

РЕЦЕНЗИЯ

на магистерскую диссертацию Ван Ячжо «Исследование и применение алгоритмов кластеризации потоков данных»

Данная диссертация посвящена методам кластеризации потоков данных, т.е. безусловно актуальной тематике, тесно связанной с различными направлениями современной информатики, в частности, с Data Mining.

Основное содержание выполненного исследования представлено в пяти главах диссертации. В первых трех главах излагаются методы кластеризации данных и потоков данных. Эти главы носят полностью реферативный характер.

В главе 4 предложен двухслойный алгоритм кластеризации, причем под слоями понимается онлайн и оффлайн обработка потоков данных. Насколько можно понять, именно в этой главе содержится основной результат, выносимый магистрантом на защиту. В главе 5 приводятся результаты тестирования двух алгоритмов – однослойного Clustream и двухслойного QuickClus. Сравнение этих результатов позволяет сделать вывод о большей эффективности двухслойного алгоритма.

Оценивая данную работу, нельзя не отметить ее серьезные недостатки. Прежде всего, это слабое владение автором русским языком, в результате чего в работе имеется большое количество орфографических, синтаксических и семантических ошибок, причем последние особенно неприятны, так как затрудняют понимание материала. Далее, неясно, какие результаты получены лично магистрантом. Он утверждает, что разработал двухслойный алгоритм QuickClus, но описания этого алгоритма нет. Указано лишь, что на слое оффлайн используется алгоритм DBSCAN, кратко описанный в главе 4. Ничего не говорится и об алгоритме Clustream. Не сказано, в какой программной среде были реализованы эти алгоритмы.

Считаю, что в связи с указанными недостатками данная диссертация может быть оценена лишь на «удовлетворительно».

20 июня 2011 г.

Рецензент
доктор физ.-мат. наук, профессор



И.Л. Братчиков