

Data - Flickr

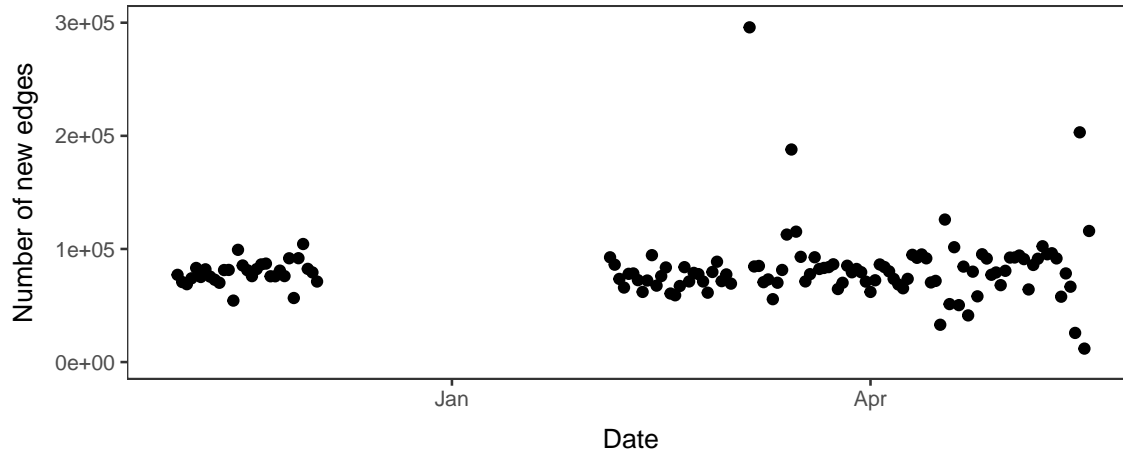
Jan Overgoor

- Source: <http://socialnetworks.mpi-sws.org/data-wosn2008.html>
- Paper: <https://people.mpi-sws.org/~amislove/publications/Growth-WOSN.pdf>

Read data

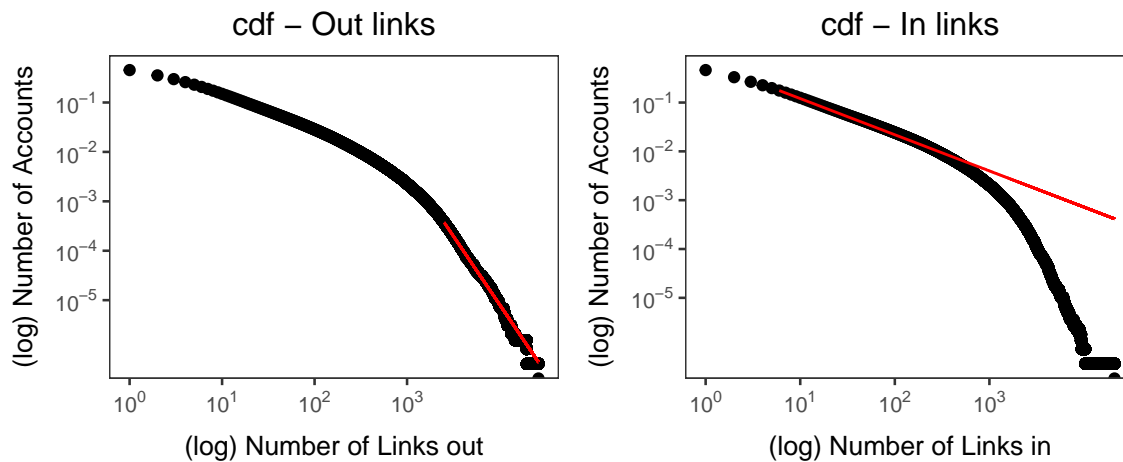
- total number of nodes: 2302925
- total number of edges: 33140018
- avg in-degree 14.3904026
- avg out-degree 14.3904026
- **why are they the same?**
 - about 60% of edges are reciprocated, so not 100%..
 - δ in/out seems reasonable in the extremes (natural limit to out, except for scammers)
 - overflow issue?

Number of connections by day



In/out degree

```
## [1] "plfit:  alpha=3.771  xmin=2516"  
## [1] "plfit:  alpha=1.741  xmin=6"
```



Jackson R

- Fitted $r^* = 0.337$
- Fitted $r = \frac{r^*}{1+r^*} = 0.252$

Model results on 2016-11-05

```
##
## =====
##                                     y
##                (1)          (2)          (3)          (4)          (5)          (6)
## -----
## log In-Degree    1.092***    1.104***                0.635***    0.488***    0.690***
##                  (0.006)    (0.007)                (0.010)    (0.010)    (0.009)
##
## Has degree       -3.584***    -4.040***                -4.174***    -4.957***    -3.613***
##                  (0.045)    (0.050)                (0.059)    (0.136)    (0.052)
##
## Reciprocal                9.192***    8.948***    9.056***    8.878***    9.562***
##                  (0.261)    (0.270)    (0.285)    (0.290)    (0.337)
##
## Is FoF                6.161***    4.724***
##                  (0.049)    (0.057)
##
## 2 Hops                                6.508***
##                                (0.147)
##
## 3 Hops                                2.597***
##                                (0.143)
##
## 4 Hops                                0.315**
##                                (0.156)
##
## 5 Hops                                -0.418**
##                                (0.187)
##
## 6+ Hops                                -1.492***
##                                (0.280)
##
## log Paths                                2.080***
##                                (0.035)
## -----
## Observations      20,001      20,001      20,001      20,001      20,001      20,001
## Log Likelihood -24,774.850 -16,504.120 -13,342.850 -9,493.637 -8,723.253 -11,209.040
## =====
## Note:                                *p<0.1; **p<0.05; ***p<0.01
## [1] "Train accuracy & 0.6351 & 0.7514 & 0.762 & 0.8552 & 0.8575 & 0.8462  \\"
## [1] "Test accuracy & 0.6495 & 0.7655 & 0.7625 & 0.8625 & 0.874 & 0.855  \\"

```

Model results on 2007-03-01

```
##
## =====
##                                     y
##          (1)          (2)          (3)          (4)          (5)          (6)
## -----
## log In-Degree    0.974***    1.007***                0.617***    0.451***    0.684***
##                  (0.005)    (0.006)                (0.008)    (0.009)    (0.008)
##
## Has degree      -3.543***    -4.010***                -4.123***    -4.680***    -3.694***
##                  (0.037)    (0.043)                (0.049)    (0.091)    (0.044)
##
## Reciprocal                9.946***    9.613***    9.712***    9.562***    11.447***
##                  (0.338)    (0.341)    (0.349)    (0.353)    (0.462)
##
## Is FoF                6.051***    4.827***
##                  (0.047)    (0.053)
##
## 2 Hops                                6.539***
##                                (0.109)
##
## 3 Hops                                2.675***
##                                (0.103)
##
## 4 Hops                                0.540***
##                                (0.113)
##
## 5 Hops                                -0.734***
##                                (0.157)
##
## 6+ Hops                                -1.359***
##                                (0.212)
##
## log Paths                                2.338***
##                                (0.039)
## -----
## Observations      20,000      20,000      20,000      20,000      20,000      20,000
## Log Likelihood -30,342.350 -19,951.770 -16,511.730 -11,620.230 -10,530.600 -14,449.670
## =====
## Note:                                *p<0.1; **p<0.05; ***p<0.01
## [1] "Train accuracy & 0.5478 & 0.7007 & 0.726 & 0.824 & 0.8309 & 0.798  \\"
## [1] "Test accuracy  & 0.5545 & 0.7055 & 0.7195 & 0.821 & 0.8225 & 0.8055  \\"

```