

Становище

относно дисертационен труд за придобиване на образователна и научна степен „доктор” в област на висшето образование „Природни науки, математика и информатика”, професионално направление 4.6 „Информатика и компютърни науки”, научна специалност 01.02.12

Автор: Ина Асенова Найденова

Тема: Приложение на Data Warehouse системи за разкриване на знания: Разреденост на данните в многомерния аналитичен модел

Изготвил становището: доц. д-р Калинка Михайлова Калоянова, СУ „Св. Кл. Охридски” - Факултет по математика и информатика

Дисертационен труд – актуалност, съдържание, резултати и постижения

Дисертационният труд „Приложение на Data Warehouse системи за разкриване на знания: Разреденост на данните в многомерния аналитичен модел” се състои от увод, 5 глави, заключение, списък на цитираната литература и 3 приложения – общо 145 страници.

Представеният дисертационен труд разглежда въпроси от една от най-актуалните съвременни области – използването на информация за анализи и вземане на решения.

Предизвикателствата пред системите, които се занимават с аналитична обработка на данните, са много и в различни направления. Сред най-важните от тях са тези, свързани с ефективността на работата на такъв тип системи. Те решават въпроси, свързани с ефективното съхранение и ефективното извличане на информацията.

Дисертацията предлага обстоен обзор на методите за ефективно съхранение и извличане на данни в аналитичните системи.

Направен е задълбочен анализ на най-разпространените техники за ефективно физическо съхранение на данните като композитно индексирание при съхранение на данните в многомерен формат, техники за компресиране на данните, разделяне на дялове, моделиране на кубове с по-малка размерност. Анализирани са съществуващите методи за намаляване на времето на отговор в аналитичните системи – пре-агрегации и материализация на сечения от куба, индексирание при релационно и многомерно съхранение на данните.

Въз основа на направения извод, че съществуващите методи не отчитат ефективно степента на запълването с информация на многомерните кубове, изследванията в дисертационния труд са насочени към използването на знанията за запълването/разредеността на многомерния куб за по-ефективна работа.

Предложен е модел на оптимизация, базиран на понятието *регулярна разреденост*. Въведена е нова класификация на видовете разреденост в n -мерен аналитичен куб, която различава *регулярната разреденост* от *случайната разреденост*. Използването на тези знания дава възможност за реализация на нови решения при анализ на данни със слаба плътност.

Систематизирани са видовете семантични връзки между стойностите на измерения в даден n -мерен аналитичен куб, водещи до появата на *регулярна*

разреденост. Предложен е подход за моделиране на кубове с по-малки размерности, на база на използването на тези връзки.

Чрез създаването на нов оригинален обект *карта на регулярна разреденост* - разширение на многомерния модел на данните, се изследват ефективното извличане, съхранение и интерпретацията на данните.

Изследвани са областите на приложение на картата – за подобряване на качеството на данните, за повишаване на производителността и намаляване на заеманото пространство на многомерните кубове, както и за някои подобрения в потребителския интерфейс и интерпретация на данните в аналитичните системи.

Предложени и подробно описани са методи за откриване на области с регулярната разреденост в n -мерен куб и за отстраняване на повторения на данни в адитивни йерархични участъци на неадитивни показатели.

Изложението в дисертационния труд е ясно и логично. Представено е формализирано представяне на многомерния куб, върху което е дефинирана *картата на регулярна разреденост*. Самата карта може представена като съвкупност от бизнес правила или графично.

В изложението са дадени голям брой примери, които целят да изяснят концепциите и начините на приложение на отделните методи.

Реализирани са експерименти с реални данни. За целта са разработени допълнителни софтуерни средства, като *Редактора за създаване на карта на регулярна разреденост*, който включва модули за откриване на некоректни данни в куб и за автоматична селекция на релевантните димензионни елементи.

Направените експерименти потвърждават работните хипотези. Получените резултати доказват ефективността на приложените методи.

Авторефератът към дисертационния труд отговаря на изискванията на съответните правилници. В него в компактен вид са представени основните резултати, постигнати от дисертационния труд.

Като самостоятелна единица той представя в достатъчно убедителен вид направеното изследване и неговите резултати.

Публикации

Докторантката е представила пет публикации. Три от публикациите са представени на международни конференции, две от публикациите са в международни списания. Две от публикациите са самостоятелни.

Всички статии са на английски език.

Публикациите са направени последователно в няколко години и отразяват съществени части от изследванията, залегнали в дисертационния труд.

До момента са забелязани две цитирания на публикациите.

Критични бележки и препоръки

Нямам съществени забележки към дисертационния труд.

Почти всички мои препоръки бяха отразени по време на работата по дисертацията.

Възможни са някои подобрения относно изложението и използваната терминология, но при липсата на установени преводи за голяма част от термините в областта, постигнатото в дисертационния труд и особено в автореферата е на много добро ниво.

Лични впечатления от кандидата

Познавам Ина Найденова от първите ѝ дни на работа като млад специалист и научен сътрудник в „Центъра за информационни технологии в съобщенията” към БТК. Работила съм с нея, докато беше асистент в катедра „Компютърна информатика” във ФМИ, имам наблюдения и върху развитието ѝ като специалист в Технологиика ЕАД. Работата ѝ винаги се е отличавала с голяма прецизност и задълбоченост, а самата тя проявява висока степен на отговорност и точност при изпълнение на поставените ѝ задачи.

Притежава добри аналитични възможности и определено има потенциал за научна и изследователска дейност.

Заклучение

В заключение считам, че оригиналността на научните и научно-приложни приноси на дисертационния труд, постигнатите практически резултати и научните публикации, които отразяват тези резултати, отговарят на изискванията на ЗРАСРБ, ПЗРАСРБ, както и на правилниците за условията и реда за придобиване на научни степени и звания на СУ „Св. Кл. Охридски” и Факултета по математика и информатика.

Затова убедено предлагам на уважаемото научно жури да присъди образователната и научната степен „**доктор**” на Ина Найденова в областта на висше образование: 4. „Природни науки, математика и информатика”, професионално направление: 4.6 „Информатика и компютърни науки”.

06.12.2012 г.
София

Рецензент:
/доц. д-р. К. Калоянова/