

Positionssensitive Dekomposition von Potenzgesetzen am Beispiel von Wikipedia-basierten Kollaborationsnetzwerken¹

Christian Stegbauer, Alexander Mehler

Institut für Gesellschafts- und Politikforschung, Fachbereich 3
Institut für Informatik, Fachbereich 8
Goethe Universität
60054 Frankfurt
stegbauer@soz.uni-frankfurt.de
a.mehler@em.uni-frankfurt.de

Abstract: Potenzgesetzliche Verteilungen spielen bei der Modellierung von Strukturen im WWW eine bedeutende Rolle. Im Kontext der *Small-World*-Debatte von Watts & Strogatz sowie von Barabási & Albert dominieren Modelle, die das Konzept der kleinen Welt zusammen mit Überlegungen zur Pareto-Verteilung von Degree-Werten zu einer Art von universalistischem Prinzip erhoben. Die bisherige Erklärung für das Zustandekommen solcher beobachteter Verteilungen greift jedoch zu kurz. *Preferential Attachment* wird in unserer Arbeit daher durch eine Art von Verteilungsmodell ersetzt, dessen Komponenten auf soziologisch begründete Positionen zurückgeführt werden können. Die dahinter stehende Idee ist, dass Arbeitsteilung in sozialen Netzwerken zu formell unterschiedlichen Positionen führt, mit denen die Entstehung empirischer Verteilungen und deren Abweichung vom Ideal eines Potenzgesetzes erklärt werden kann. Aus diesem Grund gehen wir so vor, dass wir Personen als Funktion ihrer Positionen in der Verteilung der Anzahl ihrer Kooperationen über die gesamte Wikipedia hinweg quasi einfärben. Dieses Modell, das auf die interdisziplinäre Zusammenarbeit von Soziologie und Informatik zurückgeht und folgerichtig qualitative und quantitative Ansätze integriert, lässt sich auf verwandte Verteilungen übertragen und bietet viele Ansatzpunkte für die Modellierung sozialer Netzwerke im WWW.

1 Einleitung und Begründung des Vorgehens

1.1 Disziplinenkultur und Überwindung der Gegensätze

Für die Entwicklung der Netzwerkanalyse war es wichtig, dass sich neben den dort bereits engagierten Sozialwissenschaftlern auch Mathematiker, Physiker und Informatiker diesem Wissenschaftsgebiet zuwendeten. Es wurden Modellierungen entwickelt, die Phänomene wie das von Milgram [Mi67] eingeführte Small-World-Experiment erklären konnten. Anspruch der Überlegungen von Watts & Strogatz, wie der beteiligten Physi-

¹ Danksagung: Wir danken Rüdiger Gleim für das Download und das datenbanktechnische Management der Daten.

ker/Mathematiker, ist es, nicht nur das eine Phänomen zu modellieren, sondern Erklärungen für viel weiterreichende Sachverhalte zu bieten. Ein Modell ist also für ziemlich viele Fragestellungen anwendbar, wobei dem Kontext, in dem sie stehen, offenbar keine große Bedeutung zukommt. In der Soziologie ist das meist anders. Dort gibt es tradiertermaßen [We909]² eine Scheu vor Großtheorien, die auf zwei Ursachen zurückzuführen ist: Zum einen betrifft dies die Idee von Merton [Me67], lediglich Erklärungen mittlerer Reichweite entwickeln zu wollen, weil es insbesondere auf die jeweiligen Kontexte anzukommen schien, die kulturell variabel sind (siehe auch die neuere Netzwerktheorie [Wi92] mit der dort proklamierten kulturellen Wende [Fu10]). Eine zweite Schwierigkeit, Modelle aus der Physik einfach zu übernehmen, liegt in dem die Soziologie mitbegründenden Diktum Durkheims [Du91], wonach „Soziales nur aus Sozialem“ zu erklären sei.

Kontext ist also aus einer soziologischen Perspektive etwas, was in den meisten Fällen notwendigerweise mitgedacht werden muss, um den fachlichen Ansprüchen zu genügen. Aus diesen für die Kultur der Disziplinen typischen Gegensätzen [St08] wollen wir hier eine Tugend machen. Wir verbinden eine streng formalistische Vorgehensweise mit einer, die sich in einem Kontext an Inhalten bewährt hat. Verbunden wird eine Fallanalyse aus der Wikipedia [St09] mit der eher naturwissenschaftlichen Modellierung.

1.2 Positionen und ihre Bedeutung

Im Verlauf dieses Beitrags taucht immer wieder der Begriff „Position“ auf. Positionen und positionale Systeme stellen Ordnungsprinzipien des Sozialen dar. Durch Positionen werden Wahrnehmungen und Handlungen geformt. Positionen bestimmen danach die Identität der Beteiligten, weil sie Sichtweisen und Handlungen formen und Konflikte in spezielle Bahnen lenken³. Man kann unterscheiden zwischen formalen und informellen Positionen – eine Unterscheidung, die vor allem Organisationen betrifft, die immer sowohl über eine geplante Struktur und eine sich innerhalb dieser geplanten Struktur herstellende informelle Beziehungsstruktur verfügen. Die informelle Struktur ist nie identisch mit der formellen, wohl aber hat sie einen Bezug zu dieser. Die moderne Netzwerkanalyse lässt sich insbesondere auf die Arbeiten von H. White zurückführen [Wh76], wonach Positionen in Situationen ausgehandelt werden – sie erreichen dann aber durch ihren Verhaltenszuschreibungs- und Übernahmecharakter eine gewisse Dauerhaftigkeit und Transparenz. Transparenz bedeutet, dass andere Akteure in demselben Bereich eine Vorstellung vom Verhalten von Positionen besitzen. Dies ist notwendig, um überhaupt eine Orientierung innerhalb der vielfältigen Möglichkeiten der Herstellung von Beziehungen haben zu können [Wh92]. Positionen stellen ein Stück weit Konstanz her. Sie sind ausgehandelte „Abziehbilder“ von Verhaltensweisen, die mit einer bestimmten Stellung einhergehen. Positionen abstrahieren von der Person; das bedeutet, dass Personen dem Modell nach gemäß ihrer Position handeln.

² So wandte sich schon Max Weber [We909] gegen die „energetische Kulturtheorie“ des Chemie-Nobelpreisträgers Wilhelm Ostwald. Die Auseinandersetzung hatte die fachgeschichtliche Konsequenz, dass der erste Soziologentag nicht in Leipzig, wie vorgesehen, sondern 1910 in Frankfurt stattfand.

³ Für einen kurzen Überblick über die Bedeutung von Positionen und positionalen Systemen siehe [St10].

Für Netzwerkbetrachtungen ist die positionale Analyse insbesondere deswegen so interessant, weil jede Position über ein eigenes Beziehungsmuster verfügt. Solche Beziehungsmuster lassen sich clustern, womit positionale Analyse möglich wird. Ein weiterer wichtiger soziologischer Gesichtspunkt der Beschäftigung mit Positionen ist, dass Positionen soziale Identitäten herstellen. Aus diesem Grunde ist es sinnvoll, nicht nur Knoteneigenschaften zu betrachten, sondern das gesamte Netzwerk mit seinem positionalen Gefüge zu untersuchen.

1.3 Modellinnovationen der “new Science of Networks”

Die „new science of networks“ [Ba02] hat im Wesentlichen drei neue Modelle hervorgebracht, die in Ansätzen in der sozialwissenschaftlichen Netzwerkanalyse oder ihren Vorläufern bereits vorhanden waren.

1. *Power-Law-Modelle*. Ähnliche Überlegungen aus der Ungleichverteilung von Einkommen, etwa bei Pareto oder Gini, wurden von der „new science“ in einem anderen Kontext neu erfunden (dasselbe gilt im Prinzip auch für das nachfolgend referierte Zipf'sche Gesetz). Ungleichverteilungen werden nun für Relationen untersucht. Es hat sich gezeigt, dass einige wenige Teilnehmer (Hubs) über sehr viele Beziehungen verfügen – die meisten Teilnehmer aber nur wenige Kontakte besitzen. Deutet man diesen Befund positional, dann zeigt sich, dass man bei diesem Allgemeinheitsgrad gerade mal zwei Positionen unterscheidet: den Hub und den Rest, den man dann Peripherie oder „normal“ nennen kann. Solch ungleiche Gradverteilungen wurden andernorts mit mehr inhaltlichem Anspruch als Zentrum-Peripherie Struktur bezeichnet [z.B. St01]. Die Betrachtung verfügt nur über eine geringe Komplexität, die aus der Eindimensionalität (Attribut: Grad bzw. *degree*) resultiert – das gesamte Netzwerk bleibt dabei außen vor. Der Vorteil dieses Vorgehens besteht jedoch darin, dass große Netzwerke auf diese Eigenschaft hin untersucht werden können.

2. *Small-World-Netzwerke*. Damit ist ursprünglich eine Modellierung für die Befunde eines von Stanley Milgram [Mi67] durchgeführten Experiments gemeint. Dort wurden Briefe zu einer unbekanntem, weit entfernten Person an „Duzfreunde“ weitergegeben. Es wurde dabei herausgefunden, dass die Weitergabe über erstaunlich kurze Ketten erfolgte. Da man sich dies kaum erklären konnte, kam es zu zahlreichen Nachfolgeuntersuchungen. Die Erkenntnis hierbei war, dass (neben Hubs mit vielen Verbindungen) vor allem solche Kontakte den Ausschlag für kurze Verbindungswege gaben, die über die jeweilige Nachbarschaft hinausreichten. Soziologisch wurden dazu äquivalente Beziehungen etwa in Merton und Lazarsfelds [Me68a] Unterscheidung zwischen *Locals* und *Cosmopolitans* beschrieben. Fasst man das Modell [Wa03] inhaltlich, dann sind es vor allem die *Cosmopolitans*, die über die örtlich gebundene Gruppe hinaus über heterogene Verbindungen [Bu92] verfügen. Der Hub im Small-World-Netzwerk kann über die Anzahl der Kontakte gemessen werden, wobei es auch auf die Heterogenität der Kontakte ankommt. Small-World-Modelle sind in dieser Hinsicht komplexer als das potenzgesetzliche Verteilungsmodell, das nur die Anzahl der Kontakte berücksichtigt.

3. *Preferential Attachment*. Hierbei handelt es sich um die Idee, dass vor allem im WWW Attraktivität und Sichtbarkeit eine Einheit bilden [Ba02]. Kleine Differenzen

bewirken, dass sich im Zeitverlauf die Unterschiede vergrößern bis hin zu Monopolen für bestimmte Bereiche. Auf diese Weise entstehen potenzgesetzliche Verteilungen mit quantitativ wirksamen Hubs. Dabei handelt es sich um eine Idee, die in qualitativer Hinsicht bereits als Matthäus-Prinzip bekannt war [Me68b].

Obgleich man sehen kann, dass die Ideen der „New Science of Networks“ nicht unbedingt neu waren, kann die strikte Orientierung an der Modellierung als eine Innovation angesehen werden, die in der Soziologie (mit Ausnahmen) eher selten ist. In unseren Überlegungen wenden wir diese Modelle an, wobei vor allem an das erste Modell angesetzt werden soll. Die Idee ist nun, die Modelle inhaltlich weiter anzureichern, um deren Erklärungskraft zu steigern bzw. stärker inhaltlich-soziologische Erklärungen in die Modelle einzuführen. Hierzu ist, wie oben erläutert wurde, die klassisch soziologische Orientierung am Kontext notwendig⁴. Wenn wir die Modelle inhaltlich anreichern, dann durch eine Ergänzung hinsichtlich einer positionalen Aufgliederung der im Kontext der „new science of networks“ wichtigen potenzgesetzlichen Verteilung. Unsere Hypothese besagt, dass das, was bislang als gesetzmäßiges Grundmuster von Gradverteilungen betrachtet wird, durch eine positionale Zerlegung einer viel weiterreichenden Erklärung zugeführt werden kann.

Unser Ziel ist es, die Abweichungen empirischer Gradverteilungen vom Idealbild einer potenzgesetzlichen Verteilung mit Hilfe einer externen Bestimmung von Positionen zu erklären. Auch die Verteilung der Zahl von Kollaboration⁵ in Wikipedia zeigt Abweichungen gegenüber einem Potenzgesetz. Um diese Abweichungen zu erklären, werden im übernächsten Abschnitt Bestimmungen von Positionen vorgenommen. Anschließend wird versucht, diese Positionen im Verteilungsmodell zu verorten. Hierdurch – so unsere Überlegung – müsste die Verteilung jenseits der Dichotomie „Hubs vs. Andere“ erklärbar werden. Dieses Vorgehen liefert Hinweise darauf, wie Kollaborationsprozesse ablaufen und welche Berücksichtigung unterschiedliche Beziehungsmuster bei der Konstruktion von Web-Plattformen finden müssten.

1.4 Related Work und eigener Ausgangspunkt

Es ist eine große Zahl an Untersuchungen über Kollaboration in Wikipedia im Gange oder abgeschlossen. Daher ist es kaum möglich, auf alle anderen Arbeiten hier einzugehen. Kollaboration in Wikipedia ist eine grundlegende Idee bei der Konstruktion von Artikeln. Das Zustandekommen von Wikipedia-Artikeln lässt sich einerseits durch Bezug auf die „Weisheit der Massen“ erklären. Populärwissenschaftliche Bücher [exemplarisch Su05], die dieses Prinzip thematisierten, finden sich mittlerweile viele. Auch Kritiker stellen das Prinzip der Kollaboration nicht in Frage [La06]. Selbst wissenschaftliche Autoren gehen von einer nicht-hierarchischen, auf Gleichheit beruhenden Kooperation in Wikipedia aus [z.B. Sc07; Wi06]. Zahlreiche Arbeiten beschäftigen sich zudem mit der

⁴ Schnegg [Sc06] konnte an verschiedenen Kulturen zeigen, dass das Potenzgesetz nicht für jede Kooperation gilt.

⁵ Der Begriff „Kollaboration“ wird von uns als englische Übersetzung verwendet. Im Gegensatz zum Begriff der Kooperation ist er weiter gefasst und beinhaltet auch eine Zusammenarbeit, von der die Beteiligten nicht unbedingt Kenntnis besitzen.

Motivation der Teilnehmer [z.B. Sc08; Jo08].

Untersuchungen der formalen Kollaborationsstruktur weisen auf eine große Ungleichheit in der Verteilung der Beiträge hin [St09]. Untersuchungen der Kollaborationsstruktur lassen zudem ein hohes Maß an Entropie erkennen, und zwar auf der Grundlage ungleicher Beteiligungsmuster [Ma09; MeS08]. Bei Viegas et al. [Vi07] geht es um die Visualisierung der Kooperationsstruktur – hier werden auch bezogen auf Artikel formale Positionen untersucht. Probleme der Koordination werden diskutiert [Ki09] und in Zusammenhang mit der Artikelgüte betrachtet [Ki08] – siehe hierzu auch Liu [Li09].

Insbesondere negative Beziehungen über Reverts in Wikipedia wurden von Brandes und anderen [Br09] anhand des Kollaborationsmusters in Artikeln untersucht (ähnlich Bongwon [Bo07]). Die Zusammenarbeit in bestimmten Bereichen mit dem Ziel, Autorenniquen zu identifizieren, untersuchen Rut et al. 2009 [Ru09]. Sie nutzen dazu die über Artikel und Bearbeitungen erzeugte bipartite Netzwerkstruktur.

In diesem Beitrag untersuchen wir das Kollaborationsmuster der gesamten deutschsprachigen Wikipedia (Artikel und Diskussionen inklusive). Die dabei entstehenden Kontakte über die Bearbeitung von Artikeln lässt sich mittels eines Power-Law modellieren. Dies meint, dass die Masse der Teilnehmer nur über sehr wenige, durch Koautorenschaft erzeugte (formale) Kontakte verfügt. Wenige Teilnehmer dagegen haben eine große Anzahl an Beziehungen. Hierbei muss man sich darüber im Klaren sein, dass die hier aufgezeigten Beziehungen zwar tatsächliche Kontakte beinhalten mögen, dies aber nicht im Vordergrund der Darstellung steht. Vielmehr kann man den Bezug zu anderen Teilnehmern über die Artikel als eine Art formale Beziehung bei der Konstruktion eines Artikels bezeichnen. Es wird auch nicht zwischen positiven und negativen Beziehungen unterschieden. Zudem werden über Artikel, an denen sehr viele Teilnehmer mitgearbeitet haben, auch sehr viele Beziehungen konstituiert. Dabei wächst die Zahl der auf diese Weise konstruierten Beziehungen wachsen sehr schnell an, was bedeutet, dass Teilnehmer, die vor allem an Seiten mitarbeiten, die von vielen anderen bearbeitet werden, über mehr Beziehungen verfügen, als diejenigen, die sich auf Artikel mit wenigen Autoren konzentrieren.

Die positionale Netzwerk Betrachtung, an der wir ansetzen, geht auf die Arbeiten von White und anderen zurück [Wh75; Wh76] und bezieht sich auf den Anthropologen Siegfried Nadel [Na56]. In der Zwischenzeit wurde eine Reihe von Algorithmen für die positionale Analyse entwickelt. Auf diese beziehen wir uns hier aber nicht, weil wir uns der Einfachheit halber vorwiegend an den Wikipedia-eigenen Klassifikationen von Positionen orientieren.

2 Kollaborationsnetzwerke in Wikipedia – Arbeitsteilung in Positionen

In einer größeren Untersuchung zu Wikipedia wurden verschiedene Aspekte der Kollaboration untersucht [St 09]. Wir wollen nun Erkenntnisse dieser Untersuchung auf eine Zerlegung der Potenzverteilung anwenden. Unsere Überlegung lautet, um Modelle der

Art von Potenzgesetzen zu verstehen, ist es sinnvoll, den Kontext, also mehr Wissen einzubeziehen, als dies bislang in der Forschung der Fall ist, obwohl der Vorzug allgemeinerer Modelle gerade in ihrer Übertragbarkeit auf unterschiedliche Kontexte liegt. Dieser Vorteil wird aber erkaufte durch eine kaum differenzierende Betrachtungsweise. Wir behaupten, dass eine Berücksichtigung von Kontextfaktoren in solchen Modellen eine weitergehende Erklärungskraft besitzt und möchten dies mit diesem Beitrag auf die Probe stellen.

Einen der Kontextfaktoren stellt Wikipedia als Organisation dar. Die Organisation hat im Laufe der Zeit eine Arbeitsteilung hervorgebracht, die dadurch gekennzeichnet ist, dass nicht nur eine Aufgabe zu erfüllen ist, sondern zahlreiche unterschiedliche Tätigkeiten sich herausgebildet haben. Die paradigmatische Figur in Wikipedia ist die des Artikelschreibers. Die Idee, dass alle Mitarbeiterinnen und Mitarbeiter die Phase des Artikelbearbeiters durchlaufen haben sollen, steht noch immer in Zusammenhang mit den Voraussetzungen für die Erlangung des passiven und aktiven Wahlrechts bzw. für den Aufstieg als Administrator.

In der bereits zitierten Untersuchung von Wikipedia wurde eine Operationalisierung von Positionen vorgenommen. Diese Festlegung übernehmen wir in Teilen. Dort wurden sieben Positionen unterschieden und Teilnehmer diesen Positionen zugeordnet. Es konnte gezeigt werden, dass die Positionenzuordnung fast immer ausschließlich war [St09]: von den 4.636 Teilnehmern, die einer Position aufgrund der vorgenommenen Definitionen zugeordnet wurden, kam eine Positionenüberschneidung (eine Person hat mehrere Positionen inne) in lediglich 3% der Fälle vor.

Durch die Weiterentwicklung der Inhalte und der Organisation in Wikipedia kam es zu einer Ausdifferenzierung von Positionen. Autoren konnten zu Beginn tendenziell eher Generalisten sein, heute sind sie dagegen eher Spezialisten. Generalist bedeutet, dass die Teilnehmer Artikel in verschiedenen Wissensgebieten verfassen konnten und sich darüber hinaus um die gesamte Organisation und deren Entwicklung kümmern konnten. Solche Teilnehmer kamen mit vielen unterschiedlichen Teilnehmern in Kontakt. Die Spezialisten von heute stehen eher mit anderen Autoren eines speziellen Fachgebiets in Kontakt (so die Interviewpartner unserer Studie [St09]). Für die Kollaborationsstruktur bedeutet dies, dass die Autoren ein geringeres Spektrum an Artikeln bearbeiten und hierdurch einen niedrigeren Degree-Wert (d.h. in diesem Falle: *Kollaborationsgrad*) besitzen. **Autoren** dürften damit nur in einem bestimmten inhaltlichen Bereich mit einer überschaubaren Zahl an Teilnehmern in Beziehung treten können. Das heißt, dass sie in der Verteilung der Kollaborationskontakte keinen Spitzenplatz einnehmen werden.

Administratoren sind mit einem speziellen Datenbankeintrag versehen – damit sind sie eindeutig identifizierbar. Alle Vandalismusbekämpfer sind gleichzeitig Administratoren, weil diese Position die Möglichkeit benötigt, Vandalen zu sperren, um wiederum weiteren Vandalismus vorzubeugen. Die Möglichkeit der Nutzersperrung hängt am Administratorstatus. Gleichzeitig sind Administratoren oft auch Fachautoren für ein bestimmtes Fachgebiet. Aufgrund ihres Tätigkeitsprofils, das Verwaltungsarbeiten aber auch Streitschlichtung beinhaltet, vermuten wir, dass Administratoren sowohl am oberen Ende als auch in der Mitte der Verteilung zu finden sein werden. Sie dürften über relativ

hohe Degree-Werte verfügen. Diese sollten in der Regel die Werte der reinen Artikelautoren übertreffen.

Vandalismusbekämpfer bereinigen mutwillige Zerstörungen an Artikeln – in der gesamten Bandbreite von Wikipedia. Vandalismusbekämpfer sind typische „quantitative Hubs“. Da unser Kollaborationskriterium lediglich eine gemeinsame Bearbeitung desselben Artikels vorsieht, wird der den Vandalenjägern zugeordnete Knotengrad besonders hoch sein.

Die nächste für uns interessante Gruppe von Teilnehmern ist die der **Bots**. Bots sind kleine Programme oder Skripte, die gleichförmig und häufig auftretende Aufgaben übernehmen sollen. Hierbei handelt es sich also um "nichtmenschliche Akteure", die ohne großen Aufwand viele Artikel durchforsten können. Wir erwarten, dass Bots einen Spitzenplatz in der Verteilung der Kollaborationsbeziehungen einnehmen, da sie infolge der "Mitarbeit" an vielen Artikeln über einen hohen Knotengrad verfügen müssten.

3 Experiment

In dieser Sektion unterziehen wir die Hypothese, wonach die Verteilung der Kollaborationshäufigkeiten von Wikipedianern derart in Teilverteilungen zerlegbar ist, dass die Teilverteilungen bekannten Teilnehmergruppen der Wikipedia zugeordnet werden können, einem ersten empirischen Test. Da es im Zuge dieser Verteilungsanalyse darum geht, einzelnen Ausprägungen genau einen Wikipedianer zuzuordnen, verbietet sich eine direkte Betrachtung der annahmegemäß potenzgesetzlich verteilten Wahrscheinlichkeiten $P(X=y)$, bei der y für die Anzahl der Kollaborationen steht. Vielmehr bildet die Funktion, bei der y wieder für die Anzahl der Kollaborationen steht und r für den Rang des jeweiligen Wikipedianers, den Ausgangspunkt der Betrachtung. Es zeigt sich gewissermaßen eine Zipfsche Rang-Wertverteilung [Ad00], im vorliegenden Fall genauer eine Rang-Gradverteilung, bei welcher der erste Rang vom Teilnehmer mit der höchsten Anzahl an Kollaborateuren eingenommen wird, der zweite Rang vom Teilnehmer mit der zweithöchsten Zahl an Kollaborateuren bis hin zu den untersten Rängen jener Teilnehmer, die je mit genau einem anderen Teilnehmer kollaborieren. Die Betrachtung dieser Rang-Gradverteilung verläuft quasi in Analogie zu der ursprünglich von Zipf analysierten Rang-Häufigkeitsverteilung von Wörtern [Zi72], für die Zipf ein Wert von $\alpha \sim 1$ angenommen hat (dies entspricht einem der berühmten Zipfschen Gesetze – siehe [Ra82]).

Wie in den vorangehenden Abschnitten erläutert wurde, nehmen wir ein wikibasiertes Kollaborationsnetzwerk [Me08a; Me08b] zum Ausgangspunkt, bei dem jeder Knoten genau einem registrierten Teilnehmer entspricht und bei dem immer dann zwischen zwei Knoten eine ungerichtete Kante gespannt wird, wenn es mindestens ein Dokument (z.B. einen Artikel oder ein Portal) der Wikipedia gibt, das die entsprechenden Teilnehmer gemeinsam bearbeiten. In unserem Kollaborationsmodell kodieren Kanten daher Koautorenschaft. Von den laut Wikipedia 1.220.810 registrierten Teilnehmer betrachten wir nun die Teilmenge jener 399.247 *aktiven registrierten Teilnehmer*, die in dem Speicher-auszug (d.h. Dump) namens dewiki-20100903 als aktive Teilnehmer identifizierbar sind,

indem sie sich mindestens einer Editoperation als Autor zuordnen lassen. Die Analysen sind auf keinen Namensraum beschränkt und inkludieren daher auch Pages im Namensraum *Diskussion*. Nicht registrierte Teilnehmer, die lediglich über deren IP-Adresse dem System bekannt sind, werden jedoch nicht berücksichtigt.

Diesem Ansatz folgend ergibt sich eine Verteilung der Knotengrade der Knoten in diesem Netzwerk, dass wir nach absteigendem Grad zu einer Rang-Gradverteilung ordnen und einer Verteilungsanalyse unterziehen. Hierzu berücksichtigen wir zwei Quellen für die Typisierung von knotenbildenden Teilnehmern. Unter Verwendung der Wikipedia-spezifischen Terminologie sprechen wir genauer von Benutzerteilgruppen, die durch die Zugehörigkeit zu einer oder mehreren Benutzergruppen zustande kommt (siehe Tabelle 1). Diese Benutzergruppen werden auf der Seite <http://de.wikipedia.org/wiki/Spezial:Statistik> unterschieden. Hinzu nehmen wir die in Wikipedia nicht speziell als Benutzergruppe unterschiedene Gruppe der Vandalismusbekämpfer, deren Mitglieder über die Seite <http://de.wikipedia.org/wiki/Wikipedia:Beitragszahlen/Benutzerblockaden> identifizierbar ist. Insgesamt identifizieren wir auf diese Weise 80 Vandalismusbekämpfer, die zugleich als Administratoren wie auch als Sichter fungieren (siehe Tabelle 1).

Durch Auswertung der Zugehörigkeit von Wikipedianern zu Teilnehmergruppen (formale Positionen) kann nun das Kollaborationsnetzwerk teileingefärbt und somit eine Subklassifikation der resultierenden Rang-Gradverteilung vorgenommen werden. Graphentheoretisch gesprochen haben wir damit einen einfachen Graphen vorliegen, dessen Knoten benannt und partiell typisiert sind. Der solcherart informatorisch angereicherte Graph bildet den Input der nachfolgenden Netzwerkanalyse, und zwar anstelle der üblicherweise genutzten untypisierten Graphen.

Das Ergebnis der Anpassung einer potenzgesetzlichen Funktion an die resultierende Rang-Gradverteilung aller aktiven registrierten Benutzer zeigt Abbildung 1, und zwar im Vergleich zu der entsprechenden komplementär-kumulativen Verteilung $P(X \geq y)$. Den allgemeinen theoretischen Zusammenhang von Zipfschem Gesetz $y \sim r^{-\alpha}$ und entsprechender Wahrscheinlichkeitsfunktion $P(X=y) \sim y^{-\gamma}$ erläutert Adamic [Ad00]. Danach ist der Exponent letzteren Potenzgesetzes bestimmt durch

$$\gamma = 1 + \beta = 1 + (1/\alpha) \quad (1)$$

und zwar als Funktion des Exponenten β der komplementär-kumulativen Verteilung $P(X \geq y) \sim y^{-\beta}$, für den also gilt:

$$\beta = 1/\alpha \quad (2)$$

Tabelle 1: Auflistung der untersuchten Gruppen von Mitgliedern der deutschen Wikipedia und deren Umfänge (vgl. <http://de.wikipedia.org/wiki/Spezial:Statistik>; letzter Zugriff: 08.05.2011). Die Gruppe der Vandalismusbekämpfer existiert nicht als Benutzergruppe der Wikipedia, wird aber aufgrund der Angaben auf der Seite <http://de.wikipedia.org/wiki/Wikipedia:Beitragszahlen/Benutzerblockaden> berücksichtigt.

Benutzerteilgruppe	Umfang	Gruppennamensraum
<i>aktive registrierte User</i>	399.247	
Bot	293	bot
Sichter	9.859	editor

Sichter ^ Administrator	169	editor und sysop
Sichter ^ Vandalismusbekämpfer ^ Administrator	80	editor und hunter und sysop

Tabelle 2 zeigt den beobachteten Wert für α und die entsprechenden Berechnungsergebnisse für β und γ . Die explizite Berücksichtigung von β in der vorliegenden Untersuchung hat zwei Gründe: zum einen erfolgt sie gemäß den Empfehlungen von Newman [Ne05] für die Anpassung von Potenzgesetzen. Zum anderen bestätigt sie ein Bild, das in Abbildung 1 seinen Ausdruck findet: die in Tabelle 2 wiedergegebenen Exponenten lassen in doppellogarithmischer Darstellung erkennen, dass die solcherart angepassten Funktionen stets nur den "mittleren" Teil der zugrundeliegenden beobachteten Verteilung anpassen, wobei diese Interpretation auch dann erhalten bleibt, wenn die Anpassung gemäß Newman [Ne05] an die komplementär-kumulative Verteilung erfolgt. In beiden Fällen (also im Falle der zum Ausgang gewählten Rang-Grad-Verteilung ebenso, wie im Falle der entsprechenden komplementär-kumulativen Verteilung) ergeben sich augenscheinlich mindestens drei Verteilungsbereiche, für die Potenzgesetze mit unterschiedlichen Exponenten anzunehmen sind, und zwar mit zunehmend größeren Werten, wenn man sich vom Ursprung wegbewegt. Beobachtungen dieser Art sind für das Zipfsche Gesetz bereits am Beispiel von Worthäufigkeitsverteilungen beschrieben worden und waren der Ausgangspunkt für eine Reihe von Verbesserungen [Ma61; Tu98]. Sie zeigen sich aber eben auch in dem vorliegenden Anwendungsbereich von Kollaborationsnetzwerken und geben somit Anlass zu einer differenzierteren Abbildung von Gradverteilungen – *in Abgrenzung zu den in der Literatur vielfach schlicht als einfache Potenzgesetze charakterisierten Verteilungen.*

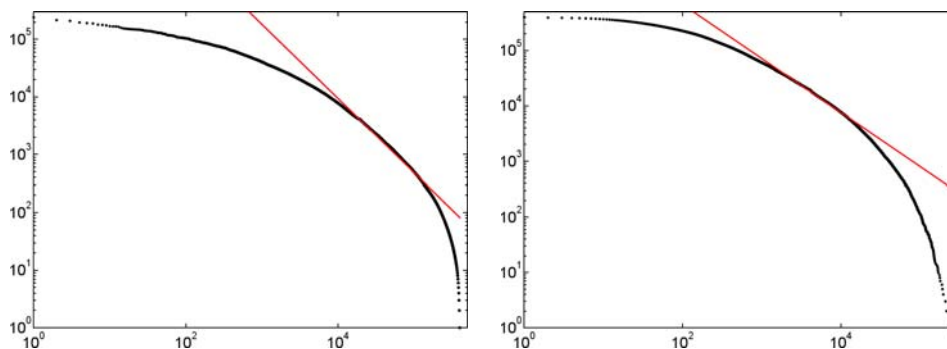


Abbildung 1: **Links:** Log-log-Plot der Verteilung der Grade aller Knoten im Kollaborationsnetzwerk der 399.247 aktiven registrierten Benutzer der Deutschen Wikipedia im Zeitraum vom 02.04.2001 bis zum 13.08.2010 auf der Grundlage des Verteilungsmodells $y \sim r^{-\alpha}$, $\alpha = 1,287$ (siehe Tabelle 2). Die Rangverteilung visualisiert 399.247 Werte (siehe Tabelle 1). **Rechts:** Anpassung des Potenzgesetzes $P(X \geq y) = Cy^{-\beta}$, $\beta = 0,976$, an die der Rang-Gradverteilung entsprechende komplementär-kumulative Verteilung.

Es sei an dieser Stelle angemerkt, dass ein Wert von $\gamma \sim 1,777$ im Bereich jener von Newman [Ne05] als "natürlich" bezeichneten Werte liegt, für die im vorliegenden Fall jedoch angenommen werden muss, dass (unabhängig vom endlichen Mittelwert der

beobachteten, empirischen Verteilung) die dazugehörige angepasste, theoretische Verteilung keinen endlichen Erwartungswert besitzt (zur formalen Begründung dieser Einschätzung siehe Newman [Ne05]). Konkret heißt das, dass man im Falle von Kollaborationsnetzwerken wie der deutschen Wikipedia *nicht* von einer mittleren Anzahl von Kooperationspartnern sprechen kann. Das wiederum bedeutet, dass es in diesem Sinne *kein* typisches Kollaborationsmuster gibt.

Tabelle 2: Ergebnisse der Anpassung von Potenzgesetzen an die Rang-Gradverteilung und an die entsprechende komplementär-kumulative Verteilung.

Verteilungstyp	Exponent	ACD	Wert gemäß Adamic [Ad00] bei bekanntem α
$y \sim r^{-\alpha}$	1,287	0,9997	
$P(X \geq y) \sim y^{-\beta}$	0,976	0,9996	$\beta \sim 0,777$
$P(X=y) \sim y^{-\gamma}$			$\gamma \sim 1,777$

Tabelle 3: Ergebnisse der Anpassung von Potenzgesetzen an die Rang-Gradverteilungen aus den Abbildungen 2-5

Benutzerteilgruppe	Exponent α	ADC	γ
Bot	0,4950	0,9897	3,02
Sichter	0,7296	0,9959	2,371
Sichter ^ Administrator	0,4849	0,9967	3,062
Sichter ^ Vandalismuskämpfer ^ Administrator	0,3237	0,9875	4,089

Betrachtet man nun die Benutzergruppen aus Tabelle 3 im Einzelnen, so ergeben sich die in den Abbildungen 2-5 wiedergegebenen Teilverteilungen bzw. die in Tabelle 3 zusammengefassten Exponenten der an diese Teilverteilungen angepassten Potenzgesetze. Dabei ist einerseits für die Gruppe der Bots wie auch für die Gruppe der Sichter bemerkenswert, dass diese jeweils über das gesamte Feld der zugrundeliegenden Rang-Gradverteilung verstreut liegen. Offenbar unterliegt ihre Verteilung einer dem zugrundeliegenden Zipfschen Gesetz vergleichbaren Gesetzmäßigkeit. Anders ausgedrückt: die Skalenfreiheit der Kollaborationsbeziehungen, wie sie die erfolgreiche Anpassung aus Tabelle 2 zum Ausdruck bringt und in Abbildung 1 visualisiert wird, scheint gleichermaßen für die Teilgesamtheiten der Bots und der Sichter zu gelten, und zwar unabhängig voneinander. In diesem Sinne wäre von einer Art Skalenfreiheit in Bezug auf positional eingegrenzte Benutzergruppen zu sprechen: *unabhängig davon, auf welche Teilgruppe man projiziert, bleibt der potenzgesetzliche Zusammenhang erhalten*. Dieser Interpretation stehen zwei Beobachtungen entgegen:

- Erstens lassen *Vandalismuskämpfer* (Abbildung 4) und *Administratoren* (Abbildung 5) eine Konzentration im oberen Teil der Verteilung erkennen, so dass man kaum von einer skalenfreien Reproduktion des Verteilungsmusters der Gesamtverteilung sprechen kann.
- Zweitens zeigt Tabelle 3, dass die an die Teilgruppen angepassten Funktionen im Hinblick auf den Exponenten α in erheblichem Maße von der Grundverteilung abweichen, und zwar so stark, dass für die Teilgesamtheiten eine zwar skalenfreie aber

dennoch entschieden eigenständige Dynamik unterstellt werden muss, die durch den Bezug auf die Gesamtverteilung unerklärt bleibt. Dies verdeutlicht zudem die letzte Spalte von Tabelle 3, welche die Werte des Exponenten γ gemäß dem Modell von Adamic [Ad00] wiedergibt. Auch hier zeigt sich eine enorme Abweichung gegenüber dem entsprechenden γ -Wert der Gesamtverteilung (siehe Tabelle 2).

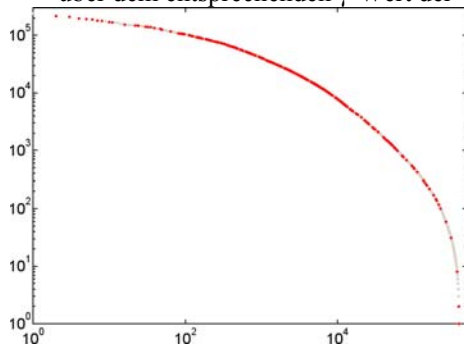


Abbildung 2: Log-log-Plot der Verteilung der Grade der **Bots** im untersuchten Kollaborationsnetzwerk: Punkte der Zielverteilung erscheinen rot vor dem Hintergrund der grau eingefärbten Restmenge der Punkte der Verteilung aus Abbildung.

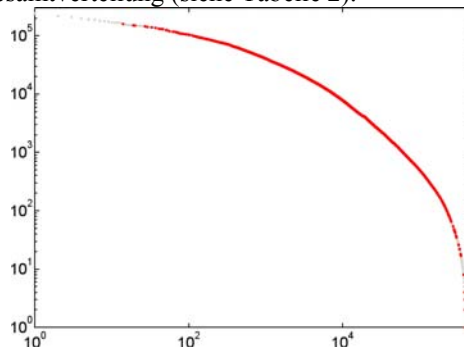


Abbildung 3: Log-log-Plot der Verteilung der Grade der **Sichter** im untersuchten Kollaborationsnetzwerk: Punkte der Zielverteilung erscheinen rot vor dem Hintergrund der grau eingefärbten Restmenge der Punkte der Verteilung aus Abbildung.

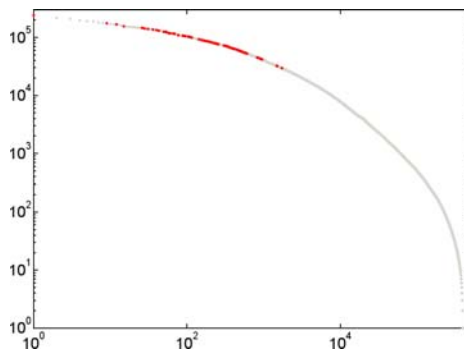


Abbildung 4: Log-log-Plot der Verteilung der Grade der **Vandalismusbekämpfer**, die zugleich Sichter und Administratoren sind, im untersuchten Kollaborationsnetzwerk: Punkte der Zielverteilung erscheinen rot vor dem Hintergrund der grau eingefärbten Restmenge der Punkte der Verteilung aus Abbildung.

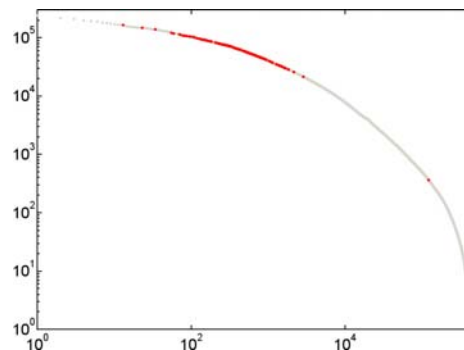


Abbildung 5: Log-log-Plot der Verteilung der Grade der **Administratoren**, die zugleich Sichter sind, im untersuchten Kollaborationsnetzwerk: Punkte der Zielverteilung erscheinen rot vor dem Hintergrund der grau eingefärbten Restmenge der Punkte der Verteilung aus Abbildung.

Diese Verteilungsanalyse steht zumindest nicht in Widerspruch zu der Eingangshypothese über die der Notwendigkeit einer gruppenspezifischen Betrachtung von Kollaborationsbeziehungen. Sie bereitet den Weg für eine tiefergehende Analyse, in der es darum geht, weit genauer als dies bisher geschehen ist, die offensichtliche Abweichung vom

Idealtypus einer potenzgesetzlichen Verteilung, wie sie in Abbildung 1 zum Ausdruck kommt, zu erklären, und zwar unter Bezug auf Knotengruppen die sich positional unterscheiden lassen. Dabei kann die Grobeinteilung in drei Verteilungsbereiche, wie sie für die quantitative Linguistik bereits gang und gäbe ist [Tu98], als Ausgangspunkt dienen, um genauere Erklärungen für Gradverteilungen zu gewinnen. Wie die in der vorliegenden Untersuchung beobachteten Gradverteilungen soziologisch genauer zu erklären sind, führt nun abschließend die folgende Sektion durch.

4. Diskussion

Es zeigt sich, dass sich die empirische Verteilung der Kollaborationen dekomponieren lässt, und zwar durch Bezug auf die Positionen der Teilnehmer. Je nachdem, welche formale Position die Teilnehmer einnehmen, differenziert sich deren Lage und deren Verteilung - teils in Abweichung vom Idealbild eines Potenzgesetzes. Dies weist darauf hin, dass Potenzgesetze die Zusammensetzung von Kollaborationsnetzwerken unzureichend darstellen. Formale Positionen stehen für Arbeitsteilung: je nach Aufgabengebiet ist die Möglichkeit, mit anderen Teilnehmern in Beziehung zu treten, unterschiedlich verteilt. Die höchsten Knotengrade werden vielfach (aber nicht ausschließlich) nur von Bots erreicht, weil diese weder zeitlichen und kognitiven Begrenzungen unterliegen. Überrascht hat uns, dass sich Bots fast über die gesamte Verteilung finden lassen. Ursache dafür ist, dass unter den etwa 300 Bots auch solche sind, die für kleinere Spezialaufgaben entwickelt wurden. Die Bots decken also eine große Bandbreite an Einsatzgebieten mit unterschiedlichen Möglichkeiten des Kontakts zu anderen Teilnehmern ab.

Als Sichter können alle engagierteren Teilnehmer bezeichnet werden. Es handelt sich um alle Mitarbeiter, die mindestens 150 Bearbeitungen vorgenommen haben. Aus diesem Grund findet man diese Position an fast allen Stellen der Verteilung. Ausnahmen sind die extremen Enden - wenn jemand die Voraussetzungen erfüllt, gehört er nicht mehr zur Masse der gelegentlichen Teilnehmer. "Normale" Mitarbeiter, wie sie durch die Sichter repräsentiert sind, decken ein breites Spektrum an Engagement ab.

Eine weitere Gruppe bilden die Administratoren. Administrator wird nur, wer sich stark engagiert. Mit dem Administratorstatus ändern sich die Aufgaben [St08] und damit auch das Teilnahmeprofil. Verwaltungsaufgaben rücken dann gegenüber dem Artikelschreiben in den Vordergrund. Ferner geht es darum, den Überblick zu bewahren, bei Streit einzugreifen und auszugleichen. Infolgedessen haben es Administratoren mit einem breiten Spektrum an Artikeln zu tun. Dies erhöht ihren Knotengrad deutlich. Die Administratoren befinden sich folgerichtig im oberen Drittel der Verteilung, was deren Ausnahmestellung sichtbar werden lässt.

Eine weitere Teilgruppe bilden die Vandalismusbekämpfer. Die Gradverteilung dieser Position ähnelt jener der Administratoren. Die Vandalismusbekämpfer kommen mit deutlich mehr Artikeln in Berührung als die normalen Administratoren, weil die Bearbeitung ganz unabhängig vom eigenen Interessensgebiet in einem weiten Spektrum erfolgt. Vandalenjäger arbeiten meist mit spezieller Software, die ihnen einen Status zwischen Bots und menschlichen Teilnehmern verleiht. Insofern ist es nicht erstaunlich, dass die

Engagiertesten von ihnen das linke obere Ende der Verteilung prägen.

5. Ausblick

In diesem Beitrag haben wir uns zunächst auf die Analyse eines kleinen Beispiels mit leicht zugänglichen Positionsbeschreibungen beschränkt. In naher Zukunft sollen in weitergehenden Analysen weitere Positionen – auch solche stärker inhaltlicher Art – berücksichtigt werden. Wir vermuten, dass sich auf diese Weise die Defizite von Potenzgesetz-orientierten Erklärungsmodellen im Hinblick auf die zugrundeliegenden Kollaborationsstrukturen beheben lassen. In diesem Sinne plädieren wir dafür, zukünftig inhaltsreichere Untersuchungen vorzunehmen, die quantitative und qualitative Methoden wie auch Erklärungsmodelle stärker miteinander verbinden.

Die positionale Ebene ist generell prägend für die Strukturierung von Netzwerken: dies betrifft zum einen die Identität der jeweiligen Teilnehmer, zum anderen bedingen Positionszugehörigkeiten aber auch Zugangsbeschränkungen (Exklusionen also), die in einem offenen Internet möglicherweise unerwünscht sind. In jedem Falle aber sind positionale Differenzierungen notwendig, um den Betrieb webbasierter Plattformen wie der Wikipedia aufrechtzuerhalten.

In künftigen Analysen planen wir eine Dekomposition der Beziehungen nach Zeiträumen der Teilnahme. In der jetzigen Version ist die gesamte Geschichte von Wikipedia zusammengefasst. Dies führt stellenweise zu sehr großen Knotengraden – besonders dort, wo viele Teilnehmer an Artikeln beteiligt sind. Dem soll durch eine zeitsensitive Betrachtung künftig entgegen gewirkt werden. Daneben wollen wir weitere Dokumenttypen in die Analyse einbeziehen [K108].

Wir sind davon überzeugt, dass Modellierungen basierend auf der Berücksichtigung von Positionen einen sehr viel weitergehenden Anwendungsbereich besitzen, als es diese Fallstudie aufzeigen konnte. Im nächsten Schritt wäre es wichtig, eine Verschiebung der Aufmerksamkeit vom individuellen Akteur mit seinen Eigenschaften (Degree), hin zu einer Analyse des gesamten Netzwerks und seiner strukturellen Zusammensetzung zu erreichen. Ein solcher Schritt würde zudem die aus soziologischer Perspektive besondere Leistung der Netzwerkforschung stärker in das Blickfeld rücken, sich nämlich dem „Dazwischen“ zuzuwenden, den Fokus also ausgehend von den Knoten hin zu den Relationen zu verschieben, zumal ein zu starkes Interesse an den Knoteneigenschaften nicht folgenlos für das gesamte Gefüge der Netzwerkforschung bliebe: Wenn es stimmt, was oben behauptet wurde, dass Positionen über Identitäten entscheiden, dann ist das, was einen Knoten bestimmt, abhängig von den Beziehungen – und nicht, wie es die knotenfixierte Sichtweise nahelegt, dass die Beziehungen von den Knoten bestimmt werden. Es ist naheliegend, dann Phänomene dieser Art anders zu modellieren und zu interpretieren.

Die Einbeziehung von Kontext in die allgemeinen Modelle kommt zudem der interdisziplinären Zusammenarbeit - hier von Soziologie und Informatik - sehr entgegen. Es kann nämlich das jeweils Interessierende miteinander in Beziehung gebracht werden, ohne die eigene fachliche Identität aufgeben zu müssen.

Literaturverzeichnis

- [Ad00] Adamic, L. A.: Zipf, power-law, Pareto – a ranking tutorial. <http://www.hpl.hp.com/research/idl/papers/ranking/>, 2000.
- [Ba02] Barabási, A.-L.: *Linked. The new science of networks*. Cambridge, Mass.: Perseus Publ., 2002.
- [B007] Bongwon S.; Chi, E. H.; Pendleton, B. A.; Kittur, A.: Us vs. Them: Understanding Social Dynamics in Wikipedia with Revert Graph Visualizations. In: IEEE Symposium on Visual Analytics Science and Technology, 2007. VAST 2007 ; Sacramento, CA, Oct. 30 - Nov. 1, 2007. Piscataway, NJ: IEEE Service Center, S. 163–170, 2007.
- [Br09] Brandes, U.; Kenis, P.; Lerner, J.; van Raaij, D.: Network analysis of collaboration structure in Wikipedia. In: Proceedings of the 18th international conference on World wide web. Madrid, Spain: ACM, S. 731–740, 2009.
- [Bu92] Burt, R. S.: *Structural holes. The social structure of competition*. Cambridge, Mass.: Harvard University Press, 1992.
- [Du91] Durkheim, É.: *Die Regeln der soziologischen Methode*. 2. Aufl. Frankfurt: Suhrkamp, 1991, zuerst 1895.
- [Fu10] Fuhse, J.; Mützel, S. (Hg.): *Relationale Soziologie. Zur kulturellen Wende der Netzwerkforschung*. Wiesbaden: VS Verlag, 2010.
- [Gr09] Gross, J. and Yellen, J.: *Graph Theory and its Applications*. Chapman & Hall/ CRC, Boca Raton, 2006.
- [Jo09] Johnson, B.: "Incentives to Contribute in Online Collaboration: Wikipedia as Collective Action". Paper presented at the annual meeting of the International Communication Association, TBA. Online verfügbar unter http://www.allacademic.com/meta/p233027_index.html, zuletzt geprüft am 03.05.2011, 2008.
- [Ki08] Kittur, A.; Kraut, R. E.: Harnessing the wisdom of crowds in Wikipedia: quality through coordination. In: Proceedings of the 2008 ACM conference on Computer supported cooperative work. San Diego, CA, USA: ACM, S. 37–46, 2008.
- [Ki09] Kittur, A.; Bryant L.; Kraut, R. E. 2009: Coordination in collective intelligence: the role of team structure and task interdependence. In: Proceedings of the 27th international conference on Human factors in computing systems. Boston, MA: ACM, S. 1495–1504.
- [Kl08] Klamma, R.; Haasler, C.: *Dynamic Network Analysis of Wikis (Proceedings of I-Know and I-Media 08)*. Graz, 2008.
- [Li09] Liu, J.; Sudha, R.: "Who Does What: Collaboration Patterns in the Wikipedia and Their Impact on Data Quality". Online verfügbar unter Available at SSRN: <http://ssrn.com/abstract=1565682>, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1565682, 2009.
- [Ma61] Mandelbrot, B.: On the theory of word frequencies and on related Markovian models of discourse. In Jakobson, R., editor, *Structure of Language and its Mathematical Aspects*, pages 190–219. American Mathematical Society, Providence, 1961.
- [Ma09] Matei, S. A.; Braun, D.; Petrache, H.: Structure of Collaboration on Wikipedia. Presentation made at Wikimania, 26.08.09, Buenos Aires. Online: <http://matei.org/ithink/2009/09/18/structure-of-collaboration-on-wikipedia/>, zuletzt geprüft am 04.05.2011, 2009.
- [Me08a] Mehler, A.: Structural similarities of complex networks: A computational model by example of wiki graphs. *Applied Artificial Intelligence*, 22(7&8):619–683; 2008.
- [Me08b] Mehler, A., Gleim, R., Ernst, A., and Waltinger, U.: WikiDB: Building interoperable wiki-based knowledge resources for semantic databases. *Sprache und Datenverarbeitung. International Journal for Language Data Processing*, 32(1):47–70; 2008.
- [MeS08] Mehler, A., Sutter, T.: Interaktive Textproduktion in Wiki-basierten Kommunikationssystemen. In: Zerfaß, A., Welker, M und J. Schmidt (Hg.): *Kommunikation, Partizipation und Wirkungen im Social Web – Weblogs, Wikis, Podcasts und Communities aus interdisziplinärer Sicht*. Köln: Halem, 267-300, 2008.

- [Me68a] Merton, R. K.: Social Theory and Social Structure. New York: Free Press, 1968.
- [Me68b] Merton, R. K.: The Matthew Effect in Science. In: Science 159: 56–63, 1968.
- [Mi67] Milgram, S.: The small world problem. Psychology Today 2: 60–67, 1967.
- [Na57] Nadel, S. F.: Theory of Social Structure. London: Routledge, 1957.
- [Ne05] Newman, M. E. J.: Power laws, Pareto distributions and Zipf's law. Contemporary Physics 46: 323–351, 2005.
- [Ra82] Rapoport, A.: Zipf's law re-visited. In Guiter, H. and Arapov, M. V., editors, Studies on Zipf's Law, pages 1–28. Brockmeyer, Bochum, 1982.
- [Ru 09] Rut J.; Schwartz, M.; Lehmann, S.: Bipartite networks of Wikipedia's articles and authors: a meso-level approach. In: Proceedings of the 5th International Symposium on Wikis and Open Collaboration. Orlando, Florida: ACM, S. 1–10, 2009.
- [Sc07] Schmalz, S.: Zwischen Kooperation und Kollaboration, zwischen Hierarchie und Heterarchie. Organisationsprinzipien und -strukturen von Wikis. In: Stegbauer, C.; Schmidt J.; Schönberger K. (Hg.): Wikis: Diskurse, Theorien und Anwendungen. kommunikation@gesellschaft, Sonderheft, 2007.
- [Sc06] Schnegg, M.: Reciprocity and the Emergence of Power Laws in Social Networks. International Journal of Modern Physics C 17: 1067–1076, 2006.
- [Sc08] Schroer, J.: Wikipedia: auslösende und aufrechterhaltende Faktoren der freiwilligen Mitarbeit an einem Web-2.0-Projekt. Berlin: Logos, 2008.
- [St01] Stegbauer, C.: Grenzen virtueller Gemeinschaft. Strukturen internetbasierter Kommunikationsforen. Wiesbaden: Westdt. Verl., 2001.
- [St08] Stegbauer, C.: "Die Invasion der Physiker": Naturwissenschaft und Soziologie in der Netzwerkanalyse. *Source*: Die Natur der Gesellschaft: Verhandlungen des 33. Kongresses der Deutschen Gesellschaft für Soziologie in Kassel 2006. Teilbd. 1 u. 2. Rehberg, K.-S. (Ed.) S. 1060–1077. Frankfurt : Campus 2008
- [St09] Stegbauer, C.: Wikipedia. Das Rätsel der Kollaboration. Wiesbaden: VS, 2009.
- [St10] Stegbauer, C.: Positionen und positionale Systeme. In: Stegbauer, C.; Häußling, R. (Hg.): Handbuch Netzwerkforschung: VS Verlag, S. 135–144, 2010.
- [Su05] Surowiecki, J.: Die Weisheit der Vielen. Warum Gruppen klüger sind als Einzelne und wie wir das kollektive Wissen für unser wirtschaftliches, soziales und politisches Handeln nützen können. München: Bertelsmann, 2005.
- [Tu98] Tuldava, J.: Probleme und Methoden der quantitativ-systemischen Lexikologie. Wissenschaftlicher Verlag, Trier, 1998.
- [Vi07] Viegas, F. B.; Wattenberg, M.; Kriss, J.; van Ham, F.: Talk Before You Type: Coordination in Wikipedia. In: Proceedings of the 40th Annual Hawaii International Conference on System Sciences: IEEE Computer Society, 2007.
- [We909] Weber, M.: Gesammelte Aufsätze zur Wissenschaftslehre. Neue Aufl. 1922.
- [Wa03] Watts, D. J.: Six Degrees. The Science of a Connected Age. New York: Norton, 2003.
- [Wa98] Watts, D. J.; Strogatz, S. H.: Collective dynamics of "small-world" networks. In: Nature 393, 6684: 440–442, 1998.
- [Wi06] Wiebrands, C.: Collaboration and communication via wiki : The experience of Curtin University Library and Information Service. (Australian Library and Information Association). Online verfügbar unter <http://hdl.handle.net/10760/8249>, 2006.
- [Wh76] White, H.; Boorman, S.; Breiger, R.: Social structure from multiple networks. I.: Block-models of roles and positions. In: American Journal of Sociology 81: 730–750, 1976.
- [Wh75] White, H. : Breiger, R.: Pattern across Networks. In: Society 12: 68–73, 1975.
- [Zi72] Zipf, G. K.: Human Behavior and the Principle of Least Effort. An Introduction to Human Ecology. Hafner Publishing Company, New York, 1972.