная вероятность (Засохло)	
болеет	нет	засохло	нет
0.1	0.9	0.1	0.9

Таблица 1.3 – Таблица условных вероятностей p(Облетело)

Состояние	Засохло		Нет	
	болеет	нет	болеет	нет
да	0.95	0.85	0.9	0.02
нет	0.05	0.15	0.1	0.98

Приведенные таблицы иллюстрируют ТУВ для трёх вершин байесовской сети доверия. Заметим, что все три таблицы показывают вероятность пребывания некоторой вершины в определённом состоянии, обусловленным состоянием её родительских вершин. Но так как вершины Болеет и Засохло не имеют родительских вершин, то их вероятности являются маргинальными, т.е. не зависят (не обусловлены) ни от чего.

На данном примере мы рассмотрели, что и как описывается очень простой байесовской сетью доверия. Современные программные средства (такие как MSBN, Hugin и др.) обеспечивают инструментарий для построения таких сетей, а также возможность использования байесовских сетей доверия для введения новых свидетельств и получения решения (вывода) за счёт пересчёта новых вероятностей во всех вершинах, соответствующих вновь введенным свидетельствам.

В нашем примере пусть известно, что дерево сбросило листву. Это свидетельство вводится выбором состояния "да" в вершине "Облетело". После этого можно узнать вероятности того, что дерево засохло. Для приведенных выше исходных данных, результаты вывода путем распространения вероятностей по БСД будут представлены в таблице 1.4.

Таблица 1.4 – Таблица результатов

Облетело (Да)	Засохло (Нет)
0.47	0.49

2 Разработанные алгоритмы пропагации свидетельств

2.1 Измененный алгоритм пропагации

Инициализация вершин в Байесовской сети:

Инициализация проходит от вершин не имеющих родительских , вниз по дереву по формуле (2.1).

Алгоритм распространения свидетельств в Байесовской сети:

Шаг 1: Проводим инициализацию вершин.

Шаг 2: Ввод свидетельства в вершину (начальная).

Шаг 2.1:Помечаем вершину как активную.

Шаг 3:Распространяем свидетельства от родительских вершин к дочерним.

Шаг 3.1:Пересчитываем активную вершину по формуле (2.1).

Шаг 3.2:Делаем вершину неактивной.

Шаг 3.3:Снимаем метку.

Шаг 3.4:Запоминаем родителя-источник свидетельства.

Шаг 3.5:Если изменение текущей вершины больше
порога делаем активными все дочерние вершины.

Шаг 3.6: Если остались активные, возвращаемся к шагу 3.1

Шаг 4: начальная вершина активна.

Шаг 4.1:Распространяем свидетельства от дочерних вершин к родительским.

Шаг 4.2:Пересчитываем активную вершину по формуле (2.2).

Шаг 4.3:Делаем вершину неактивной.

Шаг 4.4:Снимаем метку,

Шаг 4.5:Запоминаем предка-источник свидетельства.

Шаг 4.6:Если изменение текущей вершины больше порога то Шаг 4.6.1,Шаг 4.6.2

Шаг 4.6.1:Помечаем дочерние вершины (кроме источника и вершин посчитанных на шаге 3).

Шаг 4.6.2:Делаем активными все родительские вершины.

Шаг 4.7: Если остались активные, возвращаемся к шагу 4.1

Шаг 5 :Все помеченные вершины делаем активными.

Шаг 6:Распространяем свидетельства от родительских вершин к дочерним.

Шаг 6.1:Пересчитываем активную вершину по формуле (2.1).

Шаг 6.2:Делаем вершину неактивной.

Шаг 6.3:Снимаем метку.

Шаг 6.4:Запоминаем родителя-источник свидетельства.

Шаг 6.5:Если изменение текущей вершины больше порога делаем активными все дочерние вершины, за исключением вершин посчитанных на шагах 3 и 4.

Шаг 6.6: Если остались активные, возвращаемся к шагу 6.1

Алгоритм распространения свидетельств в Байесовской сети с циклами:

Шаг 1: Проводим инициализацию вершин.

Шаг 2: Ввод свидетельства в вершину (начальная).

Шаг 2.1:Помечаем вершину как активную

Шаг 3:Распространяем свидетельства от родительских вершин к дочерним.

Шаг 3.1:Пересчитываем активную вершину по формуле (2.1).

Шаг 3.2:Делаем вершину неактивной,

Шаг 3.3:Снимаем метку.

Шаг 3.4:Запоминаем родителя-источник свидетельства.

Шаг 3.5:Если вершина ранее посещалась от данного родителя-источника свидетельства, значит в сети есть цикл. Иначе Шаг 3.6.

Шаг 3.5.1: Обновляем min и max для нового значения.

Шаг 3.5.2: За новое значение берется среднее min и max.

Шаг 3.5.3: Если изменение текущей вершины больше порога

делаем активными все дочерние вершины.

Шаг 3.5.4:Переходим к Шагу 3.7.

Шаг 3.6:Если изменение текущей вершины больше порога, то

делаем активными все дочерние вершины. Дочерние вершины приравнивают min max.

Шаг 3.7: Если остались активные, возвращаемся к шагу 3.1.

Шаг 4: Начальная вершина активна. Распространяем свидетельства от дочерних вершин к родительским.

Шаг 4.1:Пересчитываем активную вершину по формуле (2.2).

Шаг 4.2:Делаем вершину неактивной.

Шаг 4.3:Снимаем метку.

Шаг 4.4:Запоминаем потомка-источник свидетельства.

Шаг 4.2:Если вершина ранее посещалась от данного потомка -источника свидетельства, значит в сети есть цикл. Иначе Шаг 4.5:

Шаг 4.2.1: обновляем min и max для нового значения

Шаг 4.2.2: за новое значение берется среднее min и max

Шаг 4.2.3: Если изменение текущей вершины больше порога делаем активными все родительские вершины. Помечаем дочерние вершины (кроме источника).

Шаг 4.2.4: Переходим к Шагу 3.7

Шаг 4.5:Если изменение текущей вершины больше порога, то помечаем дочерние вершины (кроме источника), делаем активными все родительские вершины, родительские вершины приравнивают min max.

Шаг 4.6: Если остались активные, возвращаемся к шагу 4.1

Шаг 5 :Все помеченные вершины делаем активными.

Шаг 6:Распространяем свидетельства от родительских вершин к дочерним.

Шаг 6.1:Пересчитываем активную вершину по формуле (2.1).

Шаг 6.2:Делаем вершину неактивной.

Шаг 6.3:Снимаем метку.

Шаг 6.4:Запоминаем родителя-источник свидетельства.

Шаг 6.5:Если вершина ранее посещалась от данного родителя-источника свидетельства, значит в сети есть цикл. Иначе Шаг 6.6.

Шаг 6.5.1: Обновляем min и max для нового значения.

Шаг 6.5.2: За новое значение берется среднее min и max.

Шаг 6.5.3: Если изменение текущей вершины больше порога, то делаем активными дочерние вершины за исключением вершин посчитанных на шагах 3 и 4.

Шаг 6.5.4: Шаг 6.7.

Шаг 6.6:Если изменение текущей вершины больше порога, то делаем активными дочерние вершины за исключением вершин посчитанных на шагах 3 и 4. Дочерние вершины приравнивают min max.

Шаг 6.7: Если остались активные, возвращаемся к шагу 6.1.

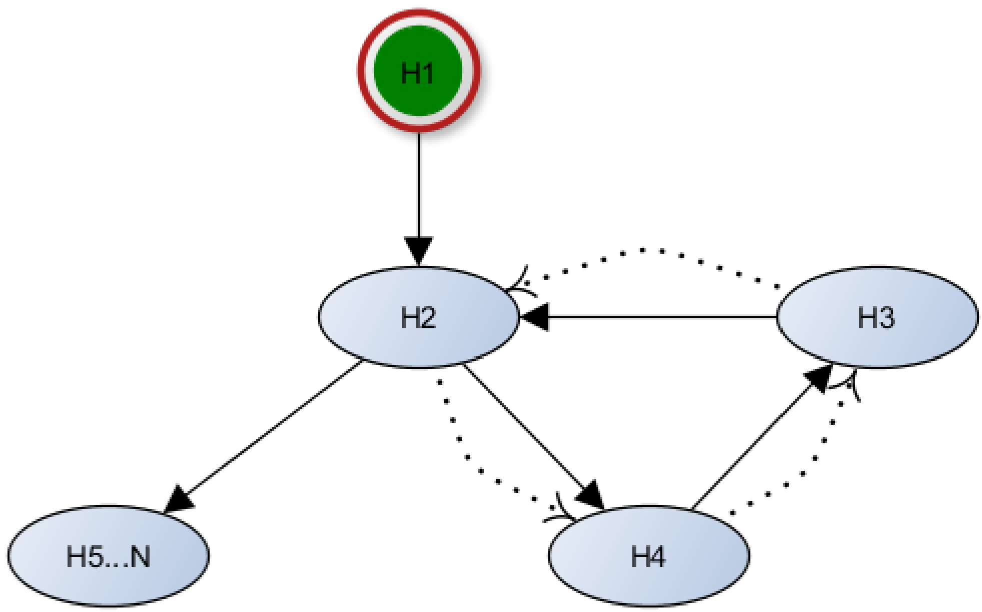

Рисунок 2.1 – Байесовская сеть с циклами

В случае если в байесовской сети возникают циклы, пример байесовской сети с циклами представлен на рисунке 2.1, будет применятся метод интервальных оценок , вершина H2 будет хранить в себе 2 вероятности

"Минимальную" и "Максимальную" в дальнейшем будет называть их min и max.

Когда свидетельства поступили от вершины H1 и вероятность вершины H2 пересчиталась по формуле (2.1), min и max будут равны этой вероятности, далее распространение свидетельств идет по вершине H4, H3 обратно к H2, здесь мы сравниваем вероятность полученную при поступлении свидетельства от вершины H3 с уже имеющейся у H2 вероятностью, и меняем наш min и max, на соответствующие вероятности, далее мы опять проходим по циклу и сравниваем пересчитанную в результате нового обхода цикла вероятность с уже имеющимися вероятностями min и max, в случае если она попадает в этот интервал, мы выходим из цикла и распространяем свидетельства дальше к вершине H5...N, если же вероятность вышла за пределы min или max мы меняем одно из них соответствующей вероятностью и продолжаем цикл до тех пор пока полученная вероятность не попадет в этот интервал.

Формулы для пересчета вероятностей используются те же что и для сетей без циклов, в зависимости откуда поступило свидетельство и куда, в случае если оно поступило от родительской к дочерней мы используем формулу (2.1), а если от дочерней к родительской формулу (2.2).

Помеченные для перерасчета вершины не могут входить в цикл, данное утверждение представлено на рисунке 2.2.

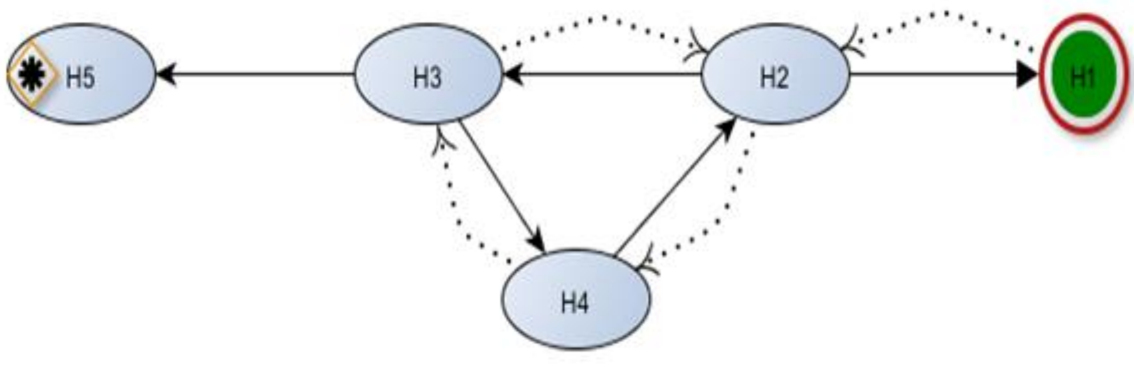

Рисунок 2.2 – Байесовская сеть с циклами и активной вершиной

Пунктирная линия на рисунке обозначает распространение свидетельств.

Вершина имеющая круглую форму и зеленую середину является вершиной получившей начальное свидетельство.

Вершина имеющая овальную форму и зеленый цвет является вершиной с уже пересчитанной вероятностью.

Вершина имеющая овальную форму с красным контуром, является вершиной для перерасчета.

Данные обозначения будут использоваться во всей работе.

2.2 Общие формулы для распространения свидетельств

В случае если распространение свидетельства идет от родительской вершины к дочерней используется данная формула (2.1), данный вид распространения свидетельств представлен на рисунке 2.3.

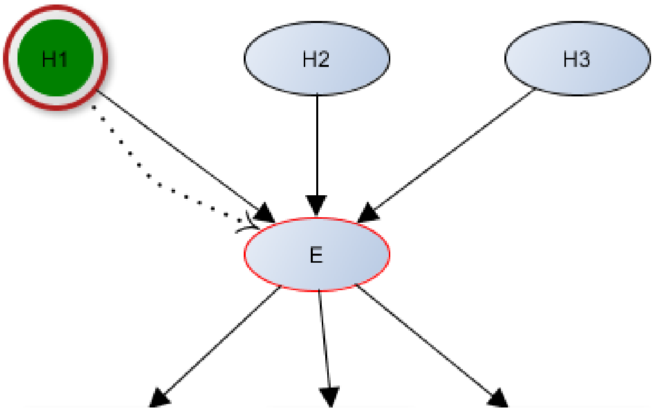

Рисунок 2.3 – Пример распределения свидетельств от дочерней к родительской

$$P(E) = \sum_i (P(E|H_i) * P(H_i)) \quad (2.1)$$

Если же свидетельство поступило от дочерней вершине к родительской используется данная формула (2.2), данный вид распределения свидетельств представлен на рисунке 2.4.

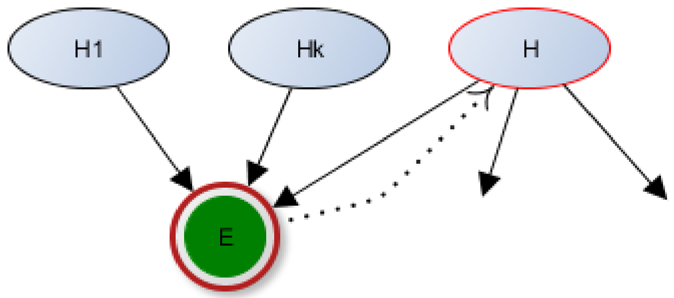

Рисунок 2.4 – Пример распределения свидетельств от дочерней к родительской

$$P(H) = P(H|E) * P(E)^{**} + P(H|\bar{E}) * P(\bar{E})^{**} \quad (2.2)$$

Где $P(E)^{**}$ и $P(\bar{E})^{**}$ это новые полные вероятности получившиеся при поступлении свидетельства из предыдущей вершины.

$$P(H|E) = \frac{\sum_{\{H1 \ldots K\}}(P(E|H \wedge \widetilde{H_1} \ldots \widetilde{H_n}) * P(H)^{*} * \prod_i^k P(\widetilde{H}))}{P(E)} \quad (2.3)$$

Где P(E) - полная вероятность посчитанная при инициализации сети.

А P^* - это вероятности родительских для них вершин, без учета нового изменения состояния E.

Переменные с $\widetilde{H}$ обозначают, что данные события могут быть истинными H, или ложными $\overline{H}$.

2.3 Применения формул для распространения свидетельств на конкретном примере

В случае если распространение свидетельства идет от родительской вершины к дочерней, для подсчета вероятностей вершины H1 на примере представленном на рисунке 2.5, используется формула (2.4).

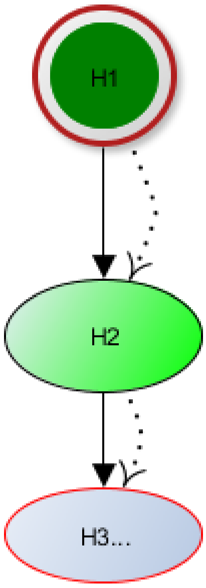

Рисунок 2.5 – Пример распространения свидетельств сверху вниз

$$P(H3\ldots) = P(H3|H2) * P(H2)^{**} \qquad (2.4)$$

Если же свидетельство поступило от дочерней вершине к родительской для подсчета вероятностей вершины H1 на примере представленном на рисунке 2.6, используется формула (2.5).

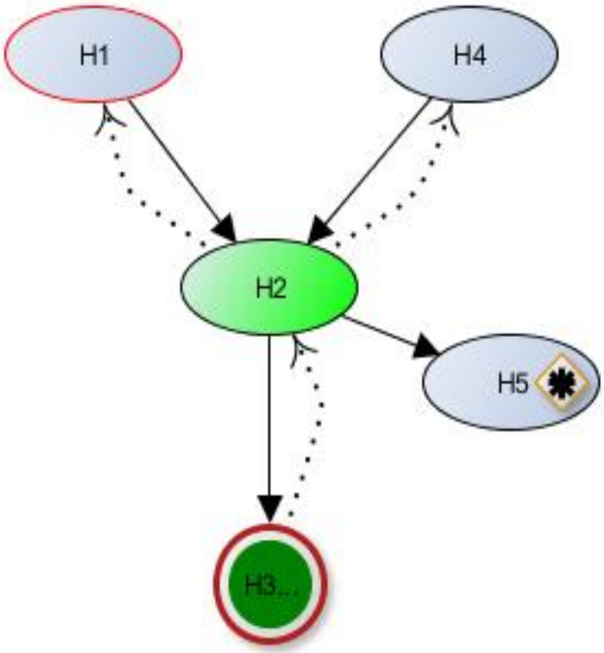

Рисунок 2.6 – Пример распространения свидетельств снизу вверх

$$P(H1) = P(H1|H2) * P(H2)^{**} + P(H1|\overline{H2}) * P(\overline{H2})^{**} \quad (2.5)$$

Где $P(H2)^{**}$ и $P(\overline{H2})^{**}$ это новые полные вероятности получившиеся при поступлении свидетельства из вершины H3.

$$P(H1|H2) = \frac{P(H2|H1 \wedge H4) * P(H1)^{*} * P(H4)^{*} + P(H2|H1 \wedge \overline{H4}) * P(H1)^{*} * P(\overline{H4})^{*}}{P(H2)} \quad (2.6)$$

$$P(H1|\overline{H2}) = \frac{P(\overline{H2}|H1 \wedge H4) * P(H1)^{*} * P(H4)^{*} + P(\overline{H2}|H1 \wedge \overline{H4}) * P(H1)^{*} * P(\overline{H4})^{*}}{P(\overline{H2})} \quad (2.7)$$

Где P(H2) - полная вероятность посчитанная при инициализации сети.

А P^{*} - это вероятности родительских для H2 вершин, без учета нового изменения состояния H2.

2.4 Оценка сложности алгоритма

Рассмотрим самый худший вариант для разработанного алгоритма пропагации свидетельств в сетях с циклами. А именно максимальное количество циклов в сети. В этом случае каждый цикл состоит из 3 вершин всего циклов: $\frac{N}{3}$, где N - это количество вершин в сети.

Эксперименты показывают что в среднем в одном цикле выполняются 2 обхода, количество дополнительных обходов можно оценить как пропорциональное количеству циклов, таким образом сложность алгоритма получается N^3.

3 Апробации алгоритмов и их сравнение

3.1 Апробация алгоритмов на сети без циклов

Графическое представление сети на которой была проведена апробация алгоритмов, представлена на рисунке 3.1.

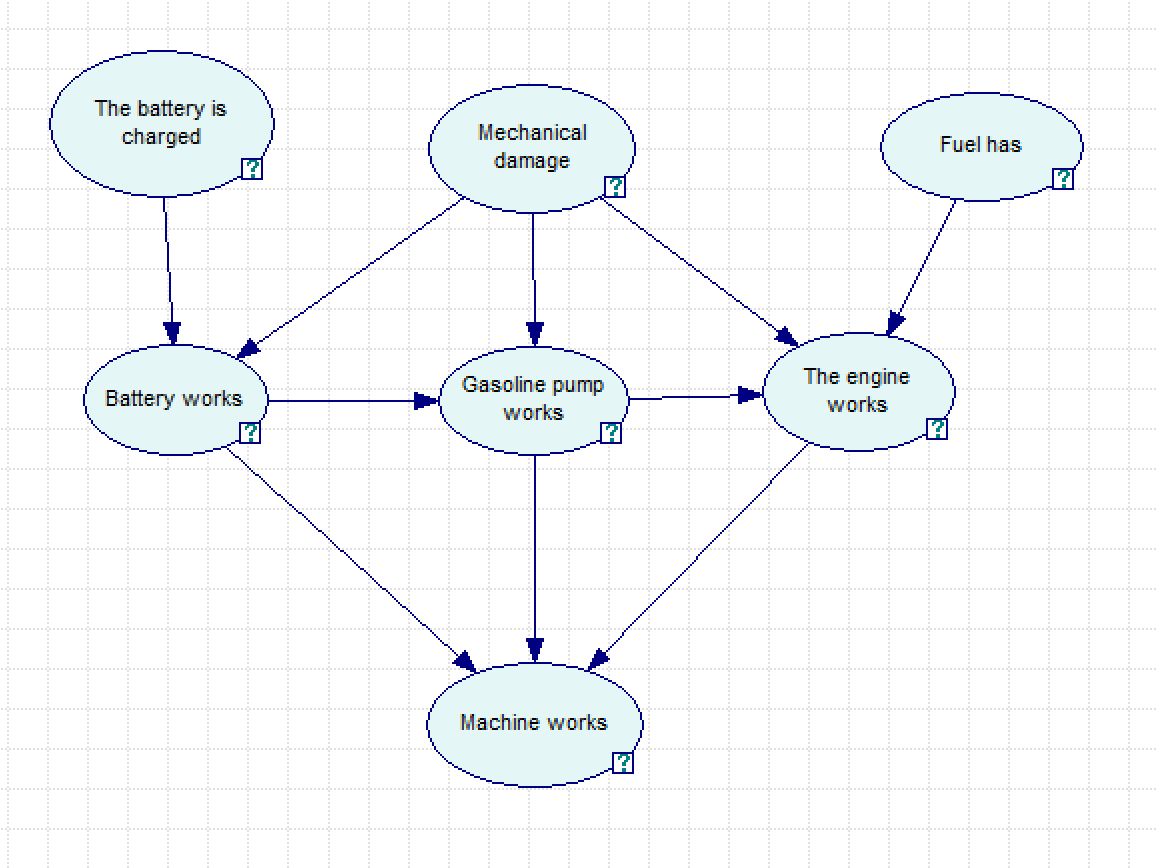

Рисунок 3.1 – Графическое представление Байесовской сети

Условные вероятности вершины "The battery is charged", "Mechanical damage", "Fuel has" представлены на рисунках 3.2, 3.3, 3.4 соответственно.

Charged	0.7
Discharged	0.3

Рисунок 3.2 – Вероятности вершины "The battery is charged"

damaged	0.8
disdamaged	0.2

Рисунок 3.3 – Вероятности вершины "Mechanical damage "

Full	0.3
Empty	0.7

Рисунок 3.4 – Вероятности вершины "Fuel has "

Вероятности вершины "Battery works", "Gasoline pump work", "The engine work", "Machine work" представлены на рисунке 3.5, 3.6, 3.7, 3.8 соответственно.

The battery is charged	Charged		Discharged	
Mechanical damage	damaged	disdamaged	damaged	disdamaged
Work	0.2	0.9	0.9	0.3
Not_Work	0.8	0.1	0.1	0.7

Рисунок 3.5 – Вероятности вершины "Battery work"

Mechanical damage	damaged		disdamaged	
Battery works	Work	Not_Work	Work	Not_Work
Gasoline_Work	0.4	0.1	0.9	0.4
Gasoline_Not_Work	0.6	0.9	0.1	0.6

Рисунок 3.6 – Вероятности вершины "Gasoline pump work"

Mechanical damage	damaged				disdamaged			
Fuel has	Full		Empty		Full		Empty	
Gasoline pump works	Gasoline_Work	Gasoline_Not_Work	Gasoline_Work	Gasoline_Not_Work	Gasoline_Work	Gasoline_Not_Work	Gasoline_Work	Gasoline_Not_Work
Engine_Work	0.4	0.3	0.2	0.1	0.8	0.4	0.4	0.3
Engine_Not_Work	0.6	0.7	0.8	0.9	0.2	0.6	0.6	0.7

Рисунок 3.7 – Вероятности вершины "The engine work "

Battery works	Work				Not_Work			
Gasoline pump works	Gasoline_Work		Gasoline_Not_Work		Gasoline_Work		Gasoline_Not_Work	
The engine works	Engine_Work	Engine_Not_Work	Engine_Work	Engine_Not_Work	Engine_Work	Engine_Not_Work	Engine_Work	Engine_Not_Work
Machine_Work	0.8	0.2	0.4	0.3	0.4	0.2	0.3	0.01
Machine_Not_Work	0.2	0.8	0.6	0.7	0.6	0.8	0.7	0.99

Рисунок 3.8 – Вероятности вершины "Machine work "

Сравнение результатов апробации алгоритмов "Genie" и алгоритма разработанного в данной работе представлены в таблицу 3.1, 3.2.

Таблица 3.1 – Сравнение результатов апробации алгоритмов

События	Вероятности полученные при инициализации сети, %		Вероятности полученные при вводе свидетельства в вершину "The battery is charged", %	
	алгоритм "Genie"	алгоритм пропагации	алгоритм "Genie"	алгоритм пропагации
The battery is charged	0.7	0.7	1	1
Mechanical damage	0.8	0.8	0,8	0.8
Fuel has	0.3	0,3	0,3	0.3
Battery works	0.47	0.47	0,34	0.34
Gasoline pump works	0,33	0. 33	0,3	0.3
The engine works	0,24	0.24	0,24	0.24
Machine works	0,23	0.23	0,19	0.19

Таблица 3.2 – Сравнение результатов апробации алгоритмов

События	Вероятности полученные при вводе свидетельства в вершину "Machine works", %	
	алгоритм "Genie"	алгоритм пропагации
The battery is charged	0,6	0.6
Mechanical damage	0,64	0.64
Fuel has	0,37	0.37

Окончание таблицы 3.2

События	Вероятности полученные при вводе свидетельства в вершину "Machine works", %	
	алгоритм "Genie"	алгоритм пропагации
Battery works	0,8	0.8
Gasoline pump works	0,58	0.58
The engine works	0,57	0.57
Machine works	1	1

3.2 Алгоритм пропагации в байесовских сетях с циклами

Графическое представление байесовской сети на которой будет проводится апробация алгоритмов представлена на рисунке 3.9.

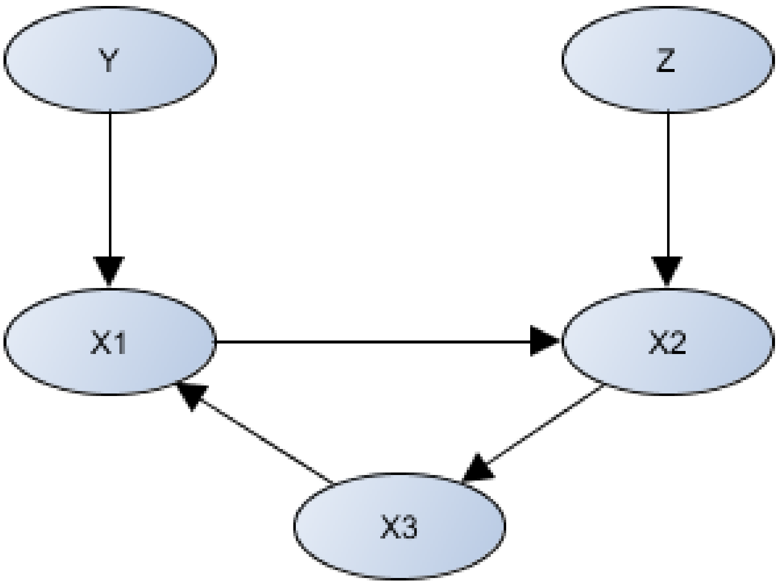

Рисунок 3.9 – Графическое представление байесовской сети

Априорные вероятности вершины Y, Z, представлены в таблицах 3.3, 3.4 соответственно.

Таблица 3.3 – Вероятности вершины Y

$P(Y)$	$P(\bar{Y})$
0,6	0,4

Таблица 3.4 – Вероятности вершины Z

$P(Z)$	$P(\bar{Z})$
0,3	0,7

Условные вероятности вершины X_1, X_2, X_3 представлены в таблицах 3.5, 3.6, 3.7 соответственно.

Таблица 3.5 – Вероятности вершины X_1

Y	$P(Y)$		$P(\bar{Y})$	
X_3	$P(X_3)$	$P(\overline{X_3})$	$P(X_3)$	$P(\overline{X_3})$
$P(X_1)$	0,2	0,4	0,25	0,1
$P(\overline{X_1})$	0,8	0,6	0,75	0,9

Таблица 3.6 – Вероятности вершины X_2

Z	$P(Z)$		$P(\bar{Z})$	
X_3	$P(X_3)$	$P(\overline{X_3})$	$P(X_3)$	$P(\overline{X_3})$
$P(X_2)$	0,6	0,2	0,2	0,9
$P(\overline{X_2})$	0,4	0,8	0,8	0,1

Таблица 3.7 – Вероятности вершины X_3

X_2	$P(X_2)$	$P(\overline{X_2})$
$P(X_3)$	0,8	0,5
$P(\overline{X_3})$	0,2	0,5

3.2.1 Подсчет вероятностей по алгоритму Николенко-Тулупьева-Сироткина

Разберем одну из ситуаций подробно [1]. Рассмотрим означивание предков *yz* . Тогда вероятности на цикле примут значения:

$P(X_1|X_3) = 0{,}2 \qquad P(X_2|X_1) = 0{,}6$

$P(X_1|\overline{X_3}) = 0{,}4 \qquad P(X_2|\overline{X_1}) = 0{,}2$

$P(X_3|X_2) = 0{,}8 \qquad P(X_3|\overline{X_2}) = 0{,}5$

Обработаем эти вероятности, как в случае изолированного цикла. Для этого выпишем и решим систему линейных уравнений:

$$\begin{cases} P(X_1) = 0{,}2P(X_3) + 0{,}4(1 - P(X_3)) \\ P(X_2) = 0{,}6P(X_1) + 0{,}2(1 - P(X_1)) \\ P(X_3) = 0{,}8P(X_2) + 0{,}5(1 - P(X_2)) \end{cases}$$

Решив указанную систему получаем, что:

$P(X_1) = 0{,}282 \qquad P(X_2) = 0{,}313 \qquad P(X_3) = 0{,}594$

Теперь по формулам:

$P(X_1X_3) = P(X_1|X_3)P(X_3)$,

$P(X_1X_2) = P(X_2|X_1)P(X_1)$,

$P(X_2X_3) = P(X_3|X_2)P(X_2)$

Получаем:

$P(X_1X_2) = 0{,}169 \qquad P(X_1X_3) = 0{,}119 \qquad P(X_2X_3) = 0{,}25$

Убедимся, что данное распределение непротиворечиво. Для этого проверим, что существует $P(X_1X_2X_3)$, удовлетворяющее следующему неравенству:

$$max\begin{Bmatrix} 0, \\ P(X_1X_2) + P(X_1X_3) - P(X_1), \\ P(X_1X_2) + P(X_2X_3) - P(X_2), \\ P(X_1X_3) + P(X_2X_3) - P(X_3) \end{Bmatrix} \leq$$

$$\leq P(X_1X_2X_3) \leq$$

$$\leq min\begin{Bmatrix} P(X_1X_2), \\ P(X_1X_3), \\ P(X_2X_3), \\ 1 - P(X_1) - P(X_2) - P(X_3) + P(X_1X_2) + P(X_1X_3) + P(X_2X_3) \end{Bmatrix}$$

Или в численном виде:

$0{,}106 \leq P(X_1X_2X_3) \leq 0{,}119$

Таким образом, мы получили, что означивание предков YZ дает непротиворечивое распределение вероятностей, на элементах цикла.

Для означивания родителей $\bar{Y}Z$ получаем вероятности на цикле:

$P(X_1|X_3) = 0{,}25 \qquad P(X_2|X_1) = 0{,}6$

$P(X_1|\overline{X_3}) = 0{,}1 \qquad P(X_2|\overline{X_1}) = 0{,}2$

$P(X_3|X_2) = 0{,}8 \qquad P(X_3|\overline{X_2}) = 0{,}5$

Решаем систему:

$$\begin{cases} P(X_1) = 0{,}25P(X_3) + 0{,}1(1 - P(X_3)) \\ P(X_2) = 0{,}6P(X_1) + 0{,}2(1 - P(X_1)) \\ P(X_3) = 0{,}8P(X_2) + 0{,}5(1 - P(X_2)) \end{cases}$$

Получаем:

$P(X_1) = 0{,}187 \qquad P(X_2) = 0{,}274 \qquad P(X_3) = 0{,}582$

Откуда:

$P(X_1X_2) = 0{,}112 \qquad P(X_1X_3) = 0{,}146 \qquad P(X_2X_3) = 0{,}219$

$0{,}07 \le P(X_1X_2X_3) \le 0{,}112$

Таким образом, мы получили, что означивание предков $\bar{Y}Z$ дает непротиворечивое распределение вероятностей, на элементах цикла.

Для означивания родителей $Y\bar{Z}$ получаем вероятности на цикле:

$P(X_1|X_3) = 0{,}2 \qquad P(X_2|X_1) = 0{,}2$

$P(X_1|\overline{X_3}) = 0{,}4 \qquad P(X_2|\overline{X_1}) = 0{,}9$

$P(X_3|X_2) = 0{,}8 \qquad P(X_3|\overline{X_2}) = 0{,}5$

Решаем систему:

$$\begin{cases} P(X_1) = 0{,}2P(X_3) + 0{,}4(1 - P(X_3)) \\ P(X_2) = 0{,}2P(X_1) + 0{,}9(1 - P(X_1)) \\ P(X_3) = 0{,}8P(X_2) + 0{,}5(1 - P(X_2)) \end{cases}$$

Получаем:

$P(X_1) = 0{,}257 \qquad P(X_2) = 0{,}72 \qquad P(X_3) = 0{,}716$

Откуда:

$P(X_1X_2) = 0{,}051 \qquad P(X_1X_3) = 0{,}143 \qquad P(X_2X_3) = 0{,}112$

$0{,}07 \le P(X_1X_2X_3) \le 0{,}112$

Таким образом, мы получили, что означивание предков $\bar{Y}Z$ дает непротиворечивое распределение вероятностей, на элементах цикла.

Для означивания родителей $\overline{YZ}$ получаем вероятности на цикле:

$P(X_1|X_3) = 0{,}25 \qquad P(X_2|X_1) = 0{,}2$

$P(X_1|\overline{X_3}) = 0{,}1 \qquad P(X_2|\overline{X_1}) = 0{,}9$

$P(X_3|X_2) = 0{,}8 \qquad P(X_3|\overline{X_2}) = 0{,}5$

Решаем систему:

$$\begin{cases} P(X_1) = 0{,}25P(X_3) + 0{,}1(1 - P(X_3)) \\ P(X_2) = 0{,}2P(X_1) + 0{,}9(1 - P(X_1)) \\ P(X_3) = 0{,}8P(X_2) + 0{,}5(1 - P(X_2)) \end{cases}$$

Получаем:

$P(X_1) = 0{,}209 \qquad P(X_2) = 0{,}754 \qquad P(X_3) = 0{,}726$

Откуда:

$P(X_1X_2) = 0{,}041 \qquad P(X_1X_3) = 0{,}182 \qquad P(X_2X_3) = 0{,}603$

$0{,}058 \le P(X_1X_2X_3) \le 0{,}041$

Но последнее неравенство не имеет решения. Следовательно, означивания родителей $\overline{YZ}$ - противоречиво. И это означивание надо исключить из байесовской сети доверия как невозможное. Таким образом, начальное распределение вероятностей Y и Z из исходного:

$P(YZ) = 0{,}209 \qquad P(\bar{Y}Z) = 0{,}754 \qquad P(Y\bar{Z}) = 0{,}726 \qquad P(\,\overline{YZ}\,) = 0{,}726$

Преобразуется в:

$P(YZ) = 0{,}25 \qquad P(\bar{Y}Z) = 0{,}17 \qquad P(Y\bar{Z}) = 0{,}58 \qquad P(\,\overline{YZ}\,) = 0$

Теперь объединим полученные на каждом из шагов вероятности. Для этого сложим вероятности, полученные в каждом из случаев, с весом равным вероятности этого случая. Тогда:

$P(X_1)$=0,282*0,25+0,187*0,17+0,257*0,58=0,251

$P(X_2)$= 0,312*0,25+0,275*0,17+0,72*0,58=0,544

$P(X_3)$= 0,594*0,25+0,582*0,17+0,716*0,58=0,663

3.2.2 Подсчет вероятностей по измененному алгоритму

Задаем пороговое значение равное E=0,05.

Проводим инициализацию вершин:

Инициализация вершины X_1:

Для родительских вершин: $P(Y) = 0{,}6$, $P(X_3)$ - неизвестно.

Вероятность вершины X_3 будем считать равной $P(X_3) = 0{,}5$, так как мы знаем что она может принять вероятность либо 0 либо 1, в теории байесовских сетей в таком случае принято брать среднее значение, тоесть 0,5.

Вероятность вершин X_1 высчитывается по формуле (2.1):

$$P(X_1) = 0{,}2 * 0{,}5 * 0{,}6 + 0{,}25 * 0{,}5 * 0{,}4 + 0{,}4 * 0{,}6 * 0{,}5 + 0{,}1 * 0{,}4 * 0{,}5 = 0{,}25$$

Инициализация вершины X_2:

$$P(X_2) = 0{,}6 * 0{,}3 * 0{,}25 + 0{,}2 * 0{,}7 * 0{,}25 + 0{,}2 * 0{,}3 * 0{,}75 + 0{,}9 * 0{,}7 * 0{,}75 = 0{,}598$$

Инициализация вершины X_3:

$$P(X_3) = 0{,}8 * 0{,}598 + 0{,}5 * 0{,}402 = 0{,}679$$

Ввод свидетельства в вершину Y:

Теперь вводим свидетельство в вершину $P(Y) = 1$, распространяем свидетельство дальше согласно алгоритму, от родительским к дочерним, тоесть к вершине X_1, пересчитываем ее вероятность по формуле (2.1) и ей заменяем вероятность посчитанную при инициализации вершин:

$$P(X_1) = 0{,}2 * 1 * 0{,}679 + 0{,}4 * 1 * 0{,}321 = 0{,}264$$

Распространяем свидетельство дальше, к вершине X_2, пересчитываем ее вероятность по формуле (2.1) и ей заменяем вероятность посчитанную при инициализации вершин:

$$P(X_2) = 0{,}6 * 0{,}3 * 0{,}264 + 0{,}2 * 0{,}7 * 0{,}264 + 0{,}2 * 0{,}3 * 0{,}736 + 0{,}9 * 0{,}7 * 0{,}736 = 0{,}592$$

Распространяем свидетельство дальше, к вершине X_3, пересчитываем ее вероятность по формуле (2.1) и ей заменяем вероятность посчитанную при инициализации вершин:

$$P(X_3) = 0{,}8 * 0{,}592 + 0{,}5 * 0{,}408 = 0{,}5$$

Распространяем свидетельство дальше, к вершине X_1, пересчитываем ее вероятность по формуле (2.1) и заменяем ей вероятность посчитанную при поступлении предыдущего свидетельства:

$$P(X_1) = 0{,}2 * 1 * 0{,}5 + 0{,}4 * 1 * 0{,}5 = 0{,}3$$

Распространяем свидетельство дальше, к вершине X_2, пересчитываем новую вероятность по формуле (2.1):

$$P(X_2) = 0{,}6 * 0{,}3 * 0{,}3 + 0{,}2 * 0{,}7 * 0{,}3 + 0{,}2 * 0{,}3 * 0{,}7 + 0{,}9 * 0{,}7 * 0{,}7 = 0{,}579$$

Так как свидетельство от вершины X_1 в вершину X_2 уже поступало, мы можем сделать вывод что начался цикл.

Создаем интервал вероятностей:

$[0{,}579 \dots 0{,}592]$

Считаем среднее:

$$\frac{0{,}579 + 0{,}592}{2} = 0{,}586$$

Сравниваем старую вероятность с новой и проверяем достаточно ли оно изменилось, чтобы имело смысл продолжать распространение свидетельств в цикле дальше:

$0{,}592 - 0{,}586 = 0{,}006$

Так как 0,006 > E продолжаем распространение свидетельств дальше, к вершине X_3, пересчитываем новую вероятность по формуле (2.1):

$P(X_3) = 0{,}8 * 0{,}586 + 0{,}5 * 0{,}414 = 0{,}676$

Так как свидетельство от вершины X_2 в вершину X_3 уже поступало, так же задаем интервальные оценки:

$[0{,}5 \dots 0{,}676]$

Считаем среднее:

$$\frac{0{,}5 + 0{,}676}{2} = 0{,}588$$

Сравниваем старую вероятность с новой и проверяем достаточно ли оно изменилось, чтобы имело смысл продолжать распространение свидетельств в цикле дальше:

$0{,}5 - 0{,}588 = 0{,}09$

Так как 0,09>E продолжаем распределение свидетельств дальше, к вершине X_1, пересчитываем новую вероятность по формуле (2.1):

$P(X_1) = 0{,}2 * 1 * 0{,}588 + 0{,}4 * 1 * 0{,}412 = 0{,}283$

Так как свидетельство от вершины X_3 в вершину X_1 уже поступало, так же задаем интервальные оценки:

$[0{,}283 \dots 0{,}3]$

Считаем среднее:

$$\frac{0{,}283 + 0{,}3}{2} = 0{,}291$$

Сравниваем старую вероятность с новой и проверяем достаточно ли оно изменилось, чтобы имело смысл продолжать распространение свидетельств в цикле дальше:

$$0{,}3 - 0{,}291 = 0{,}009$$

Так как 0,009>E продолжаем распределение свидетельств дальше, к вершине X_2, пересчитываем новую вероятность по формуле (2.1):

$$P(X_2) = 0{,}6 * 0{,}3 * 0{,}291 + 0{,}2 * 0{,}7 * 0{,}291 + 0{,}2 * 0{,}3 * 0{,}709 + 0{,}9 * 0{,}7 * 0{,}709 = \\ = 0{,}582$$

Так как свидетельство от вершины X_1 в вершину X_2 уже поступало, и текущая вершина уже имеет интервал вероятностей, мы сравниваем поступившую вероятность с интервалом , 0,582 попадает в текущий интервал $[0{,}579 \dots 0{,}592]$.

Значит средняя вероятность не изменилась, цикл закончился, распространение свидетельств завершено, интервалы заменяются на следующие итоговые вероятности вершин:

$P(X_1) = 0{,}291$

$P(X_2) = 0{,}582$

$P(X_3) = 0{,}588$

Ввод свидетельства в вершину Z с учетом вероятностей введенного ранее свидетельства в вершину Y:

Теперь вводим свидетельство в вершину $P(Z) = 1$, распространяем свидетельство дальше согласно алгоритму, от родительским к дочерним, тоесть к вершине X_2, пересчитываем ее вероятность по формуле (2.1) и ей заменяем вероятность посчитанную после ввода свидетельства в вершину Y:

$$P(X_2) = 0{,}6 * 1 * 0{,}291 + 0{,}2 * 1 * 0{,}709 = 0{,}317$$

Распространяем свидетельство дальше, к вершине X_3, пересчитываем ее вероятность по формуле (1) и ей заменяем вероятность посчитанную после ввода свидетельства в вершину Y:

$P(X_3) = 0,8 * 0,317 + 0,5 * 0,683 = 0,596$

Распространяем свидетельство дальше, к вершине X_1, пересчитываем ее вероятность по формуле (2.1) и ей заменяем вероятность посчитанную после ввода свидетельства в вершину Y:

$P(X_1) = 0,2 * 1 * 0,596 + 0,4 * 1 * 0,404 = 0,281$

Распространяем свидетельство дальше, к вершине X_2, пересчитываем ее вероятность по формуле (2.1) и заменяем ей вероятность посчитанную при поступлении предыдущего свидетельства:

$P(X_2) = 0,6 * 1 * 0,281 + 0,2 * 1 * 0,719 = 0,313$

Распространяем свидетельство дальше, к вершине X_3, пересчитываем ее вероятность по формуле (2.1) и заменяем ей вероятность посчитанную при поступлении предыдущего свидетельства:

$P(X_3) = 0,8 * 0,313 + 0,5 * 0,687 = 0,594$

Так как свидетельство от вершины X_2 в вершину X_3 уже поступало, мы можем сделать вывод, что начался цикл.

Создаем интервал вероятностей:

$[0,594 \dots 0,596]$

Считаем среднее:

$$\frac{0,594 + 0,596}{2} = 0,595$$

Сравниваем старую вероятность с новой и проверяем достаточно ли оно изменилось, чтобы имело смысл продолжать распространение свидетельств в цикле дальше:

$0,596 - 0,595 = 0,001$

Так как 0,001<E цикл завершается, распространение свидетельств завершено итоговые вероятности вершин равны:

$P(X_1) = 0,281$

$P(X_2) = 0,313$

$P(X_3) = 0,595$

3.2.3 Сравнение результатов работы алгоритма для байесовской сети с циклами

Сравнение результатов работы алгоритмов представлены в таблице 3.8.

Таблица 3.8 – Сравнение результатов апробации алгоритмов

Вершина	Алгоритм пропагации		Алгоритм Николенко-Тулупьева-Сироткина	
	инициализация	ввод парного свидетельства YZ	инициализация	ввод парного свидетельства YZ
Y	0,6	1	0,6	1
Z	0,3	1	0,3	1
X_1	0,25	0,281	0,251	0,282
X_2	0,598	0,313	0,544	0,313
X_3	0,679	0,595	0,663	0,594

Таким образом максимальное расхождение при инициализации вершин составило 0,054, а для парного свидетельства YZ 0,001.

Расхождение при инициализации связано с тем, что согласно алгоритму Николенко-Тулупьева-Сироткина означивание $\overline{YZ}$ - было исключено из сети как противоречивое, такой подход выглядит по крайней мере странным, получается что несогласованность между потомками разрешается изменением априорных вероятностей их предков.

В алгоритме разработанном в данной работе несогласованность потомков, не влияет на априорные вероятности их предков, что представляется более логичным.

ЗАКЛЮЧЕНИЕ

В ходе выполнения данной квалификационной работы были изучены теоретические источники по Байесовским сетям, рассмотрены существующие алгоритмы распространения свидетельств в сетях без циклов и с циклами.

В ходе изучения материала сделан вывод о том, что существующие алгоритмы пропагации для сетей с циклами, в большинстве своем, используют системы линейных уравнений. Количество этих систем напрямую зависит от количества циклов, и их предельная сложность сводится к факториалу от количества циклов. Главным недостатком известных алгоритмов принято следующая их особенность: если условные вероятности вершин цикла противоречат друг другу, то противоречие снимается изменением родительских, для цикла вершин.

В разработанном алгоритме распространения свидетельств в байесовских сетях с циклами несогласованность потомков не влияет на априорные вероятности их предков.

Апробации алгоритма выполнена как на искусственных примерах, так и на примерах из статей.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Тупалев, А.Л. Циклы в байесовских сетях, вероятностная семантика и отношения с соседними узлами / А. Л. Тупалев, С. И. Николенко, А. В. Сироткин. // Труды СПИИРАН. – 2006. – №3. – с. 240-263.
2. Heckerman, D. Dependency Networks for Inference Collaborative Filtering and Data Visualization // Journal of Machine Learning Research. – 2000 – №1. – p. 49–75.
3. Probabilistic Networks and Expert Systems: tutorial / R. G. Cowell, P. Dawid, S. L. Lauritzen, D. J. Spiegelhalter.– New York: Springer-Verlag, 1999. – 205 p.
4. Некруткин, В. В. Условная вероятность. Независимость событий. Испытания Бернулли / В. В. Некруткин // Теория Вероятностей : учебник / СПбГУ.– Санкт-Петербург, 2001. – с. 1-7.
5. Колмогоров, А. Н. Основные понятия теории вероятностей : учебное пособие / А. Н. Колмогоров. – Москва: Наука, 1974. – 412 с.
6. Прохоров, Ю. В. Теория вероятностей. Основные понятия. Предельные теоремы. Случайные процессы: учебное пособие / Ю. В. Прохоров, Ю. А. Розанов; Наука, – Изд. 2-е – Москва : 1973. – 488 с.
7. Лоэв, М. Теория вероятностей: учебное пособие / М. Лоэв; под. ред. Ю. В. Прохорова. – Москва : Издательство иностранной литературы, 1962. – 711 с.
8. Гмурман, В. Е. Теория вероятностей и математическая статистика: учебное пособие / В. Е. Гмурман. – Москва : 2003. – 479 с.