

Analyse und Visualisierung interdisziplinärer Publikationsnetzwerke

Martin Bauschmann¹⁾, Carolin Ahnert²⁾

- 1) Universitätsbibliothek, martin.bauschmann@bibliothek.tu-chemnitz.de, Journal Manager, Technische Universität Chemnitz, Straße der Nationen 62, 09126 Chemnitz, Deutschland
- 2) Universitätsbibliothek, carolin.ahnert@bibliothek.tu-chemnitz.de, Fachreferentin für Informatik, Soziologie und Wirtschaftswissenschaften, Technische Universität Chemnitz, Straße der Nationen 62, 09126 Chemnitz, Deutschland

Stichworte

Bibliometrie, Soziale Netzwerkanalyse, Publikationsnetzwerk, Interdisziplinarität, Visualisierung.

Keywords

Bibliometrics, Social Network Analysis, Publication Network, Interdisciplinarity, Visualisation.

Abstract (Deutsch)

Soziale Netzwerkanalysen und -visualisierungen auf Basis bibliometrischer Daten eröffnen neue Perspektiven auf komplexe Phänomene der Wissenschaftskommunikation wie etwa Interdisziplinarität, thematische Kooperationsmuster und Internationalisierung der Forschung. In der Untersuchung zeigen wir die Möglichkeiten der sozialen Netzwerkanalyse am Beispiel der instituts- bzw. fakultäts-übergreifenden Publikationstätigkeit von Angehörigen der Technischen Universität Chemnitz in den Jahren 2010 bis 2015. Interdisziplinäre Forschungsleistungen innerhalb der Hochschule werden damit sichtbar gemacht.

Aufbauend auf einer prägnanten Erläuterung der theoretischen Grundlagen zu sozialen Netzwerken und zu deren Analyse beschreiben wir eingehend die Erstellung des Publikationsnetzwerkes. Zudem geben wir graphentheoretische Interpretationsansätze für das Netzwerk. Abschließend diskutieren wir den potenziellen Nutzen des Projektes sowie methodische Erweiterungsmöglichkeiten für künftige Analysen.

Abstract (English)

Social network analysis and visualisation based on bibliometric data offers new insights into complex scientific phenomena, such as interdisciplinarity, topical patterns in science and the internationalization of research. In this paper we demonstrate the potential of social network analysis using the example of cross-institutional and cross-faculty publication activity by members of the Technische Universität Chemnitz in the period between 2010 and 2015. Thus, latent interdisciplinary research achievements within the university can be made visible.

Building on a poignant description of the theoretical basics of social network analysis, the construction and visual mapping of the publication network is described in detail. Furthermore, we present interpretive approaches for several graph analyses of the network. Finally, we discuss the potential value of the project for different target groups and possibilities for methodical enrichment.

1 Einleitung

Netzwerke sind das Gewebe, zu denen sich soziale Interaktionen zwischen Akteuren jeglicher Art formen. Die Untersuchung von Strukturen und Texturen dieser Gewebe ist Gegenstand der sozialen Netzwerkanalyse (SNA). Ihr wesentliches Forschungsinteresse liegt dabei in der Erschließung sozialer Ordnungsmuster. Aus der Sozialethnographie heraus geboren, werden die Methoden der sozialen Netzwerkanalyse inzwischen in einer Vielzahl an Disziplinen, von der Informatik und Informationswissenschaft über die Biologie und Medizin bis hin zu den Sprachwissenschaften und der Linguistik, zur Gewinnung von Erkenntnissen über Akteure und deren Beziehung angewendet (Schneeg und Lang 2002).

Auch das Publizieren als wichtiges soziales Phänomen in der Wissenschaft wird bereits seit geraumer Zeit mit den Methoden der sozialen Netzwerkanalyse beleuchtet. Das Forschungsgebiet, das sich mit der quantitativen Analyse von wissenschaftlichen Publikationen und deren Zitationen befasst, ist die Bibliometrie (Havemann 2009, S. 9).

Als Bestandteile bibliometrischer Netzwerke können prinzipiell alle bibliografischen Daten fungieren. Häufig genutzt werden Autorenangaben (Ko-Autoren-Netzwerke), Affiliations-Angaben (institutionelle Kooperationsnetzwerke), Zitationen (z.B. Ko-Zitationsnetzwerke von Autoren oder Zeitschriften) oder Wörter in Titeln und Abstracts (Ko-Wort-Analysen).

Neben der quantitativen Analyse von Netzwerken mittels Verfahren der mathematischen Graphentheorie ist ihre topologische Darstellung ein zentrales Interpretationsinstrument bibliometrischer Netzwerkanalysen. Dieses wird mit dem Schlagwort der „Wissenschaftskartierung“ (die englische Bezeichnung „Science Mapping“ ist verbreiteter) bezeichnet. Small definiert den Begriff folgendermaßen:

“A map of science is a spatial representation of how disciplines, fields, specialties, and individual documents or authors are related to one another” (Small 1999, S. 799).

Gerade in dem Aspekt der Räumlichkeit liegt dabei ein wichtiger heuristischer Nutzen: Weil der Beziehung bibliografischer Daten keine inhärente Zwei-, Drei- oder N-Dimensionalität zukommt, entspricht erst die algorithmengestützte Lokalisierung dieser Daten dem stark räumlich orientierten Denken des Menschen, das reale Entitäten häufig in ein mentales Modell aus Lage- und Standortbeziehungen zwischen symbolischen Entsprechungen der Sachverhalte überführt (ebd., S. 799).

Da Publikations- und Zitationsdaten Ausdruck sozialer Strukturen und Tendenzen innerhalb der Wissenschaftskommunikation sind, kann die Kombination von Methoden der Bibliometrie mit denen der Sozialen Netzwerkanalyse wichtige Erkenntnisse über autorenbezogene, institutionelle und thematische Zusammenhänge in der Wissenschaft zutage fördern. So erhalten latente oder unüberschaubare Phänomene in der Forschung, etwa Interdisziplinarität und Internationalität, neue Konturen. Demnach sind bibliometrische Netzwerke ganz dem Erkenntnisanspruch des übergeordneten Konzeptes der Informationsvisualisierung verpflichtet. Dessen Wesen fassen Card, Mackinlay und Shneiderman folgendermaßen definitorisch ein:

„The use of computer-supported, interactive, visual representations of abstract data in order to amplify cognition” (Card et al. 1999, S. 7).

In einem Projekt im Rahmen des Bibliometrie-Services der Universitätsbibliothek Chemnitz wurde der Versuch unternommen, die interdisziplinäre Publikationstätigkeit an der Technischen Universität Chemnitz im Zeitraum zwischen 2010 und 2015 mit den Mitteln der sozialen Netzwerkanalyse zu untersuchen und zu visualisieren.

2 Merkmale sozialer Netzwerke

Der erste Hauptbestandteil von Netzwerken, der Akteursbegriff, ist dabei in einem weit gefassten Sinne als die Gesamtheit von handelnden oder an Handlungen beteiligten Subjekten und Objekten zu verstehen. Dies umfasst einzelne Personen, Personengruppen, Organisationen bzw. Organisationseinheiten, aber auch eine Vielzahl weiterer Daseinsformen wie z.B. Gene, Tiere, Nationen, Kommunikationsinhalte und Publikationen.

In ein und demselben Netzwerk können ein Akteurstyp (unimodales Netzwerk) oder mehrere Akteurstypen (multimodales Netzwerk) gemeinsam dargestellt werden. Bei multimodalen Netzwerken bestehen Beziehungen nur zwischen Akteuren verschiedener Typen, nicht aber innerhalb eines Typs. Ein typisches Beispiel sind bimodale Netzwerke zur Darstellung von Mitgliedschaften mit Personen und Organisationen als Akteurstypen.

Ebenso vielgestaltig wie die Akteure sind die untersuchbaren Beziehungen zwischen ihnen als zweiter Bestandteil sozialer Netzwerke. Häufig in Studien anzutreffen sind etwa Verwandtschafts- und Freundschaftsbeziehungen, Mitgliedschaften, Weisungs- und Kooperationsbeziehungen sowie Wissens- und Informationsflüsse.

Eine methodisch wichtige definitorische Grobunterscheidung von Beziehungsarten ist die in gerichtete und ungerichtete Beziehungen. Gerichtete Beziehungen sind asymmetrisch, d.h. die Beziehung geht von Akteur A zu Akteur B, aber nicht umgekehrt. Ein Beispiel ist die einseitige Kommunikation von Sender zu Empfänger. Hingegen sind ungerichtete Beziehungen symmetrisch, d.h. die Richtung der Beziehung ist unerheblich. Dies trifft etwa auf Verwandtschaftsbeziehungen zu.

Weiteren Informationsreichtum erhalten soziale Netzwerke durch den dritten Hauptbestandteil, die Attribute bzw. Merkmale von Akteuren und Beziehungen. Akteursmerkmale wie Alter, Herkunft, Geschlecht eröffnen zusätzliche Analyseebenen, Beziehungsmerkmale wie Intensität oder Dauer erlauben die Gewichtung von Interaktionen und damit Aussagen, die über das dichotome Vorhandensein bzw. Nichtvorhandensein von Beziehungen hinausgehen.

3 Grundlagen der Netzwerkanalyse

3.1 Darstellung sozialer Netzwerke

Es gibt typischerweise drei Darstellungsformen von sozialen Netzwerken: als Matrix, als Liste und als Graph.

Bei der Matrixdarstellung sind alle Akteure jeweils einmal als Zeile und als Spalte in gleicher Reihenfolge abgebildet. Die einzelnen Zellen geben Auskunft über die Beziehung zwischen den beiden Akteuren deren Zeile und Spalte sich an diesem Punkt kreuzen. Es kann entweder das bloße Vorhandensein bzw. Nichtvorhandensein einer Beziehung (1 für Beziehung vorhanden, 0 für nicht vorhanden) oder die Intensität der Beziehung (jede beliebige reelle Zahl zur Abstufung der Beziehungsstärke, -dauer, -häufigkeit etc.) dargestellt werden.

Zwar gewährt die Matrixdarstellung bei kleineren Netzwerken einen Überblick über die Symmetrie der Beziehungen, bei großen Netzwerken wird sie jedoch rasch unübersichtlich.

Diese Leerstelle wird inzwischen von der Darstellungsform des Graphen effektiv gefüllt. Dabei werden die Akteure als *Knoten* (engl. *nodes*) und die Beziehungen bei Gerichtetheit als Pfeile (engl. *arcs*) und bei Ungerichtetheit als *Kanten* (engl. *edges*) in einem zweidimensionalen Raum dargestellt. Dank leistungsstarker computergestützter Mapping-Algorithmen können die strukturellen Merkmale, wie z.B. die Anzahl und Intensität von Beziehungen, zentrale Akteure und Gruppen stark verbundener Akteure, auch von komplexen Netzwerken offengelegt werden. Mit den Mitteln dynamischer Skriptsprachen, insbesondere Javascript, lassen sich die Graphen und ihre Komponenten (Akteure, Beziehung, Attribute) interaktiv zugänglich machen. Die damit verbundene Möglichkeit, den Betrachtungsfokus zwischen den Details und den ganzheitlichen, oft sehr komplexen Zusammenhängen eines Netzwerkes zu variieren, begründet das enorme Erkenntnispotenzial moderner Graphendarstellungen.

Die dritte Darstellungsform von Netzwerken, als Listen, hat sich als Ausgangsbasis für Software zur Analyse und Visualisierung von Netzwerken (z.B. Pajek, Gephi, UCINET) etabliert. Üblicherweise werden dabei die Akteure in einer Liste mit eindeutigen Identifikationsnummern und ihren zugehörigen Attributen zeilenweise dargestellt. In einer separaten Liste werden Beziehungen, ebenfalls zeilenweise, anhand der Verbindung der Identifikationsnummern von Akteuren aufgeführt.

3.2 Netzwerkstatistiken

Zur quantitativen Analyse sozialer Netzwerke und deren Knoten werden Verfahren und Metriken der mathematischen Graphentheorie herangezogen. Es gilt dabei zu beachten, dass die graphentheoretischen Konzepte traditionell auf ungewichtete unimodale Netzwerke angewendet werden. Für gewichtete und/oder multimodale Netzwerke sind zum Teil erhebliche Anpassungen der Berechnungs- und Interpretationsgrundlagen notwendig (Calero Valdez et al. 2012, 610 f.). Wichtige graphentheoretische Konzepte für ein gesamtes Netzwerk sind u.a.:

- Der *Durchmesser* (engl. *diameter*) eines Netzwerkes ist die maximale geodätische Distanz innerhalb eines Netzwerkes. Wenn die Länge des Pfades, gemessen an der Anzahl der verbindenden Kanten, zwischen zwei beliebigen Knoten als die Distanz zwischen ihnen bezeichnet wird, so ist die geodätische Distanz der kürzeste Pfad zwischen ihnen. Der Durchmesser gibt mithin eine erste Auskunft über die Breite eines Netzwerkes (Wasserman und Faust 1994, 110 f.).
- Die *Verbundenheit* (engl. *connectivity*) eines Netzwerkes bemisst sich danach, ob man von jedem Knoten über die Kanten des Netzwerkes jeden beliebigen anderen Knoten erreichen kann. Ist ein ungerichteter Graph unverbunden – man nennt solche Graphen auch unzusammenhängend –, so lassen sich isolierte *Subgraphen* (engl. *components*) identifizieren (ebd., S. 109).
- Die *Dichte* (engl. *density*) gibt Auskunft über den Grad der Verbundenheit der Knoten eines Netzwerkes, indem sie den Anteil der tatsächlichen zu der Anzahl aller möglichen Beziehungen in Relation setzt. Bei einer Dichte von 1 sind alle Knoten eines Netzwerkes miteinander verbunden. Die Dichte ist von der Größe eines Netzwerkes abhängig (ebd., S. 102).
- Die *durchschnittliche Pfadlänge* (engl. *average path length*) bezeichnet die durchschnittliche Anzahl der Kanten auf dem kürzesten Weg (geodätische Distanz) zwischen jedem beliebigen Knoten und jedem anderen beliebigen Knoten des Netzwerkes. Je höher die Zahl, desto schwächer ist die direkte Kooperation in einem Netzwerk ausgeprägt und desto wahrscheinlicher ist die Existenz zentraler Knoten, die als Mittler viele andere Knoten miteinander verbinden (Calero Valdez et al. 2012, S. 611). Die durchschnittliche Pfadlänge kann nur in verbundenen (Sub-) Graphen bestimmt werden, da die Pfadlänge zwischen den Knoten zweier unverbundener (Sub-)Graphen unendlich (∞) ist.

Wichtige graphentheoretische Konzepte für einzelne Knoten eines Netzwerkes sind u.a.:

- Der *Grad* (engl. *degree*) eines Knotens bezeichnet die Anzahl der Kanten, welche mit dem Knoten verbunden sind. Er ist damit ein Maß für die Aktivität eines Akteurs (Freeman 1978, S. 221). Ermittelt man den durchschnittlichen Grad aller Knoten des Netzwerkes (engl. *average degree*), so erhält man ein Maß für den Grad der Verflechtung innerhalb des Netzwerkes, welches im Gegensatz zur Dichte unabhängig von der Netzwerkgröße ist (Schneegg und Lang 2002, S. 37). Meist wird der Grad eines Knotens mit der Gesamtzahl der restlichen Knoten in Relation gesetzt. Das so errechnete normalisierte Maß wird als *Grad-Zentralität* (engl. *degree centrality*) bezeichnet.
- Bei der *Nähe-Zentralität* (engl. *closeness centrality*) wird für jeden Knoten berechnet, wie groß die geodätischen Distanzen zu allen anderen Knoten des Netzwerkes sind. Die Distanzen werden aufsummiert und an der Größe des Netzwerkes normalisiert. Mit Hilfe dieses Maßes lässt sich eine Aussage darüber treffen, wie gut eingebunden ein Knoten in den Informations- bzw. Interaktionsfluss eines Netzwerkes ist (Schneegg und Lang 2002, S. 38). Zentralität wird hier in der Regel als Autonomiepotenzial eines Knotens interpretiert (Freeman 1978, S. 225)).
- Die *Betweenness-Zentralität* eines Knotens bezeichnet den Anteil aller geodätischen Distanzen (kürzesten Pfade) zwischen beliebigen Knoten eines Netzwerkes, auf denen der Knoten tatsächlich liegt, normalisiert an der Gesamtzahl geodätischer Distanzen (Schneegg und Lang 2002, S. 39). Zentralität wird dabei an der Häufigkeit bemessen, mit der ein Knoten als exklusiver Mittler zwischen anderen Knoten innerhalb des Systems fungiert, und damit als Kontrollpotenzial konzeptualisiert (Freeman 1978, S. 226).

4 Interdisziplinäres Publikationsnetzwerk an der TU Chemnitz

4.1 Datenbereinigung und -aufbereitung

Die zugrunde liegenden Daten für die Netzwerkanalyse stammen aus der Universitätsbibliographie der Technischen Universität Chemnitz. In die Auswertung flossen alle Artikel in Fachzeitschriften (referiert und nicht referiert), Konferenzbeiträge (referiert und nicht referiert), Buchveröffentlichungen, Buchbeiträge und sonstige referierte Publikationen ein, die zwischen 2010 und 2015 von Angehörigen der Technischen Universität Chemnitz veröffentlicht wurden (Abruf der Daten erfolgte am 31.01.2016). Mittels eines Python-Skriptes wurde die Bibliographie nach den genannten Kriterien abgefragt und die Ergebnisse in einer Textdatei (Tabstopp-getrennt) abgelegt.

Ganz gleich aus welcher Datenbank bibliometrische Daten bezogen werden – eine Bereinigung von fehlerhaften oder mehrdeutigen Angaben ist stets notwendig. Für die Netzwerkanalyse wurde die Datenaufbereitung überwiegend mit der freien Software OpenRefine¹ durchgeführt. Diese erlaubt das halbautomatisierte Editieren, Filtern und Vereinheitlichen von Inhalten in allen Zeilen einer Spalte mit Hilfe von Transformationsausdrücken und Clusteralgorithmen. Diese Funktionen haben sich besonders bei der Vereinheitlichung von Autorennamen und der Aufspaltung von heterogenen Informationen in bestimmten Spalten (z.B. Trennung von Kostenstellen und Autorennamen in der Verfasserspalte) als nützlich erwiesen. Ergänzende Bereinigungs-schritte – etwa eine Dublettenprüfung oder der Import von Verlagsnamen aus externen Datentabellen – wurden mit Tabellenkalkulationsprogrammen durchgeführt.

Aus den bereinigten Rohdaten wurden dann – wiederum mit OpenRefine und Tabellenkalkulationsprogrammen – alle Publikationen mit mindestens zwei verschiedenen beteiligten Instituten herausgefiltert. Unter der Annahme, dass die Untergliederung der Fakultäten in Institute – wenngleich in verschiedenem Maße – einer disziplinären Differenzierung unterliegt, ist die Beteiligung von Autoren verschiedener Institute ein zwar grobgranularer, aber forschungsökonomisch handhabbarer Proxy für interdisziplinäre Ko-Publikationstätigkeiten innerhalb der Universität. Die Fakultät für Wirtschaftswissenschaften und die Fakultät für Informatik, welche keine Institutsstrukturen aufweisen, wurden nur auf Fakultätsebene ausgewertet. Die einzelnen Zentralen Einrichtungen der Universität, z.B. das Hochschulrechenzentrum oder die Universitätsbibliothek, wurden ebenfalls berücksichtigt.

Aus den gesammelten Daten wurde eine Tabelle mit den Knoten und eine Tabelle mit den Kanten des Netzwerkes als Textdateien im CSV-Format (comma-separated values) extrahiert.

4.2 Struktur des Netzwerkes

Die in Ko-Autoren-Netzwerken typischerweise genutzte unimodale Struktur mit nur einem Knotentyp, den Autoren, wurde im interdisziplinären Publikationsnetzwerk der TU Chemnitz zu einer multimodalen (trimodalen) Struktur erweitert. Durch die Integration der beiden Knotentypen Institut/Fakultät und Publikation werden Informations- und Interpretationsgehalt der Darstellung erhöht.

Wenn die Menge der Autoren als $A = \{a_1, \dots, a_n\}$, die Menge der Institute/Einrichtungen/Fakultäten als $I = \{i_1, \dots, i_n\}$, und die Menge der Publikationen als $P = \{p_1, \dots, p_n\}$ definiert werde, dann gelte für die Gesamtheit der Knoten $V: A, I, P \subseteq V$. Die Menge der Kanten E_1 , die von einem Knoten in A zu einem Knoten in I gehen, sei definiert als: $E_1(A, I) = \{(a; i) : a \in A \text{ und } i \in I\}$. Die Menge der Kanten E_2 , die von einem Knoten in A zu einem Knoten in P geht, sei definiert als: $E_2(A, P) = \{(a; p) : a \in A \text{ und } p \in P\}$. Für die Gesamtheit der Kanten E gelte: $E_1, E_2 \subseteq E$. Der ungerichtete, ungewichtete, trimodale Graph G ist nun definiert als: $G = (V, E)$.

4.3 Visualisierung des Netzwerkes

Die Visualisierung des Netzwerkes erfolgte mit der freien Software Gephi². Nach dem Import der Knoten- und Kantentabelle im CSV-Format wurde der Layout-Algorithmus *Force-Atlas 2* (Jacomy et al. 2014) verwendet, um das Netzwerk in einem zweidimensionalen Raum grafisch anzuordnen. *Force-Atlas 2* modelliert ein physikalisches System des Kräfteausgleichs: Knoten stoßen einander ab (ähnlich wie Magneten), Kanten bewirken – ähnlich wie eine gedehnte Zugfeder – die Anziehung der durch sie verbundenen Knoten (ebd.). Zusätzlich wurde die Knotengröße in die Abstoßungskräfte einbezogen,

¹ Verfügbar unter: <https://github.com/OpenRefine/OpenRefine>

² Verfügbar unter: <https://gephi.org/>

um eine möglichst überlappungsfreie Darstellung zu erhalten. Mit dieser Konfiguration ist der *ForceAtlas 2* ein leistungsstarker Algorithmus, der die schnelle und gleichermaßen präzise Anordnung großer Netzwerke erlaubt.

Die Netzwerkvisualisierung erlaubt so für Betrachter den Schluss, dass:

- Autoren umso näher beieinanderstehen, wenn sie demselben Institut/derselben Fakultät angehören,
- Autoren umso näher beieinanderstehen, je häufiger sie miteinander publiziert haben,
- aufgrund der Verbindung zwischen Autoren und ihren Instituten/Fakultäten auch Institute bzw. Fakultäten, deren Autoren häufig miteinander publiziert haben, nah zueinander stehen.

Eine Stärke von Gephi sind die sehr flexiblen Gestaltungsmöglichkeiten. So wurden einige Knoten frei mit der Maus verschoben und Farb- und Größeneinstellungen vorgenommen, um eine bessere Sichtbarkeit zu erreichen. Die Größe der Knoten macht die einzelnen Typen unterscheidbar: Institute/Fakultäten sind am größten, Autoren sind mittelgroß und Publikationen klein dargestellt. Die Größe der Autoren-Knoten ist zudem proportional zu der Anzahl der Publikationen, an denen die Autoren beteiligt sind. Institute und Autoren sind in den Farben der Fakultäten, denen sie organisatorisch zugeordnet sind, dargestellt.

Die Gesamtheit der Erkenntnismöglichkeiten aus Netzwerken eröffnet sich erst durch eine interaktive Darstellung. Gephi bietet hierzu mit dem zuschaltbaren Plugin *Sigmajs Exporter*³ einen einfachen Export-Workflow zur Überführung einer Netzwerkvisualisierung in eine browserfähige Web-Anwendung. Das interaktive Netzwerk ist verfügbar unter:

<https://www.tu-chemnitz.de/ub/publizieren/bibliometrie/network/>

Darin können einzelne Knoten des Netzwerkes gesucht und die Darstellung über Filter auf einzelne Fakultäten eingegrenzt werden. Im Netzwerk selbst werden zu jedem Knoten die zugehörige Bezeichnung und alle Verbindungen zu anderen Knoten des Netzwerkes beim Herüberfahren mit dem Mauszeiger dargestellt. Nach Anklicken eines Knotens erscheinen weitere Informationen wie etwa Publikationsdaten und Graphstatistiken in einem Seitenfester. Bestehende interdisziplinäre Publikationscluster innerhalb der Technischen Universität Chemnitz werden dadurch nicht nur sichtbar, sondern auch erkundbar.

4.4 Graphentheoretische Analyse

In der Sozialen Netzwerkanalyse besteht eine gegenläufige Abhängigkeit zwischen dem Ziel, die reiche Vielfalt und die Nuancen von sozialen Systemen mit multimodalen Netzwerkdarstellungen zu visualisieren, und dem Wunsch, aus den zugrunde liegenden Daten quantitative Netzwerkstatistiken zu aggregieren. Zwar existieren für den Großteil der in Abschnitt 3.2 vorgestellten graphentheoretischen Konzepte Modifikationen für bi- und multimodale Netzwerke (Borgatti und Halgin 2011) diese modifizierten Formeln und Algorithmen sind bisher jedoch kaum in den Funktionsumfang der bewährten Software zur Netzwerkanalyse integriert worden. Dies gilt auch für die verwendete Anwendung Gephi, die alle statistischen Berechnungen unter der Annahme ausführt, dass unimodale Daten zugrunde liegen.

Zudem spiegeln sich in dem interdisziplinären Netzwerk notwendigerweise die sehr verschiedenen Publikationspraktiken und –themen der Disziplinen sowie die Heterogenität der drei dargestellten Knotentypen, wodurch quantitative Abstraktionen erschwert werden.

Dennoch soll nicht gänzlich auf die der Netzwerkanalyse inhärente Möglichkeit der Synthese von Visualisierung und soziometrischen Daten verzichtet werden. Denn gerade diese Kombination ermöglicht es dem Betrachter, begründete Vermutungen über sichtbare strukturelle Merkmale anzustellen und damit das Verständnis analysierter Netzwerke zu vertiefen (Valdez et al., 2012, S. 612).

Das interdisziplinäre Publikationsnetzwerk an der TU Chemnitz besteht aus 31 Instituten/Fakultäten, 760 Autoren und 736 Publikationen. Der Graph zerfällt in einen großen Subgraphen, der 99,5 % aller Knoten enthält, und einen kleineren Subgraphen, der die Kooperationstätigkeit zwischen den Autoren des Instituts für Politikwissenschaft und des Instituts für Europäische Studien darstellt.

³ Verfügbar unter: <https://marketplace.gephi.org/plugin/sigmajs-exporter/>

Der große Subgraph wurde auf weitere Graphstatistiken hin ausgewertet. Nach manueller Anpassung der Berechnung⁴ an die trimodale Struktur des Netzwerkes, ergibt sich eine *Dichte (density)* von 0,007. Mithin sind 0,7 % aller möglichen Beziehungen tatsächlich im Graphen realisiert. Der Wert ist deshalb so gering, weil in dieser graphentheoretischen Abstraktion die Interdisziplinarität von den logischen Beschränkungen des trimodalen Netzwerkaufbaus überlagert wird. So ist es leicht ersichtlich, dass beispielsweise nur ein kleiner Teil der Autoren an einer bestimmten Publikation mitgewirkt hat oder dass ein Institut nur einen bestimmten Anteil der Autoren umfasst. Um die auf dem Phänomen der Interdisziplinarität beruhende Verbundenheit im Netzwerk zu bestimmen, wurde der große verbundene Subgraph des trimodalen Netzwerkes mit Hilfe eines Gephi-Plugins⁵ zur Netzwerkprojektion in ein unimodales Netzwerk (siehe Abb. 1) transformiert, das die Institute/Fakultäten, deren Autoren gemeinsam publiziert haben, miteinander in Beziehung setzt. Für dieses Netzwerk ergibt sich eine *Dichte* von 0,296. Demnach sind rund 30 % der theoretisch möglichen Beziehungen tatsächlich realisiert. Angesichts des breitgefächerten disziplinären Charakters des Netzwerkes ist dieser Wert ein erstes quantifizierbares Indiz für eine ausgeprägte interdisziplinäre Verbundenheit.

Dies lässt sich anhand des *durchschnittlichen Grades (average degree)* aller Knoten des Netzwerkes bestätigen. Der große Subgraph des trimodalen Netzwerkes gibt mit einem Wert von 5,55 hierzu noch keinen Aufschluss, da sich darin die durchschnittliche Anzahl von Autoren je Institut/Fakultät und die durchschnittliche Anzahl von Autoren je Publikation bzw. Publikationen je Autor vermischen. Aus dem unimodalen Netzwerk der Institute/Fakultäten geht nun aber hervor, dass ein Institut/eine Fakultät im Netzwerk durchschnittlich mit rund 8 (exakt: 8,28) Instituten/Fakultäten Ko-Publikationsbeziehungen aufweist.

Die *durchschnittliche Pfadlänge (average path length)* von 5,32 in dem großen Subgraphen des trimodalen Netzwerkes liegt deutlich über dem theoretischen Mindestwert in trimodalen Netzwerken von 2. Dies deutet auf die Existenz zentraler Knoten in der Gruppe der Autoren hin, welche entfernte Disziplinen miteinander verbinden. Durch das unimodale Netzwerk des Subgraphen wird diese Vermutung bestätigt. Die dafür errechnete durchschnittliche Pfadlänge von 2,06 zeigt, dass die Institute/Fakultäten in der Mehrzahl nicht direkt, sondern über ein anderes Institut/eine andere Fakultät mittelbar verbunden sind. Mit Hilfe dieser Metriken lassen sich wichtige verbindende Institute/Fakultäten sehr leicht in der interaktiven Netzwerkdarstellung identifizieren. So haben beispielsweise die Institute der Fakultät für Maschinenbau und der Fakultät für Elektrotechnik eine erkennbar weit gestreute Verteilung innerhalb des Netzwerkes.

⁴ Wenn g die Anzahl der Knoten in einem ungerichteten unimodalen Netzwerk bezeichnet, so lässt sich die maximale Anzahl der Kanten, welche zur Normalisierung der tatsächlich im Netzwerk vorhandenen Kanten verwendet wird, mit $g(g-1)/2$ berechnen. In dem trimodalen Netzwerk sind jedoch nur Kanten zwischen jeweils zwei verschiedenen Knotentypen möglich, sodass gemäß Definition der Knotenmengen in Abschnitt 4.2 die maximale Kantenzahl durch $(|I| \times |A|) + (|A| \times |P|)$ gegeben ist.

⁵ Verfügbar unter: <https://marketplace.gephi.org/plugin/multimode-networks-transformations-2/>

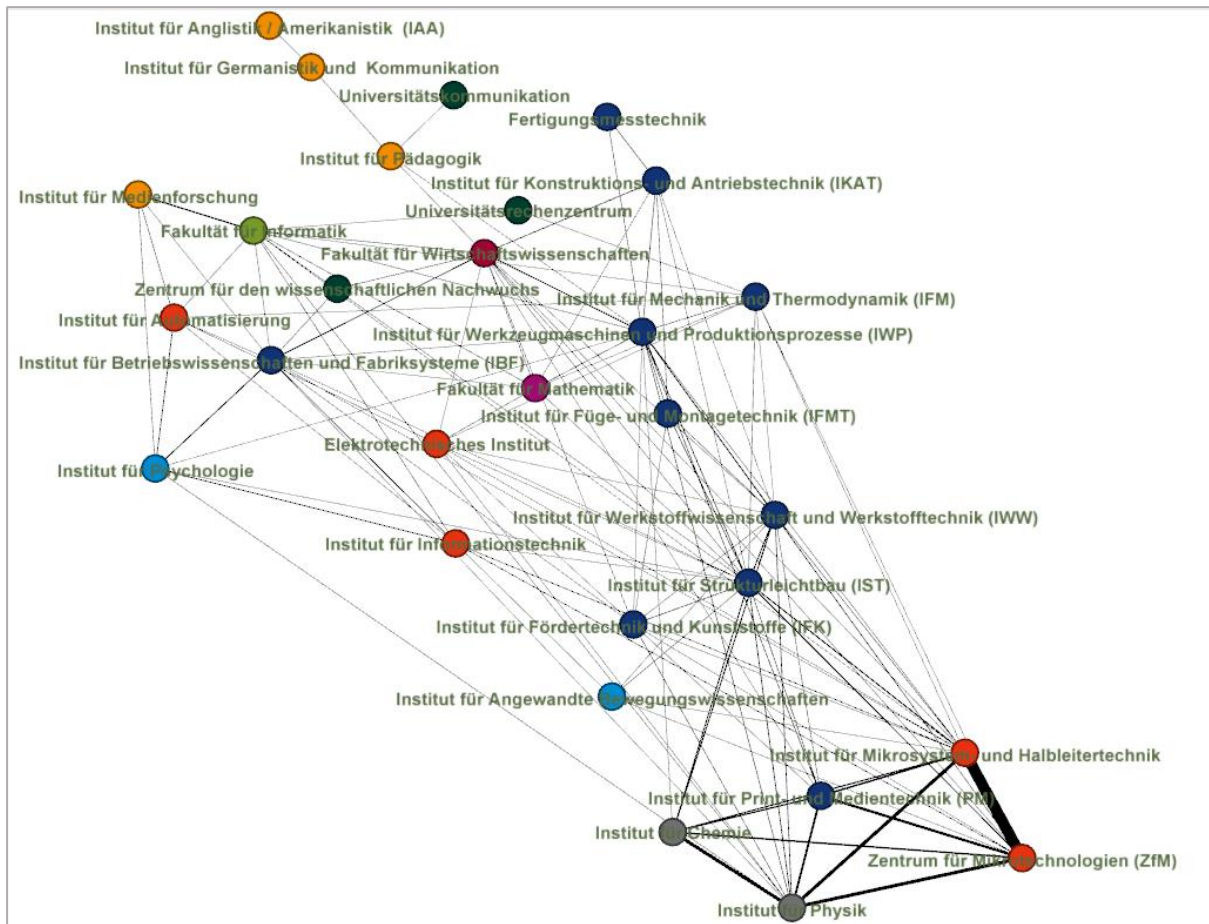


Abbildung 1 Unimodales Ko-Publikationsnetzwerk der Institute/Fakultäten

Durch die zahlreichen Vermittler-Knoten ist der große Subgraph des trimodalen Netzwerkes auch insgesamt sehr breit, was sich in einem *Graphumfang (diameter)* von 19 widerspiegelt. Die entferntesten Knoten des Netzwerkes sind also über bis zu 19 Kanten miteinander verbunden.

Auf Grund der oben geschilderten Berechnungs- und Interpretationsprobleme wurden aus den graphentheoretischen Konzepten für einzelne Knoten nur der nicht normalisierte *Grad (degree)* der Knoten ausgewertet. In der interaktiven Web-Darstellung wird der berechnete Grad eines Knotens beim Anklicken desselben in einem Seitenfenster angezeigt. So lässt sich rasch erkennen, wie viele Autoren zu einem Institut/einer Fakultät gehören, an wie vielen Publikationen ein Autor im Netzwerk mitgewirkt hat und wie viele Co-Autoren an einer Publikation beteiligt sind. Um die Konzepte der Zentralität zu verdeutlichen, wurden mit Gephi diese Werte für die Institute/Fakultäten in der unimodalen Projektion des großen trimodalen Subgraphen berechnet (siehe Tab. 1).

Tabelle 1 Beispiel

Institut/Fakultät	Fakultät	Degree Centrality	Closeness Centrality	Betweenness Centrality
Strukturleichtbau	Maschinenbau	0,68	0,68	0,09
Werkzeugmaschinen und Produktionsprozesse	Maschinenbau	0,64	0,72	0,24
Mikrosystem- und Halbleitertechnik	Elektrotechnik/ Informationstechnik	0,54	0,62	0,04
Wirtschaftswissenschaften	Wirtschaftswissenschaften	0,50	0,65	0,14

Werkstoffwissenschaft und Werkstofftechnik	Maschinenbau	0,50	0,60	0,03
Betriebswissenschaften und Fabriksysteme	Maschinenbau	0,46	0,60	0,07
Informatik	Informatik	0,43	0,58	0,08
Physik	Naturwissenschaften	0,43	0,58	0,02
Zentrum für Mikrotechnologien	Elektrotechnik/ Informationstechnik	0,43	0,57	0,02
Fördertechnik und Kunststoffe	Maschinenbau	0,39	0,56	0,01
Mechanik und Thermodynamik	Maschinenbau	0,36	0,55	0,04
Elektrotechnisches Institut	Elektrotechnik/ Informationstechnik	0,32	0,54	0,01
Print- und Medientechnik	Maschinenbau	0,32	0,52	0,00
Mathematik	Mathematik	0,29	0,54	0,01
Informationstechnik	Elektrotechnik/ Informationstechnik	0,29	0,49	0,01
Psychologie	Human- und Sozialwissenschaften	0,25	0,52	0,02
Konstruktions- und Antriebstechnik	Maschinenbau	0,25	0,50	0,01
Chemie	Naturwissenschaften	0,25	0,49	0,00
Automatisierung	Elektrotechnik/ Informationstechnik	0,25	0,45	0,01
Füge- und Montagetechnik	Maschinenbau	0,18	0,47	0,00
Angewandte Bewegungswissenschaften	Human- und Sozialwissenschaften	0,18	0,44	0,00
Pädagogik	Philosophische Fakultät	0,14	0,49	0,20
Medienforschung	Philosophische Fakultät	0,14	0,41	0,00
Zentrum für den wissenschaftlichen Nachwuchs	Zentrale Einrichtungen	0,07	0,44	0,00
Fertigungsmesstechnik	Maschinenbau	0,07	0,43	0,00
Universitätsrechenzentrum	Zentrale Einrichtungen	0,07	0,39	0,00
Germanistik und Kommunikation	Philosophische Fakultät	0,07	0,34	0,07
Universitätskommunikation	Zentrale Einrichtungen	0,04	0,33	0,00
Anglistik / Amerikanistik	Philosophische Fakultät	0,04	0,26	0,00

Die *Grad-Zentralität* (*degree centrality*) als aktivitätsbezogenes Maß spiegelt zuvorderst die Publikationskulturen der Fächer wider: die zahlreichen Institute aus dem technisch-naturwissenschaftlichen Bereich mit ihren hochfrequenten Ko-Publikationsgewohnheiten weisen erwartungsgemäß die höchsten Werte auf.

Die *Nähe-Zentralität* (*closeness-centrality*), als Kehrwert der durchschnittlichen Pfadlänge eines Knotens, ist ein Maß für die Eingebundenheit eines Instituts/einer Fakultät in Relation zur perfekten Einbindung (direkte Verbindung zu jedem anderen Knoten). Hier werden beispielsweise die sehr zentralen Positionen des Instituts für Strukturleichtbau (IST) und des Instituts für Werkzeugmaschinen und Produktionsprozesse (IWP) deutlich, deren Autoren jeweils mit Autoren aus mehr als der Hälfte der anderen Institute/Fakultäten direkt ko-publiziert haben.

Letztlich lassen sich mit Hilfe der *Betweenness-Zentralität* Knoten identifizieren, deren Wegfall die Desintegration des Netzwerkes in zwei oder mehrere disziplinäre Subgraphen zur Folge hätte. Insbesondere das Institut für Werkzeugmaschinen und Produktionsprozesse, das Institut für Pädagogik und die Fakultät für Wirtschaftswissenschaften treten hierbei als interdisziplinäre Brückenbildner hervor.

5 Resümee

Mit dem Publikationsnetzwerk wurde das Ziel verfolgt, einen neuartigen Einblick in das komplexe Phänomen der Interdisziplinarität der Forschung an der Technischen Universität Chemnitz zu geben. Durch die multimodale Struktur des Netzwerkes und dessen interaktive Darstellung wird interdisziplinäre Zusammenarbeit sowohl in ihrem fachlichen Aspekt, dargestellt durch die Institute/Fakultäten, als auch in ihrer sozialen Ausprägung, in Form der ko-publizierenden Autoren, und in ihren konkreten Ergebnissen, in Gestalt der Publikationen, greifbar. Diese reichen visuellen Erkenntnismöglichkeiten rechtfertigen auch die eingeschränkte Anwendbarkeit graphentheoretischer Analysemethoden, die mit einer multimodalen Netzwerkstruktur einhergeht.

Bereits in seiner jetzigen Form können Akteure aus der Wissenschaft und dem Wissenschaftsmanagement mit Hilfe des bibliometrischen Netzwerkes ihr Wissen über das Interaktionsgeschehen in den wissenschaftlichen Communities und über die soziale Stellung ihrer Akteure vermehren.

Es konnte bereits an anderen Publikationsnetzwerken empirisch aufgezeigt werden, dass dieses Wissen innerhalb der so untersuchten wissenschaftlichen Gemeinschaften zu einer erhöhten Aufmerksamkeit für bestehende Kooperationen führen kann (Valdez et al. 2012). Die geschärfte Selbstwahrnehmung hat wiederum eine Stärkung von Teamidentitäten zur Folge (ebd.). Auf diesem Wege werden Netzwerkanalysen zu einem effektiven Mittel der internen Erfolgsdokumentation und -steuerung, das Forschenden aufzeigt, wo sie starke Bindungen aufgebaut haben und wo bisher ungenutzte Potenziale für Zusammenarbeit liegen.

Zudem finden Sie in dem Netzwerk, gerade durch die Einbeziehung der Publikationen, ein Instrument zur Untersuchung von Themenclustern und -trends in ihrem jeweiligen Fachgebiet, an denen sie beispielsweise künftige Publikationstätigkeiten ausrichten können.

Nicht zuletzt dient die bibliometrische Netzwerkanalyse dem Wissenschaftsmanagement und der Öffentlichkeitsarbeit, indem sie die Identifikation von zentralen Akteuren und wichtigen Schnittstellen innerhalb von Kooperationsnetzwerken und die anschauliche Kommunikation von Kooperationsleistungen ermöglicht.

Um den Informationsgehalt der Netzwerkanalyse im Hinblick auf interdisziplinäre Forschungsergebnisse von Angehörigen der TU Chemnitz noch weiter anzuheben, ließe sich die Betrachtung mit Hilfe von Publikations- und Zitationsdatenbanken auch auf externe, vor allem internationale Kooperationen ausdehnen. Ebenso eine Anreicherung der Daten mit Zitationen, Publikationsorten und Daten zu Forschungsthemen aus dem Textmining von Abstracts, Keywords oder Volltexten kann aus bibliometrischer Sicht neue Erkenntnisse hervorbringen. Zudem ließen sich durch die Einbeziehung soziometrischer Daten über die kooperierenden Autoren – z.B. Alter, Sprache, Teamzugehörigkeit – weitere Einblicke über die sozialen Bedingungen und Ausprägungen interdisziplinärer Zusammenarbeit gewinnen. Bei all diesen Erweiterungsmöglichkeiten sind die potenziellen Informationszugewinne sorgsam mit den erhöhten methodischen, technischen und datenschutzrechtlichen Anforderungen an die Gewinnung und Aufbereitung der Daten abzuwägen.

Literaturverzeichnis

Borgatti, Stephen P.; Halgin, Daniel S. (2011): Analyzing affiliation networks. In: John Scott und Peter J. Carrington (Hg.): *The SAGE handbook of social network analysis*: SAGE publications, S. 417–433.

Calero Valdez, André; Schaar, Anne Kathrin; Ziefle, Martina; Holzinger, Andreas; Jeschke, Sabina; Brecher, Christian (2012): Using Mixed Node Publication Network Graphs for Analyzing Success in Interdisciplinary Teams. In: Runhe Huang et al. (Hg.): *Active media technology. 8th international conference, AMT 2012, Macau, China, December 4 - 7, 2012; Proceedings*, Bd. 7669. Berlin: Springer, S. 606–617.

Card, Stuart K.; Mackinlay, Jock D.; Shneiderman, Ben (1999): *Readings in information visualization: using vision to think*. San Francisco: Kaufmann.

Freeman, Linton C. (1978): Centrality in social networks conceptual clarification. In: *Social networks* 1 (3), S. 215–239.

Havemann, Frank (2009): *Einführung in die Bibliometrie*. Berlin: Gesellschaft für Wissenschaftsforschung.

Jacomy, Mathieu; Venturini, Tommaso; Heymann, Sebastien; Bastian, Mathieu (2014): ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. In: *PLoS one* 9 (6), e98679. DOI: 10.1371/journal.pone.0098679.

Schnegg, Michael; Lang, Hartmut (2002): Netzwerkanalyse. Eine praxisorientierte Einführung. In: *Methoden der Ethnographie* 1, S. 1–55. Online verfügbar unter <http://www.methoden-der-ethnographie.de/heft1/heft1.html>, zuletzt geprüft am 07.10.2016.

Small, Henry (1999): Visualizing science by citation mapping. In: *Journal of the American Society for Information Science* 50 (9), S. 799–813. DOI: 10.1002/(sici)1097-4571(1999)50:9<799::aid-asi9>3.3.co;2-G.

Wasserman, Stanley; Faust, Katherine (1994): *Social network analysis: Methods and applications*. Cambridge: Cambridge university press.