Experiment No.9
Social Network Analysis using R (for example: CommunityDetection Algorithm)
Date of Performance: 03/10/23
Date of Submission: 10/10/23

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 reallife — 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 associatedin 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 identifyusers 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 edgesName — user's first name
Review count — the number of reviews user has writtenYelping since — date user joined Yelp
Friends — a list containing all of the user's friends by user_idFans — number of fans user has
Elite — number of years the user has Elite status
Average stars — user's average rating of all reviews written

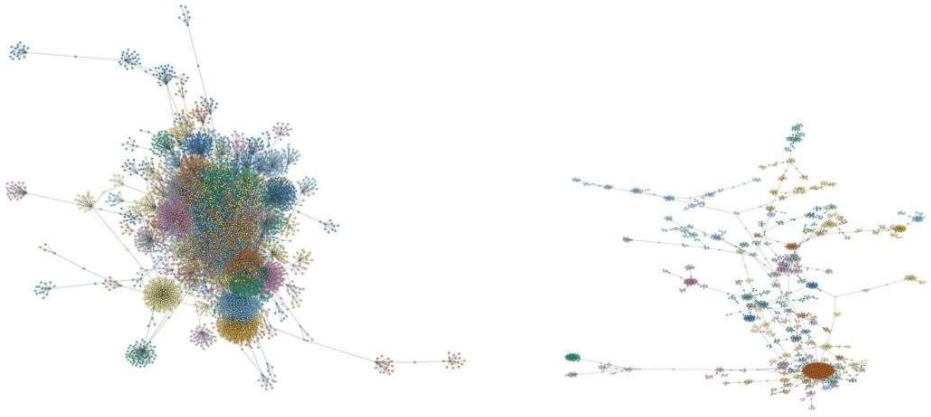
CODE:

```
#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)#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 nodessamp_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()



CONCLUSION:

Social Network Analysis (SNA) in R, particularly when applied to Community Detection Algorithms, is a powerful and versatile tool for uncovering hidden structures within complex networks. R provides a rich ecosystem of packages and libraries that make it accessible to a wide range of users, from beginners to experts. By leveraging SNA techniques in R, researchers and analysts can gain valuable insights into the relationships, clusters, and communities that exist within networks, facilitating a deeper understanding of social, organizational, and biological systems. These insights have applications in various fields, including sociology, biology, marketing, and more, making SNA in R an invaluable resource for exploring the intricate web of connections in the digital age.