

Leuchter, S., Schönbein, R. & Urbas, L. (2005). Skalenfreie Netzwerke und Benutzermodellierung. In A.B. Cremers, R. Manthey, P. Martini, & V. Steinhage (Hrsg.), *Informatik 2005. Informatik Live. Beiträge der 35. Jahrestagung der Gesellschaft für Informatik, 19.-22.09. 2005 in Bonn* (S. 271-275, Band 1). Bonn: Gesellschaft für Informatik (LNI; P-67). <http://www.safety-critical.de/doc/info2005SIP.pdf>

Skalenfreie Netzwerke und Benutzermodellierung

Sandro Leuchter*, Rainer Schönbein* & Leon Urbas**

* Fraunhofer IITB
Abteilung Interoperabilität und Assistenz
Fraunhoferstr. 1
76131 Karlsruhe
leuchter@iitb.fraunhofer.de
schoenbein@iitb.fraunhofer.de

** Technische Universität Berlin
Zentrum Mensch-Maschine-Systeme
Jebensstr. 1
10623 Berlin
leon.urbas@zmms.tu-berlin.de

Abstract: In vielen natürlichen Netzwerken kann die Herausbildung eines immer wiederkehrenden Zusammenhangs in der Häufung von Verknüpfungen zwischen Knoten beobachtet werden. Diese sog. *Skalenfreiheit* tritt insbesondere auch in sozialen Netzwerken auf, die die Grundlage vieler Benutzermodellierungsalgorithmen und Wissensmanagementansätze bilden. In diesem Beitrag wird das Phänomen der skalenfreien Netzwerke beschrieben und es werden die Auswirkungen auf Benutzermodellierungsalgorithmen und Wissensmanagementansätze anhand eines Wissensmanagementsystems (ExperOnto) sowie eines benutzeradaptiven Fachinformationsportals (useworld.net) erörtert.

1 Skalenfreie Netzwerke

Viele Zusammenhänge in natürlichen Systemen können als Netzwerke mit Hilfe von Graphen beschrieben werden. Die Knoten repräsentieren die Elemente des Systems, die Kanten die Relationen bzw. Interaktionen zwischen ihnen. Je nach untersuchtem System kann die Topologie der Netzwerke komplex sein. Um dennoch quantifizierbare Aussagen über die so repräsentierten Systeme machen zu können, werden die Netzwerke mit graphentheoretischen Konstrukten beschrieben.

Eine Systemeigenschaft, die für die Analyse nützlich ist, ist das Ausmaß an innerer Interaktion. Sie wird beschrieben durch die Wahrscheinlichkeit, dass ein Element mit k anderen Systemelementen interagiert, bzw. dass ein Knoten eine Relation mit k anderen hat. Barabási und Albert [BA99] berichten, dass diese Wahrscheinlichkeit in vielen natürlichen Systemen durch ein Potenzgesetz beschrieben werden kann: $P(k) \sim k^{-\gamma}$. Für große Netzwerke ergibt sich ein sog. skalenfreier Zustand.

Skalenfreie Netzwerke ergeben sich in unterschiedlichen Netzwerken, die interpretiert werden als Folge sozialer Interaktionen [BA99]:

- Die *Zusammenarbeit von Filmschauspielern* (Datenbasis wird nicht berichtet) folgt dem Potenzgesetz mit $\gamma = 2,3$. k wird interpretiert als Maß für Popularität: Schauspieler, die mit vielen anderen in Filmen gespielt haben, sind berühmter als die, die mit wenigen Kollegen zusammengearbeitet haben.
- Die *Verlinkung im WWW* (Daten von 1999) folgt dem Potenzgesetz mit $\gamma = 2,1$. Knoten sind Dokumente im WWW, Relationen Hyperlinks zu anderen.
- Die *Verknüpfungen im Stromnetz* im Westen der USA zwischen Generatoren, Transformatoren und Unterstationen durch Leitungen (Datenbasis 1998) folgt dem Potenzgesetz mit $\gamma \approx 4$.
- Die Zitierwahrscheinlichkeit in wissenschaftlichen Veröffentlichungen folgt dem Potenzgesetz mit $\gamma = 3$.

Das Erklärungsmodell für das Zustandekommen von so verknüpften Netzwerken ist, dass neue Knoten sich mit höherer Wahrscheinlichkeit mit Knoten verbinden, die viele Interaktionen haben, weil die Kanten die Eigenschaften repräsentieren, die den Knoten näher beschreiben. Je mehr Informationen über einen Knoten vorhanden sind, desto größer wird die Wahrscheinlichkeit, dass neue Knoten mit ihm interagieren. Das ist insbesondere bei sozialen Netzwerken ersichtlich, wenn Kanten als Popularität oder öffentliche Sichtbarkeit interpretiert werden können.

2 Auswirkung auf Benutzermodellierung

Benutzermodelle erlauben die Personalisierung von eCommerce- und Online-Informationsangeboten. Durch die serverbasierte Verfügbarkeit von Informationsdiensten, e-Learning-Angeboten und Anwendungen zur netzbasierten Zusammenarbeit ist die Möglichkeit gegeben, Benutzermodelle zu erzeugen und zu einander in Beziehung zu setzen. Folgende Beziehungen zwischen individuellen Benutzermodellen bzw. Benutzern sind möglich:

- Die Ähnlichkeit zwischen zwei Benutzern kann aus dem Vergleich der Eigenschaften des Benutzermodells geschlossen werden. Die Aussagen über die Ähnlichkeit bezieht sich dann auf die im Benutzermodell repräsentierten Informationen.
- Benutzer können durch Bewertungen anderer Benutzer ihre Wertschätzung oder ihr Vertrauen ausdrücken. Ggf. kann die Bewertung auch indirekt ermittelt werden, wenn nicht andere Benutzer, sondern deren Beiträge bewertet werden.
- Vergangene Zusammenarbeit oder andere Formen von Interaktion zwischen zwei Benutzern kann durch die Anwendung in einer Art episodischem Gedächtnis gespeichert werden.

Diese Beziehungen zwischen Benutzern werden in unterschiedlichen Klassen von Anwendungen benutzt. Insbesondere in Wissensmanagementsystemen und in *Recommender Systemen* kommen sie zum Einsatz. In aktuellen Implementierungen wird jedoch von einer Gleichverteilung der Relationen zwischen Benutzern anstatt von skalenfreien Benutzernetzwerken ausgegangen.

Eine wichtige Anwendung im Wissensmanagement ist die möglichst passgenaue multikriterielle Vermittlung von kompetenten Experten. Hier wird insbesondere die Ähnlichkeit von Benutzer- bzw. Anfrageprofilen adressiert [RAN03]. Da die Ähnlichkeit zu einem Referenzprofil festgestellt wird und nicht auf sozialen Interaktionen beruht, ist hier i.A. kein Netzwerkskalierungseffekt zu erwarten.

Auf eine Bewertung von anderen Benutzern wird in vielen Organisationen teils aus arbeitsrechtlichen, teils aus sozialpsychologischen Gründen verzichtet.

In vielen Fällen wird jedoch das Kompetenzprofil eines Benutzers aus dem episodischen Gedächtnis des Systems erzeugt. Bei gut vernetzten Experten werden besonders umfangreiche Kompetenzprofile die Folge sein. Ob diese gut vernetzten Experten aber die sind, die auch gefunden werden sollen, hängt von den Vorgaben der Organisation ab. Sie sind zwar mit hoher Wahrscheinlichkeit schon ausgelastet, können aber auf der anderen Seite wertvolle Hilfestellung leisten.

Für die Umsetzung von *Collaborative Filtering* [GN92] und *Recommender Systemen* [RV97] sind hingegen besonders die Ähnlichkeit von Benutzern und die Bewertung durch andere Benutzer relevant. Liegen über einige Benutzer weniger Informationen als über andere vor, bzw. werden Benutzer seltener bewertet, dann ist bei diesen der jeweilige Mittelwert unsicherer, denn die Varianz ist tendenziell größer.

Zusätzlich wirken Benutzer, über die viele Informationen vorliegen als „Informationshubs“. Die Wahrscheinlichkeit steigt, dass sie zu mehr anderen Benutzern ähnlich sind. Für den Fall der Ähnlichkeitsberechnung muss also der Einfluss von diesen „Power-Usern“ begrenzt werden. Das kann man am einfachsten über einen Faktor realisieren, der umso kleiner wird, je mehr Relationen im sozialen Netzwerk der betrachtete Benutzer hat.

3 Anwendungen

3.1 useworld.net

useworld.net ist eine web-basierte Informations- und Kooperationsplattform für die *scientific community* der Mensch-Maschine-Interaktionsforschung [LUR03]. Dieser Bereich ist stark interdisziplinär geprägt, so dass bei der Entwicklung mit einer sehr heterogenen Benutzergruppe zu rechnen war. Deshalb wurde Benutzermodellierung vorgesehen, um eine an der nutzerspezifischen Relevanz orientierte Adaption der Ergebnispräsentation zu implementieren.

useworld.net beinhaltet einen nutzergepflegten Katalog mit sog. InfoObjekten (Links, Dokumente, Konferenzankündigungen, Verweisen auf *special interest groups*, Stellenausschreibungen, etc.) und eine Suchmaschine. Diese beiden Komponenten können Benutzer anonym verwenden. Registrierte Benutzer haben aber zusätzlich noch Zugriff auf einen *shared workspace* basierten Kooperationsteil für geschlossene Nutzergruppen und einen persönlichen Informationsagenten, der sie per E-Mail oder Web-Seite über für sie relevante Änderungen im Portal informiert. Die Interaktion von angemeldeten Benutzern mit useworld.net wird zur Benutzermodellierung herangezogen. Insbesondere ist es auch wie bei einem *Recommender* Dienst möglich, InfoObjekte (und damit indirekt deren Besitzer) zu bewerten. Über die bewerteten Objekte und die Teile des Kataloges, in denen sie sich befinden, wird ein Interessenprofil abgeleitet. Die Überlappung von Interessenprofilen wird zur Berechnung der Ähnlichkeit von Benutzern herangezogen, die für die Bewertung nicht explizit bewerteter InfoObjekte benutzt wird.

In useworld.net werden Relationen im sozialen Netzwerk durch veröffentlichte und bewertete InfoObjekte erzeugt. Eine Analyse des Datenbestandes des seit zwei Jahren betriebenen Web-Angebotes zeigt die Charakteristik eines skalenfreien Netzwerkes: Es gibt einige wenige „Power-User“ die den größten Teil der Inhalte erstellt und den Großteil der verfügbaren Bewertungen beigesteuert haben. Das Problem an dieser Netztopologie ist, das über die Mehrzahl der Benutzer nicht genügend Informationen vorliegen, um eine Ähnlichkeit mit den Power-Usern zu berechnen. Sieht man davon ab, ist absehbar, dass die Power-User das Bewertungsbild verzerren, denn sie sind eher multidisziplinär orientiert.

3.2 ExperOnto

ExperOnto ist ein Wissensmanagementsystem, mit dem Kompetenzträger in Organisationen web-basiert recherchiert werden können, um geeignete Projektteams zusammenstellen zu können. Um einen Einsatz auch in großen, verteilten und heterogenen Organisationen zu ermöglichen, wird die in Teilorganisationen unterschiedlich ausgeprägte Terminologie durch Ontologien aufeinander abgebildet.

Die Semantik der Begriffe der verwendeten Terminologien ist der Schlüssel zu gezielter Informationsrecherche. Sie ermöglicht die intelligente Vermittlung von Informationen. Dazu wird eine geeignete Modellierung der Akteure, Situationen, Abläufe und Fakten benutzt. Kriterien für die Auswahl von Projektteammitgliedern sind: Mitarbeit an Projekten, eingesetzte Werkzeuge, Methodiken, Technologien und Modelle, die räumliche und zeitliche Verfügbarkeit, die Zugehörigkeit zu und der Rang in einer Organisationseinheit. Die ExperOnto-Ontologie modelliert dazu folgende Konzepte: (Arbeits-) Methodiken, Modelle, Werkzeuge, Projekte, Raum, Zeit, Verfügbarkeit (Konzept Kalender, Arbeitszeit etc.), Organisation, Person (Experte), Kontaktmöglichkeiten.

Die Konzepte sind insbesondere über die modellierte Person verknüpft. Die modellierten Konzepte bedingen, dass eine Art episodisches Gedächtnis herangezogen wird, um Kompetenzen zu repräsentieren.

Da das System mit Daten aus unterschiedlichen Teilorganisationen gespeist wird, ist mit mehreren Teilnetzwerken zu rechnen, die untereinander wenig verknüpft, untereinander jedoch stark vernetzt sind.

Die Auslastung der Wissensträger wird über die Modellierung der Verfügbarkeit mit herangezogen. Dadurch wird sich eine Nivellierung der Power-User, die über sehr viele Projekte mit anderen verknüpft sind, ergeben. Diese Nivellierung ist aber nur dann stabil, wenn der teilorganisationsspezifische Umgang mit dem Projektbegriff vereinheitlicht werden kann: Die Projektdauer in einer Organisationseinheit könnte z.B. grundsätzlich viel kürzer sein als in anderen. In diesem Fall hätten die Mitglieder der Organisationseinheit die Möglichkeit, in vielfältigeren Arbeitszusammenhängen repräsentiert und damit häufiger von anderen gefunden zu werden als andere Experten. An dieser Stelle müsste die Ontologie die unterschiedliche Pragmatik im Umgang mit dem Projektbegriff repräsentieren.

4 Zusammenfassung und Ausblick

Das Phänomen skalenfreier sozialer Netzwerke wirkt sich auf Methoden der Benutzermodellierung aus. Im Einzelfall ist anhand der Anwendung zu prüfen, welche Topologie beim resultierenden Netzwerk zu erwarten ist. Anschließend muss abgeschätzt werden, welche Konsequenzen sie auf die Anwendung haben kann. Gegebenfalls müssen die Algorithmen angepasst werden, um die Topologie des Netzwerkes korrekt widerzuspiegeln. Für useworld.net ist geplant, mehr Benutzer zu aktivieren, so dass weitere Daten gesammelt werden können, damit γ im Potenzgesetz größer wird. Das würde es erlauben, weitere Analysen über diesen Benutzermodellierungsansatz durchzuführen. Für ExperOnto werden Möglichkeiten untersucht werden, neben der Bedeutung der Konzepte ihren Gebrauch in der Ontologie oder im Matching-Algorithmus abzubilden. Die Sortierung der Ergebnisse von ExperOnto wird neben der Passung der Kriterien der Anfrage durch die Priorisierung des Anfragers bzw. seines Projektes beeinflusst. Das geht bis hin zu einem zentralen Scheduling der aus Anfragen resultierenden Aufträge für einen Experten.

Literaturverzeichnis

- [BA99] Barabási, A.-L.; Albert, R.: Emergence of Scaling in Random Networks. *Science*, 286, S. 509-512; 1999.
- [GN+92] Goldberg, D.; Nichols, D.; Oki, B.M.; Terry, D.: Using collaborative filtering to weave an information tapestry. *Communications of the ACM*, 35, S. (12) 61-70; 1992.
- [LUR03] Leuchter, S.; Urbas, L.; Röse, K.: Engineering and Evaluation of Community Support in useworld.net. In: *Proceedings of HCI International*, Vol. 2. Lawrence Erlbaum Publishers, Mahwah, NJ, 2003; S. 959-963.
- [RAN03] Razmerita, L.; Angehrn A.; Nabeth, T.: On the role of user models and user modeling in Knowledge Management Systems, In: *Proceedings of HCI International*, Vol. 2. Lawrence Erlbaum Publishers, Mahwah, NJ, 2003; S. 450-456.
- [RV97] Resnick, P.; Varian, H.: Recommender Systems. *Communications of the ACM*, 40, S. (3) 56-58; 1997.