

А.В. ТОРОПОВА
**ПОДХОДЫ К ДИАГНОСТИКЕ СОГЛАСОВАННОСТИ
ДАННЫХ В БАЙЕСОВСКИХ СЕТЯХ ДОВЕРИЯ**

Торопова А.В. Подходы к диагностике согласованности данных в байесовских сетях доверия.

Аннотация. Байесовские сети доверия предоставляют возможность объединения нескольких видов информации, например полученной от экспертов или статистически, позволяют работать с неполной или неточной информацией, обладают наглядностью и другими полезными свойствами. Благодаря этому они стали популярным и весьма эффективным инструментом. Однако во многих областях исследования исходные используются полученные от экспертов данные, которые могут быть не согласованы, и поэтому в некоторых задачах следует использовать инструменты для проверки их согласованности.

В работе рассмотрены примеры применения аппарата байесовских сетей доверия в медицине и здравоохранении, экологии, экономике и риск-анализе, функциональной безопасности, социологии и других предметных областях и показана необходимость разработки методов для проверки согласованности исходных данных.

Цель работы – систематизировать с помощью обзора примеры и задачи, в которых применяются байесовские сети доверия, чтобы оценить, в какой степени в этих задачах учитывается диагностика согласованности исходных данных, и насколько важным является ее применение.

Ключевые слова: диагностика согласованности данных, байесовские сети доверия.

Toropova A.V. Approaches to the Data Coherence Diagnosis in Bayesian Belief Network Models.

Abstract. Bayesian belief networks provide the ability to combine different types of information, e.g. statistical or expert data, allow working with incomplete or inaccurate information; they have clarity and other useful properties. Due to this, Bayesian belief networks have become a popular and highly effective tool in many fields of research. However, in many research areas data provided by the experts can be incoherent, and so in some tasks one should use tools to verify their coherence. The paper discusses examples of application of the Bayesian belief networks in medicine and public health, ecology, economics and risk analysis, functional safety, sociology, and other research areas, and shows the need to develop methods to check the coherence of initial data. The purpose of this work is to systematize problems and examples that illustrate the use of Bayesian belief networks by reviewing and to assess their use of data coherence diagnosis and its importance.

Keywords: data coherence diagnosis, Bayesian belief networks.

1. Введение. Байесовские сети доверия предоставляют возможность объединения нескольких видов информации, например полученной от экспертов или статистически, позволяют работать с неполной или неточной информацией, обладают наглядностью и другими полезными свойствами. Благодаря этому они стали популярным и весьма эффективным инструментом во многих областях исследования.

В работе [11] байесовские сети доверия используются для построения модели оценки интенсивности социально-значимого поведения (это может быть рискованное поведение такое, как потребление

алкоголя или незащищенные половые связи, активность в социальных сетях и др.). Однако в описанной модели возникает проблема диагностики согласованности данных, полученных от респондентов, между собой. Рассмотрим, как подобные вопросы решаются в других предметных областях.

Цель работы – систематизировать с помощью обзора примеры и задачи, в которых применяются байесовские сети доверия, чтобы оценить, в какой степени в этих задачах учитывается диагностика согласованности исходных данных, и насколько важным является ее применение.

Статья построена следующим образом: в разделе 2 приведены некоторые сведения о байесовских сетях доверия, в разделах 3 – 8 рассмотрена проблема диагностики в медицине, экологии, экономике и других областях, и в заключении сделаны выводы.

2. Байесовские сети доверия. Байесовская сеть доверия – это ациклический направленный граф (то есть граф без направленных циклов, допускаются ненаправленные циклы), вершинам которого соответствуют случайные элементы, а ребра между вершинами соответствуют условным зависимостям между элементами [12]. Каждый случайный элемент описывается функцией распределения вероятности, в данном случае – представленной в виде тензора условных вероятностей.

Приведем формальное определение [12].

Пусть X – случайный элемент, принимающий значения из множества $\{x_1, \dots, x_n\}$. Вероятность $p(X = x_i)$ будем обозначать как $p(x_i)$. Распределение вероятности означивания X будем обозначать как $P(X)$.

Направленный граф $G(V, L)$ – это пара V, L , где V – множество вершин $\{v_1, \dots, v_n\}$, а L – множество направленных ребер $\{(u, v) | u, v \in V, u \neq v\}$. Через $pa(v)$ будем обозначать множество таких вершин u , что существует ребро (u, v) , то есть множество вершин, из которых исходит ребро, направленное в v .

Формально, байесовской сетью доверия выступает пара $\langle G, P \rangle$, где G – ациклический направленный граф, а P – множество тензоров условных вероятностей $P(X | pa(X))$. Такая пара задает совместное вероятностное распределение над всеми случайными элементами, входящими в сеть, в предположении, что X независим от всех остальных элементов сети при заданном означивании его родителей, детей и родителей детей.

Важной особенностью байесовских сетей доверия как вероятностных графических моделей является правило декомпозиции, которое

обусловлено предположением об условной независимости элементов. Правило декомпозиции выглядит следующим образом: $P(X_1, \dots, X_m) = P(X_1 | \text{pa}(X_1)) \cdot \dots \cdot P(X_m | \text{pa}(X_m))$, где $P(X_1, \dots, X_m)$ – совместное вероятностное распределение всех случайных элементов модели, $P(X_i | \text{pa}(X_i))$ – вероятностное распределение случайного элемента X_i при условии означивания случайных элементов $\text{pa}(X_i)$ – родителей вершины X_i , $i = 1, \dots, m$.

Байесовская сеть доверия может быть построена как на основе экспертных оценок, так и на основе статистических данных. Экспертная информация может использоваться для установления взаимосвязей между случайными элементами и для получения оценок условных вероятностей [51]. Правило декомпозиции позволяет использовать алгоритмы вероятностного вывода [11].

3. Медицина и здравоохранение. Байесовские сети за последние пару десятков лет стали очень популярны в биомедицине и здравоохранении благодаря возможности работы с неточными знаниями, которые участвуют в диагностике заболеваний, выборе оптимальной альтернативы и прогнозирования результатов лечения, кроме того, они предлагают очень привлекательное формальное представление неточных знаний (результат объединения статистических методов для анализа данных и инструментов искусственного интеллекта) [16]. Алгебраические байесовские сети имеют логико-вероятностную семантику, и также могут быть использованы в медицинской диагностике [13].

Основные задачи биомедицины и здравоохранения – это диагностика заболеваний, прогнозирование состояния пациентов, выбор подходящего курса лечения и обнаружение функциональных взаимодействий на клеточном уровне. Рассмотрим эти задачи немного подробнее.

3.1. Диагностика заболеваний. Постановка диагноза каждому пациенту представляет собой построение гипотезы о заболевании, от которого страдает пациент, основанное на косвенных наблюдениях и диагностических тестах. Последние, впрочем, не дают стопроцентной ясности о состоянии пациента. Чтобы избежать неправильного диагноза результаты тестов должны быть рассмотрены с учетом построенной гипотезы. Байесовские сети предлагают естественную основу для такого типа рассуждений в условиях неопределенности. Большое число систем были разработаны и разрабатываются на данной основе. Наиболее известная система – это Pathfinder [31].

Формально наиболее вероятный диагноз D^* можно определить как значение из множества возможных диагнозов D , доставляющее

максимум вероятности наличия заболевания при условии конкретного набора свидетельств E , которые включают симптомы, результаты тестов и другие признаки [41]: $D^* = \arg \max_D \Pr(D | E)$.

Системы здравоохранения очень сложны и зависят от большого количества организационных, экономических и структурных факторов. Соответствующие им инструменты должны учитывать взаимодействия между элементами, которые определяют поведение таких систем, а также быть понятными в изучении и позволять их анализ с целью улучшения производительности. Поскольку многие из факторов, влияющих на производительность систем здравоохранения неточны, байесовские сети могут сыграть важную роль в их исследовании, в качестве формальной модели для представления знаний и обработки неопределенностей [16].

Байесовские сети доверия позволяют моделировать знания с неопределенностью. Аппарат байесовской сети доверия позволяет комбинировать имеющиеся статистические данные о характеристиках здоровья пациентов в дополнение к экспертной информации, которую предоставляют врачи-специалисты [57]. Кроме того, байесовские сети доверия (по сравнению с другими моделями) позволяют моделировать возможность возникновения нескольких заболеваний [4].

В настоящее время известны случаи применения байесовских сетей доверия для диагностики заболеваний рака груди [23], печени [60], рака яичников [19], рака простаты [35], зубной боли [25] и многих других.

В [23] 25 заболеваний молочной железы (11 злокачественных и 14 доброкачественных) представлены узлами байесовской сети, такие узлы содержат априорные вероятности заболевания (в зависимости от распространенности, возраста, гормональной терапии и других факторов) или возможные выводы по маммографии. Структура модели также состоит из направленных дуг, показывающих условную зависимость элементов сети. Для построения байесовской сети были использованы уже описанные в литературе сведения о зависимостях между переменными. Построенная модель смогла различить доброкачественные и злокачественные заболевания с довольно хорошим результатом (незначительная разница с показаниями специалиста-маммолога).

В [46] описывается система для диагностики слабоумия, состоящая из двух частей: DemNet, которая помогает ЛПР диагностировать вероятность слабоумия, и PathNet, идентифицирующей возможные заболевания, вызвавшие слабоумие.

В описанных задачах используется медицинская экспертная информация, ее особенностью является частичная некорректность и не-

достаточность [29], поэтому диагностика согласованности такой информации могла бы улучшить точность используемых моделей. В [45] предложены методы диагностики согласованности между собой влияния различных заболеваний.

3.2. Прогнозирование. Задача медицинского прогнозирования – определить будущий курс и исходы процессов, связанных с заболеванием [15]. Так как будущее по своей сути неопределенно, медицинское прогнозирование представляет собой рассуждение в условиях неопределенности. Еще одной важной чертой является применение знаний об изменениях в процессах заболеваний со временем. Байесовские сети в целом могут иметь понятную темпоральную структуру (рисунок 1).

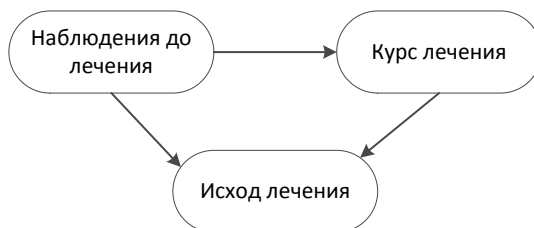


Рис. 1. Общая структура байесовских сетей при прогнозировании [41]

Формально прогноз определяется как вероятность [41]: $Pr(outcome | E, T)$, где E — доступная информация о пациенте (симптомы, результаты обследований, другие признаки), T — выбранный ход лечения. Исходом медицинского вмешательства может служить как ожидаемая продолжительность жизни пациента, так и другие аспекты качества его жизни. В качестве исхода можно рассматривать набор переменных.

Модели медицинского прогнозирования используются как на индивидуальном уровне для поддержки принятия решений о курсе лечения, так и на уровне групп пациентов для управления ресурсами в здравоохранении [58]. Прогностические байесовские сети доверия позволяют обойти трудности, с которыми сталкиваются традиционные прогностические модели. Предсказание в таких моделях обычно представляет собой одношаговый процесс и не предполагает внесения в модель дополнительных переменных, которые становятся известны с течением времени. Кроме того, в общем случае не все переменные, способные повлиять на исход, включаются в модель, например, если два предиктора зависимы, то обычно включается лишь один из них [4]. Известны примеры построения прогностических байесовских сетей в областях онкологии [30] и инфекционных заболеваний [40].

В [16] приведены репрезентационные модели, основанные на байесовских сетях доверия, применимые к конкретному случаю службы экстренной медицинской помощи. Модели построены на основе реальных данных одного испанского госпиталя с использованием алгоритмов обучения байесовских сетей. Эти модели могут быть использованы в задачах управления службами здравоохранения.

В задачах прогнозирования также крайне важно, чтобы исходные данные не имели противоречий, поэтому рекомендуется использовать средства для диагностики их согласованности. Однако в рассмотренных примерах не предложены способы проверки диагностики.

3.3. Выбор лечения. Байесовские сети доверия не предоставляют нужного уровня математической модели для выбора оптимального лечения, однако средства, используемые для решения этой проблемы часто включают байесовские сети доверия в качестве компонент систем поддержки решений для выбора оптимального лечения при данных прогнозировании [18], и в таких компонентах может проводиться диагностика согласованности исходных данных.

Также математическая модель байесовских сетей может быть расширена для включения знаний о решениях и предпочтениях. В качестве примера можно рассмотреть диаграммы влияния [41]. Как и байесовские сети – это ациклический граф. В этом графе множество вершин разделено на множество вершин моделирующих случайные величины, множество вершин, представляющих различные варианты лечения, и вершину, моделирующую предпочтения, которые надо учесть. Структура диаграммы влияния представлена на рисунке 2.

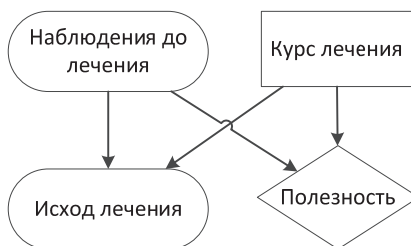


Рис. 2. Общая структура диаграммы влияния [41]

3.4.Обнаружение функциональных взаимодействий. Топологию байесовской сети доверия можно интерпретировать как представление неопределенных взаимодействий между переменными, в связи с этим в биоинформатике возрастает интерес к использованию байесовских сетей доверия для объяснения молекулярных механизмов на клеточном уровне. Например, важная исследовательская область – поиск

взаимодействий между генами на основе экспериментально полученных с помощью микрочипов данных [29].

Биологические данные часто собирают довольно долгое время, анализ временных закономерностей может показать взаимодействие переменных как функцию от времени. Байесовские сети используются для анализа таких биологических временных рядов [48].

В [34] прогнозируется взаимодействие белок – белок в геномных данных дрожжей, но в дальнейшем модель может быть усовершенствована для исследования более сложных организмов.

В [49] показано, как можно использовать байесовские сети в моделировании нейронных сетей.

Также байесовские сети доверия используют в тех областях здравоохранения, которые не относятся напрямую к лечению конкретных пациентов, например, в клинической эпидемиологии их используют для моделирования болезней, а в биоинформатике – для интерпретации данных микрочипов экспрессии генов [41].

Отметим, что в таких задачах диагностика входных данных не приведет к значительным улучшениям, поэтому в данной предметной области не рассматриваются методы проверки согласованности.

4. Экология. Также байесовские сети доверия часто используются в области экологии [22]. Байесовские сети доверия – это эффективное средство для структурирования экологических исследований. Есть два основных пути их применения [42]. Первый заключается в оценке понимания функционирования исследуемых экосистем. В таком случае исследование фокусируется на связях байесовских сетей доверия и обращается к функциональным взаимоотношениям в экосистеме или на «правилах», используемых для построения условных вероятностей для узла, и обращается к механизмам, описывающим взаимодействие факторов в определении значений переменных. Моделирование байесовских сетей доверия в основном касается таких вопросов, как «какие экологические процессы вовлечены?», «какие наиболее важны в воздействии на результаты?», «как они взаимодействуют и как прогнозирование значений переменных может оказаться полезным в экологических процессах?»

Второй путь состоит в использовании байесовских сетей доверия в оценке значений переменных, представленных узлами. Тогда исследование фокусируется на оценках, проверяющих модель и предоставляющих эмпирическую информацию, которая является количественной, полезной и относящейся к ключевым экологическим переменным. Байесовские сети доверия помогают определить сильно влияющие на результаты, но еще не очень хорошо изученные, пере-

менные, поддерживая при этом структурирование и создание средств для тестирования ответной реакции на управляющие решения.

В [28] байесовские сети доверия используются для моделирования и прогнозирования неисправностей в системе распределения питьевой воды.

В [55] – для создания модели оценки рисков угроз почвы. Чтобы определить риск уплотнения почвы, требуются или данные о поведении почвы, собрать которые требует больших затрат, или экспертные данные, которые зачастую имеют субъективный характер. В [55] предлагается модель на основе байесовских сетей доверия, объединяющая данные, полученные стандартными тестами почв, и качественные экспертные данные. В модели три результирующих вершины, дающих представление о том, насколько подвержена почва уплотнению, и двадцать девять вершин, представляющих различные характеристики почвы (текстура, содержание в почве различных веществ и др.), климатические факторы (влажность, температура, время года и др.) и другие факторы, воздействующие на почву (как используется исследуемая территория, как обрабатывается и др.), значения некоторых из этих вершин предоставляется экспертами, однако о диагностике согласованности этих данных информация отсутствует.

В [14] предложено применение куста событий и байесовской сети доверия для оценки экологической ситуации в зоне влияния потенциально химически опасных объектов и вероятности тех или иных ситуаций, связанных с его функционированием. В предложенной модели [14] содержатся вершины, отвечающие за ландшафтные условия, сезонность, время суток, метеоусловия, технологические и технических характеристики работы потенциально химически опасных объектов, условия хранения, объем и параметры загрязняющих веществ, воздействие загрязняющих веществ на персонал, население и окружающую среду, оценка экологической безопасности объекта и связи между ними (рисунок 3). О том, каким образом получены и анализируются ли исходные данные в работе не указано.

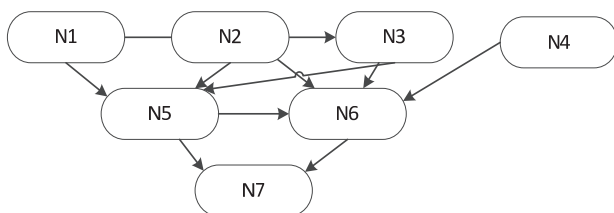


Рис. 3. Модель байесовской сети доверия: N1, N2 – ландшафтные условия, сезонность, время суток, метеоусловия; N3, N4 – характеристики работы ПХОО, условия хранения, объем и параметры загрязняющих веществ; N5, N6 – воздействие загрязняющих веществ; N7 – оценка экологической безопасности объекта [14]

Байесовские сети доверия являются хорошим инструментом для оценивания работы экосистемных служб. Так, в работе [36] они использованы для оценки работы экосистемных служб управления озерным комплексом в Бельгии.

Так как в этих задачах в основном используется смесь экспертных данных и точных технических показаний, инструменты диагностирования экспертных данных могут помочь достичь более точных результатов.

5. Функциональная безопасность. Еще одна важная сфера применения аппарата байесовских сетей доверия – функциональная безопасность. Функциональная безопасность – одна из важнейших частей многих современных технических систем. Она включает в себя процессы проектирования и оперативных мер, критически важных для систем безопасности [32]. Байесовские сети доверия используются в диагностике различных систем. В [9] описаны методы для построения экспертных систем для задач технической диагностики. В [3] предлагается методика диагностирования сложных систем, разработанная на основе апостериорного вывода в байесовских сетях доверия. В [62] описан подход к обнаружению элементов систем, вызывающих проблемы с производительностью. В [61] рассматриваются вопросы диагностики и мониторинга систем энергоснабжения. В [43] исследуется диагностика неисправностей в системах электроснабжения, в частности в бортовых системах управления. В [51] предложен метод решения проблемы диагностики газового тракта реактивных двигателей. Модель, предложенную в [38], можно использовать для диагностики производственных процессов (в качестве примера рассматривается литейная промышленность).

В [8] исследуется функционирование энергетических отраслей в условиях чрезвычайных ситуаций. Рассмотрен пример функционирования энергетической системы в ситуации похолодания: на первом этапе исследования разрабатывается онтология, описывающая взаимосвязи основных факторов в рамках угрозы похолодания; на ее основе строится БСД-модель угрозы похолодания, после этого строится событийная карта угрозы похолодания, и заключительным этапом является проведение численных расчетов на сформированных информационных моделях.

В [59] предложен графический подход для диагностики валидности предположений об условной независимости данных байесовской сети. Авторы сфокусировались на определении условной независимости и разработали графическую диагностику, которая показывает поддержанные данными предположения о независимости в структуре байесовской сети, разработали и описали метод Монте-Карло для генерации измерений неточности для согласованности данных с предположениями

об их условной независимости, то есть с помощью этого метода можно графически отобразить являются ли величины, отображающиеся вершинами байесовской сети условно независимыми или нет.

В системах, исследующих функциональную безопасность, исходные данные обычно организованы таким образом, что диагностика их согласованности не требуется.

6. Экономика и риск-анализ. Байесовские сети доверия применяются в задачах оценки экономических рисков и поддержки принятия решений в условиях неопределенности в контексте риск-менеджмента предприятий [5].

Байесовские сети доверия представляют собой удобный аппарат для анализа рисков и их количественной оценки с учетом максимального числа факторов, позволяющий кроме всего прочего наглядно отобразить структуру системы с точки зрения риск-анализа. Байесовские сети позволяют использовать как вероятности полученные аналитическим или статистическим путем, так и экспертные оценки [10].

Значение вероятности некоторого фактора в байесовской сети доверия определено изменениями в других связанных факторах. Механизм вывода основан на теореме Байеса, что делает возможным вычислить вероятность эффекта на какую-либо переменную в модели в зависимости от заданной причины. В случае с двумя связанными переменными вероятности могут быть посчитаны следующим образом:
$$P[\text{эффект}] = [P[\text{эффект} / \text{причина}] \cdot P[\text{причина}]] / P[\text{причина} / \text{эффект}],$$
где $P[\text{причина}]$ – вероятность того, что причина произойдет, $P[\text{эффект}]$ – вероятность того, что эффект возникнет, $P[\text{эффект} / \text{причина}]$ – условная вероятность эффекта при том, что причина произойдет, $P[\text{причина} / \text{эффект}]$ – условная вероятность причины, учитывая, что эффект возникнет.

Апостериорная вероятность причины, при том что эффект возникнет, может быть выведена следующим образом:
$$P[\text{причина} / \text{эффект}] = (P[\text{эффект} / \text{причина}] * P[\text{причина}]) / P[\text{эффект}].$$
Байесовские сети доверия используются для построения моделей рисков, состоящих из сценариев, основываясь на множестве известных возможных факторов риска, связанных с анализируемыми рисками. Возможные сценарии структурируются как множество взаимоисключающих и исчерпывающих элементов, которым может быть присвоено вероятностное распределение [24]. В [26] на основе байесовских сетей доверия с использованием обратной связи (feedback loop) была создана процедура для прогнозирования рисков и обнаружения их источников.

В статье [21] описан процесс построения байесовской сети доверия на основании данных, содержащих лишь маргинальные вероятности реализации риска и корреляции между различными типами рисков. В основе предложенного метода лежит решение систем уравнений, которые позволяют вычислить условные вероятности по имеющимся данным.

Вопросы анализа и моделирования операционных рисков, то есть рисков убытка в результате ошибочных внутренних процессов, действий сотрудников и систем или внешних событий, исследуются в [17]. Байесовские сети доверия позволяют корректно конструировать распределение операционных рисков в тех случаях, когда реализации рисков зависимы во времени [20] или в случаях, когда частота реализации риска и последствия реализации зависимы [7].

Информационные риски, то есть возможные угрозы информационной безопасности в любой области деятельности человека [10], также могут анализироваться с помощью байесовских сетей доверия.

В работе [37] аппарат байесовских сетей доверия использован для оценки эффективности работы персонала и организационного здоровья предприятия. В [39] показано, что при использовании этого аппарата можно не только качественно моделировать причинно-следственных связи между организационными факторами и надежности человека, но и количественно измерить эксплуатационную надежность человека, а также определить наиболее вероятные первопричины человеческих ошибок. В [44] предлагается модель для оценки факторов влияющих на бдительность рабочего, следящего и предупреждающего о приближении поездов. В [1] решалась задача оценки вероятности безошибочной работы человека-оператора, управляющего сложным технологическим объектом, тем самым определения его профессиональной надежности, как решения части общей проблемы оценки надежности человеко-машинных систем.

В [56] байесовские сети доверия используются в риск-анализе в области морского судоходства. В [6] описан метод оценки навигационной безопасности при плавании по внутренним водным путям. В [53] – при моделировании рисков природных угроз. В [2] анализируются инновационные риски, предлагается пример структуры байесовской сети для анализа среды на величину риска реализации инновационного проекта.

В задачах экономики и риск-анализа частым является интеграция различной информации, в частности статистических и полученных от экспертов данных. Данные полученные от экспертов могут содержать противоречия, поэтому следует иметь инструмент, помогающий определить насколько согласованы такие данные.

7. Социологические исследования. Применение байесовских сетей доверия в социологических исследованиях сталкивается со следующей проблемой: так как основные сведения для исследований предоставляются различными группами исследуемых, и с большой вероятностью такие данные могут быть не согласованы между собой, необходима проверка их согласованности, то есть оценка того, в какой степени исследователь может им доверять.

В [11] предложена модель для оценки интенсивности социально-значимого поведения, однако в модели не учитывается согласованность исходных данных, что может давать неточные результаты при ее исследовании. Модель представляет собой байесовскую сеть доверия с вершинами, характеризующими интенсивность поведения (*Rate*) длины интервалов между последними тремя эпизодами поведения (*t01*, *t12*, *t23*, *t_min*, *t_max*) и количество дней в исследуемом периоде (*n*) (рисунок 4).

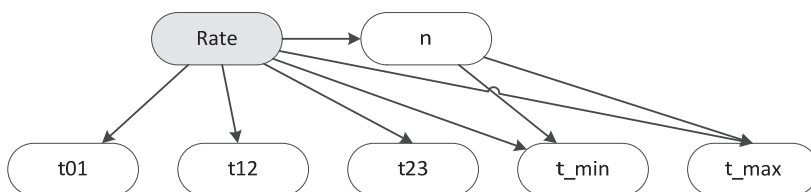


Рис. 4. Модель рискованного поведения, основанная на данных об эпизодах поведения [11]

Единственная проверка данных полученных от респондентов заключается в сравнении максимального и минимального интервалов, однако так как респонденты могут ошибаться или даже заведомо давать ложные сведения, требуется диагностика их согласованности.

Байесовские сети применяются также в анализе социальных сетей [33], в задачах, где учитываются связи между членами сетей, диагностика согласованности данных не нужна.

8. Другие области применения. Байесовские сети доверия используются и в других областях. Например, в астрономии: в [50] их используют как классификатор и средство интегрирования и исследования потоков данных, полученных с помощью специального телескопа (СТИ-II). В архитектуре: в [54] представлена модель Architecture Rationale and Element Linkage (AREL) (Рациональная Архитектура и связь элементов), которая представляет причинные взаимосвязи между архитектурными элементами и решениями в модели архитектурного дизайна, что позволяет архитекторам количественно предсказывать и диагностировать влияние изменения части требований или дизайна. В

области разработки программного обеспечения байесовские сети доверия также находят применение. Они могут быть использованы для оценки стоимости программного обеспечения на этапе его разработки [47], для обнаружения его дефектов, оставшихся необнаруженными во время тестирования [27]. Задача диагностики согласованности входных данных в этих областях не была затронута.

9. Заключение. Байесовские сети доверия – это удобный и эффективный инструмент в тех случаях, когда возникает необходимость комбинировать разные виды информации, в частности экспертную и статистическую. Экспертные данные, довольно часто используемые при сборе информации, а также в моделях принятия решений (вследствие того, что полная математическая формализация многих задач бывает невозможна из-за большой сложности), могут быть не согласованы, а в некоторых случаях и недостоверны. Поэтому актуальна задача разработки инструментария, позволяющего проводить проверку их согласованности. Первым этапом на пути к этому является исследование областей использования байесовских сетей доверия.

В данной работе приведены примеры применения аппарата байесовских сетей доверия в разных предметных областях. Большой популярностью они пользуются в частности в медицине, экологии, экономике, социологии. Указано, что в тех случаях, когда используются экспертные сведения, например в задачах медицинской диагностики и прогнозирования, риск-анализа и др., диагностика согласованности исходных данных особенно необходима. Такая диагностика позволит улучшить модель, сделать ее более надежной, оценить полученный результат. В связи с этим нужно предоставлять возможность оценивания согласованности данных в моделях, что требует разработки или адаптации соответствующих математических и программных методов.

В задачах функциональной безопасности, обнаружения функциональных взаимодействий, анализа социальных сетей данные обычно организованы таким образом, что диагностика согласованности входных данных не требуется.

В дальнейшем планируется продолжение работы над моделью, предложенной в [11], а именно решение проблемы диагностики согласованности данных в ней. Байесовские сети доверия позволяют расширять исследуемые модели новыми узлами и связями и подходят, в том числе и для оценивания согласованности данных, поэтому предлагается дополнение модели диагностическими узлами и исследование эффективности такого метода. По аналогии с этим можно будет создавать инструментарий для анализа согласованности и в более общем случае.

Литература

1. *Бабиков В.М.* Некоторые аспекты применения байесовых сетей для оценки надежности автоматизируемых человеко-машинных систем // Труды международной научно-практической конференции «Передовые информационные технологии, средства и системы автоматизации и их внедрение на российских предприятиях» АПГА-2011. Москва. 2011. С. 266–276.
2. *Белозерский А.Ю., Какатунова Т.В., Иванова И.В.* Использование аппарата нечетких байесовых сетей для оценки инновационных рисков // Транспортное дело России. 2011. № 2. С. 43–46.
3. *Дорожек И.В., Осипов Н.А.* Методика синтеза оптимальных стратегий диагностирования автоматизированных систем управления сложными техническими объектами с использованием априорной информации // Труды СПИИРАН. 2012. № 1(20). С. 165–185.
4. *Мусина В.Ф.* Байесовские сети доверия как вероятностная графическая модель для оценки медицинских рисков // Труды СПИИРАН. 2013. Вып. 24. С. 135–151.
5. *Мусина В.Ф.* Байесовские сети доверия как вероятностная графическая модель для оценки экономических рисков // Труды СПИИРАН. 2013. Вып. 25. С. 235–254.
6. *Некрасов С.Н., Прохоренков А.А.* Комбинированный метод оценки навигационной безопасности при плавании по внутренним водным путям // Вестник государственного университета морского и речного флота им. адмирала С.О. Макарова. 2011. № 1. С. 106–108.
7. *Николаева М.А., Зотова О.Ф., Агадуллина А.И.* Модели и методы управления рисками в социально-экономических системах // Управление риском. 2013. № 1(65). С. 28–34.
8. *Пяткова Е.В.* Технология комплексных исследований функционирования энергетических отраслей в условиях чрезвычайных ситуаций с применением байесовских сетей // Наука и образование: электронное научно-техническое издание. 2013. № 8. С. 293–314.
9. *Солодкий Е.М., Петроченков А.Б.* Экспертная оценка в задачах технической диагностики // Вестник Пермского национального исследовательского политехнического университета. Электротехника, информационные технологии, системы управления. 2009. № 3. С. 209–215.
10. *Студенников К.О., Лопин В.Н.* О моделировании рисков информационных систем на основе байесовских сетей // Вопросы защиты информации. 2013. № 4(102). С. 3–8.
11. *Суворова А.В.* Модели и алгоритмы анализа сверхкоротких гранулярных временных рядов на основе байесовских сетей доверия // Дисс. к. ф.-м. н. 2013.
12. *Тулупьев А. Л., Сироткин А. В., Николенко С. И.* Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах // СПб.: Изд-во С.-Петерб. ун-та. 2009. 400 с.
13. *Фильченков А.А.* Алгебраическая байесовская сеть как основа для медицинской диагностической модели // «Математическое и компьютерное моделирование в биологии и химии. Перспективы развития». Сборник трудов I Международной интернет-конференции. Казань: Из-во «Казанский университет». 2012. С. 162–166.
14. *Янников И.М., Телегина М.В., Габричидзе Т.Г.* Оценка экологической ситуации с применением методов математического моделирования // Вектор науки Тольяттинского государственного университета. 2011. № 4. С. 38–41.
15. *Abu-Hanna A., Lucas P. J F.* Prognostic Models in Medicine: AI and Statistical Approaches // Methods of Information in Medicine. 2001. vol. 40. pp. 1–5.
16. *Acid S., de Campos L.M., Fernández-Luna J.M., Rodríguez S., et al.* A comparison of learning algorithms for Bayesian networks: a case study based on data from an emer-

- gency medical service // *Artificial Intelligence in Medicine*. 2004. vol. 30. no. 3. pp. 215–232.
17. *Alexander C.* Managing operational risks with Bayesian networks // *Operational Risk: Regulation, Analysis and Management*. 2003. pp. 285–294.
 18. *Andreassen S., et al.* Using probabilistic and decisiontheoretic methods in treatment and prognosis modeling // *Artif Intell Med*. 1999. vol. 15. pp. 121–134.
 19. *Antal P., Verrelst H., Timmerman D., Moreau Y., et al.* Bayesian Networks in Ovarian Cancer Diagnosis: Potentials and Limitations // *Proceedings of the 13th IEEE Symposium on Computer-Based Medical Systems (CBMS 2000)*. 2000. pp. 103–108.
 20. *Aquaro V., Bardoscia M., Bellotti R., Consiglio A., et al.* A Bayesian Networks approach to Operational Risk // *Physica A: Statistical Mechanics and its Applications*. 2010. vol. 389. no. 8. pp. 1721–1728.
 21. *Bonafede C.E., Giudici P.* Bayesian networks for enterprise risk assessment // *Physica A: Statistical Mechanics and its Applications*. 2007. vol. 382. Issue 1. pp. 22–28.
 22. *Borsuk M.E., Stow C.A., Reckhow K.H.* A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis // *Ecological Modelling*. 2004. vol. 173. no 2. pp. 219–239.
 23. *Burnside E.S., Rubin D.L., Fine J.P., Shachter R.D., et al.* Bayesian Network to Predict Breast Cancer Risk of Mammographic Microcalcifications and Reduce Number of Benign Biopsy Results: Initial Experience // *Radiology*. 2006. vol. 240. no. 3. pp. 666–673.
 24. *Cárdenas I.C., Al-jibouri S.SH., Halman J. IM.* A Bayesian Belief Networks Approach to Risk Control in Construction Projects // 14th International Conference on Computing and Civil Engineering. Moscow. Russian Federation. The ISCCBE and Moscow State University of Civil Engineering (National Research University).
 25. *Chattopadhyay S., Davis R.M., Menezes D.D., Singh G., et al.* Application of Bayesian Classifier for the Diagnosis of Dental Pain // *J Med Syst*. 2012. vol. 36. pp. 1425–1439.
 26. *Fan C.F., Yu Y.C.* BBN-based software project risk management // *Journal of Systems and Software*. 2004. vol. 73. no 2. pp. 193–203.
 27. *Fenton N., Neil M., Marsh W., Hearty P., et al.* Predicting software defects in varying development lifecycles using Bayesian nets // *Information and Software Technology*. 2007. vol. 49. no. 1. pp. 32–43.
 28. *Francis R.A., Guikema S.D., Henneman L.* Bayesian belief networks for predicting drinking water distribution system pipe breaks // *Reliability Engineering & System Safety*. 2014. vol. 130. pp. 1–11.
 29. *Friedman N.I.R., Linal M., Nachman I., Pe'er D.* Using Bayesian network to analyze expression data // *J Comput Biol*. 2000. vol. 7. pp. 601–20.
 30. *Galan S.F., Aguado F., Diez F.J., Mira J.* NasoNet: joining Bayesian networks and time to model nasopharyngeal cancer spread // *Proceedings of the Eighth International Conference on Artificial Intelligence in Medicine in Europe (AIME 2001)*. Berlin: Springer-Verlag. 2001. LNAI 2101. pp. 207–216.
 31. *Heckerman D.E., Horvitz E.J., Nathwani B.N.* Towards normative expert systems. Part I. The Pathfinder project // *Methods of information in medicine*. 1992. vol. 31. pp. 90–105.
 32. *Hofbaur M., Sachenbacher M.* On the Role of Model-based Diagnosis in Functional Safety // *Proceedings of the 24th International Workshop on Principles of Diagnosis*. Jerusalem. 2013. pp. 65–70.
 33. *Koelle D., Pfautz J., Farry M., Cox Z., et al.* Applications of bayesian belief networks in social network analysis // *Proceedings of the 4th Bayesian Modeling Applications Workshop during the 22nd Annual Conference on Uncertainty in Artificial Intelligence*. 2006.

34. *Jansen R., Yu H., Greenbaum D., Kluger Y., et al.* A bayesian networks approach for predicting protein-protein interactions from genomic data // *Science*. 2003. vol. 302(5644). pp. 449–453.
35. *Lacave C., Diez F.J.* Knowledge Acquisition in PROSTANET – A Bayesian network for diagnosis prostate cancer // *Knowledge-Based Intelligent Information and Engineering Systems*. Lecture Notes in Computer Science. 2003. vol. 2774. pp. 1345–1350.
36. *Landuyt D., Lemmens P., D'hondt R., Broekx S., et al.* An ecosystem service approach to support integrated pond management: A case study using Bayesian belief networks — Highlighting opportunities and risks // *Journal of Environmental Management*. 2014. vol. 145. pp. 79–87.
37. *Léger A., Duval C., Farret R., Weber P., et al.* Modeling of human and organizational impacts for system risk analyses // *9th International Probabilistic Safety Assessment and Management Conference*. Hong Kong. 2008.
38. *Lewis R.W., Ransing R.S.* A semantically constrained Bayesian network for manufacturing diagnosis // *International Journal of Production Research*. 1997. vol. 35. no. 8. pp. 2171–2188.
39. *Li P.C., Chen G.H., Dai L.C., Zhang L.* A fuzzy Bayesian network approach to improve the quantification of organizational influences in HRA frameworks // *Safety Science*. 2012. vol. 50(7). pp. 1569–1583.
40. *Lucas P.J.F., De Bruijn N.C., Schurink K., Hoepelman I.M.* A probabilistic and decision-theoretic approach to the management of infectious disease at the ICU // *ArtifIntell Med*. 2000. vol. 19(3). pp. 251–279.
41. *Lucas P., van der Gaag L., Abu-Hanna A.* Bayesian networks in biomedicine and healthcare // *Artificial Intelligence in Medicine*. 2004. vol. 30. pp. 201–214.
42. *McCann R. K., Marcot B. G., Ellis R.* Bayesian belief networks: applications in ecology and natural resource management // *Canadian Journal of Forest Research*. 2006. vol. 36. no. 12. pp. 3053–3062.
43. *Mengshoel O.J., Darwiche A., Cascio K., Chavira M., et al.* Diagnosing faults in electrical power systems of spacecraft and aircraft // *Proceedings of IAAI-08*. Chicago. 2008. pp. 1699–1705.
44. *Molloy B., Balfe N., Lowe E.* Developing Bayesian Belief Networks to Support Risk-Based Decision Making in Railway Operations // *Conference Proceedings: Applied Human Factors and Ergonomics*. Poland. 2014.
45. *Nikovsky D.* Constructing Bayesian networks for medical diagnosis from incomplete and partially correct statistics // *IEEE Transactions on Knowledge and Data Engineering*. 2000. vol. 12. no. 4. pp. 509–516.
46. *Oteniya L., Cowie J., Coles R.* Diagnosis of Dementia and its Pathologies Using Bayesian Belief Networks // *Proceedings of the Eighth International Conference on Enterprise Information Systems: Databases and Information Systems Integration (ICEIS 2006)*. Cyprus. 2006.
47. *Radlinski L.* A survey of bayesian net models for software development effort prediction // *International Journal of Software Engineering and Computing*. 2010. vol. 2. no. 2. pp. 95–109.
48. *Ramoni M., Sebastiani P., Cohen P.* Bayesian clustering by dynamics // *Mach Learn* 2002. vol. 47. pp. 91–121.
49. *Rao Rajesh P.N.* Neural Models of Bayesian Belief Propagation // *The Bayesian Brain: Probabilistic Approaches to Neural Coding*. Cambridge. MIT Press. 2007. 239 p.
50. *Ritthaler M., Luger G., Young R.* Bayesian Belief Networks for Astronomical Object Recognition and Classification in CTI-II // *Astronomical Data Analysis Software and Systems XVI*. ASP Conference Series. 2007. vol. 376. pp. 413–416.
51. *Romessis C., Mathioudakis K.* Bayesian network approach for gas path fault diagnosis // *Journal of engineering for gas turbines and power*. 2006. vol. 128(1). pp. 64–72.

52. Spiegelhalter D.J., Dawid A.P., Lauritzen S.L., Cowell R.G. Bayesian Analysis in Expert Systems // *Statistical Science*. 1993. vol. 8. no. 3. pp. 219–247. URL: <http://www.jstor.org/stable/2245959> (дата доступа: 02.11.2015).
53. Straub D. Natural hazards risk assessment using Bayesian networks // *Safety and Reliability of Engineering Systems and Structures*. 2005. pp. 2535–2542.
54. Tang A., Nicholson A., Jin Y., Han J. Using Bayesian belief networks for change impact analysis in architecture design // *Journal of Systems and Software*. 2007. vol. 80. no. 1. pp. 127–148.
55. Trolborg M., Aalders I., Towers W., Hallett P.D., et al. Application of Bayesian Belief Networks to quantify and map areas at risk to soil threats: Using soil compaction as an example // *Soil and Tillage Research*. 2013. vol. 132. pp. 56–68.
56. Trucco P., Cagno E., Ruggeri F., Grande O. A Bayesian Belief Network modeling of organizational factors in risk analysis: a case study in maritime transportation // *Reliability Engineering and System Safety*. 2008. vol. 93. pp. 823–834.
57. Twardy C.R., Nicholson A.E., Korb K.B., McNeil J. Epidemiological data mining of cardiovascular Bayesian networks // *Electronic Journal of Health Informatics*. 2006. vol. 1(1). pp. 3.
58. Verduijn M., et al. Prognostic Bayesian networks I: rationale, learning procedure, and clinical use // *Journal of Biomedical Informatics*. 2007. vol. 40(6). pp. 609–618.
59. Walsh S., Whitney P. A Graphical Approach to Diagnosing the Validity of the Conditional Independence Assumptions of a Bayesian Network Given Data // *Journal of Computational and Graphical Statistics*. 2012. vol. 21. no. 4. pp. 961–978.
60. Wasyluk H., Onisko A., Druzdel M.J. Support of diagnosis of liver disorders based on a causal Bayesian network model // *Medical Science Monitor*. 2001. vol. 7. no. 1. pp. 327–332.
61. Yongli Z., Limin H., Jinling L. Bayesian network-based approach for power system fault diagnosis // *IEEE Transactions on Power Delivery*. 2006. vol. 21. pp. 634–639.
62. Zhang R., Moyle S., McKeever S., Bivens A. Performance problem localization in selfhealing, service-oriented systems using bayesian networks // *In Proceedings of the 2007 ACM symposium on Applied computing (SAC '07)*. 2007.

References

1. Babikov V.M. [Some aspects of the application of Bayesian networks to assess the reliability of the automated man-machine systems]. *Trudy mezhdunarodnoi nauchno-prakticheskoi konferentsii «Peredovye informatsionnye tekhnologii, sredstva i sistemy avtomatizatsii i ikh vnedrenie na rossiiskikh predpriiatiakh» AITA-2011* [Proceedings of the International scientific and practical conference "Advanced information technologies, tools and automation systems and their implementation at Russian enterprises» AITA-2011]. Moscow. 2011. pp. 266–276. (In Russ.).
2. Belozerskii A.Iu., Kakatunova T.V., Ivanova I.V. [The use of fuzzy bayesian networks to assess the risks of innovation]. *Transportnoe delo Rossii – Transport business in Russia*. 2011. vol. 2. pp. 43–46. (In Russ.).
3. Dorozhko I.V., Osipov N.A. [Technique of synthesis of optimal strategies for diagnostics of the automated control systems of complex technical objects with the use of aprioristic information]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2012. vol. 1(20). pp. 165–185. (In Russ.).
4. Musina V.F. [Bayesian belief networks as probabilistic graphical model for medical risk assessment]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2013. vol. 24. pp. 135–151. (In Russ.).
5. Musina V.F. [Bayesian belief networks as probabilistic graphical model for economical risk assessment]. *Trudy SPIIRAN – SPIIRAS Proceedings*. 2013. vol. 25. pp. 235–254. (In Russ.).

6. Nekrasov S.N., Prohorenkov A.A. [Combined simulation while estimation of inland waterway navigation safety]. *Vestnik gosudarstvennogo universiteta morskogo i rechnogo flota im. admirala S.O. Makarova – Vestnik gosudarstvennogo universiteta morskogo i rechnogo flota imeni admirala S.O. Makarova*. 2011. vol. 1. pp. 106–108. (In Russ.).
7. Nicolaeva M.A., Zotova O.F., Agadullina A.I. [Models and methods for risk management in the social and economic systems]. *Upravlenie riskom – Risk Management*. 2013. vol. 1(65). pp. 28–34. (In Russ.).
8. Piatkova E.V. [Technology of complex research of energy sector functioning in emergency situations with the use of Bayesian networks]. *Nauka i obrazovanie: elektronnoe nauchno-tehnicheskoe izdanie – Science and Education: electronic scientific and technical journal*. 2013. vol. 8. pp. 293–314. (In Russ.).
9. Solodkii E.M., Petrochenkov A.B. [Peer review in the problems of technical diagnostics]. *Vestnik Permskogo natsionalnogo issledovatel'skogo politekhnicheskogo universiteta. Elektrotehnika, informatsionnye tekhnologii, sistemy upravleniia – Vestnik of Perm National Research Polytechnic University. Electrical engineering, information technology, control systems*. 2009. vol. 3. pp. 209–215. (In Russ.).
10. Studennikov K.O., Lopin V.N. [About information systems risk modeling based on bayesian networks]. *Voprosy zashchity informatsii – Information Security issues*. 2013. vol. 4(102). pp. 3–8. (In Russ.).
11. Suvorova A.V. *Modeli i algoritmy analiza sverkhkorotkikh granuliarnykh vremennykh riadov na osnove baiesovskikh setei doveriia* [Models and algorithms for the analysis of granular ultrashort time series based on Bayesian belief networks]. Ph.D. thesis. 2013. (In Russ.).
12. Tulupev A. L., Sirotkin A. V., Nikolenko S. I. *Baiesovskie seti doveriia: logikoveroiatnostnyi vyvod v atclicheskikh napravlennykh grafakh* [Bayesian belief networks: the logic-probabilistic inference in the acyclic directed graph]. SPb.: Izd-vo S.-Peterb. un-ta, 2009. 400 p. (In Russ.).
13. Filchenkov A.A. [Algebraic Bayesian network as the basis for medical diagnostic model]. «*Matematicheskoe i kompiuternoe modelirovanie v biologii i himii. Perspektivy razvitiia*». *Sbornik trudov I Mezhdunarodnoi internet-konferentsii* ["Mathematical and computer modeling in biology and chemistry. Development prospects". Proceedings of the I International Internet Conference]. Kazan. «KSU ». 2012. pp. 162–166. (In Russ.).
14. Iannikov I.M., Telegina M.V., Gabrichidze T.G. [Estimation of the ecological situation with application of methods of mathematical modelling]. *Vektor nauki Toliattinskogo gosudarstvennogo universiteta – Vector of sciences*. Togliatti State University. 2011. no. 4. pp. 38–41. (In Russ.).
15. Abu-Hanna A., Lucas P. J.F. Prognostic Models in Medicine: AI and Statistical Approaches. *Methods of Information in Medicine*. 2001. vol. 40. pp. 1–5.
16. Acid S., de Campos L.M., Fernández-Luna J.M., Rodríguez S., et al. A comparison of learning algorithms for Bayesian networks: a case study based on data from an emergency medical service. *Artificial Intelligence in Medicine*. 2004. vol. 30. no. 3. pp. 215–232.
17. Alexander C. Managing operational risks with Bayesian networks. *Operational Risk: Regulation, Analysis and Management*. 2003. pp. 285–294.
18. Andreassen S., et al. Using probabilistic and decisiontheoretic methods in treatment and prognosis modeling. *Artif Intell Med*. 1999. vol. 15. pp. 121–134.
19. Antal P., Verrelst H., Timmerman D., Moreau Y., et al. Bayesian Networks in Ovarian Cancer Diagnosis: Potentials and Limitations. Proceedings of the 13th IEEE Symposium on Computer-Based Medical Systems (CBMS 2000). 2000. pp. 103–108.

20. Aquaro V., Bardoscia M., Bellotti R., Consiglio A., et al. A Bayesian Networks approach to Operational Risk. *Physica A: Statistical Mechanics and its Applications*. 2010. vol. 389. no. 8. pp. 1721–1728.
21. Bonafede C.E., Giudici P. Bayesian networks for enterprise risk assessment. *Physica A: Statistical Mechanics and its Applications*. 2007. vol. 382. Issue 1. pp. 22–28.
22. Borsuk M.E., Stow C.A., Reckhow K.H. A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. *Ecological Modelling*. 2004. vol. 173. no 2. pp. 219–239.
23. Burnside E.S., Rubin D.L., Fine J.P., Shachter R.D., et al. Bayesian Network to Predict Breast Cancer Risk of Mammographic Microcalcifications and Reduce Number of Benign Biopsy Results: Initial Experience. *Radiology*. 2006. vol. 240. no. 3. pp. 666–673.
24. Cárdenas I.C., Al-jibouri S.SH., Halman J. IM. A Bayesian Belief Networks Approach to Risk Control in Construction Projects. 14th International Conference on Computing and Civil Engineering. Moscow. Russian Federation. The ISCCBE and Moscow State University of Civil Engineering (National Research University).
25. Chattopadhyay S., Davis R.M., Menezes D.D., Singh G., et al. Application of Bayesian Classifier for the Diagnosis of Dental Pain. *J Med Syst*. 2012. vol. 36. pp. 1425–1439.
26. Fan C.F., Yu Y.C. BBN-based software project risk management. *Journal of Systems and Software*. 2004. vol. 73. no 2. pp. 193–203.
27. Fenton N., Neil M., Marsh W., Hearty P., et al. Predicting software defects in varying development lifecycles using Bayesian nets. *Information and Software Technology*. 2007. vol. 49. no. 1. pp. 32–43.
28. Francis R.A., Guikema S.D., Henneman L. Bayesian belief networks for predicting drinking water distribution system pipe breaks. *Reliability Engineering & System Safety*. 2014. vol. 130. pp. 1–11.
29. Friedman N.I.R., Linial M., Nachman I., Pe'er D. Using Bayesian network to analyze expression data. *J Comput Biol*. 2000. vol. 7. pp. 601–20.
30. Galan S.F., Aguado F., Diez F.J., Mira J. Nasonet: joining Bayesian networks and time to model nasopharyngeal cancer spread. Proceedings of the Eighth International Conference on Artificial Intelligence in Medicine in Europe (AIME 2001). Berlin: Springer-Verlag. 2001. LNAI 2101. pp. 207–216.
31. Heckerman D.E., Horvitz E.J., Nathwani B.N. Towards normative expert systems. Part I. The Pathfinder project. *Methods of information in medicine*. 1992. vol. 31. pp. 90–105.
32. Hofbauer M., Sachenbacher M. On the Role of Model-based Diagnosis in Functional Safety. Proceedings of the 24th International Workshop on Principles of Diagnosis. Jerusalem. 2013. pp. 65–70.
33. Koelle D., Pfautz J., Farry M., Cox Z., et al. Applications of bayesian belief networks in social network analysis. Proceedings of the 4th Bayesian Modeling Applications Workshop during the 22nd Annual Conference on Uncertainty in Artificial Intelligence. 2006.
34. Jansen R., Yu H., Greenbaum D., Kluger Y., et al. A bayesian networks approach for predicting protein-protein interactions from genomic data. *Science*. 2003. vol. 302(5644). pp. 449–453.
35. Lacave C., Diez F.J. Knowledge Acquisition in PROSTANET – A Bayesian network for diagnosis prostate cancer. *Knowledge-Based Intelligent Information and Engineering Systems*. 2003. LNCS 2774. pp. 1345–1350.
36. Landuyt D., Lemmens P., D'hondt R., Broeckx S., et al. An ecosystem service approach to support integrated pond management: A case study using Bayesian belief networks — Highlighting opportunities and risks. *Journal of Environmental Management*. 2014. vol. 145. pp. 79–87.

37. Léger A., Duval C., Farret R., Weber P., et al. Modeling of human and organizational impacts for system risk analyses. 9th International Probabilistic Safety Assessment and Management Conference. Hong Kong. 2008.
38. Lewis R.W., Ransing R.S. A semantically constrained Bayesian network for manufacturing diagnosis. *International Journal of Production Research*. 1997. vol. 35. no. 8. pp. 2171–2188.
39. Li P.C., Chen G.H., Dai L.C., Zhang L. A fuzzy Bayesian network approach to improve the quantification of organizational influences in HRA frameworks. *Safety Science*. 2012. vol. 50(7). pp. 1569–1583.
40. Lucas P.J.F., De Bruijn N.C., Schurink K., Hoepelman I.M. A probabilistic and decision-theoretic approach to the management of infectious disease at the ICU. *ArtifIntell Med*. 2000. vol. 19(3). pp. 251–279.
41. Lucas P., van der Gaag L., Abu-Hanna A. Bayesian networks in biomedicine and healthcare. *Artificial Intelligence in Medicine*. 2004. vol. 30. pp. 201–214.
42. McCann R. K., Marcot B. G., Ellis R. Bayesian belief networks: applications in ecology and natural resource management. *Canadian Journal of Forest Research*. 2006. vol. 36. no. 12. pp. 3053–3062.
43. Mengshoel O.J., Darwiche A., Cascio K., Chavira M., et al. Diagnosing faults in electrical power systems of spacecraft and aircraft. Proceedings of IAAI-08. Chicago. 2008. pp. 1699–1705.
44. Molloy B., Balfe N., Lowe E. Developing Bayesian Belief Networks to Support Risk-Based Decision Making in Railway Operations. Conference Proceedings: Applied Human Factors and Ergonomics. Poland. 2014.
45. Nikovsky D. Constructing Bayesian networks for medical diagnosis from incomplete and partially correct statistics. *IEEE Transactions on Knowledge and Data Engineering*. 2000. vol. 12. no. 4. pp. 509–516.
46. Oteniya L., Cowie J., Coles R. Diagnosis of Dementia and its Pathologies Using Bayesian Belief Networks. Proceedings of the Eighth International Conference on Enterprise Information Systems: Databases and Information Systems Integration (ICEIS 2006). Cyprus. 2006.
47. Radlinski L. A survey of bayesian net models for software development effort prediction. *International Journal of Software Engineering and Computing*. 2010. vol. 2. no. 2. pp. 95–109.
48. Ramoni M., Sebastiani P., Cohen P. Bayesian clustering by dynamics. *Mach Learn*. 2002. vol. 47. pp. 91–121.
49. Rao R.P.N. Neural Models of Bayesian Belief Propagation. *The Bayesian Brain: Probabilistic Approaches to Neural Coding*. Cambridge. MIT Press. 2007. 239 p.
50. Ritthaler M., Luger G., Young R. Bayesian Belief Networks for Astronomical Object Recognition and Classification in CTI-II. *Astronomical Data Analysis Software and Systems XVI*. ASP Conference Series. 2007. vol. 376. pp. 413–416.
51. Romessis C., Mathioudakis K. Bayesian network approach for gas path fault diagnosis. *Journal of engineering for gas turbines and power*. 2006. vol. 128(1). pp. 64–72.
52. Spiegelhalter D.J., Dawid A.P., Lauritzen S.L., Cowell R.G. Bayesian Analysis in Expert Systems. *Statistical Science*. 1993. vol. 8. no. 3. pp. 219–247. Available at: <http://www.jstor.org/stable/2245959> (accessed: 02.11.2015).
53. Straub D. Natural hazards risk assessment using Bayesian networks. *Safety and Reliability of Engineering Systems and Structures*. 2005. pp. 2535–2542.
54. Tang A., Nicholson A., Jin Y., Han J. Using Bayesian belief networks for change impact analysis in architecture design. *Journal of Systems and Software*. 2007. vol. 80. no. 1. pp. 127–148.

55. Troldborg M., Aalders I., Towers W., Hallett P.D., et al. Application of Bayesian Belief Networks to quantify and map areas at risk to soil threats: Using soil compaction as an example. *Soil and Tillage Research*. 2013. vol. 132. pp. 56–68.
56. Trucco P., Cagno E., Ruggeri F., Grande O. A Bayesian Belief Network modeling of organizational factors in risk analysis: a case study in maritime transportation. *Reliability Engineering and System Safety*. 2008. vol. 93. pp. 823–834.
57. Twardy C.R., Nicholson A.E., Korb K.B., McNeil J. Epidemiological data mining of cardiovascular Bayesian networks. *Electronic Journal of Health Informatics*. 2006. vol. 1(1). pp. 3.
58. Verduijn M., et al. Prognostic Bayesian networks I: rationale, learning procedure, and clinical use. *Journal of Biomedical Informatics*. 2007. vol. 40(6). pp. 609–618.
59. Walsh S., Whitney P. A Graphical Approach to Diagnosing the Validity of the Conditional Independence Assumptions of a Bayesian Network Given Data. *Journal of Computational and Graphical Statistics*. 2012. vol. 21. no. 4. pp. 961–978.
60. Wasyluk H., Onisko A., Druzdel M.J. Support of diagnosis of liver disorders based on a causal Bayesian network model. *Medical Science Monitor*. 2001. vol. 7. no. 1. pp. 327–332.
61. Yongli Z., Limin H., Jinling L. Bayesian network-based approach for power system fault diagnosis. *IEEE Transactions on Power Delivery*. 2006. vol. 21. pp. 634–639.
62. Zhang R., Moyle S., McKeever S., Bivens A. Performance problem localization in selfhealing, service-oriented systems using Bayesian networks. In Proceedings of the 2007 ACM symposium on Applied computing (SAC '07). 2007.

Торопова Александра Витальевна — младший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики, Федеральное государственное бюджетное учреждение науки Санкт-Петербургского института информатики и автоматизации Российской академии наук (СПИИРАН), аспирант математико-механического факультета, Санкт-Петербургский государственный университет (СПбГУ). Область научных интересов: байесовские сети доверия, социокomпьютинг. Число научных публикаций — 15. alexandra.toropova@gmail.com; 14-я линия В.О., д. 39, Санкт-Петербург, 199178; p.т.: +7(812)328-3337.

Toropova Aleksandra Vital'evna — junior researcher of theoretical and interdisciplinary computer science laboratory, St. Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences (SPIIRAS), Ph.D. student, Saint Petersburg State University (SPbSU). Research interests: Bayesian belief networks, social computing. The number of publications — 15. alexandra.toropova@gmail.com; 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone: +7(812)328-3337.

Поддержка исследований. Работа выполнена при частичной финансовой поддержке РФФИ (проекты 14-01-00580-а, 15-01-09001-а).

Acknowledgements. This research is supported by RFBR (grants 14-01-00580-a, 15-01-09001-a).

РЕФЕРАТ

***Торопова А.В.* Подходы к диагностике согласованности данных в байесовских сетях доверия.**

Байесовские сети доверия предоставляют возможность объединения нескольких видов информации, например полученной от экспертов или статистически, позволяют работать с неполной или неточной информацией, обладают наглядностью и другими полезными свойствами. Благодаря этому они стали популярным и весьма эффективным средством. Однако во многих областях исследования исходные данные, которые предоставляются экспертами, могут быть не согласованы, и поэтому в некоторых задачах следует использовать инструменты для проверки их согласованности.

Цель работы – систематизировать с помощью обзора примеры и задачи, в которых применяются байесовские сети доверия, чтобы оценить в какой степени в этих задачах учитывается диагностика согласованности исходных данных, и насколько важным является ее применение.

В работе рассмотрены примеры применения аппарата байесовских сетей доверия в различных областях.

Задачи медицины и здравоохранения подразделяются на четыре основных направления: диагностика, прогнозирование, лечение заболеваний и обнаружению функциональных взаимодействий между различными объектами. Диагностика согласованности исходных данных может потребоваться в первых трех, так как значительная их часть предоставляется экспертами, которые могут быть субъективны или ошибаться.

В задачах экологии в основном используется смесь экспертных данных и точных технических показаний, инструменты диагностирования экспертных данных могут помочь достичь более точных результатов.

В области функциональной безопасности систем исходные данные обычно организованы таким образом, что диагностика их согласованности не требуется.

В экономике и риск-анализе частым является интеграция различной информации, в частности статистических и полученных от экспертов данных. Данные полученные от экспертов могут содержать противоречия, поэтому следует иметь инструмент, позволяющий определить насколько согласованы такие данные.

В области социологических исследований так как часто исследования проводятся на основе опросов различных групп населения, данные предоставляемые ими могут оказаться несогласованными и требуют проверки. В задачах анализа социальных сетей, где учитываются связи между членами сетей, диагностика согласованности данных не требуется.

Таким образом, в работе показана необходимость создания и использования инструментов диагностики согласованности исходных данных в некоторых областях исследований.

SUMMARY

Toropova A.V. **Approaches to the Data Coherence Diagnosis in Bayesian Belief Network Models.**

Bayesian belief networks provide the ability to combine different types of information, e.g. statistical or expert data, allow working with incomplete or inaccurate information; they have clarity and other useful properties. Due to this, Bayesian belief networks have become a popular and highly effective tool in many fields of research. However, in many research areas data provided by the experts can be incoherent, and so in some tasks one should use tools to verify their coherence.

The purpose of this work is to systematize problems and examples using Bayesian belief networks by reviewing and to assess their use of data coherence diagnosis and its importance. The paper discusses examples of application of the Bayesian belief networks in different research areas.

Health-care and medicine problems are divided into four base areas: diagnosis, prognosis, treatment selection and discovering functional interactions between different objects. Initial data coherence diagnosis can be required in the first three, as much of this data is provided by experts who can be subjective or wrong.

The ecology problems contain a mix of an expert data and precise technical indications; expert data coherence diagnostic tools can help achieve more accurate results.

In the functional safety systems, data is mainly organized in such a way that the coherence diagnosis is not required.

In economics and risk analysis, it is common to integrate different kinds of information, in particular data that is statistical or received from the experts. The data obtained from the experts may contain contradictions, so tools for discovering incoherent data should be used.

In sociology, researchers usually retrieve initial data from the different population groups' surveys, so they can provide incoherent or even intentionally wrong data. Hence, these data is requiring a coherence diagnosis. The SNA problems consider links between network members, therefore coherence diagnosis is not necessary.

Thus, the work shows the necessity to create and use the coherence diagnostic tools of input data in different research areas.