

## МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ

УДК 519.7

### О НЕИЗБЫТОЧНОМ ПРЕДСТАВЛЕНИИ МИНИМАКСНОГО БАЗИСА СТРОГИХ АССОЦИАТИВНЫХ ПРАВИЛ

В. В. Быкова, А. В. Катаева

*Сибирский федеральный университет, г. Красноярск, Россия*

Ассоциативные правила — тип зависимостей между данными, которые отражают, какие признаки или события встречаются совместно и насколько часто это происходит. Строгие ассоциативные правила представляют интерес для тех приложений, где требуется высокая степень уверенности в установленных зависимостях между данными, например, в информационной безопасности, анализе компьютерных сетей и медицине. Чрезмерно большое число выявленных правил существенно усложняет их экспертизу и применение. Для решения этой проблемы предложен алгоритм MClose, расширяющий возможности известного алгоритма Close. Алгоритм Close формирует минимаксный базис, в котором каждое строгое ассоциативное правило имеет минимальную посылку и максимальное следствие. Однако в минимаксном базисе остаются избыточные строгие ассоциативные правила. Алгоритм MClose в процессе построения минимаксного базиса распознаёт избыточные строгие ассоциативные правила и устраняет их. Предложенный алгоритм основан на свойствах замкнутых множеств. Доказаны выводимости, аргументирующие корректность алгоритма MClose.

**Ключевые слова:** *соответствия Галуа, замкнутые множества, строгие ассоциативные правила, избыточность, минимаксный базис.*

DOI 10.17223/20710410/36/9

### ON THE NON-REDUNDANT REPRESENTATION OF THE MINIMAX BASIS OF STRONG ASSOCIATIONS

V. V. Bykova, A. V. Kataeva

*Siberian Federal University, Krasnoyarsk, Russia***E-mail:** bykvalen@mail.ru, kataeva\_av@mail.ru

Associative rules are the type of dependencies between data that reflect which features or events occur together and how often this happens. Strong associative rules are of interest for those applications where a high degree of confidence of dependencies is required. For example, they are used in information security, computer network analysis and medicine. Excessively large number of identified rules significantly complicates their expert analysis and application. To reduce the severity of this problem, we propose the MClose algorithm, which extends the capabilities of the well-known algorithm Close. The Close algorithm forms a minimax basis in which each strong associative rule has a minimum premise and a maximal consequence. However, in the

minimax basis, some redundant strong associative rules remain. The MClose algorithm recognizes and eliminates them in the process of constructing a minimax basis. The proposed algorithm is based on the properties of closed sets. Its correctness is proved by proving the reflexivity, additivity, projectivity, and transitivity properties of strong associative rules.

**Keywords:** *Galois connection, closed sets, strong association rules, non-redundant, minimax basis.*

## Введение

При поиске ассоциативных правил анализируемое множество данных обычно описывается бинарным контекстом — матрицей, строки которой отвечают объектам предметной области, а столбцы — признакам этих объектов. Единичное значение элемента матрицы трактуется как наличие у объекта соответствующего признака, а нулевое — как его отсутствие. Бинарное представление данных существенно расширяет математический аппарат для их исследования. Современные методы поиска ассоциативных правил базируются преимущественно на анализе формальных понятий и теории вероятностей [1–3]. Анализ формальных понятий, как прикладная ветвь алгебраической теории решёток, является удобным математическим аппаратом описания методов поиска ассоциативных правил [4–6].

Главная проблема при поиске ассоциативных правил — это огромное число правил, возникающих при анализе больших контекстов. Для решения этой проблемы используются различные меры значимости правил. С их помощью правила фильтруются и для анализа предъявляются только те, для которых значения мер превышают заданные пороговые значения. Подобная фильтрация, конечно, уменьшает число правил, но не решает проблему размерности полностью. Часто после фильтрации всё равно остаётся значительное число правил, при этом многие из них избыточные. Ассоциативное правило считается избыточным, если его удаление из множества правил не приводит к потере информации о связях между данными в рассматриваемой предметной области.

Множество ассоциативных правил, не содержащее избыточных (в некотором смысле) правил, принято называть базисом. Существуют различные формальные определения избыточных ассоциативных правил и методы их устранения [7]. Наиболее развиты методы устранения избыточности для строгих ассоциативных правил. Для таких ассоциативных правил имеется плотная параллель с функциональными зависимостями из теории реляционных баз данных [8]. Строгие ассоциативные правила имеют достоверность, равную единице, и представляют интерес для приложений, где требуется высокая степень уверенности в установленных зависимостях между данными, например, в информационной безопасности, анализе компьютерных сетей и медицине [9–11].

В настоящее время известен ряд алгоритмов, позволяющих строить различные базисы для строгих ассоциативных правил. Наиболее значимыми базисами являются канонический и минимаксный. Канонический базис (или базис Дюкена — Гига) состоит из минимального числа строгих ассоциативных правил, рекуррентно описываемых в терминах псевдосодержаний [12]. Канонический базис математически глубоко исследован, однако все предложенные на сегодняшний день алгоритмы его построения в большей степени представляют теоретический, чем практический интерес [13]. Минимаксный базис состоит из строгих ассоциативных правил, имеющих минимальную посылку и максимальное следствие [14]. Для минимаксного базиса имеются хо-

рошо апробированные алгоритмы, к которым относится алгоритм Close [7, 15]. Исследования показали, что в построенных канонических и минимаксных базисах остаётся некоторая избыточность, которая может быть устранена на основе выводимостей, подобных аксиомам Амстронга, известным в теории реляционных баз данных для функциональных зависимостей.

В данной работе для строгих ассоциативных правил предложен алгоритм MClose, расширяющий возможности алгоритма Close. Алгоритм MClose позволяет в процессе построения минимаксного базиса устранять избыточность с сохранением поддержки и достоверности без дополнительного обращения к исходному бинарному контексту. Доказаны выводимости, обосновывающие корректность алгоритма MClose. Доказательство этих выводимостей — основной результат работы.

### 1. Основные термины и обозначения анализа формальных понятий

Приведём термины и обозначения, традиционно применяемые в анализе формальных понятий [4–6].

Пусть для предметной области определены два непустых конечных множества  $G$  и  $M$  объектов и признаков соответственно (от немецких слов *Gegenstände* — объект, *Merkmale* — признак). Предполагаем, что все объекты в  $G$  и признаки в  $M$  различны. Пусть задано отношение  $I \subseteq G \times M$  инцидентности между  $G$  и  $M$ . Тройку  $K = (G, M, I)$  принято называть формальным контекстом (или просто контекстом) предметной области. Считаем, что существование в  $I$  пары  $(g, m)$  означает, что объект  $g$  имеет признак  $m$ , и наоборот — признак  $m$  присущ объекту  $g$ .

Выберем в  $K = (G, M, I)$  два произвольных элемента  $g \in G$  и  $m \in M$ . Определим для них два отображения  $\phi$  и  $\psi$ :

$$\phi(g) = \{m \in M : (g, m) \in I\}, \quad \psi(m) = \{g \in G : (g, m) \in I\},$$

где  $\phi(g)$  — множество признаков, присущих объекту  $g$ , а  $\psi(m)$  — множество объектов, обладающих признаком  $m$ .

Отображения  $\phi$  и  $\psi$  обобщаются на  $A \subseteq G$  и  $B \subseteq M$  следующим образом:

$$\begin{aligned} \phi(A) &= \bigcap_{g \in A} \phi(g) = \{m \in M : \forall g \in A ((g, m) \in I)\}, \\ \psi(B) &= \bigcap_{m \in B} \psi(m) = \{g \in G : \forall m \in B ((g, m) \in I)\}. \end{aligned}$$

Отсюда  $\phi(A)$  — множество признаков, общих для всех объектов из  $A$ , а  $\psi(B)$  — множество объектов, которые обладают всеми признаками из  $B$ . Отображения  $\phi$  и  $\psi$  определены так, что если  $A_1, A_2 \subseteq G$  и  $B_1, B_2 \subseteq M$ , то

$$\phi(A_1 \cup A_2) = \phi(A_1) \cap \phi(A_2), \quad \psi(B_1 \cup B_2) = \psi(B_1) \cap \psi(B_2).$$

Целесообразно положить, что  $\phi(\emptyset) = M$  и  $\psi(\emptyset) = G$ : пустому множеству объектов присущи все признаки из  $M$  и каждый объект рассматриваемого контекста  $K$  обладает пустым множеством признаков. Если для отображений  $\phi$  и  $\psi$  применить единое обозначение  $(\cdot)'$ , то формулы для  $\phi(A)$ ,  $\psi(B)$ ,  $\phi(A_1 \cup A_2)$  и  $\psi(B_1 \cup B_2)$  записываются так:

$$A' = \bigcap_{g \in A} g' = \{m \in M : \forall g \in A ((g, m) \in I)\}; \quad (1)$$

$$B' = \bigcap_{m \in B} m' = \{g \in G : \forall m \in B ((g, m) \in I)\}; \quad (2)$$

$$(A_1 \cup A_2)' = A_1' \cap A_2'; \quad (3)$$

$$(B_1 \cup B_2)' = B_1' \cap B_2'. \quad (4)$$

Если  $g \in G$  и  $m \in M$ , то обозначения  $g'$  и  $m'$  традиционно служат сокращённой формой записи множеств  $\phi(g) = \{g\}'$  и  $\psi(m) = \{m\}'$  соответственно.

Из определения отображений «'» вытекают свойства, которые формально можно выразить в виде следующих утверждений.

**Утверждение 1.** Для всякого контекста  $K = (G, M, I)$  и любых  $B_1, B_2 \subseteq M$  верны следующие свойства:

- *антимонотонность*: если  $B_1 \subseteq B_2$ , то  $B_2' \subseteq B_1'$ ;
- *экстенсивность*:  $B_1 \subseteq B_1''$ , где  $B_1'' = ((B_1)')' \subseteq M$ .

**Утверждение 2.** Для всякого контекста  $K = (G, M, I)$  и любых  $A_1, A_2 \subseteq G$  верны следующие свойства:

- *антимонотонность*: если  $A_1 \subseteq A_2$ , то  $A_2' \subseteq A_1'$ ;
- *экстенсивность*:  $A_1 \subseteq A_1''$ , где  $A_1'' = ((A_1)')' \subseteq G$ .

В силу утверждений 1 и 2, отображения  $\phi$  и  $\psi$  составляют пару соответствий Галуа между  $2^G$  и  $2^M$  — системами всех подмножеств множеств  $G$  и  $M$  соответственно, частично упорядоченными по теоретико-множественному включению [4, 5]. Известно, что для соответствий Галуа  $\phi$  и  $\psi$  справедливы равенства [5]

$$\phi(\psi(\phi(A))) = \phi(A), \quad \psi(\phi(\psi(B))) = \psi(B)$$

или, то же самое, в единых обозначениях

$$((A')')' = (A'')' = A', \quad ((B')')' = (B'')' = B'. \quad (5)$$

Двойное применение отображения «'» определяет оператор замыкания на  $2^M$  в алгебраическом смысле [4]. Ему свойственны:

- *рефлексивность*: для любого  $B \subseteq M$  всегда  $B \subseteq B''$ ;
- *монотонность*: если  $B_1 \subseteq B_2 \subseteq M$ , то  $B_1'' \subseteq B_2'' \subseteq M$ ;
- *идемпотентность*: для любого  $B \subseteq M$  всегда  $(B'')'' = B''$ .

Справедливость этих свойств вытекает из утверждений 1 и 2.

Если  $B = B''$ , то множество признаков  $B \subseteq M$  называется замкнутым относительно оператора «''» в контексте  $K$ . Множество  $B'' = \phi(\psi(B))$  можно трактовать как набор признаков, которые всегда появляются в объектах контекста  $K$  вместе с признаками из  $B$ , причём это множество является наибольшим по включению в пределах  $K$ . Очевидно, что  $(\emptyset)'' = \phi(\psi(\emptyset)) = G'$ , где  $G'$  — множество признаков, свойственных всем объектам контекста  $K$ . Если  $B' = \emptyset$ , то всегда  $B'' = \phi(\psi(B)) = \phi(\emptyset) = M$ . Если  $B' \neq \emptyset$ , то, исходя из (1)–(4), замыкание для  $B \subseteq M$  можно вычислить за один просмотр контекста  $K$  по формуле

$$B'' = \bigcap_{g \in G} \{g' : B \subseteq g'\}. \quad (6)$$

В анализе формальных понятий пара множеств  $(A, B)$ ,  $A \subseteq G$ ,  $B \subseteq M$ , таких, что  $A' = B$  и  $B' = A$ , называется формальным понятием контекста  $K = (G, M, I)$  с объёмом  $A$  и содержанием  $B$  [6]. В формальном понятии  $(A, B)$  множества  $A$  и  $B$  всегда замкнуты относительно «''» в этом контексте:  $A = A''$  и  $B = B''$ . С помощью формальных понятий возможно концептуальное (или понятийное) моделирование различных предметных областей [3, 6]. Замкнутые множества также нашли широкое применение в поиске ассоциативных правил [7, 14, 15].

## 2. Ассоциативные правила и основные меры их значимости

Ассоциативным правилом на множестве признаков контекста  $K = (G, M, I)$  называется упорядоченная пара множеств  $r = (X, Y)$ ,  $X, Y \subseteq M$ . Принято ассоциативное правило  $r = (X, Y)$  записывать в виде  $X \Rightarrow Y$ , здесь множества  $X$  и  $Y$  называют посылкой (или причиной) и заключением (или следствием) соответственно [16]. В анализе ассоциативных правил часто полагают, что посылка и заключение — непустые непересекающиеся множества. С формальных позиций эти ограничения несущественны. Применительно к заданному контексту  $K$  всякое ассоциативное правило  $X \Rightarrow Y$  количественно характеризуется с помощью поддержки  $\delta(X \Rightarrow Y)$  и достоверности  $\gamma(X \Rightarrow Y)$  [17]. Эти числовые функции определяются через понятие поддержки множества признаков.

Поддержка  $\delta(X)$  множества признаков  $X \subseteq M$  в контексте  $K = (G, M, I)$  — отношение числа объектов, которым присущи признаки  $X$ , к общему числу объектов, представленных в этом контексте:

$$\delta(X) = |X'|/|G|. \quad (7)$$

Таким образом,  $\delta(X)$  — частота встречаемости в контексте  $K$  объектов, имеющих признаки  $X$ . Из формулы (7) следует, что для любого  $X \subseteq M$  значение  $\delta(X)$  неизменно находится в естественных границах

$$0 \leq \delta(X) \leq 1. \quad (8)$$

Чем ближе значение  $\delta(X)$  к 1, тем большее число объектов рассматриваемого контекста обладает всеми признаками из  $X$ . В силу антимонотонности отображения «'» поддержка множества признаков также удовлетворяет свойству антимонотонности: для всякого контекста  $K = (G, M, I)$  и любых  $X, Y \subseteq M$  при  $X \subseteq Y$  верно неравенство

$$\delta(Y) \leq \delta(X). \quad (9)$$

Согласно (9), поддержка множества признаков не может превышать поддержки любого из его подмножеств. Так, для произвольного  $X \subseteq M$  всегда

$$0 \leq \delta(M) \leq \delta(X) \leq \delta(\emptyset) = 1.$$

Множество признаков  $X \subseteq M$  называется частым в контексте  $K = (G, M, I)$ , если его поддержка больше или равна заданному пороговому значению  $\delta_0 \in [0, 1]$ . Если  $\delta(X) \geq \delta_0$  и  $X = X''$ , то  $X$  называется частым замкнутым множеством признаков в  $K$ . Частые множества и частые замкнутые множества признаков традиционно служат основой для поиска ассоциативных правил в заданном контексте. Следует отметить, что в худшем случае число частых замкнутых множеств признаков контекста  $K$  совпадает с числом частых множеств признаков и экспоненциально зависит от  $|M|$ . Однако на практике обычно число частых замкнутых множеств значительно меньше числа частых множеств. Примеры контекстов с полиномиальным относительно  $|M|$  числом частых замкнутых множеств можно найти в [15]. Использование следующего утверждения позволяет при поиске ассоциативных правил вместо частых множеств признаков применять частые замкнутые множества и тем самым сокращать пространство поиска ассоциативных правил.

**Утверждение 3.** Для всякого контекста  $K = (G, M, I)$  и любого множества  $X \subseteq M$  поддержка  $X''$  совпадает с поддержкой множества  $X$ :

$$\delta(X'') = \delta(X).$$

**Доказательство.** Пусть  $X \subseteq M$  — произвольное множество признаков контекста  $K = (G, M, I)$ . Тогда, исходя из (5) и (7), имеем равенство

$$\delta(X'') = |(X'')'|/|G| = |X'|/|G| = \delta(X).$$

Утверждение доказано. ■

Таким образом, если  $\delta(X) \geq \delta_0$ , то  $\delta(X'') \geq \delta_0$ , т. е. замыкание частого множества признаков также является частым.

Поддержкой  $\delta(X \Rightarrow Y)$  ассоциативного правила  $X \Rightarrow Y$  относительно контекста  $K = (G, M, I)$  называется величина

$$\delta(X \Rightarrow Y) = \delta(X \cup Y) = |(X \cup Y)'|/|G|, \quad (10)$$

указывающая, какая доля объектов этого контекста имеет признаки  $X \cup Y$ . Достоверность  $\gamma(X \Rightarrow Y)$  ассоциативного правила  $X \Rightarrow Y$  относительно контекста  $K$  определяется как отношение числа объектов, обладающих всеми признаками из  $X \cup Y$ , к числу объектов, которым свойственны только признаки  $X$ :

$$\gamma(X \Rightarrow Y) = |(X \cup Y)'|/|X'|.$$

Достоверность ассоциативного правила через функцию поддержки выражается формулой

$$\gamma(X \Rightarrow Y) = \delta(X \Rightarrow Y)/\delta(X) = \delta(X \cup Y)/\delta(X). \quad (11)$$

Заметим, что достоверность определяется формулой (11) только для тех ассоциативных правил  $X \Rightarrow Y$ , для которых  $\delta(X) \neq 0$ . Если  $\delta(X) = 0$  (в контексте нет ни одного объекта, обладающего признаками  $X$ ), то, согласно (8) и (9),  $\delta(X \cup Y) = 0$ . В этом особом случае полагают  $\gamma(X \Rightarrow Y) = 1$ . Исходя из (7)–(11), достоверность ассоциативного правила  $X \Rightarrow Y$  при произвольных  $X, Y \subseteq M$  всегда находится в границах

$$0 \leq \gamma(X \Rightarrow Y) \leq 1.$$

Чем ближе значение  $\gamma(X \Rightarrow Y)$  к 1, тем с большей уверенностью можно сказать, что признаки  $Y$  появляются в объектах рассматриваемого контекста вместе с признаками  $X$ .

Ассоциативное правило  $X \Rightarrow Y$  называется минимаксным в контексте  $K = (G, M, I)$ , если для  $K$  не существует другого ассоциативного правила  $X^* \Rightarrow Y^*$ , такого, что  $X^* \subseteq X$  и  $Y \subseteq Y^*$  и

$$\delta(X^* \Rightarrow Y^*) = \delta(X \Rightarrow Y), \quad \gamma(X^* \Rightarrow Y^*) = \gamma(X \Rightarrow Y).$$

Пусть заданы контекст  $K = (G, M, I)$  и  $\delta_0, \gamma_0$  — вещественные числа из  $[0, 1]$ . Будем говорить, что  $X \Rightarrow Y$  является  $(\delta_0, \gamma_0)$ -ассоциативным правилом в  $K$ , если выполняются два условия:

$$\delta_0 \leq \delta(X \Rightarrow Y) \leq 1; \quad (12)$$

$$\gamma_0 \leq \gamma(X \Rightarrow Y) \leq 1. \quad (13)$$

Величины  $\delta_0$  и  $\gamma_0$  играют роль пороговых значений для поддержки и достоверности соответственно. При  $\delta_0 = 0$  условие (12) отражает естественные границы поддержки. Данная ситуация свидетельствует о том, что нет ограничений на частоту появления

признаков  $X \cup Y$  в  $K$ . При  $\gamma_0 = 1$  условие (13) приводит к равенству  $\gamma(X \Rightarrow Y) = 1$ . В этом случае имеем  $(\delta_0, 1)$ -ассоциативное правило, которое будем называть *строгим ассоциативным правилом*. Таким образом, строгое ассоциативное правило — правило с достоверностью 1 и любой ненулевой поддержкой. Заметим, что ассоциативные правила с нулевой поддержкой и нулевой достоверностью не имеют практической ценности и поэтому не рассматриваются.

### 3. Свойства строгих ассоциативных правил

Известен критерий наличия в контексте строгого ассоциативного правила [6]. Приведём его с доказательством.

**Утверждение 4.** Достоверность ассоциативного правила  $X \Rightarrow Y$  относительно контекста  $K = (G, M, I)$  равна 1 тогда и только тогда, когда  $X' \subseteq Y'$  (или  $Y \subseteq X''$ ).

*Доказательство.* Согласно формуле (11), равенство  $\gamma(X \Rightarrow Y) = 1$  верно тогда и только тогда, когда  $\delta(X \cup Y) = \delta(X)$ , или, то же самое, когда  $(X \cup Y)' = X'$ . В силу (4) имеем  $(X \cup Y)' = X' \cap Y'$ . Равенство  $X' \cap Y' = X'$  возможно тогда и только тогда, когда  $X' \subseteq Y'$ .

Предположим, что  $X' \subseteq Y'$ . По утверждениям 1 и 2 всегда  $Y \subseteq Y''$ , а при  $X' \subseteq Y'$  верно включение  $Y'' \subseteq X''$ . Отсюда  $Y \subseteq X''$ . Тогда в силу антимонотонности отображения «'» имеем  $(X'')' \subseteq Y'$ . С учётом (5) верно  $X' \subseteq Y'$ . ■

Заметим, что утверждение 4 тривиальным образом выполняется для  $X \Rightarrow X''$  во всяком контексте  $K$  и при любом  $X \subseteq M$ . Рассмотрим некоторые частные случаи утверждения 4, важные с точки зрения устранения избыточности в множестве строгих ассоциативных правил.

**С л у ч а й 1:** ассоциативные правила вида  $X \Rightarrow Y$  при любых  $Y \subseteq X \subseteq M$ .

В силу антимонотонности отображения «'» при  $Y \subseteq X$  справедливо включение  $X' \subseteq Y'$ . Условие утверждения 4 выполняется, поэтому  $\gamma(X \Rightarrow Y) = 1$ . Таким образом, если в ассоциативном правиле заключение является подмножеством посылки, то такое правило имеет достоверность 1 в любом контексте  $K$  с поддержкой  $\delta(X)$ . Подобные строгие ассоциативные правила не несут в себе информации о существенных отношениях между множествами признаков  $X$  и  $Y$ , кроме естественного отношения «целое и часть целого», поэтому их следует считать тривиальными и не принимать во внимание. В частности, ассоциативные правила вида  $\emptyset \Rightarrow \emptyset$ ,  $M \Rightarrow Y$  и  $X \Rightarrow \emptyset$  при любых  $X, Y \subseteq M$  относятся к тривиальным строгим правилам.

**С л у ч а й 2:** ассоциативные правила вида  $\emptyset \Rightarrow Y$  при  $\emptyset \neq Y \subseteq M$ .

Для всякого контекста  $K = (G, M, I)$  и  $Y \subseteq M$  всегда  $Y' \subseteq G$ . Кроме того,  $\psi(\emptyset) = \emptyset' = G$  и  $\delta(\emptyset) = 1$ . Поэтому для правила  $\emptyset \Rightarrow Y$  имеем

$$\gamma(\emptyset \Rightarrow Y) = \delta(\emptyset \cup Y) / \delta(\emptyset) = \delta(Y).$$

Исходя из утверждения 4, равенство  $\gamma(\emptyset \Rightarrow Y) = 1$  имеет место тогда и только тогда, когда  $\emptyset' \subseteq Y'$ . Поскольку  $\psi(\emptyset) = \emptyset' = G$ , то  $\emptyset' \subseteq Y'$  верно лишь при  $G = Y'$ . Таким образом, строгое ассоциативное правило  $\emptyset \Rightarrow Y$  при  $Y \neq \emptyset$  имеет поддержку  $\delta(\emptyset) = 1$  и отражает наличие жёсткого ограничения на контекст  $K$ : все объекты, представленные в этом контексте, обязательно обладают множеством признаков  $Y$ .

Рассмотрим строгое ассоциативное правило  $X \Rightarrow Y$ . В силу (11) его поддержка всегда совпадает с поддержкой его посылки:  $\delta(X \Rightarrow Y) = \delta(X)$ . Если  $\delta(X) \geq \delta_0$ , то также  $\delta(X \Rightarrow Y) \geq \delta_0$ . Если после какого-либо изменения правила  $X \Rightarrow Y$  результирующее правило имеет поддержку не менее  $\delta(X)$ , то говорят, что такое изменение

сохраняет поддержку исходного правила. Докажем свойства строгих ассоциативных правил, которые позволяют из одних строгих ассоциативных правил вывести другие строгие ассоциативные правила (с сохранением или без сохранения поддержки).

**Лемма 1.** Пусть в контексте  $K = (G, M, I)$  множество  $X \subseteq M$  имеет поддержку  $\delta(X) \geq \delta_0$ . Тогда для контекста  $K$  при любом  $Y \subseteq X$  всегда справедливо строгое ассоциативное правило  $X \Rightarrow Y$  с поддержкой  $\delta(X \Rightarrow Y) \geq \delta_0$ .

*Доказательство.* При  $Y \subseteq X$  в силу антимонотонности отображения «'» всегда  $X' \subseteq Y'$ . Тогда по утверждению 4 ассоциативное правило  $X \Rightarrow Y$  является строгим. В силу (11) для него  $\delta(X \Rightarrow Y) = \delta(X) \geq \delta_0$ . ■

**Лемма 2.** Если для контекста  $K = (G, M, I)$  справедливо строгое ассоциативное правило  $X \Rightarrow Y$  с поддержкой  $\delta(X)$ , то при любом  $Z \subseteq M$  для этого контекста также верно строгое ассоциативное правило  $X \cup Z \Rightarrow Y$  с поддержкой  $\delta(X \cup Z) \leq \delta(X)$ .

*Доказательство.* Воспользуемся утверждением 4 и свойством монотонности оператора замыкания. Так как  $X \Rightarrow Y$  является строгим ассоциативным правилом в  $K$ , то  $Y \subseteq X''$  и  $\delta(X \cup Y) = \delta(X)$ . При  $X \subseteq X \cup Z$  верно включение  $X'' \subseteq (X \cup Z)''$ . Следовательно,  $Y \subseteq (X \cup Z)''$ . Это означает, что для  $K$  справедливо строгое ассоциативное правило  $X \cup Z \Rightarrow Y$  и для него  $\delta(X \cup Z \cup Y) = \delta(X \cup Z)$ . Отсюда с учётом антимонотонности поддержки имеем

$$\delta(X \cup Z) = \delta(X \cup Z \cup Y) \leq \delta(X \cup Y) = \delta(X).$$

Лемма доказана. ■

Лемма 2 отражает возможность пополнения посылки для строгого ассоциативного правила, но без гарантии сохранения поддержки. Особо следует отметить случай, когда расширение посылки строгого ассоциативного правила сохраняет поддержку этого правила.

**Следствие 1.** Если для контекста  $K = (G, M, I)$  справедливо строгое ассоциативное правило  $X \Rightarrow Y$  с поддержкой  $\delta(X)$ , то при любом  $Z \subseteq Y$  для этого контекста также справедливо строгое ассоциативное правило  $X \cup Z \Rightarrow Y$  с поддержкой  $\delta(X)$ .

*Доказательство.* Если  $\gamma(X \Rightarrow Y) = 1$ , то по лемме 2 также  $\gamma(X \cup Z \Rightarrow Y) = 1$ . Значит, верны равенства

$$\delta(X \Rightarrow Y) = \delta(X \cup Y) = \delta(X), \quad \delta(X \cup Z \Rightarrow Y) = \delta(X \cup Z \cup Y) = \delta(X \cup Z).$$

Отсюда при  $Z \subseteq Y$  имеем  $\delta(X \cup Z \Rightarrow Y) = \delta(X \cup Z \cup Y) = \delta(X \cup Y) = \delta(X)$ . ■

**Лемма 3.** Пусть в контексте  $K = (G, M, I)$  множество  $X \subseteq M$  имеет поддержку  $\delta(X) \geq \delta_0$ . Если для  $K$  справедливы строгие ассоциативные правила  $X \Rightarrow Y$  и  $X \Rightarrow Z$ , то для этого контекста также справедливо строгое ассоциативное правило  $X \Rightarrow Y \cup Z$  с поддержкой  $\delta(X) \geq \delta_0$ .

*Доказательство.* По утверждению 4, поскольку  $X \Rightarrow Y$  и  $X \Rightarrow Z$  являются строгими ассоциативными правилами в  $K$ , выполняются включения  $Y \subseteq X''$  и  $Z \subseteq X''$ . Значит,  $Y \cup Z \subseteq X''$ , что достаточно для выполнимости строгого ассоциативного правила  $X \Rightarrow Y \cup Z$  в заданном контексте. В данном случае  $\delta(X \Rightarrow Y \cup Z) = \delta(X) \geq \delta_0$ , т. е. свойство аддитивности сохраняет поддержку исходных правил. ■

Легко убедиться, что свойства рефлексивности и пополнения невыполнимы для произвольных  $(\delta_0, \gamma_0)$ -ассоциативных правил. Однако следующее свойство, называемое *проективностью*, выполняется для любых  $(\delta_0, \gamma_0)$ -ассоциативных правил.

**Лемма 4.** Если для  $K = (G, M, I)$  справедливо  $(\delta_0, \gamma_0)$ -ассоциативное правило  $X \Rightarrow Y$ , то при любых  $Z \subseteq Y$  и  $Y \neq \emptyset$  для этого контекста также верно  $(\delta_0, \gamma_0)$ -ассоциативное правило  $X \Rightarrow Z$ .

**Доказательство.** Поскольку  $Z \subseteq Y$  и  $Y \neq \emptyset$ , то  $Y$  можно представить в виде  $Y = Z \cup (Y \setminus Z)$ . Тогда, согласно условию (12) и антимонотонности функции поддержки, имеем

$$\delta_0 \leq \delta(X \cup Y) = \delta(X \cup (Z \cup (Y \setminus Z))) = \delta((X \cup Z) \cup (Y \setminus Z)) \leq \delta(X \cup Z).$$

Значит,  $\delta(X \Rightarrow Z)$  удовлетворяет условию (12). Аналогично, полагая, что  $\delta(X) \neq 0$ , получаем

$$\gamma_0 \leq \gamma(X \Rightarrow Y) = \delta(X \cup Y) / \delta(X) \leq \delta(X \cup Z) / \delta(X) = \gamma(X \Rightarrow Z).$$

Следовательно,  $\gamma(X \Rightarrow Z)$  удовлетворяет условию (13). При  $\delta(X) = 0$  всегда  $\gamma(X \Rightarrow Y) = 1$  при любом  $Y$ , в том числе и  $Z \subseteq Y$ .

Применительно к строгим ассоциативным правилам лемма доказывается тривиальным образом. Если для контекста  $K$  справедливо строгое ассоциативное правило  $X \Rightarrow Y$ , то  $Y \subseteq X''$ . Значит, при любых  $Z \subseteq Y$  и  $Y \neq \emptyset$  справедливо включение  $Z \subseteq X''$ . Следовательно,  $X \Rightarrow Z$  является строгим ассоциативным правилом в  $K$ . Кроме того,  $\delta(X \Rightarrow Y) = \delta(X)$ ,  $\delta(X \Rightarrow Z) = \delta(X)$ . Если  $\delta(X) \geq \delta_0$ , то  $\delta(X \Rightarrow Z) \geq \delta_0$ . Лемма доказана. ■

Лемма 4 отражает тот факт, что правую часть всякого  $(\delta_0, \gamma_0)$ -ассоциативного правила можно «расщепить» до отдельного признака, сохраняя при этом поддержку и достоверность в заданных границах. Для строгих ассоциативных правил леммы 3 и 4 констатируют равноценность различных эквивалентных форм записи этих правил.

**Следствие 2.** Представление строгого ассоциативного правила  $X \Rightarrow Y \cup Z$  эквивалентно его представлению в виде двух строгих ассоциативных правил  $X \Rightarrow Y$  и  $X \Rightarrow Z$ , при этом  $\delta(X \Rightarrow Y \cup Z) = \delta(X \Rightarrow Y) = \delta(X \Rightarrow Z) = \delta(X)$ .

**Лемма 5.** Если для контекста  $K = (G, M, I)$  справедливы строгие ассоциативные правила  $X \Rightarrow Y$  и  $Y \Rightarrow W$  и  $\delta(X) \geq \delta_0$ , то какими бы ни были  $X, Y, W \subseteq M$ , для этого контекста также справедливо строгое ассоциативное правило  $X \Rightarrow W$  с поддержкой  $\delta(X) \geq \delta_0$ .

**Доказательство.** Воспользуемся утверждением 4. Если для контекста  $K$  верны строгие ассоциативные правила  $X \Rightarrow Y$  и  $Y \Rightarrow W$ , то верны включения  $X' \subseteq Y'$  и  $Y' \subseteq W'$ . Следовательно,  $X' \subseteq W'$ . По утверждению 4 это условие является достаточным для выполнимости строгого ассоциативного правила  $X \Rightarrow W$ . Для него  $\delta(X \Rightarrow W) = \delta(X) \geq \delta_0$ . Таким образом, поддержка результирующего правила  $X \Rightarrow W$  совпадает с поддержкой правила  $X \Rightarrow Y$ , играющего роль начала транзитивной цепочки строгих ассоциативных правил. ■

Следующая лемма обобщает лемму 5 и определяет свойство, называемое *псевдо-транзитивностью* строгих ассоциативных правил.

**Лемма 6.** Если для контекста  $K = (G, M, I)$  справедливы строгие ассоциативные правила  $X \Rightarrow Y$  и  $Y \cup Z \Rightarrow W$ , то какими бы ни были  $X, Y, Z, W \subseteq M$ , для контекста  $K$  также справедливо строгое ассоциативное правило  $X \cup Z \Rightarrow W$  с поддержкой  $\delta(X \cup Z) \leq \delta(X)$ .

**Доказательство.** Если для контекста  $K$  верны строгие ассоциативные правила  $X \Rightarrow Y$  и  $Y \cup Z \Rightarrow W$ , то  $X' \subseteq Y'$  и  $(Y \cup Z)' \subseteq W'$ . Требуется доказать, что  $(X \cup Z)' \subseteq W'$ .

По лемме 2 из строгого ассоциативного правила  $X \Rightarrow Y$  вытекает справедливость строгого ассоциативного правила  $(X \cup Z) \Rightarrow Y$ . Тогда  $(X \cup Z)' \subseteq Y'$ . В силу антимонотонности отображения «'» всегда  $(X \cup Z)' \subseteq Z'$ . Отсюда если  $(X \cup Z)' \subseteq Y'$  и  $(X \cup Z)' \subseteq Z'$ , то верно включение  $(X \cup Z)' \subseteq Y' \cap Z'$ . По формуле (4) имеем  $(Y \cup Z)' = Y' \cap Z'$ . Учитывая, что  $(Y \cup Z)' \subseteq W'$ , окончательно получаем

$$(X \cup Z)' \subseteq Y' \cap Z' = (Y \cup Z)' \subseteq W'.$$

К сожалению, свойство псевдотранзитивности не гарантирует для результирующего правила сохранение поддержки. Для него  $\delta(X \cup Z \Rightarrow W) = \delta(X \cup Z)$ , но ввиду антимонотонности поддержки  $\delta(X \cup Z) \leq \delta(X)$ . ■

Леммы 1–6 позволяют сформулировать следующую теорему.

**Теорема 1.** Для всякого контекста  $K = (G, M, I)$  и любых  $X, Y, Z, W \subseteq M$  справедливы следующие свойства строгих ассоциативных правил:

$D_1$ . *Рефлексивность:*  $X \Rightarrow X$ .

$D_2$ . *Пополнение:* если  $X \Rightarrow Y$ , то  $X \cup Z \Rightarrow Y$ .

$D_3$ . *Аддитивность:* если  $X \Rightarrow Y$  и  $X \Rightarrow Z$ , то  $X \Rightarrow Y \cup Z$ .

$D_4$ . *Проективность:* если  $X \Rightarrow Y$  и  $Z \subseteq Y$ , то  $X \Rightarrow Z$ .

$D_5$ . *Транзитивность:* если  $X \Rightarrow Y$  и  $Y \Rightarrow W$ , то  $X \Rightarrow W$ .

$D_6$ . *Псевдотранзитивность:* если  $X \Rightarrow Y$  и  $Y \cup Z \Rightarrow W$ , то  $X \cup Z \Rightarrow W$ .

Указанные в теореме 1 свойства (или выводимости)  $D_1$ – $D_6$  позволяют из некоторого множества строгих ассоциативных правил вывести многие другие строгие ассоциативные правила без дополнительного сканирования контекста. Выводимости, подобные  $D_1$ – $D_6$ , справедливы и для функциональных зависимостей, имеющих место в теории реляционных баз данных, где их принято называть аксиомами Амстронга. Из доказанных лемм следует, что выводимости  $D_1, D_3, D_4, D_5$  гарантируют сохранение поддержки: результатом применения их к строгим ассоциативным правилам с поддержкой не менее чем  $\delta_0$  всегда являются строгие ассоциативные правила с таким же порогом поддержки.

#### 4. Алгоритмы Close и MClose

Известно большое число алгоритмов поиска ассоциативных правил. Основополагающими являются алгоритмы Apriori и Close. Алгоритм Apriori основан на свойстве антимонотонности функции поддержки. Он извлекает из заданного контекста все  $(\delta_0, \gamma_0)$ -ассоциативные правила при любых заданных допустимых значениях  $\delta_0$  и  $\gamma_0$  [11]. Алгоритм Close использует свойства частых замкнутых множеств и извлекает только строгие ассоциативные правила с заданным порогом поддержки  $\delta_0$  [7, 15].

Суть классического алгоритма Close заключается в пошаговом извлечении генераторов и частых замкнутых множеств признаков [14]. Множество  $\rho \subseteq M$  называется *генератором* замкнутого множества признаков  $X \subseteq M$ ,  $X = X''$ , если  $\rho'' = X$  и не существует другого множества  $\tau \subseteq M$ , такого, что  $\tau \subset \rho$  и  $\tau'' = X$ . Другими словами, генератор замкнутого множества признаков  $X$  — наименьшее по мощности множество признаков, имеющее замыкание  $X$ . Число признаков, входящих в генератор  $\rho$ , называется *мощностью* этого генератора. Если  $|\rho| = k$ , то  $\rho$  является  $k$ -генератором.

На вход алгоритма Close подается исходный контекст  $K = (G, M, I)$  и пороговое значение  $\delta_0$ . На выходе алгоритм Close формирует минимаксный базис и записывает его в множество  $AR$ . Изначально  $AR$  считается пустым и  $k = 1$ . На первом шаге в качестве  $k$ -генераторов рассматриваются все одноэлементные подмножества множества  $M$ . Замыкание  $\rho_k''$  для генератора  $\rho_k$  вычисляется по формуле (6). Поддержка для  $\rho_k''$  находится по формуле (7). Если  $\delta(\rho_k'') \geq \delta_0$ , то по частому замкнутому множеству  $\rho_k''$  строится минимаксное строгое ассоциативное правило

$$\rho_k \Rightarrow \rho_k'' \setminus \rho_k \tag{14}$$

и сохраняется в  $AR$ . Согласно (10) и утверждению 3, для него

$$\delta(\rho_k \Rightarrow \rho_k'' \setminus \rho_k) = \delta(\rho_k'') \geq \delta_0, \quad \gamma(\rho_k \Rightarrow \rho_k'' \setminus \rho_k) = 1.$$

Тот факт, что ассоциативное правило (14) является минимаксным, следует из определения генератора.

После генерации ассоциативного правила по  $\rho_k''$  создаются кандидаты в  $(k + 1)$ -генераторы для следующей итерации. Каждый такой кандидат формируется путём объединения двух  $k$ -генераторов, обладающих одинаковыми первыми  $k - 1$  признаками; предполагается, что все признаки в  $k$ -генераторах линейно упорядочены, например лексикографически. Далее выполняется проверка, вложен ли найденный кандидат в  $\rho_k''$ . Если вложен, то он исключается из рассмотрения. После нахождения всех  $(k + 1)$ -генераторов осуществляется переход к следующей итерации. Алгоритм Close завершает работу, когда исчерпаны все генераторы.

Множество ассоциативных правил, полученных в результате работы алгоритма Close, образует минимаксный базис строгих ассоциативных правил контекста  $K$ . Корректность алгоритма Close доказана в [14]. К сожалению, минимаксный базис может содержать избыточные строгие ассоциативные правила. Алгоритм MClose с помощью выводимостей  $D_1, D_3, D_4, D_5$  расширяет возможности алгоритма Close, он в процессе построения минимаксного базиса распознаёт избыточные строгие ассоциативные правила и устраняет их.

Дадим формальное определение избыточного строгого ассоциативного правила. Пусть  $AR$  — множество строгих ассоциативных правил, каждое из которых справедливо для контекста  $K$ . Будем говорить, что строгое ассоциативное правило  $X \Rightarrow Y$  логически следует из множества  $AR$ , если оно может быть выведено из  $AR$  с помощью выводимостей  $D_1, D_3, D_4, D_5$ . Будем обозначать этот факт так:  $AR \models X \Rightarrow Y$ .

Строгое ассоциативное правило  $X \Rightarrow Y$  назовём избыточным в  $AR$ , если

$$AR \setminus \{X \Rightarrow Y\} \models X \Rightarrow Y. \tag{15}$$

Множество строгих ассоциативных правил избыточное, если оно не содержит избыточных строгих ассоциативных правил. Обозначим через CSB (Concise Strong Basis) избыточное множество минимаксных строгих ассоциативных правил. Множество CSB можно построить путём генерации минимаксных строгих ассоциативных правил (например, с помощью алгоритма Close) и устранения среди них избыточных.

Распознавание избыточного строгого ассоциативного правила в  $AR$  основано на проверке логического следования (15). Алгоритм такой проверки использует понятие замыкания множества признаков относительно множества  $AR$  и является полиномиальным относительно  $|M|$  и  $|AR|$ . Замыканием множества  $X \subseteq M$  относительно  $AR$

(обозначается  $X^+$ ) называется множество всех признаков  $m \in M$ , таких, что верно логическое следование  $AR \models X \Rightarrow m$ . Заметим, что неизменно  $X^+ \subseteq M$ . Из выводимостей  $D_1, D_3, D_4$  вытекает справедливость следующего критерия: логическое следование  $AR \models X \Rightarrow Y$  имеет место тогда и только тогда, когда  $Y \subseteq X^+$ . Отсюда

$$AR \models X \Rightarrow X^+, \quad AR \models X \Rightarrow X^+ \setminus X.$$

Чтобы убедиться в справедливости (15), достаточно вычислить  $X^+$  относительно  $AR \setminus \{X \Rightarrow Y\}$  и проверить включение  $Y \subseteq X^+$ . Если  $Y \subseteq X^+$ , то строгое ассоциативное правило  $X \Rightarrow Y$  избыточно в  $AR$ , иначе оно не является избыточным.

Алгоритм вычисления  $X^+$  целиком базируется на выводимостях  $D_1, D_3, D_4, D_5$  и сводится к выполнению следующих действий. Сначала полагается  $X^+ = X$ . Далее осуществляется просмотр правил из  $AR$  и пополнение  $X^+$  по следующему принципу: если для правила  $Y \Rightarrow Z \in AR$  верно включение  $Y \subseteq X^+$ , то множество  $Z$  добавляется к  $X^+$ . Этот процесс повторяется до тех пор, пока изменяется  $X^+$ . Поскольку множества  $M$  и  $AR$  конечные, то процесс вычисления  $X^+$  конечен.

Заметим, что процесс исключения избыточных строгих ассоциативных правил не требует доступа к контексту  $K$ , и поэтому время его выполнения незначительно по сравнению со временем получения частых замкнутых множеств признаков. Для того чтобы исключить добавление заведомо избыточного строгого ассоциативного правила в  $AR$ , необходимо всякий раз после построения  $\rho_k''$  выполнять следующие действия. Если посылка  $\rho_k$  найденного правила (14) не равна  $\rho_k''$ , то найти замыкание  $\rho_k^+$  относительно вычисленного множества  $AR$ . Если  $\rho_k^+ = \rho_k''$ , то это минимаксное ассоциативное правило является избыточным, иначе оно включается в  $AR$ .

После завершения генерации минимаксных строгих ассоциативных правил необходим дополнительный просмотр результирующего множества  $AR$  с целью обнаружения оставшихся избыточных правил. Такие правила вполне возможны: они неизбыточные по отношению к ранее выявленным правилам, однако после пополнения  $AR$  новыми правилами могут оказаться избыточными. Таким образом, по построению результирующее множество  $AR$  состоит из минимаксных строгих ассоциативных правил и является неизбыточным. Заметим, что оперативное удаление избыточных правил сдерживает рост мощности  $AR$  и снижает время выполнения алгоритма.

Алгоритмы Apriori, Close и MClose сравнивались по числу сгенерированных строгих ассоциативных правил и времени работы. Эксперименты осуществлялись на компьютере с процессором Intel® Core™ i5 CPU & 2.30 ГГц и ОЗУ размером 4 Гбайт. Эксперименты выполнялись на контекстах, сгенерированных случайным образом.

Результаты экспериментов представлены в таблице. Для всякого анализируемого контекста  $K = (G, M, I)$  указаны  $|G|$  — число объектов,  $|M|$  — число признаков,  $\sigma = n/(|G| \cdot |M|)$  — плотность контекста, где  $n$  задаёт число единичных элементов матрицы инцидентности  $I$ . Контекст из 10000 объектов сформирован многократным копированием контекста, состоящего из 500 объектов.

Из таблицы видно, что алгоритмы Close и MClose эффективнее алгоритма Apriori как по числу извлечённых строгих ассоциативных правил, так и по времени работы. Алгоритм MClose по времени работы сопоставим с алгоритмом Close. Однако алгоритм MClose более чем в 2 раза уменьшает мощность минимаксного базиса, формируемого алгоритмом Close. Перспективны исследования, направленные на устранение избыточности для ассоциативных правил с любыми поддержками и достоверностями.

## Результаты экспериментов

Характеристика контекста $K = (G, M, I)$			Число извлечённых строгих ассоциативных правил / время, мс		
$ G $	$ M $	$\sigma$	Apriori	Close	MClose
20	10	0,38	1797 / 17562	45 / 250	22 / 297
30	10	0,39	2029 / 18347	46 / 374	19 / 412
30	10	0,55	15438 / 187202	69 / 390	20 / 484
50	10	0,53	27769 / 375178	46 / 78	13 / 124
500	10	0,53	27769 / 376154	42 / 124	13 / 168
10000	10	0,53	27769 / 378400	42 / 671	13 / 872

## ЛИТЕРАТУРА

1. Витяев Е. Е., Демин А. В., Пономарев Д. К. Вероятностное обобщение формальных понятий // Программирование. 2012. № 5. С. 18–34.
2. Городецкий В. И., Самойлов В. В. Ассоциативный и причинный анализ и ассоциативные байесовские сети // Тр. СПИИРАН. 2009. Вып. 9. С. 13–65.
3. Кузнецов С. О. Автоматическое обучение на основе анализа формальных понятий // Автоматика и телемеханика. 2001. № 10. С. 3–27.
4. Биркгоф Г., Барти Т. Современная прикладная алгебра. СПб.: Лань, 2005. 400 с.
5. Гуров С. И. Булевы алгебры, упорядоченные множества, решетки: определения, свойства, примеры. М.: Книжный дом «ЛИБРОКОМ», 2013. 352 с.
6. Ganter B. and Wille R. Formal Concept Analyses: Mathematical Foundations. Springer Science and Business Media, 2012. 314 p.
7. Pasquier N., Bastide Y., Taoui R., and Lakhal L. Generating a condensed representation for association rules // J. Intelligent Inform. Systems. 2005. V. 24. No. 1. P. 29–60.
8. Майер Д. Теория реляционных баз данных. М.: Мир, 1987. 608 с.
9. Батура Т. В. Модели и методы анализа компьютерных социальных сетей // Программные продукты и системы. 2013. № 3. С. 130–137.
10. Платонов В. В., Семенов П. О. Методы сокращения размерности в системах обнаружения сетевых атак // Проблемы информационной безопасности. Компьютерные системы. 2012. № 3. С. 40–45.
11. Payaraja M. and Meyyappan T. Mining medical data to identify frequent diseases using Apriori algorithm // Pattern Recognition, Informatics and Mobile Engineering (PRIME), IEEE, 2013. P. 194–199.
12. Duquenne V. and Obiedkov S. A. Attribute-incremental construction of the canonical implication basis // Ann. Math. Artif. Intelligence. 2007. V. 49. No. 1. P. 77–99.
13. Rudolph S. Some notes on pseudo-closed sets // LNCS. 2007. V. 4390. P. 151–165.
14. Zaki M. J. and Hsiao C.-J. Efficient algorithms for mining closed itemsets and their lattice structure // IEEE Trans. Knowledge Data Eng. 2005. V. 17. No. 4. P. 462–478.
15. Uno T., Asai T., Uchida Y., and Arimura H. An efficient algorithm for enumerating closed patterns in transaction databases // LNCS. 2004. V. 3245. P. 16–31.
16. Zhang C. and Zhang S. Association Rules Mining. Springer, 2002. 240 p.
17. Geng L. and Hamilton H. J. Interestingness measures for data mining: a survey // ACM Computing Surveys. 2006. V. 38. No. 3. Article 9.

## REFERENCES

1. Vityaev E., Demin A. V., and Ponomaryov D. K. Probabilistic generalization of formal concepts. Programming and Computer Software, 2012, no. 5, pp. 219–230.

2. *Gorodetskiy V. I., Samoylov V. V.* Assotsiativnyy i prichinnyy analiz i assotsiativnye bayesovskie seti [Association and Casual rule mining using associative Bayesian networks]. Tr. SPIIRAN, 2009, iss. 9, pp. 13–65. (in Russian)
3. *Kuznetsov S. O.* Machine learning on the basis of formal concept analysis. Automation and Remote Control, 2001, vol. 6, iss. 10, pp. 1543–1564.
4. *Birkhoff G. and Bartee T. C.* Modern Applied Algebra. N. Y., McGraw-Hill, 1970. 431 p.
5. *Gurov S. I.* Bulevy algebrы, uporyadochennyye mnozhestva, reshetki: opredeleniya, svoystva, primery [Boolean Algebras, Ordered Sets, Lattices: Definitions, Properties, Examples]. Moscow, LIBROKOM Publ., 2013. 352 p. (in Russian)
6. *Ganter B. and Wille R.* Formal Concept Analyses: Mathematical Foundations. Springer Science and Business Media, 2012. 314 p.
7. *Pasquier N., Bastide Y., Taoui R., and Lakhal L.* Generating a condensed representation for association rules. J. Intelligent Inform. Systems, 2005, vol. 24, no. 1, pp. 29–60.
8. *Maier D.* Theory of Relational Databases. Computer Science Pr., 1983. 640 p.
9. *Batura T. V.* Modeli i metody analiza komp'yuternykh sotsial'nykh setey [Social networks analysis models and methods]. Programmnye Produkty i Sistemy, 2013, no. 3, pp. 130–137. (in Russian)
10. *Platonov V. V. and Semenov P. O.* Metody sokrashcheniya razmernosti v sistemakh obnaruzheniya setevykh atak [Dimension reduction in network attacks detection systems]. Information Security Problems. Computer Systems, 2012, no. 3, pp. 40–45. (in Russian)
11. *Ilayaraja M. and Meyyappan T.* Mining medical data to identify frequent diseases using Apriori algorithm. Pattern Recognition, Informatics and Mobile Engineering (PRIME), IEEE, 2013, pp. 194–199.
12. *Duquenne V. and Obiedkov S. A.* Attribute-incremental construction of the canonical implication basis. Ann. Math. Artif. Intelligence, 2007, vol. 49, no. 1, pp. 77–99.
13. *Rudolph S.* Some notes on pseudo-closed sets. LNCS, 2007, vol. 4390, pp. 151–165.
14. *Zaki M. J. and Hsiao C.-J.* Efficient algorithms for mining closed itemsets and their lattice structure. IEEE Trans. Knowledge Data Eng., 2005, vol. 17, no. 4, pp. 462–478.
15. *Uno T., Asai T., Uchida Y., and Arimura H.* An efficient algorithm for enumerating closed patterns in transaction databases. LNCS, 2004, vol. 3245, pp. 16–31.
16. *Zhang C. and Zhang S.* Association Rules Mining. Springer, 2002. 240 p.
17. *Geng L. and Hamilton H. J.* Interestingness measures for data mining: a survey. ACM Computing Surveys, 2006, vol. 38, no. 3, article 9.