Exercise 2: ORGB 672

Lakshya Agarwal

2024-03-21

Contents

1 Setup

```
library(tidygraph)
```

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

```
##
## Attaching package: 'tidygraph'
```

```
## The following object is masked from 'package:stats':
##
##     filter
```

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.2     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.3     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.2
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks tidygraph::filter(), stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(igraph)
```

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

```
##
## Attaching package: 'igraph'
##
## The following objects are masked from 'package:lubridate':
##
##     %--%, union
##
## The following objects are masked from 'package:dplyr':
##
##     as_data_frame, groups, union
##
## The following objects are masked from 'package:purrr':
##
##     compose, simplify
##
## The following object is masked from 'package:tidyr':
##
##     crossing
```

```
##
## The following object is masked from 'package:tibble':
##
##     as_data_frame
##
## The following object is masked from 'package:tidygraph':
##
##     groups
##
## The following objects are masked from 'package:stats':
##
##     decompose, spectrum
##
## The following object is masked from 'package:base':
##
##     union
```

```
library(ggplot2)
library(vroom)
```

```
##
## Attaching package: 'vroom'
##
## The following objects are masked from 'package:readr':
##
##     as.col_spec, col_character, col_date, col_datetime, col_double,
##     col_factor, col_guess, col_integer, col_logical, col_number,
##     col_skip, col_time, cols, cols_condense, cols_only, date_names,
##     date_names_lang, date_names_langs, default_locale, fwf_cols,
##     fwf_empty, fwf_positions, fwf_widths, locale, output_column,
##     problems, spec
```

```
library(scales)
```

```
##
## Attaching package: 'scales'
##
## The following object is masked from 'package:vroom':
##
##     col_factor
##
## The following object is masked from 'package:purrr':
##
##     discard
##
## The following object is masked from 'package:readr':
##
##     col_factor
```

```
library(ggraph)
library(ggtext)
library(ggrepel)
```

```
library(ggforce)
library(ggthemes)
library(patchwork)
library(qualpalr)
```

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

```
# setwd("./Exercise 2")
```

2 Code

2.1 Create a graph with the specified edges

```
edges <- vroom("./edge_list.csv", delim = ",")
```

```
## Rows: 16 Columns: 2
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (2): from, to
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
graph <- as_tbl_graph(edges, directed = FALSE)

graph
```

```
## # A tbl_graph: 10 nodes and 16 edges
## #
## # An undirected simple graph with 1 component
## #
## # Node Data: 10 x 1 (active)
##    name
##    <chr>
##  1 1
##  2 2
##  3 A
##  4 B
##  5 D
##  6 5
##  7 C
##  8 3
##  9 6
## 10 4
## #
## # Edge Data: 16 x 2
##    from    to
##   <int> <int>
```

```
## 1     1     2
## 2     2     3
## 3     3     7
## # i 13 more rows
```

2.2 Calculate centrality measures for {A, B, C, D}

```
centrality_measures <- graph %>%
  activate(nodes) %>%
  mutate(
    degree = centrality_degree(),
    closeness = centrality_closeness(),
    betweenness = centrality_betweenness(),
    label = paste(
      "Name: ", name,
      "\nDegree: ", round(degree, 2),
      "\nCloseness: ", round(closeness, 2),
      "\nBetweenness: ", round(betweenness, 2)
    ),
    color = case_when(
      name %in% c("A", "B", "C", "D") ~ "red",
      TRUE ~ "black"
    )
  )

centrality_measures
```

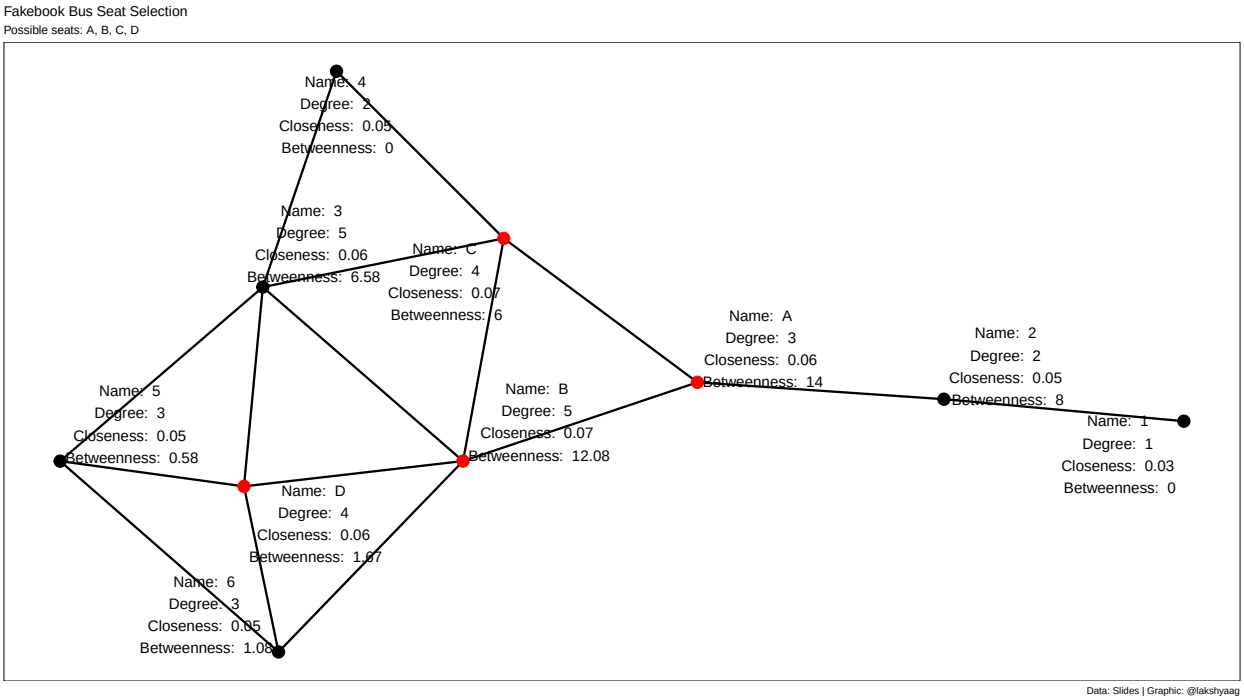
```
## # A tbl_graph: 10 nodes and 16 edges
## #
## # An undirected simple graph with 1 component
## #
## # Node Data: 10 x 6 (active)
##    name  degree closeness betweenness label                                color
##    <chr>  <dbl>     <dbl>       <dbl> <chr>                                <chr>
##  1 1          1    0.0333       0     "Name:  1 \nDegree:  1 \nCloseness:~ black
##  2 2          2    0.0455       8     "Name:  2 \nDegree:  2 \nCloseness:~ black
##  3 A          3    0.0625      14     "Name:  A \nDegree:  3 \nCloseness:~ red
##  4 B          5    0.0714      12.1   "Name:  B \nDegree:  5 \nCloseness:~ red
##  5 D          4    0.0588       1.67  "Name:  D \nDegree:  4 \nCloseness:~ red
##  6 5          3    0.0476       0.583 "Name:  5 \nDegree:  3 \nCloseness:~ black
##  7 C          4    0.0667       6     "Name:  C \nDegree:  4 \nCloseness:~ red
##  8 3          5    0.0625       6.58  "Name:  3 \nDegree:  5 \nCloseness:~ black
##  9 6          3    0.0526       1.08  "Name:  6 \nDegree:  3 \nCloseness:~ black
## 10 4          2    0.05         0     "Name:  4 \nDegree:  2 \nCloseness:~ black
## #
## # Edge Data: 16 x 2
##    from    to
##   <int> <int>
## 1     1     2
## 2     2     3
## 3     3     7
## # i 13 more rows
```

2.3 Plot network with labels

```
network_plot <- centrality_measures %>%
  ggraph(layout = "stress") +
  geom_edge_link(width = 1) +
  geom_node_point(
    size = 5,
    aes(color = color)
  ) +
  scale_color_manual(values = c("black", "red")) +
  geom_node_text(aes(label = label), size = 5, repel = TRUE) +
  theme_void() +
  theme(
    legend.position = "none",
    panel.border = element_rect(fill = NA)
  ) +
  plot_annotation(
    title = "Fakebook Bus Seat Selection",
    subtitle = "Possible seats: A, B, C, D",
    caption = "Data: Slides | Graphic: @lakshyaag"
  )

network_plot
```

```
## 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.
```



```r
ggsave("fakebook_bus_network_plot.png", network_plot,
  width = 16, height = 9, dpi = 300
)
```

3 Fakebook Bus Seat Selection Discussion

As a new intern at Fakebook, choosing where to sit on the company bus is an important decision for fostering informal connections with coworkers.

3.1 Seat A

- **Degree**: 3
- **Closeness**: 0.06
- **Betweenness**: 14

3.1.1 Benefits

- **Central Role**: With the highest betweenness, consistently choosing Seat A will place the intern in a position where many paths pass through, making them a key connector in the network.
- **Influence**: The intern could potentially influence the flow of information and communication by taking a central seat.

3.1.2 Drawbacks

- **Responsibility**: The central role might come with the pressure of being involved in many interactions.
- **Overload**: There's a risk of becoming overwhelmed by the need to facilitate conversations.

3.2 Seat B

- **Degree**: 5
- **Closeness**: 0.07
- **Betweenness**: 12.08

3.2.1 Benefits

- **Social Hub**: Seat B has the highest degree, offering the opportunity to directly connect with more people.
- **Accessibility**: High closeness means that the intern is, on average, closer to everyone else, facilitating easier communication and relationship-building.

3.2.2 Drawbacks

- **Less Privacy**: More connections might mean less personal space and more interruptions.
- **High Engagement**: The intern might be expected to engage with a larger number of people regularly due to high proximity.

3.3 Seat C

- **Degree**: 4
- **Closeness**: 0.07
- **Betweenness**: 6

3.3.1 Benefits

- **Balanced Interaction**: Seat C offers a good balance of direct connections and centrality without the intensity of Seat B.
- **Strategic Position**: Moderately high closeness and betweenness suggest a strategic position for networking without being overwhelmed.

3.3.2 Drawbacks

- **Less Prominent**: Not as central as Seat A or as connected as Seat B, which might limit networking opportunities.

3.4 Seat D

- **Degree**: 4
- **Closeness**: 0.06
- **Betweenness**: 1.67

3.4.1 Benefits

- **Quiet Networking**: Seat D allows for networking opportunities with a reasonable number of direct connections and a lower profile in the network.
- **Less Pressure**: Lower betweenness means less responsibility in facilitating interactions.

3.4.2 Drawbacks

- **Peripheral Role**: Less centrality might make it harder for the intern to establish themselves as a key player in the network.
- **Indirect Influence**: The intern may have less influence on the overall flow of communication due to a lower betweenness centrality.

3.5 Conclusion

In conclusion, the intern's choice should align with their networking style and desired level of interaction. If the goal is to be a central figure, Seat A or B might be the best choice. Seat C offers a balanced approach, while Seat D is not recommended for the intern's purposes.