

ОТЗЫВ

на автореферат диссертации Цымблера Михаила Леонидовича «Интеллектуальный анализ данных в СУБД», представленной на соискание ученой степени доктора физико-математических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей в диссертационный совет Д 212.298.18 при ЮУрГУ

Диссертация Цымблера Михаила Леонидовича посвящена актуальной проблеме интеграции интеллектуального анализа данных в реляционные СУБД с применением высокопроизводительных вычислительных систем. По объективным для нашей страны причинам проблема решается путем внедрения фрагментарного параллелизма в код открытой последовательной СУБД PostgreSQL, доработанной под руководством автора до прототипа параллельной СУБД Pargre SQL. Эта СУБД попала в 2013 г. в пятерку лидеров рейтинга TPC-C среди параллельных СУБД для кластеров.

Теоретические исследования в диссертации связаны, прежде всего, с созданием и реализацией параллельных алгоритмов решения ряда задач интеллектуального анализа данных для современных высокопроизводительных систем с узловыми многоядерными акселераторами: кластеризация, поиск шаблонов, анализ временных рядов (главы 2, 3). Теоретические вопросы пресловутой интеграции рассмотрены в главах 4, 5. Здесь предлагаются и исследуются свои алгоритмы.

С профессиональной точки зрения, важным достоинством работы является разработка набора алгоритмов и их реализация в составе подсистемы pgMining интеллектуального анализа данных, которая может быть интегрирована в разные открытые СУБД с перспективой дальнейшего расширения в будущем.

Работа в целом прошла серьезную апробацию. Ее основные результаты опубликованы в достаточном числе статей, изданных в журналах, рекомендованных ВАК и проиндексированных в Scopus и Web of Science.

МОИ ЗАМЕЧАНИЯ ПО АВТОРЕФЕРАТУ. Они касаются «бросающихся в глаза» отдельных нюансов выполненного исследования и текстовых формулировок и никоим образом не влияют на положительную оценку работы в целом.

1. ЮУрГУ располагает весьма мощной кластерной платформой – суперкомпьютер «Торнадо ЮУрГУ». И если конечной задачей исследование в будущем ставить создание готового к применению программного продукта, то логично было бы выполнять все разработки именно для этой платформы. Вместе с тем, параллельный алгоритм DDCapriory реализован на платформе IBM Cell (см. С.24 автореферата). Безусловно, архитектура СВЕА весьма привлекательна. Но программирование для нее своеобразно, и она едва ли входит в состав «Торнадо».

Аналогичная ситуация – с алгоритмами MDD и PRAM, реализованными на суперкомпьютере ИВМиМГ СОРАН (см. С.35,40 автореферата).

2. Говоря о результатах проведенных экспериментов, автор всегда отмечает, что предложенные алгоритмы опережают по быстродействию аналоги конкурентов. Спрашивается: какие аналоги и какие конкуренты?

3. В автореферате было бы полезно, на наш взгляд, все же акцентировать рассматриваемые (подразумеваемые) в работе приложения Data Mining (правда, на С.12 автореферата имеется упоминание о социальных сетях). Особенности конкретной предметной области должны накладывать, по нашему мнению, свои ограничения на эффективность тех или иных алгоритмов (если не вести речь просто о реализации некоторых операторов языка DMQL).

МОЙ ГЛАВНЫЙ ВЫВОД. Судя по автореферату, работа Михаила Леонидовича Цымблера «Интеллектуальный анализ данных в СУБД» представляет собой завершенное диссертационное исследование, имеющее серьезное теоретическое и прикладное значение. Она вполне удовлетворяет требованиям ВАК по докторским диссертациям. Автор этой работы безусловно заслуживает присвоения ему ученой степени доктора физико-математических наук по специальности 05.13.11 – Математическое и программное обеспечение вычислительных машин, комплексов и компьютерных сетей.

Руководитель Республиканского научного семинара
«Методы моделирования» Академии наук Республики
Татарстан, профессор КНИТУ-КАИ им.А.Н. Туполева
докт.физ.-мат.наук., профессор

В.А. Райхлин

420111, г. Казань, ул. К.Маркса,10
varaikhlin@gmail.com
8 903 306 51 24

23.12.2019

