УДК 004.8

А.А. Фильченков, А.Л. Тулупьев, А.В. Сироткин

КОМПАРАТИВНЫЙ АНАЛИЗ КЛИК МИНИМАЛЬНЫХ ГРАФОВ СМЕЖНОСТИ АЛГЕБРАИЧЕСКИХ БАЙЕСОВСКИХ СЕТЕЙ

Фильченков А.А, Тулупьев А.Л., Сироткин А.В. **Компаративный анализ клик минимальных графов смежности алгебраических байесовских сетей.**

Аннотация. Алгебраические байесовские сети представляют собой логико-вероятностную графическую модель систем знаний с неопределенностью и могут быть применимы в обработке статистических данных и машинном обучении. Важную роль в их работе играет вторичная структура, представляемая в виде графа смежности. Данная статья вводит классификацию клик минимальных графов смежности в зависимости от числа их детей, а также числа вхождения в них числа особых ребер. Получено восемь различных типов клик, для которых были получены и обоснованы оценки числа зависимых от них компонент (феодов и жил).

Ключевые слова: алгебраические байесовские сети, вторичная структура, машинное обучение, вероятностно-графические модели систем знаний, глобальная структура.

Filchenkov A.A, Tulupyev A.L., Sirotkin A.V. **Algebraic Bayesian networks minimal join graph cliques comparative analysis.**

Abstract. Algebraic Bayesian networks are probabilistic-logic graphic models of knowledge systems with uncertainty and can be applied to statistic data processing and machine learning. Secondary structure usually represented as join graph is crucial for its work. The article represents minimal join graphs cliques classification on the number of their vertexes and the number of special edges they contain. Eight different clique types are designed and estimations of number of components depended on them (feuds and sinews) are obtained and proven.

Keywords: algebraic Bayesian networks, secondary structure, machine learning, probabilistic graphical knowledge models, global structure.

1. Введение. Алгебраические байесовские сети (АБС) представляют собой логико-вероятностную графическую модель систем знаний с неопределенностью [1–2, 15, 25] и могут быть применимы в обработке статистических данных и машинном обучении [9, 16]. От других вероятностных графических моделей, в первую очередь от байесовских сетей доверия, АБС отличает возможность работы с интервальными оценками вероятностей, заданных на множестве пропозициональных формул. АБС осуществляют логико-вероятностные априорный и апостериорный выводы (как локальные, так и глобальные), а также позволяют поддерживать непротиворечивости сети (как локальную, так и глобальную).

Выделяют первичную и вторичную структуру АБС [7, 15, 17]. Первичная структура представляется собой максимальные фрагменты знаний, а вторичная структура характеризует связь между фрагментами знаний и обычно представляется в виде графа над ними (так называемого графа смежности).

Вторичная структура АБС является важным объектом для изучения, поскольку именно она обусловливает как эффективность работы алгоритмов вывода (с точки зрения времени работы), так и саму возможность их применения, поскольку для некоторых вторичных структур работа АБС затруднена [11–14, 16]. Особую роль в представлении вторичной структуры играют минимальные графы смежности, которые являются особым подвидом графов смежности [4–8]. Именно в данном множестве, которое совпадает с множеством нередуцируемых графов смежности [3], содержатся все ациклические графы смежности, которые являются самыми эффективными для работы АБС [10], если они вообще существуют для данного набора максимальных фрагментов знаний.

Предложены схемы алгоритмов по синтезу данного множества. Одновременно с этим там же было предложено улучшение алгоритма, основанное на рассмотрении различных типов клик [18]. Дальнейшие результаты анализа клик минимальных графов смежности [18] позволяют углубить классификацию клик для получения новых улучшений алгоритмов синтеза множества минимальных графов смежности.

Цель данной работы — предложить расширенную классификацию клик АБС, учитывающую последние результаты их анализа.

2. Основные определения и обозначения. *Граф* — пара $\langle V, E\rangle$, где V — множество вершин графа, а E — множество ребер, каждое из которых является неупорядоченной парой (v_i, v_j), $i \neq j$, v_i, $v_j \in V$. Для удобства будем через V и E обозначать функции от графа, возвращающие множество его вершин и множество его ребер соответственно:

$$V(G') = V';\ E(G') = E', \qquad \text{где } G' = \langle V', E'\rangle.$$

Для внесения ясности определим: $\subseteq$ — (нестрогое) включение:

$$(A \subseteq B) \Leftrightarrow \big((\forall a \in A) a \in B\big).$$

В противоположность этому $\subset$ — строгое включение:

$$(A \subset B) \Leftrightarrow \big((\forall a \in A, a \in B) \exists b \in B : b \notin A\big).$$

Алфавитом будем называть множество атомарных пропозициональных формул

$$A = \{x_1, \dots, x_n\}.$$

Словом тогда будем называть некое подмножество алфавита.

Первичная структура, или множество главных конъюнктов максимальных ФЗ, вошедших в байесовскую сеть, — такое множество слов $V^* = \{V_i \subseteq A\}_{i=1}^{i=m}$, что

- оно не содержит несобственное подмножество алфавита:
$$V_i \neq A, V_i \neq \emptyset;$$
- никакое слово полностью не содержит никакое другое слово:
$$\forall\, i \neq j\ (V_i \not\subset V_j) \text{ и } (V_j \not\subset V_i).$$

Вторичная структура, или граф максимальных фрагментов знаний — ненаправленный граф, вершины которого соответствуют элементам множества главных конъюнктов максимальных ФЗ, вошедших в АБС, а ребра возможны только между теми вершинами, для которых пересечение соответствующих МФЗ непусто.

С этого момента под словом «граф» мы будем понимать именно граф максимальных фрагментов знаний, если не оговорено обратное.

Введем понятия веса для вершины, для ребра и для подграфа. *Вес* $W(V_i)$ *вершины* $V_i \in V(G)$ — множество атомов алфавита, вошедших в V_i:
$$W(V_i) = \{x_i | x_i \in V\}.$$

Вес $W(\{V_i, V_j\})$ ребра $\{V_i, V_j\} \in E(G)$, графа G определяется как пересечение весов тех вершин, которые соединены этим ребром:
$$W(\{V_i, V_j\}) = W(V_i) \cap W(V_j).$$

Вес $W(H)$, *подграфа* $H \subseteq G$ — наибольшее по включению слово, которое входит в веса всех его вершин:
$$W(H) = \bigcap\nolimits_{V \in H} W(V).$$

Под *магистральным путем* $B: V_b \dashrightarrow V_e$ от вершины V_b до вершины V_e, пересечение весов которых непусто, будем понимать такой путь от вершины V_b до вершины V_e, что вес любой принадлежащей ему вершины содержит пересечение весов начальной и конечной вершин:
$$B: V_b \dashrightarrow V_e = P: V_b \rightsquigarrow V_e,$$
такой, что
$$\forall V_i \in B W(V_b) \cap W(V_e) \subset W(V_i).$$

Граф *магистрально связен*, если между каждой парой несовпадающих вершин, веса которых содержат общие элементы, существует магистральный путь. Будем обозначать множество магистрально связных графов как $\mathbf{BCG}$.

Благодаря введенным понятиям, *граф смежности* определяется как магистрально связный граф МФЗ. При этом он не обязательно связен (например, граф из двух вершин ab и cd, в котором нет ребер, тем не менее, является магистрально связным).

Дерево смежности — это граф смежности, являющийся деревом.

Минимальный граф смежности — граф смежности, число ребер в котором минимально (рис. 1).

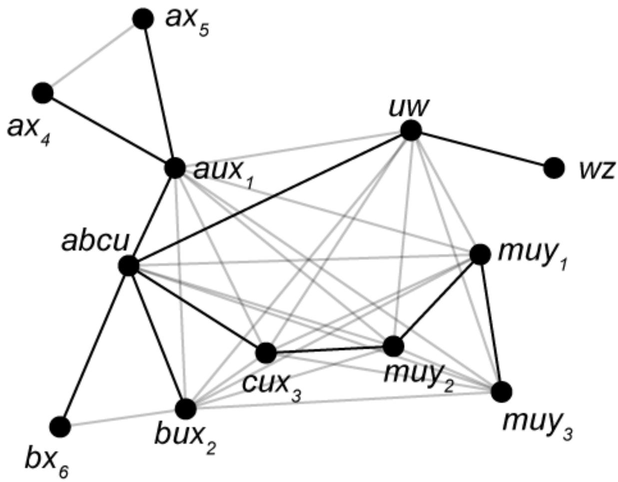

Рис. 1. Максимальный граф смежности и минимальный граф смежности. Графы смежности, построенные на вершинах $abcu$, ax_4, ax_5, aux_1, bux_2, cux_3, muy_1, muy_2, muy_3, uw, wz. Минимальный граф смежности состоит из черных ребер, максимальных граф смежности — из черных и серых.

Максимальный граф смежности $G_{\max}$ — наибольший по числу ребер граф смежности (рис. 1). Так как в графе МФЗ возможны не все ребра, а только те, которые соединяют вершины, пересечение весов которых непусто, то максимальный граф смежности вовсе необязательно совпадает с полным подграфом. Доказано [18], что для заданного множества вершин существует и при этом единственный максимальный граф смежности.

3. Сужение, граф клик и сжатие. *Сужение* $G \downarrow U$ ненаправленного графа G на слово U — это ненаправленный граф, в который входят только те вершины и ребра исходного графа G, веса которых содержат или равны U:

$$G \downarrow U = \langle \{V_i | V_i \in V(G), U \subseteq W(V_i)\}, \{E_i | E_i \in E(G), U \subseteq W(E_i)\} \rangle.$$

Значимое слово графа G — слово, являющееся весом какого-либо ребра графа G.

Значимое сужение $G \downarrow U$ — сужение графа G на значимое для графа G слово U.

Клика веса U — значимое сужение $G_{\max} \downarrow U$ (рис. 2).

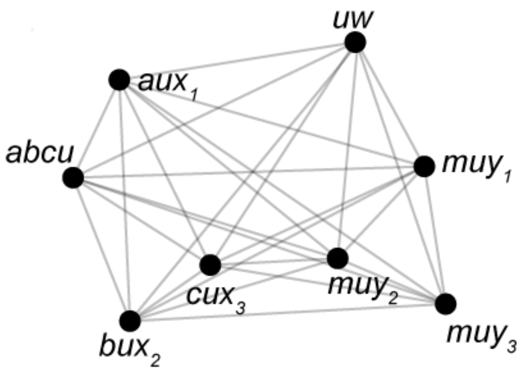

Рис. 2. Клика a.
Клика, образованная сужением максимального графа с рис. 1 на вес u.

Доказано [18], что если сужение G на произвольное непустое слово непусто, то оно является полным подграфом. Таким образом, введенная нами клика является кликой в общепринятом смысле слова, т.е. полным подграфом. Будем обозначать множество всех клик графа $G_{\max}$ через Clique**.**

Граф клик — направленный граф, вершинами которого являются клики из множества Clique (рис. 3). 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:

$$G_{\text{Clique}} = \langle \text{Clique}, E_{\text{Clique}} \rangle, \text{ где } P, Q \in \text{Clique}, (P, Q) \in E \Leftrightarrow$$
$$\Leftrightarrow Q \subset P \text{ и } \nexists R \in \text{Clique}: Q \subset R \subset P.$$

Доказано [18], что граф клик не содержит направленных циклов.

Будем называть клики, в которые входят ребра из клики C в графе клик, *сыновьями клики* C.

Сильное сужение $G \not\downarrow U$ — значимое сужение $G \downarrow U$, из которого удалили все ребра веса U (рис. 4):

$$G \not\downarrow U = \langle \{V_i | V_i \in V(G), U \subseteq W(V_i)\}, \{E_i | E_i \in E(G), U \subset W(E_i)\} \rangle.$$

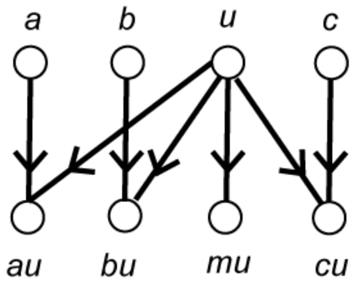

Рис. 3. Граф клик.
Граф клик для графа на рис. 1.

Сильное сужение графа $G_{\max} \mathrel{\rlap{\downarrow}{\,\,\,}} U$ представляет собой компоненты связности, на которые разбивается сужение $G_{\max} \downarrow U$ путем удаления ребер веса U. Такие компоненты будем называть владениями (рис. 5).

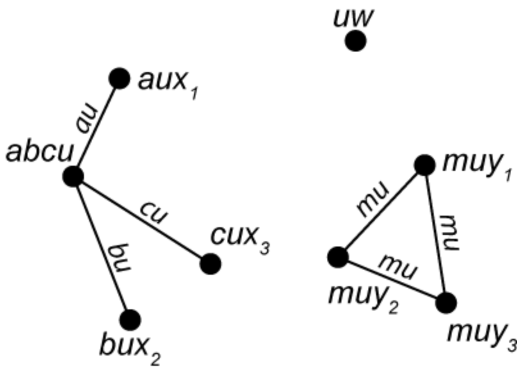

Рис. 4. Сильное сужение.
Сильное сужение графа с рис. 1 на вес u. Оно состоит из 3–х владений.

Доменная вершина D_U клики U — вершина, принадлежащая клике U и не принадлежащая ни одному из ее сыновей (рис. 5).

Вассал V_U клики U — множество вершин, принадлежащих какому-либо одному сыну U по графу клик (рис. 5).

Два вассала V_U^i и V_U^j называются *братьями* (будем обозначать это через $\leftrightarrow$), 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* (будем обозначать это через $\leftrightarrow$), 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 $V_U^{w_2}, \dots, V_U^{w_{n-1}}$ — брат $V_U^{w_n}$, а $V_U^{w_n}$ — брат V_U^j:

$$V_U^i \leftrightarrow V_U^j \Leftrightarrow \exists\{V_U^{w_1}, \dots, V_U^{w_n}\}: V_U^b \leftrightarrow V_U^{w_1}; V_U^{w_n} \leftrightarrow V_U^e \text{ и}$$
$$\forall i < n \;\; V_U^{w_i} \leftrightarrow V_U^{w_{i+1}}.$$

Полусиблинговый путь между двумя полусиблинговыми вассалами V_U^i и V_U^j — такой набор вассалов $\{V_U^{w_1}, \dots, V_U^{w_n}\}$, что $V_U^b \leftrightarrow V_U^{w_1}$; $V_U^{w_n} \leftrightarrow V_U^e$ и $\forall i < n \;\; V_U^{w_i} \leftrightarrow V_U^{w_{i+1}}$.

Братство B_U клики U — непустой набор вассалов $\{V_U^1, \dots, V_U^l\}$ клики U, такой, что с каждым вассалом в братство входят все его полусиблинги и только они (рис. 5):

$$B_U = \{V_U^i \big| (V_U^i \in B_U) \& (V_U^i \leftrightarrow V_U^j) \Rightarrow V_U^j \in B_U; V_U^i, V_U^j \in B_U \Rightarrow$$
$$\Rightarrow V_U^i \leftrightarrow V_U^j\}.$$

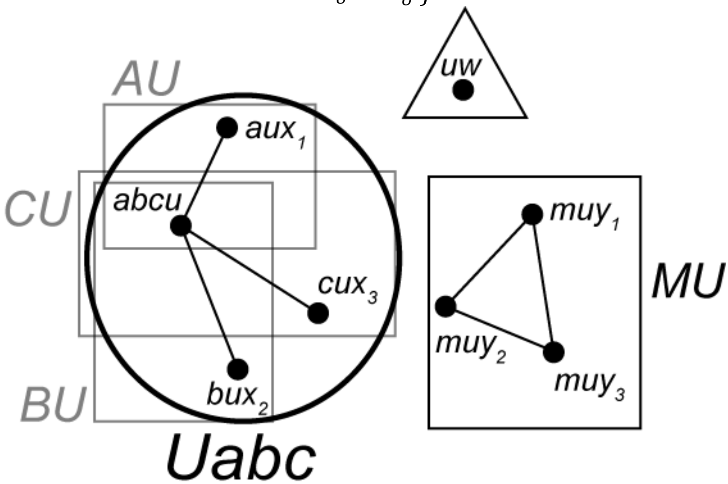

Рис. 5. Владения.

Треугольником обозначена доменная вершина uw; прямоугольником обозначен вассал MU, состоящий из вершин muy_1, muy_2 и muy_3; овалом обозначено братство $Uabc$, состоящее из выделенных полупрозрачными прямоугольниками вассалов AU, BU и CU; вассалы AU, BU и CU приходятся друг другу братьями.

Теорема (классификация владений) [24]. Любое владение любого сильного сужения $G \not\downarrow U$ является либо доменной вершиной, либо вассалом, либо братством U.

В [23] приводятся три определения *сжатия*, задающие эту операцию на различных множествах:

1. Сжатие σ_U компоненты связности $P_U^i \subseteq G \not\downarrow U$ в вершину f_i — отображение на множестве подмножеств вершин, сопоставляющее множеству вершин P_U^i вершину f_i (рис. 6);

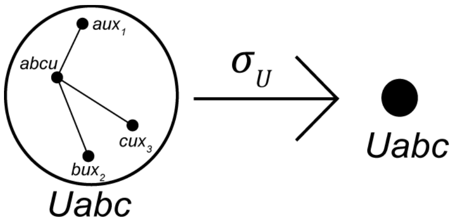

Рис. 6. Сжатие владения.
Сжатие братства $Uabc$ в вершину.

2. Сжатие σ_U множества ребер

$$E_{i,j} \subset (G \downarrow U) \backslash (G \not\downarrow U),$$

соединяющих владения P_U^i и P_U^j в ребро $e_{i,j}$, — отображение на множестве подмножеств ребер, сопоставляющее множеству ребер $E_{i,j}$ ребро $e_{i,j}$, соединяющее вершины $f_i = \sigma_U(P_U^i)$ и $f_j = \sigma_U(P_U^j)$ и имеющее кратность, равную $|E_{i,j}|$ (рис. 7);

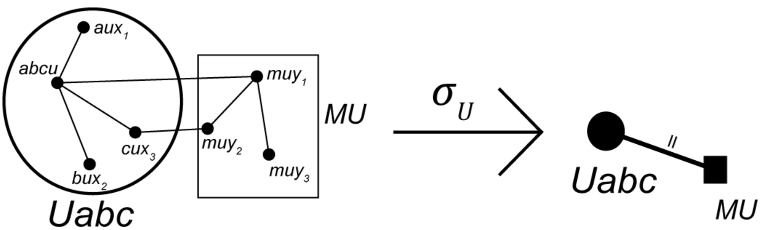

Рис. 7. Сжатие множества ребер.
Сжатие множества ребер, соединяющих владения $Uabc$ и MUв ребро кратности 2, соединяющее соответствующие вершины.

3. *Сжатие* σ_U графа смежности G в граф K_U — отображение на множестве графов, сопоставляющее графу смежности G, являющему графом смежности, граф K_U, вершинами которого являются владения сильного сужения $G \Downarrow U$, а ребро между двумя вершинами f_1 и f_2 графа K_U существует, если существует ребро в графе G между вершинами, f_1 и f_2, принадлежащими соответствующим владениям P_U^1 и P_U^2. Кратность такого ребра (f_1, f_2) равна числу всех ребер, соединяющих вершины изP_U^1 и P_U^2 (рис. 8).

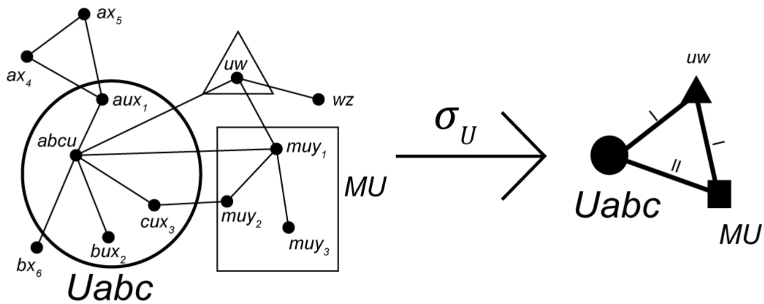

Рис. 8. Сжатие графа.

Феод — f_i из определения выше — вершина, получившаяся путем сжатия какого-то владения (рис. 6).

Курия веса U — граф K_U из определения сжатия — ненаправленный граф с кратными ребрами, полученный сжатием значимого сужения $G \downarrow U$ (рис. 8).

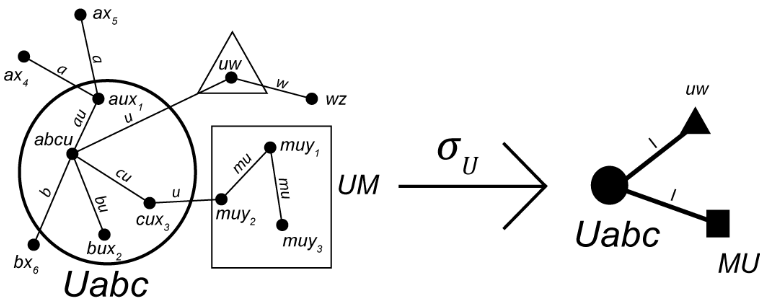

Рис. 9. Оммаж.

Сжатие минимального графа, построенного над теми же вершинами, что и на рис. 8, до оммажа.

Оммаж H_U — курия K_U, являющаяся деревом, все ребра которой имеют кратность, равную единице (рис. 9). Доказано[23], что любой минимальный граф смежности сжимается до оммажа.

4. Классификация клик.

Пояснение. Теорема о множестве минимальных графов смежности говорит о том, что для того, чтобы перебрать все возможные минимальные графы смежности, достаточно перебрать все возможные комбинации жил по одной для каждой клики. Заметим, что бывают клики, для которых существует всего одна жила. Было бы удобно не выбирать такие жилы каждый раз для всех комбинаций жил, а объединить их в одном дереве, что уменьшило бы число элементов, по которым ведется перебор, а, значит, ускорило бы работу программы.

Определение 1. *Собственное ребро клики* — ребро, принадлежащее клике, вес которого совпадает с весом клики.

Сильное сужение на вес U получается путем удаления из сужения на тот же вес собственных ребер веса U. Собственные ребра названы так, потому что они не принадлежат никакому вассалу клики. Собственные ребра — это ребра, которые сжимаются до ребер соответствующего оммажа.

Замечание 1. Собственные ребра клики и только они соединяют следующие элементы:

- две доменные вершины;
- доменную вершину с вершиной какого-либо сына клики;
- вершины, принадлежащие разным сыновьям клики.

Утверждение 1. Ребра, соединяющие вершины различных владений одной клики, являются собственными.

Доказательство. Ребра, соединяющие вершины различных владений — это те ребра, которые не входят в строгое сужение, потому что они соединяют компоненты связности строго сужения, а, значит, их вес в точности равен весу сужения, т.е. это собственные ребра.

Утверждение 2. Если клика сжимается до одного феода, то имеет ровно одну, причем пустую жилу.

Доказательство. Курия, получающаяся вследствие такого сжатия, состоит ровно из одной вершины, и, соответственно, не имеет ребер, а, значит, любая жила будет пустой. Всего существует ровно одна пустая жила.

Пояснение. Будем классифицировать клики по двум признакам: 1) по наличию у них детей и 2) по числу собственных ребер.

Определение 2. По числу собственных ребер множество клик можно поделить на следующие:

- *безреберные клики* — C_0 — клики, у которых нет собственных ребер;
- *однореберные клики* — C_1 — клики, у которых ровно одно собственное ребро;
- *многореберные клики* — C_n — клики, у которых более одного собственного ребра.

С формальной точки зрения, безреберные клики не являются кликами по определению, поскольку введенная нами клика — это значимое сужение, т.е. сужение на вес какого-либо ребра. Так как у безреберных клик нет собственных ребер, то в графе вообще не существует ребер с их весом, т.е. сужение, которое безреберные клики представляют собой, не является значимым, а потому и не является кликой. Однако, с точки зрения полноты классификации, мы должны учитывать и их тоже. Кроме того, понятие псевдоклики (которое мы введем далее) используется для построения *замкнутого снизу* графа клик и позволяет выявлять возможность существования деревьев смежности по указанному графу клик.

Определение 3. По наличию детей множество клик можно поделить на следующие (табл. 1):

- *бездетные* — C^- — клики, у которых нет детей;
- *родительские* — C^+ —клики, у которых есть дети.

Табл. 1. Классификация клик

По ребрам \ По детям	Бездетные	Родительские
Безреберные	C_0^-	C_0^+
Однореберные	C_1^-	C_1^+
Многореберные	C_n^-	C_n^+

Помимо рассмотренных признаков важно также дополнительно ввести классификацию по числу феодов, в которые они сжимаются

Определение 4. По числу феодов, в которые сжимаются клики, они делятся на:

- *моноклики* — клики, сжимающиеся до одного феода.
- *стереоклики* — клики, сжимающиеся до более чем одного феода.

Это несколько отличается от использованных в работе [18] терминологии и вызвано тем, что в контексте расширенной системы терминов классификации, которую мы применяем здесь, подобные названия будут более понятны.

Утверждение 3. Если у клики есть дети, то число ее вершин больше двух.

Доказательство. Так как у клики есть сын, то она имеет хотя бы две вершины (иначе клика состояла бы только из одной вершины). Так как клика не может совпадать с сыном, то в ней должны быть еще какие-нибудь вершины, поэтому их больше двух.

Определение 5. *Псевдоклика* — C_0^- — сужение, не имеющее собственных ребер и не имеющее детей.

Утверждение 4. Псевдоклика состоит ровно из одной вершины, а потому, формально, не является кликой.

Доказательство. Если бы у псевдоклики имелись две вершины, то они должны быть соединены хотя бы одним ребром, которое являлось бы собственным, так как детей у псевдоклики нет.

Определение 6. *Моноклика-0* — C_0^+ — сужение, не имеющее собственных ребер, но имеющее детей.

Утверждение 5. Моноклика-0 имеет всего один феод.

Доказательство. Действительно, если бы моноклика-0 имела несколько феодов, они были бы соединены ее собственными ребрами, которых у нее нет.

Определение 7. *Биклика* — C_1^- — клика, имеющее ровно одно собственное ребро и не имеющее детей.

Утверждение 6. Биклика состоит ровно из двух вершин.

Доказательство. Так как у биклики нет сыновей, все входящие в нее ребра — ее собственные, а значит, собственные ребра клик образуют полный граф. Поскольку одно ребро может быть только у полного графа из двух вершин, то в биклике всего две вершины.

Определение 8. *Моноклика-1* — C_1^+ — клика, имеющая ровно одно собственное ребро и имеющая детей.

Утверждение 7. Моноклика-1 сжимается в один феод и состоит из двух вассалов, каждый из которых имеет по k вершин и которые пересекаются по $k - 1$ вершине.

Доказательство. Докажем сначала, что моноклика-1 сжимается до одного феода. Пойдем от противного: пускай какая-то моноклика-1 сжимается до хотя бы двух феодов, т.е. состоит из хотя бы двух владений. Так как у моноклики-1 есть дети, которые состоят не менее чем из двух вершин, то хотя бы одно из этих владений имеет не менее двух вершин, и существуют как минимум два собственных ребра, которые соединяют эти две вершины с какой-либо вершиной какого-либо из других владений. Итак, моноклика-1 сжимается до одного феода. Этот феод должен быть братством (так как только в этом случае у клики будут собственные ребра), и должна существовать всего одна пара вершин, соединенная собственным ребром, т.е. братство состоит всего из двух владений с одинаковым числом вершин, которые пересекаются по всем вершинам, кроме одной с каждой стороны.

Определение 9. *Бездетная поликлика* (возможно также название *бездетная стереоклика*) — C_n^- — клика, имеющая более одного собственного ребра, но не имеющая детей.

Определение 10. *Родительская поликлика* — C_n^+ — клика, имеющая более одного собственного ребра и имеющая детей.

Определение 11. *Моноклика-*n — mC_n^+ — клика, имеющая более одного собственного ребра и имеющая детей, но состоящая ровно из одного феода.

Определение 12. *Родительская стереоклика* — pC_n^+ — клика, имеющая более одного собственного ребра, имеющая детей и состоящая более чем из одного феода.

На основе введенной классификации и полученных результатов мы можем составить сводную таблицу с характеристиками различных видов клик (табл. 2).

Табл. 2. **Характеристики различных видов клик**

Сужение	Обозначение	Является ли кликой	Есть ли дети	Число собственных ребер	Число вершин	Число феодов	Число жил
Псевдоклика	C_0^-	Нет	Нет	0	1	0	0
Моноклика-0	C_0^+	Нет	Да	0	> 2	1	1
Биклика	C_1^-	Да	Нет	1	2	2	1
Моноклика-1	C_1^+	Да	Да	1	> 2	1	1
Бездетная стереоклика	C_n^-	Да	Нет	> 1	> 2	> 2	> 1
Родительская поликлика	C_n^+	Да	Да	> 1	> 2	$\forall$	$\forall$
Моноклика-n	mC_n^+	Да	Да	> 1	> 2	1	1
Родительская стереоклика	pC_n^+	Да	Да	> 1	> 2	> 1	> 1

Замечание 2. Моноклика-0, моноклика-1 и моноклика-n и только они являются монокликами. Бездетная стереоклика и родительская стереоклика и только они являются стереокликами.

5. Заключение. В статье разработана классификация клик, основанная на числе *собственных ребер* — ребер, вес которых совпадает с весом клики, и числе сыновей клики. Дополнительно произведена классификация по числу феодов, до которых сжимается клика (т.е. по числу ее владений). Выделено шесть различных типов клик, в классификацию попало также два сужения (*псевдоклика* и *моноклика-0*), которые формально не являются кликами, но обладают определенными сходными свойствами.

Для всех типов клик выведены число феодов, до которых они сжимаются, и число жил, которые в них содержатся.

Введенная классификация и полученные результаты развивают теорию алгебраических байесовских сетей (в частности, более раннюю классификацию, рассматривавшую всего два типа клик в зависимости от числа ребер, в них входящих) и являются основой для разработки улучшенных алгоритмов синтеза множества минимальных графов смежности, которые будут учитывать специфику различных типов клик, что позволит ускорить их работу и вывод результатов [19–22].

Литература

1. *Городецкий В.И.* Алгебраические байесовские сети — новая парадигма экспертно-вычислительных систем // Юбилейный сборник трудов институтов Отделения информатики, вычислительной техники и автоматизации РАН. Т. 2. М.: РАН, 1993, С. 120–141.
2. *Городецкий В.И., Тулупьев А.Л.* Формирование непротиворечивых баз знаний с неопределенностью // Известия РАН. Сер. Теория и системы управления. 1997. № 5. С. 33–42.
3. *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.
4. *Павельчук А.В., Тулупьев А.Л., Тотмянина С.А.* Подход к объектно-ориентированному представлению данных алгебраических байесовских сетей в java-коде и реляционных СУБД // Региональная информатика-2008 (РИ-2008). XI Санкт-Петербургская международная конференция. Санкт-Петербург, 22–24 октября, 2008 г.: Материалы конференции / СПОИСУ. СПб.: 2009. С. 68–76.
5. *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.
6. *Тотмянина С.А., Павельчук А.В., Тулупьев А.Л.* Алгебраические байесовские сети: структуры данных в СУБД и Java- коде // Интегрированные модели, мягкие вычисления, вероятностные системы и комплексы программ в искусственном интеллекте. Научно-практическая конференция студентов, аспирантов, молодых ученых

и специалистов (Коломна, 26–27 мая 2009 г.). Научные доклады. В 2-х т. Т. 2. М.: Физматлит, 2009, С. 123–131.
7. *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).
8. *Тулупьев А.Л.* Алгебраические байесовские сети: реализация логико-вероятностного вывода в комплексе java-программ // Труды СПИИРАН. СПб.: Наука, 2009. Вып. 8. С. 191–232.
9. *Тулупьев А.Л.* Алгебраические байесовские сети: система операций глобального логико-вероятностного вывода // Информационно-измерительные и управляющие системы. 2010. № 11. С 65–72.
10. *Тулупьев А.Л.* Ациклические 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.
11. *Тулупьев А.Л.* Байесовские сети: логико-вероятностный вывод в циклах. СПб.: Изд-во С.-Петербургского ун-та, 2008. 140 с. (Элементы мягких вычислений.)
12. *Тулупьев А.Л.* Непротиворечивость оценок вероятностей в алгебраических байесовских сетях. Вестник СПбГУ. Сер. 10. 2009. Вып. 3. С. 144–151.
13. *Тулупьев А.Л.* Преобразование ациклических байесовских сетей доверия в алгебраические байесовские сети // Известия высших учебных заведений: Приборостроение. 2009. № 3. С. 21–23.
14. *Тулупьев А.Л.* Согласованность данных и оценка вероятности альтернатив в цикле стохастических предпочтений // Известия высших учебных заведений: Приборостроение. 2009. № 7. С. 3–8.
15. *Тулупьев А.Л., Николенко С.И., Сироткин А.В.* Байесовские сети: логико-вероятностный подход. СПб.: Наука, 2006. 607 с.
16. *Тулупьев А.Л., Сироткин А.В.* Алгебраические байесовские сети: принцип декомпозиции и логико-вероятностный вывод в условиях неопределенности // Информационно-измерительные и управляющие системы. 2008. № 10. т. 6. С. 85–87.
17. *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.
18. *Фильченков А.А., Тулупьев А.Л.* Структурный анализ систем минимальных графов смежности Труды СПИИРАН. 2009. Вып. 11. С. 104–127.
19. *Фильченков А.А*, Алгоритм построения множества минимальных графов смежности при помощи клик владений// Труды СПИИРАН. 2010. Вып. 2 (13). [в печати]
20. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи клик собственников владений// Труды СПИИРАН. 2010. Вып. 4 (15). [в печати]
21. *Фильченков А.А.* Алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик // Труды СПИИРАН. 2010. Вып. 1 (12). [в печати]
22. *Фильченков А.А.*. Алгоритм построения множества минимальных графов смежности при помощи самоуправляемых клик собственников// Труды СПИИРАН. 2010. Вып. 3 (14). [в печати]
23. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Особенности анализа вторичной структуры алгебраической байесовской сети // Труды СПИИРАН. 2010. Вып. 1 (12). [в печати]

24. *Фильченков А.А., Тулупьев А.Л., Сироткин А.В.* Структурный анализ клик минимальных графов смежности // Вестник Тверского государственного университета. Сер. Прикладная математика. 2010. [в печати].
25. *Gorodetsky V.I., Drozdgin V.V., Jusupov R.M.* Application of Attributed Grammar and Algorithmic Sensitivity Model for Knowledge Representation and Estimation // Artificial Intelligence and Information, Control Systems of ROBOTSA. Amsterdam: Elsevier Science Publishers B. V., 1984, P. 232–237.

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 — 9. 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 — 9. 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. Tulupyev.

Тулупьев Александр Львович — д.ф.-м.н., доцент; 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 информатики в социокультурных исследованиях, применение методов биостатистики и математического моделирования в эпидемиологии, технология разработки программных комплексов с СУБД. Число научных публикаций — 220. 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.. Associate 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 — 220. 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.

Сироткин Александр Владимирович — младший научный сотрудник лаборатории теоретических и междисциплинарных проблем информатики СПИИРАН. Область научных интересов: алгебраические байесовские сети: вычислительные аспекты логико-вероятностного вывода в условиях неопределенности. Число научных публикаций — 50. avs@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.

Sirotkin Alexander Vladimirovich — junior researcher, Theoretical and Interdisciplinary Computer Science Laboratory, SPIIRAS. Research interests: algebraic Bayesian networks, algorithms of probabilistic-logic inference under uncertainty. The number of publications — 50. avs@iias,spb.su, www.tulupyev.spb.ru; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-3337, fax +7(812)328-4450.

Поддержка исследования. Работа выполнена при финансовой поддержке РФФИ, проект № **09-01-00861**-а «Методология построения интеллектуальных систем поддержки принятия решений на основе баз фрагментов знаний с вероятностной неопределенностью».

Рекомендовано лабораторией ТиМПИ, заведующий лабораторией д-р физ.-мат. наук, доцент А.Л. Тулупьев.
Статья поступила в редакцию 10.12.2010.

РЕФЕРАТ

Фильченков А.А, Тулупьев А.Л., Сироткин А.В. **Компаративный анализ клик минимальных графов смежности алгебраических байесовских сетей.**

Цель работы — предложить расширенную классификацию клик АБС, учитывающую новые результаты в анализе клик, включающие в себя классификацию *владений* — компонент связности клик, из которых исключили ребра их веса, и известные свойства этих владений.

Предложена классификация клик, основанная на наличии детей: клики поделились на *бездетные* и *родительские*, а также основанная на числе *собственных ребер* — ребер, вес которых совпадает с весом клики: клики поделились на *одно*- и *многореберные*, а также к классификации были добавлены сужения, которые по классификации стали *безреберными кликами*. Последнее сделано для того, чтобы не оговариваться каждый раз, когда нужно рассматривать одновременно клики и сужения, которые не являются кликами, но могут быть получены как подграфы максимального графа смежности, а, кроме того, для того, чтобы в дальнейшем использовать замкнутый снизу граф клик — граф клик, к которому добавлены сужения, состоящие ровно из одной вершины. Оба указанных типа сужений укладываются в понятие *безреберных клик*.

Помимо этого предложена классификация клик, основанная на числе владений (или, что то же самое, числа феодов, до которых сжимается клика): клики поделены на *моноклики*, сжимаемые до феода ребра, и *стереоклики*, сжимаемые до нескольких феодов.

На основании рассмотренных классификаций предложена сводная классификация, которая рассматривала восемь различных типов клик: *псевдоклика*, *моноклика-0*, *биклика*, *моноклика-1*, *бездетная стереоклика*, *родительская поликлика*, *моноклика-n*, *родительская стереоклика*.

Выявлены свойства всех рассматриваемых клик (и сужений), установлено число вершин и жил для каждого типа клики (сужения).

Полученные результаты развивают теорию алгебраических байесовских сетей и, в частности, теорию глобальной структуры алгебраических байесовских сетей. На основе введенной классификации могут быть предложены улучшенные алгоритмы для построения множества минимальных графов смежности, учитывающих специфику разных типов клики, и поэтому работающих быстрее.

SUMMARY

Filchenkov A.A, Tulupyev A.L., Sirotkin A.V. **Algebraic Bayesian networks minimal join graph cliques comparative analysis.**

The goal of this work is to suggest minimal join graph cliques extended classification that will pay respect to the new achievements in clique analysis including classification of *possessions* — e.g. connection components of cliques, edges of the weight the same with those were excluded from them, and the known properties of that possessions.

Clique classification is suggested that is based on the presence of children: clique set is divided to classes of *childless cliques* and *parent cliques;* also based on the number of *proper edges* — e.g. the edges which weight equals to the clique weight: clique set is divided to *one-edged cliques* and *many-edged cliques,* also the classification base was extended on some narrowings types, that are classified as edgeless cliques. The extension was made not to give more word for the cases when one need to deal both with the cliques and with narrowings that are maximal join graph subgraphs and also for further exploitation of a term closed from below clique graph — the clique graph that was contatins all the narrowings with the only vertex. Both the narrowing types are described with the term *edgeless cliques*.

Also cliques classification is suggested that is based on the number of possessions (or on the number of the feuds the clique being compressed into): clique set was divided to *monocliques*, that are compressed into the only feud, and *stereocliques*, that are compressed to a number of feuds.

Consolidated classification is suggested that is based on the considered classifications that observes 8 different clique types: *pseudoclique, monoclique-0, biclique, monoclique-1, childless stereoclique, parent polyclique, monoclique-n, parent stereoclique*.

Properties for all the cliques discussed (and the narrowings) are obtained, number of vertices and number of sinews are estimated for every the type.

The results develop algebraic Bayesian network theory and algebraic Bayesian network global structure theory in particular. New improved minimal join graph set synthesis algorithms that would pay attention to the different clique types features (and thus working faster) could be suggested on the base of the represented classification.