

DISS. ETH NO. 23124

Society in Graphs

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH

(Dr. sc. ETH Zurich)

presented by

BARBARA KELLER

MSc, ETH Zürich, Switzerland

born on 22.04.1982

citizen of
Switzerland

accepted on the recommendation of
Prof. Dr. Roger Wattenhofer, examiner
Prof. Dr. Eric Goles, co-examiner
2015

Abstract

Individuals are seldom completely independent and should be considered as part of their social surroundings. Their actions both influence their social network and are influenced by the network. This phenomenon can be seen in online social networks formed on online platforms as well as in offline social networks formed, for example, by a group of scholars. In order to understand the fundamental effects of influence in social systems, we model social networks by the means of a well understood mathematical model, a graph.

The first part of this dissertation focuses on the impact of gender on the formation of a network of PhD students and their supervisors. We find evidence for homophilic behavior as well as for the existence of a *glass ceiling* in a network from the data of co-authorship. From over 1.3 millions of authors of scientific articles in computer science, we extract the student-supervisor relationship graph and analyze its properties.

Furthermore, we introduce mathematical formulations for the occurrence of a glass ceiling and an influence inequality in a network. We establish a network forming process integrating three observed characteristics of this network, namely a smaller entry rate for women, preferential attachment, and homophilic behavior. We prove that these three conditions are sufficient to produce a glass ceiling in the network. We also show, that if one of these three characteristic is missing, the glass ceiling according to the mathematical definition does not occur.

In the second part of this dissertation we examine how opinions evolve in networks. Every node in the network has an initial opinion and the nodes can observe the opinions of their neighbors. We assume a simplistic setting where the nodes are influenced by the opinions of their neighbors and always change their opinion to the opinion of the majority of their neighbors. We study several different variations of this model and investigate how long the system takes to reach a stable state. For asynchronous networks we find unweighted graphs which take $\Omega(n^2)$ steps until convergence. For unweighted synchronous networks we show graphs with a convergence time of $\tilde{\Omega}(n^2)$, where the $\tilde{\Omega}(\cdot)$ notation hides polylogarithmic factors. Additionally we show that allowing the influence to be weighted increases the convergence time dramatically to $\Omega(2^n)$.

Zusammenfassung

Menschen sind selten komplett unabhängig und sollten als Teil ihrer sozialen Umgebung betrachtet werden. Ihre Handlungen beeinflussen einerseits Ihre soziale Umgebung und sind andererseits selbst durch das soziale Netz der betreffenden Person beeinflusst. Man findet dieses Phänomen sowohl in sozialen Netzwerken welche auf online Plattformen geformt werden, als auch in herkömmlichen sozialen Netzwerken, wie man sie zum Beispiel unter Schülern findet. Um die fundamentalen Effekte von Einfluss in sozialen Systemen besser zu verstehen, modellieren wir soziale Netzwerke mit einem in der Mathematik gut bekannten Werkzeug, dem Graph.

Der erste Teil dieser Dissertation beschäftigt sich hauptsächlich mit den Auswirkungen welches das Geschlecht von Studenten auf die Bildung des Netzwerkes von Doktoranden/Doktorandinnen und ihrer Doktorväter/Doktormütter hat. In einem Netzwerk, aus Daten über Koautorenschaft, zeigen wir sowohl Hinweise für homophiles Verhalten als auch für die Existenz einer *gläsernen Decke*. Aus diesem Datenset, welches über 1.3 Millionen Autoren von wissenschaftlichen Artikeln aus dem Bereich Informatik enthält, haben wir einen Graph extrahiert, welcher die Studenten-Leiter Beziehung darstellt und studieren dessen Eigenschaften.

Ausserdem führen wir mathematische Formulierungen für die Erscheinung einer gläsernen Decke und einer ungleichen Verteilung von Einfluss in einem Netzwerk ein. Wir konstruieren einen Netzwerkbildungsprozess welcher drei beobachtete Eigenschaften, nämlich weniger Frauen, bevorzugte Anschliessung und homophiles Verhalten integriert. Zusätzlich beweisen wir, dass diese drei Eigenschaften ausreichend sind um eine gläserne Decke in einem Netzwerk hervorzurufen. Wir zeigen auch, dass falls eine dieser drei Eigenschaften wegfällt, sich keine gläserne Decke nach mathematischer Definition bildet.

Im zweiten Teil dieser Dissertation untersuchen wir, wie sich Meinungen in Netzwerken entwickeln. Jeder Knoten im Netzwerk hat eine ursprüngliche Meinung und die Knoten können die Meinung ihrer Nachbarn sehen. Wir nehmen ein einfaches Modell an, in welchem die Knoten beeinflusst sind von den Meinungen ihrer Nachbarn und ihre Meinung immer an die Meinung der Mehrheit Ihrer Nachbarn anpassen. Wir ana-

lysieren verschiedene Varianten dieses Modells und erforschen wie lange ein solches System braucht, um einen stabilen Zustand zu erreichen. In asynchronen Netzwerken finden wir ungewichtete Graphen welche $\Omega(n^2)$ Schritte benötigen um sich zu stabilisieren. Für ungewichtete, synchrone Netzwerke zeigen wir Graphen welche eine Stabilisierungszeit von $\tilde{\Omega}(n^2)$ benötigen, wobei die $\tilde{\Omega}(\cdot)$ Notation polylogarithmische Faktoren versteckt. Zusätzlich zeigen wir, dass wenn der Einfluss zwischen den Knoten gewichtet sein kann, sich die Stabilisierungszeit tiefgreifend erhöht zu $\Omega(2^n)$.