Сети Байеса как инструмент моделирования сложных природных систем с высоким уровнем неопределенности

В. Н. Таран
Крымский федеральный университет имени В.И. Вернадского
г. Симферополь, Республика Крым, Россия
victoriya_yalta@ukr.net

***Аннотация*. На современном этапе развития науки ученые занимаются исследованием больших сложных систем, которые плохо структурируются, имеют большой уровень неопределенности и непредсказуемости протекания тех или иных процессов, осложняются присутствием в данных системах факторов риска появления непредвиденных ситуаций, что позволяет их отнести к классу сложных задач анализа динамических процессов случайной природы. Классические хорошо изученные методы моделирования и прогнозирования не дают достоверных результатов, что заставляет искать новые способы анализа и исследования сложных систем. Примером высокорезультативных методов прогнозирования являются сети Байеса, основанные на экспертных оценках специалистов и на априорных и апостериорных данных наблюдений. Сети доверия Байеса позволяют делать вероятностный прогноз как результирующих показателей, так и показывает возможные альтернативы при воздействии некоторых управляющих факторов. Также позволяют автоматически строить цепочки зависимостей между факторами разных уровней, которые опираются на результаты множественных наблюдений. Сети Байеса используются в различных областях исследований, в том числе: в медицине, метеорологии, исследовании природных катастроф.**

Ключевые слова: сложные системы; сети Байеса; моделирование; неопределенность; прогнозирование уровня катастроф

I. Введение

Катастрофы, в том числе, природные катастрофы причиняют непоправимый вред не только человеку, но и всей Земле, уничтожая природу и ареалы обитания разнообразных обитателей фауны.

Если мы сможем предвидеть активизацию сложных природных процессов, имеющих катастрофические последствия, таких как: цунами, землетрясения, оползни, обвалы, природные пожары, ураганы, обработка прибрежной зоны моря, снежные лавины, экстремальные температуры воздуха [1], то появится возможность минимизировать эти последствия путем проведения специальных защитных мер, что дает возможность защитить результаты деятельности человека, а также его жизнь.

Человечество дано ведет наблюдения за природными процессами, фиксирует собранные данные, пытается из систематизировать и на их основе проводить прогностические расчеты. Построены модели для различных процессов, предложены способы их применения и рассчитаны возможные ограничения и ошибки. Однако добиться прогноза, удовлетворяющего хотя бы на 80%, практически не удается. Чаще всего это связано с высоким уровнем неопределенности, который носят природные процессы.

Полученные из различных компонентов системы данные часто неоднородны: разные параметры могут быть измерены с разной точностью, а также они могут представлять разные пространственные и временные масштабы [2]. Мы хотим использовать современные подходы к моделированию и прогнозированию сложных природных катастрофических процессов, которые, в отличие от классических методов, используют качественные и количественные оценки, позволяют учитывать неопределенность и экспертные оценки, а также дают возможность уточнять риски и их уровень. Такому подходу соответствуют байесовские сети доверия.

Целью данного исследования является анализ характеристик процессов протекания сложных природных процессов с высоким уровнем неопределенности, ведущих к катастрофам и нарушающих экосистему земного шара, а также моделирование их протекания с помощью сети доверия Байеса.

Теория байесовских сетей в последнее время стала очень популярной в решении различных прикладных задач в различных областях науки и производства [3], а также в решении проблем, связанных с неопределенностью [4].

II. Анализ характеристик протекания сложных природных процессов

A. Выделение факторов, которые влияют на активизацию катастрофических процессов

Лучшее понимание протекания природных процессов и причин их активизации может помочь обществам в планировании стратегий управления, которые, скорее всего, приведут к достижению и поддержанию хорошего состояния экосистемы [2]

На наш взгляд наиболее существенным фактором, который способствует активизации природных процессов, носящих катастрофический характер, является солнечная активность и ее 11-летние циклы. Можно наблюдать при повышении общего фона солнечной радиации и солнечной активности резкое повышение не только катастроф природного характера, но и общественных волнений, которые могут заканчиваться человеческими жертвами. Именно этот фактор многие ученые рассматривают как ключевой в исследовании процессов природного характера, а также указывается на его неоспоримое влияние [1,5,6,7].

Второй фактор можно выбрать как имеющиеся данные по недавно начавшейся или продолжающейся активизации рассматриваемого природного процесса на данной территории, а также в непосредственной близости от рассматриваемого ареала моделирования и прогнозирования. Таким образом, рассматривается лаговый фактор, так называемый фактор запаздывания или авторегрессии.

Третьим фактором мы выберем общее повышение активности сопутствующих природных процессов. Например, увеличение или появление сейсмической активности, выпадение аномального количества осадков или наоборот их продолжительное отсутствие, которое приводит к засухе и уничтожению растительного покрова за счет его высыхания или пожароопасной ситуации. Также это может быть наблюдение в ближайших районах ураганов, бурь или смерчей. В общем случае этот фактор назовем метеорологический фактор или погодный.

Четвертым фактором выберем некоторый фактор регулирования (или так называемого управления) в виде вложенных средств в защитные сооружения, т.е. это могут быть укрепительные мероприятия [1, 5, 6, 7], или мероприятия по оповещению и предупреждению населения, или долгосрочное строительство, направленное на защиту сооружений, дорог или конструкций, или посадка лесных массивов для укрепления склонов, защиты надземного покрова от пересыхания и вымывания поверхностного слоя почвы.

Значения каждого из факторов можно определить, как «Да» или «Нет», т.е. наблюдается резкое повышение солнечной активности: «Да» или «Нет», началась активизация природных процессов: «Да» или «Нет», и.п.

Для фактора «Вложенные средства» определим следующие значения: «Мало», «Средне», «В полном объеме»

B. Определение прогнозируемых показателей

Рассмотрим показатели, для которых можно и нужно получить прогноз. Это может быть несколько результирующих показателей.

Во-первых, нас будет интересовать, следует ли ожидать активизацию катастрофических природных процессов и на сколько разрушительны будут последствия этих процессов: «Разрушений мало, и они несерьезные», «Разрушений много, но несерьезные», «Разрушений мало, но последствия катастрофические», «Разрушений много, и последствия катастрофические». Этот показатель назовем «Природные риски».

Во-вторых, нас интересует, в течение какого периода времени могут произойти катастрофические природные процессы: «Неделя», «Месяц», «Не произойдут». Назовем этот показатель «Временные риски».

В-третьих, немаловажный вопрос: а сколько денежных средств потребуется на покрытие разрушений и устранение последствий катастрофических природных процессов. Этот показатель назовем «Денежные затраты» и он будет принимать значения: «Малые», «Средние», «Огромные».

III. Построение сети доверия Байеса

A. Понятие сети доверия Байеса и ее преимущества использования

Моделирование и прогнозирование сложных (опасных) природных процессов классическими методами не дает достоверных данных, на которые можно было бы опираться при формировании возможных сценариев развития рассматриваемых процессов и принятия управленческих решений [6]. Поэтому рассмотрим сеть доверия Байеса.

Этот новый подход может лучше выявить взаимосвязи между причинами, а также может обновить вероятности по мере развития условий окружающей среды [8], построить наиболее точные и корректные суждения с максимальным учетом доступной информации [4].

Сети Байеса – это мощный и эффективный математический инструмент исследования и воспроизведения реальной картины процессов, который целесообразно применять для решения задач вероятностного прогнозирования и оценки рисков [9, 10].

Модель, основанная на байесовской сети убеждений, позволяет комбинировать как статистические данные, так и экспертные предположения о природе поведения и взаимосвязях между элементами [11, 12]. Байесовские сети являются одним из представлений баз знаний с неопределенностью [13]

Сети Байеса позволяют за счет выявления структурных отношений между атрибутами и последующей расстановки приоритетов в процессе обучения и самообучения решать задачи максимизации и оптимизации с наивысшей оценкой [14]. Авторы [15] использовали сети Байеса для визуализации и пространственной оптимизации сценариев развития выбранных экосистем.

Интегрированные Байесовские сети (IBN) становятся полезными инструментами в анализе рисков, управлении рисками и науке принятия решений для планирования ресурсов и управления окружающей средой [16]. В [17] говорится, что байесовская сеть в основном используется в качестве статистической схемы для вероятностного прогнозирования, которая может представлять причинно-следственные связи между переменными.

Итак, Байесовская сеть представляет собой вероятностную графическую модель и задается парой (G, P). <G> – это ациклический граф, вершины которого представляют переменные, например, случайные, и имеют несколько состояний. Ориентированные ребра графа показывают отношение прямой зависимости одной переменной от другой. Каждому ребру соответствует таблица P условных вероятностей перехода из состояний вершины верхнего уровня в состояние вершины нижнего уровня. Для вершин верхнего уровня задаются безусловные вероятности, полученные по результатам наблюдений [1, 5–7, 9, 10]. Этот метод способен гибко настраиваться на структурные изменения во внешней среде моделируемого процесса [18].

B. Построение структуры сети доверия Байеса и моделирование

Мы предложили структуру сети доверия Байеса, которая состоит из двух уровней. На первом уровне расположены факторы (их четыре), от которых зависят результирующие показатели. На втором уровне мы расположим три результирующих показателя: «Риски Активизации», «Риски Разрушений» и «Риски Времени». На третьем уровне расположим вершину «Денежные затраты». Соединим вершины верхнего уровня с вершинами среднего, а вершины среднего с вершиной нижнего с помощью стрелок-отношений, показывающих зависимость между ними (рис. 1). Для моделирования используем программу Netica.

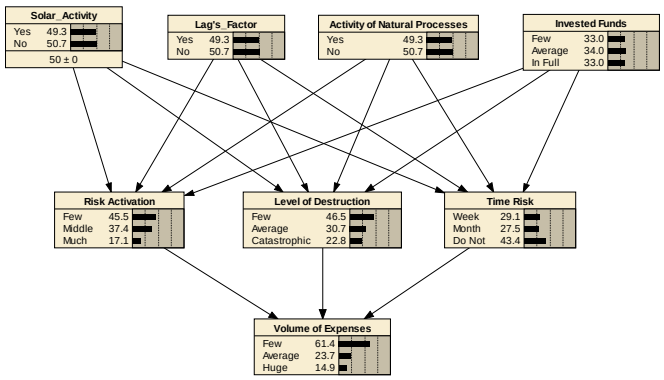
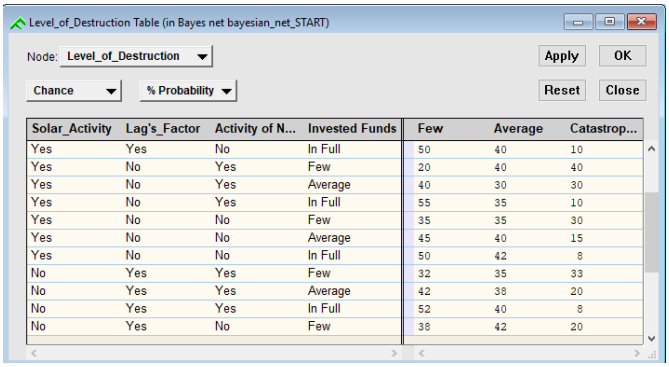


Рис. 1. Сеть доверия Байеса для моделирования и прогнозирования сложных природных процессов

При этом для определения зависимости требуется заполнить условные вероятности перехода между вершинами как причинно-следственные связи. Заполняются эти таблицы с помощью экспертных оценок специалистов соответствующей области (рис. 2). Таким образом, байесовская сеть обучается на условных и безусловных вероятностях или на экспертных оценках.

Для моделирования выберем начальные значения верхних вершин-факторов. Пусть, например, Солнечная активность увеличилась, а остальные факторы – нет, при этом вложения денежных средств отсутствовала или были на низком уровне (рис. 3)



Solar_Activity	Lag's_Factor	Activity of N...	Invested Funds	Few	Average	Catastrop...
Yes	Yes	No	In Full	50	40	10
Yes	No	Yes	Few	20	40	40
Yes	No	Yes	Average	40	30	30
Yes	No	Yes	In Full	55	35	10
Yes	No	No	Few	35	35	30
Yes	No	No	Average	45	40	15
Yes	No	No	In Full	50	42	8
No	Yes	Yes	Few	32	35	33
No	Yes	Yes	Average	42	38	20
No	Yes	Yes	In Full	52	40	8
No	Yes	No	Few	38	42	20

Рис. 2. Обучение сети доверия Байеса

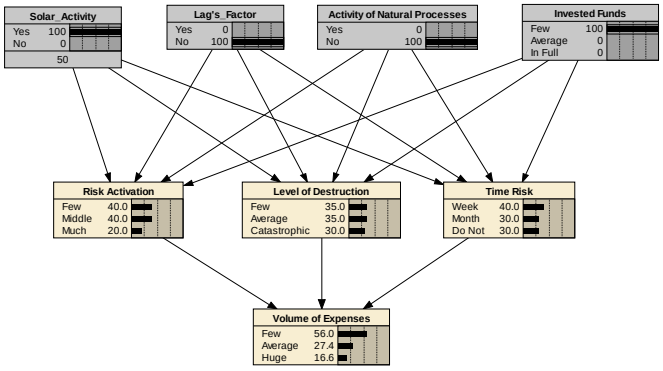


Рис. 3. Моделирование с помощью сети доверия Байеса

Полученная модель позволяет получить прогноз, при каких начальных вложениях общая сумма затрат окажется минимальной (рис. 4). Иначе говоря, моделирование можно проводить традиционным способом «сверху-вниз» и, наоборот, «снизу-вверх», опираясь на априорные характеристики модели.

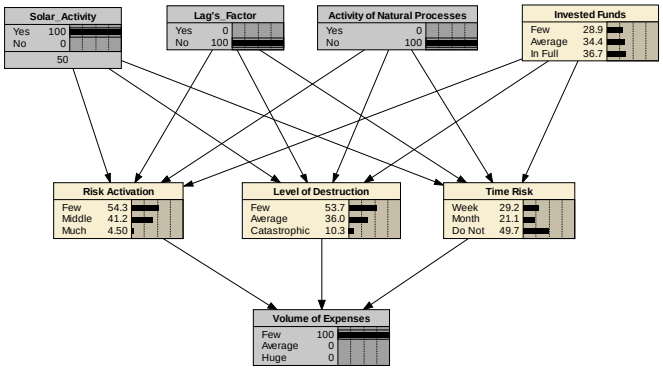


Рис. 4. Моделирование с помощью сети доверия Байеса для верхней вершины «Предварительное вложение средств»

Мы получили в первом случае следующий результат. Вероятность появления большого количества стихийных бедствий невысока (20 %), а вот при этом уровень катастрофических разрушений может достигнуть 30 %, которые произойдут в течение ближайшей недели с вероятностью 40 %. Общие затраты на покрытие ущерба от разрушений распределились следующим образом:

небольшие – 56 %, средние – 27 % и огромные – 17 %. Таким образом, можно сделать вывод, что при отсутствии вложения денежных средств в мероприятия, способствующие противостоять значительным разрушениям, общий катастрофический ущерб может доходить до 17 %.

IV. Заключение

Таким образом, чтобы учесть неопределенность и случайно проявляющиеся риски активации сложных природных процессов, можно использовать байесовские доверительные сети при моделировании и прогнозировании катастрофических последствий природных процессов [5].

Полученная модель может быть использована в управлении окружающей средой для демонстрации того, при каком пороговом значении происходят значительные изменения. В отличие от многих других многомерных статистических методологий, представленный подход может обрабатывать отсутствующие данные, а также данные различных типов, от полностью количественных до присутствия-отсутствия, в одном и том же анализе [2].

Применяемые примеры показывают, что протестированный метод может быть полезен при оценке факторов, потенциально влияющих на состояние индикаторов, и какими могут быть соответствующие пороговые значения этих переменных, что позволяет ее использовать в управлении окружающей средой для демонстрации того, при каком пороговом значении состояние индикатора может реагировать на давление или комбинацию давлений.

В отличие от многих многомерных статистических методологий, байесовские классификаторы могут обрабатывать отсутствующие данные и данные различных типов, от полностью количественных до присутствия-отсутствия, в одном и том же анализе [2]

Таким образом, при использовании системы поддержки принятия решений и создании единого технического хранилища для обработки информации и прогнозирования можно предусмотреть количество и характер (то есть последствия) активации сложных опасных природных процессов; время, в течение которого они могут происходить, а также оптимизировать затраты на реализацию мер по преодолению катастрофических последствий экзогенных процессов [5].

Список литературы

[1] Таран В.Н. Моделирование природных катастрофических процессов Южного берега Крыма с помощью сети Байеса. Auditorium. 2016. № 3 (11). С. 47-54.

[2] A. Lehikoinen, J. Olsson, L. Bergström, U. Bergström, A. Bryhn, R. Fredriksson and L. Uusitalo, Evaluating Complex Relationships Between Ecological Indicators and Environmental Factors in the Baltic Sea: a Machine Learning Approach. Ecological Indicators. 2019. Т. 101. С. 117-125. DOI: 10.1016/j.ecolind.2018.12.053

[3] P. Seisenbekova, A. Shayakhmetova and M. Othman, The Use of The Bayesian Approach in the Formation of The Student's Competence in the ICT Direction. 2019 IEEE International Conference on Automatic Control and Intelligent Systems, I2CACIS 2019 - Proceedings 2019. С. 85-90. DOI:10.1109/I2CACIS.2019.8825060

[4] V.N. Taran, Bayesian Networks for Modeling Complex Systems In: 2017 IEEE II International Conference on Control in Technical Systems (CTS), pp. 240-243. (2017) DOI: 10.1109/CTSYS.2017.8109535

[5] Таран В.Н. Моделирование сложных (опасных) природных процессов при помощи сети доверия Байеса. Прикаспийский журнал: управление и высокие технологии. 2019. № 2 (46). С. 90-100. DOI: 10.21672/2074-1707.2019.46.2.090-100

[6] V. Taran, Bayesian Belief Networks for Modeling Complex Natural Processes. Научная парадигма цивилизации в XXI веке: капитализм, социализм и четвертая технологическая революция Сборник трудов Десятого международного конгресса передовых наук (Конгресс WARP). Под редакцией В.И. Бархатова, Д.А. Плетнёва. 2019. С. 191-197.

[7] W. Wang, K. Shen, C. Dong, B. Wang, Q. Wang, F. Khan, Failure Probability Analysis of the Urban Buried Gas Pipelines Using Bayesian Networks. Process Safety and Environmental Protection: Transactions of the Institution of Chemical Engineers, Part B. 2017. Т. 111. С. 678-686. DOI:10.1016/j.psep.2017.08.040

[8] N.A. Kharitonov, A.L. Tulupyev, Algebraic Bayesian networks: sequestered fusion of knowledge patterns under information deficiency conditions. *Scientific and Technical Journal of Information Technologies, Mechanics and Optics*, 2019, vol. 19, no. 4, pp. 641–649 (in Russian). doi: 10.17586/2226-1494-2019-19-4-641-649

[9] Трухан С.В., Бидюк П.И. Применение сетей Байеса к построени. моделей оценки риска актуарные процессов. ScienceRise. 2016. Т. 8. № 2 (25). С. 6-14. DOI: 10.15587/2313-8416.2016.74962

[10] A.N. Terentyev, L.A. Korshevnyuk, P.I. Bidyuk, Bayesian Network as Instrument of Intelligent Data Analysis. Journal of Automation and Information Sciences. 2007. Т. 39. № 8. С. 28-38. DOI: 10.1615/JAutomatInfScien.v39.i8.40

[11] N.A. Kharitonov, A.A. Zolotin, A.L. Tulupyev, Software Implementation of Algorithms for Maintaining Consistency in Algebraic Bayesian Networks. In: 2017 IEEE Proceedings of the XXI International Conference on Soft Computing and Measurements SCM`2018, 19-22. DOI: 10.17586/2226-1494-2018-18-6-1108-1117

[12] S.V. Goriainov, A.Yu. Kuznetsov, E.V. Tushkanov, O.V. Kuznetsova and E.B. Romanova, Analysis of Bayesian Mathematical Systems for Pattern Recognition for Identifying Objects on Hyperspectral Images. In: 2017 IEEE Proceedings of the XXI International Conference on Soft Computing and Measurements SCM`2018, 69-72. DOI: 10.1109/SCM.2017.7970489

[13] A.A. Zolotin and A.L. Tulupyev, Matrix-Vector Algorithms for Global A Posteriori Output in Algebraic Bayesian Networks. In: 2017 IEEE Proceedings of the XXI International Conference on Soft Computing and Measurements SCM`2018, 45-48. DOI: 10.1109/SCM.2017.7970483

[14] Maiyar L.M., Tiwari M.K., Cho S., Kiritsis D., Thoben K.-D. Optimising Online Review Inspired Product Attribute Classification Using the Self-Learning Particle Swarm-Based Bayesian Learning Approach. International Journal of Production Research. 2019. Т. 57. № 10. С. 3099-3120. doi.org/10.1080/00207543.2018.1535724

[15] C. Li J. Tang, L. Zeng, C. Zhang, H. Ran and Z. Zhou, How to Optimize Ecosystem Services Based on a Bayesian Model: a Case Study of Jinghe River Basin. Sustainability. 2019. Т. 11. № 15. С. 4149.

[16] Marcot B.G., Penman T.D. Advances in Bayesian network modelling: Integration of modelling technologies. *Environmental Modelling and Software*, 2019, vol. 111, pp. 386–393. doi: 10.1016/j.envsoft.2018.09.016

[17] Ji Yae Shin, Muhammad Ajmal, Yoo Jiyoung, Kim Tae-Woong, A Bayesian Network-Based Probabilistic Framework for Drought Forecasting and Outlook. Advances in Meteorology. Volume 2016, Article ID 9472605, http://dx.doi.org/10.1155/2016/9472605

[18] Азарнова Т.В., Аснина Н.Г., Проскурин Д.К., Полухин П.В. Формирование структуры байесовской сети процесса тестирования надежности информационных систем. Вестник Воронежского государственного технического университета. 2017. Т. 13. № 6. С. 45-51.