

# Data - Flickr

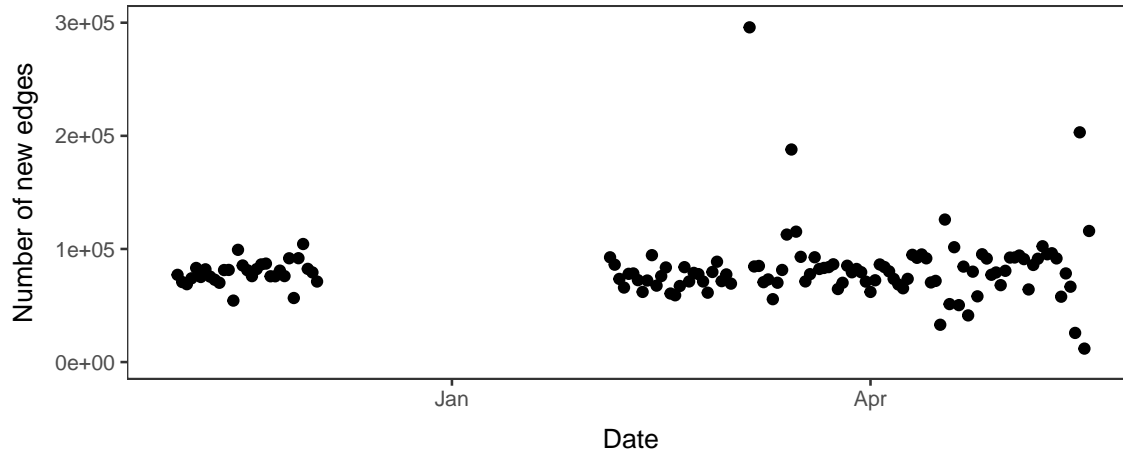
*Jan Overgoor*

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

Read data

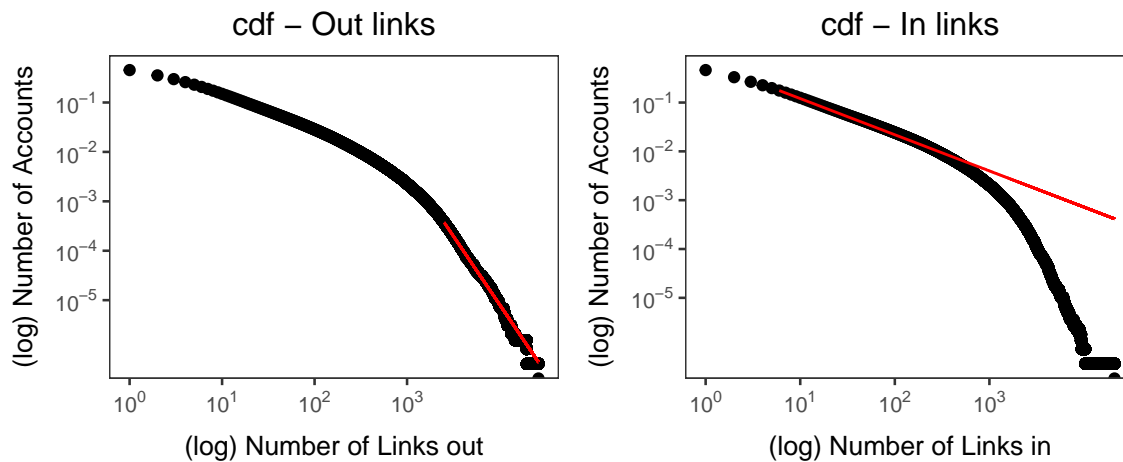
- total number of nodes: 2302925
- total number of edges: 33140018
- avg in-degree 14.3904026
- avg out-degree 14.3904026
- **why are they the same?**
  - about 60% of edges are reciprocated, so not 100%..
  - $\delta$  in/out seems reasonable in the extremes (natural limit to out, except for scammers)
  - 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    1.061***                0.513***    0.432***    0.496***    0.425***
##                  (0.007)                (0.009)    (0.010)    (0.011)    (0.010)
##
## Reciprocal       8.702***    8.948***    8.564***    8.548***    8.779***    9.869***
##                  (0.256)    (0.270)    (0.277)    (0.286)    (0.348)    (0.420)
##
## Is FoF                6.161***    4.605***
##                  (0.049)    (0.054)
##
## 2 Hops                                3.860***
##                                (0.078)
##
## 3 Hops                                -0.128*
##                                (0.066)
##
## 4 Hops                                -2.582***
##                                (0.090)
##
## 5 Hops                                -3.444***
##                                (0.135)
##
## 6+ Hops                               -4.674***
##                                (0.240)
##
## log Hops                                -9.119***
##                                (0.120)
##
## log Paths                                2.761***
##                                (0.040)
## -----
## Observations      20,001      20,001      20,001      20,001      18,516      20,001
## Log Likelihood -19,035.280 -13,342.850 -11,714.450 -10,210.870 -7,050.743 -11,586.390
## =====
## Note:                                                     *p<0.1; **p<0.05; ***p<0.01
## [1] "Train accuracy:"
## [1] "0.7457 & 0.2677 & 0.8348 & 0.8261 & 0.8414 & 0.8653"
## [1] "Test accuracy:"
## [1] "0.7615 & 0.2605 & 0.8487 & 0.8415 & 0.8534 & 0.8756"
```

## Model results on 2007-03-01

```
##
## =====
##                                     y
##                               (1)   (2)   (3)   (4)   (5)   (6)
## -----
## log In-Degree    0.944***           0.471***  0.363***  0.430***  0.421***
##                  (0.006)           (0.008)  (0.009)  (0.010)  (0.008)
##
## Reciprocal       9.607***   9.613***   9.285***   9.362***   9.883***  12.671***
##                  (0.337)   (0.341)   (0.345)   (0.356)   (0.472)  (0.532)
##
## Is FoF           6.051***   4.685***
##                  (0.047)   (0.050)
##
## 2 Hops           4.183***
##                  (0.072)
##
## 3 Hops           0.212***
##                  (0.060)
##
## 4 Hops          -2.159***
##                  (0.073)
##
## 5 Hops          -3.604***
##                  (0.128)
##
## 6+ Hops         -4.416***
##                  (0.184)
##
## log Hops                -8.936***
##                          (0.108)
##
## log Paths                3.192***
##                          (0.043)
## -----
## Observations      20,000      20,000      20,000      20,000      17,565      20,000
## Log Likelihood -23,548.080 -16,511.730 -14,822.580 -12,789.410 -7,819.970 -15,116.650
## =====
## Note:                                     *p<0.1; **p<0.05; ***p<0.01
## [1] "Train accuracy:"
## [1] "0.6886 & 0.1979 & 0.7955 & 0.7868 & 0.8126 & 0.8236"
## [1] "Test accuracy:"
## [1] "0.6975 & 0.1945 & 0.7979 & 0.7801 & 0.8073 & 0.8288"
```