

УДК 004.93

Т. А. Зайко

Запорізький національний технічний університет

## ПОБУДОВА БАЗ ЧИСЕЛЬНИХ АСОЦІАТИВНИХ ПРАВИЛ

**Т. А. Зайко Побудова баз чисельних асоціативних правил.** *Вирішено задачу побудови баз чисельних асоціативних правил. Запропоновано метод синтезу баз асоціативних правил, у якому виконується фаззифікація транзакційної бази даних, обчислюються граничні значення підтримки, використовуються критерії для оцінювання непрямих асоціацій, що знижує ступінь участі користувача в процесі пошуку асоціативних правил. Проведено експерименти по дослідженню розробленого методу.*

*Асоціативне правило, база даних, нечітка логіка, підтримка, транзакція, функція належності.*

**Т. А. Зайко Построение баз численных ассоциативных правил.** *Решена задача построения баз численных ассоциативных правил. Предложен метод синтеза баз ассоциативных правил, в котором выполняется фаззификация транзакционной базы данных, вычисляются пороговые значения поддержки, используются критерии для оценивания косвенных ассоциаций, что понижает степень участия пользователя в процессе поиска ассоциативных правил. Проведены эксперименты по исследованию разработанного метода.*

*Ассоциативное правило, база данных, нечеткая логика, поддержка, транзакция, функция принадлежности.*

**T. Zayko Database constructiona of quantitative association rules.** *The problem of construction of base of quantitative association rules is solved. A method for the synthesis of association rules bases, which consists of the fuzzification of transactional database, the calculation of thresholds of support, the criteria to evaluate the indirect associations, which reduces the degree of user involvement in the process of mining association rules is proposed. The experiments on the investigation of the developed method are conducted.*

*Association rule, database, fuzzy logic, support, transaction, membership function.*

### Вступ. Постановка завдання побудови баз чисельних асоціативних правил

Останнім часом для виявлення нових взаємозв'язків у великих наборах даних широкого розповсюдження набули асоціативні правила [1]. Відомі методи індукції асоціативних правил [1–4] призначені, як правило, для роботи з бінарними даними та дозволяють виявляти залежності типу «Якщо відбулася сукупність подій  $X$ , то відбудеться і набір подій  $Y$ ». Проте реальні задачі класифікації, діагностування та розпізнавання образів, як правило, пов'язані з необхідністю обробки великих масивів даних, що приймають не тільки бінарні значення, а і значення з певного діапазону [5].

Тому актуальною є розробка методу синтезу баз асоціативних правил, що дозволяє витягати нові закономірності з вибірок даних, що приймають чисельні значення.

Метою цього дослідження є розробка методу синтезу баз чисельних асоціативних правил.

Нехай задано базу транзакцій  $D$ :

$$D = \{T_1, T_2, \dots, T_{N_D}\},$$

у якій кожний елемент  $T_j$ ,  $j = 1, 2, \dots, N_D$  містить інформацію про деякі взаємозалежні події;  $N_D = |D|$  – кількість елементів (транзакцій) у наборі даних  $D$ ;

$T_j = \{t_{1j}, t_{2j}, \dots, t_{N_{item,j}}\} \subseteq I$  –  $j$ -а транзакція бази  $D$ , що представляє собою список елементів  $t_{ij}$  з конкретним значенням числових атрибутів;  $t_{ij}$  –  $i$ -й елемент  $j$ -ї транзакції  $T_j$ ,  $i = 1, 2, \dots, N_{|T_j|}$ ;  $N_{|T_j|}$  – кількість елементів в  $j$ -й транзакції  $T_j$ ;  $I = \{\tau_1, \tau_2, \dots, \tau_{N_I}\}$  – множина

можливих змінних (ознак), які можуть входити в список елементів кожної транзакції  $T_j$ ,

$j = 1, 2, \dots, N_D$  набору даних  $D$ ;  $\tau_a$  –  $a$ -й елемент множини  $I$ ,  $a = 1, 2, \dots, N_I$ ;  $N_I = |I|$  – кількість елементів множини  $I$ .

У випадку, якщо база транзакцій  $D$  містить окрім бінарних, ще й дійсні змінні, елементи  $t_{ij}$  транзакції  $T_j$  представляються кортежем:

$$t_{ij} = \langle \tau_{ij}; v(\tau_{ij}) \rangle,$$

де  $\tau_{ij}$  – ознака із множини  $I$ , що відповідає елементу  $t_{ij}$ ;  $v(\tau_{ij})$  – значення ознаки  $\tau_{ij}$  в транзакції  $T_j$ ,  $v(\tau_{ij}) \in \Delta_{ij} = [\tau_{ij \min}; \tau_{ij \max}]$ ;  $\tau_{ij \min}$  і  $\tau_{ij \max}$  – мінімальне та максимальне значення з діапазону можливих значень  $\Delta_{ij}$  ознаки  $\tau_{ij}$ .

Тоді задача синтезу бази чисельних асоціативних правил на основі заданої транзакційної бази даних  $D$  полягає в тому, що необхідно побудувати набір (базу) чисельних асоціативних правил БП у вигляді імплікацій  $\langle X, v(X) \rangle \rightarrow \langle Y, v(Y) \rangle$ , у яких набори  $X$  та  $Y$  не перетинаються [1, 2, 6–8]:

$$\text{БП: } \langle X, v(X) \rangle \rightarrow \langle Y, v(Y) \rangle : X \subset I, Y \subset I, X \cap Y = \emptyset,$$

де  $v(X)$  та  $v(Y)$  – множини значень ознак, що належать множинам  $X$  і  $Y$ , відповідно.

### Побудова баз чисельних асоціативних правил

Для можливості видобування асоціативних правил з транзакційних баз даних  $D$ , що містять чисельні атрибути, такі атрибути перетворюються до формату, доступного для застосування відомих методів пошуку асоціативних правил [3]. При цьому потрібно виконувати розбиття чисельних ознак на непересічні інтервали, кожний з яких розглядається потім як новий атрибут. У розробленому методі синтезу баз чисельних асоціативних правил пропонується використовувати підхід на основі теорії нечітких множин [9, 10], що дозволяє розбивати вихідні ознаки на нечіткі інтервали та працювати з кожною ознакою, а не з окремими інтервалами її розбиття.

Тому на початковому етапі запропонованого методу виконується фазифікація бази транзакцій  $D$ , тобто приведення всіх її чисельних значень до нечіткого виду:  $D \rightarrow \text{Fuzzy}D$ . Таке перетворення дозволить виділити нечіткі терми кожної ознаки для можливості виконання подальшого видобування асоціативних правил.

При пошуку асоціативних правил важливою характеристикою, що використовується в процесі їх індукції, є підтримка наборів елементів, а також її граничне значення, яке задається, як правило, користувачем у якості параметра методу. У розробленому методі побудови баз чисельних асоціативних правил підтримку транзакції  $T_j$  будемо розраховувати як перетинання функцій належності ознак, що входять у транзакцію  $T_j$ :

$$\text{supp}(T_j) = \bigcap_{\tau_a \in T_j} \mu_a(T_j),$$

де  $\mu_a(T_j)$  – значення функція належності  $a$ -ї ознаки, обчислене для її значення в транзакції  $T_j$ .

Тоді підтримка набору  $X$  визначається як сума підтримок усіх транзакцій, що містять ця множина:

$$\text{supp}(X) = \sum_{X \subseteq T_j} \text{supp}(T_j) = \sum_{X \subseteq T_j} \bigcap_{\tau_a \in T_j} \mu_a(T_j).$$

Зважаючи підтримку набору  $X$ , що враховує оцінки індивідуальної інформативності ознак, які входять у набір даних, визначимо в такий спосіб:

$$\text{wsupp}(X) = \text{supp}(X) \sum_{\tau_a \in X} w_a,$$

де величина  $\sum_{\tau_a \in X} w_a$  визначає оцінку інформативності набору ознак  $X$ .

Зважена підтримка асоціативного правила  $X \rightarrow Y$  може бути визначена за формулою:

$$w\text{supp}(X \rightarrow Y) = \text{supp}(X \cup Y) \sum_{\tau_a \in X \cup Y} w_a.$$

Будемо вважати набір  $X$  зваженим набором, що часто зустрічається, якщо буде виконуватися умова:

$$w\text{supp}(X) \geq w\text{minsupport},$$

де  $w\text{minsupport}$  – граничне (мінімально припустиме) значення зваженої підтримки.

Важливо відзначити, що в деяких випадках крім наборів, що часто зустрічаються,  $X$  важливими для видобування нових знань про досліджувані об'єкти або процеси є нечасті набори елементів, що дозволяють виявляти непрямі асоціації.

Якщо два набори елементів  $X$  і  $Y$  суттєво залежать від наявності третього набору  $Z$ , тоді будемо вважати, що пара  $X$  і  $Y$  побічно пов'язана за набором  $Z$ :  $X \xrightarrow{Z} Y$ . Наявність такого зв'язку будемо визначати, виходячи з істинності таких умов:

1) значення зваженої підтримки набору  $X \cup Y$  є меншою за мінімальну припустиму:

$$w\text{supp}(X \cup Y) < \beta_{w\text{supp}(X \cup Y)},$$

де  $\beta_{w\text{supp}(X \cup Y)}$  – граничне значення зваженої нечіткої підтримки між наборами  $X$  і  $Y$  – величина, яка вказує на те, що набори  $X$  і  $Y$  зустрічаються не часто. Величину  $\beta_{w\text{supp}(X \cup Y)}$  можна встановити в такий спосіб:  $\beta_{w\text{supp}(X \cup Y)} = w\text{minsupport}$ ;

2) існує непустий набір  $Z$  ( $\exists Z \neq \emptyset$ ), для якого виконуються умови:

$$\begin{cases} w\text{supp}(X \cup Z) \geq \beta_{w\text{supp}(Z)}; \\ w\text{supp}(Y \cup Z) \geq \beta_{w\text{supp}(Z)}; \end{cases} \quad \text{і} \quad \begin{cases} w(X, Z) \geq w_{\text{min}}; \\ w(Y, Z) \geq w_{\text{min}}; \end{cases}$$

де  $\beta_{w\text{supp}(Z)}$  – граничне значення зваженої нечіткої підтримки між деяким набором і набором  $Z$ , що є ключовим для появи пари наборів  $X$  і  $Y$ , – величина, яка вказує на те, що набори  $X$  і  $Y$  зустрічаються часто при наявності множини  $Z$ . Величину  $\beta_{w\text{supp}(Z)}$  доцільно встановити в такий спосіб:  $\beta_{w\text{supp}(Z)} \geq \beta_{w\text{supp}(X \cup Y)}$ ;  $w(X, Z)$  і  $w(Y, Z)$  – значення критерію оцінювання взаємозв'язку між множинами  $X$  і  $Z$ , а також  $Y$  і  $Z$ , відповідно;  $w_{\text{min}}$  – мінімально припустиме значення критерію оцінювання взаємозв'язку між множинами елементів бази транзакцій.

У якості критерію  $w(X, Z)$  доцільно використовувати такий:

$$w(X, Z) = \frac{p(X \cap Z)}{\sqrt{p(X)p(Z)}},$$

де  $p(X)$ ,  $p(Z)$ ,  $p(X \cap Z)$  – імовірність появи наборів  $X$ ,  $Z$  і  $X \cap Z$  в базі даних  $D$ .

Таким чином, використання запропонованих вище критеріїв і їх граничних значень дозволить видобувати набори, які не тільки часто зустрічаються, але й набори, що рідко виникають у вихідній базі даних, однак, є цікавими та дозволяють виявляти нові знання про досліджувані об'єкти або процеси.

Після виконання підготовчих процедур, пов'язаних з перетворенням транзакційної бази даних і визначенням граничних значень підтримки й вірогідності, виконується видобування асоціативних правил і побудова на їх основі бази правил. У запропонованому методі з метою скорочення простору пошуку використовується властивість антимонотонності підтримки [3] при створенні нових наборів кандидатів.

Для побудови бази чисельних асоціативних правил задається транзакційна база даних  $D$ , яка містить транзакції  $T_j$  із числовими значеннями атрибутів  $\tau_a$ , вибирається за

необхідністю набір функцій належності  $\mu$ , використовуваних для розбиття діапазонів  $\Delta_{ak}$  значень чисельних ознак на нечіткі інтервали, визначаються значення мінімальної зваженої підтримки  $w\text{minsupport}$  та зваженої вірогідності  $w\text{minconfidence}$ , а також інших граничних значень ( $\beta_{w\text{supp}(X \cup Y)}$ ,  $\beta_{w\text{supp}(Z)}$ ,  $w_{\text{min}}$ ), необхідних для роботи методу.

Потім кожне  $j$ -е чисельне значення  $\tau_{aj}$   $a$ -ї ознаки  $\tau_a$  в транзакції  $T_j$  перетворюється до нечіткого значення  $f\tau_{aj}$ :

$$f\tau_{aj} = \sum_{k=1}^{N_{\text{еіо. } a}} \frac{\mu_{ak}(\tau_a \in T_j)}{|\Delta_{ak}|},$$

де  $\mu_{ak}(\tau_a \in T_j)$  – функція належності  $a$ -ї ознаки  $k$ -му терму, обчислена для значення ознаки  $\tau_a$  в транзакції  $T_j$ ;  $|\Delta_{ak}|$  – ширина  $k$ -го діапазону розбиття  $a$ -ї ознаки.

Після цього обчислюється потужність кожного  $k$ -го діапазону розбиття  $a$ -ї ознаки:

$$C\Delta_{ak} = \sum_{j=1}^{N_D} \mu_{ak}(\tau_a \in T_j),$$

і знаходиться максимальне значення такої величини для кожної  $a$ -ї ознаки:

$$\max C\Delta_a = \max_{k=1, 2, \dots, N_{\text{еіо. } a}} C\Delta_{ak}, a = 1, 2, \dots, |I|,$$

а також відповідний до величини  $\max C\Delta_a$  інтервал розбиття  $\max \Delta_a$ , який в подальшому процесі видобування асоціативних правил буде використовуватися для подання нечітких характеристик елемента  $\tau_a$ .

Для кожного інтервалу  $\max \Delta_a$ ,  $a = 1, 2, \dots, |I|$  обчислюється зважена підтримка  $w\text{supp}(\max \Delta_a)$  за формулами, наведеними вище. Усі інтервали  $\max \Delta_a$ , значення зваженої підтримки яких не менші мінімально припустимого граничного значення  $w\text{minsupport}$ , заносяться в масив  $FI_1$ , що містить одноелементні набори, які часто зустрічаються:

$$FI_1 = \{\max \Delta_a \mid w\text{supp}(\max \Delta_a) \geq w\text{minsupport}\}.$$

Інтервали з малими значеннями зважених підтримок  $w\text{supp}(\max \Delta_a)$  заносяться в масив  $RI_1$  одноелементних наборів, що рідко зустрічаються:

$$RI_1 = \{\max \Delta_a \mid w\text{supp}(\max \Delta_a) < w\text{minsupport}\}.$$

У випадку, якщо множина  $FI_1$  є порожньою, метод припиняє свою роботу, оскільки згенерувати асоціативні правила, що часто зустрічаються та є достовірними, не представляється можливим.

Потім на основі поточної множини  $FI_d$   $d$ -елементних наборів генерується множина  $C_{d+1}$   $(d+1)$ -елементних кандидатів в набори, що часто зустрічаються. При цьому аналогічно методу Аргіогі [2, 5] для зменшення кількості кандидатів на  $(d+1)$ -й ітерації використовується властивість антимонотонності підтримки, яка полягає в тому, що підтримка будь-якої множини елементів  $X$  не перевищує значення мінімальної підтримки будь-якої його підмножини  $Y \subset X$  [2, 5]. Тому на етапі генерації множини кандидатів  $C_{d+1}$  відтинаються (не створюються й не заносяться в  $C_{d+1}$ ) ті набори, які не можуть стати, такими, які часто зустрічаються, що визначається на основі інформації про набори з низькими значеннями підтримки  $w\text{supp}$ , розрахованими на попередніх етапах і розташованих в множині  $RI$ . Таким чином, при створенні нової множини  $C_{d+1}$  кандидатів використовується ідея про те, що у набору, який потенційно є таким, що часто зустрічаються, усі підмножини також повинні бути такими, що часто зустрічаються (значення всіх підтримок підмножин повинне бути не нижче граничного значення). Отже, кандидат  $X$ , що містить підмножину  $Y \subset X$ , яка була відкинута на попередніх етапах як

така, що не часто зустрічається ( $Y \in RI$ ), не включається в наступну множину  $C_{d+1}$  кандидатів у набори, що часто зустрічаються.

Після формування множини  $C_{d+1}$  для кожного набору  $X = \{\tau_1, \tau_2, \dots, \tau_{d+1}\} \in C_{d+1}$  ( $|X| = d + 1$ ) обчислюється його нечітка характеристика для  $j$ -ї транзакції  $T_j$ :

$$\mu_X(T_j) = \bigcap_{a: \tau_a \in X} \mu_a(\tau_a \in T_j),$$

далі визначається зважена підтримка набору  $X$ :

$$wsupp(X) = \sum_{X \in T_j, T_j \in D} \mu_X(T_j) \sum_{\tau_a \in X} w_a.$$

Якщо значення  $wsupp(X)$  є не меншим за мінімально припустимий поріг  $wminsupport$ , множина  $X$  заноситься в масив  $FI_{d+1}$  наборів, що часто зустрічаються, елементів, а якщо ні, то у масив наборів  $RI_{d+1}$ , що рідко зустрічаються.

У випадку, якщо  $FI_{d+1} \neq \emptyset$ , виконуються дії, аналогічні описаним вище. А якщо ні, то вважається, що подальше генерування наборів, що зустрічаються часто, є неможливим. Тому далі виконується видобування асоціативних правил із прийнятним рівнем вірогідності.

Асоціативні правила будемо генерувати виходячи з того, що:

$$wconf(X \rightarrow Y) = \frac{wsupp(X \rightarrow Y)}{wsupp(X)} \geq wminconfidence, X \cap Y = \emptyset.$$

Масив усіх наборів, що часто зустрічаються, знайдених раніше, може бути сформований як сукупність масивів  $FI_C$ :

$$FI = \bigcup_{C=1}^d FI_C.$$

Для кожного набору  $A \in FI$  та кожної його підмножини  $X \in A$  виконуються перевірки:

$$\frac{wsupp(A)}{wsupp(X)} \geq wminconfidence \quad \text{та} \quad \frac{wsupp(A)}{wsupp(A \setminus X)} \geq wminconfidence.$$

Нехай  $Y = A \setminus X$ . Тоді, якщо виконується перша умова, то генерується асоціативне правило  $X \rightarrow Y$ . Якщо виконується друга умова, то генерується правило  $Y \rightarrow X$ . При невиконанні обох умов, генерації правила для  $A \in FI$  й  $X \in A$  не відбувається.

Після цього виконується пошук цікавих рідких вигляду  $X \xrightarrow{Z} Y$ . Для цього формується множина  $RI$ :

$$RI = \bigcup_{C=1}^d RI_C,$$

і для кожного його її  $A \in RI$  виконуються наступні дії:  $X = A_{|A|}$  – останній елемент множини  $A$ ;  $Y = A_{|A|-1}$  – передостанній елемент множини  $A$ ;  $Z = A \setminus (X \cup Y)$ . Тоді будемо витягати асоціативні правила вигляду  $X \xrightarrow{Z} Y$  при виконанні таких умов:

$$\begin{cases} wsupp(X \cup Y) < \beta_{wsupp(X \cup Y)}; \\ (wsupp(X \cup Z)) \cap (wsupp(Y \cup Z)) \geq \beta_{wsupp(Z)}; \\ w(X, Z) \cap w(Y, Z) \geq w_{min}. \end{cases}$$

Після видобування імплікацій вигляду  $X \rightarrow Y$  й  $X \xrightarrow{Z} Y$  на їх основі синтезується база асоціативних правил, яка описує досліджувані об'єкти й процеси.

Таким чином, запропонований метод синтезу баз чисельних асоціативних правил передбачає використання критеріїв для оцінювання непрямих асоціацій, що знижує ступінь участі користувача в процесі пошуку асоціативних правил, зменшує ймовірність

видобування правил, що некоректно описують досліджувані об'єкти та процеси, а також дозволяє витягати набори, які не тільки часто зустрічаються, але і рідкі цікаві правила.

Для дослідження ефективності запропонованого методу побудова баз чисельних асоціативних правил оцінимо його обчислювальну складність  $O$ . Видобування асоціативних правил пов'язане з побудовою множини наборів, які часто зустрічаються,  $FI$ , що у свою чергу вимагає визначення значень підтримок кожного з кандидатів, максимальна кількість яких не перевищує  $|I|^2$ . Складність цього процесу становить  $O_{FI}(|I|^2)$  операцій. Процес видобування асоціативних правил із множини  $FI$  передбачає обробку кожної підмножини  $A \in FI$ , на що буде потрібно  $O_{\text{вс}}(|I|^2)$  операцій. Тому обчислювальна складність запропонованого методу складе:  $O = O_{FI}(|I|^2) + O_{\text{вс}}(|I|^2) = O(|I|^2)$ .

Оскільки залежність елементарних операцій від розміру вхідних даних є квадратичною, можна зробити висновок про те, що запропонований метод є обчислювально ефективним.

### Експерименти й результати

Для дослідження властивостей і характеристик запропонованого методу побудови баз чисельних асоціативних правил було виконано програмну реалізацію методу мовою програмування C#.

Експериментальне дослідження розробленого методу виконувалося за допомогою тестових транзакційних баз даних. Результати досліджень наведено в табл. 1, у якій використовуються такі позначення:  $N_D$  – кількість транзакцій  $T_j$  у базі  $D$ ;  $|I|$  – кількість елементів (ознак), з яких могли формуватися транзакції;  $\overline{|T_j|}$  – середня кількість ознак у транзакціях бази  $D$ ;  $|БП|$  – кількість витягнутих асоціативних правил у синтезованій базі правил БП;  $t$  – час роботи методу.

Таблиця 1.

Результати експериментальних досліджень

№	Характеристики бази транзакцій $D$			Результати синтезу баз асоціативних правил	
	$N_D$	$ I $	$\overline{ T_j }$	$ БП $	$t$
1	10000	100	10	202	0,12
2	10000	500	10	246	0,41
3	10000	1000	10	281	0,92
4	50000	100	10	489	0,53
5	50000	500	20	519	2,27
6	50000	1000	30	621	5,82
7	100000	1000	10	802	6,13
8	100000	5000	20	858	31,05
9	100000	10000	30	992	67,58

Як видно з таблиці, час роботи запропонованого методу суттєво залежить від кількості  $|I|$  елементів у базі  $D$ , що підтверджує оцінку обчислювальної складності  $O(|I|^2)$  методу. Крім того, результати експериментів показали, що кількість згенерованих правил  $|БП|$  збільшується зі збільшенням параметрів  $N_D$ ,  $|I|$  і  $\overline{|T_j|}$  бази транзакцій  $D$ . Це обумовлено великою кількістю різних елементів у множині  $FI$  наборів, які часто зустрічаються, що дозволило згенерувати більше чисельних асоціативних правил.

Таким чином, результати експериментів показали, що розроблений метод дозволяє витягати із заданого набору транзакцій чисельні асоціативні правила, сформовані не тільки на основі наборів, що часто зустрічаються, але й на основі наборів, які рідко виникають у вихідній базі даних та є цікавими і дозволяють виявляти нові знання про досліджувані об'єкти або процеси.

### Висновки

У роботі вирішено актуальну задачу автоматизації побудови баз чисельних асоціативних правил.

Наукова новизна роботи полягає в тому, що запропоновано метод синтезу баз чисельних асоціативних правил, у якому виконується фазифікація транзакційної бази даних, обчислюються граничні значення підтримки, використовуються критерії для оцінювання непрямих асоціацій, що знижує ступінь участі користувача в процесі пошуку асоціативних правил та дозволяє виявляти нові знання про досліджувані об'єкти або процеси.

Практична цінність отриманих результатів полягає в тому, що на основі запропонованого методу розроблено програмне забезпечення, що дозволяє виконувати побудову баз чисельних асоціативних правил.

Робота виконана в рамках держбюджетної науково-дослідної теми Запорізького національного технічного університету «Інтелектуальні інформаційні технології автоматизації проектування, моделювання, керування та діагностування виробничих процесів і систем» (номер державної реєстрації 0112U005350).

1. Zhang C. Association rule mining: models and algorithms / C. Zhang, S. Zhang. – Berlin : Springer-Verlag. – 2002. – 238 p.

2. Zhao Y. Post-mining of association rules: techniques for effective knowledge extraction / Y. Zhao, C. Zhang, L. Cao. – New York : Information Science Reference. – 2009. – 372 p.

3. Зайко Т. А. Анализ методов построения ассоциативных правил в интеллектуальных системах / Т. А. Зайко, А. А. Олейник, С. А. Субботин // Комп'ютерно-інтегровані технології: освіта, наука, виробництво. – 2012. – № 9. – С. 34–42.

4. Зайко Т. А. Методи пошуку асоціативних правил в транзакційних базах даних / Т. А. Зайко, А. О. Олійник, С. О. Субботин // Сучасні проблеми і досягнення в галузі радіотехніки, телекомунікацій та інформаційних технологій : VI міжнародна науково-практична конференція, 19–21 вересня 2012 р. : матеріали конференції. – Запоріжжя, 2012. – С. 186–187.

5. Интеллектуальные информационные технологии проектирования автоматизированных систем диагностирования и распознавания образов : монография / [С. А. Субботин, Ан. А. Олейник, Е. А. Гофман, С. А. Зайцев, Ал. А. Олейник ; под ред. С. А. Субботина]. – Харьков : ООО "Компания Смит", 2012. – 317 с.

6. Lian W. An efficient algorithm for finding dense regions for mining quantitative association rules / W. Lian, D. W. Cheung, S. M. Yiu // Computers & Mathematics With Applications. – 2005. – Vol. 50, № 3. – P. 471-490.

7. Khan M. S. Weighted Association Rule Mining from Binary and Fuzzy Data / M. S. Khan, M. Mueyba, F. Coenen // Lecture Notes in Computer Science. – 2008. – Vol. 5077. – P. 200-212.

8. Koh Y. S. Rare Association Rule Mining and Knowledge Discovery / Y. S. Koh, N. Rountree. – New York : Information Science Reference. – 2009. – 320 p.

9. Гибридные нейро-фаззи модели и мультиагентные технологии в сложных системах : монография / [В. А. Филатов, Е. В. Бодянский, В. Е. Кучеренко и др. ; под общ. ред. Е. В. Бодянского]. – Дніпропетровськ : Системні технології, 2008. – 403 с.

10. Encyclopedia of artificial intelligence / Eds.: J. R. Dopico, J. D. de la Calle, A. P. Sierra. – New York : Information Science Reference, 2009. – Vol. 1–3. – 1677 p.