

КЛАСТЕРИЗАЦІЯ МУРАШИНИМИ АЛГОРИТМАМИ

М.О.Філінюк

Науковий керівник – доц., д.т.н. Штовба С. Д.

Кластеризація – це процес розбиття заданої вибірки об'єктів на множини, що не перетинаються зі схожими властивостями, що називаються кластерами. Розбиття проходить таким чином, щоб кожен кластер складався зі схожих об'єктів, а об'єкти різних кластерів суттєво відрізнялись.

В даній статті пропонується огляд переваг та недоліків нового методу кластеризації – кластеризація мурашиними алгоритмами. Основною ідеєю мурашиних алгоритмів є природна поведінка мурах, основі якої лежить їхня самоорганізація. Незважаючи на хаотичну поведінку кожного окремого мурахи, поведінка всього мурашника є достатньо розумною. При обміні інформацією, мурахи використовують феромони, чим більша концентрація феромонів на шляху, тим вища імовірність того, що мураха обере саме цей шлях. Кластеризація на основі мурашиного алгоритму має два підходи: на основі роботи мурашиного алгоритму і на основі природних особливостей мурах.

Перший підхід базується на роботі мурашиного алгоритму. В результаті його роботи отримують граф зв'язності, з деякою кількістю феромонів на ребрах. Кластеризація починається з ребра, з максимальною кількістю феромонів. Далі обирається спосіб кластеризації: на основі довжини зв'язків або на основі кількості феромонів на зв'язках. В першому випадку, обмеженням є сумарна довжина зв'язків від однієї з двох вершин, що розглядається. В другому випадку обмежена сума вершин, що обернено пропорційна кількості феромонів на ребрах, що входять в граф.

Другий підхід базується на природній особливості мурах чистити свої гнізда від тіл, тим самим формуючи кладовища з трупів. На протязі кластеризації мурахи представляються як програмні агенти, що у випадковому порядку переміщуються в області, що називається графом. Об'єкти (в природі - тіла) можуть довільно розміщуватись в середовищі, а агенти можуть довільно обирати, переміщати та залишати їх. Основний механізм, що лежить в основі цього виду кластеризації є більше притягування нових об'єктів великими кластерами, ніж малими. Це додатній зворотній зв'язок, що призводить до створення більших кластерів. Точна індивідуальна поведінка мурахи фактично і здійснює цей додатній зворотній зв'язок.

Відсутність централізованого контролю системою дозволяє може зробити цей тип алгоритму підходящим для застосування у широких проблемних задачах кластеризації, де дані не можуть бути доступні у повному обсязі. Наприклад, у задачах з географічного картування, у задах з просторовими та часовими обмеженнями.