freshmen social network

Yun-Tsz Tsai

Data handling

The original data (a) comes in the format of a matrix, (b) is a realtively small data set with 38 nodes, and therefore I decided to keep it the way it is and do not turn it into an edge list. In addition, the original data contains label that divided the ties between individual into 5 categories, with 1~3 represent positive relationship, 4 represents neutral, and 5 represents negative ties. After some explorations, for period 4, I only keep the positivve ties for the sake of easier interpretation.

Since the analysis aims to examine the differences of the friendship network at different periods of time, I do not remove any isolated nodes for comparison.

Initial exploration of the network

Friendship Network at Time 1
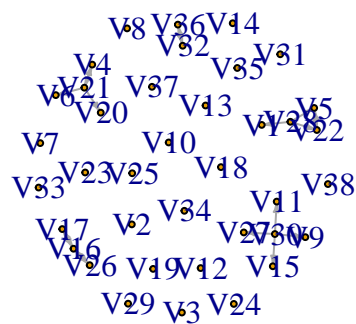
Friendship Network at Time 4
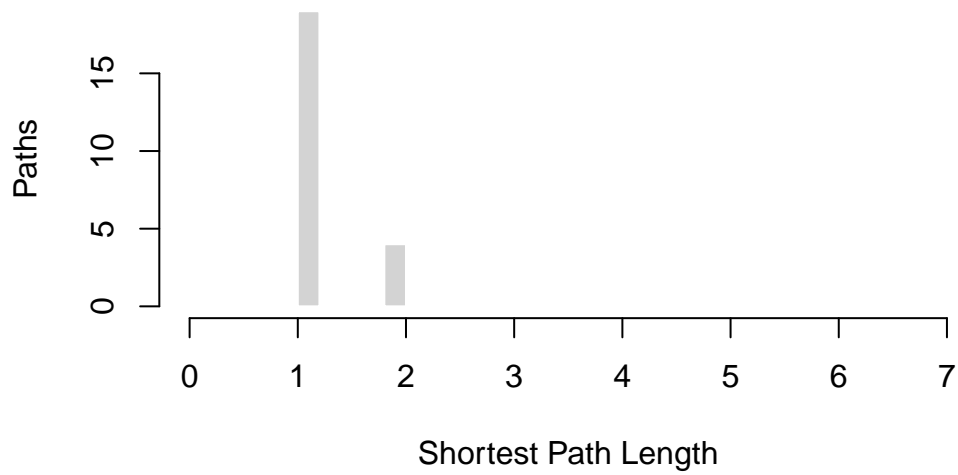

Descriptive data for the network

The in-degree and out-degree are identical for both networks at time 1 and time 4. However, most figures at time 4 are much higher than those at time 1, except for reciprocity. The density increases slightly from 0.01 to 0.09, meaning that after a certain period of time, individuals in this network have more positive relationship with one another. However, the desity is still low in general.

```
# A tibble: 7 x 3
  ``                  Time_1 Time_4
  <I<chr>>             <dbl>  <dbl>
1 in degree (mean)       0.5   3.39
2 in degree (sd)        0.65   2.33
3 out degree (mean)      0.5   3.39
4 out degree (sd)       0.98   4.23
5 density               0.01   0.09
6 reciprocity           0.53   0.43
7 transitivity           0.2   0.42
```

There is no differences between the in-degree and out-degree shortest paths in terms of distribution for both Time 1 and Time 4. However, the resason for 1 being the shortest and most frequnt length of path at Time 1 is that there is not many edges then and individuals are mostly closely connected to each other. At Time 4, individuals nominated more individuals whom they have a positive relationship with, which increased the density of the network as well as the approachability of individuals. Thus, we see the distribution of shortest paths deviates more compared to that of Time 1.
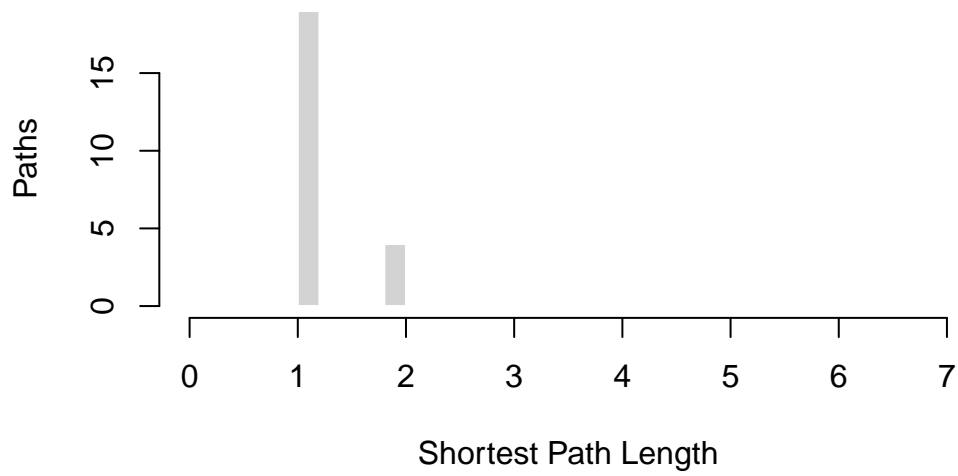
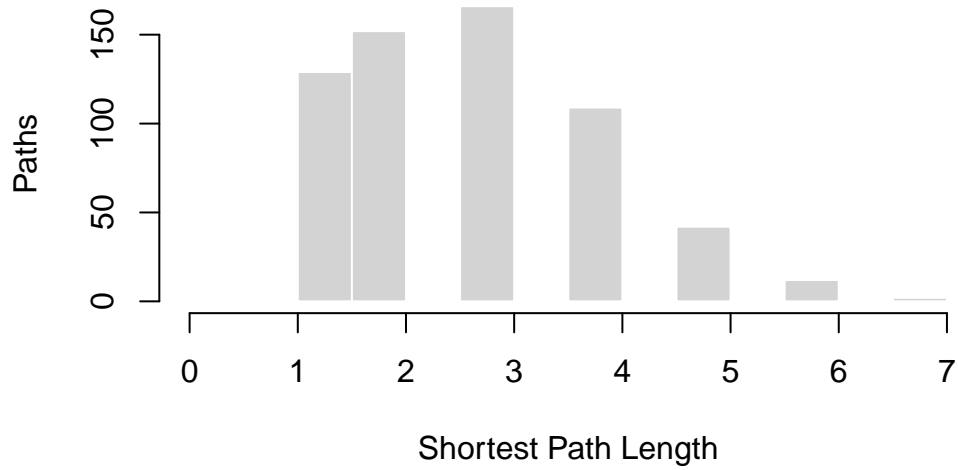# Shortest Path Lengths

at Time 1 –in degree

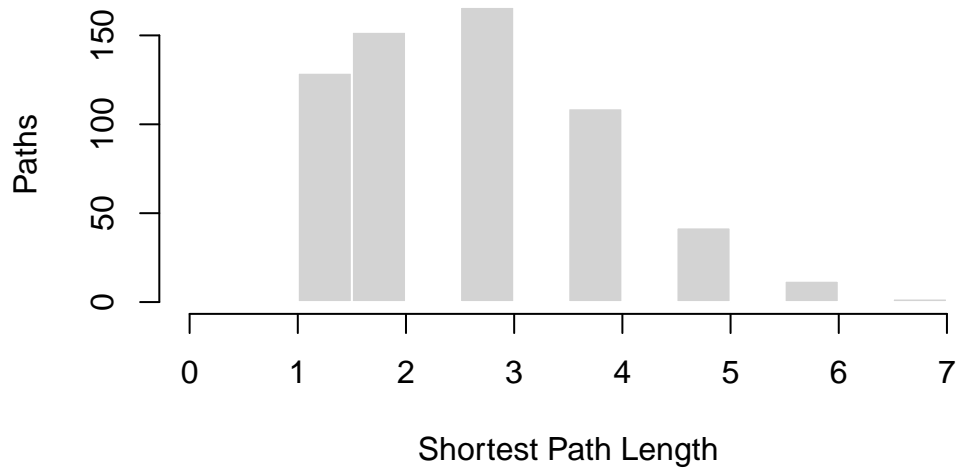# Shortest Path Lengths

at Time 1 – out degree

Shortest Path Lengths

at Time 4 –in degree

Shortest Path Lengths
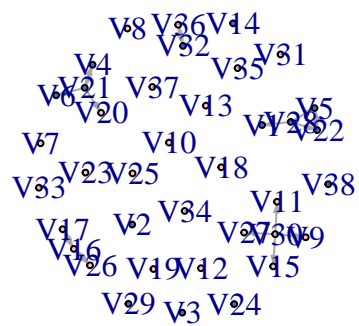
at Time 4 – out degree

4

		N
gender	Female	24
	Male	14
program	2-year	15
	4-year	23
smoke	No	20
	Yes	18
drug	No	22
	Yes	16

Descriptive data for nodes level variables

Visualization of node attributes

Friendship Network at Time 1
blue: male, pink: female

V8 V36 V14
V32 V31
V4 V35
V21 V37 V6 V5
V6 V20 V13 V28 V22
V7 V10
V33 V23 V25 V18 V38
V17 V2 V34 V11
V16 V27 V30 V9
V26 V19 V12 V15
V29 V3 V24

Friendship Network at Time 4
blue: male, pink: female

Assortativity

```
EX index at Time 1 =   -0.404288
```

```
EX index at Time 4 =   -0.1461519
```

It looks like as time went by, the tendency of assortativity decreases slightly, with both boys and girls nominating the opposite sex more compared to Time 1.

As for the measurements for robustness and vulnerability of the network, one can use the number of nodes in a network as the denominator and see if removing the node from the graph causes any noticeable impact (e.g. a significant decrease of network density) on the original network structure. If so, increases the numerator by 1 and repeat the process after examining all the nodes.