



# Gut vernetzt

Wie sich Netzwerke  
besser verstehen und  
erzeugen lassen

*von Dirk Eidemüller*



Das Gebiet der synthetischen Netzwerke boomt. Mithilfe solcher im Computer simulierten Netze werden heute so unterschiedliche Dinge wie die Verknüpfung der Neuronen im Gehirn, der Datenverkehr im Internet oder Stromnetze untersucht. Ein Forscherteam um Ulrich Meyer vom Institut für Informatik der Goethe-Universität hat nun Standardverfahren zur Erstellung solcher Netze einen wichtigen Schritt vorangebracht.

Unser ganzes Leben ist eingebettet in die unterschiedlichsten Arten von Netzwerken. Dazu zählen nicht nur die sozialen Beziehungen, die heute über die sozialen Medien von Algorithmen mitbestimmt werden. Die Produkte, die wir für unseren Lebensunterhalt kaufen, benötigen ein komplexes Geflecht aus Stoffströmen und Logistik, bevor sie im Regal landen. Auch der elektrische Strom aus unserer Steckdose hängt an einem komplexen Netz, das mit dem Ausbau erneuerbarer Energien laufend erweitert und umgestaltet wird.

Je komplexer Netzwerke sind, desto schwerer lässt sich vorhersagen, wie sich Informationen verbreiten oder wie sich Schwankungen oder Störungen auf den Fluss von Waren oder Strom auswirken. Weil die Grundstruktur aller Netze jedoch gleich ist – sie bestehen aus Elementen (Knoten), die über Verbindungen (Kanten) miteinander verknüpft sind –, können alle Netzwerke in abstrakter Weise dargestellt und analysiert werden. Hierzu haben die theoretische Informatik und die Mathematik das mächtige Werkzeug der Graphentheorie geschaffen. Ein zentraler Bestandteil der aktuellen Forschung hierzu sind die synthetischen Netzwerke, die im Computer erzeugt werden und die keine direkte Abbildung eines realen Netzwerks sind. Bei dieser Form der Darstellung werden die Elemente, also Dinge, Akteure oder Ereignisse, in Form von Knotenpunkten simuliert, die über Kanten miteinander verbunden werden. Dies sieht so ähnlich aus wie die Verbindung von Nervenzellen über Synapsen. Aus diesem Grund ist die Hirnforschung einerseits eine wichtige Inspiration für die theoretische Informatik – und synthetische Netzwerke sind andererseits ein probates Mittel zur Untersuchung echter Nervensysteme. Mit solchen Netzwerken kann man aber auch Infra-

strukturen simulieren, etwa Straßen-, Strom- und Wassernetze.

#### Tests am künstlichen Netz

»Es gibt viele Zwecke, für die man synthetische Netzwerke braucht«, sagt Ulrich Meyer, Professor für Algorithm Engineering an der Goethe-Universität und Fellow am Frankfurt Institute of Advanced Studies (FIAS). Bei manchen Anwendungen spielt der Datenschutz eine Rolle, weshalb man etwa die Kontakte zwischen Menschen in sozialen Medien oft nur mit simulierten Daten untersucht. Oder Forscherinnen und Forscher wollen unterschiedliche Szenarien durchspielen, die real noch gar nicht vorliegen – etwa wenn man neue Funktionen bei sozialen Medien einführen und vorher die möglichen Verknüpfungen zwischen Menschen untersuchen will. Manchmal ist auch die Gewinnung der exakten Daten enorm schwierig – etwa bei der Verknüpfung von Neuronen im Nervensystem. Auch in der Physik benötigt man für viele Fragestellungen große Mengen synthetisch erzeugter Daten, um damit Algorithmen zu trainieren. Es kommt auch vor, dass bei manchen Netzwerkanalysen die realen Datenmengen so riesig sind, dass es lange dauern würde, sie in einen Supercomputer einzuspeisen.

Doch ein solches Elektronengehirn kann ein vergleichbares synthetisches Netzwerk in kurzer Zeit selbst erzeugen. Denn Supercomputer ver-

#### AUF DEN PUNKT GEBRACHT

- Wer die Netze beispielsweise von Strom, Finanzen, Viren oder Waren besser verstehen oder sicherer machen will, muss testen, wie sich Informationen ausbreiten oder Störungen und Schwankungen auswirken.
- Wenn technische oder Datenschutzgründe die Analyse echter Netzwerke kompliziert machen, können sie durch synthetische Netze simuliert werden.
- Forschende an der Goethe-Universität arbeiten an synthetischen Netzwerken, die keine Supercomputer benötigen, sondern auf handelsüblichen PCs laufen können.



## ZUR PERSON

**Ulrich Meyer**, Jahrgang 1971, promovierte 2002 in Informatik an der Universität des Saarlandes und dem Max-Planck-Institut für Informatik. Nach Stationen in Ungarn und den USA wurde er 2007 als Professor für Algorithm Engineering an die Goethe-Universität berufen. Von 2014 bis 2022 war er Sprecher des DFG-Schwerpunktprogramms »Algorithmen für große Daten«. Seine aktuelle Forschung umfasst sowohl theoretische als auch experimentelle Aspekte der Verarbeitung großer Datensätze mit fortschrittlichen Berechnungsmodellen.

[umeyer@ae.cs.uni-frankfurt.de](mailto:umeyer@ae.cs.uni-frankfurt.de)



**Manuel Penschuck**, Jahrgang 1988, promovierte 2021 an der Goethe-Universität über die skalierbare Erzeugung von Zufallsgraphen. Als Postdoc arbeitet er jetzt an der Goethe-Universität mit dem Schwerpunkt auf parallele Graphen-Algorithmen in Anwesenheit von Speicherhierarchien für große Netzwerke.

[mpenschuck@ae.cs.uni-frankfurt.de](mailto:mpenschuck@ae.cs.uni-frankfurt.de)

fügen nicht nur über Tausende von Prozessoren, was extrem schnelle Rechenoperationen ermöglicht, sondern besitzen auch einen riesigen Arbeitsspeicher, der für den Zugriff und die Verarbeitung der vielen Giga- bis Terabyte an Daten benötigt wird, die größere Netzwerke erreichen können. Der Nachteil: Die Rechenzeit auf Supercomputern ist teuer und nur begrenzt verfügbar. »Wir arbeiten deshalb mit unserem Forschungsteam schon seit Jahren daran, die

Verfahren zur Erzeugung synthetischer Netze zu optimieren«, sagt Meyer. Das Ziel: die Arbeit mit größeren synthetischen Netzwerken auch auf handelsüblichen Computern nutzbar zu machen. Bisher sei dafür aber vor allem der Arbeitsspeicher der Computer mit 8 bis 16 Gigabyte zu klein, erklärt Meyer. Damit hakt es an einem entscheidenden Punkt, denn dann muss der Computer die Daten vom Arbeitsspeicher auf die Festplatte auslagern, was bei unstrukturierten Zugriffsmustern die Ladezeiten deutlich erhöht. Das bremst den Rechner extrem aus – und zwar um mehrere Größenordnungen – und kann die Arbeit mit solchen Netzwerken schnell unmöglich machen. »Wir arbeiten deshalb mit unserem Forschungsteam schon seit Jahren daran, die Standardverfahren zur Erzeugung synthetischer Netzwerke zu optimieren«, sagt der Informatiker.

## PC statt Supercomputer

Zum Team gehört auch Manuel Penschuck, der über die Erzeugung synthetischer Netzwerke promoviert hat und nun als Postdoc in der Frankfurter Gruppe arbeitet. Wie er betont, ist die Skalierbarkeit solcher Verfahren entscheidend. »Wenn es gelingt, die Erzeugung großer Netzwerke auch auf normalen Computern mit vertretbarem Zeitaufwand zu gestalten, dann könnten damit viele Anwendungen möglich werden, mit denen wir heute noch gar nicht rechnen«, sagt er. Das war schon in der Vergangenheit häufig so: Neue Anwendungen folgen den technischen Möglichkeiten. Konnte sich bei den ersten Personal Computern noch niemand vorstellen, wozu Privatleute jemals mehr als ein paar Megabyte an Speicherplatz benötigen könnten, so benötigen heute Videospiele, Bildbearbeitung und Videoschnitt viele Gigabyte an Daten.

Dass beim Erzeugen synthetischer Netzwerke so viele unstrukturierte Zugriffe entstehen, liegt an der Architektur der Netzwerke: »Wenn man einen Knotenpunkt neu zu einem Netzwerk hinzufügt, kann dieser in vielen Netzwerkmodellen mit zufälligen anderen Knotenpunkten aus dem bestehenden Netz verknüpft werden«, erklärt Penschuck. »Liegt dieser alte Knotenpunkt nun nicht im aktuellen Arbeitsspeicher vor, sondern befindet sich auf der Festplatte, muss er dort erst ausgelesen werden.« Die Frankfurter Forscher haben deshalb spezielle Systeme entwickelt, mit denen sie den gesamten Prozess unempfindlicher gegenüber den langsamen Festplatten- oder SSD-Zugriffen machen können.

## Leistung sparen durch schlaue Strukturen

»Unser Verfahren beruht auf einer intelligenten Strukturierung der Daten«, sagt Meyer. Dazu



werden die Knoten etwa nach Anzahl ihrer bestehenden Verbindungen vorsortiert und dann in strukturierten Blöcken im Speicher abgelegt. Außerdem lässt sich ein hinreichend komplexes Netzwerk in Unterteile aufteilen. Das Hinzufügen neuer Knotenpunkte und Verbindungen lässt sich dann parallelisieren – also auf mehrere Prozessorkerne im Computer verteilen. Heutige Computer besitzen zur schnelleren Bearbeitung verschiedener Aufgaben mehrere Prozessorkerne auf dem zentralen Chip. Das erhöht auch die Arbeitsgeschwindigkeit bei der Generierung von synthetischen Netzwerken entsprechend.

»Mithilfe all dieser Kniffe lässt sich der Einbruch der Rechengeschwindigkeit umgehen, sobald das Netzwerk nicht mehr in den Computerspeicher passt«, sagt Penschuck. »Das erkaufen wir uns zwar dadurch, dass unser Verfahren – bedingt durch die Strukturierung der Daten – bei kleineren Netzwerken etwas langsamer ist als das jeweilige Standardverfahren zur Erzeugung synthetischer Netzwerke. Aber während die Standardverfahren bei großen Netzwerken grob um einen Faktor Tausend langsamer werden, sobald der Arbeitsspeicher mit den Netzwerkdaten gefüllt ist, verlieren die von uns entwickelten Netzwerk-Generatoren kaum an Geschwindigkeit.«

### Netzwerke für die nächste Epidemie

In Zukunft wollen die Forscher die synthetischen Netzwerke leichter allgemein zugänglich

machen. Bislang erfordert der Umgang mit ihnen noch ein hohes Maß an Kompetenz in Informatik und eine gewisse Einarbeitung. Gerade bei vielen gesellschaftlich relevanten Fragestellungen sollen solche Netzwerke künftig einfacher zur Anwendung kommen – sei es bei der Simulation von Infektionsketten in der Epidemiologie oder bei der Analyse von Kontakten in sozialen Netzwerken im Rahmen der sogenannten Sozioinformatik. »Deshalb wollen wir eine Art Werkzeugkoffer für synthetische Netzwerke entwickeln, mit dem man auch ohne große Vorkenntnisse arbeiten kann und der dennoch alle wichtigen Instrumente zur Verfügung stellt«, erklärt Meyer.

Dieser Werkzeugkoffer wird nicht nur den neuen Netzwerk-Generator enthalten, sondern auch viele weitere Verfahren, die von der weltweiten Forschungsgemeinde in den vergangenen Jahren entwickelt wurden. Dieses Softwarepaket kann dann auf allen möglichen Gebieten eingesetzt werden – in der Epidemiologie und Sozioinformatik, aber auch in der Physik und der Analyse von Stromnetzen. »Wir sind selbst gespannt, zu welchen Zwecken die Anwender solche Netzwerke noch nutzen werden«, so Meyer. Die Möglichkeit, mit großen Netzwerken an handelsüblichen PCs zu arbeiten, ohne Rechenzeit am Supercomputer beantragen (und bezahlen) und sensible Daten herausgeben zu müssen, dürfte für Forscherinnen und Forscher aus den unterschiedlichsten Gebieten interessant sein. ●

**Komplexe Versorgung:**  
Im Control Center des Unternehmens 50Hertz steuern und überwachen Mitarbeitende das Stromübertragungsnetz im Norden und Osten Deutschlands.

### Der Autor

Dirk Eidemüller  
dirk.eidemueller@gmx.de  
(siehe Seite 23)