

А.А. ФИЛЬЧЕНКОВ, К.В. ФРОЛЕНКОВ, А.Л. ТУЛУПЬЕВ  
**УСТРАНЕНИЕ ЦИКЛОВ ВО ВТОРИЧНОЙ СТРУКТУРЕ  
АЛГЕБРАИЧЕСКОЙ БАЙЕСОВСКОЙ СЕТИ  
НА ОСНОВЕ АНАЛИЗА ЕЕ ЧЕТВЕРТИЧНОЙ СТРУКТУРЫ**

---

*Фильченков А.А., Фроленков К.В., Тулупьев А.Л. Устранение циклов во вторичной структуре алгебраической байесовской сети на основе анализа ее четвертичной структуры.*

**Аннотация.** Алгебраическая байесовская сеть (АБС) — одна из логико-вероятностных графических моделей баз фрагментов знаний с неопределенностью. Алгоритмы глобального логико-вероятностного вывода АБС могут применяться при условии ацикличности её вторичной структуры — графа смежности. Существующий метод преобразования графа смежности в дерево смежности ограниченно применим. Цель работы — предложить новые методы преобразования цикличной АБС к ацикличной, основывающиеся на структурной теореме о циклах минимальных графов смежности. В работе предложено два метода устранения циклов и доказана их корректность.

**Ключевые слова:** алгебраические байесовские сети, четвертичная структура, вероятностные графические модели систем знаний, глобальная структура, ацикличность первичной структуры.

*Filchenkov A.A., Frolenkov K.V., Tulupyev A.L. Algebraic Bayesian Network Secondary Structure Cycles Elimination Based on its Quaternary Structure Analysis.*

**Abstract.** Algebraic Bayesian Network (ABN) is one of the logical and probabilistic graphical models of bases of knowledge patterns uncertainty. Algorithms for global logical and probabilistic inference in ABN can be applied only under the condition of acyclicity of its secondary structure — join graph. The existing method for join graph transformation into join tree application is restrictedly applicable. The goal of the work is to offer new methods for the transformation cyclic ABN into an acyclic one, based on a structure theorem for minimal join graphs cycles. Two methods for eliminating cycles are proposed and their correctness is proved.

**Keywords:** algebraic Bayesian networks, quaternary structure, machine learning, probabilistic graphical knowledge models, global structure, primary structure acyclicity.

---

**1. Введение.** Алгебраическая байесовская сеть (АБС) — одна из логико-вероятностных графических моделей систем знаний с неопределенностью. АБС позволяют использовать интервальные оценки вероятности для представления неопределенности в знаниях, что выгодно отличает их от прочих вероятностно-графических моделей [10–15, 18, 19].

АБС может быть применима [28] как одна из математических моделей анализа информационной безопасности [2, 3]. В частности, АБС могут использоваться [7] для оценки защищённости системы от социоинженерных атак [1, 17]. Обработка интервальных оценок также оказывается востребованной в анализе гранулярных данных о рекордных интервалах между последними эпизодами [8, 9].

Одним из наиболее значимых нерешенных вопросов в теории АБС является разработка методов обучения сети. У АБС выделяют несколько глобальных структур, из которых наиболее важными являются две: первичная, представляющая набор фрагментов знаний (ФЗ), математической моделью которых выступает идеал конъюнктов, с заданными на них оценками вероятности истинности, и вторичная — особый граф (*граф смежности*), построенный над первичной структурой. Существует несколько подзадач, которые необходимо решить в рамках задачи обучения АБС [21, 22]. Одной из таких задач является преобразование первичной структуры, над которой невозможно построение ациклического графа смежности (будем называть такие первичные структуры *циклическими*), к стохастически эквивалентной первичной структуре, над которой построение ациклического графа смежности возможно (такие первичные структуры будем называть *ациклическими*). Требование ацикличности первичной структуры возникает в связи с особенностями алгоритмов логико-вероятностного вывода, которые определены только для дерева смежности [10, 13, 1819].

В 2008 году был предложен метод преобразования первичной структуры с построенным над ней графом смежности к первичной структуре с ациклическим графом смежности [16], однако он применим лишь к обособленным циклам. Поэтому задача преобразования циклической первичной структуры к ациклической по-прежнему актуальна.

Над заданной первичной структурой можно построить целое множество минимальных графов смежности. Известно, что для ациклической первичной структуры множество ациклических графов смежности совпадает с множеством минимальных графов смежности [4]. Исследование минимальных графов смежности активно развивалось в последнее несколько лет [4, 5, 20, 23–27, 29–37], одним из результатов работы в этом направлении стала теорема о циклах в минимальных графах смежности [30], в которой сформулированы необходимые и достаточные условия ацикличности минимального графа смежности над заданной первичной структурой. Эти условия сформулированы относительно четвертичной структуры АБС. Формулировка этой теоремы открывает возможности для разработки новых методов устранения циклов.

*Цель работы* — предложить новые методы преобразования циклической первичной структуры к ациклической, основывающиеся на теореме о циклах минимальных графов смежности.

**2. Основные определения.** Будем пользоваться системой терминов, введенной в [14, 32, 33]. В работе [30] была сформулирована теорема о необходимых и достаточных условиях ацикличности первичной структуры, которая будет являться основной для методов, предлагаемых в данной работе. В данном разделе мы приведем лишь необходимые для формулировки данной теоремы понятия.

*Значимая клика*  $U$  — значимое сужение максимального графа смежности на вес  $U$ . Множество значимых клик обозначается как  $\text{Clique}$ .

*Практика* — ненаправленный граф, являющийся сужением максимального графа смежности на пустой вес (т. е. фактически совпадающий с ним).

*Замкнутое сверху множество значимых клик* — объединение  $\text{Clique}$  с одноэлементным множеством, содержащим практику.

В рамках данной работы для краткости будем *кликкой* называть элемент множества  $\text{Clique}$  или практику.

*Родительский граф над замкнутым сверху множеством значимых клик* — направленный граф, вершинами которого являются клики. Ребро из вершины  $P$  в вершину  $Q$  проведено, если клика  $P$  содержит клику  $Q$ , и не существует клики  $R$ , такой, что клика  $P$  содержит клику  $R$  и клика  $R$  содержит клику  $Q$ . Тогда клика  $P$  — *отец* клики  $Q$ , а  $Q$  — *сын*  $P$ .

*Вассал* — множество вершин, входящих в какого-либо сына клики  $U$ .

Два вассала  $V_U^i$  и  $V_U^j$  называются *братьями*, если их пересечение непусто:

$$V_U^i \leftrightarrow V_U^j \Leftrightarrow V_U^i \cap V_U^j \neq \emptyset.$$

Два вассала  $V_U^i$  и  $V_U^j$  называются *полусиблингами*, если существует такой упорядоченный набор вассалов  $\{V_U^{w_1}, \dots, V_U^{w_n}\}$ , что  $V_U^i$  — брат  $V_U^{w_1}$ ,  $V_U^{w_1}$  — брат  $V_U^{w_2}$ , ...,  $V_U^{w_{n-1}}$  — брат  $V_U^{w_n}$ , а  $V_U^{w_n}$  — брат  $V_U^j$ .

*Полусиблинговый путь* между двумя родственными вассалами  $V_U^i$  и  $V_U^j$  — такой набор вассалов  $\{V_U^{w_1}, \dots, V_U^{w_n}\}$  из определения полусиблингов, что  $V_U^i \leftrightarrow V_U^{w_1}$ ,  $V_U^{w_1} \leftrightarrow V_U^{w_2}$  и  $\forall i < n$   $V_U^{w_i} \leftrightarrow V_U^{w_{i+1}}$ .

*Братский цикл* — полусиблинговый цикл  $S_1, S_2, \dots, S_n$  такой, что пересечение всех этих вассалов непусто.

*Небратский цикл* — полусиблинговый цикл, не являющийся братским.

**Теорема (о циклах минимальных графов смежности) [30].** Любой минимальный граф смежности, построенный над заданным набором вершин, является ациклическим тогда и только тогда, когда не существует полусиблинговых небратских простых циклов.

**3. Методы устранения обособленных циклов.** Этот метод был предложен в [16].

Метод устранения обособленных циклов предполагает, что дана первичная структура и для нее построен граф смежности, который содержит цикл. Более того, цикл предполагается не содержащим дуг.

Если цикл четной длины, все его вершины разбиваются на пары так, чтобы сумма их номеров в цикле по каждой паре была одинакова, после чего ФЗ из каждой пары подвергаются слиянию (рис. 1).

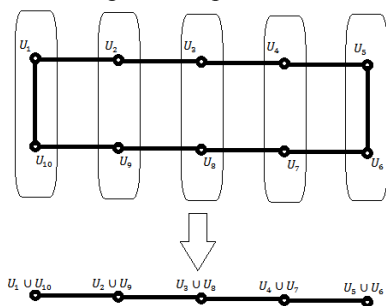


Рис. 1. Пример работы алгоритма преобразования цикла в цепь смежности для чётного числа ФЗ.

Для циклов нечетной длины производятся те же действия, за исключением того, что вместо одной пары слиянию подвергается 3 фрагмента (рис. 2). Результатом применения этого метода будет граф, удовлетворяющий ограничениям, накладываемым на графы смежности. Кроме того, АБС, полученная таким образом, глобально непротиворечива тогда и только тогда, когда глобально непротиворечива исходная АБС [14].

К сожалению, область применимости описанного метода ограничена. В предложенной формулировке он позволяет устранять лишь обособленные циклы. Известно [6, 12, 18], что локальные алгоритмы, имеют экспоненциальную (относительно количества атомов фрагмента знаний) сложность по времени и памяти. Это компенсируется малыми размерами ФЗ. Но при слиянии двух или нескольких ФЗ это свойство

может исчезнуть. В силу вышесказанного необходимо предложить более генерализованные методы устранения циклов.

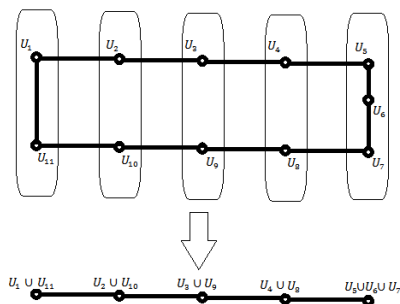


Рис. 2. Пример работы алгоритма преобразования цикла в цепь смежности для нечётного числа ФЗ.

#### 4. Методы устранения цикличности первичной структуры.

Если предыдущий метод предполагал, что даны первичная структура и построенная над ней вторичная, то рассматриваемые в этом разделе методы исходят из того, что дана лишь первичная структура, и задача состоит в том, чтобы преобразовать ее к ациклической первичной структуре. Эта идея значительно отличается от предыдущего метода, в котором первичная структура менялась в зависимости от уже построенной вторичной структуры, поэтому одна и та же первичная структура могла быть преобразована к двум разным первичным структурам, если над ней были построены разные графы смежности.

**4.1. Наивный метод устранения циклов через добавление фрагмента знаний.** Пусть дана первичная структура АВС. Необходимые и достаточные условия наличия в ней циклов были сформулированы в приведенной выше теореме о циклах. Поэтому основная идея этого и рассматриваемого далее метода состоит в том, чтобы устранить все небратские полусиблинговые простые циклы, следствием чего будет являться возможность построения ациклической вторичной структуры.

Рассмотрим небратский полусиблинговый простой цикл. Допустим, он состоит из значимых клик весов  $U_1, U_2, \dots, U_n$ . Поскольку цикл небратский, то не существует фрагмента знаний, который бы содержал вес  $U_{i=1}^n U_i$ . Добавим такой ФЗ в первичную структуру (он бу-

дет максимальным) и удалим те МФЗ, которые перестали быть максимальными. Рассматриваемый цикл перестал быть небратским.

Описанный метод может эффективно применяться в ряде примеров, для которых преобразование цикла в цепь смежности даёт худшие результаты. Однако было замечено, что такое действие не всегда ведет к сокращению количества циклов.

Для иллюстрации рассмотрим следующий пример: для набора фрагментов знаний  $\{ace, acd, be, bcd\}$  минимальный граф смежности и соответствующий родственный граф выглядят, как показано на рис 3.

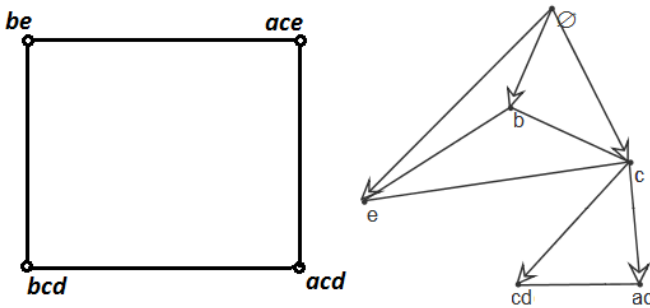


Рис. 3. МГС и семейный граф для  $\{ace, acd, be, bcd\}$ .

$bce$  — полусиблинговый небратский цикл. Следует добавить вершину веса  $bce$ . В результате получаем (рис 4).

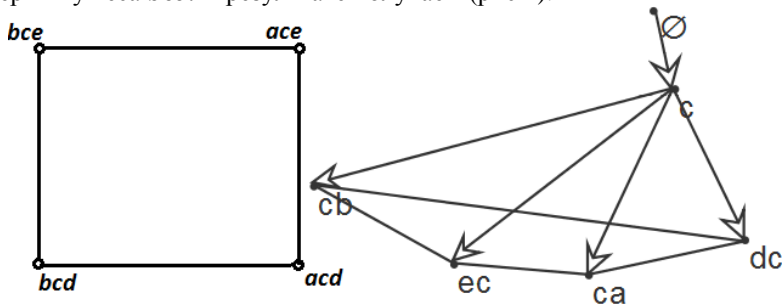


Рис. 4. МГС и родительский граф после добавления  $bce$ .

То есть не любая операция добавления  $U_{i=1..n} U_i$  для весов всех входящих в цикл клик уменьшает количество циклов. Несмотря на это,  $\sum(H_0 - h_i)$  (где  $H_0$  — высота родительского графа,  $h_i$  — расстояние от цикла с номером  $i$  до праплики) по всем циклам в родительском графе

уменьшается при каждом таком действии. Таким образом, последовательно устраняя циклы, спускаясь вниз по родительскому графу (т.е. выбирая небратские полусиблинговые простые циклы для клик, соответствующих вершинам графа), после перебора всех клик будет построена ациклическая первичная структура, а это значит, что метод применим.

**4.2. Метод устранения циклов через однократное добавление фрагмента знаний.** В случаях, подобных приведенному примеру, для применения последнего метода требуется последовательная перестройка родительского графа с учетом изменений первичной структуры, что также может оказаться затратным. Поэтому был предложен модифицированный метод: рассмотрим небратский полусиблинговый простой цикл. Допустим, он состоит из значимых клик весов  $U_1, U_2, U_3, \dots, U_n$ . Построим вес, который объединяет веса всех потомков всех этих клик  $\bigcap_{i=1..n, j=1..d_i} D_{i,j}$  где  $D_{i,j}$  — вес  $j$ -го потомка клики  $U_i$ . Добавим ФЗ соответствующего веса в первичную структуру и удалим все МФЗ, которые перестанут быть максимальными. Так мы устраним все небратские полусиблинговые простые циклы и не добавим ни одного нового.

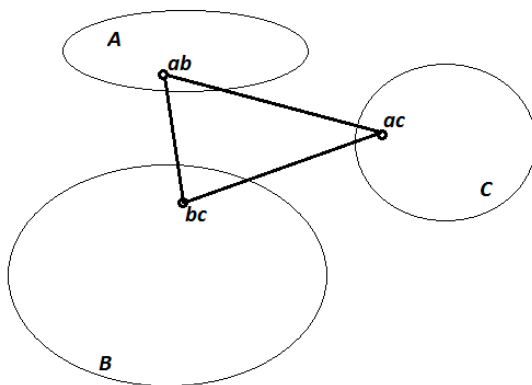


Рис.5 Схема графа смежности, стягиваемого в единственную вершину.

Этот метод, как и первый из описанных, эффективен по времени, но также имеет ряд примеров исходных данных, построенные по которым АВС содержат слишком крупные фрагменты знаний, что приводит к неэффективности. Например, для графа, состоящего из 3 компонент  $A, B$  и  $C$ , все элементы которых содержат соответствующие

переменные, и связанных, как показано на рисунке 5, результирующая АБС состоит всего из одного ФЗ.

**5. Заключение.** В работе были рассмотрены три метода построения ациклической вторичной структуры для заданной первичной структуры. Первый из рассматриваемых методов, предложенный в 2008 году [16], состоит в устранении циклов в заданной вторичной структуре путем объединения фрагментов знаний, входящих в цикл, таким образом, чтобы они этот цикл превратились в цепь. Данный метод применим только к обособленным циклам, точнее — к АБС, содержащим лишь один цикл.

Два других — новых — метода, предложенных в данной работе, опираются на структурную теорему о циклах в минимальных графах смежности [30] и преобразуют произвольную первичную структуру в ациклическую. Разница между этими двумя методами состоит в том, что первый из них требует постоянного перестроения первичной структуры путем последовательного добавления фрагментов знаний, тогда как второй предполагает перестроение первичной структуры ровно столько раз, сколько существует небратских полусиблинговых простых циклов в четвертичной структуре соответствующей АБС. Последние два метода применимы для произвольных первичных структур, а не только для обособленных циклов.

Развитие компаративного анализа указанных методов затрудняется тем, что пока не выработано критериев для сравнения. Полученные первичные структуры необходимо сравнивать между собой, и основным кандидатом в критерии является размер фрагментов знаний в этих структурах. Однако оценка, соответствующая рассматриваемому кандидату в критерии, не формализована, что пока не позволяет количественно оценить эффективность или неэффективность предлагаемых методов.

Помимо этого следует указать также на то, что исходная задача — построение ациклической первичной структуры, также требует более тщательной, строгой, полной формализации, отсутствие которой сдерживает исследования в соответствующем направлении. Отсутствие выделенного и формализованного инварианта, который должен быть сохранен в первичной структуре после ее преобразования к ациклической, приводит к тому, что пока невозможно указать критерий «допустимости» способов изменения первичной структуры, а также выполнить строгое обоснование корректности таких способов.



## Литература

1. *Ванюшичева О.Ю., Тулупьева Т.В., Пащенко А.Е., Тулупьев А.Л., Азаров А.А.* Количественные измерения поведенческих проявлений пользователя, ассоциированных с социоинженерными атаками. // Труды СПИИРАН. 2011. Вып. 19. С. 34–47.
2. *Котенко И.В., Степашкин М.В., Юсупов Р.М.* Модели и методы информационной безопасности математические модели, методы и архитектуры для защиты компьютерных сетей: аналитический обзор перспективных направлений исследований по результатам международного семинара MMM-ACNS-2005 // Труды СПИИРАН. 2006. Т. 2. № 3. С. 11-29.
3. *Котенко И.В., Юсупов Р.М.* Перспективные направления исследований в области компьютерной безопасности // Защита информации. Инсайд. 2006. № 2. С. 46.
4. *Опарин В.В., Тулупьев А.Л.* Синтез графа смежности с минимальным числом ребер: формализация алгоритма и анализ его корректности // Труды СПИИРАН. СПб.: Наука, 2009. Вып. 11. С. 142–157.
5. *Опарин В.В., Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Матроидное представление семейства графов смежности над набором фрагментов знаний // Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. 2010. Вып. 4. С. 73–76
6. *Сироткин А.В.* Модели, алгоритмы и вычислительная сложность синтеза согласованных оценок истинности в алгебраических байесовских сетях // Информационно-измерительные и управляющие системы. 2009. №11. С. 32–37.
7. *Сироткин А.В., Тулупьев А.Л., Фильченков А.А., Пащенко А.Е., Тулупьева Т.В., Мусина В.Ф.* Особенности вероятностных графических моделей комплекса «информационная система – персонал» для оценки его защищенности от социоинженерных атак // Научная сессия НИЯУ МИФИ-2012 (30 января–4 февраля 2012 г., Москва). Аннотации докладов. В 3 т. Т. 3: Стратегические информационные технологии в атомной энергетике и промышленности. Проблемы информационной безопасности в системе высшей школы. Экономические и правовые проблемы инновационного развития атомной отрасли. Образование в национальном исследовательском ядерном университете. М.: НИЯУ МИФИ, 2012. С. 80.
8. *Суворова А.В., Пащенко А.Е., Тулупьева Т.В.* Оценка характеристик сверхкороткого временного ряда по гранулярным данным о рекордных интервалах между событиями // Труды СПИИРАН. 2010. Вып. 12. С. 170–181.
9. *Суворова А.В., Тулупьев А.Л., Пащенко А.Е., Тулупьева Т.В., Красносельских Т.В.* Анализ гранулярных данных и знаний в задачах исследования социально значимых видов поведения // Компьютерные инструменты в образовании. №4. 2010. С. 30–38.
10. *Тулупьев А.Л.* Алгебраические байесовские сети: глобальный логико-вероятностный вывод в деревьях смежности: Учеб. пособие. СПб.: СПбГУ; ООО Издательство «Анатолия», 2007. 40 с. (Сер. Элементы мягких вычислений).
11. *Тулупьев А.Л.* Алгебраические байесовские сети. Логико-вероятностный подход к моделированию баз знаний с неопределенностью. СПб.: СПИИРАН, 2000. 292 с.
12. *Тулупьев А.Л.* Алгебраические байесовские сети: локальный логико-вероятностный вывод: Учеб. пособие. СПб.: СПбГУ; ООО Издательство «Анатолия», 2007. 80 с. (Сер. Элементы мягких вычислений).
13. *Тулупьев А.Л.* Адциклические алгебраические байесовские сети: логико-вероятностный вывод // Нечеткие системы и мягкие вычисления. Научный журнал Российской ассоциации нечетких систем и мягких вычислений. 2006. Том 1, № 1. С. 57–93.

14. *Тудупьев А.Л.* Байесовские сети: логико-вероятностный вывод в циклах. СПб.: Изд-во С.-Петербургского ун-та, 2008. 140 с. (Элементы мягких вычислений).
15. *Тудупьев А.Л.* Согласованность данных и оценка вероятности альтернатив в цикле стохастических предпочтений // Известия высших учебных заведений: Приборостроение. 2009. №7. С. 3–8
16. *Тудупьев А.Л., Абрамян А.К.* Логико-вероятностный вывод в направленном БСД-цикле // Труды СПИИРАН. 2007. Вып. 4. С. 87–118.
17. *Тудупьев А.Л., Азаров А.А., Пащенко А.Е.* Информационные модели компонент комплекса «Информационная система – персонал», находящегося под угрозой социоинженерных атак // Труды СПИИРАН. 2010. Вып. 3 (14). С. 50–57.
18. *Тудупьев А.Л., Николенко С.И., Сироткин А.В.* Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006. 607 с.
19. *Тудупьев А.Л., Сироткин А.В., Николенко С.И.* Байесовские сети доверия: логико-вероятностный вывод в ациклических направленных графах. СПб.: Изд-во С.-Петербург. ун-та, 2009, 400 с.
20. *Тудупьев А.Л., Столяров Д.М., Ментюков М.В.* Представление локальной и глобальной структуры алгебраической байесовской сети в Java-приложениях // Труды СПИИРАН. 2007. Вып. 5. СПб.: Наука, 2007. С. 71–99.
21. *Тудупьев А.Л., Фильченков А.А.* Иерархия глобальных структур в задачах автоматического обучения алгебраических байесовских сетей // VI Международная научно-практическая конференция «Современные информационные технологии и IT-образование». (12–14 декабря 2011 г., Москва.) Сборник научных трудов. М.: 2011. С. 490–496.
22. *Тудупьев А.Л., Фильченков А.А., Вальтман Н.А.* Алгебраические байесовские сети: задачи автоматического обучения // Информационно-измерительные и управляющие системы. 2011. № 11, т. 9. С. 57–61.
23. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик // Труды СПИИРАН. 2010. Вып. 1 (12). С. 119–133.
24. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик-собственников // Труды СПИИРАН. 2010. Вып. 3 (14) С. 150–169.
25. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи клик владений // Труды СПИИРАН. 2010. Вып. 2 (13). С. 119–133.
26. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи клик-собственников владений // Труды СПИИРАН. 2010. Вып. 4 (15). С. 193–212.
27. *Фильченков А.А.* Алгоритмы построения третичной структуры алгебраической байесовской сети // Труды СПИИРАН. 2011. Вып. 17. С. 197–218.
28. *Фильченков А.А.* Математическое моделирование диагностической модели защищенности информационной системы на основе комбинирования неполной и неточной аналитической информации // VII Санкт-Петербургская межрегиональная конференция «Информационная безопасность регионов России (ИБРР-2011)». (26–28 октября 2011 г., Санкт-Петербург.) Материалы конференции. СПб.: СПОИСУ, 2011. С. 175–176.
29. *Фильченков А.А., Тудупьев А.Л.* Алгоритм выявления ацикличности первичной структуры алгебраической байесовской сети по ее четвертичной структуре // Труды СПИИРАН. 2011. Вып. 4(19). С. 128–145.
30. *Фильченков А.А., Тудупьев А.Л.* Анализ циклов в минимальных графах смежности алгебраических байесовских сетей // Труды СПИИРАН. 2011. Вып. 17. С. 151–173.

31. *Фильченков А.А., Тулупьев А.Л.* Понятие торакса в применении к исследованию графов смежности алгебраических байесовских сетей // Труды СПИИРАН. 2011. Вып. 16. С. 186–205.
32. *Фильченков А.А., Тулупьев А.Л.* Структурный анализ систем минимальных графов смежности // Труды СПИИРАН. 2009. Вып. 11. С. 104–127.
33. *Фильченков А.А., Тулупьев А.Л.* Третьичная структура алгебраической байесовской сети // Труды СПИИРАН. 2011. Вып. 18. С. 164–187.
34. *Фильченков А.А., Тулупьев А.Л.* Совпадение множеств минимальных и нeredуцируемых графов смежности над первичной структурой алгебраической байесовской сети // Вестник Санкт-Петербургского государственного университета. Серия 1. Математика. Механика. Астрономия. 2012. Вып. 2. С. 65–74.
35. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Мощность множества минимальных графов смежности // Труды СПИИРАН. 2010. Вып. 4 (15). С. 136–161.
36. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Особенности анализа вторичной структуры алгебраической байесовской сети // Труды СПИИРАН. 2010. Вып. 1 (12). С. 97–118.
37. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Структурный анализ клик минимальных графов смежности // Вестн. Тверск. гос. ун-та. Сер.: Прикладная математика. 2011. №20. С. 139–151.

**Фильченков Андрей Александрович** — аспирант кафедры информатики математико-механического факультета С.-Петербургского государственного университета (СПбГУ), младший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики СПИИРАН. Область научных интересов: автоматическое обучение вероятностных графических моделей. Число научных публикаций — 50. [aaafil@mail.ru](mailto:aaafil@mail.ru), СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450. Научный руководитель — А.Л. Тулупьев.

**Filchenkov Andrey Alexandrovich** — PhD student of Computer Science Department, SPbGU, junior researcher, Theoretical and Interdisciplinary Computer Science Laboratory, SPIIRAS Research area: machine learning of probabilistic graphical models. The number of publications — 50. [aaafil@mail.ru](mailto:aaafil@mail.ru), SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450. Scientific advisor — A.L. Tulupjev.

**Фроленков Константин Владиславович** — студент кафедры информатики Санкт-Петербургского государственного университета (СПбГУ). Область научных интересов: обучение алгебраических байесовских сетей. [frolenk@mail.ru](mailto:frolenk@mail.ru). Научный руководитель — А.Л. Тулупьев.

**Frolenkov Konstantin Vladislavovich** — a student of Computer Science Department, SPbGU. Research area: machine learning of Algebraic Bayesian Networks. [frolenk@mail.ru](mailto:frolenk@mail.ru). Scientific advisor — A.L. Tulupjev.

**Тулупьев Александр Львович** — д.ф.-м.н., профессор; заведующий лабораторией теоретических и междисциплинарных проблем информатики СПИИРАН, доцент кафедры информатики математико-механического факультета С.-Петербургского государственного университета (СПбГУ). Область научных интересов: представление и обработка данных и знаний с неопределенностью, применение методов математики и ин-

форматики в социокультурных исследованиях, применение методов биостатистики и математического моделирования в эпидемиологии, технология разработки программных комплексов с СУБД. Число научных публикаций — 250. ALT@iias.spb.su, www.tulupyev.spb.ru; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-3337, факс +7(812)328-4450.

**Tulupyev Alexander Lvovich** — PhD in Computer Science, Dr. of Sc. Professor; Head of Theoretical and Interdisciplinary Computer Science Laboratory, SPIIRAS, Associate Professor of Computer Science Department, SPbSU. Research area: uncertain data and knowledge representation and processing, mathematics and computer science applications in socio-cultural studies, biostatistics, simulation, and mathematical modeling applications in epidemiology, data intensive software systems development technology. Number of publications — 250. ALT@iias.spb.su, www.tulupyev.spb.ru; SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450.

**Поддержка исследования.** Работа выполнена при финансовой поддержке РФФИ, проекты № **09-01-00861-а** «Методология построения интеллектуальных систем поддержки принятия решений на основе баз фрагментов знаний с вероятностной неопределенностью» и № **12-01-00945-а** «Развитие теории алгебраических байесовских сетей и родственных им логико-вероятностных графических моделей систем знаний с неопределенностью».

Рекомендовано ТимПИ СПИИРАН, зав. лаб. А.Л. Тулупьев, д.ф.-м.н., проф.  
Работа поступила в редакцию 01.05.2012.

## РЕФЕРАТ

### **Фильченков А.А., Фроленков К.В., Тулупьев А.Л. Устранение циклов во вторичной структуре алгебраической байесовской сети на основе анализа ее четвертичной структуры.**

Алгебраическая байесовская сеть (АБС) — одна из логико-вероятностных графических моделей баз фрагментов знаний с неопределенностью. Первичной структурой АБС называется набор фрагментов знаний, а вторичной структурой — граф смежности, построенный над первичной.

Алгоритмы глобального логико-вероятностного вывода АБС могут меняться при условии ацикличности её вторичной структуры. Однако не над любой первичной структурой АБС возможно построить ацикличную вторичную структуру (такие первичные структуры называются *циклическими*). Если же над первичной структурой возможно построение ациклической вторичной структуры (такие первичные структуры называются *ациклическими*), то множество ациклических вторичных структур, которые возможно построить над ней, совпадает с множеством минимальных графов смежности.

Известен метод преобразования циклической первичной структуры к ациклической, но он применим лишь для графов смежности с обособленными циклами. Цель работы — предложить новые методы преобразования циклической первичной структуры АБС к ациклической, основывающиеся на структурной теореме о циклах минимальных графов смежности. Теорема описывает необходимые и достаточные условия цикличности первичной структуры.

Рассмотрен известный метод преобразования обособленного цикла в графе смежности в цепь смежности. Однако его использование ограничено, кроме того, оно может привести к неэффективности результирующей АБС из-за возрастания количества атомов в составе некоторых из ее фрагментов знаний. Приведена формулировка теоремы о циклах минимальных графах смежности, которая утверждает, что минимальный граф смежности ациклический тогда и только тогда, когда четвертичная структура АБС не содержит особых циклов (небратских полусиблинговых простых циклов, НППЦ).

Предложен новый метод, основывающийся на добавлении нового фрагмента знаний, устраняющего каждый НППЦ. Приведен пример использования данного алгоритма, который показывает, что данный алгоритм может эффективно применяться только к классу наборов исходных данных, поскольку в общем случае требует многоэтапной перестройки и анализа структур АБС.

Предложена модификация описанного метода, которая позволяет повысить его эффективность за счет отказа от перестройки. Показано, что в некоторых случаях работа модифицированного алгоритма также может привести к неэффективности результирующей АБС.

Исследования в устранении циклов АБС сдерживаются отсутствием формальных критериев сравнения результирующих ациклически первичных структур и недостаточной степенью формализованности задачи устранения циклов.

## SUMMARY

### *Filchenkov A.A., Frolenkov K.V., Tulupyev A.L. Algebraic Bayesian Network Secondary Structure Cycles Elimination Based on its Quaternary Structure Analysis.*

Algebraic Bayesian network (ABN) is one of the logical and probabilistic graphical models of knowledge patterns bases. The set of knowledge patterns is called the primary structure. The secondary structure is a join graph built over given primary one.

The algorithms for global ABN probabilistic logical inference can only be applied if the ABN secondary structure is acyclic. However there are ABN primary structures which do not permit to synthesize an acyclic secondary structure (such primary structures are called cyclic). If it is possible to synthesize an acyclic secondary structure on the primary structure (such primary structures are called acyclic), then the set of acyclic secondary structures coincides with the set of minimal join graphs.

Method for transforming a cyclic secondary structure into acyclic one is known, but it is only applicable to join graph containing solitary cycles. The purpose of the paper is to offer new methods for transforming cyclic ABN primary structures into acyclic ones based on the structural theorem on minimal join graphs cycles. The theorem shows necessary and sufficient conditions for ABN primary structure cyclicity.

The known method for transforming a solitary cycle in join graph into a join chain is considered. However its usage is limited, moreover it can result in inefficiency of derived ABN due to increase of quantity of atoms in certain of its knowledge patterns. The wording of the theorem on minimal join graphs cycles, which claims that a minimal join graph is acyclic if and only if the ABN quaternary structure contains no special cycles (unbrotherly halfsibling prime cycles, UHPC), is given.

A new method based on a new knowledge pattern addition with the view of UHPC removal is offered. An example of the method application is shown which displays the algorithm being efficiently applicable only to a class of input data since in the general case it requires multistep ABN structures rebuilding and analysis.

A modification of the above method is offered, which allows increasing its efficiency due to escaping the rebuilding. Shown, that in certain cases application of the modified algorithm can also result in inefficiency of derived ABN.

Researches of ABN cycles removal are held back by absent of formal criteria for comparison derived acyclic primary structures and lack of formalization of removal cycles problem.