ANALYSIS OF BITCOIN OTC TRUST NETWORK

DATASET INFORMATION

Data source - https://snap.stanford.edu/data/soc-sign-bitcoinotc.html

The dataset used is a who trusts whom network of people who trade using Bitcoin on Bitcoin OTC platform. It contains 5,881 nodes and 35,592 edges in total.

Members of Bitcoin OTC rate other members in a scale of -10 (total distrust) to +10 (total trust) in steps of 1.

The dataset is arranged as – Source, Target, Rating, Time;

where

Source: rater

Target: ratee

Rating: source's rating for the target

Time: time of the rating

TOOLS USED

→ R Studio for processing the data

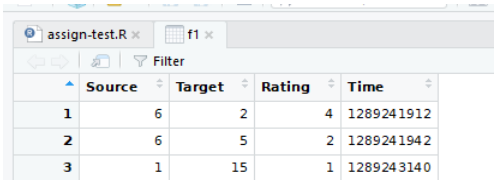
→ iGraph package for visualizing the network

OPERATIONS PERFORMED

For the purpose of our operations, the first 100 records were taken into consideration out of 35,592 records.

Initially records with rating value 5 and above were identified and their respective index numbers were stored in a list. Then the ratee with highest number of raters was deduced with the help of this list and the records.

COLLECTION OF THE FIRST 100 RECORDS

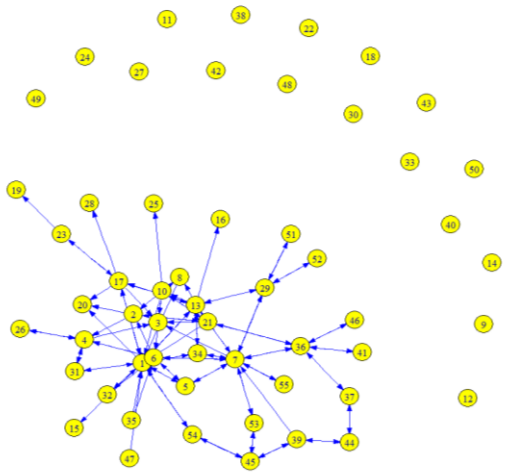


	Source	Target	Rating	Time
1	6	2	4	1289241912
2	6	5	2	1289241942
3	1	15	1	1289243140

List f1 stores all the 35,592 records.

```
> library(igraph)

> l <- c()

> for (j in 1:100) {
+      val1 <- f1[[j, 1]]
+      val2 <- f1[[j, 2]]
+      l <- c(l,val1)
+      l <- c(l,val2)
+      }

> network1 <- graph(edges = l)
> V(network1)$color = "yellow"
```

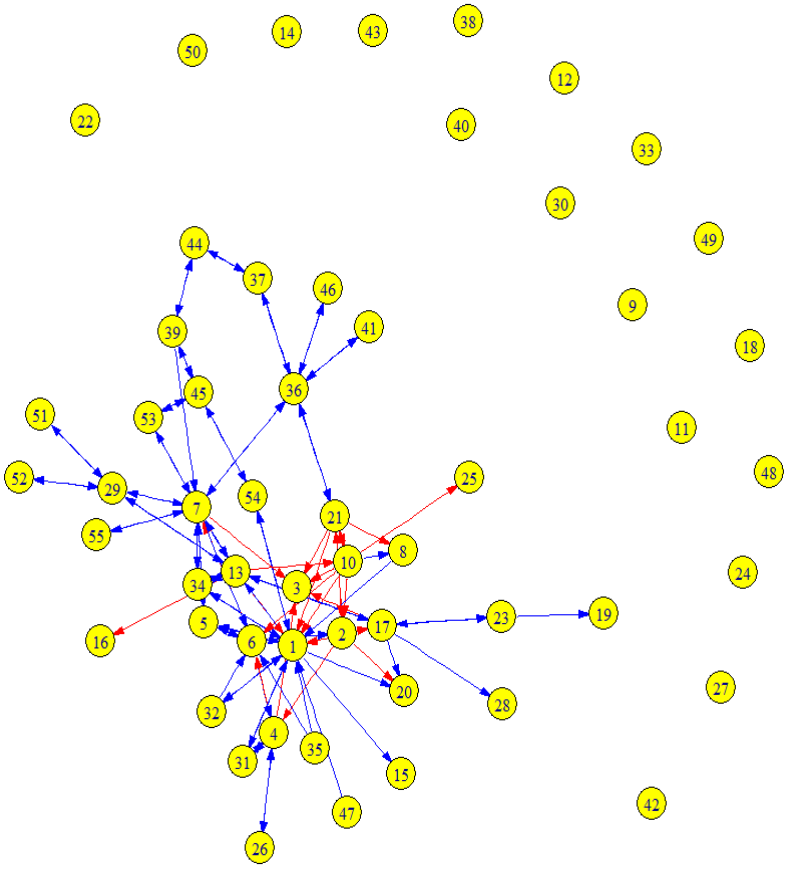
```
> V(network1)$size = 8
> E(network1)$arrow.size = 0.1
> E(network1)$color = "Blue"


> show(network1)


IGRAPH D--- 55 100 --
+ attr: color (v/c), size (v/n), arrow.size (e/n), color (e/c)
+ edges:
 [1]  6-> 2  6-> 5  1->15  4-> 3 13->16 13->10  7-> 5  2->21  2->20 21-> 2
[11] 21-> 1 21->10 21-> 8 21-> 3 17-> 3 17->23 10-> 1 10-> 6 10->21 10-> 8
[21] 10->25 10-> 2 10-> 3  4->26 26-> 4  5-> 1  5-> 6  5-> 7  1-> 5  6-> 4
[31]  4-> 6  2-> 4 17->28 17->13 13->17 13->29 29->13 17->20  4->31 31-> 4
[41] 32-> 6 13-> 1  7->34 34-> 7 32-> 1  1->32  1->34 34-> 1 34->13 13->34
[51]  6-> 7  7-> 6  1->17  1->31 31-> 1 35-> 6  1->13 36->37 37->36 35-> 1
[61] 17-> 1  8-> 1  7->29  1->20 37->44 44->37 39->45 39-> 7 39->44 44->39
[71] 23->17 23->19 36->46 46->36 47-> 1 13-> 7  7->13 29->51 51->29 29->52
+ ... omitted several edges


> plot(network1)
```



IDENTIFICATION OF RECORDS WITH RATING > 4

Records with rating > 4 are stored in list named ‘r’.

```
> r <- c()


> for (j in 1:100) {
+      val1 <- f1[[j, 3]]
+      if (val1 > 4) {
+          r <- c(r,j)
+          }
+      }


> show(r)


[1]  4 5 6 8 9 10 11 12 13 14 15 17 18 19 21 22 23 31 32 42 51 53 61
[24] 89


> for (j in 1:length(r)) {
+      E(network1)[E(network1)[r[j]]]$color = "Red"
+      }

> plot(network1)
```



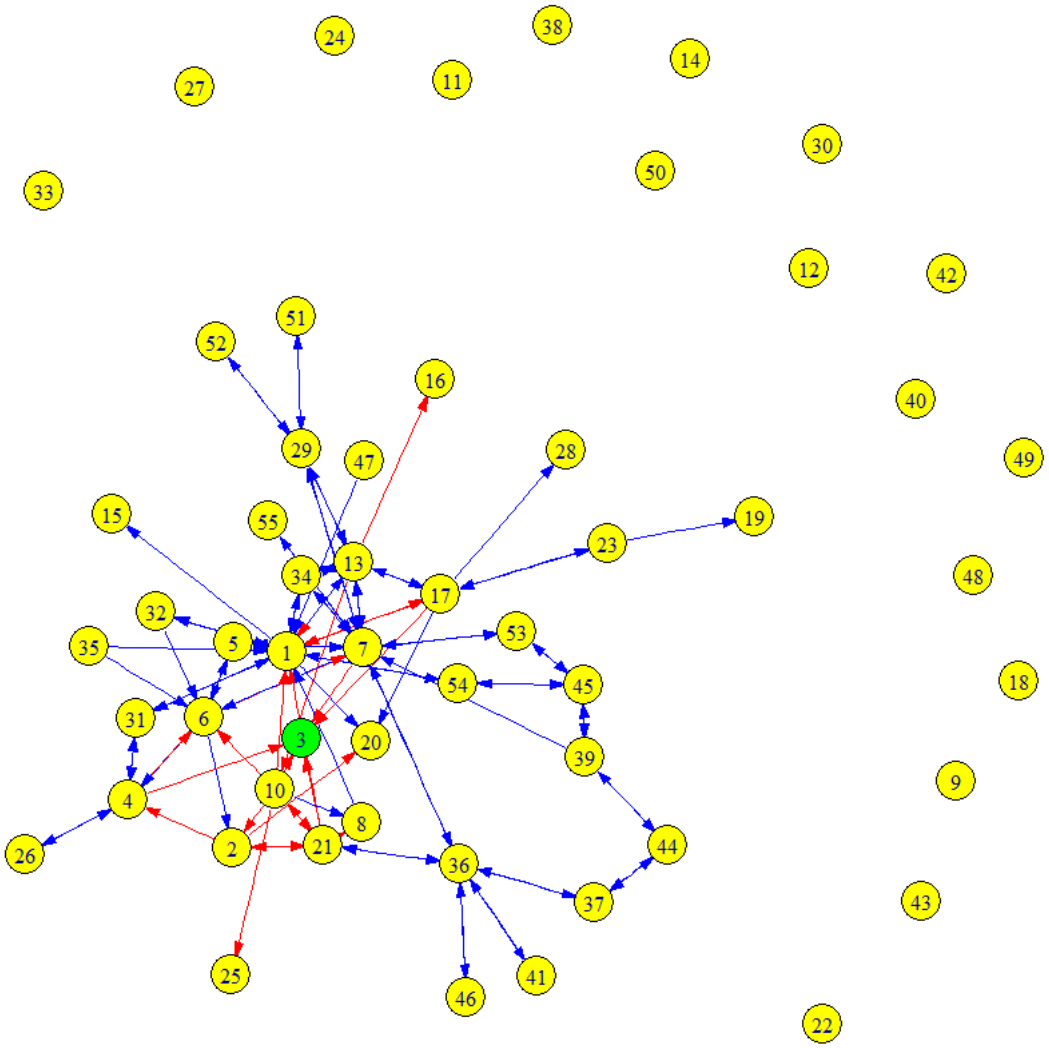
	Source	Target	Rating	Time
1	6	2	4	1289241912
2	6	5	2	1289241942
3	1	15	1	1289243140
4	4	3	7	1289245277
5	13	16	8	1289254254
6	13	10	8	1289254301
7	7	5	1	1289362700
8	2	21	5	1289370557
9	2	20	5	1289370622
10	21	2	5	1289380982
11	21	1	8	1289441411
12	21	10	8	1289441438
13	21	8	9	1289441451
14	21	3	7	1289441526
15	17	3	5	1289442313
16	17	23	1	1289489668
17	10	1	8	1289555656
18	10	6	7	1289555731
19	10	21	8	1289555746
20	10	8	1	1289555768
21	10	25	10	1289555842

IDENTIFICATION OF RATEE WITH MAXIMUM RATERS

Total number of connections of ratees with their corresponding raters is stored in the list named ‘ratee_cn’.

```
> ratee_cn = c()
> ratee = 0
> rater = 0

> for (j in 1:length(r)) {
+         k <- j
```

```
+          rater <- 0
+          ratee <- f1[[r[k],2]]
+          for (k in j:length(r)) {
+              if (ratee == f1[[r[k],2]]) {
+                  rater = rater + 1
+                  }
+              }
+          ratee_cn <- c(ratee_cn, rater)
+          }

> show(ratee_cn)

 [1] 5 1 2 2 1 2 4 1 1 4 3 3 2 1 1 1 2 1 1 2 1 1 1 1

> max(ratee_cn)

[1] 5
> match(c(5), ratee_cn)
[1] 1
> f1[[r[1],2]]
[1] 3
> V(network1)[3]$color = "Green"
> plot(network1)
```


38
24
14
27
11
30
50
33
12
42
51
52
16
40
29
47
28
49
15
55
19
23
13
34
17
48
32
53
35
5
1
7
18
54
45
31
6
3
20
39
9
10
4
8
44
26
2
21
36
37
43
25
46
41
22

CONCLUSION

It was observed from the output that ratee number ‘3’ is the most trusted ratee as per the above mentioned conditions. It has been offered rating value 5 and above by five raters.

The experiment was performed on only 100 records. If more number of records are taken into consideration then there would be possibility of getting a different number as the result for most trusted ratee.