

## Ex2\_Network\_Analysis

---

#Loading libraries

```
# Install the readxl package if you haven't already
#install.packages("readxl")

if (!requireNamespace("gt", quietly = TRUE)) install.packages("gt")
library(readxl)
```

```
## Warning: package 'readxl' was built under R version 4.3.3
```

```
library(gt)
```

```
## Warning: package 'gt' was built under R version 4.3.3
```

```
library(igraph)
```

```
## Warning: package 'igraph' was built under R version 4.3.3
```

```
##
```

```
## Attaching package: 'igraph'
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##      decompose, spectrum
```

```
## The following object is masked from 'package:base':
```

```
##
```

```
##      union
```

```
library(ggraph)
```

```
## Loading required package: ggplot2
```

```
## Warning: package 'ggplot2' was built under R version 4.3.3
```

```
library(knitr)
```

```
## Warning: package 'knitr' was built under R version 4.3.3
```

## Loading the dataset

```
fakebook = read_excel("E:/One Drive/OneDrive/Desktop/Ex2_Dataset.xlsx")
```

## Computing the Degree Centrality, Betweenness Centrality and Closeness Centrality and

creating a table to summarize them for each node

```
# Preprocess the dataset: Identify and consolidate reciprocal connections
g_directed <- graph_from_data_frame(fakebook, directed = TRUE)
g_undirected <- as_undirected(g_directed, mode = "collapse")

# Calculate Centrality Measures for the preprocessed (undirected for simplicity) graph
degree centrality <- degree(g_undirected) # For undirected, degree considers all connections
betweenness centrality <- betweenness(g_undirected)
closeness centrality <- closeness(g_undirected)

# Nodes of interest - Adjust as needed
nodes_of_interest <- c("A", "B", "C", "D")

# Subset centrality measures for specific nodes (A, B, C, and D)
degree centrality_subset <- degree centrality[names(degree centrality) %in% nodes_of_interest]
betweenness centrality_subset <- betweenness centrality[names(betweenness centrality) %in% nodes_of_interest]
closeness centrality_subset <- closeness centrality[names(closeness centrality) %in% nodes_of_interest]

# Combine into a dataframe
centrality_df_subset <- data.frame(
  node = names(degree centrality_subset),
  degree = degree centrality_subset,
  betweenness = betweenness centrality_subset,
  closeness = closeness centrality_subset
)

# Creating a table to present values
gt_table <- centrality_df_subset %>%
  gt() %>%
  tab_header(
    title = "Centrality Measures for Nodes",
    subtitle = "This table showcases degree, betweenness, and closeness centrality."
  ) %>%
  cols_label(
    node = "Node",
    degree = "Degree Centrality",
    betweenness = "Betweenness Centrality",
    closeness = "Closeness Centrality"
  ) %>%
  tab_options(
    heading.background.color = "gray"
    # Removed unsupported options for this example
  ) %>%
```

```
data_color(
  columns = vars(degree, betweenness, closeness),
  colors = scales::col_numeric(
    palette = c("lightblue", "blue"),
    domain = NULL
  )
)
```

```
## Warning: Since gt v0.9.0, the 'colors' argument has been deprecated.
## * Please use the 'fn' argument instead.
## This warning is displayed once every 8 hours.
```

```
## Warning: Since gt v0.3.0, 'columns = vars(...)' has been deprecated.
## * Please use 'columns = c(...)' instead.
```

```
gt_table <- gt_table %>%
  tab_style(
    style = list(
      cell_text(size = px(12))
    ),
    locations = cells_body(columns = TRUE)
  )
```

```
## Warning: Since gt v0.3.0, 'columns = TRUE' has been deprecated.
## * Please use 'columns = everything()' instead.
```

```
# Print the gt table
gt_table
```

## Centrality Measures for Nodes

This table showcases degree, betweenness, and closeness centrality.

Node	Degree Centrality	Betweenness Centrality	Closeness Centrality
A	3	14.000000	0.06250000
B	5	9.033333	0.07142857
C	5	8.600000	0.07142857
D	5	3.266667	0.06250000

## Visualizing the network graph

```
g = g_undirected

# Create a new label that combines the node name and centrality values
node_labels <- paste(V(g)$name,
  "\nDegree: ", round(degree centrality, 3),
  "\nBetweenness: ", round(betweenness centrality, 3),
  "\nCloseness: ", round(closeness centrality, 3), sep="")
```

```

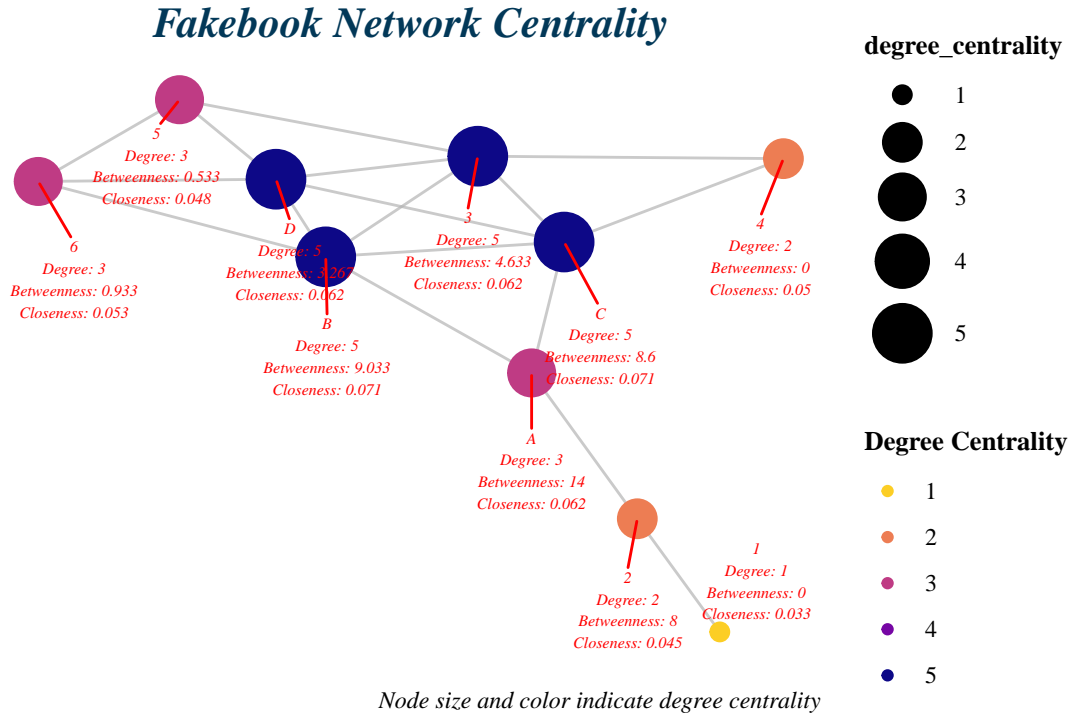
# Create the visualization with improved colors and a legend
p <- ggraph(g, layout = 'fr') +
  geom_edge_link(edge_alpha = 0.7, edge_width = 0.5, edge_color = "gray70") +
  geom_node_point(aes(size = degree centrality, color = degree centrality), show.legend = TRUE) +
  geom_node_text(aes(label = node_labels), vjust = 1.8, color = "red", size = 2, repel = TRUE,
    fontface = "italic", family = "serif") +
  scale_size_continuous(range = c(3, 10)) +
  scale_color_viridis_c(option = "C", end = 0.9, direction = -1, guide = 'legend') +
  ggtitle("Fakebook Network Centrality") +
  theme_graph(base_family = "serif") +
  theme(
    plot.title = element_text(size = 16, face = "bold.italic", hjust = 0.5, color = "#023858"),
    plot.background = element_rect(fill = "white", color = NA),
    legend.position = "right",
    legend.title = element_text(size = 10, face = "bold"),
    legend.text = element_text(size = 9),
    plot.margin = margin(30, 30, 30, 30)
  ) +
  labs(caption = "Node size and color indicate degree centrality", color = "Degree Centrality")
p

```

```

## Warning: Using the 'size' aesthetic in this geom was deprecated in ggplot2 3.4.0.
## i Please use 'linewidth' in the 'default_aes' field and elsewhere instead.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_lifecycle_warnings()' to see where this warning was
## generated.

```



## Saving to JPG

```
ggsave("E:/Fakebook_Network.jpg",
  p,
  width = 10,
  height = 5,
  dpi = 300,
  device = "jpg"
)
```

## Useful insights

### Seat A: The Connector

With a high betweenness centrality, seat A is a significant connector between different areas of the bus. This implies that while it may not have the highest number of immediate connections, it sits on paths connecting various groups, acting as a bridge. This position would be beneficial for someone interested in a diverse range of conversations or being a point of information exchange. The potential downside is that it may not provide the deepest connections with any particular group due to its bridging role.

### Seat B: The Network Hub

Seat B has a high degree of centrality and is close to the other highly central nodes (C and D). Its closeness centrality is also the highest, which suggests it can interact with many others more directly and quickly. This seat is excellent for someone looking to be well-integrated into the bus network, forming numerous casual connections.

#### Seat C: The Inner Circle

Similar to seat B, it has a high degree and closeness centrality but slightly lower betweenness centrality. This suggests it's well-connected but slightly less critical as a bridge. This seat would allow you to form many connections and be at the center of the social network on the bus.

#### Seat D: The Local Node

Like B and C, it has a high degree centrality but the lowest betweenness centrality. This suggests that while it's well-connected, it's not as crucial for connecting different parts of the bus network. It might be better for forming strong connections within a more localized area of the bus.

---

## Factors impacting decision:

Based on the centrality measures and their implications:

If goal is to maximize influence and information access, and you thrive in dynamic, fast-paced social settings, choose Seat B. If you aim to strategically position yourself as a connector and understand diverse perspectives without being overwhelmed by the central social buzz, choose Seat A. If you prefer a balance between being at the social forefront and having meaningful interactions, choose Seat C. If depth and quality of connections are your priority, and you have a more reserved or targeted networking style, choose Seat D.

---

## Node Selection:

In order to make the most out of the networking opportunity on the Fakebook bus with a balance of influence and information flow without the pressure of being the main hub, I would choose Seat C as it stands out as the optimal choice. It provides a high degree of centrality for networking breadth and closeness centrality for quick information access while slightly reducing the pressure of being a bridge within the network, which falls on Seat A.

Benefits: I am well-placed to create robust connections and still be a part of various group conversations. Also it is suitable for those who want the benefits of centrality without the intense pressure of being the primary hub.

Drawbacks: I may still face the challenge of managing multiple interactions, and there could be an expectation for me to disseminate information due to my position.