



Fachbereich Informatik

Bachelorarbeit

im Studiengang Informatik

Datenaufbereitung und Netzwerkanalysen in den deutschen Fußball-Ligen

Thomas Keller

09. August 2010

Betreuer: Dipl.-Inf. Darko Obradovic

Prüfer: Prof. Dr. Andreas Dengel

2. Prüfer: Prof. Dr. Didier Stricker

Erklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Thomas Keller

Zusammenfassung

In dieser Arbeit wird eine soziale Netzwerkanalyse im Bereich Fußball durchgeführt. Im Anschluss an eine Einführung über diese Technik werden zwei bisherige Studien über diese Thematik vorgestellt. Die eine beschäftigt sich mit der brasilianischen Liga, die andere mit der niederländischen Nationalmannschaft. Aus den Vereinen, Spielern, Trainern und Beratern der ersten drei deutschen Bundesligen werden anschließend verschiedene Netzwerke erstellt, die visualisiert und analysiert werden. Ein Spielernetzwerk verbindet alle Spieler, die schon einmal zusammengespielt haben, um beispielsweise die zentralsten Spieler zu ermitteln. Darüber hinaus entstehen bipartite Netzwerke, die aus zwei verschiedenen, disjunkten Mengen bestehen. Dies betrifft beispielsweise die Verbindung von Spielern zu Beratern, was zu interessanten Erkenntnissen über die Beraterstruktur bei den Vereinen führt. Zudem zeigt ein Netzwerk, das aus dem Transferfluss der Bundesligavereine besteht, die häufigen eher lokalen Vereinswechsel ihrer Spieler.

Abstract

This bachelor thesis is about a social network analysis in the domain football. After an introduction about this technique, two earlier studies about this topic are presented. One of them analysed the brazilian league, the other was about the dutch national team. Subsequently different networks are created from the teams, players, trainers and advisers of the first three german football leagues and are visualized and analysed. A player network connects all persons, who played together in a team yet, for instance in order to determine the most central players. Furthermore bipartite networks, which consist of two different disjoint sets, are investigated. For example the connection between players and advisers shows interesting results about the structure of advisers in the teams. Besides that a network demonstrates the transfer flow of the football teams and points out that players often move to clubs, which are in the same region than the previous team.

Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | Einleitung | 1 |
| 2 | Datensammlung | 3 |
| 2.1 | Entity-Relationship-Diagramm | 3 |
| 2.2 | Auslesen der Daten | 4 |
| 3 | Grundlagen | 7 |
| 3.1 | Verschiedene Graphtypen | 7 |
| 3.2 | Metriken für klassische Graphen | 7 |
| 3.3 | Metriken für bipartite Graphen | 9 |
| 3.4 | Zentralität | 10 |
| 3.5 | Zusammenhängende Teilgruppen | 12 |
| 4 | Frühere Studien | 15 |
| 4.1 | Brasilianische Liga | 15 |
| 4.2 | Niederländische Nationalmannschaft | 17 |
| 5 | Netzwerkanalysen in den Bundesligen | 21 |
| 5.1 | Spielernetzwerk | 21 |
| 5.1.1 | Normales Spielernetzwerk | 21 |
| 5.1.2 | Modifiziertes Spielernetzwerk | 24 |
| 5.2 | Vereinsnetzwerke | 26 |
| 5.2.1 | Flow-Netzwerk | 26 |
| 5.2.2 | Ego-Netzwerk | 28 |
| 5.3 | Netzwerk Spieler-Berater | 30 |
| 5.4 | Netzwerk Verein-Berater | 32 |
| 5.5 | Netzwerk Spieler-Trainer | 34 |
| 5.6 | Netzwerk Verein-Trainer | 36 |
| 6 | Netzwerkanalysen im europäischen Fußball | 39 |
| 6.1 | Spielernetzwerk | 39 |

| | | |
|----------|------------------------------|-------------|
| 6.2 | Flow-Netzwerk | 40 |
| 7 | Fazit | 45 |
| 7.1 | Zusammenfassung | 45 |
| 7.2 | Ausblick | 45 |
| | Abbildungsverzeichnis | ix |
| | Tabellenverzeichnis | xi |
| | Literaturverzeichnis | xiii |

Kapitel 1

Einleitung

Fußball ist eine der weltweit bedeutendsten Sportarten und man hat bei der Weltmeisterschaft 2010 erneut gesehen, wie groß das Interesse der meisten Menschen an diesem Sport ist. Auch in Deutschland sind die meisten Stadien in der Bundesliga sehr gut gefüllt und die TV-Übertragungen haben meist sehr hohe Einschaltquoten. In dieser Arbeit werden einige Aspekte der deutschen Fußballligen mit der Technik der sozialen Netzwerkanalyse aufgezeigt.

Die soziale Netzwerkanalyse ist eine Methode aus den Sozial- und Verhaltenswissenschaften. Hierbei bringt man Akteure und deren Beziehungen in ein soziales Netzwerk, um anschließend bestimmte Analysen durchführen zu können (9).

Die ersten drei Fußball-Ligen der Saison 2009/2010 in Deutschland werden mit ihren Vereinen, Spielern, Trainern und Beratern von einer Internetseite ausgelesen, um hiermit soziale Netzwerke zu erstellen. Beispielsweise werden alle aktuellen Spieler der ersten drei Ligen in einem Netzwerk verbunden, wenn sie jemals zusammen in einem Verein gespielt haben. So kann etwa die Zentralität bestimmter Spieler oder die Profis, die mit den meisten anderen Spielern schon zusammen gespielt haben, ermittelt werden.

Zu Beginn dieser Arbeit wird in Kapitel 2 die Datensammlung und Aufbereitung vom Fußball-Portal <http://www.transfermarkt.de> beschrieben. Anschließend werden in Kapitel 3 Grundlagen aus der Graphentheorie erklärt, um die verschiedenen Metriken der sozialen Netzwerkanalyse zu verstehen. Weiterhin werden in Kapitel 4 frühere Durchführungen einer sozialen Netzwerkanalyse im Bereich Fußball vorgestellt. Der Hauptteil der Arbeit in Kapitel 5 besteht aus der eigentlichen Analyse ausgewählter Aspekte im deutschen Fußball. Kapitel 6 besteht aus einem Einblick in eine mögliche Erweiterung auf die europäische Ebene. Den Abschluss bildet Kapitel 7 mit einem kurzen Fazit und einem Ausblick auf mögliche weitere Studien zu dieser Thematik.

Kapitel 2

Datensammlung

Die Daten der ersten drei Ligen werden von einer Internetseite ausgelesen. Das Fußball-Portal <http://www.transfermarkt.de> wird ausgewählt, da man daraus etliche Informationen über die Spieler und Trainer der Bundesligavereine ermitteln kann.

2.1 Entity-Relationship-Diagramm

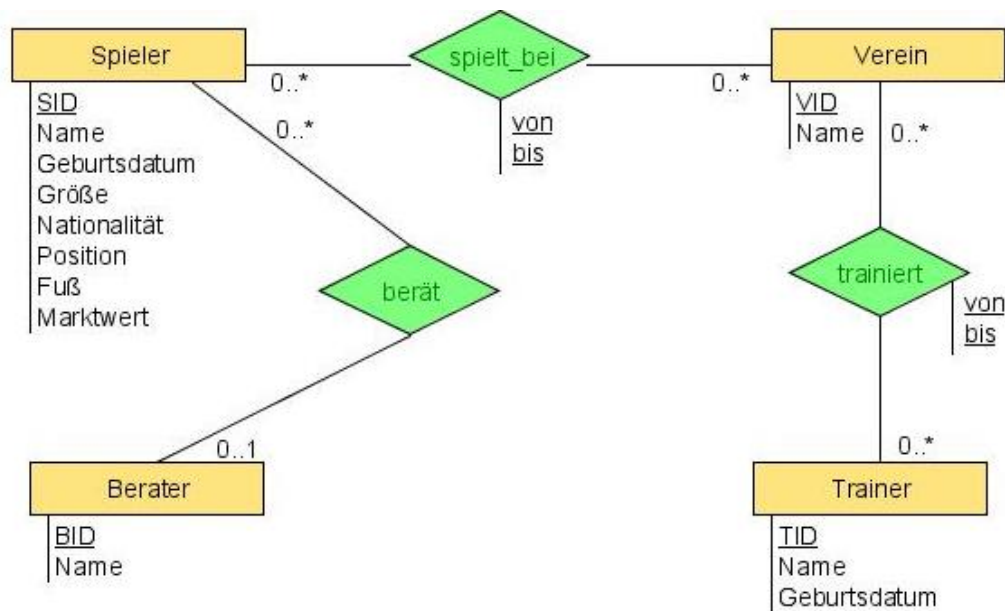


Abbildung 2.1: Entity-Relationship-Diagramm

Zu Beginn wird ein Entity-Relationship-Diagramm (deutsch: Gegenstands-Beziehungs-Modell) entworfen, das aus einzelnen Objekten (Entitäten) besteht, die über Beziehungen (Relationships) miteinander verbunden sind (1). Diese wiederum besitzen Kardinalitäten, die angeben, an wie vielen Beziehungen eine Entität teilnehmen kann. Die Objekte werden mit Hilfe von Attributen genauer beschrieben.

Das obige Entity-Relationship-Diagramm (Abbildung 2.1) besteht aus vier Entitäten. Neben den Entitäten Spieler, Trainer und Verein werden die Berater als weitere Untersuchungsgruppe eingeführt. Für alle vier Mengen wird eine eindeutige ID als Primärschlüssel hinzugefügt. Beziehungen bestehen zuallererst zwischen Spielern und Vereinen. Hierbei spielt der Zeitraum des Verbleibs bei einem Verein eine wichtige Rolle für die spätere Analyse. Um Spieler, die beispielsweise verliehen und nach der Rückkehr wieder nach kurzer Zeit weiterverkauft wurden, nicht in der Datenbank zu speichern, werden nur Stationen, die länger als 80 Tage dauern, gesichert. In die Untersuchung werden nur Aufenthalte bei Vereinen aus den ersten drei Bundesligen einbezogen. Aus Vereinfachungsgründen wird für alle Spielertransfers als Transferzeitpunkt im Sommer der 01.07. und im Winter der 01.01. angegeben. Für die Beziehung zwischen einem Trainer und einem Verein kann man problemlos das exakte Datum speichern, da es hierbei keine Leihgeschäfte gibt. Die dritte Beziehung besteht zwischen den Entitäten Spieler und Berater. Hierbei ist zu beachten, dass man nur die jeweils aktuelle Beraterfirma eines Spielers vom Fußball-Portal ermitteln kann.

Zur Entität Spieler kann man auf der Internetseite einige Attribute finden. Diese reichen von Name, Geburtsdatum, Größe, Nationalität, Position, starkem Fuß bis zu einem geschätzten Marktwert. Im Bereich der Trainer kann man ebenfalls Name, Geburtsdatum und Nationalität finden. Bei Vereinen und Beratern waren leider keine für die vorliegende Untersuchung verwertbaren Informationen zu entdecken.

2.2 Auslesen der Daten

Als Grundlage für das Auslesen der Daten wird das Entity-Relationship-Modell mit Hilfe von SQL-DDL (Data Definition Language) in ein Datenbankschema überführt. Bei der dazu notwendigen Abbildung ins Relationenmodell ist zu erwähnen, dass, neben den neuen Relationen Spieler, Trainer, Berater und Verein, auch „spielt_bei“ und „trainiert“ hinzukommen, da diese n:m-Beziehungen sind. Da jeder Spieler bei transfermarkt.de maximal einem Berater zugeordnet ist, reicht hinsichtlich der Relation Spieler-Berater ein zusätzliches Attribut BID in der Relation Spieler als Fremdschlüssel auf die

Relation Berater aus. Anschließend werden nun die Bundesligavereine der Seite <http://www.transfermarkt.de> durchgegangen, um die einzelnen Spieler und Trainer mit ihren Attributen herauszulesen und mit Hilfe der Java-Datenbankschnittstelle JDBC (Java Database Connectivity) in eine Datenbank zu speichern. Dadurch hat man nun die Daten für die späteren Analysen gesammelt und aufbereitet.

Kapitel 3

Grundlagen

Da die Daten aus der Datenbank in eine Netzwerkrepräsentation überführt werden sollen, werden nun Grundlagen aus der Graphentheorie und Metriken der sozialen Netzwerkanalyse für die späteren Untersuchungen erläutert. Diese stammen weitgehend aus (4).

3.1 Verschiedene Graphtypen

Ein klassischer **Graph** $G=(V,E)$ besteht aus einer Menge von Knoten V und Verbindungen (Kanten) zwischen je zwei Knoten E . Im Gegensatz dazu besteht ein **bipartiter Graph** aus zwei disjunkten Knotenmengen, wobei ausschließlich Links zwischen Knoten aus verschiedenen Knotenmengen existieren dürfen. Die zwei Knotenmengen werden als Top-Knoten und Bottom-Knoten unterschieden. Weiterhin besitzt ein **gerichteter Graph** Kanten, bei denen Pfeile die Richtung angeben. **Gewichtete Graphen** zeichnen sich dadurch aus, dass ihre Kanten mit einem Wert versehen sind.

3.2 Metriken für klassische Graphen

Hier werden einige Metriken für klassische Graphen vorgestellt. Die Anzahl der Knotenmenge V wird mit n abgekürzt, m steht für die Anzahl der Links.

Grad

Unter dem Grad (englisch: degree) eines Knotens versteht man die Anzahl der Nachbarknoten. In diesem Zusammenhang kann man den **Durchschnittsgrad** eines Graphen mit der Formel $k=(2m/n)$ berechnen. Mit der Metrik

der **Gradverteilung** ist die Wahrscheinlichkeit gemeint, mit der ein zufällig ausgewählter Knoten einen bestimmten Grad besitzt.

Dichte

Die Dichte eines Graphen G (englisch: density) wird mit der Formel

$$\rho(G) = \frac{2m}{n(n-1)}$$

ausgedrückt. Sie definiert den Quotienten aus der tatsächlichen Anzahl der Links in einem bestimmten Graphen und der Anzahl aller möglicher Links. Dies entspricht der Wahrscheinlichkeit, dass zwei zufällig ausgewählte Knoten verlinkt sind.

Distanz

Als Distanz zwischen zwei Knoten bezeichnet man die minimale Anzahl an Links, um von einem Knoten zum anderen Knoten zu gelangen. Somit kann man die **durchschnittliche Distanz** eines Graphen aus den einzelnen Entfernungen berechnen. Das Maximum der Distanzen aller Knotenpaare wird als **Diameter** bezeichnet. Ein Graph wird **zusammenhängend** genannt, wenn man von jedem Knoten jeden anderen Knoten erreichen kann.

Clusterkoeffizient

Der Clusterkoeffizient macht eine Aussage über den Grad der Verlinkung in einem Graphen. Im Falle eines zufällig ausgewählten Knotens bestimmt er die Wahrscheinlichkeit, dass zwei Nachbarn miteinander verlinkt sind. Sei t die Anzahl der Verbindungen der Nachbarknoten und $\deg(v)$ der Grad, so kann der Clusterkoeffizient eines Knotens v folgendermaßen ausgedrückt werden:

$$cc(v) = \frac{2t}{\deg(v) \cdot (\deg(v) - 1)}$$

Dies entspricht der Dichte der Nachbarschaft eines Knotens und wird als lokaler Clusterkoeffizient bezeichnet. Den globalen Clusterkoeffizienten erhält man dementsprechend aus dem Durchschnitt aller lokalen Koeffizienten eines Graphen.

Assortativity coefficient

Dieser Koeffizient, der in (6) definiert wurde, hat den Wertebereich $-1 \leq r \leq 1$ und gibt an, ob die Knoten eines Graphens eher mit Knoten ähnlichen ($r > 0$)

oder sehr verschiedenen Grads ($r < 0$) verbunden sind. Beispielsweise bedeutet ein Ergebnis von 1, dass alle Knoten nur Nachbarn haben, die denselben Grad besitzen. Der Assortativity coefficient ist durch die nachfolgende Formel definiert, wobei j_i und k_i die Gradwerte der Knoten an den Enden des i-ten Links mit $i=1\dots L$ sind.

$$r = \frac{L^{-1} \sum_i j_i k_i - [L^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}{L^{-1} \sum_i \frac{1}{2}(j_i^2 + k_i^2) - [L^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}.$$

3.3 Metriken für bipartite Graphen

Die Grundmetriken für bipartite Graphen sind Erweiterungen der Metriken für klassische Graphen.

Grad

Den Durchschnittsgrad kann man für die Top-Knoten und Bottom-Knoten bestimmen. Insgesamt stimmen aber sowohl Durchschnittsgrad und Gradverteilung mit der klassischen Graphentheorie überein.

Dichte

Die Dichte in einem bipartiten Graphen wird mit der Gleichung

$$\rho(G) = \frac{m}{n_A n_B}$$

ermittelt, wobei m für die Anzahl der Links und n_A und n_B jeweils für die Anzahl der beiden disjunkten Knotenmengen A und B stehen.

Clusterkoeffizient

Beim Clusterkoeffizienten für bipartite Graphen kann die klassische Formel nicht verwendet werden, weil die Nachbarn eines Knotens nie verlinkt sein können, da nur Knoten aus disjunkten Mengen verbunden sein dürfen. Um Nachbarschaftszusammenhänge auch bei bipartiten Graphen feststellen zu können, wird folgende Formel für Knotenpaare u und v derselben Knotenmenge benutzt:

$$cc(u, v) = \frac{|N(u) \cap N(v)|}{|N(u) \cup N(v)|}$$

Somit teilt man die Anzahl der gemeinsamen Nachbarn zweier Knoten durch die Vereinigung beider Nachbarn. Aus keinen gemeinsamen Nachbarn folgt

ein Wert von 0, die gleiche Nachbarschaft führt zu einem Ergebnis von 1. Je nach Nähe zu 0 oder 1 kann man so die Ähnlichkeit der Nachbarschaften der beiden Akteure erkennen. Nun wird der Durchschnitt aller Clusterkoeffizienten eines Knotens gebildet, wobei nur Paare verwendet werden, bei denen mindestens ein gemeinsamer Nachbar vorhanden ist.

$$cc(u) = \frac{\sum_{v \in N(N(u))} cc(u, v)}{|N(N(u))|}$$

Der Clusterkoeffizient eines Graphen wird demnach mit folgender Formel berechnet, wobei n_A und n_B für die Anzahl der beiden disjunkten Knotenmengen A und B stehen.

$$cc(G) = \frac{n_A cc(A) + n_B cc(B)}{n_A n_B}$$

Zudem gibt es zwei verschiedene Varianten des Clusterkoeffizienten. Bei sehr unterschiedlichen Gradwerten kann es vorkommen, dass trotz Ähnlichkeit der Nachbarschaften ein sehr geringer Wert ermittelt wird. Beispielsweise kommt bei einem Knoten mit hohem Grad und einem Knoten mit niedrigem Grad eine kleine Zahl heraus, obwohl die Nachbarn eines Knotens eine Teilmenge der anderen Nachbarmenge sein können. Dies wird durch folgende Veränderung verhindert:

$$cc_{min}(u, v) = \frac{|N(u) \cap N(v)|}{\min(|N(u)|, |N(v)|)}$$

Die nachfolgende Umkehrung von $cc_{min}(u, v)$ führt ausschließlich zu hohen Ergebnissen, wenn beide Nachbarschaften ähnliche Gradwerte haben und eine starke Überlappung besteht. $cc_{max}(u, v)$ erreicht nur das Resultat 1, wenn beide Nachbarschaften identisch sind.

$$cc_{max}(u, v) = \frac{|N(u) \cap N(v)|}{\max(|N(u)|, |N(v)|)}$$

3.4 Zentralität

Die Zentralität macht eine Aussage über die Wichtigkeit einzelner Knoten. Es gibt verschiedene Definitionen von Zentralität, die im Folgenden behandelt werden. Man unterscheidet zwischen Degree Zentralität, Closeness Zentralität und Betweenness Zentralität. Die Zentralitätsdefinitionen stammen aus (9).

Degree Zentralität

Die Degree Zentralität ist die einfachste Zentralitäts-Definition. Der zentralste Akteur ist in diesem Fall der Knoten mit den meisten Nachbarknoten. Daher kann die obige Definition des Grads wieder benutzt werden. Um nur Werte zwischen 0 und 1 zu erhalten, wird der einzelne Degree-Wert durch die maximal mögliche Anzahl an Nachbarknoten g geteilt.

$$C'_D(n_i) = \frac{d(n_i)}{g-1}.$$

Closeness Zentralität

Im Fall der Closeness Zentralität geht es um die durchschnittliche Distanz zu allen anderen erreichbaren Knoten. Hierbei kann der zentralste Akteur am schnellsten mit den anderen Akteuren interagieren. Die Closeness eines Knotens wird mit der Distanz d folgendermaßen definiert:

$$C_C(n_i) = \left[\sum_{j=1}^g d(n_i, n_j) \right]^{-1}$$

Deshalb kann $C_C(n_i)$ maximal $(g-1)^{-1}$ werden, wenn der Knoten zu allen anderen Akteuren eine Verbindung besitzt. Somit wird der Closenesswert standardisiert, indem man die obige Formel mit $g-1$ multipliziert:

$$C'_C(n_i) = \frac{g-1}{\sum_{j=1}^g d(n_i, n_j)}$$

Betweenness Zentralität

Die Betweenness Zentralität sagt aus, inwieweit ein bestimmter Knoten die Kontrolle über die Pfade zwischen nichtbenachbarten Knoten besitzt. Ein Akteur besitzt eine maximale Betweenness Zentralität, wenn er auf allen kürzesten Pfaden zwischen unterschiedlichen Knoten liegt. Wenn man die Anzahl der kürzesten Pfade zwischen zwei Knoten j und k als g_{jk} und die Häufigkeit des Auftretens eines Knotens n_i in den kürzesten Pfaden mit $g_{jk}(n_i)$ bezeichnet, wird die Betweenness Zentralität eines Knotens n_i folgendermaßen definiert, wobei j und k unterschiedlich zu i sein müssen:

$$C_B(n_i) = \sum_{j < k} \frac{g_{jk}(n_i)}{g_{jk}}$$

Da das Maximum $(g-1)(g-2)/2$ ist, wird der obige Wert standardisiert, indem man durch das Maximum teilt.

$$C'_B(n_i) = \frac{C_B(n_i)}{(g-1)(g-2)/2}$$

Im Endeffekt geht es bei der Degree- und der Closeness-Zentralität um den Grad der Abhängigkeit eines Akteurs von den anderen Akteuren. Wenn der jeweilige Akteur direkte bzw. kurze Wege zu vielen anderen Akteuren besitzt, ist er recht wenig auf andere angewiesen. Im Gegensatz dazu misst die Betweenness-Zentralität eines Akteurs, inwieweit die anderen Akteure von diesem abhängig sind (2).

3.5 Zusammenhängende Teilgruppen

Zusammenhängende Teilgruppen sind Teilgruppen von Akteuren, die relativ stark verknüpft sind. Es gibt verschiedenen Definitionen von zusammenhängenden Teilgruppen, die im Folgenden vorgestellt werden und aus (9) stammen. In der späteren Arbeit wird nur die erste Cliques-Definition benutzt, da die verwendeten Tools bei der Berechnung der anderen Definitionen aufgrund der großen Datenmenge überfordert waren.

Clique

Eine Clique eines Graphen ist ein maximaler, vollständiger Teilgraph, der aus mindestens drei Knoten besteht. Dies bedeutet, dass alle Knoten einer Clique benachbart sind und keine Teilmenge einer anderen Clique selbst als Clique bezeichnet wird. Damit nicht alle benachbarten Knotenpaare in diese Definition fallen, werden nur Mengen gezählt, die drei oder mehr Knoten beinhalten.

n-Clique

Bei der n-Clique geht es darum, dass die einzelnen Mitglieder sich recht schnell erreichen können. Beispielsweise bedeutet eine 2-clique, dass alle Akteure eine kürzeste Verbindung besitzen, die maximal 2 ist. Somit wird schnell klar, dass eine 1-Clique eine normale Clique ist. Das Problem bei n-Cliques besteht darin, dass sie nicht unbedingt zusammenhängend sein müssen. Dies liegt daran, dass die kürzesten Pfade nicht zwingend zum Teilgraphen gehören müssen, was dazu führt, dass der Diameter größer als n sein kann oder die n-Clique sogar nicht zusammenhängend ist.

n-Clan, n-Club

Um die obige Problematik der n-Cliques zu beheben, gibt es zwei verschiedene Möglichkeiten: Bei den n-Clans werden alle n-Cliques ausgeschlossen, die

einen Diameter haben, der größer als n ist. Ein n -Club ist ein maximaler Teilgraph, der den Diameter n besitzt.

Kapitel 4

Frühere Studien

Hier werden zwei Studien präsentiert, die sich jeweils mit sozialer Netzwerkanalyse im Bereich Fußball beschäftigt haben.

4.1 Brasilianische Liga

In dieser im Oktober 2004 veröffentlichten Arbeit geht es um die Analyse der brasilianischen Liga (7). Alle 13411 Spieler, die von 1971-2002 jemals dort gespielt haben, wurden zusammen mit den 127 Mannschaften zuerst in ein bipartites Netzwerk gebracht. Zu Beginn der Arbeit wurde die Anzahl der Tore pro Spiel jedes Vereins erläutert. Der beste Verein (Sao Caetano) schaffte 1,73 Tore pro Spiel, Colatina erreichte mit nur 0,22 Toren pro Spiel einen Negativrekord. Interessant war, dass der Durchschnitt fast genau 1 ist. Anschließend wurde die Gradverteilung im bipartiten Netzwerk ermittelt. Durchschnittlich spielte ein Spieler bei 1,37 Vereinen, wobei Dadá Maravilha mit 11 Vereinen die geringste Vereinstreue bewies. Abbildung 4.1 zeigt sowohl die Wahrscheinlichkeitskurve der Spieler ($P(N)$ über die Anzahl der Vereine N) als auch die der Vereine ($P(S)$ über die Anzahl der Vereine N). Während man $P(N)$ mit $10^{-0,38N}$ nähern konnte, gab es für $P(S)$ keine Näherung, was mit der geringen Zahl der Vereine begründet wurde.

Dann wurde ein Netzwerk analysiert, das ausschließlich aus Spielern besteht, die verbunden sind, wenn sie jemals zusammengespielt haben. Insgesamt gab es 315566 Kanten bei den 13411 Knoten. In Abbildung 4.2 wurden die Ergebnisse des Brazilian Soccer Player (BSP) Netzwerks und zweier Zufallsnetzwerke verglichen. Der durchschnittliche Grad war 47,1, was bedeutet, dass ein Spieler im Durchschnitt mit etwa 47 Spielern zusammengespielt hat. Der sehr hohe Clusterkoeffizient von 0,79 zeigt einen gewaltigen Unterschied zu den sehr niedrigen Clusterkoeffizienten der beiden Zufallsgraphen. Dieses

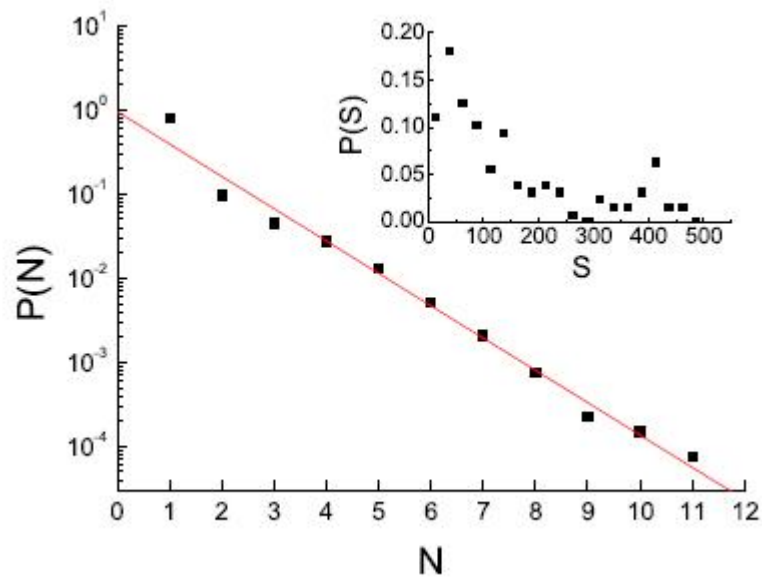


Abbildung 4.1: Gradverteilung im brasilianischen Netzwerk aus (7)

Phänomen zeigt sich bei sehr vielen sozialen Netzwerken. Weiterhin wurde eine durchschnittliche Distanz von 3,29 ermittelt, was recht wenig erscheint, da die Analyse mehr als 30 Saisons umfasst.

| | BSP | Erdős-Rényi | Configuration |
|-----------|---------|-------------|---------------|
| v | 13,411 | 13,411 | 13,411 |
| e | 315,566 | 315,443 | 345,294 |
| \bar{k} | 47.1 | 47.1 | 51.5 |
| C | 0.790 | 0.004 | 0.008 |
| A | 0.12 | 0.00 | 0.46 |
| D | 3.29 | 2.84 | 2.85 |

Abbildung 4.2: Vergleich des brasilianischen Netzwerks mit Zufallsgraphen mit Anzahl der Knoten v , Anzahl der Kanten e , Durchschnittsgrad \bar{k} , Clusterkoeffizient C , Assortativity coefficient A , durchschnittliche kürzeste Distanz D aus (7)

Sehr interessant ist auch der in Abbildung 4.3 gezeigte zeitliche Verlauf der verschiedenen Metriken. Die stetige Zunahme des durchschnittlichen Grads im Laufe des Beobachtungszeitraums wurde sowohl mit einer längeren

Dauer eines Spielers im Profifußball als auch mit der zunehmenden Transfer-rate erklärt. Diese geringere Vereinstreue ist auch ein Grund für den abneh-menden Clusterkoeffizienten. Darüber hinaus wurde der vermehrte Wechsel der stärksten Spieler ins Ausland als weitere Ursache dafür angesehen.

| | 1975 | 1980 | 1985 | 1990 | 1995 | 2002 |
|-----------|--------|---------|---------|---------|---------|---------|
| v | 2,490 | 6,420 | 8,797 | 10,329 | 11,629 | 13,411 |
| e | 48,916 | 128,424 | 181,293 | 219,968 | 254,371 | 315,566 |
| \bar{k} | 39.3 | 40.0 | 41.2 | 42.6 | 43.7 | 47.1 |
| C | 0.84 | 0.83 | 0.82 | 0.81 | 0.80 | 0.79 |
| A | 0.02 | 0.06 | 0.06 | 0.07 | 0.08 | 0.12 |
| D | 3.17 | 3.35 | 3.39 | 3.27 | 3.28 | 3.29 |

Abbildung 4.3: zeitlicher Verlauf der Metriken im brasilianischen Netzwerk aus (7), Bedeutungen der Variablen wie in Abbildung 4.2

4.2 Niederländische Nationalmannschaft

In dieser Arbeit wurde das holländische Nationalteam mittels sozialer Netzwerkanalyse genauer untersucht (3). Hierzu wurden alle Spieler, die jemals für die Nationalelf der Niederlande gespielt haben, in ein Netzwerk gebracht und verbunden, wenn sie mindestens ein Spiel gemeinsam gespielt haben. Abbildung 4.4 zeigt die Ergebnisse der Analyse. Insgesamt 691 Spieler wurden mit 10450 Links verbunden, was zu einer Dichte von 0,044 führt. Der recht hohe Clusterkoeffizient (0,75) und die recht niedrige durchschnittliche kürzeste Distanz (4,49) verdeutlichen, wie auch das brasilianische Netzwerk, das Kleine-Welt-Phänomen (englisch: small world phenomenon). Dieses besagt, dass man in realen sozialen Netzwerken, trotz sehr geringer Dichte, sehr schnell von jeder Person zu einer anderen Person gelangt.

Harry Dénis hat mit 117 anderen Nationalspielern zusammengespield und besitzt den höchsten Grad aller Spieler. Dagegen weist Edwin van der Sar, der insgesamt die meisten Länderspiele für die Niederlande absolviert hat, nur einen Grad von 97 auf. Darüber hinaus hat Harry Dénis neben den meisten Nachbarn im Netzwerk auch den geringsten Clusterkoeffizienten (0,17). Die größte Distanz von einem Knoten zum anderen liegt bei 11, wobei dies bei 324 Knotenpaaren der Fall ist. Im Bereich der Zentralität ist Roel Wiersma der zentralste Spieler, wenn man von der Closeness ausgeht. Der Wert liegt

bei 0,32, wobei man beachten muss, dass hier der Kehrwert, also die durchschnittliche Distanz, als Closeness bezeichnet wird. Die höchste Closeness besitzen die Spieler, die um das Jahr 1950 gespielt haben, was plausibel ist, wenn man bedenkt, dass die niederländische Nationalelf etwa 100 Jahre besteht.

| Metric | Value |
|---------------------------|--------|
| Number of nodes | 691 |
| Number of links | 10450 |
| Density | 0.044 |
| Average degree | 30.25 |
| Assortativity coefficient | -0.063 |
| Average distance | 4.49 |
| Diameter | 11 |
| Eccentricity | 8.59 |
| Clustering coefficient | 0.75 |
| Algebraic connectivity | 0.16 |

Abbildung 4.4: Metriken im niederländischen Netzwerk aus (3)

In Abbildung 4.5 wird der Verlauf der Ergebnisse der verschiedenen Metriken erläutert. Sowohl der durchschnittliche Grad als auch die durchschnittliche Distanz und die größte Distanz steigen nahezu linear an. Während die Dichte immer kleiner wird, bleibt der Clusterkoeffizient fast identisch, wobei eine geringe Abwärtstendenz vorhanden ist.

| Metric | 1926 | 1946 | 1966 | 1986 | 2008 |
|---------------------------|--------|-------|-------|-------|--------|
| Number of nodes | 181 | 282 | 427 | 556 | 691 |
| Number of links | 1956 | 3170 | 5190 | 7575 | 10450 |
| Density | 0.12 | 0.080 | 0.057 | 0.049 | 0.044 |
| Average degree | 21.61 | 22.48 | 24.31 | 27.25 | 30.25 |
| Assortativity coefficient | - 0.17 | -0.16 | -0.16 | -0.11 | -0.063 |
| Average distance | 2.32 | 2.70 | 3.37 | 3.88 | 4.49 |
| Diameter | 4 | 6 | 8 | 10 | 11 |
| Eccentricity | 3.48 | 4.58 | 6.16 | 7.54 | 8.59 |
| Clustering coefficient | 0.77 | 0.76 | 0.76 | 0.75 | 0.75 |
| Algebraic connectivity | 1.13 | 0.68 | 0.31 | 0.21 | 0.16 |

Abbildung 4.5: zeitliche Entwicklung der Metriken im niederländischen Netzwerk aus (3)

Kapitel 5

Netzwerkanalysen in den Bundesligen

Die gesammelten Daten aus den ersten drei deutschen Bundesligen werden nun aus der Datenbank in verschiedene Netzwerke gebracht und anschließend analysiert. Als unterstützende Bibliotheken und Programme für die Analysen werden JUNG¹, UCINET² und PAJEK³ verwendet.

5.1 Spielernetzwerk

5.1.1 Normales Spielernetzwerk

Zuerst werden die 1461 Spieler aus der Datenbank jeweils als Knoten in ein Netzwerk gebracht und verbunden, wenn sie jemals bei einem der 56 Profivereine der Saison 2009/2010 zusammengespelt haben. Insgesamt entstehen so 34538 Links. Einen Überblick über die verschiedenen Metriken findet man in Tabelle 5.1.

Der Graph ist zusammenhängend, da man von jedem Spieler alle anderen Spieler erreichen kann. Dies geschieht, bei Benutzung des schnellsten Pfades, in maximal vier Schritten, was wiederum das in Kapitel 4.2 erläuterte Kleine-Welt-Phänomen verdeutlicht.

Bei der Gradverteilung in Abbildung 5.1 findet man zwischen dem Grad 24 und 32 die drei größten Maxima. In diesem Intervall befinden sich die Spieler, die nur bei einem oder maximal zwei Vereinen gespielt haben. Genau 57 Spieler besitzen 27 Nachbarn und damit stellt dieser Punkt das globale Ma-

¹<http://jung.sourceforge.net/>

²<http://www.analytictech.com/ucinet/>

³<http://vlado.fmf.uni-lj.si/pub/networks/pajek/>

| Metrik | Wert |
|------------------------------------|--------------------------------|
| Anzahl Knoten | 1461 |
| Anzahl Links | 34538 |
| Durchschnittsgrad | 47,3 |
| höchster Grad | 123 (Albert Streit) |
| Diameter | 4 |
| größte durchschn. Distanz | 3,242 (15 Sp. von Heidenheim) |
| kleinste durchschn. Distanz | 2,064 (Alexander Voigt) |
| größte Closeness | 0,484 |
| kleinste Closeness | 0,308 |
| durchschnittliche Distanz | 2,50 |
| Clusterkoeffizient | 0,63 |
| Assortivity coefficient | 0,21 |
| maximaler durchschn. Grad (Verein) | 64,769 (FC Augsburg) |
| minimaler durchschn. Grad (Verein) | 26,294 (Werder Bremen II) |
| größte Betweenness | 0,0088 (Christian Mikolajczak) |

Tabelle 5.1: Metriken im Spielernetzwerk

ximum dar. Ab diesem Punkt geht der Graph stets bergab, wobei zwischen Grad 42 und 65 die Schwingungen weitgehend konstant sind. Ab Grad 85 sind bei den einzelnen Gradwerten meist nur einzelne Spieler, die sehr wechselwillig waren und auch schon etwas älter sind, zu finden. Dadurch, dass er schon bei fünf verschiedenen Bundesligisten gespielt hat, besitzt Albert Streit mit 123 den höchsten Grad.

Wenn man für jeden Profiverein den Durchschnittsgrad seiner aktuellen Spieler berechnet, sieht man, dass tendenziell die Zweitligisten den höchsten Durchschnittsgrad haben und die Drittligisten den geringsten. Vor allem die vier zweiten Mannschaften der Erstligavereine liegen sehr weit unten, was aber den Grund hat, dass deren Spieler mittlerweile fast alle unter 23 Jahre alt sein müssen und so normalerweise wenig Erfahrungen bei anderen Profivereinen haben. Ganz vorne liegt der FC Augsburg mit 64,8, was für eine recht hohe Zahl an Spielern spricht, die in mehreren Vereinen gespielt haben. Auch der 1. FC Kaiserslautern liegt mit einem Durchschnittsgrad von 56,6 recht weit vorne auf Platz fünf. Erklären lässt sich dies dadurch, dass man dort in letzter Zeit wieder vermehrt auf Spieler setzt, die schon in Deutschland gespielt haben, um die längere Eingewöhnungszeit eines aus dem Ausland kommenden Spielers zu vermeiden und die Verständigung in der Mannschaft aufgrund vorhandener Deutschkenntnisse zu verbessern.

Die kleinste durchschnittliche Distanz zu den anderen Spielern hat Alex-

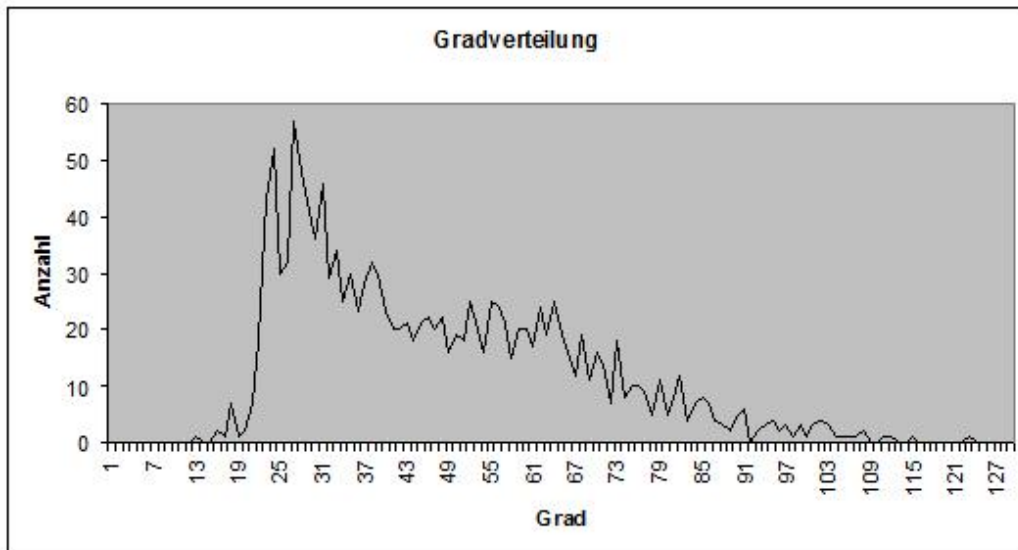


Abbildung 5.1: Gradverteilung im Spielernetzwerk

ander Voigt mit 2,06 Schritten (Closeness von 0,48), der in der Saison 2009/2010 beim FSV Frankfurt spielte (Tabelle 5.2).

| Spieler | Verein | Closeness | d. Distanz |
|-----------------|---------------|-----------|------------|
| Alexander Voigt | FSV Frankfurt | 0,484 | 2,064 |
| Christian Rahn | Fürth | 0,479 | 2,086 |
| Florian Heller | Mainz | 0,475 | 2,107 |
| Tom Starke | Duisburg | 0,474 | 2,111 |
| Tom Geissler | Koblenz | 0,474 | 2,112 |

Tabelle 5.2: Closeness im Spielernetzwerk

Das Spielen bei mehreren Vereinen in der ersten als auch in der zweiten Bundesliga kann die Ursache der kurzen Entfernungen zu allen anderen Knoten sein. Insgesamt 15 Spieler vom 1. FC Heidenheim aus der dritten Bundesliga verzeichnen demgegenüber mit durchschnittlichen 3,42 Schritten (Closeness von 0,31) die längste Distanz. Dieser Verein und ein Großteil seiner Spieler spielte in der Saison 2009/2010 im ersten Jahr in der dritten Liga und ist somit im Vergleich zu anderen Mannschaften schlechter vernetzt. Durchschnittlich beträgt die kürzeste Distanz zwischen zwei Spielern 2,50, was sehr gering ist und wiederum das Kleine-Welt-Phänomen zeigt. Darüber hinaus steht hierfür auch ein hoher globaler Clusterkoeffizient von 0,634, was auch daran liegt, dass Spieler, die recht kurz bei maximal einem Profiverein

gespielt haben, einen lokalen Clusterkoeffizienten von 1,0 haben.

Trotz der höchsten Closeness Zentralität bei Alexander Voigt hat dieser Spieler nicht die höchste Betweenness (Tabelle 5.3). Christian Mikolajczak von Dynamo Dresden hat den Höchstwert mit 0,00884. Dieser Spieler hat schon bei Vereinen in allen Profiligen gespielt und liegt deshalb auf den meisten kürzesten Pfaden zwischen zwei Spielern.

| Spieler | Verein | Betweenness |
|-----------------------|-------------------|--------------------|
| Christian Mikolajczak | Dynamo Dresden | 0,0088 |
| Danny Schwarz | Bayern München II | 0,0086 |
| Albert Streit | Schalke 04 | 0,0082 |
| Tobias Rathgeb | VFB Stuttgart II | 0,0077 |
| Benjamin Schüßler | Holstein Kiel | 0,0076 |

Tabelle 5.3: Betweenness im Spielernetzwerk

Ähnlich wie das brasilianische Netzwerk in Abschnitt 4 hat das deutsche Spielernetzwerk einen positiven assortivity coefficient (0,21). Dieses Ergebnis stimmt mit der Analyse in (6) überein, bei der alle sozialen Netzwerke, im Gegensatz zu den technischen und biologischen, einen positiven Koeffizienten hatten.

Bei der Analyse der Cliques ist festzustellen, dass die vier größten Cliques exakt aus den vier Mannschaften mit den größten Kadern bestehen. Trainer und Manager Felix Magath beendete die Saison bei Schalke 04 mit 36 Spielern, die die größte Clique bilden. Danach folgen TuS Koblenz (33), Hannover 96 (32) und die erste Mannschaft von Werder Bremen (31). Es existieren sieben verschiedene Cliques mit jeweils 30 Mitgliedern, von denen zwei den 1. FC Kaiserslautern betreffen. Während eine Clique aus der Rückrundemannschaft der Saison 2009/2010 gebildet wird, umfasst die andere die Hinrundemannschaft. 27 Spieler befinden sich in beiden Mengen, während sich die Winterzugänge (De Wit, Steinhöfer, Knaller) und Winterabgänge (Dzaka, Pinheiro, Hesse) auf unterschiedliche Mengen verteilen.

5.1.2 Modifiziertes Spielernetzwerk

Im Folgenden wird das Spielernetzwerk leicht verändert, um eine detailliertere Gruppenbildung festzustellen. Zwei Spieler werden nur dann verbunden, wenn sie in mindestens zwei Vereinen zusammengespielt haben. Dies hat zur Folge, dass 1007 der 1461 betrachteten Spieler einen Grad von 0 haben (Tabelle 5.4).

| Grad | Anzahl |
|------|--------|
| 0 | 1007 |
| 1 | 207 |
| 2 | 118 |
| 3 | 67 |
| 4 | 28 |
| 5 | 11 |
| 6 | 9 |
| 7 | 5 |
| 8 | 3 |
| 9 | 3 |
| 10 | 2 |
| 11 | 1 |

Tabelle 5.4: Gradverteilung im modifizierten Spielernetzwerk

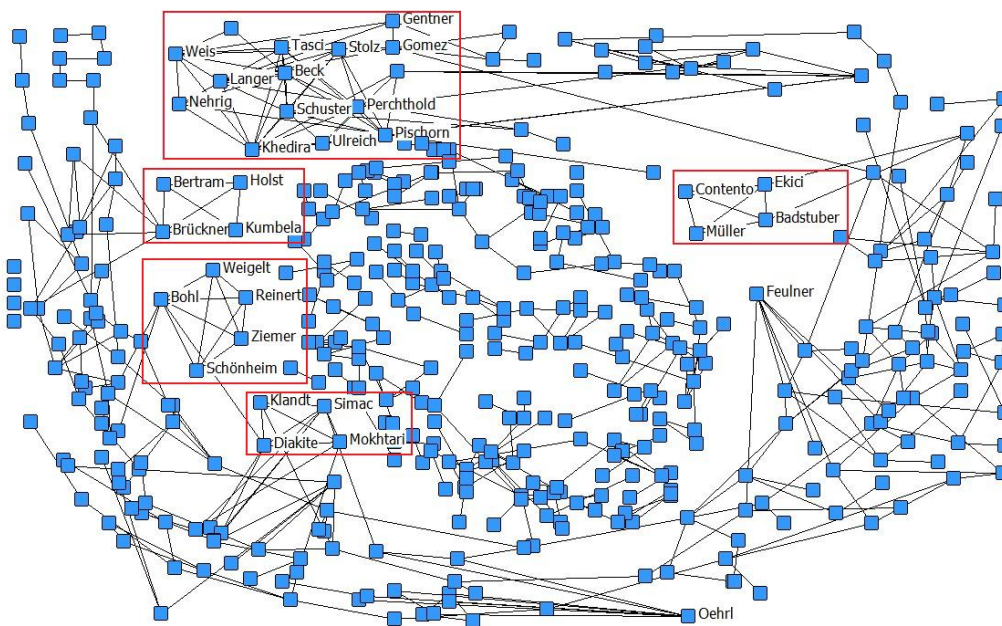


Abbildung 5.2: Modifiziertes Spielernetzwerk

Wie zu erwarten war, haben die Spieler den größten Grad, die zusammen bei den Amateuren und anschließend bei den Profis eines Vereins gespielt haben. In der Saison 2009/2010 haben die Amateure aus Dortmund, München, Stuttgart und Bremen in der dritten Liga gespielt und sind deshalb

aufgeführt. Interessant ist, dass die sechs Spieler mit den Gradwerten 9 bis 11 alle beim VFB Stuttgart I und beim VFB Stuttgart II gespielt haben. Daraus könnte man schließen, dass der VFB Stuttgart sehr viel auf seine Jugendspieler setzt, die offensichtlich den Weg von der zweiten Mannschaft in die Profimannschaft geschafft haben und immer noch in einer der drei Profiligen spielen. Sami Khedira hat mit 11 die meisten Nachbarn und zählt zusammen mit seinen Nationalmannschaftskollegen Andreas Beck (10) und Serdar Tasci (9) zu den am besten vernetzten Spielern in diesem Netzwerk.

In Abbildung 5.2 findet man das modifizierte Spielernetzwerk, das aus den 454 Spielern besteht, die mindestens einen Nachbarn haben. Die oben beschriebene gute Vernetzung beim VFB Stuttgart ist im oberen Teil der Abbildung zu erkennen. Die größte Clique besteht aus den sieben Spielern Khedira, Tasci, Beck, Stolz, Pischorn, Schuster und Perchthold. Weitere Gruppen, bei denen alle Spieler direkt verbunden sind, stammen aus dem markierten Stuttgarter Bereich im oberen Teil des Bildes. Weiter rechts im Bild befinden sich die Spieler des FC Bayern München, die vorher bei der zweiten Mannschaft dieses Vereins gespielt haben, darunter die beiden Senkrechtstarter der Saison 2009/2010 Thomas Müller und Holger Badstuber. Somit verdeutlicht diese Grafik, dass der FC Bayern in dieser Saison vermehrt auf die eigene Jugend gesetzt hat. Eine weitere Clique bilden die ehemaligen Talente des 1. FC Kaiserslautern (Bohl, Reinert, Schönheim, Ziemer), die nicht den großen Sprung in die 1. Bundesliga geschafft haben und gemeinsam mit Benjamin Weigelt mittlerweile beim SV Wehen Wiesbaden ihr Geld verdienen. Dieser Verein hat auch mit der markierten Gruppe mit den Spielern Klandt, Simac, Mokhtari und Diakite zu tun, da diese Spieler ihn alle verlassen haben und jetzt beim FSV Frankfurt spielen. Diese sehr häufigen regionalen Wechsel werden später noch weiter beleuchtet. Die Profis Holst, Bertram, Brückner und Kumbela haben alle eine gemeinsame Zeit bei Rot-Weiß Erfurt und dem SC Paderborn.

5.2 Vereinsnetzwerke

Im Folgenden werden zwei verschiedene Vereinsnetzwerke vorgestellt.

5.2.1 Flow-Netzwerk

In diesem Netzwerk werden die 56 Vereine als Knoten untergebracht. Eine gerichtete und gewichtete Kante zeigt die Anzahl der Vereinswechsel aller 1461 geführten Spieler von einem Verein zu einem anderen Verein an. Es ist zu beachten, dass auch das Zurückkommen eines Leihspielers als Transfer

gezählt wird, sofern der Spieler in der darauffolgenden Saison wieder bei seinem alten Verein spielt.

In Abbildung 5.3 sind alle Kanten aufgeführt, bei denen das Gewicht mindestens drei ist. Alle Vereine, die nun keinen Nachbarn mehr haben, und die Kantengewichte werden aus Übersichtsgründen vernachlässigt. Man erkennt, dass nur wenige Vereine mehr als drei Nachbarn in diesem Netzwerk haben, was zu der These führt, dass manche Vereine gerne von bestimmten anderen Clubs Spieler kaufen. Auffällig ist, dass dies häufig Vereine sind, die aus nahegelegenen Städten und Regionen kommen. Neben den markierten Verbindungen der ersten und zweiten Mannschaften eines Clubs sind dies beispielsweise die bayerischen Vereine aus Ingolstadt, Regensburg und Burghausen oder der FSV Frankfurt und Wehen Wiesbaden. Auch die SpVgg Unterhaching kauft seine Spieler vermehrt von zwei Münchener Vereinen. Darüber hinaus verdeutlichen die Verbindungen Mainz-Koblenz und Paderborn-Bielefeld diese These.

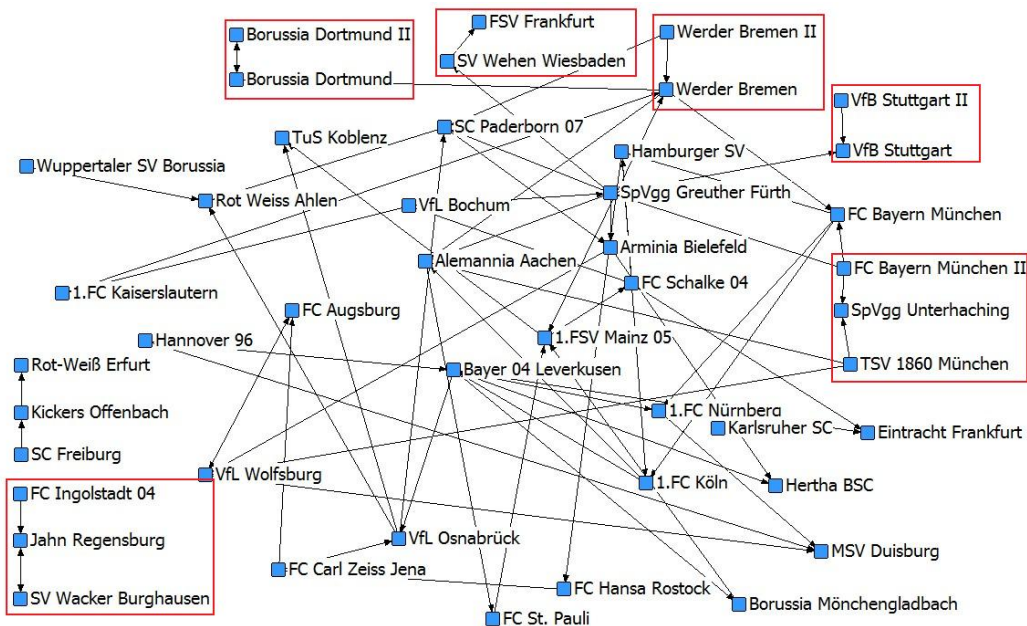


Abbildung 5.3: Vereinsnnetzwerk

Man findet in Abbildung 5.4 dasselbe Netzwerk, jedoch mit angezeigten Kantengewichten, die größer als drei sind. Wie in Abschnitt 5.1 werden die vielen Wechsel von der zweiten in die erste Mannschaft des VfB Stuttgart mit einem Wert von 24 gezeigt. Im Gegensatz dazu gibt es bei Borussia Dortmund nur 10 Spieler, die den Sprung in das Erstligateam geschafft haben, wobei vier den umgekehrten Weg gemacht haben. Das größte Kantengewicht eines

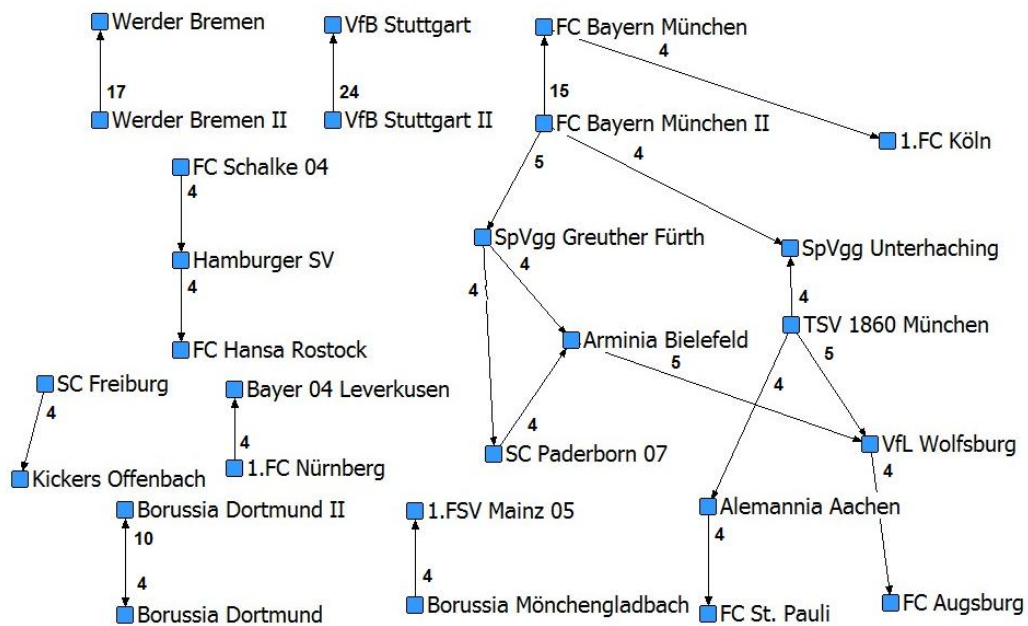


Abbildung 5.4: Vereinsnnetzwerk mit Kantengewichten

Vereinswechsels liegt bei 5 und kommt dreimal vor, wobei der VfL Wolfsburg zweimal das einkaufende Team ist. Dieser Club weist eine Vorliebe für Spieler der Vereine TSV 1860 München und Arminia Bielefeld auf.

Bei diesem Netzwerk bietet es sich an, den PageRank-Algorithmus anzuwenden. Hierbei wird die Wichtigkeit eines Knotens anhand seiner Verlinkungen bestimmt. Somit kann man die Knoten anhand ihrer Gewichte sortieren.(8) Dieser Algorithmus wurde leicht umgeschrieben, um gewichtete Kanten zu berücksichtigen. Der VfL Osnabrück (1,70), Greuther Fürth (1,60) und der 1.FSV Mainz 05 (1,51) liegen in der Rangliste vorne. Diese haben somit die höchsten Transferaktivitäten der Vereine aus den deutschen Bundesligen. Sehr weit unten liegen, wie zu erwarten war, die vier zweiten Mannschaften. Den letzten Platz belegt der am schlechtesten vernetzte Verein 1. FC Heidenheim (0,32), der erst ein Jahr in der dritten Bundesliga ist.

5.2.2 Ego-Netzwerk

In diesem gerichteten und gewichteten Netzwerk werden Anhaltspunkte dafür aufgezeigt, wie gut ein bestimmter Verein einen anderen Verein kennt und umgekehrt. Durch eine gemeinsame Zeit bei einem Verein sind einem Fußballer normalerweise die Stärken und Schwächen eines anderen Spielers

bekannt. Der jeweilige Trainer kann sich so Informationen über die Spieler eines kommenden Gegners erhoffen, um eine möglichst optimale Vorbereitung auf das Aufeinandertreffen zu gewährleisten.

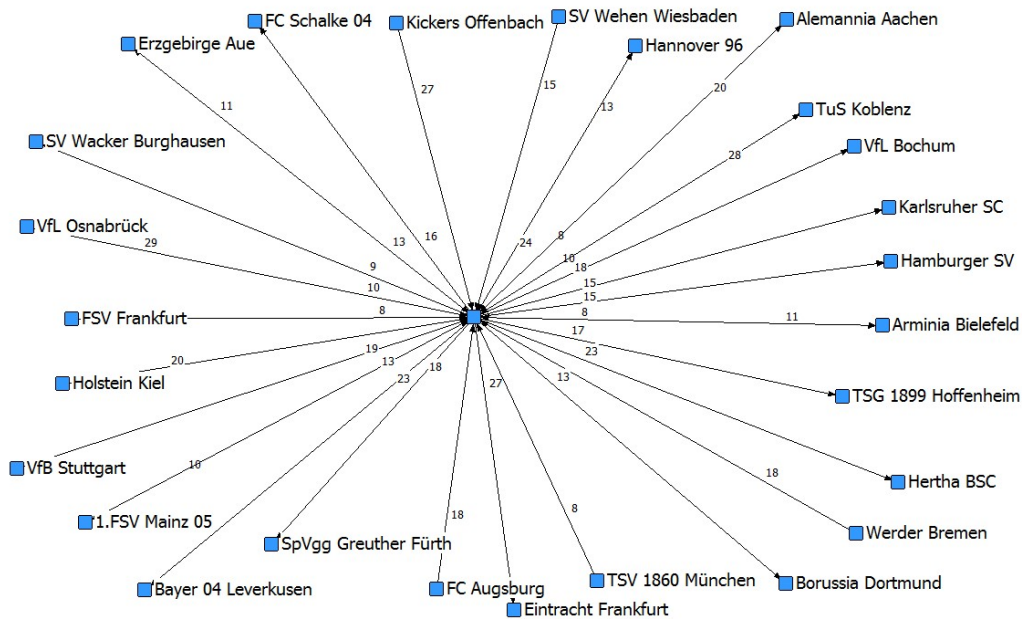


Abbildung 5.5: Ego-Netzwerk des 1. FC Kaiserslautern

Ein Ego-Netzwerk besteht aus einem zentralen Knoten (ego) und anderen Akteuren, die mit diesem verbunden sind.⁽⁹⁾ In Abbildung 5.5 wird ein Ego-Netzwerk des 1. FC Kaiserslautern dargestellt, wobei nur die Kanten angezeigt werden, bei denen dieser Verein beteiligt ist. Aus Übersichtsgründen sind nur die Vereine dargestellt, bei denen die größten Kantenwerte auftreten. Deshalb ist aus dem Graph auch nicht ersichtlich, dass der 1. FC Heidenheim der einzige betrachtete Verein ist, mit dem der 1. FC Kaiserslautern keine einzige Spielerbekanntschaft hat.

Interessanterweise kennen die Spieler des VfL Osnabrück die meisten Spieler des 1. FC Kaiserslautern, der der Gegner in der ersten Runde im DFB-Pokal der Saison 2010/2011 ist. Dies sollte aber den 1. FC Kaiserslautern nicht allzusehr beunruhigen, da dessen Spieler Ricky Pinheiro nur in der Rückrunde der Saison 2009/2010 nach Osnabrück ausgeliehen war, jedoch mittlerweile wieder zurückgekehrt ist. Dennoch wird hier eine mögliche praktische Verwendung von Netzwerken im Fußball aufgezeigt.

5.3 Netzwerk Spieler-Berater

Dieses bipartite Netzwerk besteht aus den disjunkten Mengen Spieler und Berater. Eine Verbindung besteht genau dann, wenn eine Beraterfirma aktuell einen Spieler berät. Bei 1461 Spielern ergeben sich nur 1298 Links, woraus sich schließen lässt, dass 163 Spieler entweder keinen Berater haben oder aus diversen Gründen nicht in der Datenquelle aufgeführt wurden. Beispielsweise wurden die Beziehungen bei der Datensammlung nicht berücksichtigt, bei denen ein Spieler von einem Familienangehörigen beraten wurde, damit solche Sonderfälle die Analyse nicht verfälschen. Die Dichte ist mit 0,0033 sehr gering.

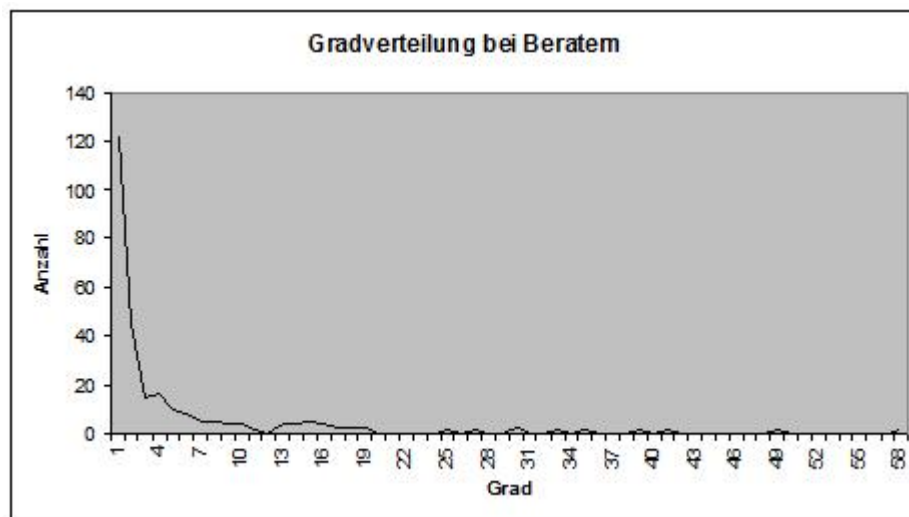


Abbildung 5.6: Gradverteilung der Berater

Eine Beraterfirma berät im Durchschnitt lediglich 4,86 Spieler, was aber vor allem an den vielen Einzelpersonen liegt, die nur für einen oder zwei Spieler zuständig sind (Abbildung 5.6).

In Tabelle 5.5 sind die zehn Beraterfirmen aufgeführt, die die meisten Spieler der drei deutschen Profiligen beraten. Die Stars and Friends GmbH hat mit 58 Spielern in der Saison 2009/2010 mehr Spieler als alle anderen Firmen beraten. Danach folgen SportsTotal und Jörg Neubauer.

Sehr interessant ist die Frage, inwieweit sich Spieler, die von derselben Firma beraten werden, bei bestimmten Vereinen häufen. Die Rogon AG hatte um das Jahr 2002 zwölf Spieler des 1. FC Kaiserslautern beraten, worauf manche den Absturz dieses Vereins in den folgenden Jahren zurückführten. Dazu schreibt „der Westen“ auf seiner Internetseite: Der Satz „Man spiele nicht, wenn man den falschen Berater habe.“ galt auf dem Betze angeblich als

| Berater | Anzahl |
|--------------------------------------|---------------|
| Stars & Friends GmbH | 58 |
| SportsTotal | 49 |
| Jörg Neubauer | 41 |
| Pro Profil GmbH | 39 |
| Karl M. Herzog Sportmanagement | 35 |
| Gerster/Koutsoliakos | 33 |
| IMG - International Management Group | 30 |
| Soccertalk Sportmanagement | 30 |
| ROGON Sportmanagement GmbH & Co. KG | 27 |
| Steffen Schneekloth | 25 |

Tabelle 5.5: Anzahl der Spieler von Beratern

geflügeltes Wort.(10) Auch andere Vereine, wie der 1. FC Saarbrücken, hatten mit der Häufung von Spielern bei Rogon Schwierigkeiten (5). Mittlerweile spielt kein einziger Spieler, der von Rogon beraten wird, in Kaiserslautern. Die Stars and Friends GmbH berät mit 5 Spielern die meisten, darauf folgen SportsTotal und Dr. Michael Becker mit jeweils drei Spielern (Tabelle 5.6).

| Berater | Anzahl |
|----------------------|---------------|
| Stars & Friends GmbH | 5 |
| SportsTotal | 3 |
| Dr. Michael Becker | 3 |
| Kögl & Partner GmbH | 2 |
| act SPORTS GROUP | 2 |

Tabelle 5.6: Anzahl der Spieler von Beratern beim 1. FC Kaiserslautern

Insgesamt liegt der 1. FC Kaiserslautern immer noch recht weit oben, wenn man die Häufung von Spielern einer Beraterfirma in einem Verein betrachtet. Tabelle 5.7 zeigt, dass der FSV Frankfurt acht Spieler beschäftigt, deren Spielerberater Gerster/Koutsoliakos sind. Energie Cottbus und Holstein Kiel folgen mit einer Anzahl von sieben Spielern, was ebenfalls recht hoch ist. Sehr auffällig ist auch, dass die vier Beraterfirmen, welche die meisten Spieler in einem einzelnen Verein betreuen, bei keinem anderen Verein mehr als drei Spieler vertreten. Dies deutet darauf hin, dass bestimmte Vereine bei Transfers Spieler mit bestimmten Beraterfirmen bevorzugen. Bei Schalke 04 gab es Bedenken wegen der Vielzahl der Spieler, die von Rogon

beraten wurden, da man die früheren Probleme in Kaiserslautern vor Augen hatte. Trotz der Aussage von Uli Hoeness „Wenn ich Schalke einen guten Rat geben darf: Der Verein sollte diese Zusammenarbeit reduzieren.“(10) im Jahr 2008, spielen immer noch sechs Spieler von Rogon in Gelsenkirchen.

| Verein | Anzahl | Berater |
|----------------------|--------|--------------------------------|
| FSV Frankfurt | 8 | Gerster/Koutsoliakos |
| Energie Cottbus | 7 | Karl M. Herzog Sportmanagement |
| Holstein Kiel | 7 | Steffen Schneekloth |
| FC Schalke 04 | 6 | ROGON Sportmanagement |
| Alemannia Aachen | 5 | Stars & Friends GmbH |
| 1. FC Kaiserslautern | 5 | Stars & Friends GmbH |
| Hannover 96 | 5 | Stars & Friends GmbH |

Tabelle 5.7: Maximale Spieleranzahl einer Beraterfirma in den Vereinen

5.4 Netzwerk Verein-Berater

Das bipartite Netzwerk Verein-Berater besteht aus 969 Links zwischen den Mengen Verein und Berater und hat die Dichte 0,065. Im Durchschnitt hat ein Verein 17,3 verschiedene Berater, ein Berater berät im Mittel Spieler von 3,6 verschiedenen Vereinen. Die Stars and Friends GmbH berät Spieler von 34 verschiedenen Vereinen und ist somit nicht nur für die meisten Spieler, sondern auch für die meisten Vereine zuständig. Hertha BSC Berlin ist mit 22 verschiedenen Beratern verknüpft und hat damit die meisten Links. Dies spricht dafür, dass dieser Verein den Einkauf eines Spielers nicht von der Beraterfirma abhängig macht. Weiterhin hat die Stars and Friends GmbH mit 2,35 die kürzeste durchschnittliche Distanz (und damit auch die höchste Closeness-Zentralität) zu den anderen Knoten. Anschließend folgen die Pro Profil GmbH (2,49) und Sports Total (2,52). Der Verein FC Augsburg besitzt mit 2,69 die kürzeste durchschnittliche Distanz bei den Vereinen, gefolgt vom FC Schalke 04 (2,70) und Hansa Rostock (2,71).

Im Folgenden werden zum ersten Mal in dieser Arbeit die verschiedenen Definitionen für den Clusterkoeffizienten in bipartiten Netzwerken benutzt (Tabelle 5.8).

Der Standard-Clusterkoeffizient liegt bei 0,19, wobei der Wert der Berater mit 0,21 größer ist als der Vereinskoeffizient mit 0,1. Es besteht insgesamt ein großer Unterschied zwischen den verschiedenen Definitionen, da man ein Ergebnis von 0,77 bei dem Beraterkoeffizienten erhält, wenn man im Nenner nur

| Clusterkoeffizienten | Wert |
|----------------------|------|
| Vereine | 0,1 |
| Berater | 0,21 |
| Gesamt | 0,19 |
| Vereine (Min) | 0,19 |
| Berater (Min) | 0,77 |
| Gesamt (Min) | 0,67 |
| Vereine (Max) | 0,16 |
| Berater (Max) | 0,23 |
| Gesamt (Max) | 0,22 |

Tabelle 5.8: Clusterkoeffizienten im Netzwerk Verein-Berater

das Minimum der Anzahl der Nachbarn verwendet. Daraus könnte man folgern, dass viele Berater im Wesentlichen bei denselben Vereinen ihrer Arbeit nachgehen, wobei manche Berater mit wenigen, andere mit vielen Vereinen in Kontakt stehen. Man muss jedoch bedenken, dass es sehr viele Beraterfirmen gibt, die nur sehr wenige Spieler in den Bundesligen vertreten. Beispielsweise ist Franck Ribéry der einzige Spieler in den deutschen Profiligen, der von dem französischen Berater Alain Migliaccio vertreten wird. Folglich wird der Clusterkoeffizient bei Paaren, bei denen dieser Berater vertreten ist, immer 1,0, da, wie bereits oben erwähnt, nur Paare in die Berechnung eingehen, bei denen der Schnitt nicht leer ist. Beim Vereinskoeffizienten gibt es keine so großen Abweichungen, da dieses Problem hier nicht auftritt, weil beim Verein mit dem niedrigsten Grad immer noch 9 verschiedene Berater zuständig sind. Insgesamt sieht man aber an diesem Beispiel die Problematik der Clusterkoeffizienten bei bipartiten Graphen.

Interessant an diesem Netzwerk ist auch, dass die Stars and Friends GmbH nicht nur die größte Degree- und Closeness-Zentralität hat, sondern auch die höchste Betweenness mit 0,11 besitzt. Damit ist diese Beraterfirma ganz klar der zentralste Akteur im Netzwerk. Wie schon bei der Closeness, kommen danach die Pro Profil GmbH (0,07) und Sports Total (0,06). Bei Betrachtung der Vereine erkennt man, dass Hertha BSC Berlin nicht nur den höchsten Grad, sondern auch die höchste Betweenness (0,03) hat. Bei der Closeness war der Verein jedoch nicht unter den ersten drei.

5.5 Netzwerk Spieler-Trainer

Bei diesem bipartiten Netzwerk gibt es 2715 Links zwischen den Mengen Spieler und Trainer, was zu einer Dichte von 0,033 führt. Alexander Baumjohann, der aktuell beim FC Schalke 04 spielt, erreicht mit sieben verschiedenen Trainern, die am Ende der Saison 2009/2010 eine Profimannschaft in Deutschland trainiert haben, das Maximum. Der Verein aus Gelsenkirchen beschäftigt darüber hinaus mit Felix Magath den Trainer, der die meisten aktuellen Spieler trainiert hat. Insgesamt kommt er auf eine Zahl von 118, was darauf zurückzuführen ist, dass er in den letzten Jahren etliche Vereine trainiert hat (z.B. Schalke, Wolfsburg, Bayern München). In Tabelle 5.9 kann man erkennen, welche vier Erstligatrainer die meisten Spieler der drei Profiligen trainiert haben. Dies deutet darauf hin, dass sich diese Trainer im deutschen Fußball etabliert haben und deshalb, auch wegen ihrer Erfahrung, bevorzugt von Erstligavereinen verpflichtet wurden. Der Spielerdurchschnittsgrad befindet sich bei 1,858, während der Trainerdurchschnittsgrad bei etwa 48,5 liegt.

| Name | Verein | Degree |
|------------------|-------------------|--------|
| Felix Magath | Schalke 04 | 118 |
| Friedhelm Funkel | Hertha BSC Berlin | 104 |
| Dieter Hecking | 1. FC Nürnberg | 86 |
| Ralf Rangnick | TSG Hoffenheim | 79 |
| Wolfgang Wolf | Kickers Offenbach | 78 |

Tabelle 5.9: maximale Degrees bei Trainern

Der Graph ist nicht zusammenhängend, was aber an der Homepage liegt, die ausgelesen wurde. Einige Spieler waren bei einer der 56 Mannschaften aufgeführt, obwohl sie offiziell für die jeweilige zweite Mannschaft gespielt und damit möglicherweise bei keinem der in diesem Netzwerk berücksichtigten Trainer trainiert haben. Nach dem Weglassen dieser Spieler ist das übrige Netzwerk zusammenhängend.

Weiterhin liegt der Diameter bei 8 und tritt immer zwischen zwei verschiedenen Spielern auf. Dies bedeutet, dass man den Umweg über vier Trainer machen muss, um zu einem bestimmten anderen Fußballer zu gelangen. Vor allem tritt dieser Wert bei Knotenpaaren auf, bei denen eine Mannschaft aus den oberen beiden Ligen und eine aus der dritten Liga kommt. Sehr viele Spieler brauchen zu Spielern des 1. FC Heidenheim 8 Schritte, da diese, wie schon im Spielernetzwerk gesehen, wenig vernetzt sind.

Felix Magath hat mit einer durchschnittlichen Distanz von 3,0 die kürzeste, was natürlich auch am höchsten Grad liegt. Auch die nachfolgenden Trainer Jürgen Klopp (3,1) und vor allem Friedhelm Funkel (3,2) sind schon recht lange im deutschen Fußball tätig. Die größte Closeness haben drei Spieler, die schon über 32 Jahre alt und schon lange im deutschen Fußball tätig sind. Christoph Dabrowski (VFL Bochum), Jacek Krzynowek (Hannover 96) und Diego Klimowicz (VFL Bochum) haben alle im Durchschnitt etwa 3,5 Schritte, um zu anderen Spielern zu gelangen.

| Name | Degree | Closeness | Betweenness |
|--------------------|--------|-----------|-------------|
| Felix Magath | 118 | 0,33 | 0,13 |
| Jürgen Klopp | 72 | 0,32 | 0,07 |
| Friedhelm Funkel | 104 | 0,31 | 0,09 |
| Hans-Jürgen Boysen | 68 | 0,30 | 0,07 |
| Dieter Hecking | 86 | 0,29 | 0,06 |

Tabelle 5.10: Zentralität bei den Trainern

Wie schon beim Verein-Berater-Netzwerk gibt es einen Akteur, der bei allen Zentralitätswerten vorne liegt (Tabelle 5.10). Felix Magath besitzt mit 0,13 auch die größte Betweenness. Vielleicht sind seine Erfolge in Wolfsburg und Schalke auch damit zu erklären, dass er schon mit sehr vielen Spielern zusammengearbeitet und viele Kontakte in den deutschen Ligen hat. Friedhelm Funkel (0,09), der ebenfalls schon sehr lange im Trainergeschäft ist, und der Trainer des FSV Frankfurt Hans-Jürgen Boysen (0,07) liegen dahinter. Letzterer befindet sich auch deshalb auf Platz 3, weil er mit dem Heidenheimer Spieler Christian Gmünder in der Saison 2001/2002 eine gemeinsame Zeit beim FC Sandhausen hatte und der 1. FC Heidenheim sehr schlecht vernetzt ist. Die schlechte Vernetzung des 1. FC Heidenheim führt auch dazu, dass Christian Gmünder, der als einziger Heidenheimer Spieler bei einem im Netzwerk befindlichen Trainer (außer dem aktuellen Trainer) trainiert hatte, den größten Betweenness-Wert (0,02) besitzt. Lukas Podolski und Bradley Carnell haben die zweit- bzw. drittgrößte Betweenness der Spieler.

Der Clusterkoeffizient beträgt bei den Spielern 0,47 und bei den Trainern nur 0,02, was insgesamt zu einem Wert von 0,45 führt. Die recht hohe Zahl bei den Spielern hat natürlich den Grund, dass sie zumindest in der letzten Saison alle bei demselben Trainer gespielt haben. Bei einem Spielerdurchschnittsgrad von 1,858 muss sich so ein großer Koeffizient ergeben. Im Gegensatz dazu ist es offensichtlich, dass verschiedene Trainer selten die gleichen Spieler trainiert haben.

5.6 Netzwerk Verein-Trainer

Hier werden nun die Trainer mit ihren bisher trainierten Vereinen in Beziehung gebracht.

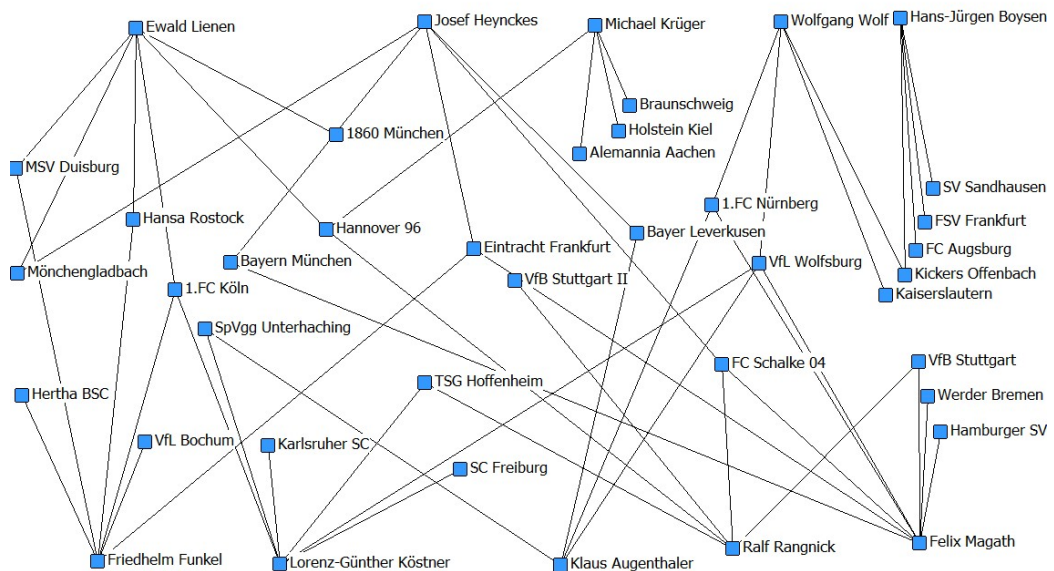


Abbildung 5.7: Netzwerk Verein-Trainer

Abbildung 5.7 zeigt einen Teil des bipartiten Netzwerks, bei dem die Trainer, die den größten Grad haben, am oberen und unteren Bildrand aufgeführt sind. Insgesamt entstehen 131 Links zwischen den 56 Trainern und 56 Vereinen und somit ergibt sich eine Dichte von 0,042. Borussia Mönchengladbach hat mit sechs Beziehungen zu den aktuellen Trainern den Höchstwert. Felix Magath hat, wie beim Spieler-Trainer-Netzwerk, den höchsten Grad der Trainer, da er schon acht verschiedene Profivereine in seiner Karriere trainiert hat. Er liegt damit weit über dem Trainerdurchschnittsgrad von 2,4. Der Vereinsdurchschnittsgrad ist nahezu gleich mit etwa 2,3.

Da in diesem Netzwerk nur die Trainer enthalten sind, die am Ende der Saison 2009/2010 Trainer waren, ist das Netzwerk nicht zusammenhängend. Es gibt einige Vereine und Trainer, die nur einen Nachbarn haben und so hat es wenig Sinn, z.B. den Clusterkoeffizienten zu analysieren.

Die kürzeste durchschnittliche Distanz zu den anderen Knoten hat Ewald Lienen (1860 München), der schon fünf verschiedene Profivereine trainiert hat, mit etwa 4,0 Schritten. Weitere Trainer folgen mit Felix Magath (4,2) und Jupp Heynckes (4,3). Dies ist darauf zurückzuführen, dass diese drei

Trainer schon bei vielen Vereinen in der Bundesliga trainiert haben und demzufolge auch besser vernetzt sind. Wie schon beim Grad hat Borussia Mönchengladbach die geringste durchschnittliche Distanz aller Vereine (4,1), gefolgt vom 1. FC Köln (4,3).

Der Verein Borussia Mönchengladbach hat mit 0,17 auch die größte Betweenness und sogar eine knapp höhere Zahl als Felix Magath (0,16) und Ewald Lienen (0,14). Insgesamt ist dieses Netzwerk jedoch zu klein, um signifikante Ableitungen zu ermöglichen.

Kapitel 6

Netzwerkanalysen im europäischen Fußball

Dieses Kapitel bietet einen Einstieg in eine Netzwerkanalyse des europäischen Fußballs. Aus sechs europäischen Ligen wurden die Vereine mit ihren aktuellen Spielern und zugehörigen Beratern vom Fußball-Portal ausgelesen. Neben der Bundesliga waren dies die englische Premier League, die spanische Primera División, die italienische Serie A, die französische Ligue 1, sowie die niederländische Eredivisie. Die zum Teil unvollständigen Informationen über die Karriere der Trainer führte zum Weglassen dieser Menge.

6.1 Spielernetzwerk

In diesem Netzwerk sind 2899 Spieler der genannten sechs Ligen aufgeführt. Verbindungen bestehen zwischen zwei Spielern, wenn diese eine gemeinsame Zeit bei einem Verein hatten. Tabelle 6.1 zeigt einige Metriken an.

Der schon 35-jährige Argentinier Hernán Crespo besitzt mit 125 den größten Grad. Dies dürfte auch daran liegen, dass bei ihm insgesamt zehn Vereinswechsel verzeichnet sind. Sowohl der Spieler mit der höchsten Closeness (Mohamed Sissoko) als auch der Fußballer mit der größten Betweenness (Jérémy Bréchet) haben schon in vier verschiedenen Ländern gespielt. Insgesamt kann man an Diameter, Durchschnittsgrad und durchschnittlicher Distanz erkennen, dass die Spieler etwas schlechter vernetzt sind als im deutschen Netzwerk.

| Metrik | Wert |
|------------------------------------|-------------------------------------|
| Anzahl Knoten | 2899 |
| Anzahl Links | 59987 |
| Durchschnittsgrad | 41,385 |
| höchster Grad | 125 (Hernán Crespo, FC Parma) |
| Diameter | 6 |
| kleinste durchschnittliche Distanz | 2,416 (Mohamed Sissoko, Juv. Turin) |
| größte Closeness | 0,414 |
| durchschnittliche Distanz | 3,04 |
| Clusterkoeffizient | 0,66 |
| Assortativity coefficient | 0,28 |
| größte Betweenness | 0,0115 (Jérémy Bréchet, Sochaux-M.) |

Tabelle 6.1: Metriken im europäischen Spielernetzwerk

6.2 Flow-Netzwerk

Bei diesem Flow-Netzwerk geht es darum, den Transferfluss zwischen den sechs Ländern darzustellen. Abbildung 6.1 zeigt die Anzahl der Transfers zwischen Deutschland und den übrigen Ligen an.

Die niederländische Ehrendivision wurde als sechste Liga ausgewählt, da man zur Zeit den Eindruck hat, dass sehr viele Spieler von der Ehrendivision in die Bundesliga wechseln. Die Zahl 24 ist der höchste Wert und bestätigt diese Vermutung. Nur die Häufung der Wechsel von Deutschland nach England (23) kommt an diesen Wert heran. Recht wenig Austausch besteht vor allem zwischen der Bundesliga und der Ligue 1 aus Frankreich. In Abbildung 6.2 wird der Transferfluss zwischen den fünf nicht-deutschen Ligen verdeutlicht.

Man erkennt sofort, dass sehr wenige Spieler sich dafür entscheiden, in die Niederlande zu gehen. Aus Deutschland (13) und England (9) kommen noch die meisten Spieler. Dies dürfte daran liegen, dass ein Wechsel in die Ehrendivision nicht so attraktiv ist, da das Niveau und das Gehalt in den anderen Ligen höher sind. Fußballer, die in den Niederlanden herausragen, gehen vor allem nach England, Deutschland und Spanien.

Weiterhin fällt der große gegenseitige Spieleraustausch von Spanien mit England und Italien auf, der größere Dimensionen als der Spieleraustausch zwischen diesen Ländern und Deutschland hat. Dieser Unterschied kann nicht allein darauf zurückgeführt werden, dass in der Bundesliga, anders als in Spanien, England und Italien, nur 18 und nicht 20 Mannschaften spielen. Da auch Häufigkeit der Spielertransfers innerhalb der deutschen ersten Bundesliga ge-

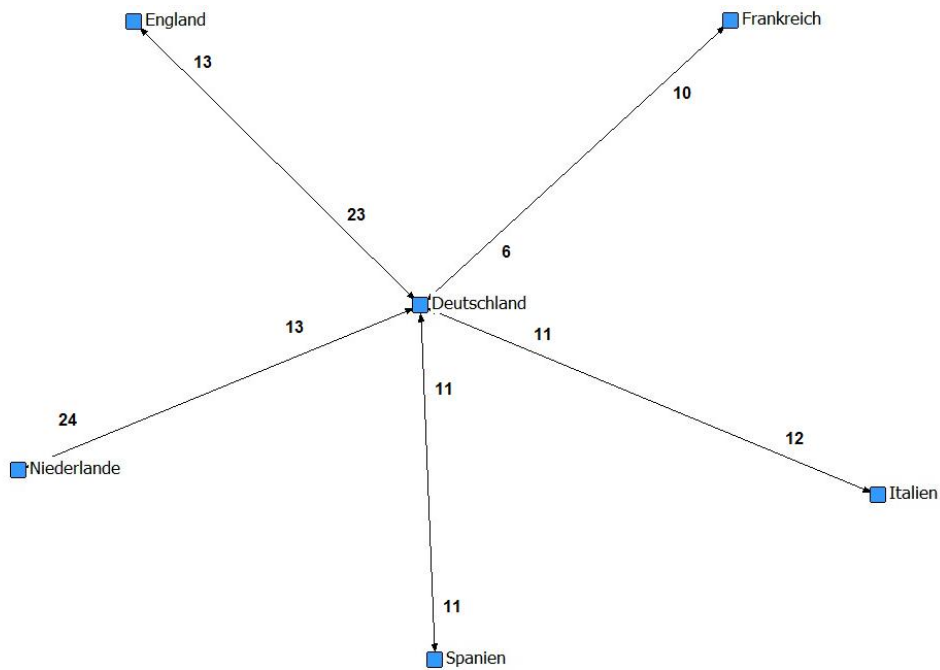


Abbildung 6.1: Ego-Netzwerk Deutschlands

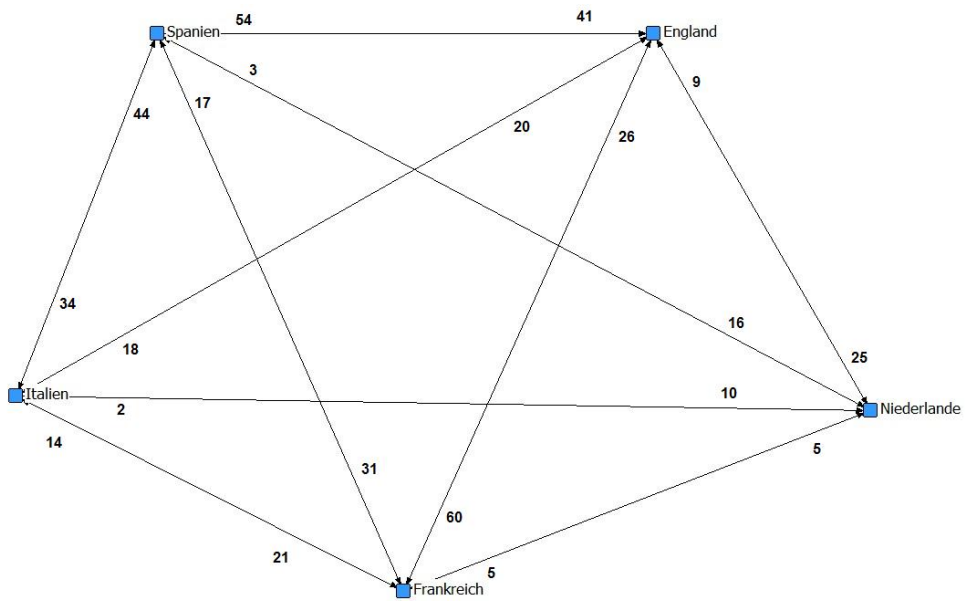


Abbildung 6.2: Flow-Netzwerk von Europa

ringer ist als in den ersten Ligen der in die Betrachtung einbezogenen anderen europäischen Länder, kommen vielmehr folgende Erklärungen in Betracht: Denkbar sind eine größere Vereinstreue der Spieler in Deutschland oder eine verstärkte Einbindung der deutschen Spieler aus der eigenen Jugend bzw. aus tieferen Spielklassen. Zudem könnte der Transferfluss zwischen der Bundesliga und nicht aufgeführten Ligen, zum Beispiel aus Osteuropa, größer sein. Dies könnten Ansatzpunkte für zukünftige Untersuchungen sein.

Von Frankreich nach England sind 60 Spielerwechsel vorhanden, was das höchste Kantengewicht ist. Somit geht ein Spieler aus der Ligue 1 viel häufiger in die Premier League als nach Deutschland oder Italien.

Abbildung 6.3 zeigt den Transferfluss der Vereine in Europa. Es sind nur Kanten gezeichnet, die das Gewicht 4 überschreiten. Da keiner der deutschen Vereine einen intensiven Austausch mit einem anderen europäischen Verein hatte und der deutsche Fluss in Abschnitt 5.2.1 schon beleuchtet wurde, sind die deutschen Vereine nicht noch einmal aufgeführt. Die neun Kantengewichte, die größer als fünf sind, werden angezeigt. Von Arsenal London sind 9 Wechsel nach Birmingham aufgeführt, was das Maximum darstellt. Darüber hinaus tauschen die beiden Rotterdamer Vereine ihre Spieler recht oft. Bei Beteiligung französischer Mannschaften gab es nur ein Vereinspaar, das mehr als vier Transfers in eine Richtung tätigte. Im Vergleich dazu fällt der doch sehr hohe Wechsel innerhalb der übrigen Ligen, insbesondere der Serie A, auf. Größere Transferaktivitäten zwischen Mannschaften aus verschiedenen Ländern kommen nur zweimal vor, wobei Real Madrid jeweils beteiligt ist. Zum einen kaufte dieser Verein in den letzten Jahren einige Spieler aus Liverpool, zum anderen wechselten recht viele Fußballer von Real Madrid zu Inter Mailand.

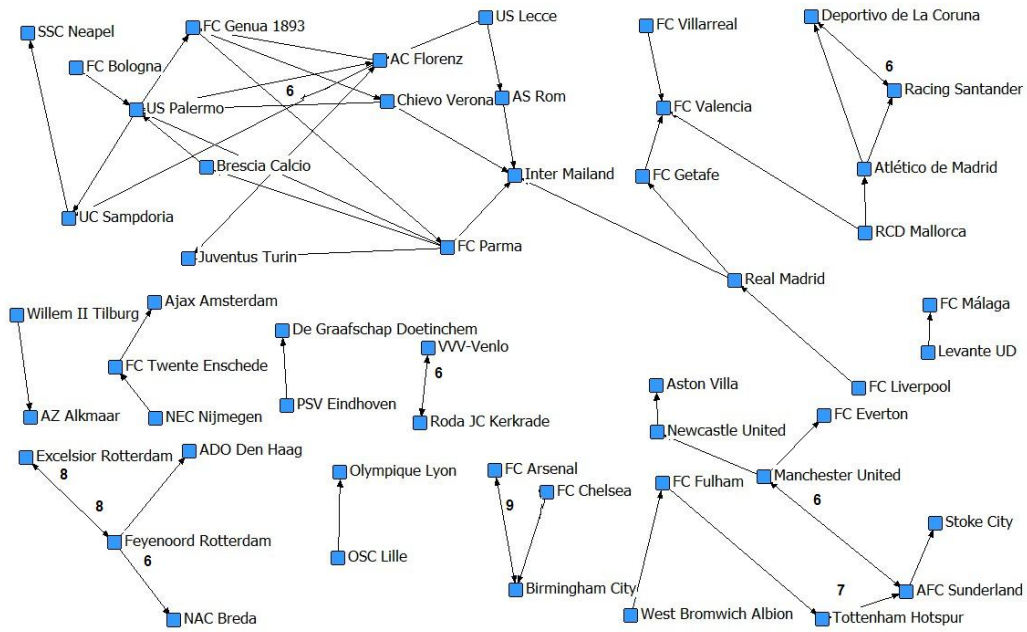


Abbildung 6.3: Flow-Netzwerk der europäischen Vereine

Kapitel 7

Fazit

7.1 Zusammenfassung

In dieser Arbeit wurden zahlreiche Aspekte einer Netzwerkanalyse im deutschen Profifußball aufgezeigt. Es ergaben sich einige interessante Erkenntnisse, wie beispielsweise die immer noch vorhandene Spezialisierung von Beratern auf bestimmte Vereine. Weiterhin wurde gezeigt, dass regionale Vereinswechsel verhältnismäßig oft vorkommen. Darüber hinaus wurden Unterschiede in der Einbindung der Spieler der zweiten Mannschaften bei den vier Vereinen festgestellt, bei denen die zweiten Mannschaften in der Saison 2009/2010 in der dritten Bundesliga gespielt haben. Im Bereich der Zentralität fielen häufig die gleichen Namen auf, die hier die Spitzenwerte erreichten. Felix Magath tauchte bei dieser Thematik bei den Trainern sehr häufig auf, während Albert Streit bei den Spielern der Auffälligste war. Insgesamt wurde verdeutlicht, dass man in dieser Art von Netzwerken recht schnell von einem Akteur zu einem beliebigen anderen gelangen kann, wie das Spielernetzwerk mit einem Durchmesser von 4 gezeigt hat.

7.2 Ausblick

Diese Studie spezialisierte sich auf die Spieler und Trainer, die am Ende der Saison 2009/2010 in einer der ersten drei deutschen Bundesligen gespielt haben. Zudem wurden nur Stationen der Spieler bei den 56 Vereinen dieser Spielzeit aufgeführt. Dieser Datensatz könnte stark erweitert werden. Zum einen könnte man auch tiefere Ligen mit in Betracht ziehen, zum anderen könnten alle Bundesliga-Spieler mit allen ihren bisherigen Stationen ab einer bestimmten Zeit erfasst werden. Vor allem bei der zuletzt genannten Erweiterung könnte man analysieren, inwieweit sich bestimmte Gegebenheiten im

deutschen Fußball verändert haben oder gleichgeblieben sind. Die Untersuchung der europäischen Vereine, die in Kapitel 6 begonnen wurde, könnte ausgeweitet werden, indem man alle Ligen in Europa berücksichtigt. Da in dieser Arbeit nur die aktuellen Berater berücksichtigt wurden, könnte man frühere Berater in die Untersuchungen einbeziehen. Problematisch wäre vor allem beim letzten Punkt das Auffinden einer Datenquelle, aus der man vollständige und korrekte Daten erhält. Bei den in dieser Arbeit ausgelesenen 1461 Spielern gab es schon zwei, bei denen die Transfergeschichte nicht vollständig war. Weiterhin waren im europäischen Netzwerk sehr unvollständige Informationen bei den Trainerstationen vorhanden. Solche Schwierigkeiten würden sich bei größeren Datensätzen und wohl auch bei weniger erfolgreichen ausländischen Ligen vergrößern. Da viele Tools schon große Probleme mit der Größe des Spielernetzwerks hatten, müsste man sich bei einer Vergrößerung der Datensätze auf noch mehr Komplikationen einstellen. Wenn man diese beiden Punkte in den Griff bekommt, könnten zukünftige Analysen weitere interessante Erkenntnisse hervorbringen.

Abbildungsverzeichnis

| | | |
|-----|--|----|
| 2.1 | Entity-Relationship-Diagramm | 3 |
| 4.1 | Gradverteilung im brasilianischen Netzwerk | 16 |
| 4.2 | Vergleich des brasilianischen Netzwerks mit Zufallsgraphen . . | 16 |
| 4.3 | zeitlicher Verlauf der Metriken im brasilianischen Netzwerk . . | 17 |
| 4.4 | Metriken im niederländischen Netzwerk | 18 |
| 4.5 | zeitliche Entwicklung der Metriken im niederländischen Netz- werk | 19 |
| 5.1 | Gradverteilung Spielernetzwerk | 23 |
| 5.2 | Modifiziertes Spielernetzwerk | 25 |
| 5.3 | Vereinsnetzwerk | 27 |
| 5.4 | Vereinsnetzwerk mit Kantengewichten | 28 |
| 5.5 | Ego-Netzwerk des 1. FC Kaiserslautern | 29 |
| 5.6 | Gradverteilung der Berater | 30 |
| 5.7 | Netzwerk Verein-Trainer | 36 |
| 6.1 | Ego-Netzwerk Deutschlands | 41 |
| 6.2 | Flow-Netzwerk von Europa | 41 |
| 6.3 | Flow-Netzwerk der europäischen Vereine | 43 |

Tabellenverzeichnis

| | | |
|------|---|----|
| 5.1 | Metriken im Spielernetzwerk | 22 |
| 5.2 | Closeness im Spielernetzwerk | 23 |
| 5.3 | Betweenness im Spielernetzwerk | 24 |
| 5.4 | Gradverteilung im modifizierten Spielernetzwerk | 25 |
| 5.5 | Anzahl der Spieler von Beratern | 31 |
| 5.6 | Anzahl der Spieler von Beratern beim 1. FC Kaiserslautern . . | 31 |
| 5.7 | Maximale Spieleranzahl einer Beraterfirma in den Vereinen . . | 32 |
| 5.8 | Clusterkoeffizienten im Netzwerk Verein-Berater | 33 |
| 5.9 | maximale Degrees bei Trainern | 34 |
| 5.10 | Zentralität bei den Trainern | 35 |
| 6.1 | Metriken im europäischen Spielernetzwerk | 40 |

Literaturverzeichnis

- [1] P. P.-S. Chen. The entity-relationship model—toward a unified view of data. *ACM Trans. Database Syst.*, 1(1):9–36, 1976.
- [2] D. Jansen. *Einführung in die Netzwerkanalyse: Grundlagen, Methoden, Forschungsbeispiele*. Cambridge University Press, dritte edition, 2006.
- [3] R. Kooij, A. Jamakovic, F. van Kesteren, T. de Koning, I. Theisler, and P. Veldhoven. The dutch soccer team as a social network. *Netherlands Organization for Scientific Research (NWO)*, 2009.
- [4] M. Latapy, C. Magnien, and N. D. Vecchio. Basic notions for the analysis of large two-mode networks. *Social Networks 30, ScienceDirect*, 2004.
- [5] U. Muras. Freundliche Übernahme des 1. fc kaiserslautern. *Welt Online*, 2002.
- [6] M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89(20):208701, Oct 2002.
- [7] R. N. Onody and P. A. de Castro. Complex network study of brazilian soccer players. *Phys. Rev. E*, 70(3):037103, Sep 2004.
- [8] L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the web. Technical report, Stanford University, 1998.
- [9] S. Wasserman and K. Faust. *Social Network Analysis*. Cambridge University Press, sechzehnte edition, 2007.
- [10] K. Wille. "die drahtzieher im fußball": 75000 euro für eine handynummer. *Der Westen*, 2008.

