

Sunday, November 26. 2006

GSM-basierte Verspätungserkennung

Das schöne an Technik ist der Umstand, das die meisten Systeme im Grunde genommen Multiple-Use-Techniken sind. Sie haben im Grunde genommen einen Hauptanwendungszweck, aber können für eine Vielzahl anderer Dinge verwendet werden. Keiner wird bestreiten, das das GSM-Netz in Deutschland eigentlich zur Kommunikation gedacht ist, sozusagen dessen Hauptzweck ist. Ich finde, das das Netz dafür viel zu schade ist. Es ist im Grunde genommen etwas sehr viel besseres. Es ist ein deutschlandweit verteiltes System von Funkbaken mit einer schier unglaublichen Zahl an Sensoren. Und das kann man ausnutzen. Diese Idee ist alles andere als neu. Sie wurde schon für eine Vielzahl von Ideen eingesetzt. So existieren Ideen der Stauerkennung über die Bewegungsmuster von Mobiltelefonen.

Eine solche Idee schwebte mir auch vor, als ich vor einigen Tagen über eine Verspätungserkennung für Züge nachdachte. Nicht das es so etwas noch nicht geben würde (Auch wenn ich während der Zugfahrt den Thrill eines Heureka-Erlebnisses hatte), das Tool Anorak macht soetwas. Aber mir schwebt da noch etwas anderes vor:

Im Grunde bewegt sich in jedem Moment ein Schwarm von Mobiltelefonen durch das Netz der Bahn. Warum ist das wichtig? Das interessante an roamingfähigen Mobilfunknetzen ist jetzt der Punkt, das ein Handy recht genau weiss wo es ist. Also jetzt nicht auf den Meter genau. Aber es weiss, welche mit welcher Mobilfunkstation es gerade verbunden ist. Eine Bahnreise ist jetzt eine Bewegung durch verschiedene Zellen des Mobilfunknetzes. Als Beispiel: Auf dem Weg zwischen Hamburg und Berlin werden beispielweise die Zellen 1 bis 200 durchquert. Soweit so logisch. Wenn ich eine gegebene gleichbleibende Strecke habe, die zu einem Zeitpunkt 0:00 startet und 1:33:00 endet, dann muesste ich um 0:00 in Zelle 0, um 1:33:00 in Zelle 200 sein, und vor allen Dingen ich bin zu bestimmten Zwischenzeiten in den Zwischenzellen.

Weiterhin sei folgende Definition von "pünktlicher Zug" gegeben: Wenn ein Zug seine Zielzelle nach Punkt 1:33:00 erreicht sei dieser pünktlich. Unpünktlichkeit kann man auch ganz einfach definieren. Der Zeitunterschied zwischen dem Zeitpunkt wann der Zug eine Zelle erreicht hat und wann er diese wirklich erreicht, ist der Wert der Unpünktlichkeit. Dadurch weiss ich zu jedem Zeitpunkt der Fahrt, wieviel Verspätung der Zug hat.

Anorak implementiert genau das, geht aber einen wesentlichen Schritt nicht. An einem Ort zu spät sein, heisst nicht notwendigerweise, das der Zug sein Ziel nicht pünktlich erreicht. Ein Zug der Hamburg fünf Minuten zu spät verlässt, kann Berlin durchaus pünktlich erreichen, wenn der Zugführer wenig materialsparend fährt. Deswegen möchte ich eine weitere Komponente einführen: Die Pünktlichkeitswahrscheinlichkeit oder in wieviel Prozent aller Fälle war ein Zug pünktlich, der eine Zelle C zu einem Zeitpunkt t erreicht hat. Daraus kann sich dann nämlich ergeben, das ein Zug der in Hamburg mit 5 Minuten Verspätung startet, mit 80% Wahrscheinlichkeit Berlin Hbf noch pünktlich erreicht, ein Zug der auf Höhe Spandau fünf Minuten Verspätung aufweist, aber meinetwegen nur eine Pünktlichkeitswahrscheinlichkeit von 20% hat. Beziehungweise das der Zug in einem Fall 0 Minuten verspätung haben wird, und im anderen Fall eine Wahrscheinliche verspätung von 5 Minuten.

Nun funktionieren statistische Voraussagen immer dann am besten, wenn ich eine möglichst grosse Anzahl von Werten erhoben hab, auf denen ich meine Auswertungen stützen kann. Und da kommt jetzt der Handyschwarm ins Spiel. Nehmen wir an, das eine hinreichende Anzahl von Handys ständig im Netz unterwegs sind. Nehmen wir an, das jedes dieser Handys über eine IP-Verbindung den Zeitpunkt, den Zug und die Funkzelle, die gerade erreicht wird, an ein zentrales Element liefert (hinreichend anonymisiert, aber autorisiert, die Fragestellung ist hier, wie man die Integrität der Datenbasis sicherstellt.). Dieses zentrale Element kann jetzt auf Basis dieser Daten ein Modell für die Strecke errechnen, mit dem sich für jeden Eintrittspunkt in eine Zelle die entsprechende Wahrscheinlichkeit des pünktlichen Eintreffens und des wahrscheinlichen Zeitpunkts des Eintreffens ermitteln lässt.

Die Genauigkeit dieses System hängt natürlich davon ab, wieviele Instanzen der dezentralen Komponente es gibt. Je mehr, und je weiträumiger Verteilt, desto besser wird zwangsläufig die Datenbasis.

Für mich wäre das die MobileWeb2.0-Anwendung mit Schwarmintelligenz (cool, zwei Buzzwords in einem Halbsatz), die für mich mal ein wenig echten Nutzen hätte. Man wäre nicht mehr darauf angewiesen, das einem die Bahn mitteilt, das man verspätet ist. Das System wäre vollkommen autark. Ausserdem. Im Grunde müsste das gleiche Modell auch für Fahrten mit dem Kraftfahrzeug funktionieren. Die Fahrstrecke müsste dann aber zusätzlich in eine Abfolge von Funkzellen zerlegt werden. Aufwendiger, weil die Freiheitsgrade eines Autofahrers höher sind. Aber durch Zerlegung der Strecke, in Steckensegmente und beschränkung auf Autobahnen sicherlich genauso machbar.

Blog Export: c0t0d0s0.org, http://www.c0t0d0s0.org/

Posted by Joerg Moellenkamp in Braindump, German at 12:19

Die Bahn AG wird wohl ewig die Bundesbahn bleiben!

Gute Idee, auch wenn ich selber der Schwarmintelligenz nicht mehr so euphorisch gegenüber eingestellt bin wie einst. Auf Basis welcher Datenerhebungen und Modelle die Disponenten in der Netzleitzentrale die Vorhersagen vom Computer serviert bekommen, steht da leider nicht: <http://www.zeit.de/2006/48/Bahn-2>
Anonymous on Nov 26 2006, 13:19

grins die werden ewig bei mir so heissen ... Aber ich änder das gleich mal ...

Ich denke mal, die Bahn hat durch die Linienförmige Zugbeeinflussung bessere Daten, aber an die kommt man ja nicht ran ...
ausserdem denke ich, das für eine Verspätungserkennung die GSM-Daten ausreichend sind ...
Anonymous on Nov 26 2006, 13:36

Kennst Du schon diese Seite?
http://reiseauskunft.bahn.de/bin/ts/bhftafel.exe/dn?L=vs_dbdialog&rt=1&
Da publiziert die Bahn die aktuellen Zugverspätungen.

Aber die Idee mit dem statistischen Verspätungsmodell finde ich extrem cool.
Kannst Du noch einen Link auf dieses Anorak-Tool geben? Ich finde das leider nicht im Netz.

Eine andere Idee, die neulich in den de.etc-bahn-Newsgruppen aufkam, war, die statistischen Verspätungsinformationen dafür zu nutzen, gleich bei der Verbindungsauskunft zu errechnen, mit welcher Wahrscheinlichkeit alle Anschlüsse erreicht werden, und was die wahrscheinlichste Ankunftszeit ist.
(man bekommt mit Tricks aus der o.g. URL auch andere Tage raus als den heutigen).

Ich glaube, Die Bahn bietet noch viele Ideen für lustige Web2.0-Anwendungen.
Anonymous on Nov 26 2006, 19:15