

Das Rucksackproblem der Bundesliga-Manager — optimierte Entscheidungen bei der Kaderzusammenstellung

Janick Frasch, Dennis Janka, Robert Kircheis, Sebastian Sager

Interdisziplinäres Zentrum für Wissenschaftliches Rechnen
Universität Heidelberg
INF 368, 69120 Heidelberg

Zusammenfassung

Die mathematische Disziplin *Optimierung* beschäftigt sich damit, beweisbar beste Lösungen für Fragestellungen in Technik, Gesellschaft, Naturwissenschaften oder Wirtschaft zu finden. Ihr sind große Erfolge zu verdanken im Hinblick auf Effizienzsteigerungen, Kostensparnisse und Sicherheitsverbesserungen, die das Leben in der modernen Wissensgesellschaft jeden Tag entscheidend verbessern.

In diesem Essay wollen wir uns der Optimierung auf eine spielerische Art und Weise nähern, um grundlegende Vorgehensweisen anschaulich erklären zu können. Der Zugang erfolgt über das sogenannte *Kicker Manager Spiel*, einem prototypischen Managerspiel für interessierte Fußballfreunde. Mehrere hunderttausend Menschen, europaweit gar mehrere Millionen, versuchen Jahr für Jahr, ihre Kollegen und Freunde zu übertrumpfen, indem sie bessere Entscheidungen als diese treffen. Beste Voraussetzungen, um einmal zu diskutieren, was denn eine optimale Entscheidung bei der Kaderzusammenstellung wäre, und wie man diese berechnen kann.

Als Motivation, sich mit der Methodologie Optimierung zu befassen, stellen wir eine Optimierungssoftware mit komfortablen Ein- und Ausgabemöglichkeiten frei zur Verfügung.

1 Einleitung

Das *Kicker Manager Spiel* (KMS) wurde erstmals 1991 vom Magazin *Der Kicker* durchgeführt. Seit dieser ersten Postkarten-basierten Version hat sich die Anzahl der Teilnehmer bis zum Jahre 2011 stetig auf mehr als 600.000 in den verschiedenen Spielvarianten erhöht. Zudem gibt es ein komfortables Online-Interface. Gleichzeitig gibt es eine Vielzahl verwandter Spiele in Deutschland und in Europa, beispielsweise von SportBild, Sport1, Comunio, Fantasy Premier League (UK) oder Fanta Calcio (I). Im folgenden betrachten wir das KMS in der *interactive* Version für die 1. Bundesliga als Prototypen, welches mit 215.000 Teilnehmern in der Saison 2011/2012 die beliebteste Variante war. Wir betonen allerdings, dass die grundsätzliche Vorgehensweise leicht auf andere Spielvarianten übertragbar ist. Beim KMS hat jeder Mitspieler¹ ein Budget zur Verfügung, mit dem am Anfang der Bundesligasaison ein Kader von 22 Spielern zusammengestellt wird. Jeder einzelne Spieler hat einen fixen, vom *Kicker* festgelegten Marktwert. Zusätzlich gibt es einige Beschränkungen, beispielsweise eine maximale Anzahl von Spielern aus einem Verein.

Aus dem einmal gekauften Kader werden pro Spieltag 11 Spieler aufgestellt. Je nach deren Leistung an dem real stattfindenden Bundesligaspieltag erhalten diese Spieler Punkte vom *Kicker*, beispielsweise für geschossene Tore oder überzeugende Abwehrarbeit. Jeder Mitspieler erhält nun die gesammelten Punkte der von ihm aufgestellten 11 Spieler. Wer am meisten über die 34 Bundesligaspieltage sammeln kann, gewinnt das Managerspiel.

¹Aus Faulheit verwenden wir die männliche Bezeichnung auch für die erfreulich zahlreichen weiblichen Managerinnen. Man beachte den Unterschied zwischen Mitspieler (= Manager) und Spieler (= Fussballer).

Da das Managerspiel von Jahr zu Jahr wiederholt wird, liegt es nahe, sich an den erreichten Punktzahlen des vergangenen Jahres zu orientieren. Dieses ist die Basis des von uns gewählten Optimierungsansatzes. Wir interessieren uns für die Frage, wie die Lösung des folgenden Problems aussieht:

Wähle 22 Spieler aus, so dass der Kader im Sinne der KMS Spielregeln zulässig ist und die Summe der ggfs. gewichteten Vorjahrespunkte maximiert wird.

Diese Aufgabenstellung entspricht dem Prototyp eines Optimierungsproblems und ist ein klassisches Beispiel für Entscheidungsfindung. Allerdings enthält schon die Formulierung des Optimierungsproblems offensichtlich mindestens eine Annahme, die zweifelhaft ist. Der Fußball wäre kaum so interessant, wenn die neue Saison tatsächlich einfach ein Abbild der vergangenen Saison wäre. Wir relativieren dieses später noch durch die Einführung von Gewichtungsfaktoren, die die Integration von persönlichen Präferenzen und Visionen erlauben. Doch das grundlegende Problem bleibt bestehen: wir kennen die Zukunft nicht. In diesem Sinne sollen optimierte Lösungen also nur als Entscheidungshilfe dienen und keinesfalls dem Anspruch genügen, tatsächlich das KMS zu gewinnen.

In Sektion 2 formulieren wir ein mathematisches Optimierungsproblem, dessen Lösung die obige Problemstellung beantwortet. In Abschnitt 3 beschreiben wir eine Optimierungssoftware, die wir kostenfrei zur Verfügung stellen und erläutern insbesondere, warum selbst die schnellsten Computer nicht einfach alle Möglichkeiten ausprobieren können. In Abschnitt 4 geben wir beispielhaft optimale Lösungen für einige Szenarien an.

2 Mathematisches Optimierungs-Modell

Ein Blick auf die Problemstellung in der Einleitung legt die weiteren Schritte einer mathematischen Annäherung nahe. Zum einen müssen Aussagen darüber getroffen werden, ob ein Spieler in den Kader aufgenommen wird oder nicht. Am einfachsten ist dies mit Hilfe von *Variablen* möglich, die nur den Wert 0 oder 1 annehmen können, sogenannten Binärvariablen. So gilt beispielsweise $x_{\text{Neuer}} = 1$, falls Manuel Neuer vom FC Bayern München Teil unseres Kadern ist, ansonsten gilt $x_{\text{Neuer}} = 0$. Um die Schreibarbeit zu verringern, schreiben wir kürzer x_i , wobei nun i für einen speziellen Spieler steht. Die Menge aller Spieler bezeichnen wir als \mathcal{S} . Der Ausdruck $i \in \mathcal{S}$ bedeutet also, dass i repräsentativ für einen der ca. 500 zur Auswahl stehenden Spieler steht. Ähnliche Men-

Menge	Beschreibung
\mathcal{S}	Gesamtmenge aller Spieler
T	Menge aller Torhüter
A	Menge aller Abwehrspieler
M	Menge aller Mittelfeldspieler
S	Menge aller Stürmer
\mathcal{V}	Menge aller Vereine
V_j	Menge aller Spieler bei Verein j , $j \in \mathcal{V}$

Tabelle 1: Mengen des MILPs.

gen benutzen wir auch, um kurz schreiben zu können, dass wir alle Torwarte, Verteidiger, Mittelfeldspieler oder Angreifer meinen, vergleiche Tabelle 1. Es gilt also insbesondere, dass die Menge

\mathcal{S} die Vereinigungsmenge der Torwart-, Verteidiger-, Mittelfeldspieler- und Stürmer-Mengen ist, also $\mathcal{S} = T \cup A \cup M \cup S$.

Dieser Trick erlaubt uns eine kompakte Schreibweise der *Summe der Vorjahrespunkte*. Wenn wir für jeden Spieler $i \in \mathcal{S}$ die Vorjahrespunkte mit p_i bezeichnen, so ist unter Verwendung des Summenzeichens \sum die zu maximierende Summe

$$\sum_{i \in \mathcal{S}} p_i \cdot x_i.$$

Da die Variablen x_i nur die Werte 0 oder 1 annehmen können, gibt die Summe also genau die Summe der Vorjahrespunkte aller Spieler in unserem Kader an. Die Regeln des KMS schränken die Kaderzusammenstellung ein. So ist ein Kader nur dann zulässig, wenn

(KMS1) der Kaderwert 42,5 Mio EUR nicht übersteigt,

(KMS2) er genau 3 Torhüter, 6 Abwehrspieler, 8 Mittelfeldspieler und 5 Stürmer enthält,

(KMS3) er max. 4 Spieler pro Bundesligamannschaft enthält.

Zudem gibt es noch Einschränkungen für die aufgestellte Elf pro Spieltag:

(KMS4) max. 3 Spieler pro Bundesligamannschaft in der ersten Elf,

(KMS5) 1 Torwart und ein 3-5-2, 4-5-1, 4-4-2, 4-3-3 oder 3-4-3 System (Abwehr - Mittelfeld - Sturm).

Nach diesen Definitionen können wir nun das Optimierungsproblem mathematisch in einer Standardform als sogenanntes *mixed-integer linear program* (MILP) formulieren. Für jeden Spieler $i \in \mathcal{S}$ bezeichnen wir die Transferkosten mit k_i (in Mio EUR) und die Punktwerte der Vorsaison mit p_i .

Unser Ausgangsproblem, einen Kader von 22 Spielern nach den KMS Regeln (KMS1-3) zusammenzustellen, sieht dann wie folgt aus:

$$\max_{x_i, i \in \mathcal{S}} \sum_{i \in \mathcal{S}} p_i \cdot x_i \quad (1a)$$

unter den Bedingungen dass

$$\sum_{i \in \mathcal{S}} k_i \cdot x_i \leq 42,5 \quad (1b)$$

$$\sum_{i \in T} x_i = 3, \quad \sum_{i \in A} x_i = 5, \quad \sum_{i \in M} x_i = 8, \quad \sum_{i \in S} x_i = 6, \quad (1c)$$

$$\sum_{i \in V_j} x_i \leq 4 \quad \forall j \in \mathcal{V}, \quad (1d)$$

$$x_i \in \{0, 1\} \quad \forall i \in \mathcal{S}. \quad (1e)$$

Die Zielfunktion (1a) zielt auf einen Kader mit maximalem Punktwert in der Vorsaison ab. Nebenbedingung (1b) stellt sicher, dass das verfügbare Budget von 42,5 Mio EUR eingehalten wird (KMS1). Nebenbedingungen (1c) stellen sicher, dass ein vollständiger Kader zusammengestellt wird (KMS2); (1d) beschränkt die Spieler auf maximal 4 von jedem Verein (KMS3).

Die verwendeten Konzepte sind prototypisch für die Formulierung von Optimierungsproblemen: zuerst die Definition von Variablen, dann die einer zu maximierenden *Zielfunktion* und zu

erfüllender *Nebenbedingungen* als Funktionen eben dieser Variablen. Dies erlaubt, für jeden beliebigen Kader einen *Vektor* von Variablen x_i anzugeben, die ihn darstellen. Durch Einsetzen in (1b-1d) kann man dann überprüfen, ob der Kader zulässig im Sinne des KMS ist und welchen Wert die Zielfunktion (1a) hat.

Um die individuelle Entwicklung einzelner Spieler und Vereine im Optimierungsmodell berücksichtigen zu können, gibt es zwei Möglichkeiten die Vorjahrespunkte zu gewichten. Zum einen ist jeder Spieler mit einem individuellen Faktor e_i ausgestattet. Hier würde man einen Wert größer als 1 wählen wenn man erwartet, dass ein Spieler im Vergleich zum Vorjahr deutlich mehr Punkte erzielt, etwa z.B. Thomas Kraft, der in der vergangenen Saison 12 Spiele für den FC Bayern München absolvierte, nun aber bei Hertha BSC voraussichtlich Stammtorhüter ist und somit wohl deutlich mehr Spiele absolvieren wird, in denen er die Chance hat Punkte zu sammeln. Einen Wert kleiner als 1 würde man für Spieler wählen, bei denen man erwartet, dass sie in der kommenden Saison weniger Punkte als im Vorjahr erzielen. Neben dem individuellen Faktor gibt es noch die Möglichkeit eine Prognose für das Abschneiden eines Vereins im Vergleich zum letzten Jahr in die Optimierung einfließen zu lassen. Hat ein Team in der letzten Saison besonders gut bzw. schlecht gespielt und man erwartet für diese kommende Saison eine Verschlechterung bzw. Verbesserung, und damit einhergehend auch eine schlechtere bzw. bessere Bewertung der Spieler, so kann man dies abbilden durch einen Faktor $t_{j(i)}$, der gleich ist für alle Spieler $i \in \mathcal{S}$ eines Teams $j \in \mathcal{V}$. Durch Hinzunahme dieser beiden Gewichtungsfaktoren ändert sich die Zielfunktion (1a) somit zu

$$\max_{x_i, i \in \mathcal{S}} \sum_{i \in \mathcal{S}} p_i \cdot e_i \cdot t_{j(i)} \cdot x_i. \quad (2)$$

Im Moment wird durch das Optimierungsproblem noch einen ausgeglichener Kader aus 22 ähnlich starken Spielern bestimmt. Eine solche Zusammenstellung ist allerdings eher ungewöhnlich für das KMS, da man z.B von drei Torhütern nur einen pro Spieltag aufstellen kann und es deshalb sinnvoller sein kann einen sehr starken Torhüter und zwei günstige Ergänzungsspieler zu kaufen (sofern der stärkste gekaufte Torhüter regelmäßig spielt) anstelle von drei ähnlich starken Torhütern. Gängiger ist beim KMS daher die Taktik einen Kader aus 11 bzw 14 teuren Stammspielern zusammenzustellen und mit günstigen Spielern (Marktwert von etwa 0.2 Mio. EUR) aufzufüllen. Diese Taktik lässt sich ebenfalls in unserem Optimierungsproblem anwenden. Dazu werden neue Binärvariablen s_i eingeführt, die angeben, ob ein Spieler Stammspieler ist. In der Zielfunktion strebt man dann eine Maximierung der Vorjahrespunkte dieser Spieler an, also

$$\max_{x_i, s_i, i \in \mathcal{S}} \sum_{i \in \mathcal{S}} p_i \cdot e_i \cdot t_{j(i)} \cdot s_i.$$

anstelle von (2). Zusätzlich muss man noch sichergestellt werden, dass jeder Stammspieler auch überhaupt gekauft wird. Dies geschieht durch Hinzunahme der Nebenbedingungen

$$s_i \leq x_i \quad \forall i \in \mathcal{S}.$$

Durch zusätzliche Nebenbedingungen kann sichergestellt werden, dass möglichst viele der Stammspieler auch gemeinsam im KMS aufgestellt werden können. Wir wählen hierfür die folgenden Nebenbedingungen:

$$1 \leq \sum_{i \in T} s_i \leq 2 \quad 3 \leq \sum_{i \in A} s_i \leq 6 \quad 3 \leq \sum_{i \in M} s_i \leq 6 \quad 1 \leq \sum_{i \in S} s_i \leq 4$$

Je nach KMS-Taktik kommt noch eine Bedingung hinzu, die die Anzahl der Stammspieler im eingekauften Kader beschränkt, etwa auf 14:

$$\sum_{i \in \mathcal{S}} s_i = 14.$$

3 Die Optimierungssoftware

Die nun erreichte Formalisierung unserer ursprünglichen Fragestellung aus Abschnitt 1 erlaubt einen direkten brute-force Ansatz zur Optimierung: das Ausprobieren und Vergleichen sämtlicher Kombinationen. Ähnlich wie in der berühmten Geschichte des indischen Maharadscha mit den verdoppelten Reiskörnern pro Schachfeld ist die Anzahl der Möglichkeiten jedoch astronomisch hoch. Es gibt

$$\binom{500}{22} = \frac{500 \cdot 499 \cdot 498 \cdot \dots \cdot 479}{1 \cdot 2 \cdot 3 \cdot \dots \cdot 22} > \frac{479^{22}}{22^{21}} \approx 6 \cdot 10^{30}$$

verschiedene Möglichkeiten, 22 Spieler aus den ca. 500 vorhandenen auszuwählen. Wenn wir annehmen, dass eine auf 10GHz getaktete CPU pro Takt gleich 1000 verschiedene Kaderzusammenstellungen auswerten und vergleichen kann; und dass wir einen Parallelrechner mit 60 Millionen Prozessoren verwenden dürfen; so ergibt sich immer noch eine Rechenzeit von

$$\frac{6 \cdot 10^{30}}{10 \cdot 10^9 \cdot 1000 \cdot 60 \cdot 10^6} \text{s} = \frac{10^{10}}{60 \cdot 60 \cdot 24 \cdot 365} \text{a} \approx 317 \text{a}$$

mehr als 300 Jahren!

Erfreulicherweise erlauben die methodischen Fortschritte von Optimierungsverfahren, die beweisbar beste Lösung in einem Bruchteil einer Sekunde zu berechnen, und dies auf einem handelsüblichen PC oder Notebook. Hierbei kommt uns zugute, dass das Optimierungsproblem (1) eine spezielle Struktur aufweist: die des *Rucksack* (engl.: knapsack) Problemes. Eine anschauliche Erläuterung des Rucksack-Problemes ist ein Dieb, der sich bei einer gegebenen Menge von Gegenständen für die wertvollsten entscheiden muss. Hierbei muss er berücksichtigen, dass er nur ein gewisses Gewicht oder Volumen transportieren kann. Abstrakt gesehen entspricht dies genau der Managerentscheidung, mit einem limitierten Budget die wertvollsten Spieler auszuwählen.

Diese unscheinbare Problemformulierung tritt als Subproblem in vielerlei mathematischen und wirtschaftlichen Anwendungen auf und wird seit mehr als einem Jahrhundert ausgiebig studiert. Moderne Optimierungssoftware kann diese spezielle Problemklasse daher besonders effizient behandeln.

Da wir unser Problem, einen bestmöglichen Kader einzukaufen, bereits in der MILP-Standardform aufgeschrieben haben, können wir es nun mit geringem Implementierungsaufwand von einer solchen Optimierungssoftware einlesen und lösen lassen. Dabei steht uns eine Vielzahl an kommerziellen, aber auch kostenlosen Programmen zur Verfügung. Wir entscheiden uns, aufgrund seiner besonders einfachen Problembeschreibungssprache, für das GNU Linear Programming Kit (GLPK), das unter [2] kostenlos sowohl für Linux als auch für Windows zum Download bereit steht.

Nach Installation dieser Software kann sich der interessierte Leser unter [1] eine Implementierung unseres Modells herunterladen und diese ausprobieren. Darin findet man nach dem Entpacken für Windows-Nutzer eine `kickermanager.exe` und für Linux-Nutzer ein Python-Script, das mit dem Aufruf `python kickermanager.py` ausgeführt werden kann. Nach Aufruf eines dieser Programme werden im Unterordner `results` drei Textdateien `best11.txt`, `best14.txt` und `best22.txt` erstellt, die den nach Vorjahrespunkten besten 11er, 14er und 22er Kader enthalten. In der Datei `data/teamweights.csv`, die man z.B. mit Microsoft Excel, aber auch mit einem normalen Texteditor verändern kann, lassen sich die Gewichtungsfaktoren der 18 Bundesligateams individuell einstellen (standardmäßig 1). In eigenen `csv`-Dateien für jede Spielergruppe (`defender.csv`, usw.) können zudem die Gewichtungsfaktoren jedes einzelnen Spielers (standardmäßig ebenfalls 1) im Hinblick auf seine potentielle Entwicklung in der kommenden Saison verändert werden.

4 Ergebnisse

Zu drei verschiedenen KMS-Strategien, bei denen je 11, 14 oder 22 Stammspieler gekauft werden wurden von uns nach Vorjahrespunkten optimale Kader berechnet. Dabei wurden die individuellen Gewichtungen einzelner Spieler und Mannschaften in einer selbsternannten Expertenrunde bestimmt. Der Kader mit 22 Stammspielern besteht dann aus: *Starke, Trapp, ter Stegen - Dante, de Jong, Barth, Balogun, Noveski, Schmelzer - Balitsch, Rupp, Reus, Höger, Holtby, Töre, Pinto, Tiffert - Allagui, Esswein, Schieber, Lakic, Shechter*. Dieser Kader ist beim KMS unter dem Namen *IWR Team22* angemeldet und öffentlich einsehbar.

Für den Kader mit 14 Stammspielern ergab sich (Stammspieler sind dabei fett gedruckt): *Schwabke, **Trapp, Starke** - Gulde, Vaaßen, **Noveski, Fuchs, Dante, Schmelzer** - Buchtman, **Holtby, Pinto, Ilicovic, Tiffert Reus, Klement, Sala - Uth, Derstroof, Cisse, Lakic, Schieber***. Dieser Kader nimmt unter dem Benutzernamen *IWR Team14* am KMS Interactive teil.

Der dritte Kader mit 11 Stammspielern besteht aus: *Unnerstall, Radlinger, **Trapp** - Balogun, Brooks, **Noveski, Hummels, Schmelzer, Schulze - Holtby, Hegeler, Reus, Tiffert, Schulz (K'Lautern), Stefanovic A., Yildirim, Buchtman - Lakic, Ayik, Cisse, Schürle, Polter***. Dieser Kader ist unter *IWR Team11* angemeldet.

Alle hier verwendeten Gewichtungen können unter [1] eingesehen werden.

5 Zusammenfassung und Ausblick

Interessant ist es natürlich, die Einflüsse der Präferenzen auf das Optimierungsergebnis zu betrachten. Dies überlassen wir dem geeigneten Leser als Einstieg in unser Optimierungstool. Auch für a posteriori Analysen ist der framework geeignet. So könnte man untersuchen, ob es eine der Bundesliga-Abschlusstabelle entsprechende Gewichtung der Mannschaften gibt, die gute Ergebnisse erzielt hätte. Auch das Kicker-Bewertungssystem als solches lässt sich analysieren und ggfs. verbessern. Unserer Ansicht nach offen ist auch die statistische Korrelation zwischen Spieler- und Teamperformance in einer Saison.

In diesem kurzen essay haben wir exemplarisch an einem Managerspiel den Einsatz von Optimierungsmethoden aufgezeigt. Wir hoffen, damit einen Beitrag zur Überbrückung der leider immer größer werdenden Distanz zwischen aktueller Forschung und interessierter Öffentlichkeit zu leisten.

Literatur

- [1] J. Frasc, D. Janka, R. Kircheis, and S. Sager. Kicker Manager Spiel Software. <http://mathopt.de/KMS>.
- [2] Andrew Makhorin. GNU Linear Programming Kit, version 4.34. <http://www.gnu.org/software/glpk/glpk.html>.