Vidyavardhini's College of Engineering & Technology

Department of Computer Engineering

Experiment No. 6
Social Network Analysis using R (for example: Community Detection Algorithm)
Date of Performance: 05/10/2023
Date of Submission: 12/10/2023

Aim: Social Network Analysis using R (for example: Community Detection Algorithm)

Theory:

Online social platforms have enabled people around the world to interact with each other and build relationships with others they share common interests with. This can be observed in real life — naturally, we tend to develop and maintain relationships with others that are similar to us. People with similar interests tend to gravitate towards each other and become associated in communities — clusters or groups of people that share similar traits with each other. Since people tend to cluster with others similar to them, we can use community detection to identify users with a high number of degrees (connections) and see how far their reach can travel in the network.

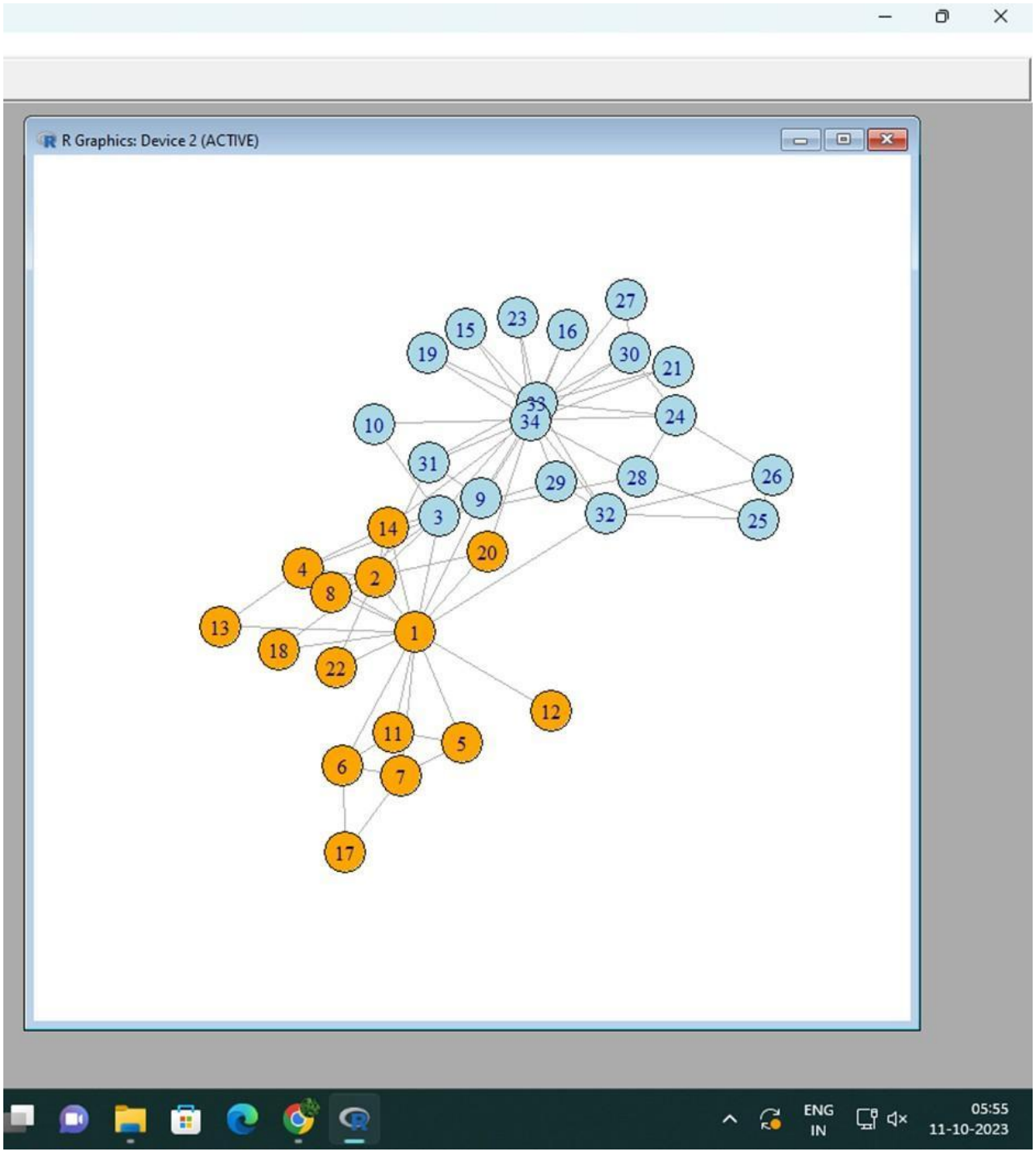
- User Data Extraction — Since we are only interested in user data, we will only extract the following variables:
- User_id — Yelp user ID; this is needed to make nodes and edges
- Name — user's first name
- Review count — the number of reviews user has written
- Yelping since — date user joined Yelp
- Friends — a list containing all of the user's friends by user_id
- Fans — number of fans user has
- Elite — number of years the user has Elite status
- Average stars — user's average rating of all reviews written

CODE:

```
1ibrary(igraph)
gizvan <- function(G) { c=
decompose. graph (G) l =
length(c)
v <= vector() while(1==1){
 x <-E(G)
```

```
y <- edge_betweenness (G) z
<- which.max(y)
edge <- x[z]
a <- ends (G,z[1]) [1]
b <- ends(G,z[1]) [2] v
<- c(v,a,b)
G <- delete_edges (G, edge) c =
decompose.graph (G)
l= length(c)
}
if(l==2)(
paths <- shortest.paths (G) for(i
in 1:lengea(V(G)))(
if (paths[a, i] !=Inf) {
V(G) [i]$color = "lightblue"
}
e1se{
V(G) [i]$color = "orange"
}
}
G <- G + edge(v)
plot(G)
}
return(c)
}
g <- read.graph("C:/Users/admin/Desktop/ComunityDetection/karate.gml",format = "gml")
plot(g)
c <- girvan(g)
```

Output:


R Graphics: Device 2 (ACTIVE)
27
15
23
16
19
30
21
33
34
24
10
31
29
28
26
9
3
32
25
14
20
4
2
8
13
1
18
22
12
11
5
6
7
17
ENG
IN
05:55
11-10-2023

Conclusion:

In this experiment, the primary focus was on conducting Social Network Analysis using the R programming language. The specific task involved the implementation of a Community Detection Algorithm on a social network graph. This algorithm was responsible for extracting user data and the connections between them from a social platform like Yelp. To accomplish this, the experiment leveraged the igraph library in R to detect communities within the network. The Girvan-Newman algorithm was put into action, which operates by iteratively removing edges with the highest betweenness centrality. As a result of these efforts, a network visualization emerged, with communities color-coded to offer valuable insights into how users naturally group together based on their interactions and shared characteristics.