

Big Data módszerek és ajánló rendszerek

Benczúr András, Fekete Zsolt,
MTA SZTAKI

A feladat célja Web, közösségi tartalmak (Twitter, Last.fm, stb) feldolgozása területén jelentkező extrém méretű információfeldolgozási és modellezési problémák megoldása. A tipikus feladat olyan módszerek kifejlesztése és alkalmazása, amelyeket korábban hatékonysági korlátok miatt nem tartottak megvalósíthatónak. A kutatás elméleti korlátok, közelítési hibabecslések, modellezés és heurisztikák ötvözete, amelyek célja gyakorlati problémák megoldása és valós adatok kezelése.

A kiemelten javasolt kutatási területek a következők, illetve ezek kombinációja:

1. Gráf és mátrix eljárások elosztott architektúrájú sokgépes rendszereken (Hadoop, GraphLab, Stratosphere, stb)
2. Ajánló rendszerek;
3. Web és közösségi tartalmak vizsgálata és modellezése;
4. Emberi, felhasználói viselkedés kutatása.

A témához kapcsolódó diszciplínák: algoritmusok elmélete, kiemelten gráfok, illetve párhuzamos és elosztott algoritmusok; gépi tanulás statisztikai és lineáris algebrai módszerei.

A jelentkezővel szemben támasztott elvárások:

- Programozási gyakorlat Python, C++ vagy Java programnyelveken
- Algoritmusok alaposabb ismerete
- Angol nyelvismeret.