

Hochleistungsrechnen in der Chemometrie: Parallelisierung von Algorithmen zur Analyse chemischer Daten

- Diplomarbeit Thomas Beutlich -

Ziele/Aufgaben:

Im Rahmen dieser Diplomarbeit sollen zwei ausgewählte und häufig verwendete iterative Algorithmen der explorativen Datenanalyse von großen vollbesetzten Matrizen für einen Einsatz auf Parallelrechnern untersucht werden.

Folgende Algorithmen sind vorgesehen:

- NIPALS-Algorithmus (unvollständige Singulärwertzerlegung bzw. Hauptkomponentenanalyse)
- Fuzzy k-means Algorithmus (unüberwachte Clusteranalyse)

Im Rahmen der Diplomarbeit sind die folgenden Aufgaben formuliert:

- parallele (MPI-) Implementierung der genannten Algorithmen unter Verwendung von PLAPACK
- Untersuchung der Ergebnisse der Parallelisierung (Laufzeitverhalten, Geschwindigkeitsgewinne, Skalierbarkeit)
- Optimierung bzgl. des Laufzeitverhaltens
- Vergleich der parallelen Algorithmen mit bekannten MATLAB m-files
- Schaffung einer Schnittstelle (Script) zwischen MATLAB und diesen parallelen Algorithmen

Nutzung der Ergebnisse durch:

- Prof. Dr. rer. nat. habil. R. Salzer
Institut für Analytische Chemie