

1 Ucinet, Netdraw, Pajek?

Es gibt eine Unzahl von open source, freien und kommerziellen Programmen zur Netzwerkanalyse. In vielen Fällen handelt es sich eher um Programmbibliotheken als um Anwendungen, oder die Funktionalität ist sehr stark eingeschränkt. In dieser und der nächsten Woche lernen wir zwei bzw. drei wichtige Programme kennen, mit denen Sie alle gängigen (und viele ungewöhnliche) Analysen durchführen können.

Ucinet: Der Dinosaurier unter den Programmen. Ucinet ist seit vielen Jahren verfügbar (aktuelle Version: 6) und enthält buchstäblich hunderte von Prozeduren, die im Lauf der Zeit hinzugefügt wurden.

- Vorteile: Kann praktisch alles, liest und schreibt verschiedene Formate, stabil, weitverbreitet, gut für die Eingabe, Verwaltung und Konvertierung von kleinen und mittelgroßen Datensätzen
- Nachteile: Kommerziell (aber sehr günstig für Studierende und Universitäten, 60 Tage ohne Registrierung nutzbar), keine Visualisierung (aber Schnittstellen zu Netdraw und Pajek), veraltetes Handbuch (aber aktuelle Hilfefunktion), Benutzeroberfläche

NetDraw: Freeware, mit Ucinet gebündelt aber auch separat erhältlich. Liest Ucinet- und Pajek-Formate. Schwerpunkt auf graphischer Darstellung, enthält aber auch einige Analyseprozeduren.

- Vorteile: Frei, intuitiv zu bedienen, Vielzahl von Funktionen, stabil, Integration mit Ucinet.
- Nachteile: Benutzeroberfläche, nicht für große Datensätze geeignet.

Pajek: Freeware, (relativ) neues Programm, von Mathematikern entwickelt, statistische Analyse *und* Visualisierung von Netzwerken.

- Vorteile: Frei, nach Eingewöhnung intuitiv zu bedienen, sehr große (und wachsende) Anzahl von Funktionen, schnell, kann mit großen und sehr großen Datensätzen arbeiten
- Nachteile: Oberfläche erfordert Eingewöhnung, wichtige Funktionen bei de Nooy et al. gut erklärt, aber neue Funktionen kaum dokumentiert, ungewöhnliche (graphentheoretisch-mathematische) Perspektive der Entwickler.

2 Datensätze

- Streik-/Sägemühle
- Carter-Administration
- Politische Bücher

- Politische Blogs
- (Kursdaten)

3 Ucinet kennenlernen

Ucinet ist gewöhnungsbedürftig. Die Oberfläche erinnert an Windows 95-Zeiten und die Zuordnung der Menüs ist nicht sofort nachvollziehbar. Beim Arbeiten mit Ucinet folgt man aber immer demselben Zyklus: Daten werden in einer Prozedur eingeladen, verarbeitet und wieder als Datei gespeichert. Dateien sind also nicht permanent „offen“.

3.1 Format

Ucinet liest eine Vielzahl von Formaten ein, u. a. Excel und Pajek. Ucinets natives Format ist insofern ungewöhnlich, als jedes Netzwerk durch zwei Dateien (mit den Extensionen `##h` und `##d`) beschrieben wird.

3.2 Diary Data Carter Administration

Der Datensatz `carteryear1` enthält die (dokumentierten) Treffen zwischen Präsident Carter (`presat`) und verschiedenen Ministern und Beratern während des ersten Jahres von Carters Präsidentschaft. Ein zweiter Datensatz zeigt dasselbe Netzwerk vier Jahre später.

Starten Sie `ucinet`. Öffnen Sie den Spreadsheet-Editor (mit `ctrl-S` oder dem zweiten Button von links). Öffnen Sie den Datensatz `carteryear1`. Was sehen Sie? Wieviele Knoten gibt es? Ist das Netzwerk binär oder valued? Sind die `ties` gerichtet?

3.3 Dichte, Zentralität, Zentralisation

Wieviele Verbindungen kann es in diesem Netzwerk maximal geben? Wieviel Prozent dieser Verbindungen sind realisiert? Wählen Sie aus dem Menü `Network - Cohesion - Density - New Density`. Als Input-Datensatz müssen Sie wieder `carter1` angeben (falls noch nicht voreingestellt). Überrascht Sie das Ergebnis?

Wir kennen drei verschiedene Konzepte der Zentralität eines Knotens: `degree`, `closeness` und `betweenness`. Auf Basis dieser drei Maße kann man außerdem den Grad der Zentralisation des Gesamtnetzwerkes bestimmen (Vergleich mit maximal möglicher Ungleichverteilung von Zentralität). Alle diese (und einige andere) Maßzahlen sind in `ucinet` implementiert. Bestimmen Sie die entsprechenden Werte mit `Network - Centrality - Degree`, `Network - Centrality - Closeness` bzw. `Network - Centrality - Freedman Betweenness - Node Betweenness`.

Angenommen Sie wollen wissen, wie die Akteure CSAT und ZBAT miteinander verbunden sind. Wählen Sie `Network - Paths`.

Haben Sie eine Vorstellung davon, wie das Kommunikationsnetzwerk insgesamt aussehen muß?

Wählen Sie das Menü `Visualize - Netdraw`. Falls notwendig, öffnen Sie den Datensatz über `File - Open - Ucinet dataset - network`. Entspricht die Grafik Ihren Erwartungen? Können Sie erklären, warum das Netzwerk hochgradig, aber nicht hundertprozentig zentralisiert ist?

Gehen Sie zurück nach `ucinet`. Bestimmen Sie jetzt für den Datensatz `carteryear4` die drei Varianten von Zentralität und Zentralisation (das Netzwerk ist etwas kleiner geworden). Wie hat sich das Netzwerk verändert? Welche graphische Struktur erwarten Sie. Öffnen Sie den Datensatz nun in `Netdraw`. Entspricht das Bild Ihren Erwartungen? Wie interpretieren Sie die Veränderungen?

3.4 Subgruppen

Zu beiden Zeitpunkten bildet das Netzwerk eine einzige Komponente. Wie können Sie dies in `Netdraw` graphisch erkennen? In `Ucinet` können Sie diesen optischen Eindruck mit `Network - Regions - Components - Simple Graph` bestätigen. Gibt es innerhalb dieser Komponente Subgruppen? In `Netdraw` können Sie sich sehr einfach k -Cores anzeigen lassen: `Analysis - K-Cores`. Die Knoten werden jetzt entsprechend eingefärbt. Wenn Sie mit der rechten Maustaste auf einen Knoten klicken, sehen Sie u. a. die Zugehörigkeit zum Core.

Eine Clique (im engen Sinn) ist eine Subgruppe, in der alle Mitglieder direkt miteinander verbunden sind. `Ucinet` verfügt über einen effizienten Algorithmus zur Identifikation von Cliques. Wählen Sie `Network - Subgroups - Cliques` (wiederum mit dem `carteryear4`-Datensatz). Betrachten Sie die Ausgabe. Sind die beiden Cliques Cliques im engeren Sinne? Ist ein Teil einer der beiden Cliques eine Clique im engeren Sinne? $N(=2)$ -Cliques und -Clans sind für diesen Datensatz nicht sehr informativ – warum?