Assignment 1

Student ID : 22223696

2023-01-31

1. Basic measures

```
library(igraph)
```

```
## Warning: package 'igraph' was built under R version 4.1.3
```

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

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

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

```
library(kableExtra)
```

```
## Warning: package 'kableExtra' was built under R version 4.1.3
```

```
dib_graph<-read.graph("dib2.graphml",format="graphml")
```

1.1 Give the number of nodes and edges

```
cat("num vertices:", vcount(dib_graph), "\n")
```

```
## num vertices: 8969
```

```
cat("num edges :", ecount(dib_graph), "\n")
```

```
## num edges : 46750
```

1.2 Is the network strongly or weakly connected. If neither, what is the distribution of component sizes.

1. Strongly connected components

```
strong_component = as.data.frame(table(factor(components(dib_graph, mode="strong")$csize)))
names(strong_component)[1] = "Component Size"
kable(strong_component)
```

Component Size	Freq
1	3024
2	180
3	25
4	5
5	1
6	1
5479	1

Answer: The table above shows strongly connected components in the directed network. There is one component with 5479, 6 and 5 nodes and a distribution of component with sizes varying from 1 to 4.

2. Weakly connected components

```
weak_component = as.data.frame(table(factor(components(dib_graph, mode="weak")$csize)))
names(weak_component)[1] = "Component Size"
kable(weak_component)
```

Component Size	Freq
2	30
3	11
4	1
8872	1

Answer: The table above shows weakly connected components in the directed network. There is one component with 8872 and 4 nodes with 30 weakly connected components of size 2 and 11 components with size 3.

1.3 What is the diameter of the network ?

```
cat("The diameter of the network is : ", diameter(dib_graph, directed = T, unconnected = TRUE, weights=N
```

```
## The diameter of the network is :  18
```

1.4 What is the average path length of the network ?

```
cat("The average path length of the network :", mean_distance(dib_graph, directed = T), "\n")
```

```
## The average path length of the network : 6.017593
```

1.5 What is the clustering coefficient of the network ?

```
cat("The clustering coeff of the graph is :", transitivity(dib_graph, type="localaverage"), "\n")
```

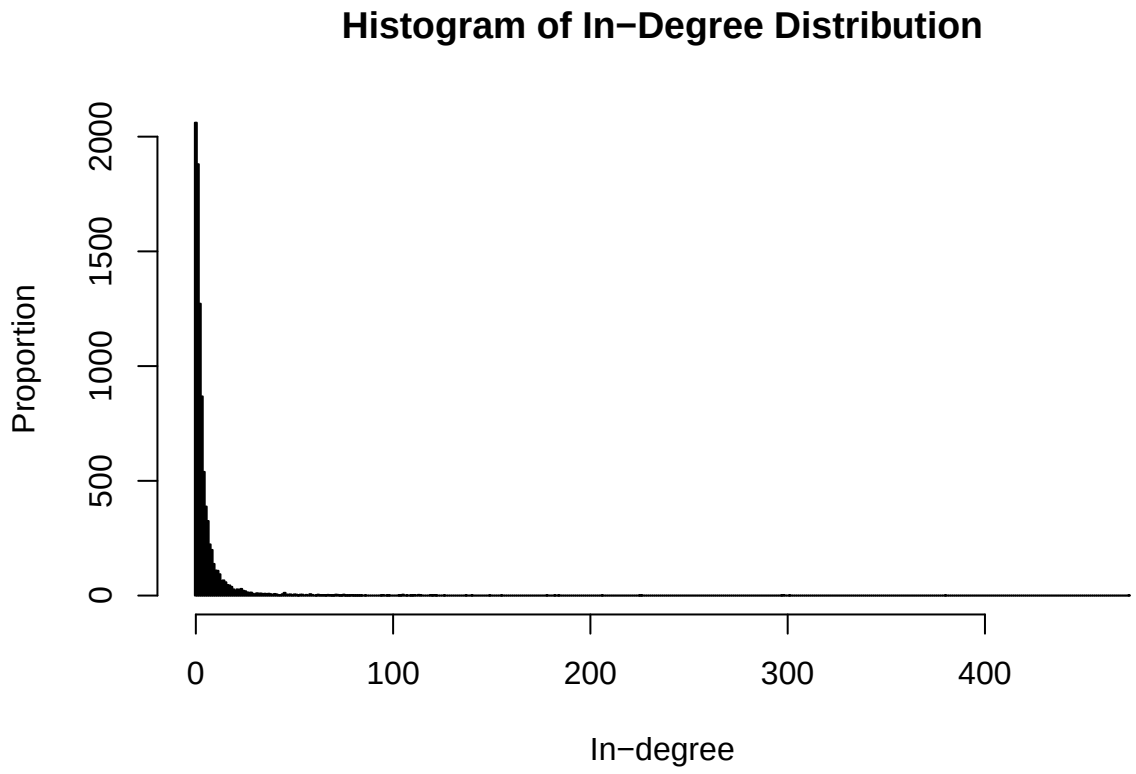
```
## The clustering coeff of the graph is : 0.2300017
```

1.6 What is the in- and out-degree distribution ?

```
deg<-degree(dib_graph, mode = "in")
cat("The in-degree distribution of the graph varies from ", min(deg), "to ", max(deg))
```

```
## The in-degree distribution of the graph varies from  0 to  473
```

```
hist(deg,
     breaks=(min(deg)-1):(max(deg))+0.5,
     xlab = "In-degree",
     ylab = "Proportion",
     main = "Histogram of In-Degree Distribution",
     border="black",
     col="white",
     )
```




```
deg<-degree(dib_graph, mode = "out")
cat("The in-degree distribution of the graph varies from ", min(deg), "to ", max(deg))
```

```
## The in-degree distribution of the graph varies from  0 to  58
```

```
hist(deg,
     breaks=(min(deg)-1):(max(deg))+0.5,
     xlab = "Out-degree",
     ylab = "Proportion",
```

```
     main = "Histogram of Out-Degree Distribution",
     border="black",
     col="white")
```