УДК 004.89

МАРКОВСКИЕ ИДЕИ В БАЙЕСОВСКИХ СЕТЯХ

Ахметова А.М.[1], Абдилдаева А.А.[1], Литвиненко Н. Г.[1], Литвиненко А. Г.[2]

[1]Казахский Национальный Университет имени Аль-Фараби, г. Алматы
[2]Lehrstuhl für Mathematics for Uncertainty Quantification
ardak_66@mail.ru
ORCID ID: https://orcid.org/0000-0001-8360-5504
ORCID ID: https://orcid.org/0000-0002-6381-9350
ORCID ID: https://orcid.org/0000-0002-0576-8305
ORCID ID: https://orcid.org/0000-0001-5427-3598

Аннотация. В статье рассматриваются байесовские сети (в дальнейшем БС) с дополнительными ограничениями, обусловленными идеями марковских сетей (в дальнейшем МС). Графовые модели, описываемые подобными сетями, будем называть байесовскими марковскими сетями (БМС). Данные модели описывают многие реальные задачи с различными видами неопределенностей, имеющих различные причинно-следственные связи. Ограничения, накладываемые марковским свойством, во многих случаях позволяют значительно упростить расчеты в байесовских сетях при наличии свидетельств. В то же время практика показывает, что большинство моделей, отражающих реальные процессы и построенных на аппарате байесовских сетей в действительности обладают и марковским свойством.

В статье описаны идеи расчетов в байесовских сетях с ограничениями, накладываемыми марковским свойством. Описаны различия при расчетах в байесовских сетях без марковских ограничений и с марковскими ограничениями.

Работа написана в рамках грантового финансирования AP19679142 «Поиск оптимальных решений в байесовских сетях в моделях с линейными ограничениями и линейными функционалами. Разработка алгоритмов и программ» (2023-2025гг.) МОНВ РК.

Ключевые слова: Байесовская сеть, Марковская сеть, Марковское случайное поле, Марковские свойства, графическая модель, свидетельство, распространение свидетельств.

Введение

Познакомиться с теорией байесовских сетей можно в [1, 2, 3]. Подобные задачи часто имеют причинно-следственные связи между различными элементами. В основе данных графических моделей обычно используются направленные ацикличные графы. Вершинами в БС являются переменные, носящие вероятностный характер. Между переменными существуют различные причинно-следственные связи. Теория марковских сетей хорошо описана в [4, 5, 6]. Для иллюстрации излагаемого материала в статье используется программный комплекс HUGIN EXPERT. Ознакомится с работой данной программы можно в [7]. В статье будет использован известный учебный пример ASIA. Пример придуман и впервые описан в [8].

Байесовские сети позволяют решать различные достаточно сложные задачи, имеющие различные виды неопределенностей. Подобные задачи часто имеют причинно-следственные связи между различными переменными. Теория БС позволяет рассчитать значения некоторых переменных на основе известных или ранее рассчитанных переменных используя причинно-следственные связи. Если одна или несколько переменных получили некоторые значения, то говорят, что переменные получили свидетельства. Наличие свидетельств резко усложняет расчеты в БС. Поиск хороших, быстрых алгоритмов для расчетов в БС со свидетельствами достаточно актуальная проблема. Дополнительное условие марковости во многих случаях может значительно упростить алгоритмы. На примере будет рассмотрена возможность подобного упрощения.

Постановка задачи

Краткое описание байесовских сетей и марковских сетей.

Пусть задан граф $\mathbf{G} = <\mathbf{V}, \mathbf{E}>$, где $\mathbf{V}$ – множество вершин графа, $\mathbf{E}$ – множество ребер графа. Ориентированный, ацикличный граф $\mathbf{G}$ называется байесовской сетью если каждой вершине $\mathbf{v} \in \mathbf{V}$ поставлена в соответствие случайная величина $\mathbf{X_v}$, а каждому ребру $\mathbf{e} = (\mathbf{u}, \mathbf{v}) \in \mathbf{E}$ ставится вероятностная зависимость случайной величины $\mathbf{X_v}$ от случайной величины $\mathbf{X_u}$.

Вершина u называется родительской для вершины v, если ребро $(\mathbf{u}, \mathbf{v}) \in \mathbf{E}$. Множество всех родительских вершин v обозначается через $\mathbf{parents}(\mathbf{X_v})$. Для вычисления вероятности существует цепное правило для байесовских сетей:

$$\boldsymbol{P}(\mathbf{X_1}, \mathbf{X_2}, \mathbf{X_3}, \dots, \mathbf{X_n}) = \prod_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{n}} \boldsymbol{P}(\mathbf{X_i}\,|\mathbf{parents}(\mathbf{X_v}))$$

Цепное правило позволяет разложить совместное распределение в произведение условных распределений. Очевидно, что при расчете соблюдается принцип Маркова: переменная зависит только от родительских переменных и не важно каким образом родительские переменные получили свое значение.

Марковская сеть это графовая модель, представленная неориентированным графом, в которой множество случайных величин обладает Марковским свойством. Марковское свойство определяется одним из трех эквивалентных способов:

Свойство пар: Любые две несмежные переменные условно независимы с учетом всех других переменных

Локальное свойство: переменная условно независима от всех других величин, с учетом своих соседей

Глобальное свойство: Любые два подмножества переменных условно независимы с учетом разделяющего подмножества

В байесовских сетях играет важное понятие d-разделенности. Суть данного понятия – попытаться выделить из байесовской сети некоторую подсеть, расчеты в которой не зависят от вершин, не лежащих в данной подсети. Это позволит существенно снизить объемы расчетов во всей сети. Ниже приведено определение d-разделенности.

Путь S между вершинами a и b называют d-разделённым или блокированным множеством вершин Z тогда и только тогда, когда:

S содержит цепь $i \to z \to j$ или разветвление $i \leftarrow z \to j$ такие, что z принадлежит Z или

S содержит инвертированное разветвление (коллайдер) $i \to z \leftarrow j$ такие, что z не принадлежит Z и у вершины z нет потомков принадлежащих Z.

Понятие d-разделенности можно распространить и на непересекающиеся подмножества A и B. Два подмножества A и B называются d-разделённым, если любые два элемента из этих подмножеств d-разделены.

Рассмотрим известный в литературе по БС учебный пример ASIA. В данном примере рассматриваются несколько вариантов задания свидетельств. Проводим анализ решений. Рассматриваем способ решения при условии марковости. Сравниваем два способа построения алгоритмов.

Описание учебного примера (рисунок 1).

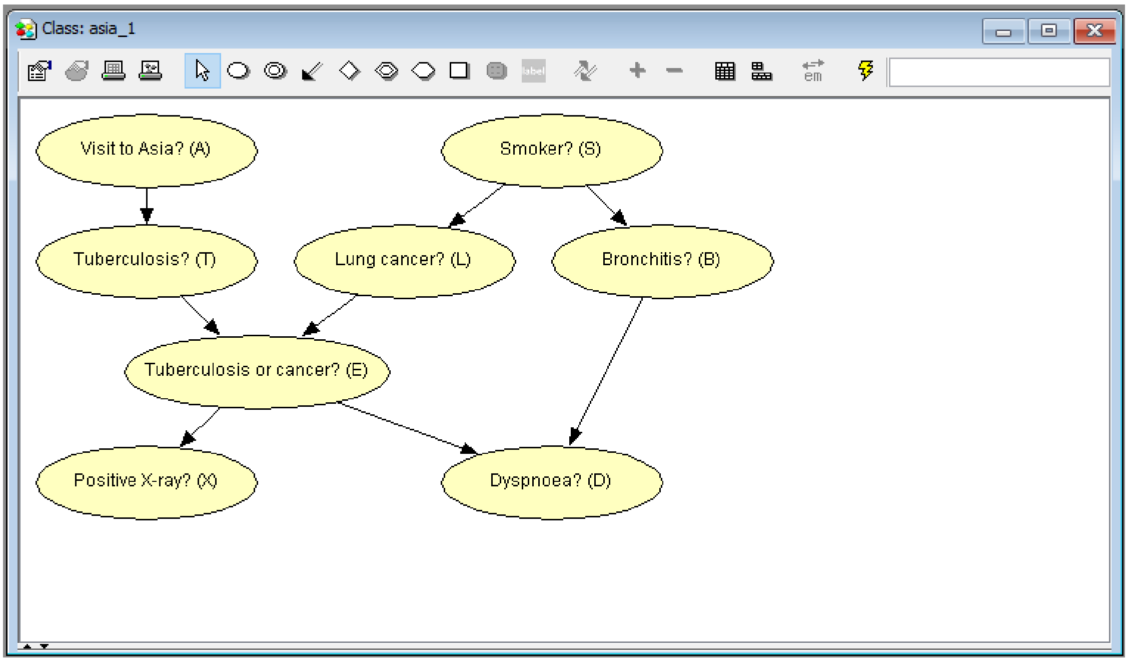

Рисунок 1 - Учебная байесовская сеть ASIA

Исследуется здоровье человека. Вершина А обозначает был или нет исследуемый в Азии. В таблице 1 определена вероятность данного факта. От этого в значительной степени зависит вероятность заболевания туберкулезом (Т). В таблице 2 определена вероятность заболевания. Вершина S и таблица 3 определяют курит или нет человек. Курение может способствовать заболеванию раком (L), заболеванию бронхитом (B) или и тем и другим одновременно (E) (таблица 4, таблица 5 и таблица 6). Результат рентгена (X) грудной клетки может показать с определенной вероятностью (таблица 7) заболевание раком или туберкулезом, но различить эти болезни не может. Рак, туберкулез и бронхит с определенной вероятностью вызывают одышку (таблица 8).

Visit to Asia? (A)

Edit Functions View

yes	0.01
no	0.99

Таблица 1

Tuberculosis? (T)

Edit Functions View

Visit to Asia? (A)	yes	no
yes	0.05	0.01
no	0.95	0.99

Таблица 2

Smoker? (S)

Edit Functions View

yes	0.5
no	0.5

Таблица 3

Lung cancer? (L)

Edit Functions View

Smoker? (S)	yes	no
yes	0.1	0.01
no	0.9	0.99

Таблица 4

Bronchitis? (B)

Edit Functions View

Smoker? (S)	yes	no
yes	0.6	0.3
no	0.4	0.7

Таблица 5

Tuberculosis or cancer? (E)

Edit Functions View

Tuberculosi...	yes		no	
Lung cance...	yes	no	yes	no
yes	1	1	1	0
no	0	0	0	1

Таблица 6

Positive X-ray? (X)

Edit Functions View

Tuberculosis or cancer? (E)	yes	no
yes	0.98	0.05
no	0.02	0.95

Таблица 7

Dyspnoea? (D)

Edit Functions View

Bronchitis? (B)	yes		no	
Tuberculosis or cancer? (E)	yes	no	yes	no
yes	0.9	0.8	0.7	0.1
no	0.1	0.2	0.3	0.9

Таблица 8

Методы исследования

Рассмотрим стандартные расчеты данного примера (рисунок 2). Предполагаем, что свидетельства отсутствуют. Вероятности переменных рассчитываются последовательно, начиная с маргинальных (независимых) переменных. Маргинальными в нашем примере являются переменные A и S. Маргинальные переменные отнесем к нулевому уровню. В первую очередь рассчитываются переменные, зависящие лишь от маргинальных переменных. Это переменная T, зависящая от маргинальной переменной A и переменные L и B, зависящие от переменной S. Переменные S, L, B отнесем к первому уровню. Далее мы рассчитываем переменные второго уровня. Эти переменные могут зависеть от переменных нулевого и первого уровня. E единственная переменная второго уровня, зависящая от переменных первого уровня T и L. Затем рассчитываем переменные третьего уровня. Эти переменные могут зависеть от переменных нулевого, первого и второго уровня. Переменная X зависит от переменной второго уровня E. Переменная D зависит от переменной второго уровня E и переменной первого уровня B.

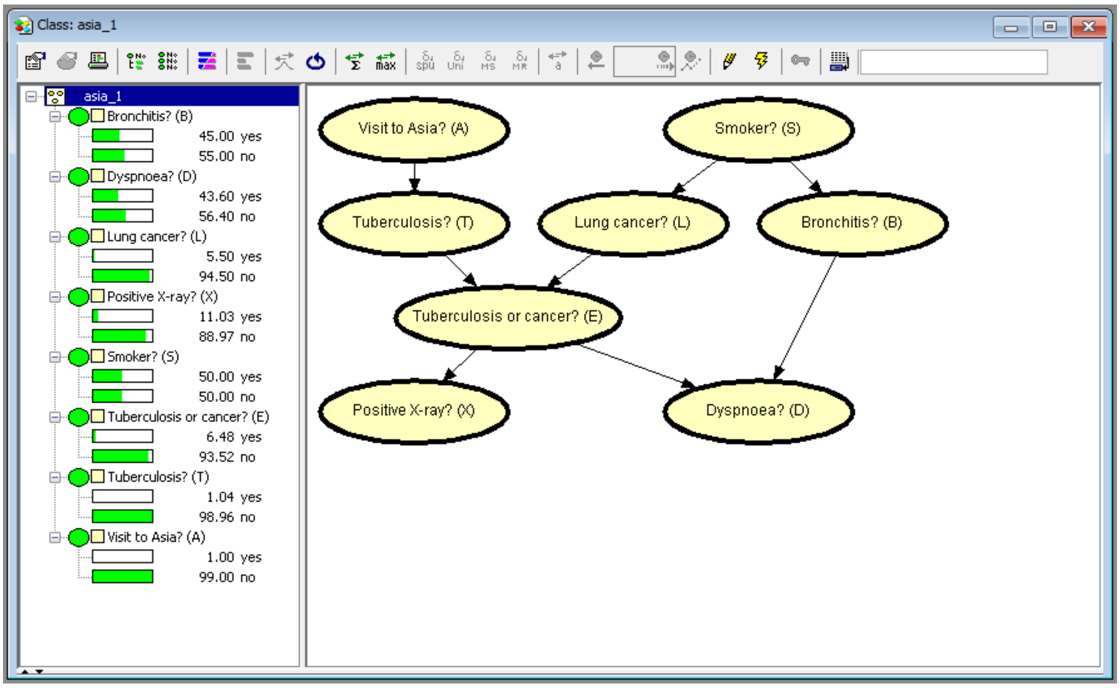

Рисунок 2 - Пример стандартного расчета

Решение находится слева. Из решения видно, например, что вероятность заболевания туберкулезом или раком (E) при заданных маргинальных переменных равна 6.4%. А вероятность иметь одышку (D) равна 43.6%. Заметим, что в расчетах мы использовали только соседние переменные, т.е. реально и существенно использовали марковское свойство.

Ситуация с расчётами существенно меняется при получении свидетельств. Допустим мы сделали рентген, давший положительный результат. Решение мы видим на рисунке 3. Наиболее популярный алгоритм расчетов в байесовских сетях состоит в следующих построениях:

1. Построение доменного графа.
2. Построение морального графа.
3. Построение триангулярного графа.
4. Построение дерева смежности.
5. Построение дерева сочленений.

Разработка и реализация подобных алгоритмов достаточно сложна. Хотелось бы найти возможности определить и в данном случае марковское свойство и найти способ разделить переменные на группы и проводить расчеты тем же способом, как и в случае отсутствия свидетельств.

При наличии свидетельств, свидетельства становятся в каком-то смысле определяющими в расчетах. В нашем случае будем считать переменную X переменной нулевого уровня. Переменная E будет переменной первого уровня, зависит только от переменной нулевого уровня. Переменные T, L, D будут переменными второго уровня. Переменные A, S будут переменными третьего уровня. Переменная S будет переменной четвертого уровня, зависит от переменных второго и третьего уровня.

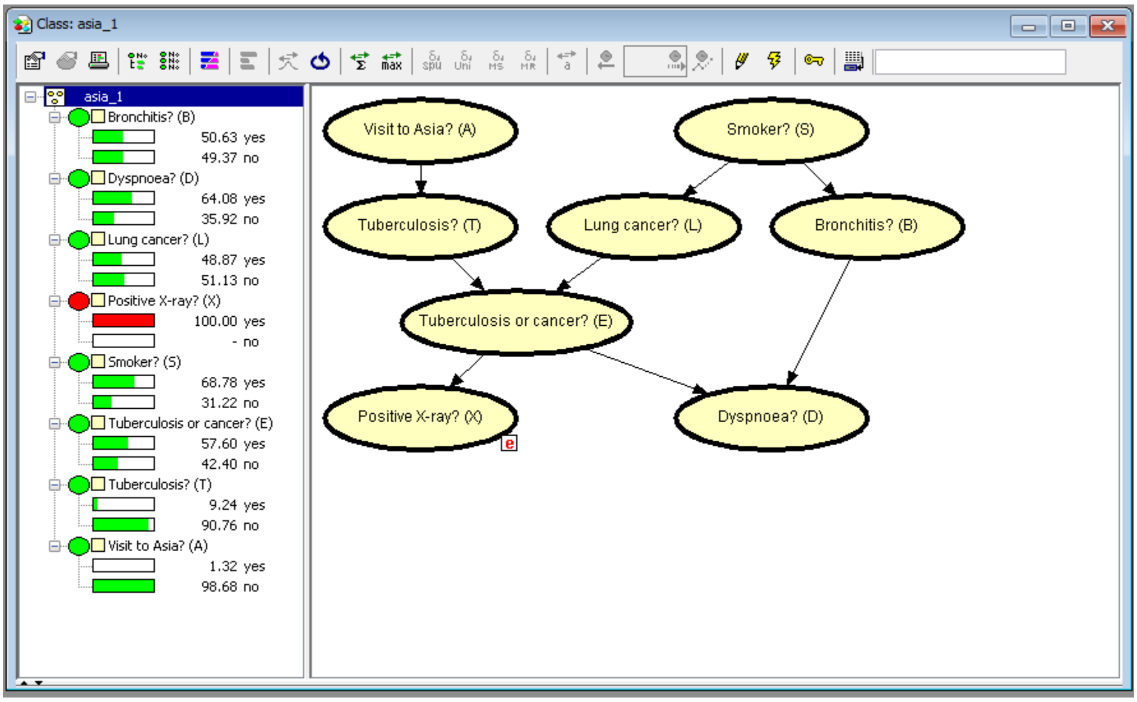

Рисунок 3 - Решение при получении свидетельств

Расчеты, в силу структур таблиц условных вероятностей, можно проводить используя теорему Байеса.

Результаты и их обсуждения

В статье рассматриваются возможности использования свойства марковости для упрощения расчетов в байесовских сетях. Решая одни проблемы, мы получаем новые проблемы:

Обоснование использования свойства Маркова.

Как использовать существующие таблицы вероятностей с недостаточно удобными структурами.

Как корректно обосновать использование теоремы Байеса.

Необходимо доказать, что при нескольких свидетельствах и смене направлений некоторых ребер не возникает циклов.

Исследование возникновения противоречий в байесовских сетях при некорректном, несовместимом получении нескольких свидетельств. Пути решения возможных противоречий.

Тем не менее рассматриваемый подход представляется достаточно перспективным и требует дальнейшего исследования.

Заключение

В статье рассматриваются возможности использования марковости при расчетах в байесовских сетях со свидетельствами. Цель – упростить алгоритмы расчетов. Работа выполнена в рамках грантового проекта AP19679142 «Поиск оптимальных решений в байесовских сетях в моделях с линейными ограничениями и линейными функционалами. Разработка алгоритмов и программ» (2023-2025гг.) МОНВ РК.

Список литературы

[1] Pearl J. Probabilistic Reasoning in Intelligent Systems. San Francisco: Morgan Kaufmann Publishers, 1988. 552 p

[2] Литвиненко Н.Г., Литвиненко А.Г., Мамырбаев О.Ж., Шаяхметова А.С. Байесовские сети. Теория и практика. Алматы: Институт информационных и вычислительных технологий, 2020. 197 p. ISBN 978-601-332-888-1.

[3] Robert G. Cowell, A. Philip Dawid, Steffen L. Lauritzen, David J. Spiegelhalter. Probabilistic Networks and Expert Systems. Springer, 1999. 3321 p. ISBN 0-387-98767.

[4] Li, S. Z. Markov Random Field Modeling in Image Analysis. Springer, London, 2009.

[5] Kindermann Ross; Snell, J. Laurie. Markov Random Fields and Their Applications (англ.). American Mathematical Society, 1980. ISBN 0-8218-5001-6.

[6] Rue, Håvard; Held, Leonhard. Gaussian Markov random fields: theory and applications (англ.). CRC Press, 2005. — ISBN 1584884320.

[7] HUGIN Graphical User Interface. Documentation. Release 9.1. 2021.

[8] Lauritzen, S.L. and Spiegelhalter, D.J. Local Computations with Probabilities on Graphical Structures and Their Application to Expert Systems. Journal of the Royal Statistical Society. Series B, 50, 157-224. 1988.

БАЙЕСТІК ЖЕЛІЛЕРДЕГІ МАРКОВТІК ИДЕЯЛАР

Ахметова А.М.[1], Абдилдаева А.А. [1], Литвиненко Н. Г. [1], Литвиненко А. Г. [2]

[1]әл-Фараби атындағы Қазақ ұлттық университеті, Алматы қ-сы
[2]Белгісіздікті сандық бағалау үшін математика кафедрасы, Германия
ardak_66@mail.ru
ORCID ID: https://orcid.org/0000-0001-8360-5504
ORCID ID: https://orcid.org/0000-0002-6381-9350
ORCID ID: https://orcid.org/0000-0002-0576-8305
ORCID ID: https://orcid.org/0000-0001-5427-3598

Аңдатпа. Мақалада Марков желілерінің (бұдан әрі МС) идеяларына байланысты қосымша шектеулермен Байес желілері (бұдан әрі БС) қарастырылады. Ұқсас желілермен сипатталған графикалық модельдерді Байес Марков желілері (BMS) деп атаймыз. Бұл модельдер әртүрлі себептік байланыстары бар әртүрлі белгісіздіктері бар көптеген нақты міндеттерді сипаттайды. Марков қасиеті қойған шектеулер көптеген жағдайларда дәлелдер болған кезде Байес желілеріндегі есептеулерді едәуір жеңілдетеді. Сонымен қатар, тәжірибе көрсеткендей, нақты процестерді көрсететін және аппаратта салынған Байес желілерінің модельдерінің көпшілігі Марковтық қасиетке ие.

Мақалада Марков қасиетімен шектелген Байес желілеріндегі есептеу идеялары сипатталған. Марков шектеулері жоқ және Марков шектеулері бар Байес желіліреіндегі есептеулердегі айырмашылықтар сипатталған.

Жұмыс AP19679142 "Сызықтық шектеулер мен сызықтық функционалдығы бар модельдерде Байес желілерінде оңтайлы шешімдерді табу. Алгоритмдер мен бағдарламаларды құру" (2023-2025жж.) гранттық қаржыландыру аясында орындалған.

Кілттік сөздер: Байес желісі, Марков желісі, Марковтың кездейсоқ өрісі, Марков қасиеттері, графикалық модель, айғақтар, айғақтардың таралуы.

MARKOV IDEAS IN BAYESIAN NETWORKS

Akhmetova A.M. [1], Abdildayeva A.A. [1], Litvinenko N. G. [1], Litvinenko A. G. [2]

Al-Farabi Kazakh National University, Almaty
Lehrstuhl für Mathematics for Uncertainty Quantification, Germany
ardak_66@mail.ru
ORCID ID: https://orcid.org/0000-0001-8360-5504
ORCID ID: https://orcid.org/0000-0002-6381-9350

Марковские идеи в байесовских сетях
Ахметова А.М., Абдилдаева А.А., Литвиненко Н. Г., Литвиненко А. Г.

ORCID ID: https://orcid.org/0000-0002-0576-8305
ORCID ID: https://orcid.org/0000-0001-5427-3598

Abstract. The article deals with Bayesian networks (hereinafter BS) with additional limitations due to the ideas of Markov networks (hereinafter MS). Graph models described by such networks will be called Bayesian Markov networks (BMS). These models describe many real-world problems with different types of uncertainties having different cause-and-effect relationships. The limitations imposed by the Markov property in many cases make it possible to significantly simplify calculations in Bayesian networks in the presence of evidence. At the same time, practice shows that most models reflecting real processes and built on the apparatus of Bayesian networks actually have the Markov property.

The article describes the ideas of calculations in Bayesian networks with restrictions imposed by the Markov property. The differences in calculations in Bayesian networks without Markov constraints and with Markov constraints are described.

The work was written within the framework of grant funding AP19679142 "Search for optimal solutions in Bayesian networks in models with linear constraints and linear functionals. Development of algorithms and programs" (2023-2025) MONV RK.

Keywords: Bayesian network, Markov network, Markov random field, Markov properties, graphical model, evidence, dissemination of evidence.

Сведения об авторах:

Рус: Ахметова А.М. - PhD, и.о. ассоциированного профессора кафедры Искусственного интеллекта и Big Data, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан.

Қаз: Ахметова А.М. - PhD, Жасанды интеллект және Big Data кафедрасының қауымдастырылған профессор м.а., әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.

Англ: Akhmetova A.M. - PhD, acting Associate Professor of the Department of Artificial Intelligence and Big Data, Al-Farabi Kazakh National University, Almaty, Kazakhstan.

Рус: Абдилдаева А.А. - PhD, ассоциированный профессор кафедры Искусственного интеллекта и Big Data, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан.

Қаз: Абдилдаева А.А. - PhD, Жасанды интеллект және Big Data кафедрасының қауымдастырылған профессоры, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.

Англ: Abdildayeva A.A. - PhD, Associate Professor of the Department of Artificial Intelligence and Big Data, Al-Farabi Kazakh National University, Almaty, Kazakhstan.

Рус: Литвиненко Н.Г. - магистр, Казахский национальный университет имени аль-Фараби, Алматы, Казахстан.

Қаз: Литвиненко Н.Г.- магистр, әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.

Англ: Litvinenko N. G.- master, Al-Farabi Kazakh National University, Almaty, Kazakhstan.

Рус: Литвиненко Александр- доктор естественных наук кафедры математики количественной оценки неопределенности, Германия

Қаз: Литвиненко Александр – жаратылыстану ғылымдарының докторы, белгісіздікті сандық бағалау математика кафедрасы, Германия

Англ: Litvinenko Alexander - Dr. rer. nat. Chair of Mathematics for Uncertainty Quantification, Germany.