

Adatbányászat alkalmazása

Kármán Péter Elemér

Konzulens: dr Pataricza András

Feladatkiírás:

A számítógépes rendszerek megbízhatósági tulajdonságainak különböző mérési módszerekkel történő elemzése jellemzően nagy, de önmagában kevés jelentéssel bíró adathalmazt eredményez, amelyből külön eljárással kell kinyerni a rendszerre vonatkozó megbízhatósági paramétereket. A feladat ezen adatfeldolgozás megvalósítása adatbányászati eszközök felhasználásával.

A félévi munka:

- 1) A különböző adatbányászati feladatok és az őket megoldó algoritmusok megismerése Bodon Ferenc: Adatbányászati algoritmusok c. tanulmánya alapján.
- 2) Az adatbányászati funkciók alkalmazási lehetőségeinek megismerése OLTP rendszerek megbízhatósági benchmark-tesztjeinek kiértékelésére a Data Mining Approach to Identify Key Factors in Dependability Experiments c. cikk alapján.
- 3) A webszerverek megbízhatósági tesztelésére alkalmas WEB-DB és az alapjául szolgáló SPECWeb99 benchmark komponenseinek (mért paraméterek, szimulált hibák) és működésének (mérési környezet, hibainjektálás menete) megismerése a Dependability Benchmarking of Web-Servers c. cikk alapján.
- 4) Az IBM DB2 Universal Database v7.1 adatbázis, valamint rá épülve, az IBM DB2 Intelligent Miner for Data v8.1 adatbányász szoftver telepítése, az adatbányász funkciók kipróbálása egy minta-adatbázison, valamint az Apache webszerver SPECWeb-alapú benchmark teszteléséből kapott megbízhatósági paraméterek alkotta adathalmazon.

Felhasznált irodalom:

DB2 UDB V7.1 Database Administration Certification Guide

- 5) Az IBM DB2 Warehouse Manager, a Crystal Enterprise9, az IBM Tivoli Enterprise Data Warehouse V1.2 és az IBM Tivoli Service Level Advisor telepítése, a webszerver benchmark eredmények áthelyezése az SLA adatbázisába, valamint az SLA-generálás és analízis, illetve trend-vizsgálat működésének megismerése a megfelelő Redbook alapján.

Felhasznált irodalom:

Implementing Tivoli Data Warehouse 1.2 Redbook
Introducing IBM Tivoli Service Level Advisor Redbook