

Konzept und Implementierung eines echtzeitfähigen Model Management Systems

am Beispiel zur Überwachung von Lastprognosen für den Intraday Stromhandel

Yvonne Hegenbarth

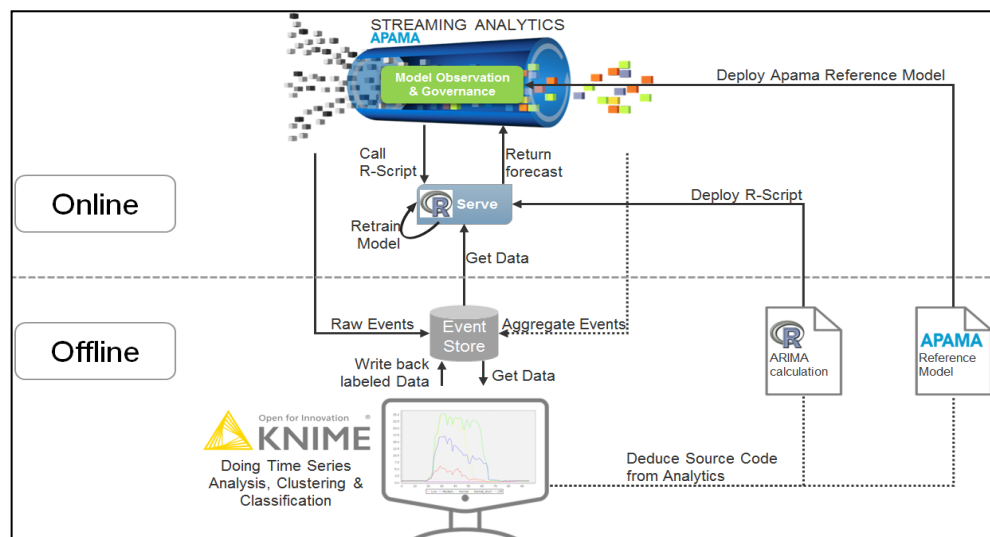


Abbildung 1 • Architektur des echtzeitfähigen Model Management Systems

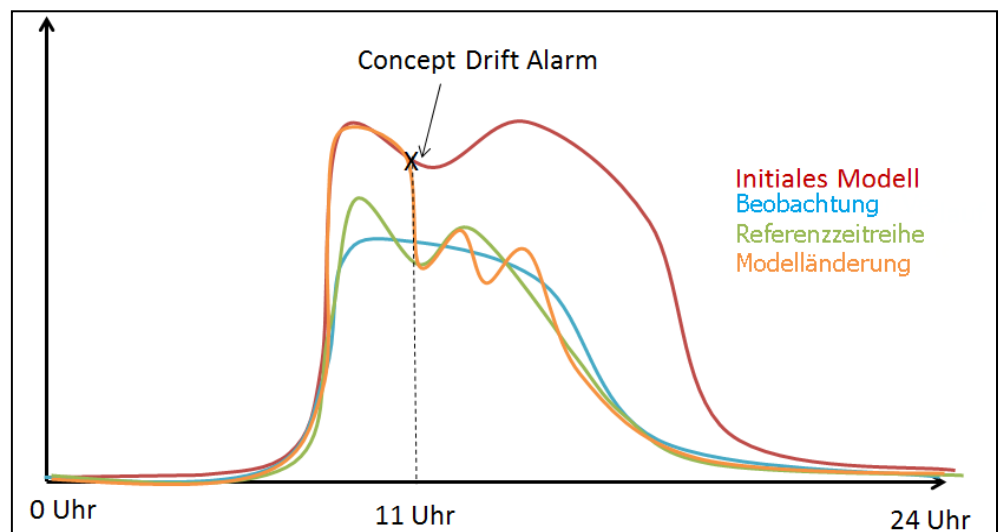


Abbildung 2 • Illustration einer Modelländerung nach einem Concept Drift Alarm

Motivation

Zur Gewährleistung der Stromstabilität in Deutschland müssen Verteilernetzbetreiber darauf achten, dass zur gleichen Zeit sowohl Strom produziert als auch entnommen wird. Die Erstellung von *vorhersagenden Modellen* soll unter anderem dabei helfen den Stromverbrauch des Folgetages zu prognostizieren.

In der Regel wird bei dem Prozess zur Berechnung von vorhersagenden Modellen mehrere zehn bis hundert Modelle erstellt. Zur Verwaltung und zum Reproduzieren von vorhersagenden Modellen muss ein *Model Management System* aufgesetzt werden. Die erzeugten Modelle werden persistiert und stehen für die Wiederverwendung zur Verfügung. Mit Hilfe von Visualisierungstools werden Benutzer in den Prozess zur Erstellung von vorhersagenden Modellen integriert. Dadurch wird vorausgesetzt, dass der Mensch als Überwacher agiert und aktiv die Parameter des Modells anpasst, falls die Vorhersagen zu stark von der Realität abweichen.

In der bevorstehenden Arbeit sollte ein *echtzeitfähiges Model Management System* (eMMS) entwickelt werden, das nicht den Persistenz-Ansatz verfolgt und den Menschen als Überwacher entlastet. Basierend auf einer Modelländerung kann entsprechend auf dem Energiemarkt im kontinuierlichen *Intraday Handel* überflüssiger Strom verkauft oder auf Grund mangelnder Reserven Strom gekauft werden.

Zielsetzung

Ziel der Arbeit war, ein eMMS zu entwickeln und dabei den Mensch als Akteur und Überwacher zu entlasten. Die Arbeit fokussierte sich nicht auf die Erstellung eines vorhersagenden Modells mit minimalem Vorhersagefehler. In dieser Arbeit wurde vielmehr erörtert, ob ein auf *Complex Event Processing* (CEP) basiertes Model Management möglich ist.

Vorgehensweise

Die in der Arbeit vorgestellte Architektur (vgl. Abbildung 1) sieht sowohl eine *CEP-Engine* (hier „Apama Streaming Analytics“), zur Modellwartung und -steuerung, als auch ein *Rserve* zur statistischen Berechnung vorhersagender Modelle vor. Weiterhin wird ein Datenbankmanagementsystem (hier *Event-Store*) benötigt, zur Verwaltung historischer Messwerte und ein Tool, zur Untersuchung und Aufbereitung der historisch hinterlegten Daten (hier *KNIME*).

Zur Untersuchung wurden von der EWE AG reale Datensätze von sogenannten *Registrierenden Leistungsmessungen*, mit unterschiedlichem Verbrauchsverhalten, zur Verfügung gestellt.

Das eMMS ist in zwei voneinander losgelöste Prozesse unterteilt: *Offline-* und *Online-Analyse*.

Die Offline-Analyse beschreibt den Prozess zur Datenanalyse historischer Daten mit Hilfe von Data Mining Techniken und die Erstellung von *beschreibenden Modellen* – Clustering und Klassifikation.

Das Clustering-Modell (*Referenzmodell*) beschreibt formähnliche Zeitreihen (*Referenzzeitreihen*) und ein Klassifikator (*Entscheidungsregeln*) prognostiziert basierend auf kalendrischen Merkmalen, die für den folgenden Tag zu erwartende Referenzzeitreihe. Basierend auf den Entscheidungsregeln, wird im weiteren Verlauf ein Trainingsdatensatz mit den vier aktuellsten Zeitreihen, die in der Referenzzeitreihe enthalten sind, gewählt. Dieser Trainingsdatensatz wird für die Berechnung der vorhersagenden Modelle zur Zeitreihenprognose (*ARIMA-Modelle*) benötigt.

Die Online-Analyse umfasst sowohl die *heuristische* Berechnung von vorhersagenden Modellen als auch eine Kombination aus einer *periodischen* und *schwellwertbasierten* Wartung. Über ein fortlaufendes Zeitfenster betrachtet wird eine Modellabweichung (*Concept Drift*) erkannt und basierend auf dem aktuellen Zeitreihenverlauf ein neues vorhersagendes Modell mit Hilfe des Referenzmodells berechnet (vgl. Abbildung 2).

Ergebnisse

Das im Rahmen dieser Arbeit entwickelte Verfahren hat die Machbarkeit zur Überwachung und Steuerung von vorhersagenden Modellen mit CEP Technologie gezeigt und den Mensch als Akteur und Überwacher weitestgehend entlastet.

Bei allen drei Datensätzen konnte eine bessere Vorhersagegenauigkeit mit der Modelländerung erzielt werden. Die Auswertungen bei zwei der Verbrauchsdaten ergaben eine geringe Verbesserung, wohingegen beim dritten Datensatz eine erhebliche Besserung nachgewiesen wurde (Vorhersagefehler halbiert).

Zusammenfassend ist das in der Arbeit implementierte eMMS im Stande drei wesentliche Aspekte zu behandeln:

1. Unvorhergesehene Zeitreihen (Anomalien) werden frühzeitig vom System erkannt und ein neues vorhersagendes Modell wird erstellt.
2. Das entwickelte System ist bei der Verschiebung von Clusterzentren im Referenzmodell bei gleichbleibenden Entscheidungsregeln robust.
3. Auch im Zuge einer Entwicklung im Datenstrom (*Concept Evolution*) mit sich ändernden Entscheidungsregeln, werden weiterhin angemessene Vorhersagen getroffen.

Ausblick

Die im Zusammenhang der Entwicklung aufgetretenen und noch zu untersuchenden Aspekte sind wie folgt:

Zum einen die Generierung von den beschreibenden Modellen (Clustering und Klassifikation) als *Predictive Model Markup Language* (PMML) zur einfacheren Überführung der Modelle in die CEP-Engine.

Zum anderen kann zur Minimierung des Berechnungsaufwandes von vorhersagenden Modellen entweder die Parameterkombinationen eingeschränkt werden, oder mit Approximationsalgorithmen (z.B. mit einfachen Modellen wie *Mean* oder *Naïv*) die Modellberechnung überbrückt werden.

Schließlich soll auf die Implementierung einer *clusterbasierten Änderungserkennung* im Datenstrom hingewiesen werden, zur automatisierten Concept Drift Erkennung sowie zur Analyse eines neuen Verhaltens in der Zeitreihe als Concept Evolution, um einen neuen Trainingsprozess der Offline-Analyse zu starten.

h_da

HOCHSCHULE DARMSTADT
UNIVERSITY OF APPLIED SCIENCES

fbi

FACHBEREICH INFORMATIK

In Kooperation mit

software AG

Kontakt

Yvonne Hegenbarth
Software AG
Uhlandstraße 9
64291 Darmstadt-Eberstadt
E-Mail: Yvonne.Hegenbarth@softwareag.com

Referentin

Prof. Dr. Inge Schestag
Hochschule Darmstadt, Fachbereich Informatik
Birkenweg 7
64295 Darmstadt
E-Mail: Inge.Schestag@h-da.de

Korreferent

Prof. Dr. Markus Döhring
Hochschule Darmstadt, Fachbereich Informatik
Schöfferstraße 8a
64295 Darmstadt
E-Mail: Markus.Doehring@h-da.de