

Optimierung und Simulation Serverkonsolidierung via Bin Packing

B. Görder, P. Huhn, M. Kolonko, J. Lahmann

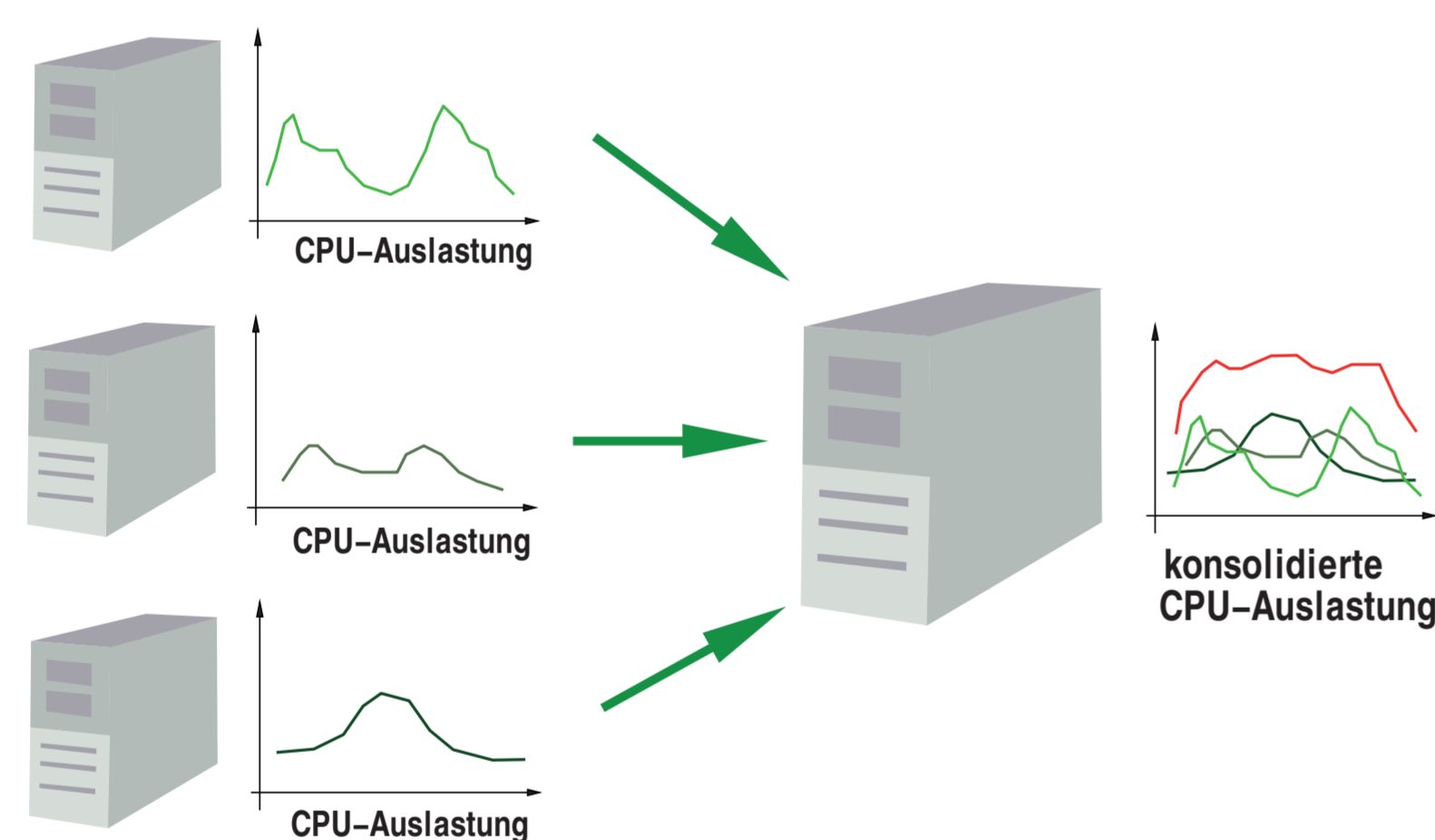
Problemstellung

Serverkonsolidierung

Rechner und Server müssen so ausgelegt sein, dass ihre Ressourcen die Belastungen zu Spitzenzeiten bewältigen können. Diese Belastungsspitzen fallen jedoch im allgemeinen sehr selten und zeitversetzt an, so dass sich meist eine durchschnittliche Auslastung von weniger als 20% ergibt. Dies bietet Einsparmöglichkeiten.

Ein Konzept zur besseren Nutzung von Rechnerressourcen liegt in der Konsolidierung von Servern. Mehrere Ausgangsrechner werden als virtuelle Maschinen auf einen Server abgebildet, dessen Kapazität aber wesentlich unter der Summe der Kapazitäten der Ausgangsrechner liegt. Dabei wird ausgenutzt, dass die Belastungsspitzen der Ausgangsrechner meist zeitversetzt auftreten und die Leistungsspitzen einzelner Rechner mit Phasen niedriger Recherauslastung anderer Rechner kompensiert werden können.

Ziel ist es, die Ausgangsrechner so zusammenzustellen, dass sie von einer möglichst geringen Anzahl von Zielsystemen ersetzt werden können.



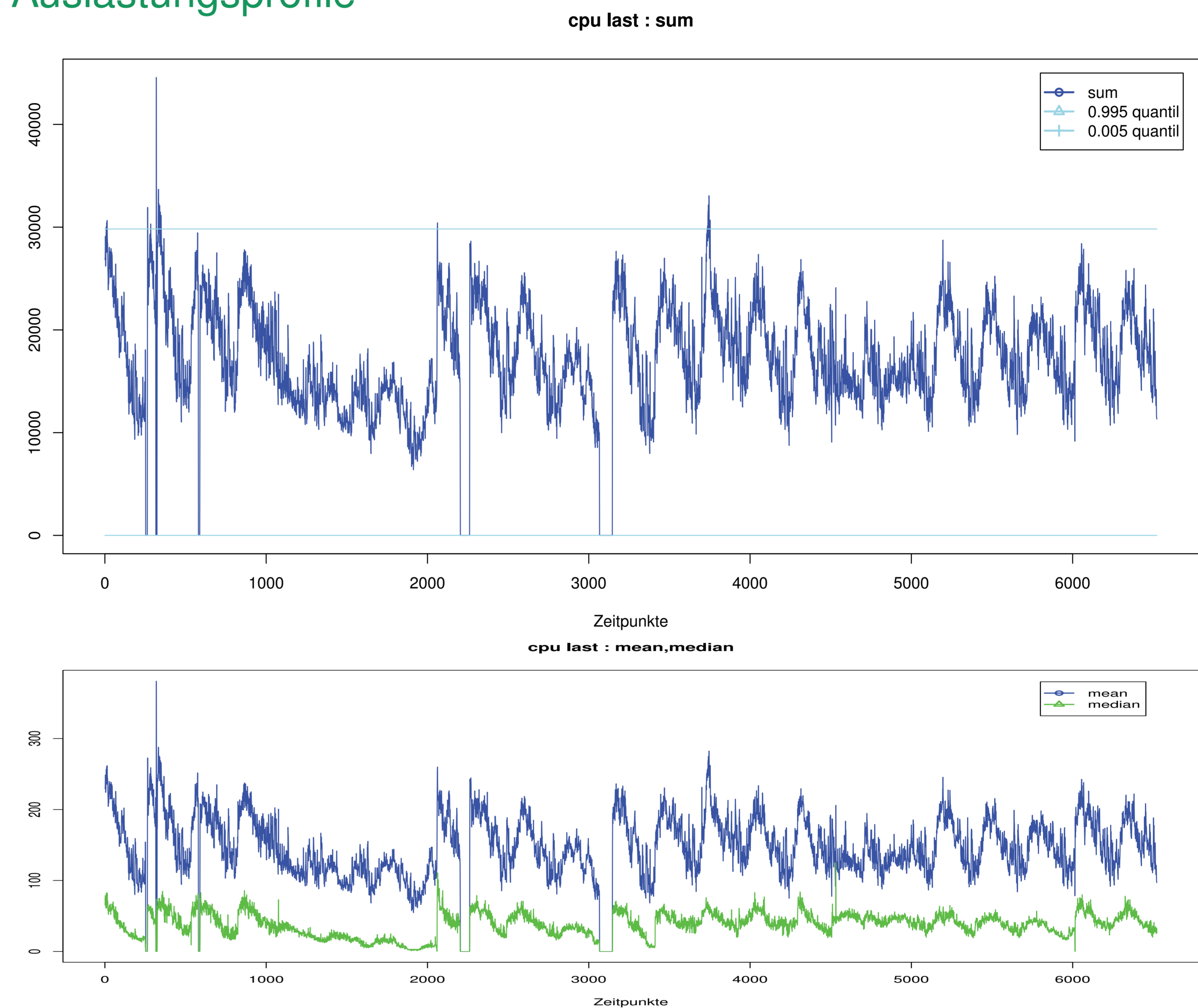
Math. Modellierung

Bin Packing Problem

Jeder Ausgangsrechner i wird als "einzupackendes Objekt" angesehen, das in ein Zielsystem j ("Bin") gepackt wird, so dass die Leistung des Rechners i vom Server j als virtuelle Maschine erbracht wird. Zu jedem Zeitpunkt t benötigt ein Ausgangsrechner Ressourcen im Umfang von r_i^t , und die Summe der von den auf einem Zielsystem virtualisierten Rechnern benötigten Ressourcen soll zu keinem Zeitpunkt die entsprechende Kapazität des Zielsystems überschreiten. Die Anzahl der Zielsysteme, die zum "Einpacken" aller Ausgangsrechner benötigt werden, soll minimiert werden.

Beim Vorliegen einer einzelnen Kapazitätsrestriktion liegt ein sogenanntes **Bin Packing Problem** vor. Bei einer Vielzahl von Kapazitätsrestriktionen für $t = 1, 2, \dots$ spricht man von einem **Multidimensionalen Bin Packing Problem** oder **Vector Bin Packing Problem**.

Auslastungsprofile



Ergebnisse

Klassifikation und Komplexität

Bin Packing Probleme sind **NP-schwere** Optimierungsprobleme. Das multidimensionale Bin Packing Problem ist sogar **APX-schwer**, d.h. es existieren keine polynomialen Approximationsverfahren – falls $P \neq NP$ gilt.

Fitting-Heuristiken

Bei den Fitting-Heuristiken werden die Ausgangsrechner nacheinander, sortiert nach bestimmten Kriterien, einem der bereits vorhandenen Zielsysteme unter Berücksichtigung der Ressourcenanforderungen und weiterer Auslastungskriterien zugeordnet. Falls auf keinem Zielsystem mehr ausreichend Kapazität vorhanden ist, wird ein weiteres eingerichtet und der Ausgangsrechner diesem zugewiesen.

Lokale Suchverfahren

Liegt bereits eine nicht optimale Lösung, also eine Zuordnung von Ausgangsrechnern zu Zielsystemen, vor, so kann diese iterativ verbessert werden. Dazu wird wiederholt das am geringsten ausgelastete Zielsystem aufgelöst und eine zufällige Anzahl von Ausgangsrechnern aus den übrigen Zielsystemen entfernt und eine neue Lösung mittels FirstFit für die nun wieder zuzuweisenden Rechner generiert.

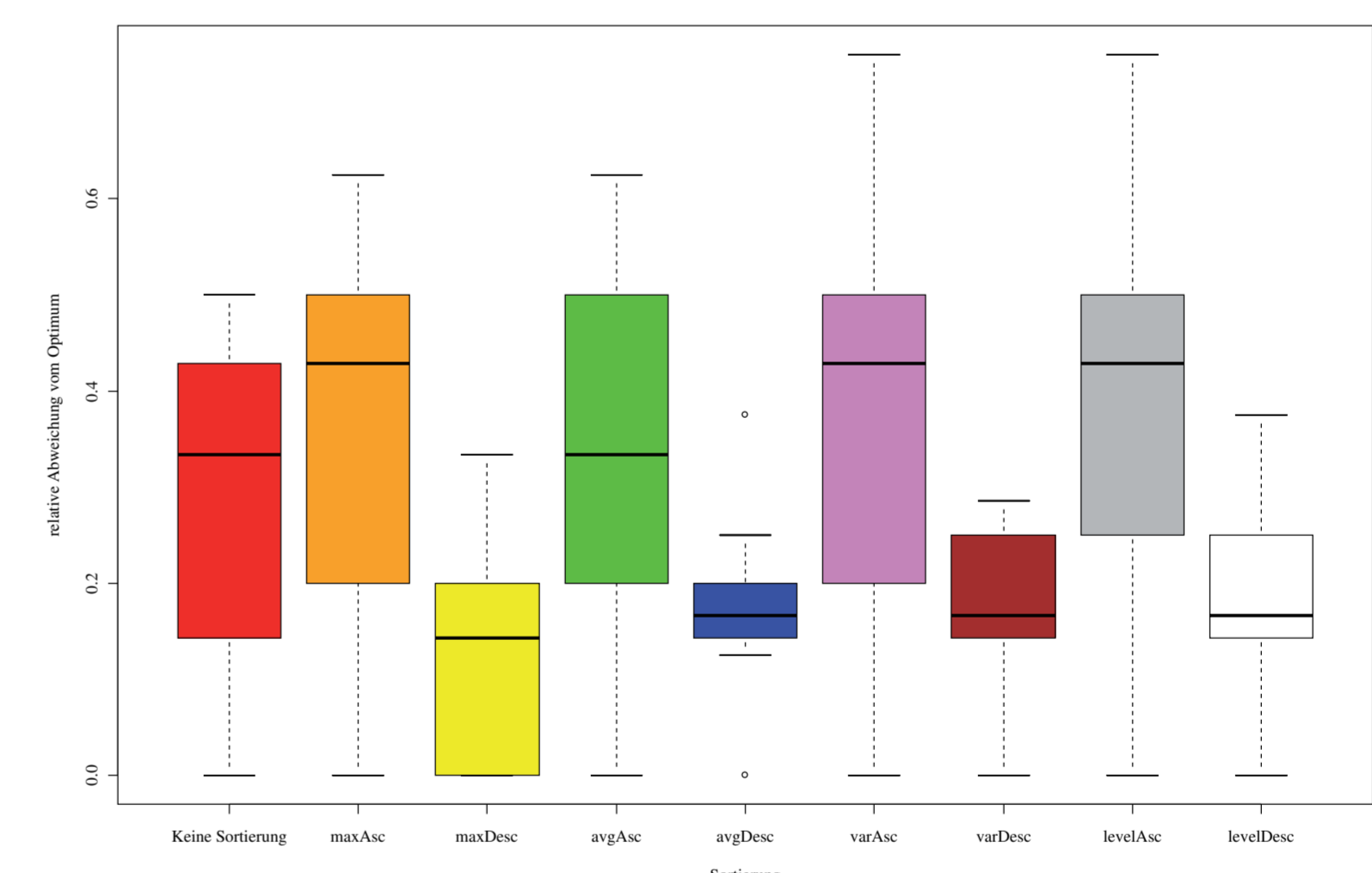


Abbildung 1: Fitting-Heuristiken: Lösungsqualität bei unterschiedlicher Vorsortierung der Ausgangsrechner

Verfahren	mittlere #Zielsysteme	mittlere #Iterationen
FirstFit	11	1
Random Search	8,95	989
First Choice		
Hill Climbing	8,06	225,01
Simulated Annealing	8,0	54,93

Abbildung 2: Lokale Suchverfahren: Durchschnittliche Ergebnisse

Optimalitätsnachweis

Die Optimalitätsprüfung der Lösungen erfolgt durch Vergleich mit unteren Schranken für die benötigte Anzahl von Zielsystemen.