Redes y Sistemas Complejos
Cuarto Curso del Grado en Ingeniería Informática
Seminario 1: Herramientas de Análisis de Redes

Oscar Cordón García
Dpto. Ciencias de la Computación e Inteligencia Artificial
ocordon@decsai.ugr.es

HERRAMIENTAS PARA EL ANÁLISIS DE REDES Y SISTEMAS COMPLEJOS (1)

La creciente necesidad de minería y visualización de redes complejas ha derivado en el desarrollo de una gran colección de herramientas software

Algunas son más técnicas y están dirigidas a programadores mientras que otras son más intuitivas y, por tanto, más adecuadas para usuarios de las ciencias sociales

La mayoría permiten el cálculo de medidas locales y globales de la red, su visualización gráfica y la detección de comunidades y, en general:

- Creación de redes,
- Visualización y manipulación de redes,
- Análisis cualitativo y cuantitativo/estadístico de redes,
- Detección de comunidades,
- Análisis predictivo (modelos de contagio, predicción de enlaces, etc.)

Los más populares pueden ser Gephi, Pajek, UCInet, NodeXL y las distintas bibliotecas desarrolladas en R (igraph, sna, tnet, statnet y NetworkX)

Combe D, et al. A comparative study of social network analysis tools. Soc Netw 2010, 2:1–12

Paquetes Software de Visualización y cálculo de medidas básicas:

Herramientas intuitivas, con GUI y sin necesidad de conocimientos de programación (dirigidas a usuarios de áreas generales)

Gephi: http://gephi.org/
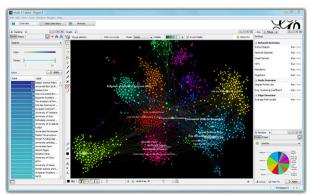
Será el que usemos nosotros:
Descarga: https://gephi.org/users/download/
Tutorial: http://gephi.org/users/quick-start/

Gephi is an interactive visualization and exploration platform for all kinds of networks and complex systems, dynamic and hierarchical graphs, endowed with a three-dimensional render engine to display real-time evolving networks

Developed in Java. Runs on Windows, Linux and Mac OS X. Gephi is open-source

Bastian M., Heymann S., Jacomy M. (2009). Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media
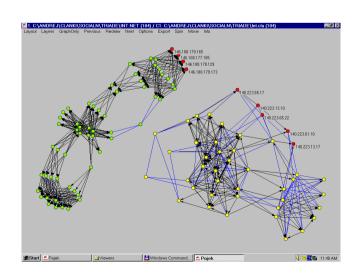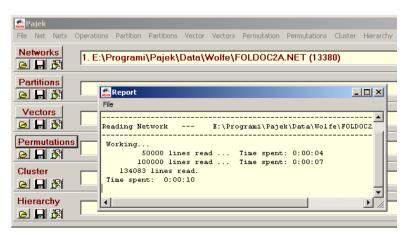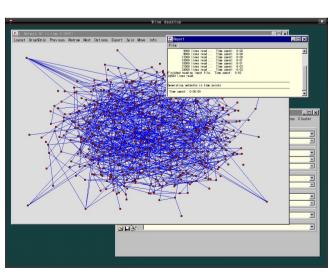
- **Pajek**: http://pajek.imfm.si/doku.php

Pajek (Slovene word for Spider) is a program, for Windows, for analysis and visualization of large-scale networks. It is freely available, for noncommercial use

Very extensive functionality via drop-down menus

Batagelj V., Mrvar A.: Pajek - Analysis and Visualization of Large Networks. in Jünger, M., Mutzel, P., (Eds.) Graph Drawing Software. Springer, Berlin 2003. pp. 77-103
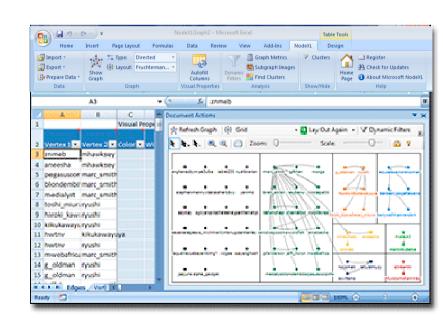
- **NodeXL**: http://nodexl.codeplex.com/

NodeXL is a freely available add-in to Microsoft Excel 2007 and 2010 for the overview, discovery and exploration of networks

With NodeXL, you can enter a network edge list in a worksheet, click a button and see your graph, all in the familiar environment of the Excel window

It does not require any programming skills and it is very user friendly

Not suitable for the analysis of large networks

Smith M et al. Analyzing (social media) networks with NodeXL. In: Fourth International Conference on Communities and Technologies. New York, NY: ACM 2009, 255–264

Herramientas de Programación:

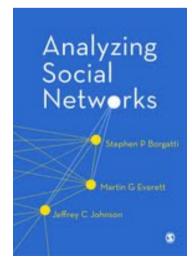Herramientas más técnicas dirigidas a usuarios con conocimientos de programación de ordenadores:

- **UCInet**: https://sites.google.com/site/ucinetsoftware/home

  UCINET 6 for Windows is a commercial software package for the analysis of social network data. It was developed by Lin Freeman, Martin Everett and Steve Borgatti. It comes with the NetDraw network visualization tool

  Especially suitable for statistical and matricial analyses
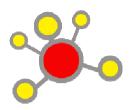
  The program can be downloaded and used for free for 90 days

  Borgatti SP, Everett MG, Freeman LC. Ucinet for Windows: Software for social network analysis. Harv Anal Technol 2002, 2006

- **Igraph**: http://igraph.sourceforge.net/

Igraph is a free software package for creating and manipulating undirected and directed graphs. It includes implementations for classic graph theory problems like minimum spanning trees and network flow, and also implements algorithms for some recent network analysis methods, like community structure search

It contains functions for generating regular and random graphs, manipulating graphs, assigning attributes to vertices and edges. It can calculate various structural properties, graph isomorphism, includes heuristics for community structure detection, supports many file formats. The R and Python interfaces support visualization

Igraph runs on most modern machines and operating systems, and it is tested on MS Windows, Mac OSX and various Linux versions. The software you need for installing igraph depends on whether you want to use the C library, the R package or the Python extension; and may vary depending on your platform

HERRAMIENTAS PARA EL ANÁLISIS DE REDES Y SISTEMAS COMPLEJOS (7)

- **NetworkX**: http://networkx.github.io/

NetworkX is a Python language software package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks

Extensive functionality:

- scales to large networks by taking advantage of existing C, Fortran libraries for large matrix computations
- Nodes can be "anything" (e.g. text, images, XML records). Edges can hold arbitrary data (e.g. weights, time-series)
- Generators for classic graphs, random graphs, and synthetic networks
- Standard graph algorithms. Network structure and analysis measures

Open source. Well tested. Additional benefits from Python: fast prototyping, easy to teach, multi-platform

Hagberg AA, Schult DA, Swart PJ. Exploring network structure, dynamics, and function using NetworkX. In: Seventh Python in Science Conference; 2008, 11–15
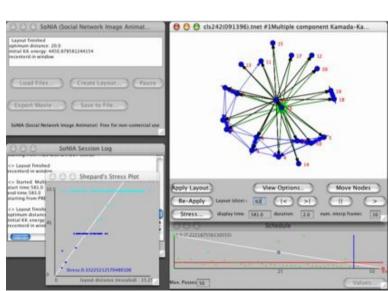
Herramienta Mixta y Especializada:

SoNIA (http://www.stanford.edu/group/sonia/) is a Java-based package for visualizing dynamic network data. In addition to information about the relations between entities there is also information about when these relations occur, or at least the relative order in which they occur

Our intention is to aid the user in constructing "meaningful" layouts and export the resulting images or "movies" of the network, along with information about the techniques and parameter settings used to construct the layouts, and some form of statistic indicating the "accuracy" or degree of distortion present in the layout

It also works as a platform for the development, testing, and comparison of various static and dynamic layout techniques (modular OpenSource project)

Bender-deMoll S and McFarland DA. The Art and Science of Dynamic Network Visualization. Journal of Social Structure (2006) 7:2

NetLogo: Modelado de Procesos Dinámicos en Redes

NetLogo (http://ccl.northwestern.edu/netlogo/) is a multi-agent programmable modeling environment for simulating natural and social phenomena. You can download it free of charge

It is authored by Uri Wilensky and developed at the Center for Connected Learning (CCL) and Computer-Based Modeling, NorthWestern University

NetLogo is particularly well suited for modeling complex systems developing over time. Modelers can give instructions to thousands of "agents" all operating independently. This allows us to explore the connection between the micro-level behavior of individuals and the macro-level patterns emerging from their interaction

NetLogo lets students "play" with simulations, exploring their behavior under various conditions. It is also an authoring environment enabling developers to create their own models

Emplearemos distintas simulaciones realizadas en NetLogo en las prácticas de la asignatura. Podéis descargarlo en https://ccl.northwestern.edu/netlogo/oldversions.shtml