

ПОШУК АСОЦІАТИВНИХ ПРАВИЛ ПРИ ІДЕНТИФІКАЦІЇ НАДЗВИЧАЙНИХ СИТУАЦІЙ ЯК ЗАДАЧА DATA MINING

Вознюк А.В.

Науковий керівник – д.т.н., проф. Перевозніков С. І.

Всі відхилення на протязі процесу розвитку надзвичайної ситуації (НС), що можуть змінити хід її розвитку та вплинути на якість вихідного результату знаходяться в відомих рамках, які, можуть бути визначені в процесі аналізу НС. Звичайно, що в таких умовах Data Mining здатна дати кращі результати ніж, наприклад лише аналітичні методи. Так, як при використанні алгоритмів Data Mining, непрямим чином, враховуються всі фактори, що можуть порушити звичайний розвиток надзвичайної ситуації.

Для пошуку частих наборів для асоціативних правил використовується модифікований алгоритм побудови FP-дерева (кожен шлях від кореню до листка якої відповідає деякій частій транзакції). Застосування алгоритму побудови FP-дерев дасть змогу досягнути таких переваг: заміна транзакцій бази даних у компактну структуру, що забезпечить ефективне та повне вилучення частих предметних наборів; при побудові FP-дерева виконується процедура розділення та захвату, яка дозволяє виконати декомпозицію поставленої задачі на множину більш простих; дозволяє уникнути ресурсоемкої процедури генерації кандидатів, що є характерним для більшості інших алгоритмів.

Стандартний алгоритм побудови FP-дерева змінений таким, чином, щоб дозволяти знаходити узагальненні асоціативні правила. Нехай $s_{i,l-1}$ є предком елементу $s_{i,l}$. Наведемо вигляд модифікації алгоритму. Нехай є частина транзакція $\langle s_{1,3}, s_{5,2} \rangle$. В стандартному алгоритмі будується гілка дерева, що містить цих два елементи, в модифікованому – кожна вершина гілки міститиме вектор, що складається з елементу, який входить в часту транзакцію та усіх його предків. Вигляд побудованих гілок дерев наведено на рисунку 1.

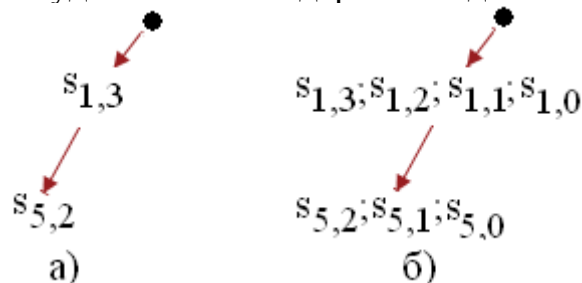


Рисунок 1 – Вигляд гілки FP-дерева а) в стандартному алгоритмі б) при вищенаведеній модифікації

Отже, такий підхід дасть змогу знаходити узагальнені асоціативні правила без необхідності обходу усіх транзакцій сховища даних стільки раз, скільки рівнів ієрархії предків елементів необхідно включити в дерево, тобто ефективніше стандартних алгоритмів пошуку асоціативних правил.