



Protokoll zum Vortrag Hromkovic, Zufall als Quelle der Effizienz

Der Vortrag behandelte das Problem des Abgleichs extrem großer Datenbestände mittels zufällig gewählter Primfaktor-Restklassenbildung.

Eingangs erwähnte der Vortragende die beiden Zufallshypothesen (unvollständiges Wissen oder inhärente Eigenschaft des Universums).

Darauf aufbauend schilderte er das Problem des Abgleichs großer Datenbestände auf verteilten Speichern und machte plausibel, dass bei Übermittlung eines der involvierten Datenbestands die Übertragungsfehler-rate so hoch wäre, dass das Vertrauen in dieses „sichere“ aber ineffiziente deterministische Verfahren beschränkt sein muss.

Alternativ dazu stellte er vor, dass man durch

- Wahl einer beliebigen (zufällig ausgewählten) Primzahl kleiner der numerischen Interpretation des binären DB-Inhalts, $P < \text{num}(\text{DB})$,
- Restklassenbildung $\text{num}(\text{DB}) \text{ MOD } P = R1$
- Übertragung von P und $\text{num}(\text{DB}) \text{ MOD } P$
- Restklassenbildung von $\text{num}(\text{DB}2) \text{ MOD } P = R2$ und
- Vergleich dieser beiden Reste

diese Frage wesentlich effizienter (beschränkte Übertragungsmenge) und relativ sicher beantworten kann.

Antworten:

$R1 \neq R2 \Rightarrow$ DB-Replikation enthält Fehler; definitiv korrekter Schluss

$R1 = R2 \Rightarrow$ DB-Replikation ist mit hoher Wahrscheinlichkeit korrekt. Die Wahrscheinlichkeit konvergiert mit steigender DB-Größe gegen 1.

Aufbauend auf diesen Ergebnissen wurde gezeigt, warum es Primzahlen P (Zeugen) gibt, die fälschlicherweise Identität anzeigen und wie diese (theoretisch) ermittelt werden können. (Alle P mit P ist Primfaktor von $(\text{DB}1 - \text{DB}2)$).

Es wurde gezeigt, dass der Quotient schlechte Zeugen / alle Zeugen $< 2 \ln n / n$ ist und dieser Quotient mit steigendem n strikt gegen 0 konvergiert. Somit läßt sich bei bekanntem n die Fehlerwahrscheinlichkeit berechnen.

Schließlich wurden Methoden diskutiert, wie man unter Kenntnis dieses Verfahrens die Fehlerwahrscheinlichkeit weiter reduzieren kann (Replikation mit unterschiedlichen Primzahlen, künstliche Ausdehnung des Datenraums von DB).

Die Veranstaltung verlief stark interaktiv mit sehr vielen Fragen des Vortragenden an das Publikum. Richtige Antworten wurden teils mit einer Mozartkugel, teils mit einem Lehrbuch (nach Schwierigkeitsgrad) „vergütet“.

Dem Vortrag wohnten 29 Personen bei. Wer sich in diese Thematik vertiefen möchte, ist eingeladen, Kapitel 6, Der Zufall und seine Rolle in der Natur, oder: Zufall als Quelle der Effizienz in der Algorithmik aus dem Buch Juraj Hromkovic, Sieben Wunder der Informatik: Eine Reise an die Grenze des machbaren mit Aufgaben und Lösungen, B.G. Teubner Verlag, Wiesbaden, 2006, nachzulesen.

