Zurich University of Applied Sciences

Graph Drawing

Oliver Dürr

Brown Bag Semiar 16 Apr.

Übersicht

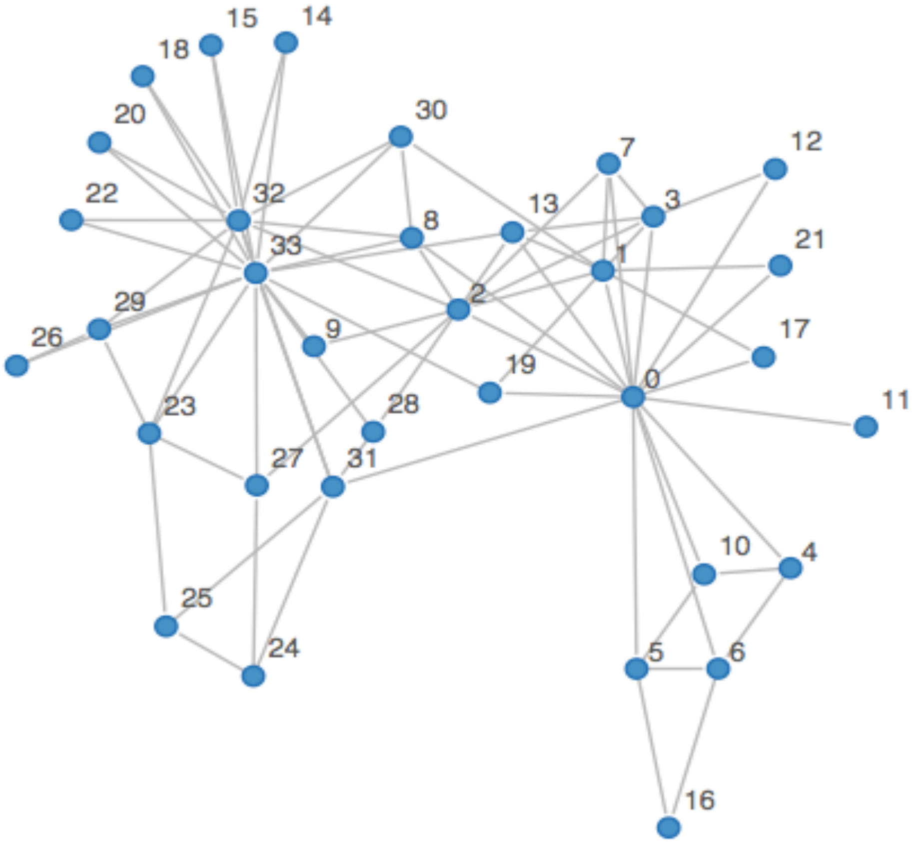
Zurich University of Applied Sciences

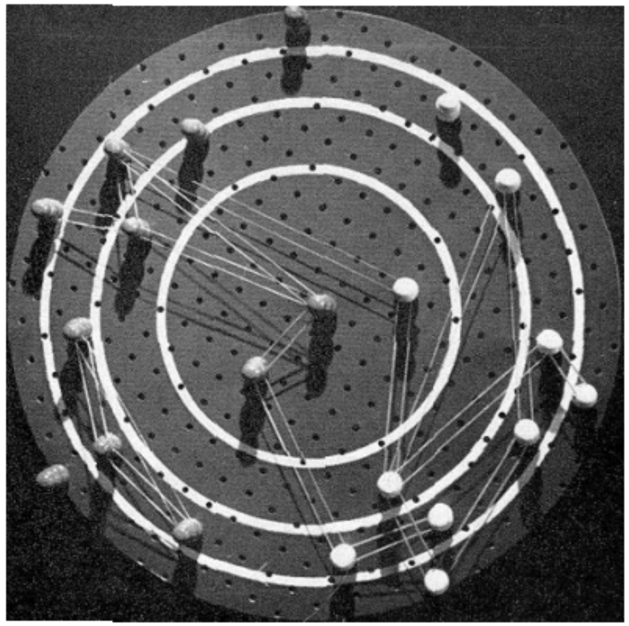
- Beispiele für Netzwerke
- Was sind Netzwerke
- Überblick Netzwerkanalyse
- Clustering auf Netzwerken
- Graph-Drawing
  - Problemstellung
  - Force-Directed Methoden
  - Optimierung I (Barnes-Hut Approximation)
  - Optimierung II (Multilevel)

Beispiele für Netzwerke

Social Networks (Zachary Karate Club)

Zurich University of Applied Sciences

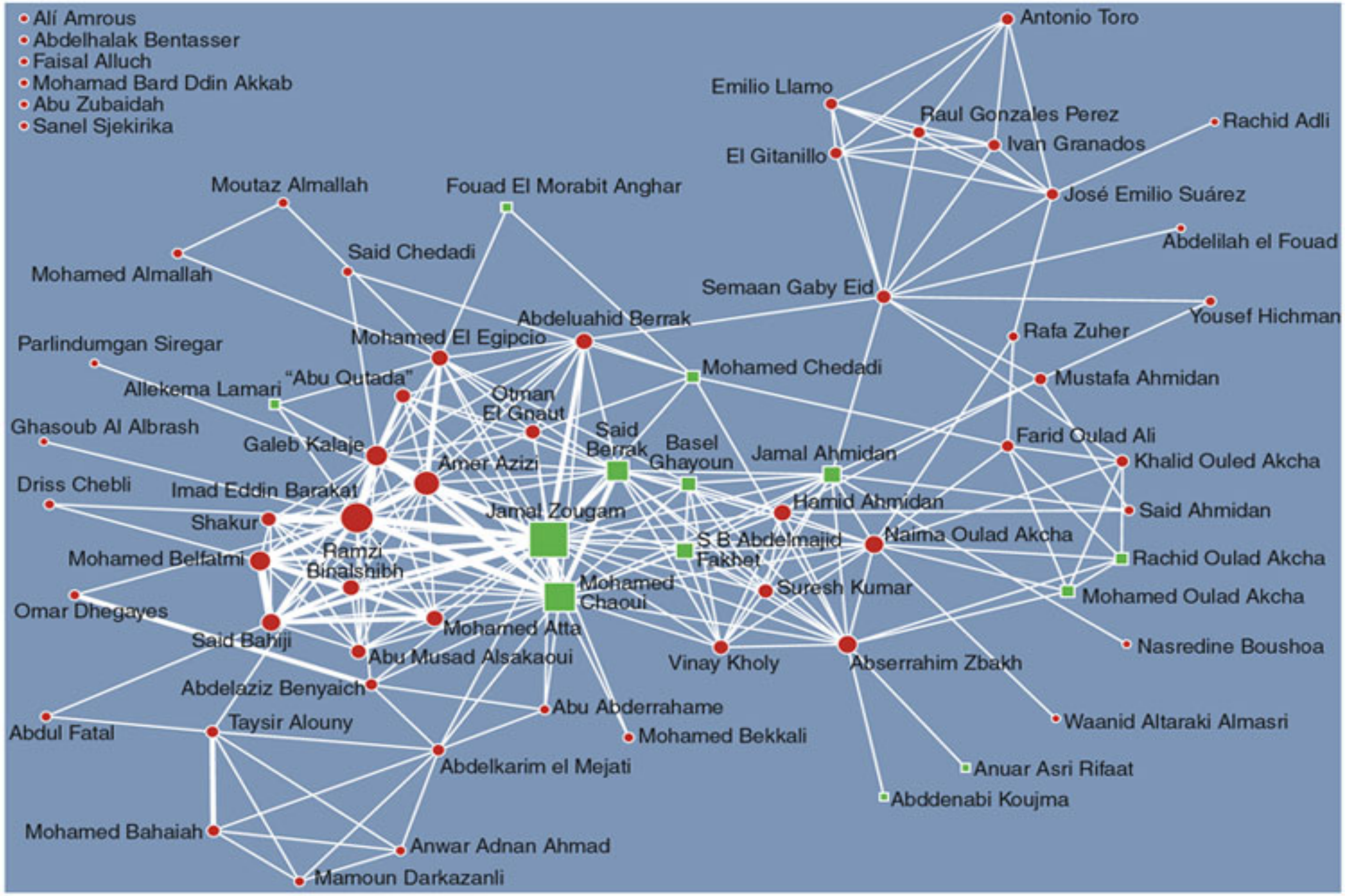




(c) McKenzie's board for manual layout [Nor52]

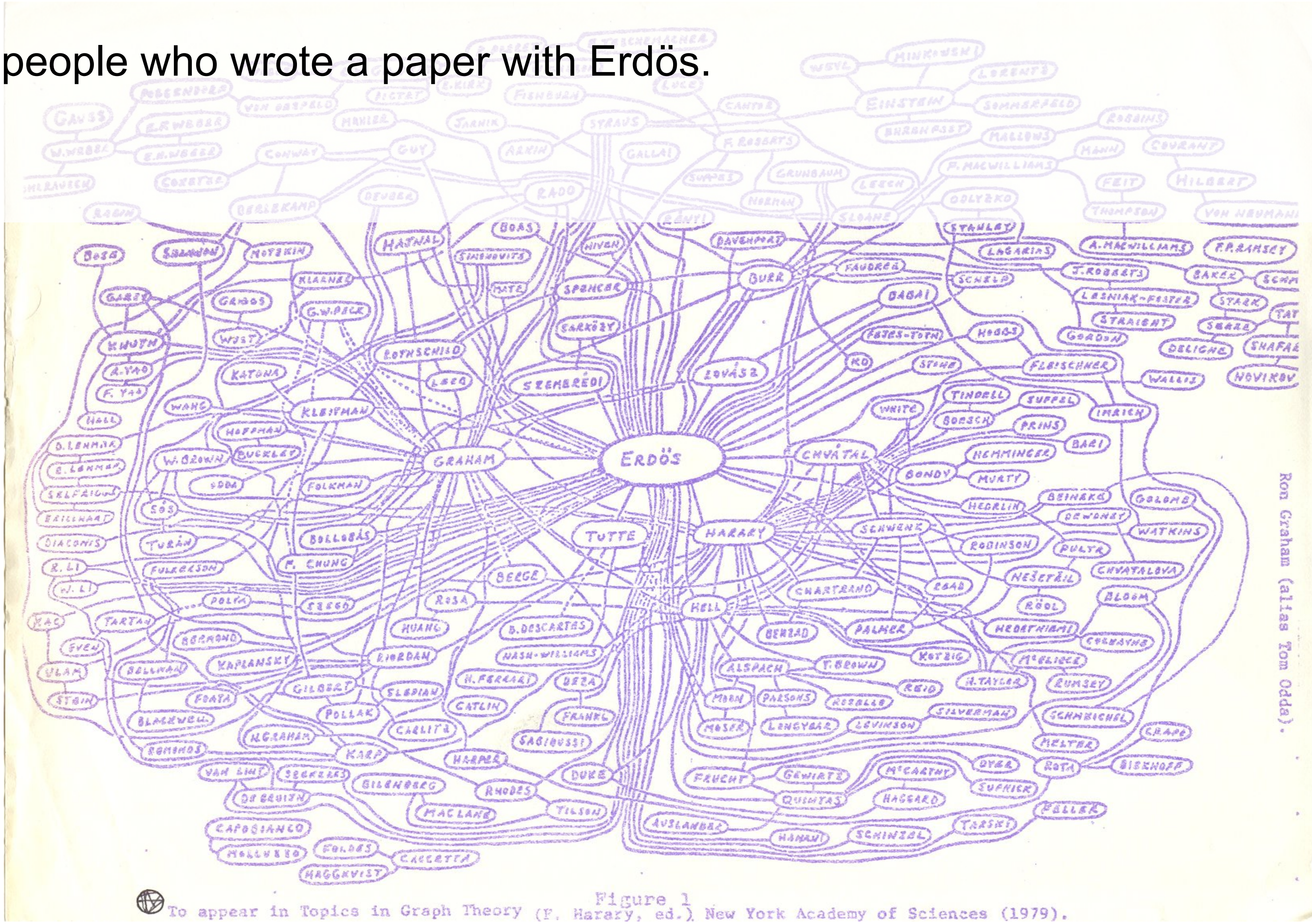
Terror Networks

From [American Scientist Article]()

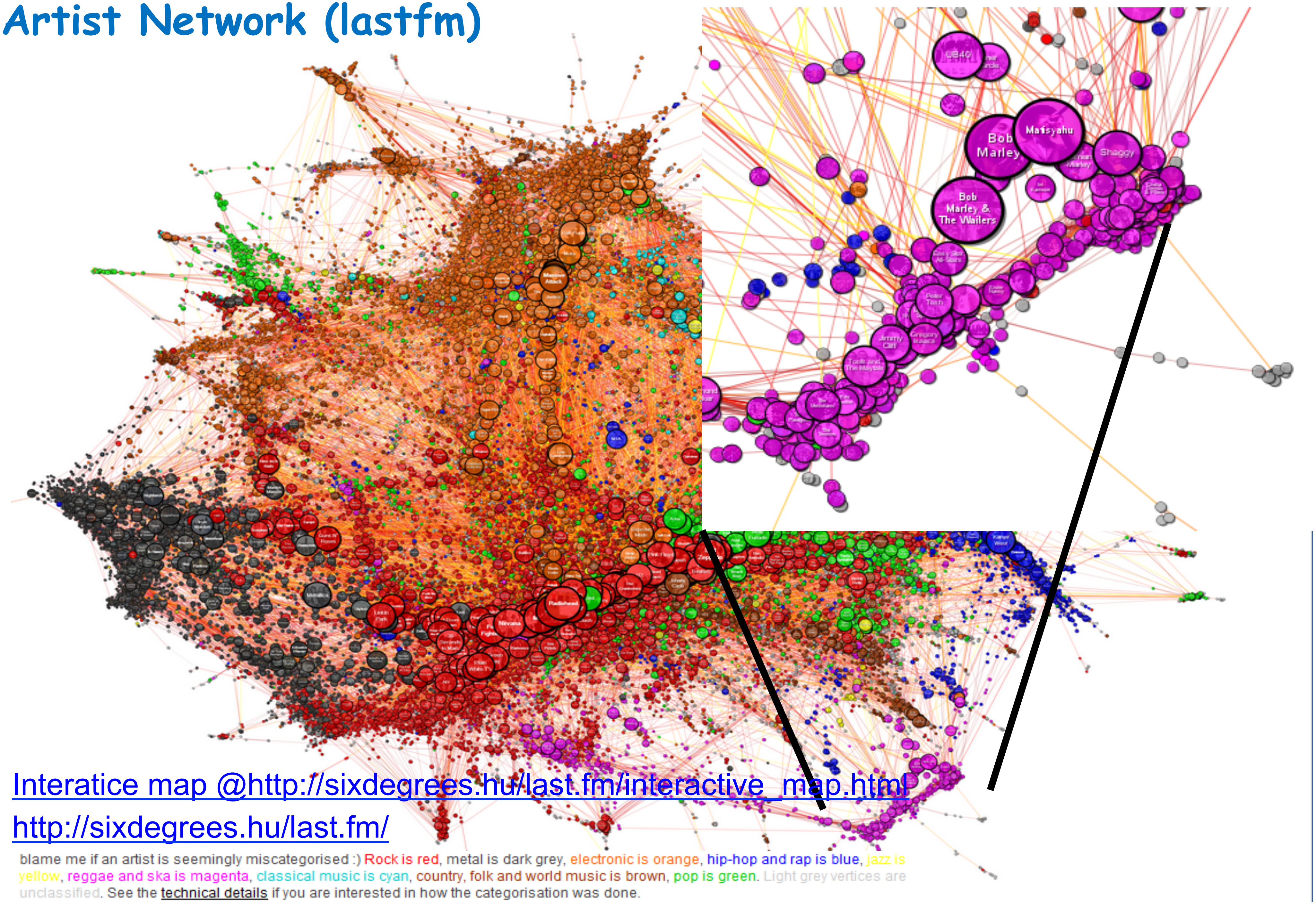


Authorship network

All people who wrote a paper with Erdös.



Artist Network (lastfm)



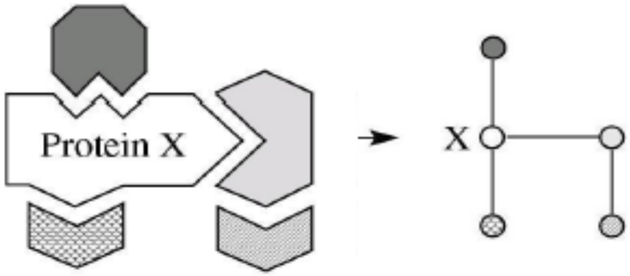
[Interatice map @http://sixdegrees.hu/last.fm/interactive_map.html](http://sixdegrees.hu/last.fm/interactive_map.html)

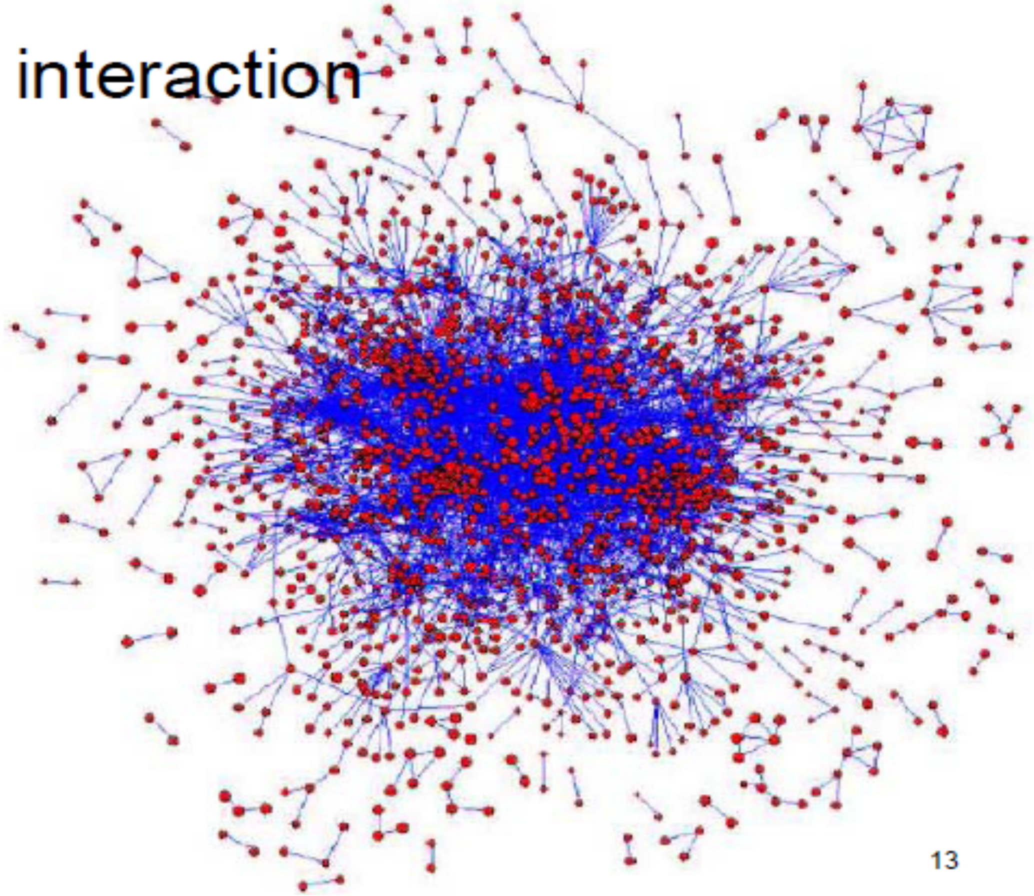
http://sixdegrees.hu/last.fm/

Protein Protein Interaction Network

- **Biological nets**

  E.g., Protein-protein interaction (PPI) networks



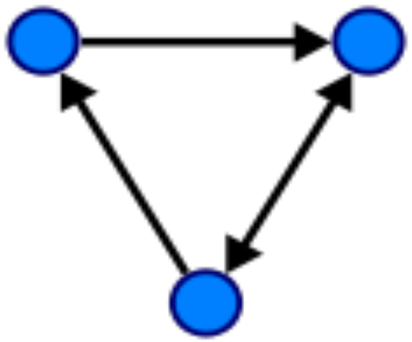


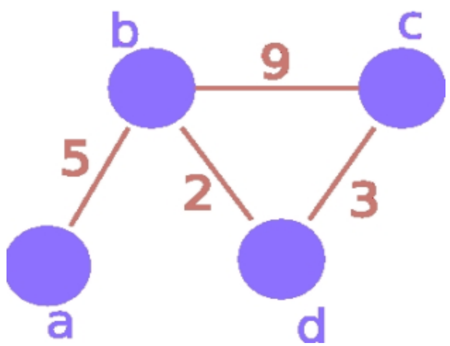
Definition von Netzwerken

Introduction

What is a network (or graph)?

- A set of nodes (vertices) and edges (links)
- **Edges** describe a relationship between the **nodes**
- Some generalizations (see later)
- Edges can be (un-)directed, or (un-)weighted

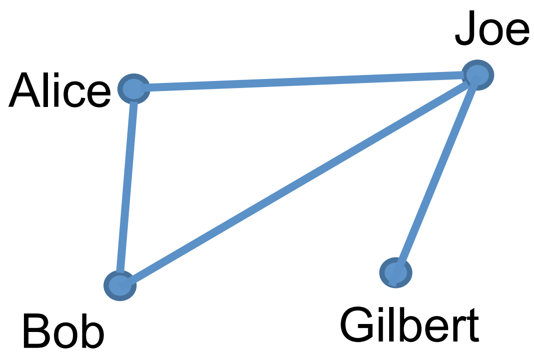




Beschreibungen

Liste: (undirected graph)

Alice	Bob
Alice	Joe
Bob	Joe
Joe	Gilbert



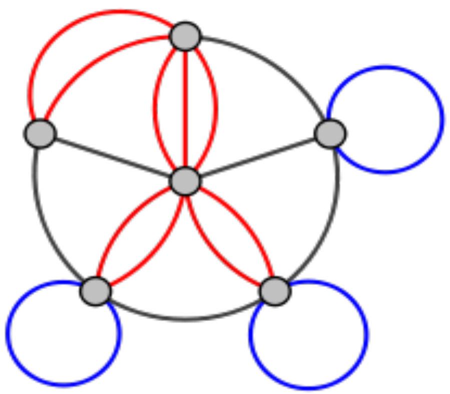
Adjacency matrix

	Alice	Bob	Joe	Gilbert
Alice	0	1	1	0
Bob	1	0	1	0
Joe	1	1	0	1
Gilbert	0	0	1	0

For undirected and unweighted graphs, adjacency matrix is symmetric and consists of 0‘s and 1‘s.

Verallgemeinerungen

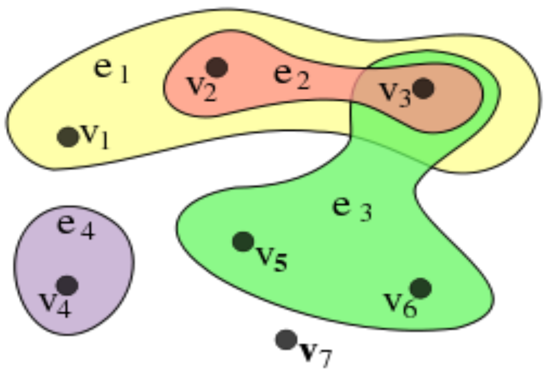
Pseudograph / Multigraph



Hat self-loops und mehrfache Kanten

Hypergraph

Eine Kante verbindet mehrere Knoten



Im folgenden nur behandeln wir nur einfache Graphen

Kurzer Überblick: Netzwerkanalyse

Network Analysis (Overview, biased selection)

Zurich University of Applied Sciences

All these networks (social-, actors, genes, …) are similar and can analyzed with the same tools:

Analysis

- General Graph Theory
  - Shortest Path,…
- **Graph Drawing**
  - Layout Algorithms
- Global Properties
  - Diameter of Graph,
- Centrality Measures
  - E.g. Google's PageRank
- **Community Detection**
  - E.g. Friendships
- Modules / Motifs
  - Pattern in the Network
- Comparison
  - How similar are (sub)networks / Chemoinformatik
- Dynamics on Graphs
  - E.g. spreading of rumors / viruses

Properties of typical networks

- Small worldness
  - Six Steps of Seperation
- Scale-freeness
  - There are hubs of abitrary size (Barabasi 1999)

Generation

- From data
  - Correlation with thresholding
  - Graphical Models (e.g Bayesian networks)
- Random
  - Erdös, Barabasi

Software

- Standalone Tools: Cytoscape, Gephi
- R: **igraph**
- Libraries: Jung, Pajec, Graphviz, **prefuse**

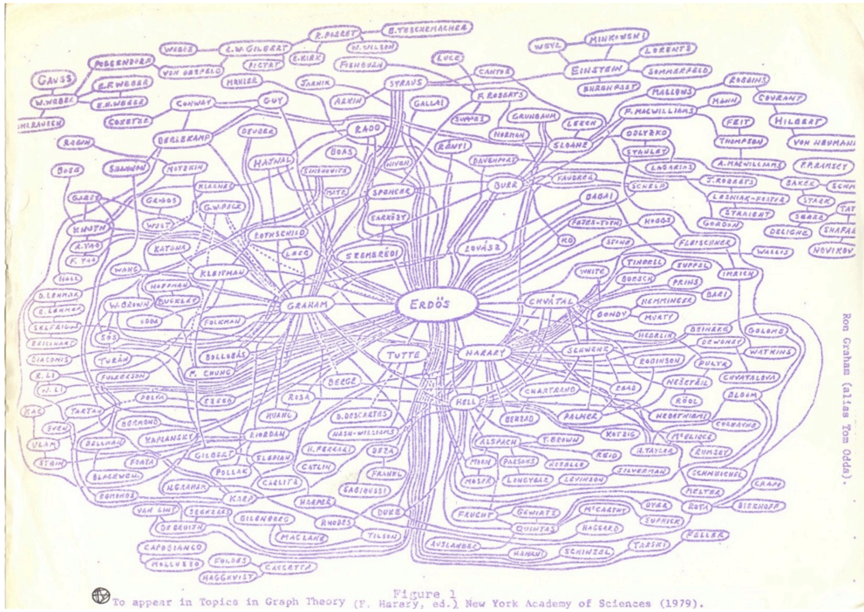
Surprising properties of networks I

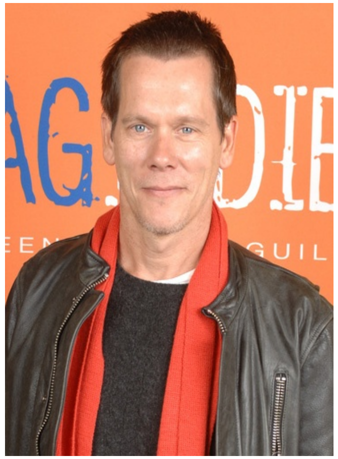
Zurich University of Applied Sciences

(Most) “real world” networks have in common that it just takes a few steps to reach any node

(small-worldness) / “6 degrees of separation”

Average path length ***L~Log(N)***





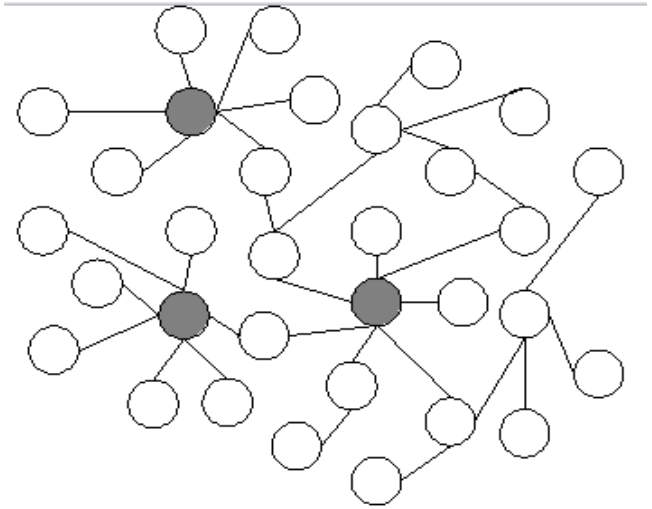
Kevin Bacon at place 507

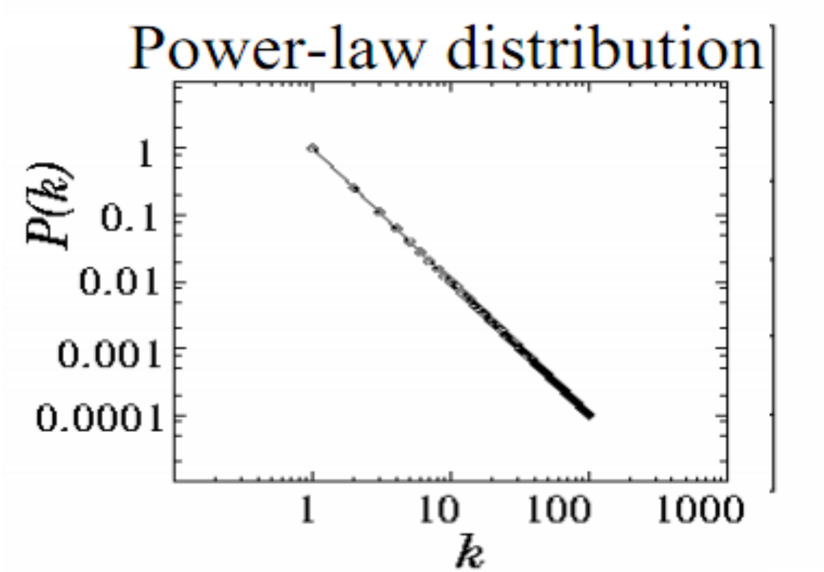
Surprising properties of networks II

Zurich University
of Applied Sciences

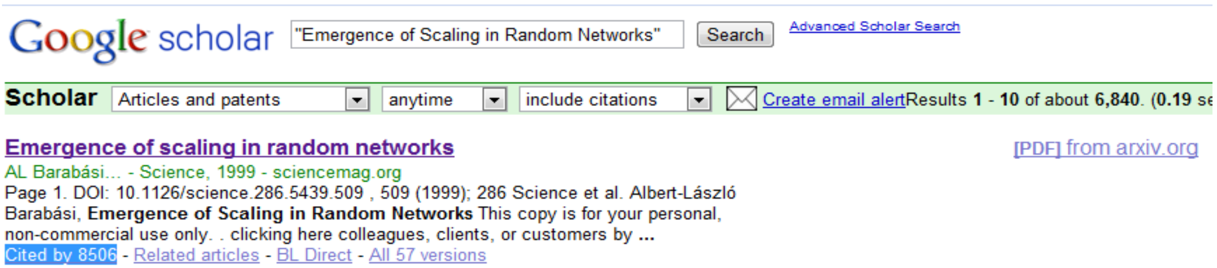
Degree Distribution:

Number of edges k



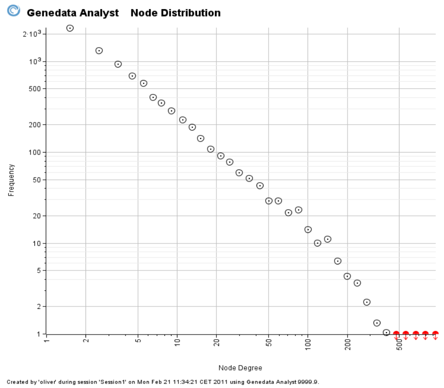


$$P(k) \sim ck^{-\gamma}$$

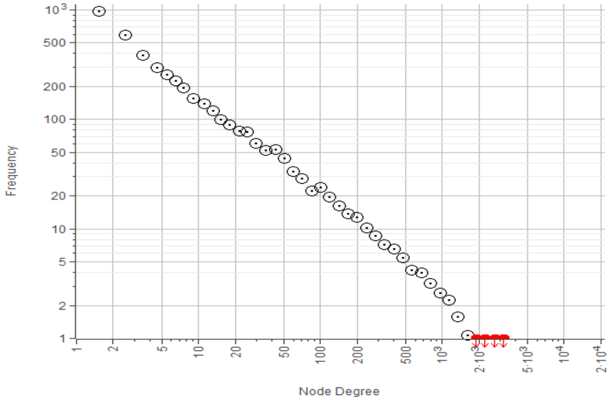


Emergence of scaling in random networks
[PDF] from arxiv.org
AL Barabási... - Science, 1999 - sciencemag.org
Page 1. DOI: 10.1126/science.286.5439.509 , 509 (1999); 286 Science et al. Albert-László Barabási, **Emergence of Scaling in Random Networks** This copy is for your personal, non-commercial use only. . clicking here colleagues, clients, or customers by ...
Cited by 8506 - Related articles - BL Direct - All 57 versions

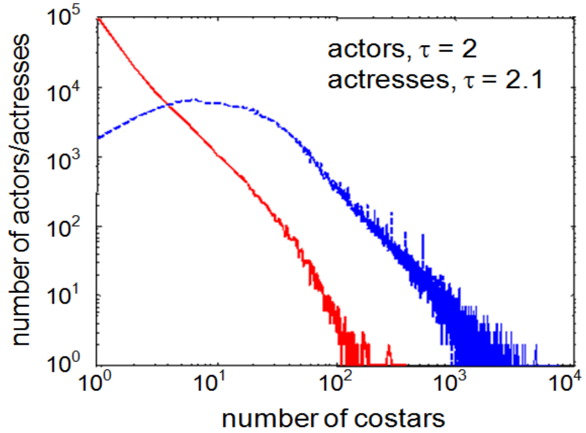
Examples of networks



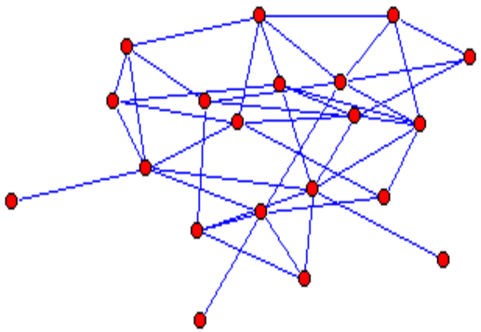
PPI network



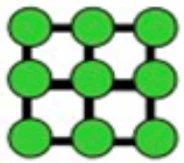
Correlation network



IMDb costarring

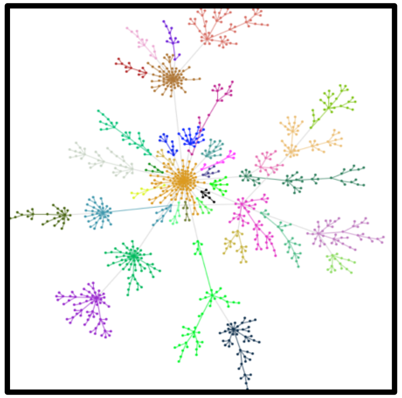


Random Graph (Erdös)
“Small world”
Not scale-free



$$\langle l \rangle \propto N^{1/2}$$

2D Lattice
No “small world”
Not scale-free



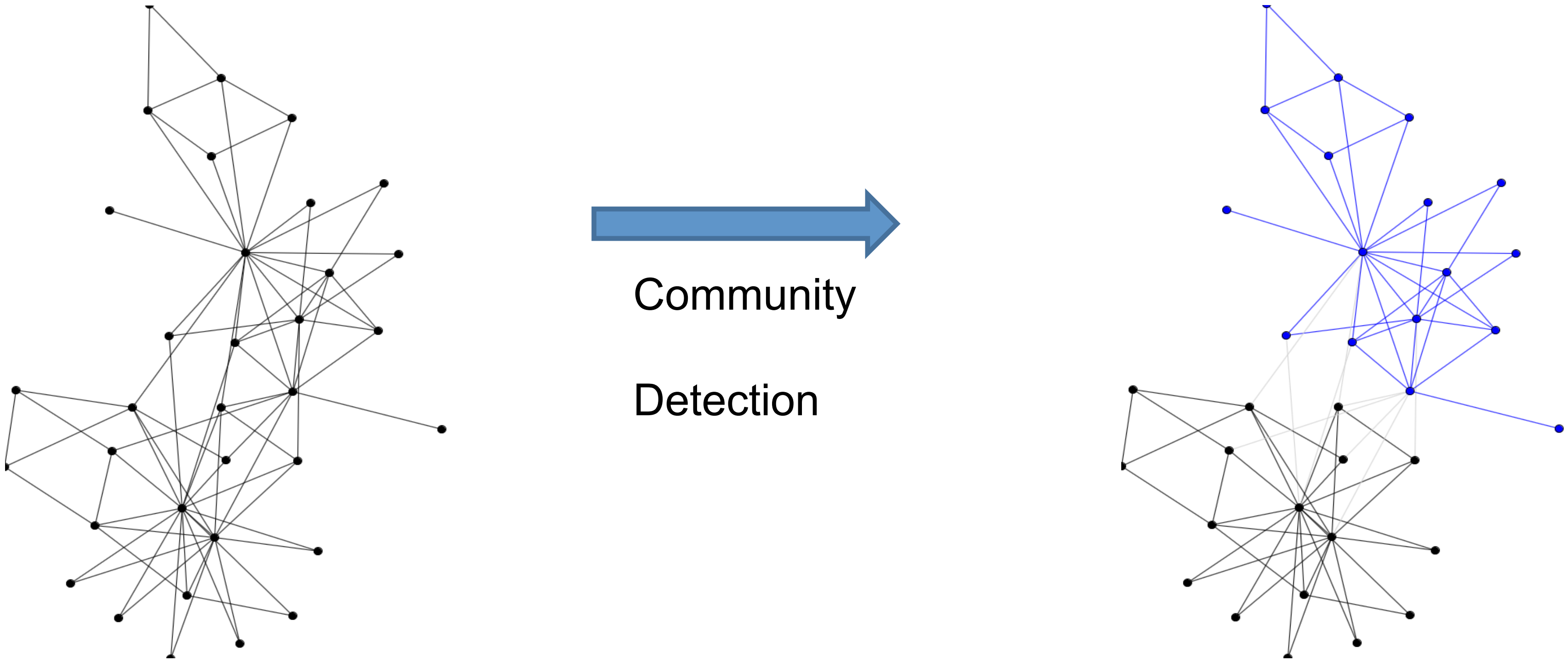
Random Graph (Barabasi)
“Small world”
Scale-free

Graph clustering / Community detection

Zurich University of Applied Sciences

Task:

- Decompose network into parts
- Biological relevance:
  - Given a gene network (PPI, correlation,…) find parts of genes with act together (modules)
- Notion: Many ties (edges) within a community, few between communities



Community detection

Zurich University of Applied Sciences

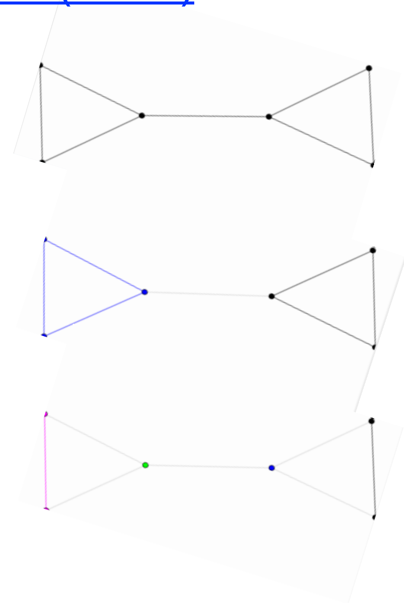
Community detection

- Dozens of different fields
- Hundreds of measures
- Thousands of algorithms

Popular measure: **modularity**

[Newman and Girvan, Phys. Rev. E 69, 026113 (2004)](#)

$$Q = \sum_{\upsilon,\mu} \left(\frac{A_{\mu\upsilon}}{2m} - \frac{k_{\upsilon} \cdot k_{\mu}}{(2m)^2} \right) \delta(c(\mu), c(\upsilon))$$



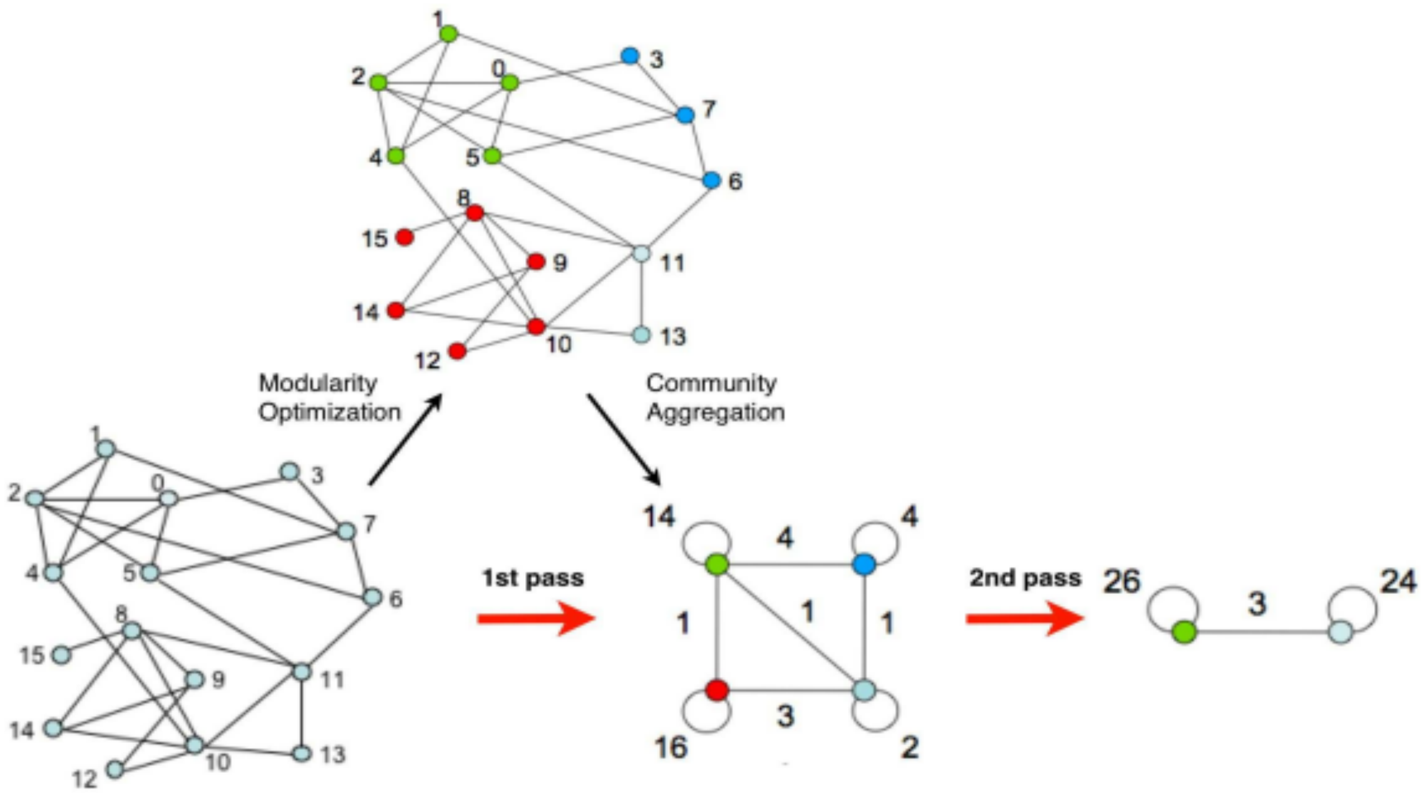
Q=0	1 Community
Q=0.35	2 Communities
Q=0.03	4 Communities

Direct Optimization of Q?

NP complete (Brandes 2008)

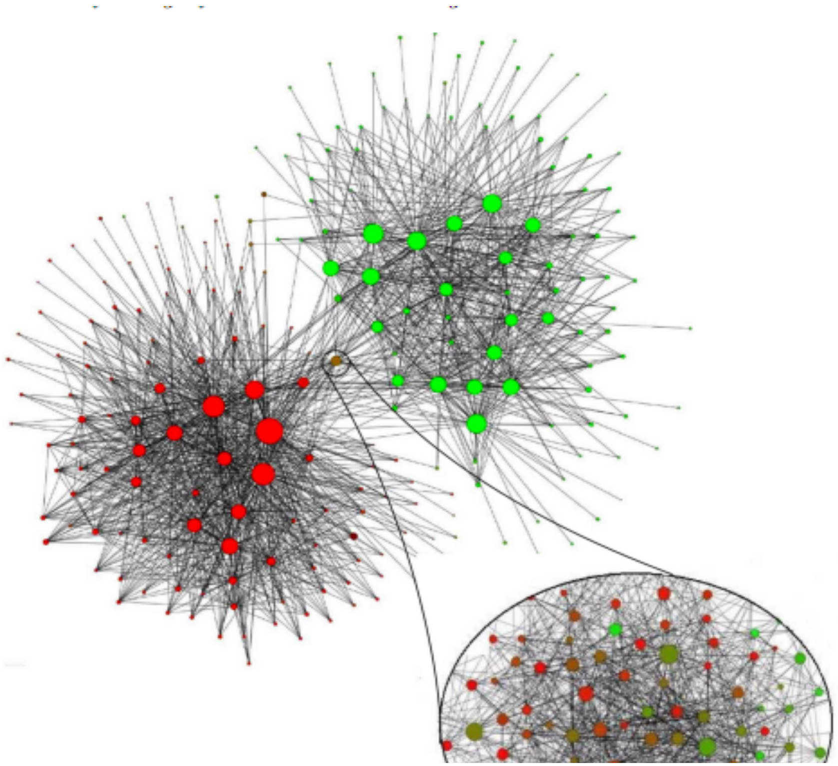
Heuristic Multiscale Optimization of Modularity

Multiscale Heuritic of Blondel et al. 2008 http://arxiv.org/abs/0803.0476v2



Heuristic Multiscale Optimization of Modularity

Multiscale Heuristic of Blondel et al. 2008 http://arxiv.org/abs/0803.0476v2



Network of 2’000’000
mobile phone customers.
Connected if people had a phone call.

Guess the country…

…it’s Belgium

	Karate	Arxiv	Internet	Web nd.edu	Phone	Web uk-2005	Web WebBase 2001
Nodes/links	34/77	9k/24k	70k/351k	325k/1M	2.6M/6.3M	39M/783M	118M/1B
CNM	.38/0s	.772/3.6s	.692/799s	.927/5034s	-/-	-/-	-/-
PL	.42/0s	.757/3.3s	.729/575s	.895/6666s	-/-	-/-	-/-
WT	.42/0s	.761/0.7s	.667/62s	.898/248s	.56/464s	-/-	-/-
Our algorithm	.42/0s	.813/0s	.781/1s	.935/3s	.769/134s	.979/738s	.984/152mn

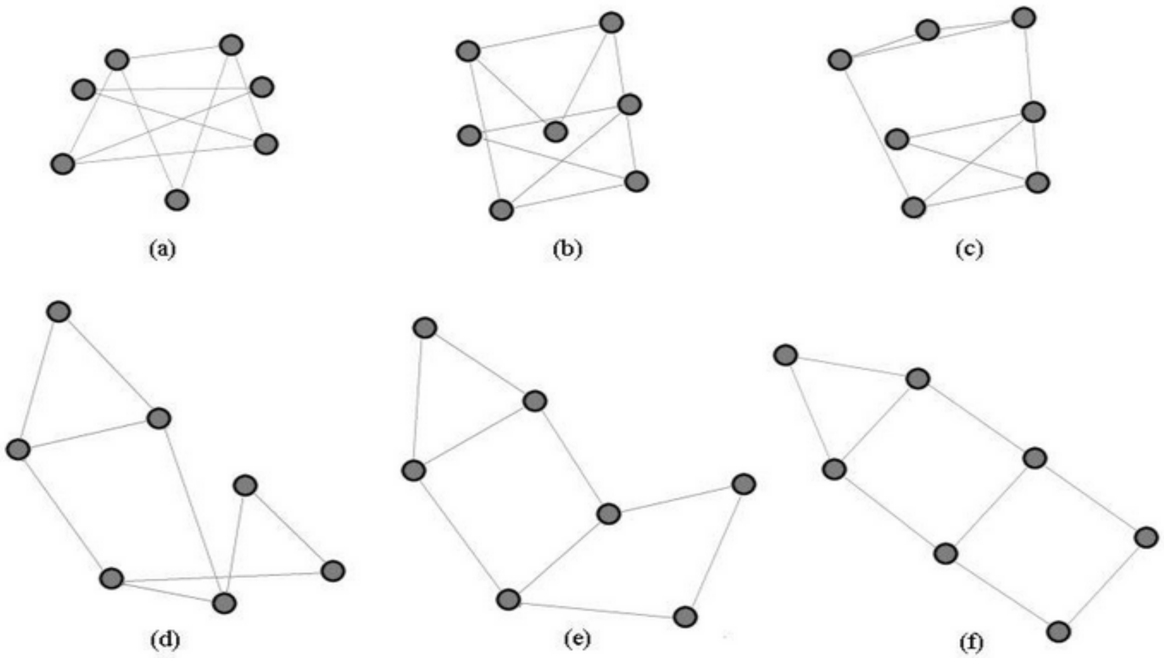
Zurich University
of Applied Sciences

Graph Drawing

Graph Drawing (Definition)

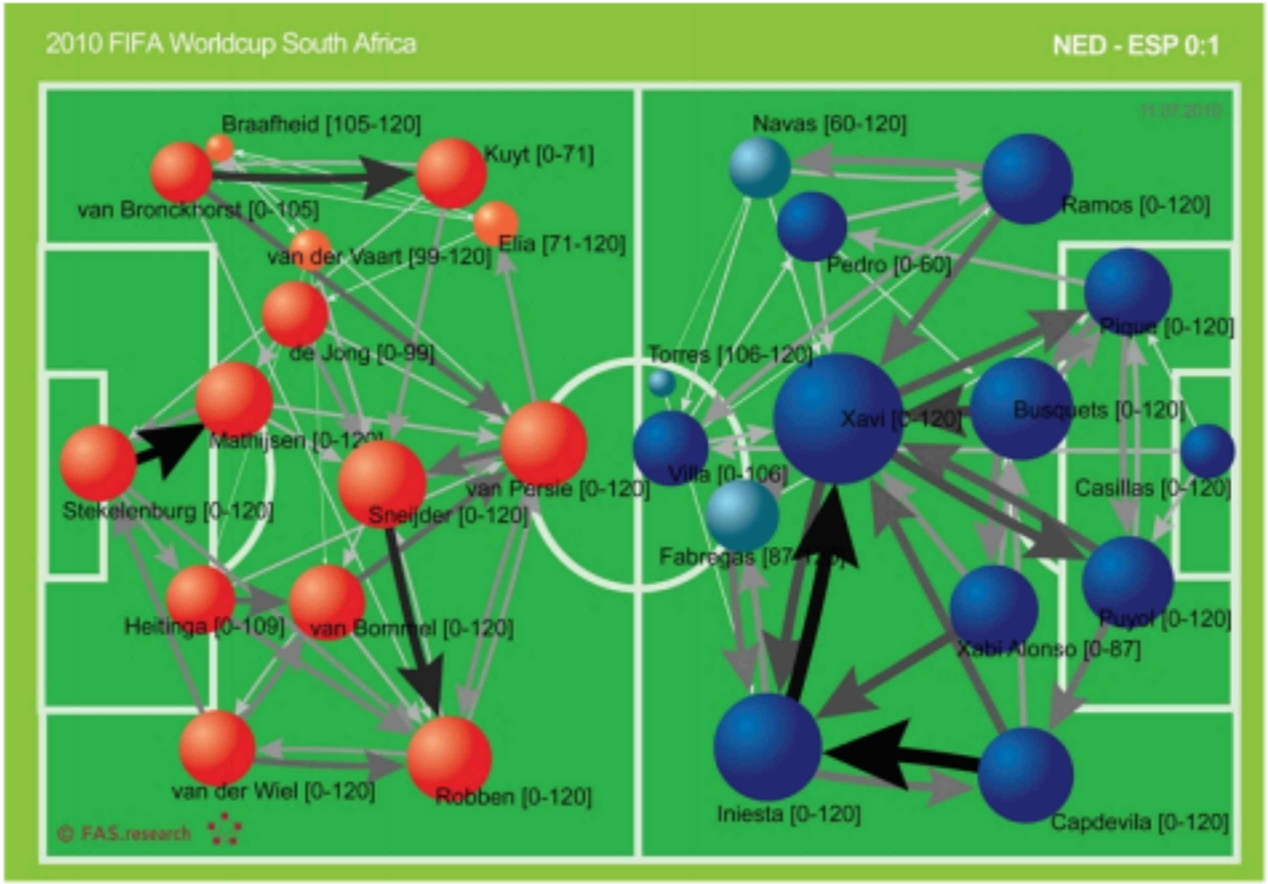
Die Art wie man einen Graph zeichnet ist eine spezielle Darstellung (layout).

Beispiel a-f immer der selbe Graph



Masse für die Qualität: Crossing, Längenverteilung der Kanten, Symetrie, ...

Graph has natural layout

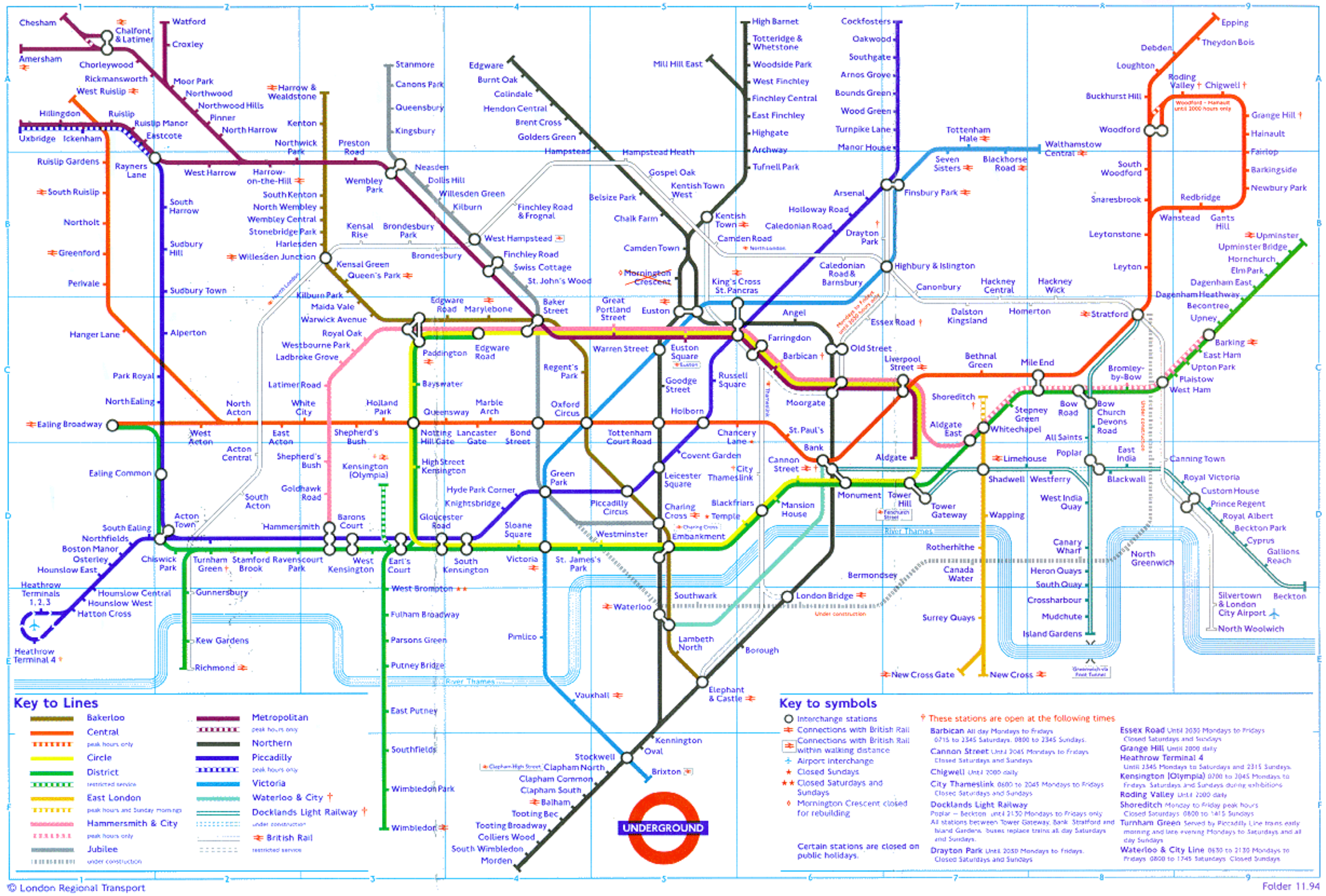


(b) Passes among players during the FIFA World Cup 2010 final. Layout according to (assumed) tactical lineup [PNK10]

Randes in http://cs.brown.edu/~rt/gdhandbook/chapters/social.pdf

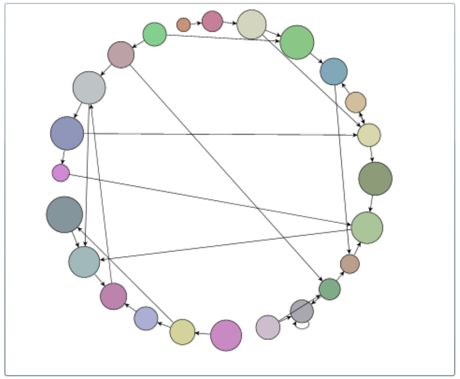
Graph has natural layout

Zurich University of Applied Sciences

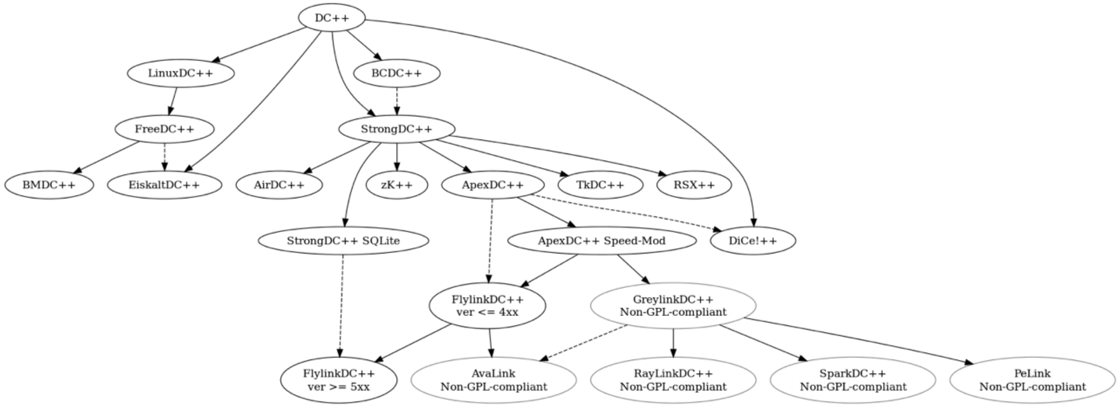


Kinds of Layout (Not Force Directed)

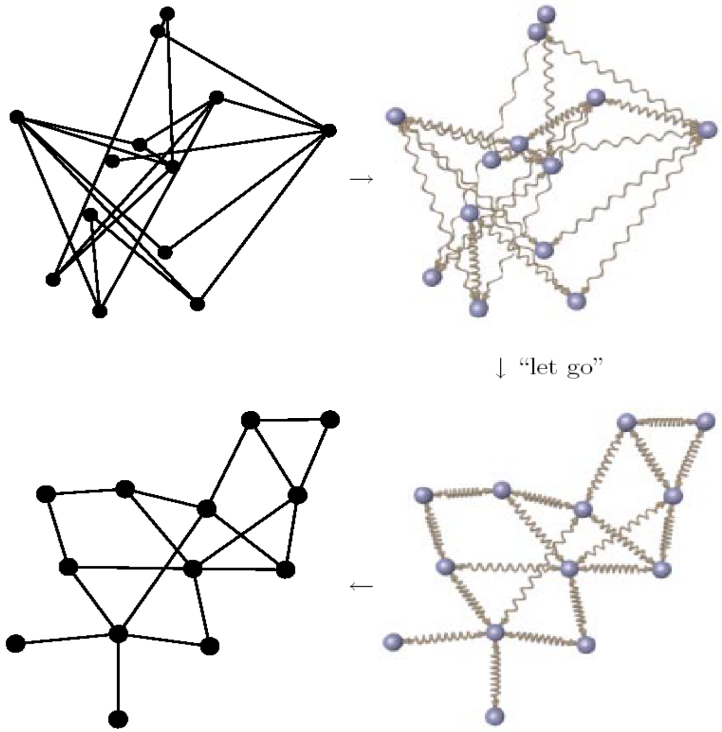
Circular

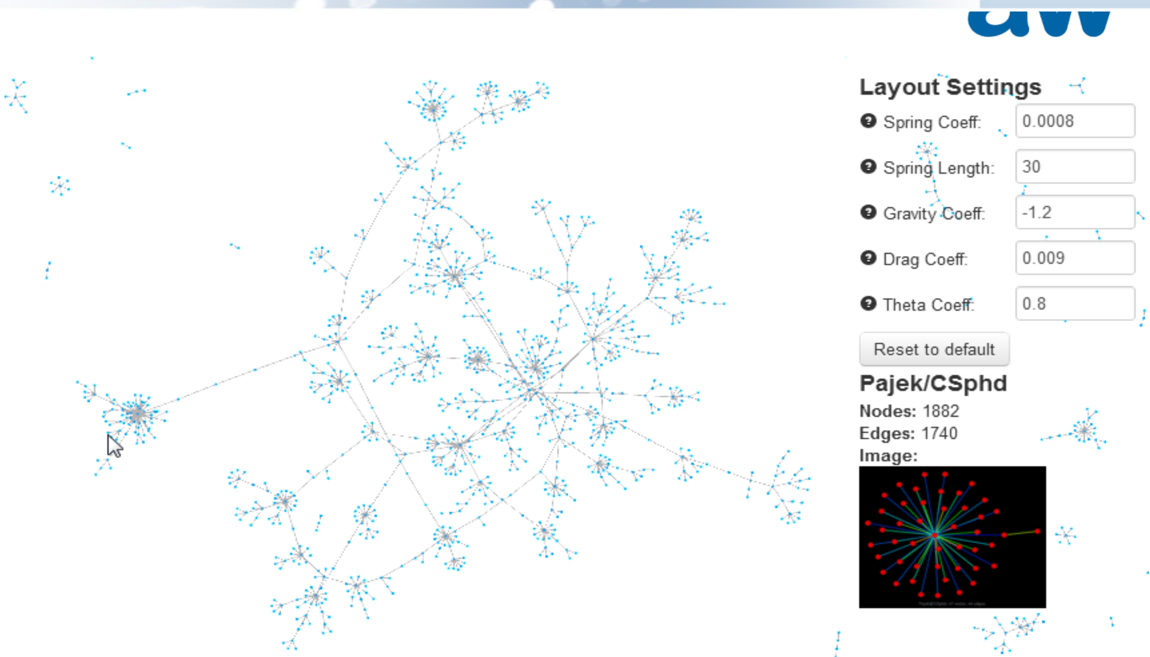


Layered (good for DAG)



Force Directed Layout





http://www.yasiv.com/graphs#Pajek/CSphd

Forces:

– Vertex-Vertex Repulsion: Electrically charged particles (nodes)
– Edges: Attraction: Springs that connect particles via edges;
– (Viscosity to damp movements)

See also:

http://vimeo.com/4356593
http://vimeo.com/3206267

3 historical models

Eades Spring embedder (1984)

- Springs and artificial repulsion forces (logarithmic)

Fruchterman-Reingold (1991)

- Forces artificial springs and repulsion
- Optimization: Steepest decent, step size ~ to temperatur

Kamanda and Kawai (1989)

- Only Springs but with rest length

Details of the Intercation (Springs)

Zurich University
of Applied Sciences

Spring Forces:

Hook's law **||x(from(ei)) - x(to(ei))||** used in prefuse library

FR found that they had more success using a quadratic spring force
||x(from(ei)) - x(to(ei))||^2 better for local minima backed by Jiggle [Jiggle DISS]

Eades **log(||x(from(ei)) - x(to(ei))|| / x0)**,

- Logarithmic spring force leads to an unaesthetically high degree of variance in the edge lengths [Joggle DISS]

Kamada and Kawai (**only Springs**)

- Rest length of the corresponding spring is proportional to **shortestPath(vi, vj)**
- **Force (1 / shortestPath(vi, vj)) • | ||xi - xj|| - c • shortestPath(vi, vj) |**
- Problem n^2 Springs

Details of the interactions (V-V Forces)

Vertex-Vertex Forces

Kamanda Kawai

(no such force) just springs with rest length

Eades, prefuse-library

1/||xi - xj||^2

Fruchterman Reingold

1/||xi - xj||

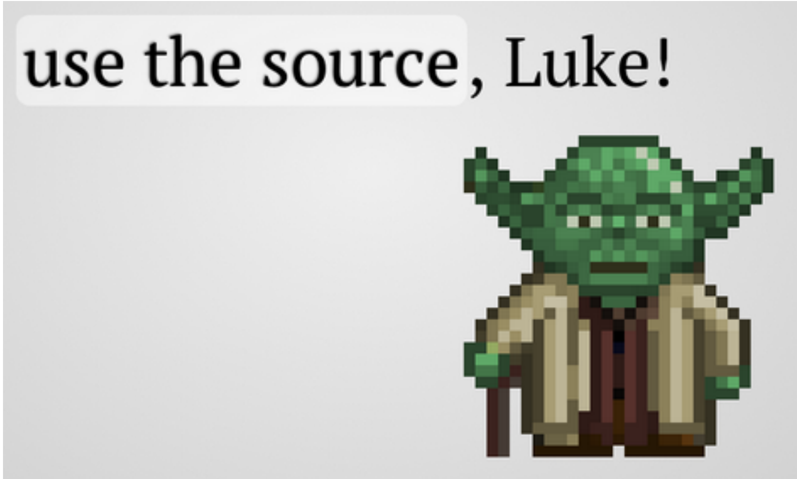
Side remark (Column interaction in 2d log(**||xi - xj|| ➔ Force 1/r** as in FR)

Detail: Forces

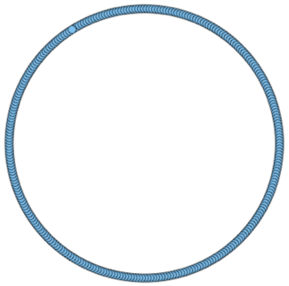
Zurich University
of Applied Sciences

Additional Term in source

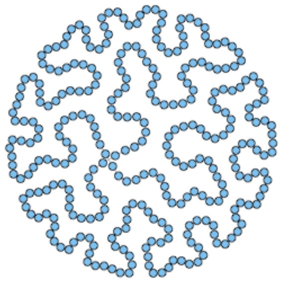
Fruchterman-Reingold in R



Expected



Observe



File: lauout.c method igraph_layout_fruchterman_reingold

```
yd=MATRIX(*res, j, 1)-MATRIX(*res, k, 1);
zd=MATRIX(*res, j, 2)-MATRIX(*res, k, 2);
ded=sqrt(xd*xd+yd*yd+zd*zd);  /*Get dyadic euclidean distance*/
if (ded != 0) {
  xd/=ded;                      /*Rescale differences to length 1*/
  yd/=ded;
  zd/=ded;
  /*Calculate repulsive "force"*/
  rf=frk*frk*(1.0/ded-ded*ded/repulserad);
} else {
  /* ded is exactly zero. Use some small random displacement. */
  xd=RNG_NORMAL(0,0.1);
  yd=RNG_NORMAL(0,0.1);
  zd=RNG_NORMAL(0,0.1);
  rf=RNG_NORMAL(0,0.1);
}
MATRIX(dxdydz, j, 0)+=xd*rf;      /*Add to the position change vector*/
MATRIX(dxdydz, k, 0)-=xd*rf;
MATRIX(dxdydz, j, 1)+=yd*rf;
```

An additional undocumented V-V force with r^2

Keeps disconnected graphs together. Switched of by setting parameter repluserad high.

Demo Force Directed

```
# Demo dynamic of layout
g <- graph.ring(100, directed=FALSE)
wt <- multilevel.community(g)
V(g)$color <- wt$membership
l <- layout.random(g)
for (i in 1:100) {
  l <- layout.fruchterman.reingold(g, params=list(niter=5, start=l, repulserad=1e30))
  #l <- layout.fruchterman.reingold(g, params=list(niter=4000, start=l))
  plot(g,layout=l,vertex.size=3, vertex.label=NA, main=paste0("FR ", i))
  Sys.sleep(1);
}
l <- layout.fruchterman.reingold(g, params=list(niter=5000, start=l, repulserad=1e30))
plot(g,layout=l,vertex.size=3, vertex.label=NA, main=paste0("FR ", i))
```

Running Times

Number of Edges: **n**

Time per Iteration θ**(n^2)**

Number of Iterations until convergence (poorly understood) but generally assumed to be **~ n**

Eades / Fruchterman Reingold θ**(n^3)**

Kamanda Kawei

- Different naïve θ**(n^3)**
- Spezialized approach θ**(nm log n)**

Pros and Cons force directed

Zurich University of Applied Sciences

- **Pros (from wikipedia)**
- Good-quality results
  - At least for graphs of medium size (up to 50–100 vertices)… following criteria: **uniform edge length, uniform vertex distribution, symmetry**.
- Flexibility
  - Force-directed algorithms can be easily adapted and extended to fulfill additional aesthetic criteria.
- Simplicity
  - Typical force-directed algorithms are simple and can be implemented in a few lines of code. Other classes of graph-drawing algorithms, like the ones for orthogonal layouts, are usually much more involved.
- Interactivity
  - …This makes them a preferred choice for dynamic and online graph-drawing systems.
- Strong theoretical foundations
  - …statisticians have been solving similar problems in multidimensional scaling (MDS) since the 1930s, and physicists also have a long history of working with related n-body problems –

Ways to do Graph Drawing (Force Directed)

Zurich University of Applied Sciences

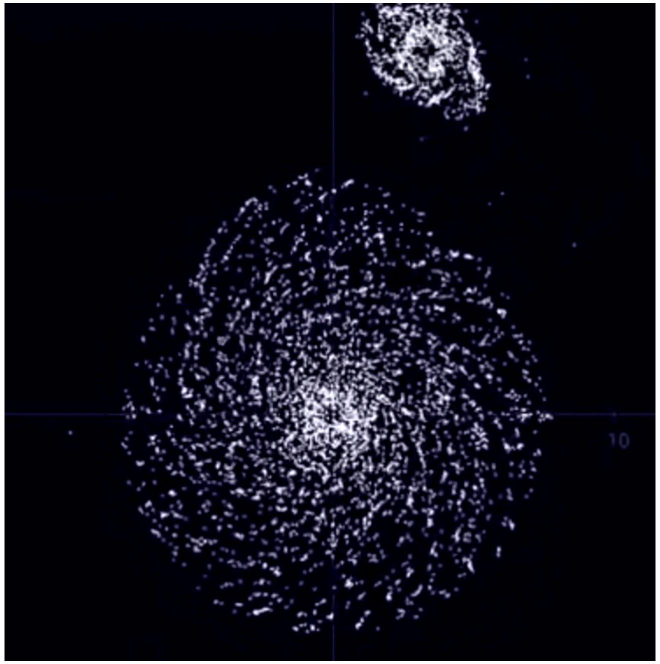
Cons (from wikipedia)

- High running time
  - The typical force-directed algorithms are in general considered to have a running time equivalent to O(n3), where n is the number of nodes of the input graph. This is because the number of iterations is estimated to be O(n), and in every iteration, all pairs of nodes need to be visited and their mutual repulsive forces computed O(n2).
    Solution: Barnes-Hut Approximation.
- Poor local minima
  - …The problem of poor local minima becomes more important as the number of vertices of the graph increases. For example, using the Kamada–Kawai algorithm[10] to quickly generate a reasonable initial layout and then the Fruchterman–Reingold algorithm[11] to improve the placement of neighbouring nodes. Another technique to achieve a global minimum is to use a **multilevel approach**.

Lets tackle the cons…

How to draw it faster I (Barnes-Hut)

two galaxies colliding (Nice Movie)



Particels which are far enough away can be approximated by the center of mass. Stored on the nodes in the search tree

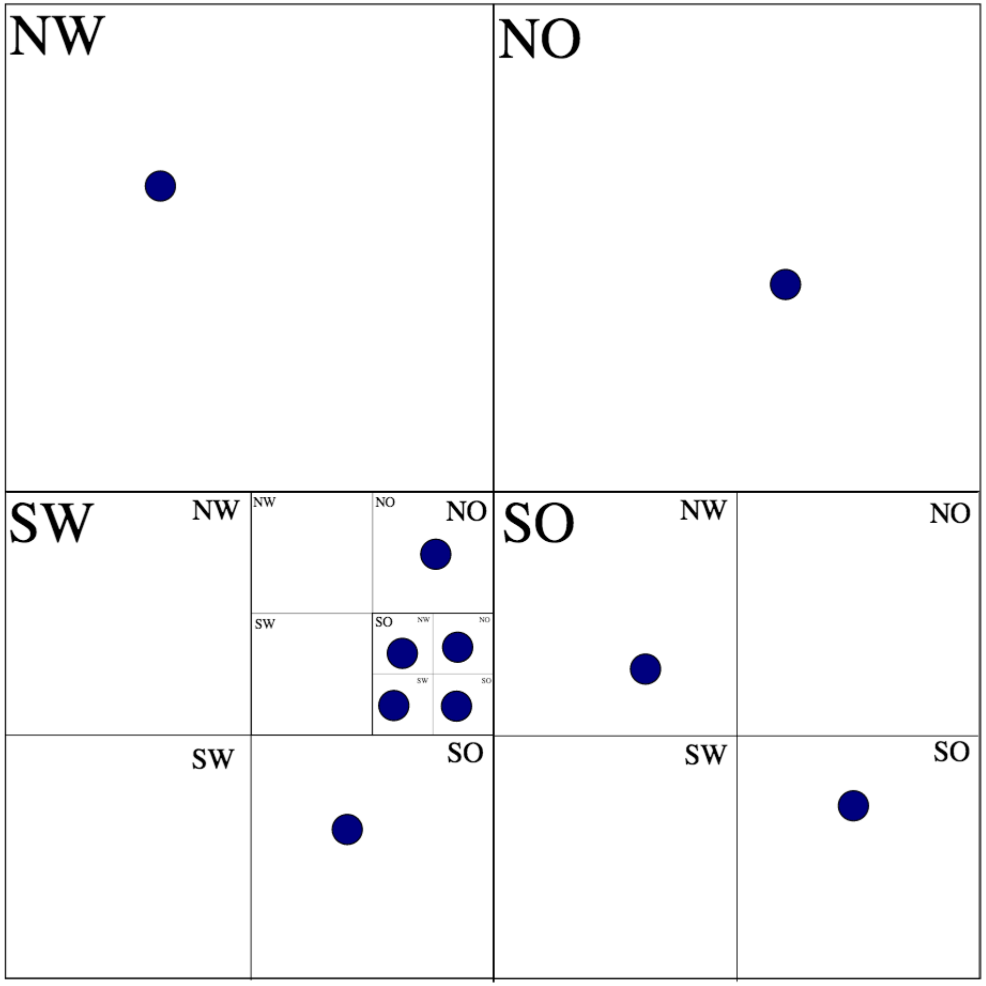
Has been used for simulating e.g. many particel problems.

How to draw it faster I (Barnes-Hut)

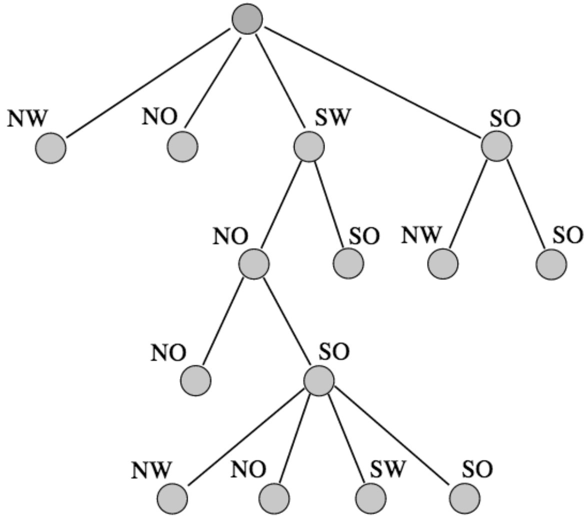
Zurich University
of Applied Sciences

Barnes-Hut Approximation (from:O(n^2) to O(n log n))

In space



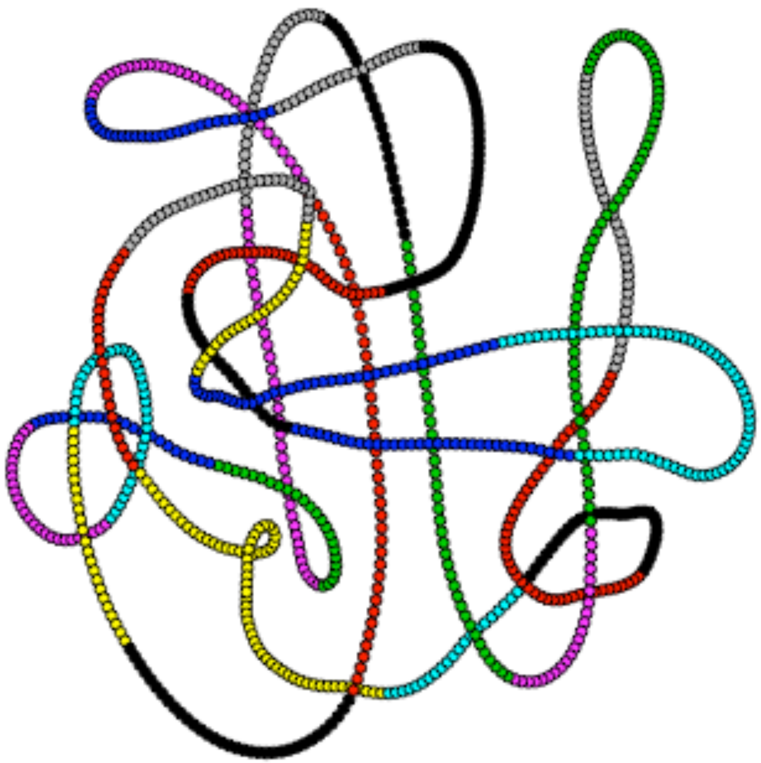
The search Tree (quadtree)



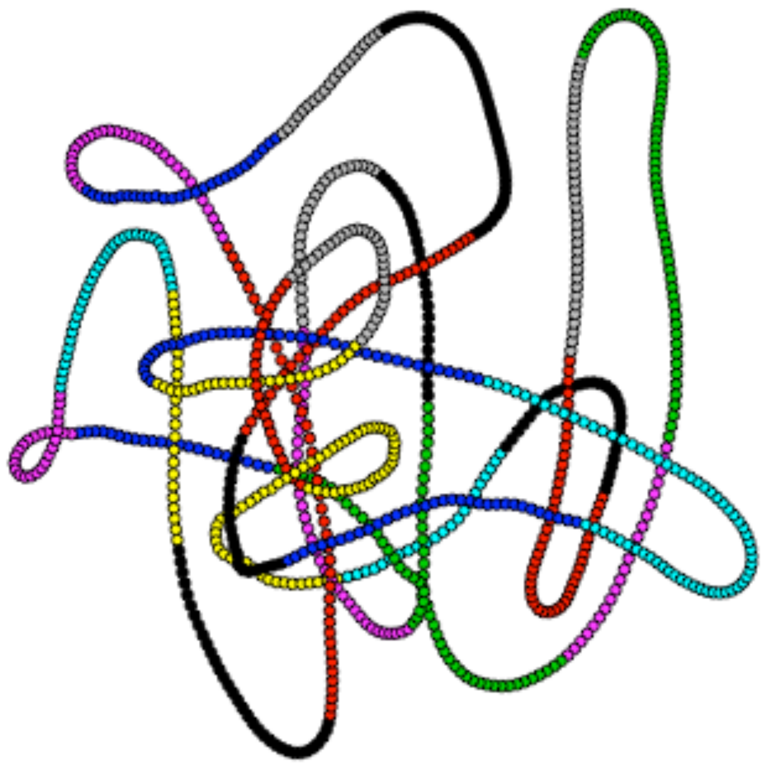
Barnes Hut vs. no approximation

Zurich University
of Applied Sciences

No BH 30 sec



BH 0.8 5.5 sec



Similar layout in 1/6 of the time.

Scaling n log(n) vs. n^2

How to draw it faster II (Multilevel)

Zurich University
of Applied Sciences

Problem: unfolding takes quite some time

Simple Force Directed
http://www.youtube.com/watch?v=-K5zTCrQ_wc&feature=plcp

Idea: Decompose graph into various small parts and do a successively layout.
Walshaw 2000

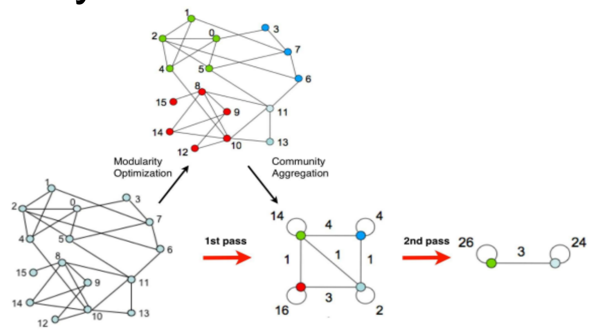
- Input: $G^0 = (V;E)$ with random initial placements

> Multilevel: Coarsen $G^0 \rightarrow (G^0, G^1, G^2 \ldots G^{k-1}, G^k)$
> For i=k to i=0,
>   Compute the layout of G^i /* On GPU */
>   Interpolate the initial positions of G^{i-1}

- Output: $G^0 = (V'; E)$ with final placements

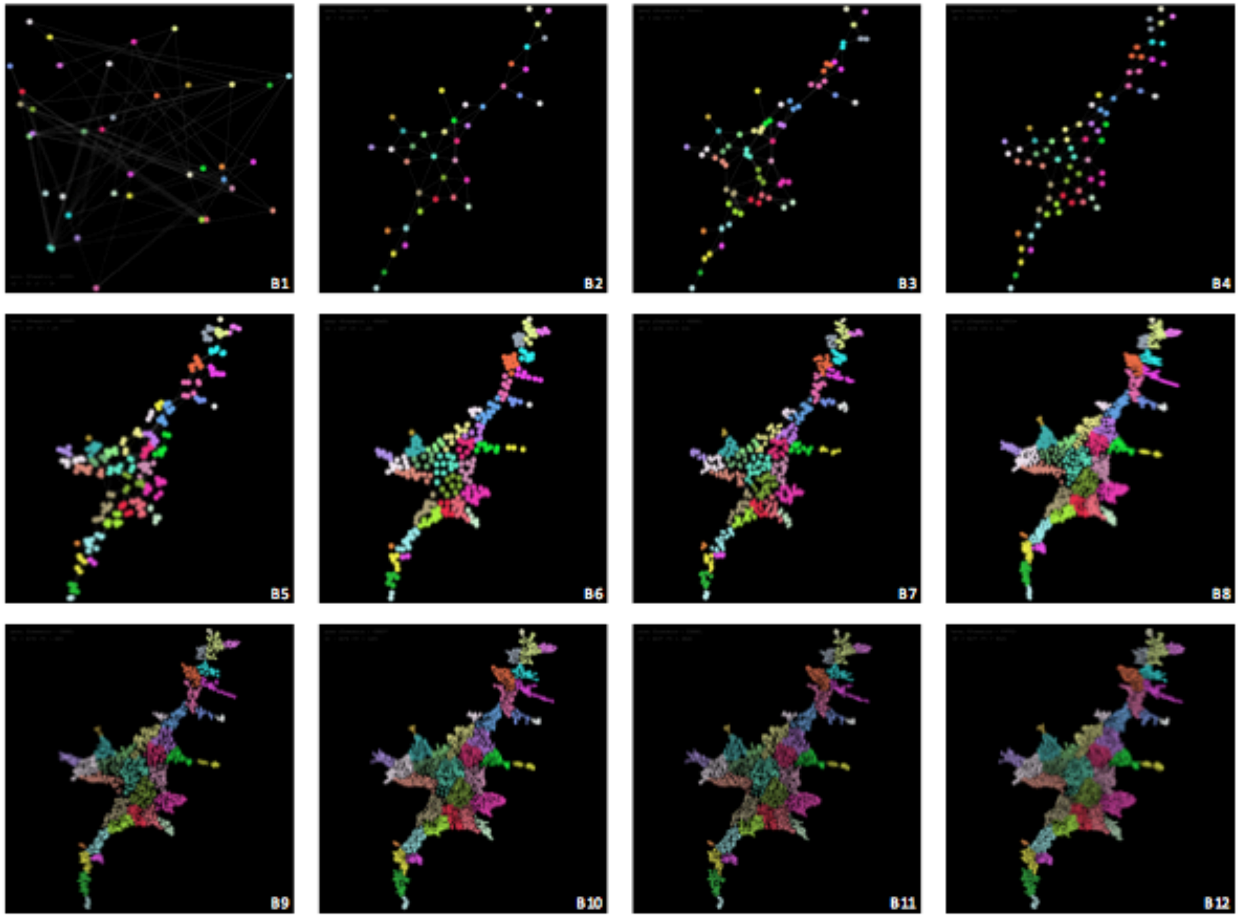
Different ways to coarsen:
See e.g. An Experimental Evaluation of Multilevel Layout Methods
Bartel et al 2012

Why not take to coarsen?



ML layout algorithm at work

Zurich University
of Applied Sciences



Multilevel Layout
http://www.youtube.com/watch?v=J_wkNESO65k&feature=plcp