



Experiment No.7
Social Network Analysis using R (for example: Community Detection Algorithm)
Date of Performance:
Date of Submission:



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:

```
#remove users with no friends
```

```
sample <- subset(user_df, friends != "None")
```

```
#make a subset; we only need to retain data of users with some social  
activity sub <- subset(sample, year == 2005 & review_count >= 2 &  
no_of_friends >= 2)
```

CSL702: Big Data Analytics Lab



```
#make links (nodes and edges)
sample_friends <- sub %>%
select(user_id, friends) sample_users <-
strsplit(sample_friends$friends, split=",")
sample_dat<-data.frame(user_id=rep(sample_friends$user_id,

sapply(sample_users, length)), friends = unlist(sample_users))

#network is still too big, take a random sample of
100k nodes samp_net <- sample_n(sample_dat,
100000)
#make network

network <- graph.data.frame(samp_net) network_s <- simplify(network)
net_deg <- degree(network_s) all_degree <- degree(network, mode = 'all')
#graph user with max degrees sub_all <- subcomponent(network_s,
which(all_degree == max(all_degree)), 'all') g_sub <-
induced_subgraph(network_s, sub_all)

#communities

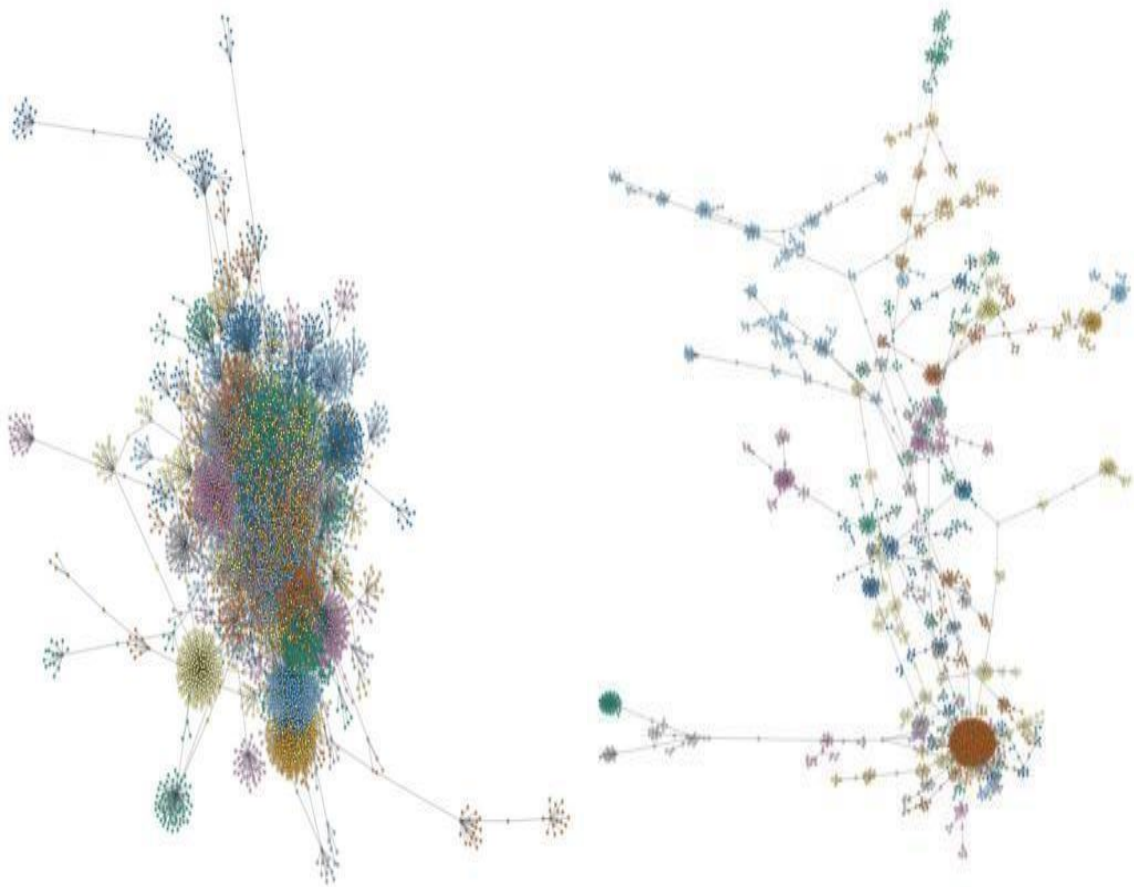
graph.com <- fastgreedy.community(as.undirected(g_sub))

V(g_sub)$color <- graph.com$membership + 1

#create pdf graph for high resolution (try zooming in!)
pdf("communities2005.pdf", 10,10)
plot(g_sub, vertex.color = V(g_sub)$color, vertex.size = 1,
vertex.label = NA, vertex.frame.color = adjustcolor("#41424c",
alpha.f = 0.25), edge.arrow.size = 0.1, edge.color =
adjustcolor("#41424c", alpha.f = 0.20), edge.width = 1.5,
```



```
edge.arrow.mode=0, layout=layout_with_lgl, asp = 0.9, dpi=300  
)  
dev.off()
```



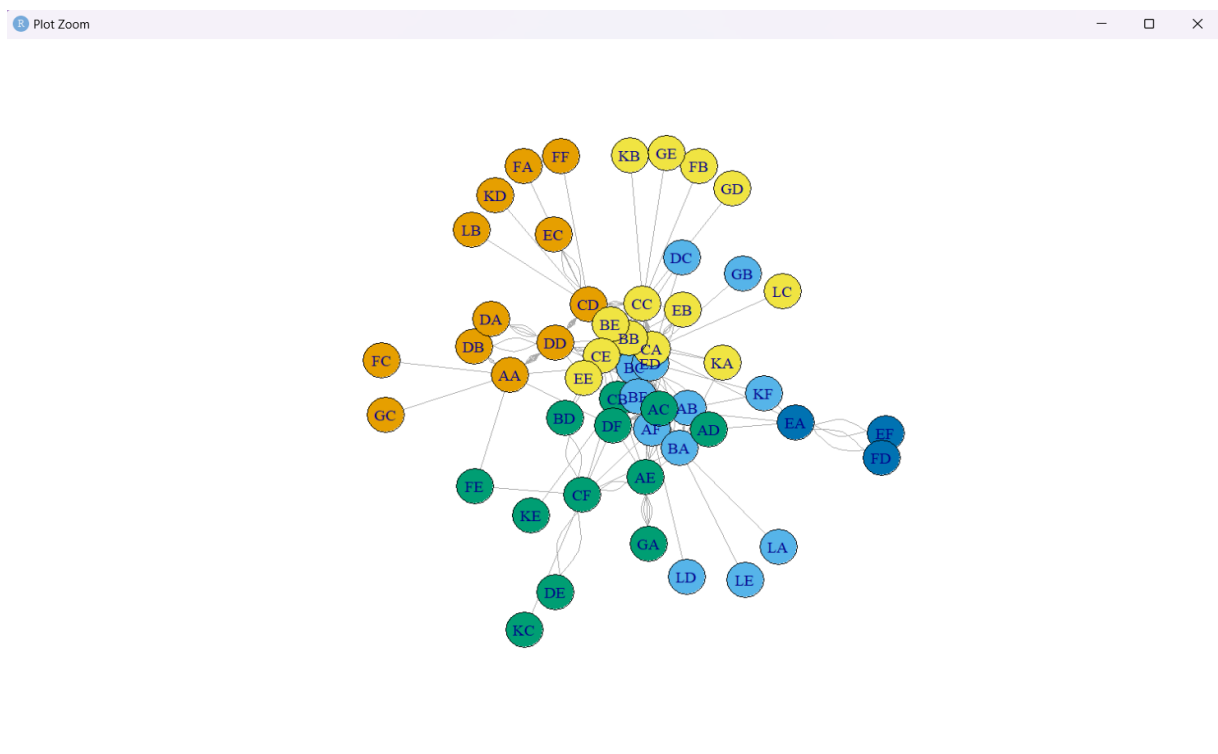
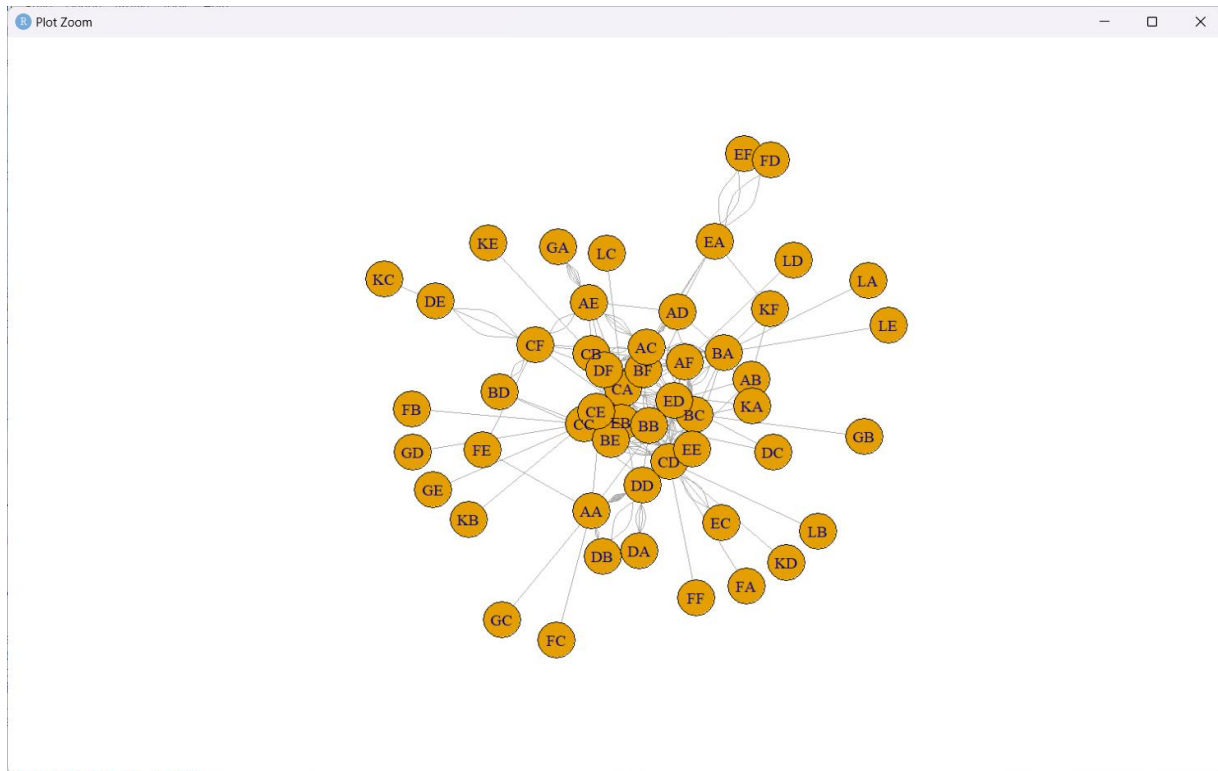


Code:

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Analysis.r
Source on Save Run
1 install.packages("igraph")
2 library(igraph)
3
4 edges <- read.csv("C:\\Users\\Prathmesh\\Downloads\\data.csv")
5
6 social_network <- graph.data.frame(edges, directed = FALSE)
7
8 plot(social_network, layout = layout_fruchterman_reingold)
9
10 communities <- cluster_louvain(social_network)
11
12 plot(social_network, layout = layout_fruchterman_reingold,
13      vertex_color = membership(communities))
14
15 community_metrics <- data.frame(
16   community = 1:length(communities),
17   size = sizes(communities),
18   density = sapply(communities, function(community) {
19     graph_density(subgraph(social_network, v(social_network)
20                           [community]))
21   })
22 )
23 print(community_metrics)
24
```



Output:





Conclusion:

Social Network Analysis (SNA) using R, with community detection algorithms, is approach for gaining insights into the structure and relationships within a social network. Community detection helps uncover hidden subgroups within a network, revealing patterns of association that may not be apparent at first glance. It provides insights into the structure and organization of the network, such as which nodes tend to interact more frequently with one another. Visualizing communities within a social network can help researchers and analysts gain a better understanding of the network's structure. Nodes within the same community are typically coloured similarly for easy identification.

In this experiment, we successfully analyze social network using community detection approach.