

Analyzing a pharma company network: are friends still competing?

Michael Haas

Institut für Computerlinguistik
Ruprecht-Karls-Universität Heidelberg

14. November 2011

- ▶ Der Erfolg eines Medikaments in einer Studie bestimmt Marktzulassung und Verkaufszahlen.
- ▶ Negative Studienergebnisse können Firmen-Insider (Anteilseigner oder Vorstandsmitglieder) schädigen.
- ▶ Womöglich haben Insider mehrerer Firmen ein Interesse daran, dass ihre Firmen ihre Produkte nicht vergleichen.

Vergleichen befreundete Medikamentenhersteller ihre Produkte in medizinischen Studien?

Wieso Netzwerkanalyse?

- ▶ Daten bieten sich zur Modellierung als (soziales) Netzwerk an
- ▶ Netzwerk-Paradigma erlaubt intuitives Verständnis des multi-partiten Netzwerks
- ▶ Vielfältige Analyse- und Wiederverwendungsmöglichkeiten

Vergleichen befreundete Medikamentenhersteller ihre Produkte in medizinischen Studien?

- ▶ Aufbau des Netzwerks
 - ▶ Abschöpfen der Datenquellen
 - ▶ Verknüpfung der Daten zwischen Quellen
- ▶ Analyse des Netzwerks
 - ▶ Welche Firmen sind befreundet?
 - ▶ Welche Firmen vergleichen ihre Medikamente?
 - ▶ Vergleichen befreundete Firmen ihre Produkte weniger oft?

Aufbau des Netzwerks

Wir müssen wissen:

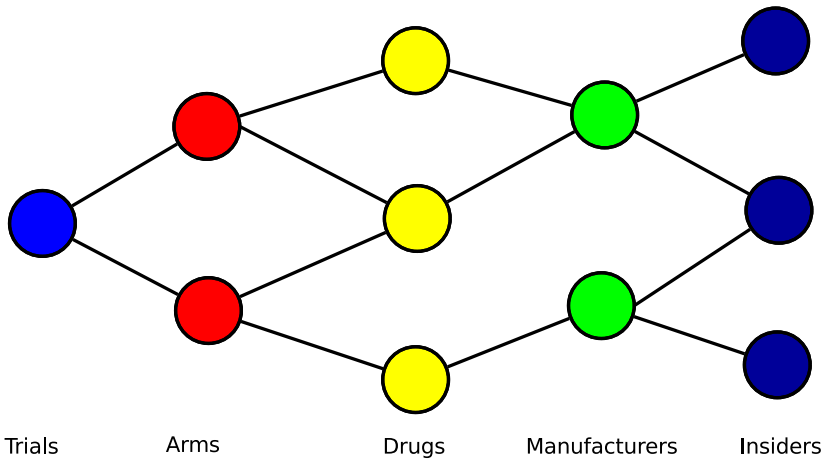
- ▶ Welche Medikamente in klinischen Studien verglichen werden
 - ▶ [ClinicalTrials.gov](https://clinicaltrials.gov)
- ▶ Welche Insider (Anteilseigner und Vorstandsmitglieder) in welchen Firmen sitzen
 - ▶ US-Handelsaufsichtsbehörde SEC
- ▶ Welcher Hersteller welche Medikamente produziert
 - ▶ [DrugBank](https://www.drugbank.ca)

Verknüpfung der Daten

Entitäten aus Datenquellen müssen verknüpft werden: Named Entity Resolution

- ▶ Medikamente aus ClinicalTrials.gov mit Herstellern aus DrugBank (“Etanercept + Levofloxacin + placebo prednisone capsules” ↔ DB00005, DB01137)
- ▶ Hersteller aus DrugBank mit Firmen aus der Insider-Datenbank der SEC (“Warner chilcott inc” ↔ “Warner Chilcott CORP”)

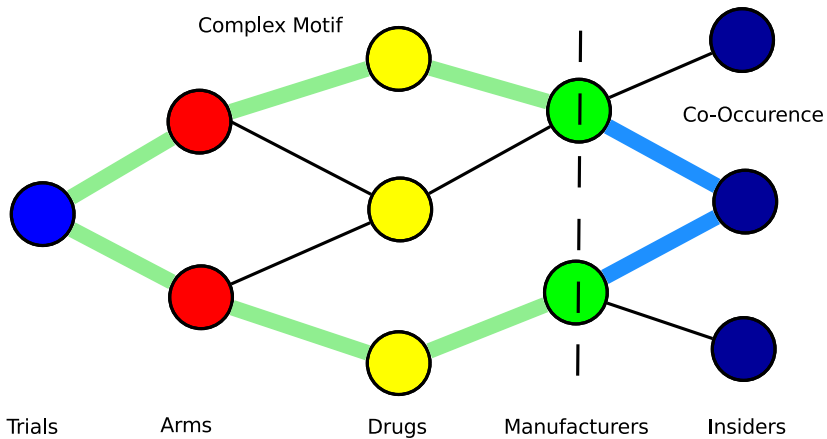
→ Jeweils pattern-basierte Lösung, $F \approx 0.8$



Projektion auf Hersteller

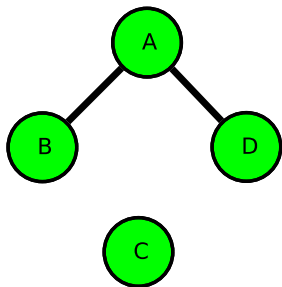
Auf der Menge der Hersteller interessiert uns:

- ▶ Wie verbinden die gemeinsamen Insider die Hersteller?
- ▶ Wie verbinden die gemeinsamen Medikamententests die Hersteller?

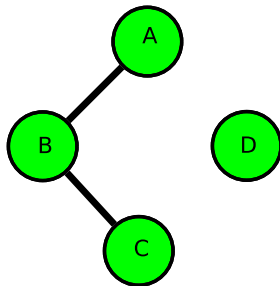


Zwei neue Netzwerke:

- ▶ **trial projection** Hersteller verknüpft, die Medikamente vergleichen
- ▶ **insider projection** Hersteller verknüpft, die gemeinsame Insider haben



Trial projection



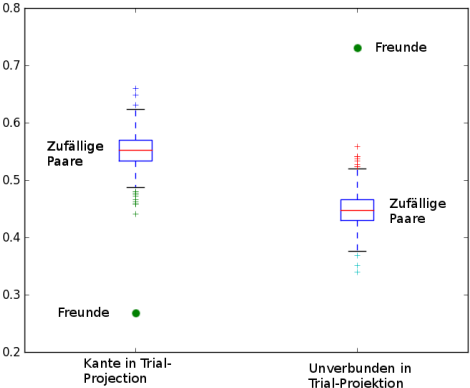
Insider projection

Vergleichen befreundete Medikamentenhersteller ihre Produkte in medizinischen Studien?

- ▶ Befreundete Hersteller: Menge alle mit einer Kante verbundenen Firmen aus insider projection
- ▶ Wie oft testen diese befreundeten Hersteller gegeneinander, verglichen mit zufälligen Paaren von Herstellern?

Boxplot

Vergleichen befreundete Medikamentenhersteller ihre Produkte in medizinischen Studien?



Vergleichen befreundete Medikamentenhersteller ihre Produkte in medizinischen Studien?

Befreundete Firmen haben weniger häufig wenigstens eine Studie, in der sie ihre Produkte vergleichen, verglichen mit zufälligen Paaren von Firmen.

Mögliche Erklärungen:

- ▶ Insider beeinflussen klinische Studien
- ▶ Insider wählen Firmen aus unterschiedlichen Branchen

Alle Beobachtung auf Netzwerk relevant?

Sind die Verbindungen zwischen Firmen wirklich relevant?

- ▶ Ein Insider in vielen Firmen: Insider verteilt seinen Einfluss auf viele Firmen
- ▶ Eine Firma mit vielen Insidern: Einfluss eines einzelnen Insiders gering

Test gegen Null-Model

- ▶ Test der Verbindungen zwischen Firmen im Netzwerk gegen Fixed Degree Sequence Model (Zweig, 2011)
- ▶ nur signifikante Verbindungen werden beibehalten
→ auch auf beschnittenen Projektionen gleiches Muster sichtbar.

Vergleichen befreundete Medikamentenhersteller ihre Produkte in medizinischen Studien?

- ▶ Beobachtungen auf dem Netzwerk erhärten die Hypothese
- ▶ Beobachtung auch bei verschiedenen Signifikanz-Leveln weiterhin sichtbar
- ▶ Zur endgültigen Klärung fehlt: Bevorzugen Insider Firmen aus verschiedenen Sub-Branchen?
→ Netzwerkanalytischer Ansatz auf vorhandenen Daten
- ▶ Erkenntnis: Arbeit mit Open Data, vor allem quellenübergreifend, schwieriger als notwendig