

Forschungsmodul: Komplexe Systeme

Bericht zur Vorlesung vom 29. November 2007 von Jan-Philip Gehrcke

Abstraktion komplexer Systeme

In den Naturwissenschaften ist es oftmals so, dass sehr große Systeme nur durch Betrachtung von Erwartungswerten bestimmter Größen ausreichend beschrieben werden können. Paradebeispiel ist die statistische Mechanik in der Physik. Eine detaillierte Betrachtung der einzelnen Bestandteile ($\sim 10^{24}$), um makroskopische Ergebnisse zu bekommen, ist hier nahezu unmöglich. Die Resultate kommen durch Mittelung zustande und ohne wirklich zu wissen, was innerhalb des Systems exakt passiert und wie die Einzelteile wechselwirken.

Nun ist es aber so, dass es auch Systeme gibt, deren Verhalten nur unter genauer Berücksichtigung der inneren Wechselwirkung zu bestimmen und vorherzusagen ist. Diese Systeme gehören zur Klasse der komplexen Systeme. Auch hier gibt es ein besonders prominentes Beispiel: Innerhalb der Neurowissenschaften kann man das Verhalten z.B. des Menschen sicherlich nicht durch das Mitteln über irgendwelche Zustände berechnen. In biologischen Organismen ist die Charakteristik der Verschaltung der Elemente - also deren Wechselwirkung untereinander - und nicht die Anzahl der Bauteile entscheidend.

Denkbar sind viele mögliche Systeme bzw. Netzwerke. Die soziale und wirtschaftliche Interaktion zwischen Menschen, Gruppierungen, Unternehmen, Organisationen und auch der Aufbau des Internets ist ein komplexes System. Will man ein komplexes Netzwerk abstrahieren, muss man es auf seine Topologie beschränken. Dabei kann man sich verschiedene Typen von Netzen vorstellen. Die einzelnen Knoten (oder auch Agenten) können verschieden gewichtete Verbindungen (Kanten) in verschiedene Richtungen aufweisen. Klassifizieren kann man zum Beispiel die Menge der Graphen $G(N, M)$ mit N Knoten und M Kanten, was dem mikroskopischen Ensemble in der statistischen Mechanik entsprechen könnte oder die Menge der Graphen $G(N, p)$ mit N Knoten und der Wahrscheinlichkeit p , dass zwischen zwei Agenten eine Kante existiert.

Letzteres entspräche dem kanonischen Ensemble. Diesen letzteren Ansatz wählten auch Paul Erdős und Alfréd Rényi 1960 in ihrem Modell zur Beschreibung komplexer Systeme. Im Erdős-Rényi Modell (welches nur aus Knoten und Kanten besteht) sind alle Kopplungen/Kanten - entsprechend der klassischen Thermodynamik und dem Gibbs'schen Postulat - gleich wahrscheinlich.

Für sehr große Netzwerke (also im Limes $N \rightarrow \infty$) soll nun untersucht werden, ob bestimmte Eigenschaften des Graphen G in Abhängigkeit der Wahrscheinlichkeit p existieren, welche ihrerseits von N abhängen darf und soll. Die Frage ist also ab welcher Wahrscheinlichkeit $p(N)$ eine bestimmte Eigenschaft in einem Graphen G sicher auftritt. Diese kritische Wahrscheinlichkeit sei p_c .

Als Beispiel sei ein Graph $G(N, p)$ gewählt. Es soll untersucht werden, wieviele Subgraphen F mit K Knoten und l Kanten dieser besitzt. Der Erwartungswert $\langle Y \rangle$ für die Anzahl der Subgraphen berechnet sich dann wie folgt:

$$\langle Y \rangle = \binom{N}{K} \frac{k!}{a} p^l \approx \frac{N^k p^l}{a}$$

Der Binomialkoeffizient gibt die Anzahl der Möglichkeiten K Knoten aus N zu wählen. Diese können verschieden permutiert sein (Faktor $k!$). Gleichwertige Subgraphen (Isomorphismen) werden über den Quotienten a kompensiert. Dass l Kanten im Subgraphen existieren, ist durch die Wahrscheinlichkeit p^l gegeben. Nimmt man an, dass $p(N) = cN^{-K/l}$ ist, ergibt sich:

$$\langle Y \rangle = \frac{c^l}{a} = \lambda$$

Die Anzahl der Subgraphen mit einer gewählten Eigenschaft ist für $N \rightarrow \infty$ poissonverteilt. Unter dieser Bedingung lässt sich zeigen, dass bei angenommener Wahrscheinlichkeitsverteilung $p(N)$ mindestens ein Subgraph F in G liegt. Daher ist $p(N) = p_c(N)$.

Mit gleicher Betrachtungsweise lässt sich bestimmen, für welche z bei $p(N) = N^z$ welche Typen von Subgraphen erstmalig auftreten. Ab $z = -2$ ergibt sich die erste Kante überhaupt. Erhöht man z , ergeben sich immer stärker vernetzte Subgraphen. Ab $z = -1/2$ kommen die ersten voll vernetzten Subgraphen zustande, in denen jeder Knoten mit jedem verbunden ist.

Immer noch im Limes $N \rightarrow \infty$ lässt sich die "Small World"-Eigenschaft nachweisen. Der der größte Abstand D von einem Knoten zu einem beliebigen anderen innerhalb des Systems, auch Durchmesser genannt, skaliert mit dem Logarithmus der Systemgröße:

$$D \approx \log(N)$$

So ist es nicht verwunderlich, dass alle Menschen auf der Erde über ca. sechs "Hops" miteinander verbunden sind, wie in den 60er Jahren gezeigt wurde.

Beruhend auf der Annahme der Poissonverteilung der mittleren Kantenzahl k pro Knoten lässt sich eine Betrachtung durchführen, wieviele mögliche Optionen Z_m es für den "Endpunkt" nach m Schritten ausgehend von einem Startpunkt im Netzwerk gibt. Diese Betrachtung führt auf die Rekursion

$$\langle Z_{m+1} \rangle = \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle} Z_m ; \quad Z_1 = \langle k \rangle .$$

Für ein nach dem Erdős-Rényi Modell abstrahierten Netzwerk lassen sich - wie man sieht - eine Reihe von Eigenschaften bestimmen. Dazu gehören desweiteren z.B. die mittlere kürzeste Weglänge oder die größte Zusammenhangskomponente.

Für reale Netzwerke ist die Annahme, dass alles gleichwahrscheinlich ist, jedoch falsch. Am Beispiel des Internets ist eindrucksvoll zu sehen, dass es ganz zentrale Knoten gibt, an denen sich die Kanten anhäufen, weil sie nahezu jeden Agenten bedienen (müssen). Die Verteilung der Kanten folgt in realen Netzwerken, die also nicht zufällig aufgebaut sind, eher Potenzgesetzen. In logarithmischer Auftragung ist der Unterschied zur Poissonverteilung sehr deutlich zu sehen.

Real tauchen nicht wenige Knoten mit vergleichsweise riesiger Kantenzahl auf - ein Fall, den die Poissonverteilung eigentlich ausschließt. Diese realen Systeme verhalten sich allesamt recht ähnlich. So wurde gezeigt, dass man den Aufbau des Internets mit der Zusammensetzung von Schauspielergruppierungen in Filmen und den Zitationsverbindungen in wissenschaftlichen Publikationen vergleichen kann (überall liegen Potenzgesetze vor).

Die wichtigste Schlussfolgerung ist, dass reale Systeme - egal welcher Art - zwar komplex sind, aber alles andere als zufällig aufgebaut.