Very short introduction into the Bielefeld view on NEMO

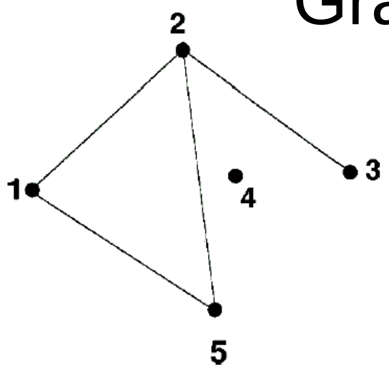
www.AndreasKrueger.de/networks

IRU day

Dublin 1.7.2009

1

Graph G=(V,E)



- KNOTEN, node, vertex, actor, point, individual ...
- KANTE, edge, line, bond, tie, connection...

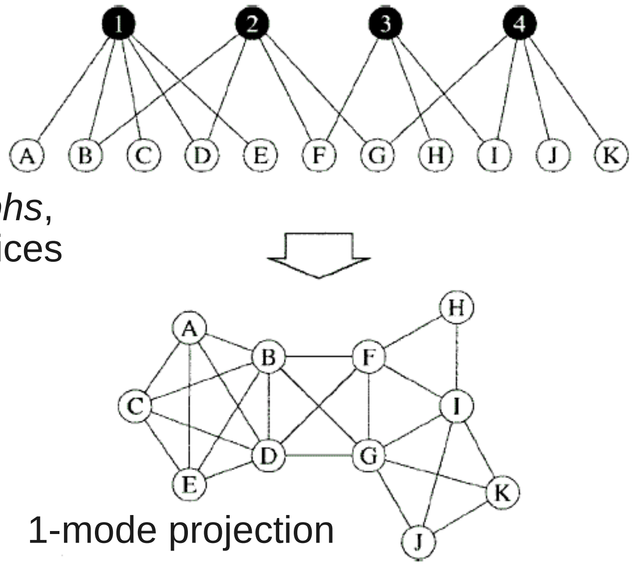
V $\subseteq Z^+$ elements represented by dots
e.g. **V**={1,2,3,4,5}

E $\subseteq \binom{V}{2}$, elements represented by lines
e.g. **E**={ (1,2) , (1,5) , (2,3) , (2,5) }

Neighbours:
x~y := {$\exists$ **e**$\in$**E** with e=(x,y) or e=(y,x)}

2

Bipartite Graphs



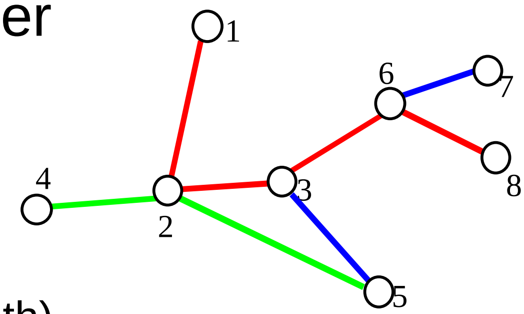
Up to now we have only seen so called *1-mode graphs*, i.e. there is **one** type of vertices

Now imagine for example 4 **films** (black) and 11 playing **Actors** (white).

From the 2-mode graph we can generate a 1-mode graph by projection (under information loss)

FIG. 14. A schematic representation of a bipartite graph, such as the graph of movies and the actors who have appeared in them. In this small graph we have four movies, labeled 1 to 4, and eleven actors, labeled *A* to *K*, with edges joining each movie to the actors in its cast. The bottom figure shows the one-mode projection of the graph for the eleven actors. After Newman, Strogatz, and Watts (2001).

Pfade, Durchmesser

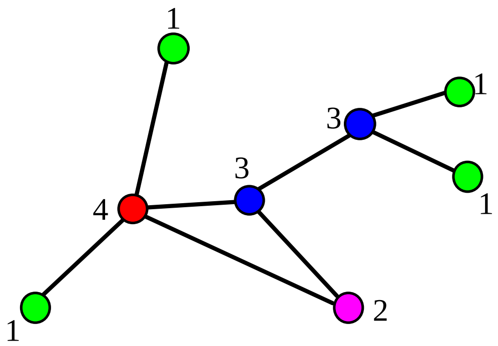


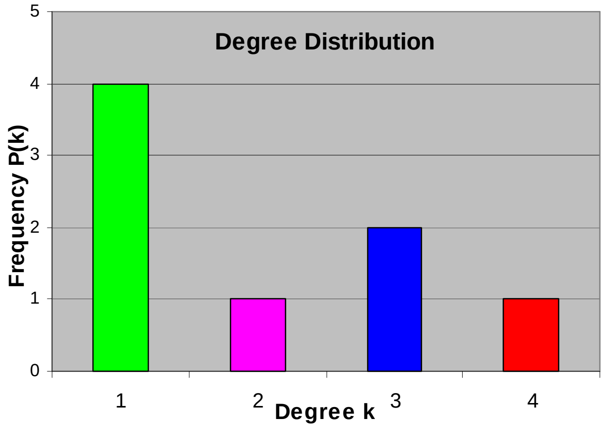
- pathlength (geodesic path)
  - **Shortest** connection between 2 nodes
  - Example pathlength(1,8) = 4
- global graph-properties
  - *Diameter* = **longest** geodesic path (here 4)
  - *characteristic pathlength* = **average** of all paths (i,j)

degree of a node

$k_x = \deg(x) = |N_1(x)|$
$= \text{number of } N_1\text{-Neighbours of node } x$

P(k) = *Degree-Distribution (frequency)*
= **number of nodes** with deg=k





Degree Distribution of ER G(N,p) is ~ Poisson

The average is good estimator
for the whole distribution (bellshaped)

$$\langle k \rangle = (N-1)p$$
$$= (N-1)\frac{M}{N(N-1)/2} = \frac{2M}{N}$$
$$= \mu$$

The degree has a binomial distribution. For N>>1 it becomes Poissonian:

$$P(k) = e^{-\mu}\frac{\mu^k}{k!}$$

with an exponential tail for large k

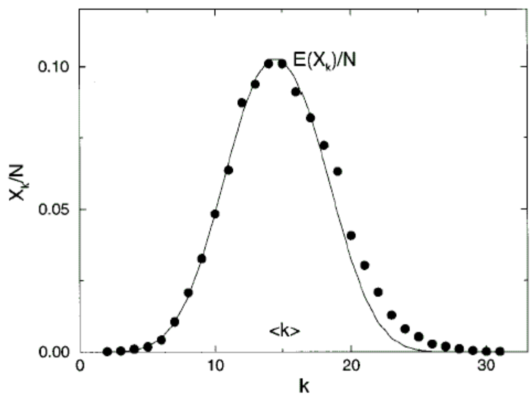
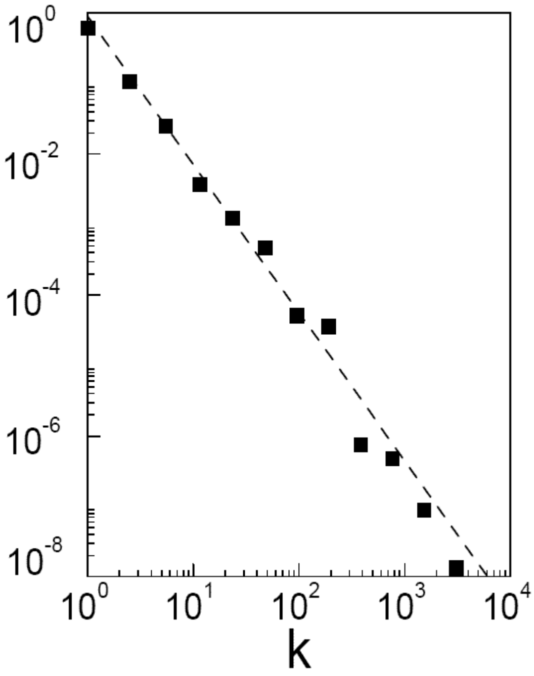


FIG. 7. The degree distribution that results from the numerical simulation of a random graph. We generated a single random graph with $N = 10\,000$ nodes and connection probability $p = 0.0015$, and calculated the number of nodes with degree k, X_k. The plot compares X_k/N with the expectation value of the Poisson distribution (13), $E(X_k)/N = P(k_i = k)$, and we can see that the deviation is small.

Empirical Property 1: scale free



WWW-Ausschnitt N=325729
kmean=5.46 gamma=2.1 (condmat9910332)

In MEASURED networks, the degree distribution is **not** Poissonian (with exponential tail) for large k

but "fat tail"
→ falling **power-law**

$$P(k) = \frac{1}{k^\gamma}$$

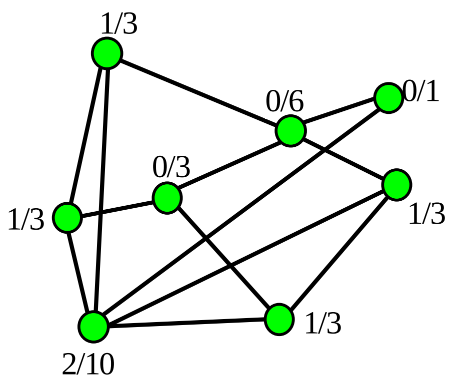
$$\gamma \sim 2.5$$

An average <k> doesn‘t really make sense here = no *built-in scale*

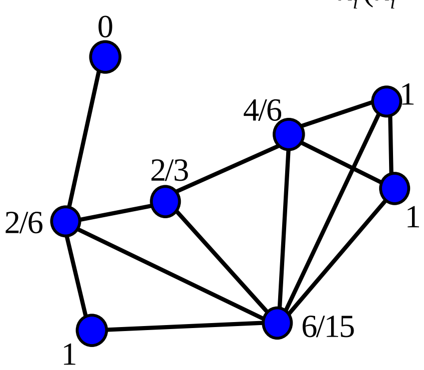
→ „scale-free“

high clustering

$$C_i = \frac{\#T_i}{k_i(k_i-1)/2}$$



C=0.1917



C=0.6333

In both cases M=13 and N=8, but in the *right* picture many more friends are themselves direct friends to each other

! “Empirical Networks” have a significantly **higher clustering-coefficient** than ErdosRenyi-RandomGraphs !

O-graph

Organisations Projection

graph characteristic	FP1	FP2	FP3	FP4
# vertices: N	2500	6135	9615	20873
(N for larg. comp.)	(2038)	(5875)	(8920)	(20130)
N outside larg.comp.	462	260	695	743
# edges: M	9557	64300	113693	199965
(# edges M larg.comp.)	(9410)	(64162)	(113219)	(199182)
mean degree: $\bar{d}$	7.65	20.96	23.65	19.16
($\bar{d}$ larg.comp.)	(9.23)	(21.84)	(25.39)	(19.79)
maximal degree: $d_{\max}$	140	386	648	649
mean triangles per vertex: $\triangle$	22.90	169.70	244.91	146.04
($\triangle$ larg.comp.)	(27.97)	177.16	263.84	151.26
maximal triangle-number	966	5295	15128	10730
cluster coefficient: $\bar{C}$	0.57	0.72	0.72	0.79
($\bar{C}$ larg. comp.)	(0.67)	(0.74)	(0.75)	(0.81)
number of components	369	183	455	467
diameter of largest component	9	7	9	10
mean path length: λ of l.c.	3.70	3.27	3.32	3.59
exponent of degree distribution	-2.1	-2.0	-2.0	-2.1
variance of degree exponent	0.4	0.3	0.3	0.3
exponent of org-size distr.	-2.1	-1.9	-1.7	-1.8
variance of size exponent	0.5	0.3	0.5	0.3
mean # projects per org: $\mathbb{E}(\lvert O\rvert)$	2.40	4. 87	5.6	6.24
maximal size (max $\lvert O\rvert$)	130	82	138	172

TABLE II: Basic network properties of FP1–4 organizations projection.

9

P-graph

Projects Projection

graph characteristic	FP1	FP2	FP3	FP4
# vertices: N	3283	3884	5528	9087
(N for larg. comp.)	(2764)	(3662)	(5027)	(8566)
N outside larg.comp.	519	222	501	521
# edges: M	51217	94527	202358	348542
(# edges M larg.comp.)	(50940)	(94471)	(202306)	(348474)
mean degree: $\bar{d}$	31.20	48.68	73.20	76.71
($\bar{d}$ larg. comp.)	(36.86)	(51.60)	(80.49)	(81.36)
maximal degree: $d_{\max}$	282	387	917	771
mean triangles per vertex: $\triangle$	774.41	871.19	1970.30	2034.31
($\triangle$ larg.comp.)	919.53	923.98	2167.05	2158.03
maximal triangle-number	12903	11125	37247	41141
cluster coefficient: $\bar{C}$	0.67	0.54	0.44	0.47
($\bar{C}$ larg.comp.)	(0.75)	(0.57)	(0.48)	(0.50)
number of components	369	183	455	467
diameter of largest component	9	7	10	9
mean path length: λ of l.c.	3.24	2.80	2.72	2.80
exponent of degree distribution	(-0.8, -3.4)	(-0.7, -3.3)	(-0.6, -3.7)	(-0.3, -2.2)
variance of degree exponent	(0.4, 3.6)	(0.3, 1.7)	(0.3, 1.4)	(0.2, 0.6)
exponent of proj-size distr.	-3.59	-2.9	-3.2	-4.1
variance of size exponent	0.6	0.4	0.2	0.3
mean # orgs per project: $\mathbb{E}(\lvert P\rvert)$	3.15	3.08	3.22	2.71
maximal size (max $\lvert P\rvert$)	20	44	73	54

TABLE III: Basic network properties of FP1–4 projects projection.

10

Um empirische und generierte Graphen zu vergleichen …

… sehen wir immer *3 Diagramme zusammen*

Empirical Network: „FP3“

3rd framework programme

O=9615 P=5529 M=31380

Project-Sizes: min=1 max=73 mean=5.6
Orgas-Sizes: min=1 max=138 mean=3.2

RandomSet Network

~ same #orgs, #projs as FP3

~ same project sizes as FP3

bipMolloyReed Network

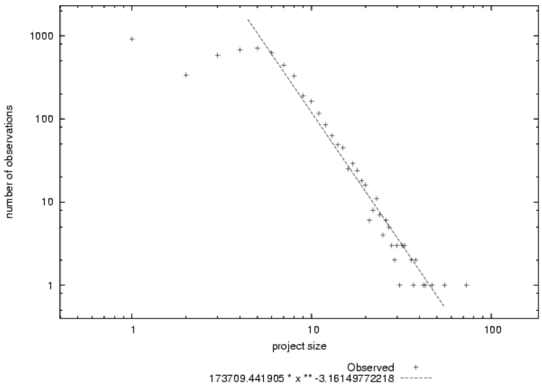
~ same #orgs, #projs as FP3

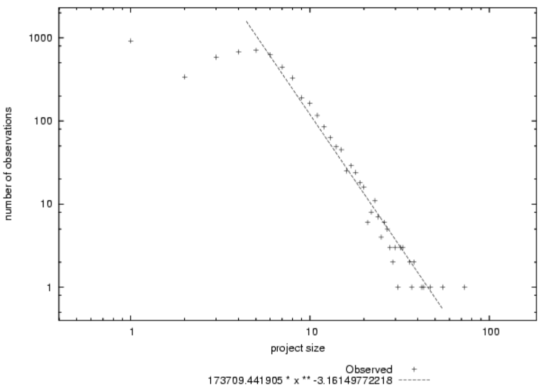
~ same project sizes as FP3
~ same organization sizes as FP3

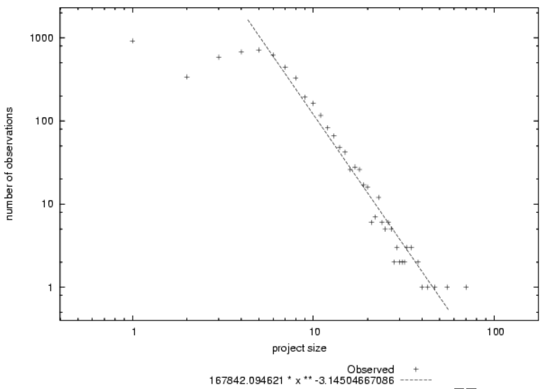
Project **Sizes:**

identical

Because the empirical project sizes are the **inputs** for both simulations



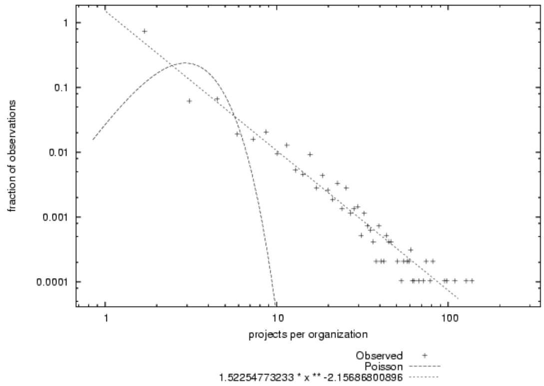


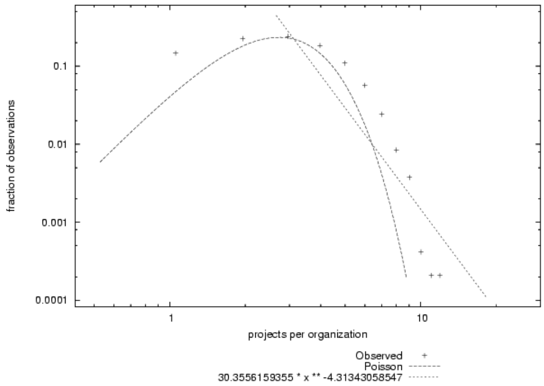


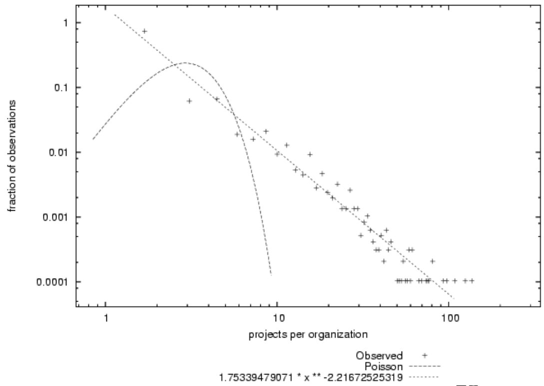
Organization **Sizes**:

RandomSetModel:
Exponential Decay ~ Poisson Distrib.
predicted by theory!

FP3 <-> MolloyReed
identical because input



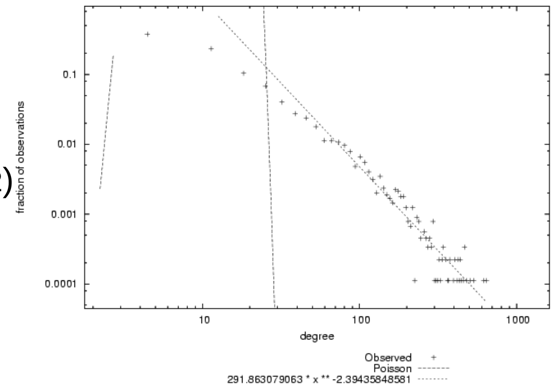


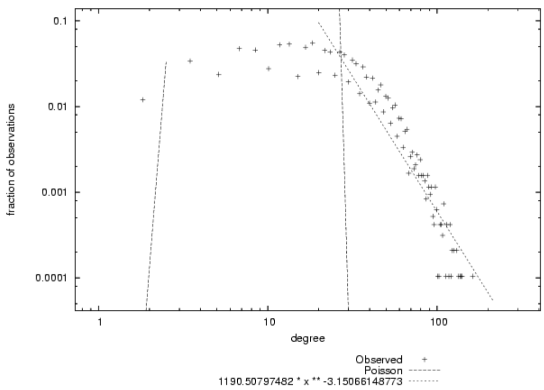


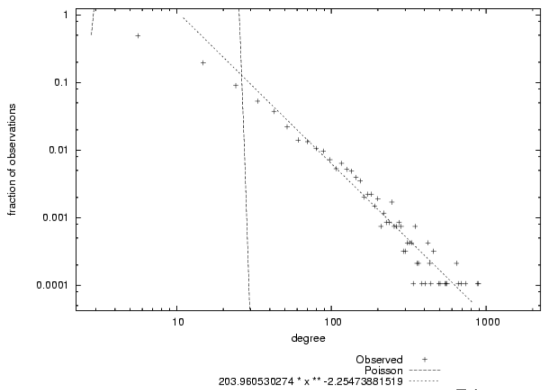
Organization Graph:
Projection onto organisations

RandomSet:
fat tail *but* much steeper (exponent >3.2)

MolloyReed:
similiar to empirical FP3;
both are ~ on a straight power law
with exponent 2.3 - 2.4





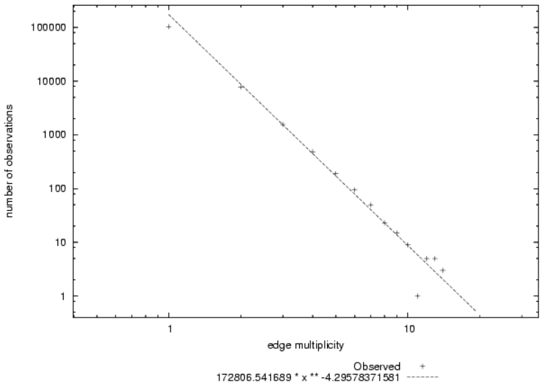


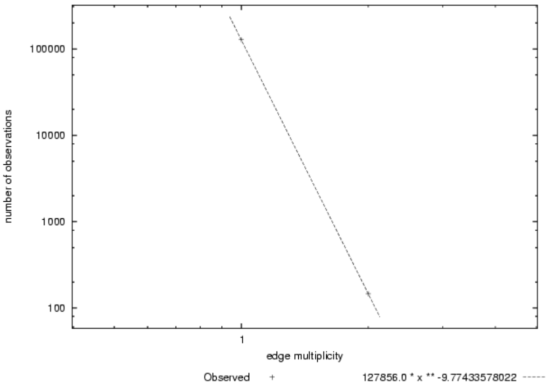
Organization Graph: Edge Multiplicities

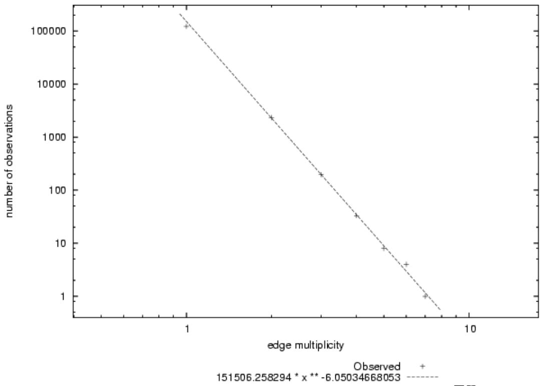
Empirical FP3:
highest multiplicity 14

RandomSet:
only 1 and 2

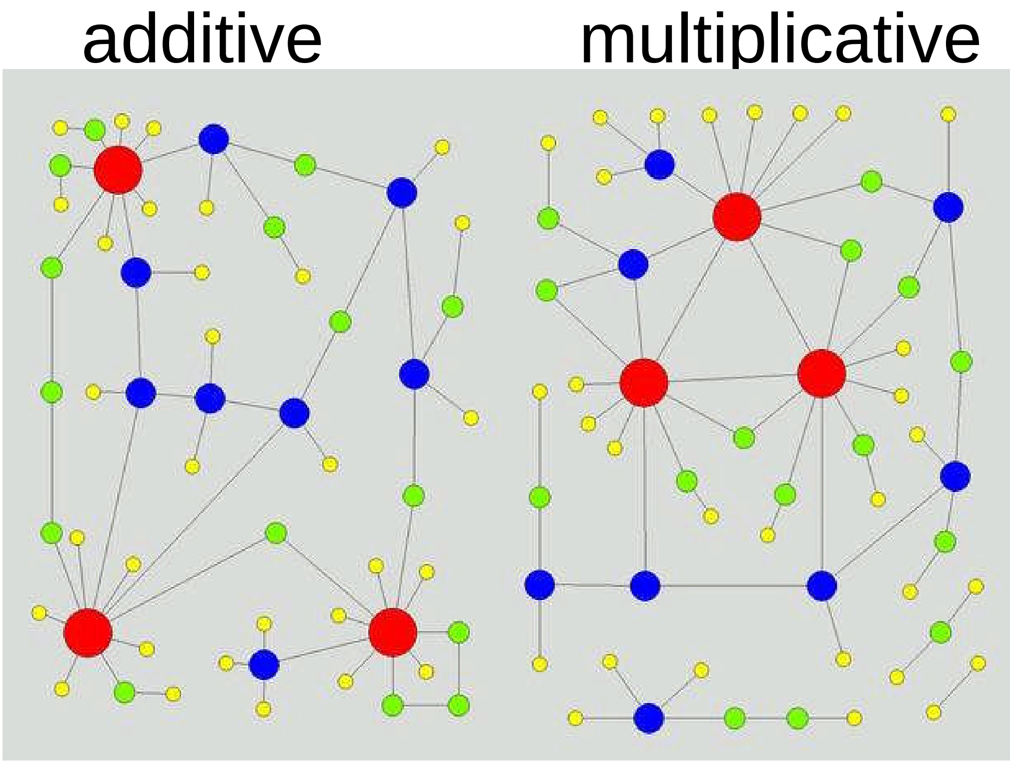
MolloyReed:
highest multiplicity only 7





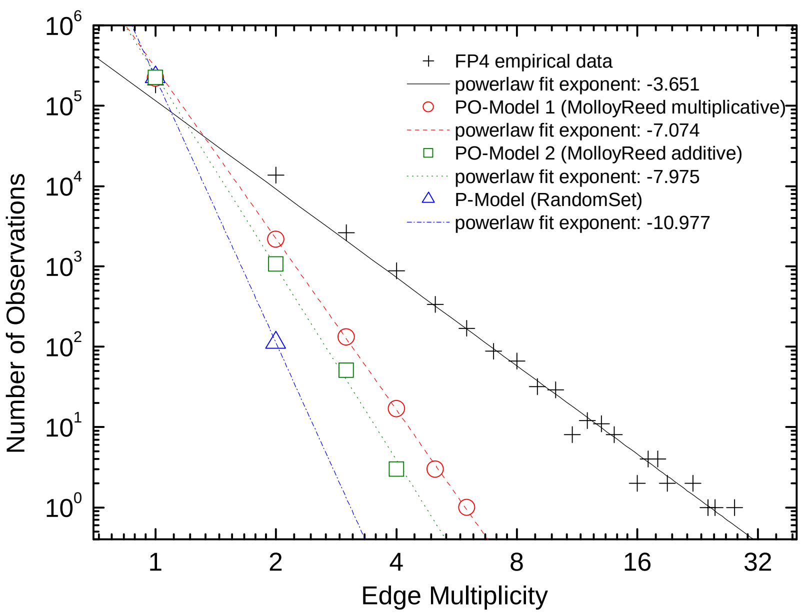


16



17

Multiplicative vs. Additive: OrganisationsGraph Edge Multiplicity



18