

Kurzfassung der Diplomarbeit von Binh Luong

Komplexe Zeitreihenanalyse

Die vorliegende Diplomarbeit beschäftigt sich mit den existierenden Algorithmen von Zeitreihenanalyse (ARIMA und Exponential Smoothing). Folgende Fälle werden untersucht, um die beiden Algorithmen zu verbessern und zu erweitern:

- Behandlung einer Periodizität mit nicht stabiler Periodenlänge
Häufig gibt es bei täglich erhobenen Daten eine monatliche Periodizität, zum Beispiel bei Gehaltseingang auf einem Konto. Die Länge der Periode schwankt dabei zwischen 28 und 31 Tagen.
- Behandlung mehrerer Periodizitäten
Seien zum Beispiel regelmäßige Stromverbrauchs-, Ladenverkaufs- oder Verkehrsstromdaten bekannt. Dann gibt es häufig mehrere Perioden verschiedener Länge.
- Behandlung einer Periodizität mit nicht-ganzzahliger Periodenlänge
Bei wöchentlichen Daten gibt es normalerweise eine jährliche Periodizität. Die Länge der Periode ist in diesem Fall nicht ganzzahlig ($365.25 / 7 \approx 52.179$) und wird anschließend von ARIMA oder Exponential Smoothing gerundet. Dies führt dann zur falschen Vorhersagen.

Das Thema wird zunächst theoretisch untersucht. Das beinhaltet die Suche nach bisherigen Veröffentlichungen und die Analyse existierender Verfahren. Wo es keine geeigneten Lösungen gibt, soll eine Lösung entworfen und detailliert beschrieben werden.

Im praktischen Teil werden die Verfahren unter IBM NZA (Netezza Analytics) implementiert. Dabei kann die existierende Implementierung von FFT (Fast Fourier Transform), ARIMA und Exponential Smoothing verwendet werden.

Neuer Code wird mit Stored Procedure (SQL) und UDFs (C++) implementiert.

Dauer: 6 Monaten

Institut: Angewandte Informatik und Formale Beschreibungsverfahren (AIFB)

Betreuer: Prof. Dr. Rudi Studer (KIT)

Dr. Christoph Lingenfelder und Dr. Boris Charpiot (IBM Deutschland GmbH)

Dr. Achim Rettinger (KIT) und Dipl. – Inform. Benedikt Kämpgen.