**Network Visualization of Facebook users**

**Analytical report**

1. **Executive summary**

This is a report on a social network analysis using R, whereby Facebook users are analyzed based on sex, group, and frind count. The following packages were used *igraph, network, sna, visNetwork, threejs, and ggraph.* The analysis involves plotting of a whole network, several network grouped into groups and also grouped by sex.The main findings were that most female users had served as hubs (arrows pointing outwards) and male users served as authorities (arrows pointing inwards). The user with highest number of friends were male.

1. **Importation of the two csv files and the creation of a directed igraph network from these files.**

The two datasets were loaded into R and the following are the output;

# Load and inspect the datasets  
nodes=read.csv ("Fb\_Nodes.csv", header = T)   
head(nodes)

## id group sex friend\_count  
## 1 SE F female 73  
## 2 PH G male 53  
## 3 PD G female 363  
## 4 LN G female 64  
## 5 MB F female 19  
## 6 AP C female 99

links=read.csv("FB\_Links.csv",header = T)  
head(links)

## from to  
## 1 LSF TSL  
## 2 GL TSL  
## 3 SLL TSL  
## 4 CSH TSL  
## 5 DSF TSL  
## 6 CS TSL

The output of the igraph object is as follows:

# create the igraph object  
net <- graph\_from\_data\_frame(d=links, vertices=nodes, directed=T)   
net

## IGRAPH f625049 DN-- 93 323 --   
## + attr: name (v/c), group (v/c), sex (v/c), friend\_count (v/n)  
## + edges from f625049 (vertex names):  
## [1] LSF->TSL GL ->TSL SLL->TSL CSH->TSL DSF->TSL CS ->TSL SE ->TSL ALL->LSF  
## [9] GL ->LSF DL ->LSF SLL->LSF ADL->LSF AL ->LSF CSH->LSF AAE->LSF DSF->LSF  
## [17] CS ->LSF SE ->LSF SL ->LSF BCL->LSF BL ->LSF MBM->JLT BB ->JLT AC ->JLT  
## [25] KM ->JLT BS ->JLT SS1->SS2 LBW->SS2 TR ->LRG JLG->LRG DSG->LRG TR ->DSG  
## [33] JLG->TR KM ->MD BS ->MD LL ->MD CWB->MD JK ->LBW JK ->DM RW ->DM   
## [41] NW ->DM TB1->DM JK ->TB1 RW ->TB1 NW ->TB1 SMH->TB1 ME ->TB1 TM ->TB1  
## [49] DSW->TB1 DK ->TB1 JK ->SMH NW ->SMH ME ->SMH TM ->SMH DSW->SMH RBW->SMH  
## [57] DK ->SMH TB2->SMH JK ->SS1 ASS->SS1 JK ->ASS RW ->JK ME ->JK TM ->JK   
## + ... omitted several edges

1. **Network attribute inspection as described in task 2.**

The output from the E and V functions display the edges and vertices of the network, while the data frame shows the vertices (nodes), and edges (links).

# Use the E and V function to inspect the attributes of the network  
  
E(net) # The edges of the "net" object

## + 323/323 edges from f625049 (vertex names):  
## [1] LSF->TSL GL ->TSL SLL->TSL CSH->TSL DSF->TSL CS ->TSL SE ->TSL ALL->LSF  
## [9] GL ->LSF DL ->LSF SLL->LSF ADL->LSF AL ->LSF CSH->LSF AAE->LSF DSF->LSF  
## [17] CS ->LSF SE ->LSF SL ->LSF BCL->LSF BL ->LSF MBM->JLT BB ->JLT AC ->JLT  
## [25] KM ->JLT BS ->JLT SS1->SS2 LBW->SS2 TR ->LRG JLG->LRG DSG->LRG TR ->DSG  
## [33] JLG->TR KM ->MD BS ->MD LL ->MD CWB->MD JK ->LBW JK ->DM RW ->DM   
## [41] NW ->DM TB1->DM JK ->TB1 RW ->TB1 NW ->TB1 SMH->TB1 ME ->TB1 TM ->TB1  
## [49] DSW->TB1 DK ->TB1 JK ->SMH NW ->SMH ME ->SMH TM ->SMH DSW->SMH RBW->SMH  
## [57] DK ->SMH TB2->SMH JK ->SS1 ASS->SS1 JK ->ASS RW ->JK ME ->JK TM ->JK   
## [65] DK ->JK TB2->JK ECW->JK RW ->ECW NW ->RW ME ->RW TM ->RW TB2->RW   
## [73] AC ->MBM BS ->MBM KW ->MBM SE ->LZ ARL->LZ PW ->LZ DZB->LZ CW1->LZ   
## + ... omitted several edges

V(net) # The vertices of the "net" object

## + 93/93 vertices, named, from f625049:  
## [1] SE PH PD LN MB AP SM CB TB2 RBW NW AL DB CWB BS AC BW DZM KW   
## [20] VST TM CR AAE HRS LRS SL JJ TBC LE BCL SLL TNB MGE JLG DGT DK WT DSW  
## [39] ME MWB BD LL BB LSM KM SP LCP ADE KH ECS KT GB RBH JU JZ DSM BL   
## [58] CG CSH GGA ADL DL LBF GL CS DSF ALL CW1 DZB PW ARL LZ MBM RW ECW JK   
## [77] ASS SS1 SMH TB1 DM LBW MD TR DSG LRG SS2 JLT LSF TSL LBC LLB DW

V(net)$group # Vertex attribute "group"

## [1] "F" "G" "G" "G" "F" "C" "C" "W" "W" "W" "W" "F" "M" "M" "M" "M" "M" "M" "M"  
## [20] "W" "W" "F" "F" "F" "F" "F" "C" "C" "C" "F" "F" "G" "G" "G" "G" "W" "W" "W"  
## [39] "W" "M" "M" "M" "M" "M" "M" "S" "S" "S" "S" "S" "S" "H" "H" "H" "H" "M" "F"  
## [58] "F" "F" "F" "F" "F" "F" "F" "F" "F" "F" "B" "B" "B" "B" "B" "M" "W" "W" "W"  
## [77] "W" "W" "W" "W" "W" "W" "W" "G" "G" "G" "W" "M" "F" "F" "C" "C" "H"

# Find nodes and edges by attribute:  
# (that returns oblects of type vertex sequence/edge sequence)  
V(net)[group=="B"]

## + 5/93 vertices, named, from f625049:  
## [1] CW1 DZB PW ARL LZ

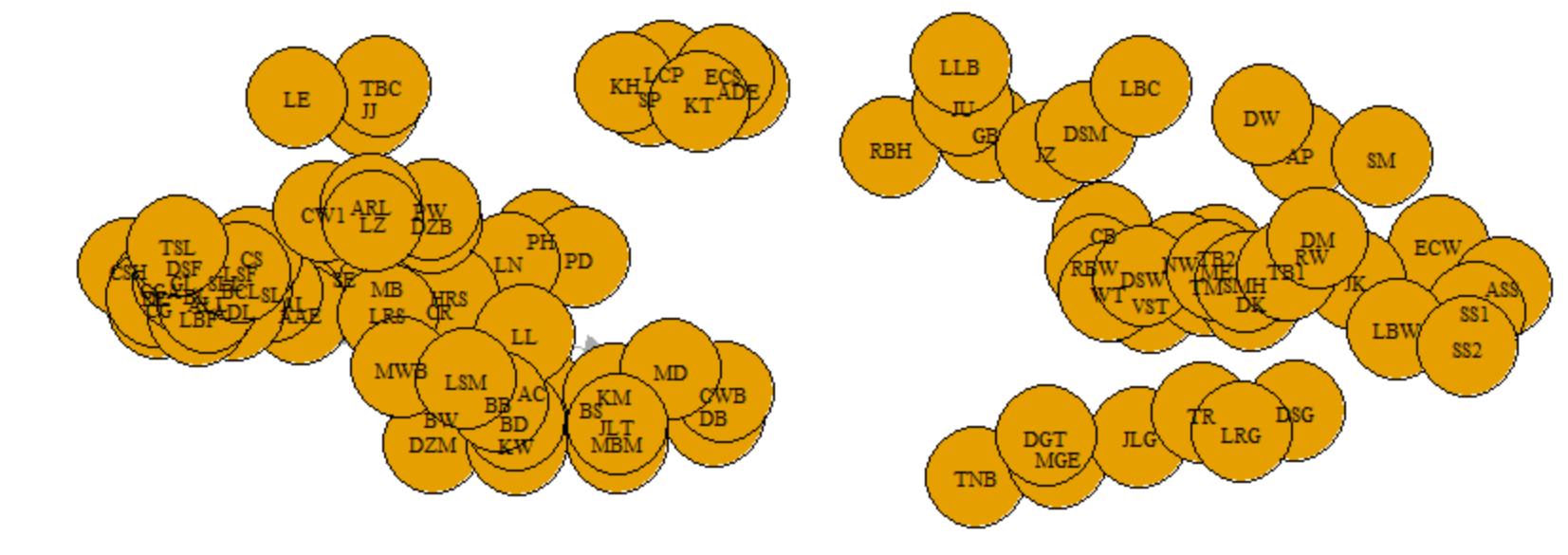
# use the as\_data\_frame function to describe the nodes and link  
df\_edge=as\_data\_frame(net, what="edges")  
head(df\_edge)

## from to  
## 1 LSF TSL  
## 2 GL TSL  
## 3 SLL TSL  
## 4 CSH TSL  
## 5 DSF TSL  
## 6 CS TSL

df\_vertices=as\_data\_frame(net, what="vertices")  
head(df\_vertices)

## name group sex friend\_count  
## SE SE F female 73  
## PH PH G male 53  
## PD PD G female 363  
## LN LN G female 64  
## MB MB F female 19  
## AP AP C female 99

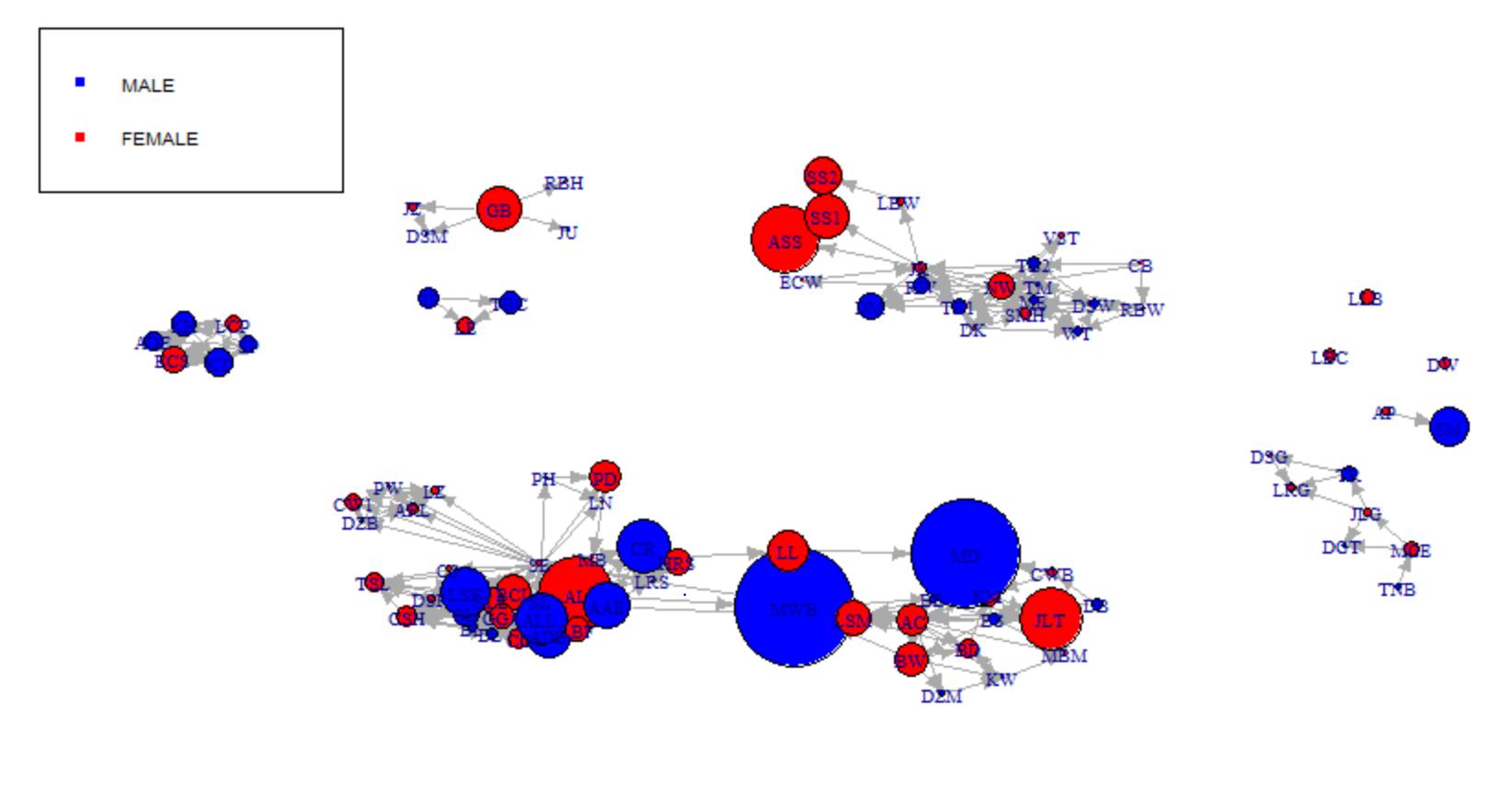
The plot of the network is as follows;



**Observation and explanation**

The nodes are too close to each other hence difficult to identify the label of the node, also the arrows are not visible.

1. **Re-plotting of the network with adjusted attribute settings as described in task 3.**

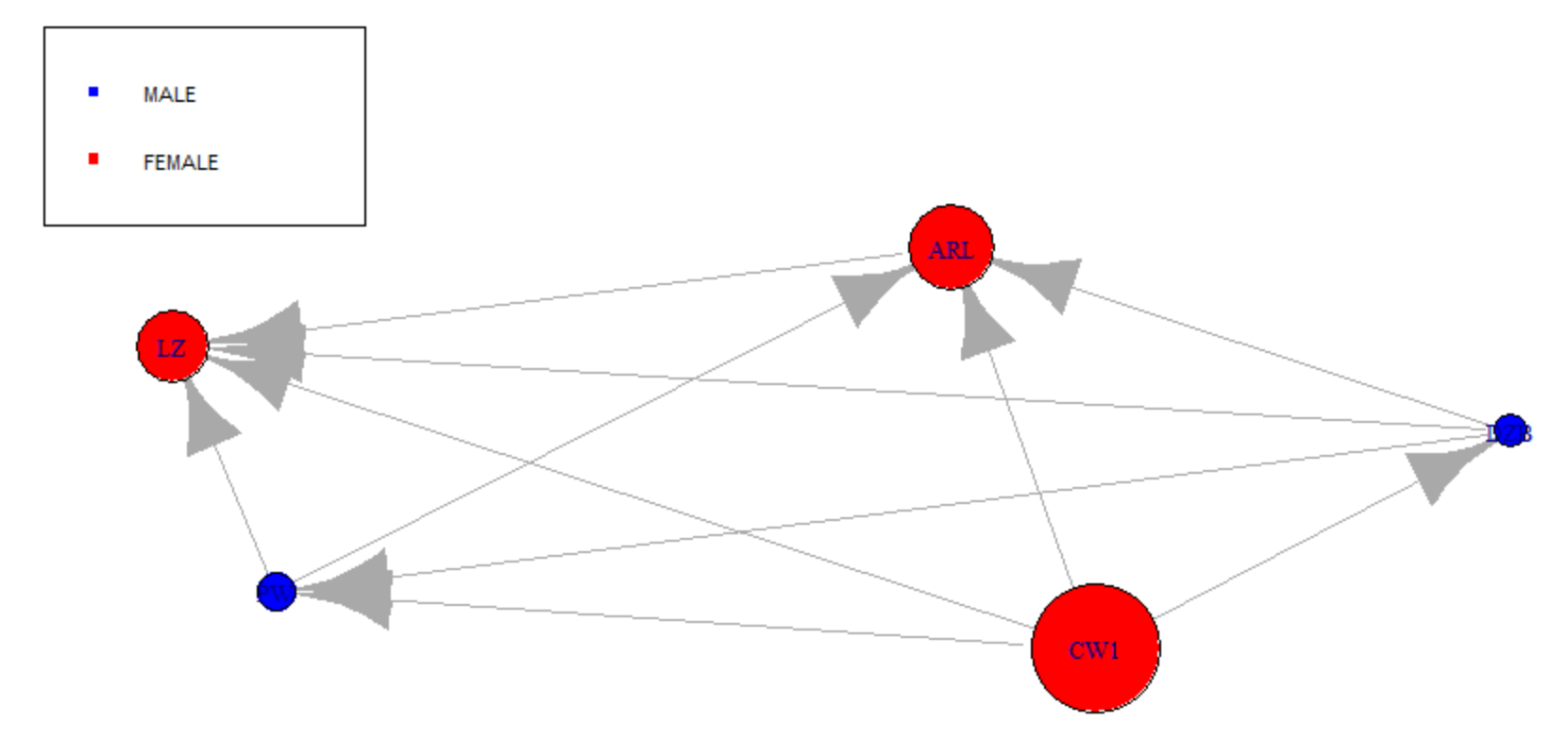


**Observation and explanation**

Previously you could not see the direction of the arrows clearly, we could not tell from which sex each of the nodes belong to, and also we could not compare the nodes based on size (friend count). But after the adjusting the settings then the node with the highest friend count is labelled has the largest node and its clearly labeled MWB.

1. **Plotting a series of networks where each plot contains only Facebook users that belong within a particular group, as described in task 4.**

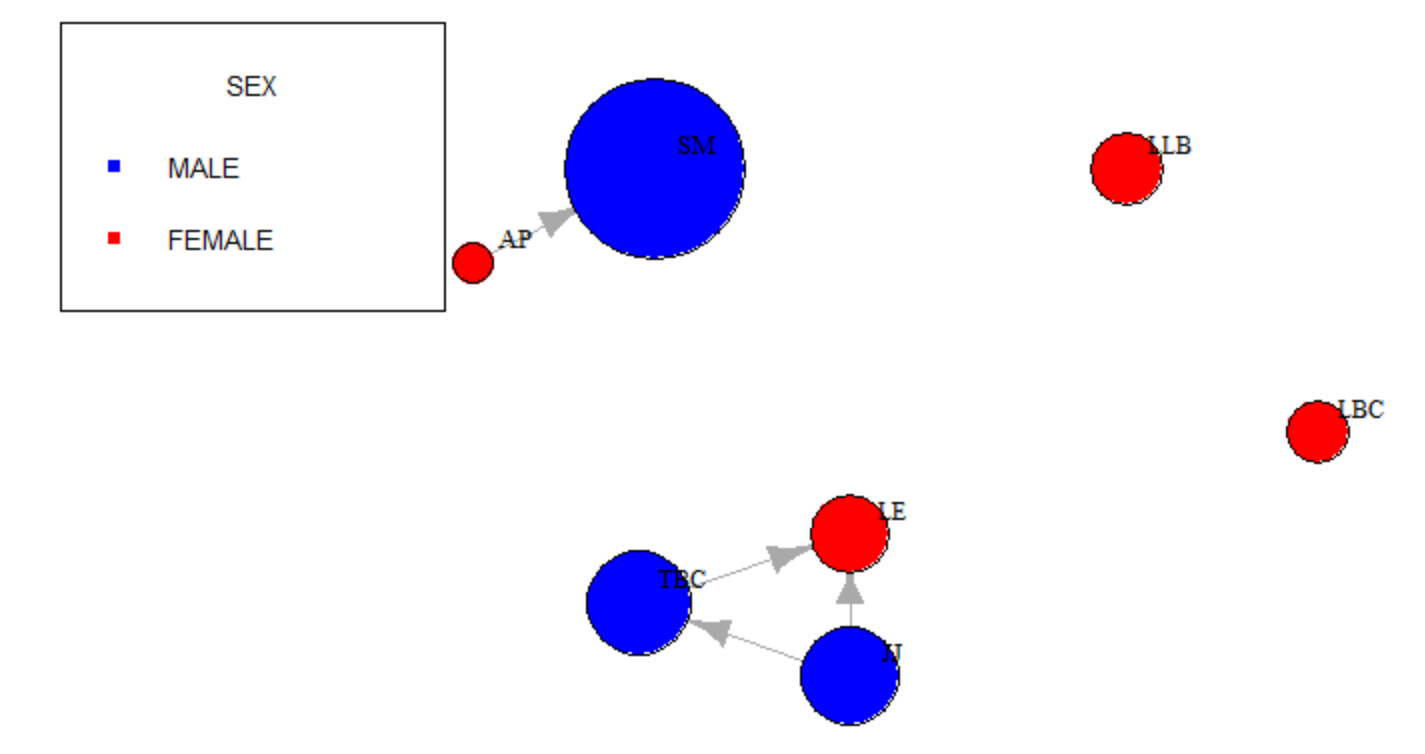
GROUP B



**Observation and explanation**

From group B the user CW1was a good hub but a poor authority since it had many arrows pointing outwards and none pointing towards it. It also had the highest number of friends.

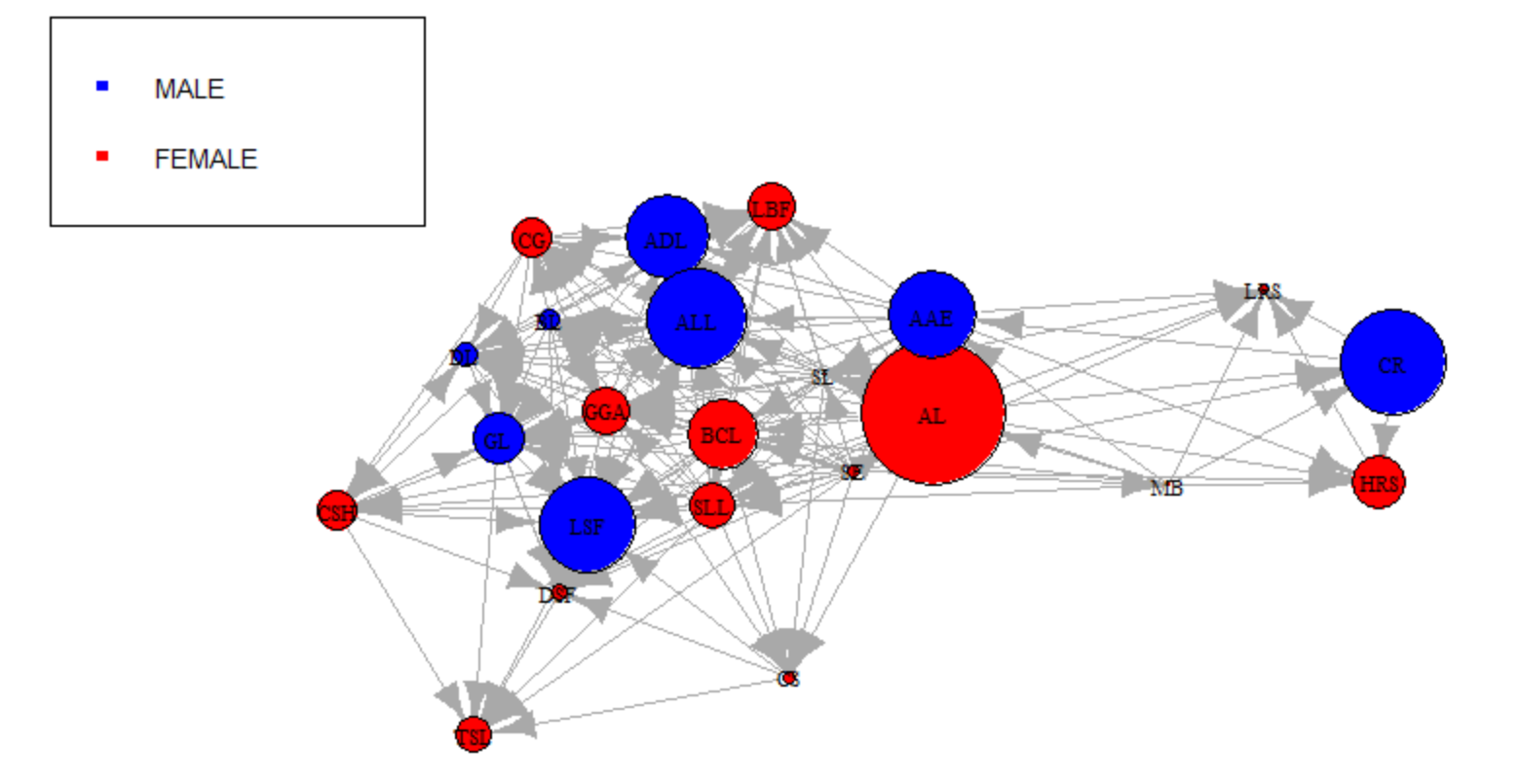
GROUP C



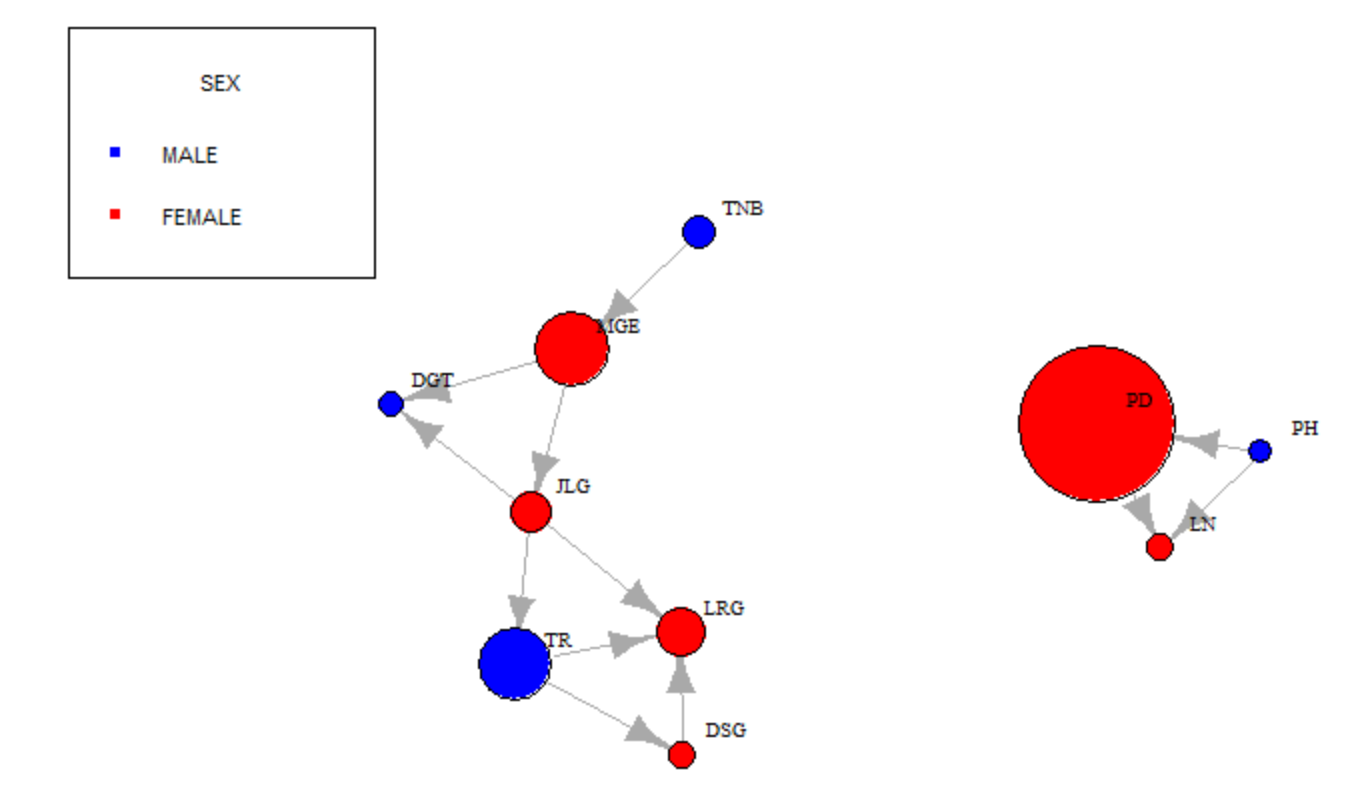
**Observation and explanation**

User JJ was a good hub connected to 2 users (TBC & LE) the most number of users. User SM had the highest number of friends but poor hub. User LLB and LBC had no connections.

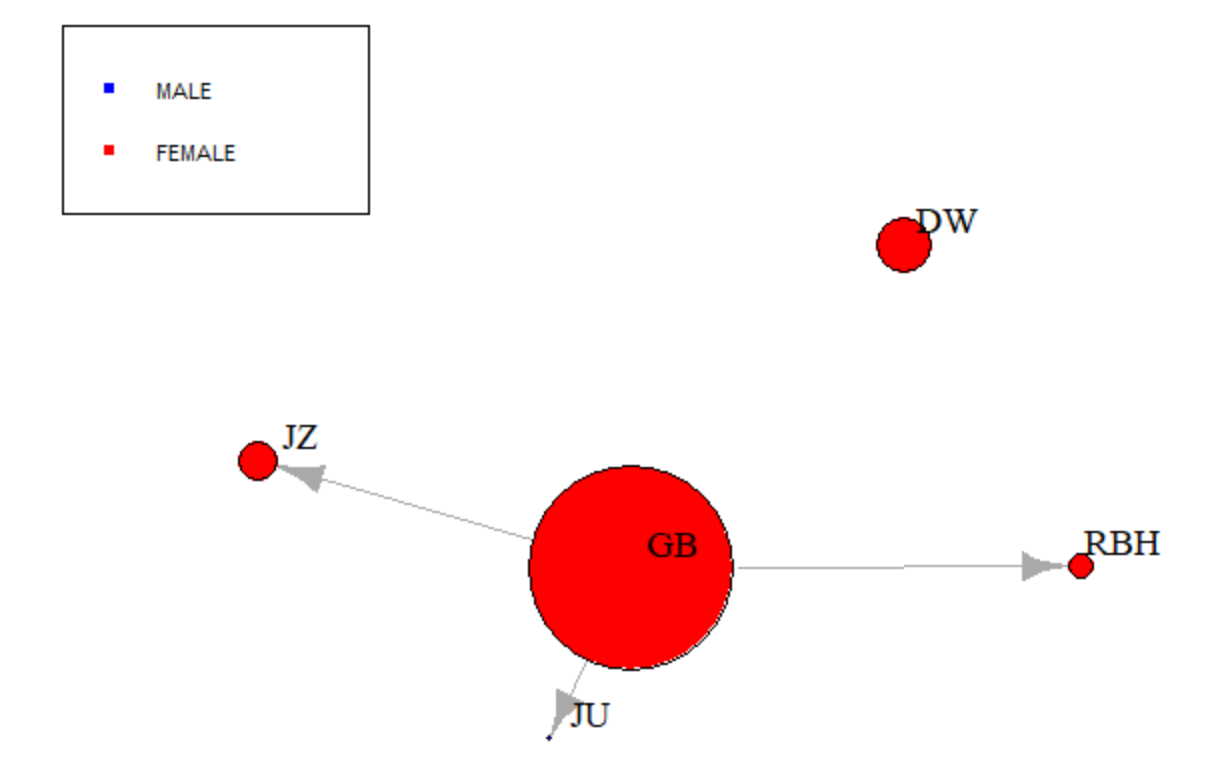
GROUP F



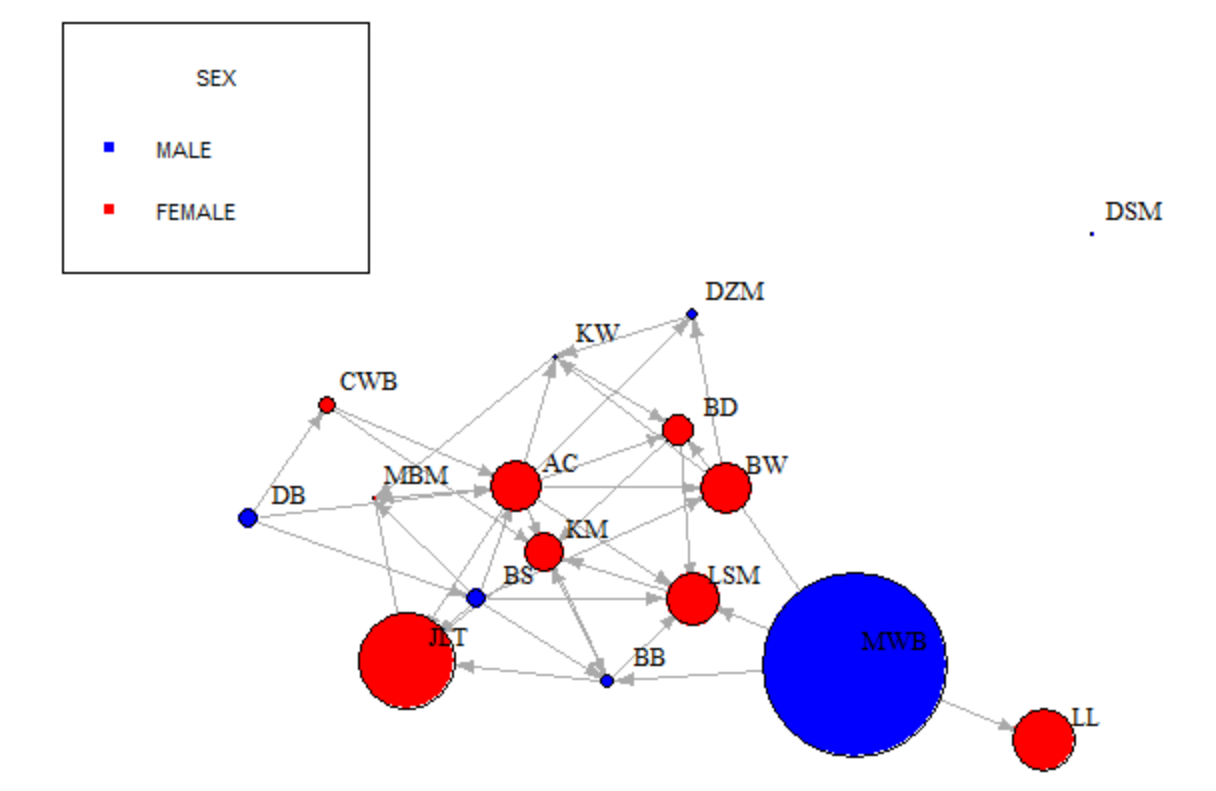
**GROUP G**



**GROUP H**



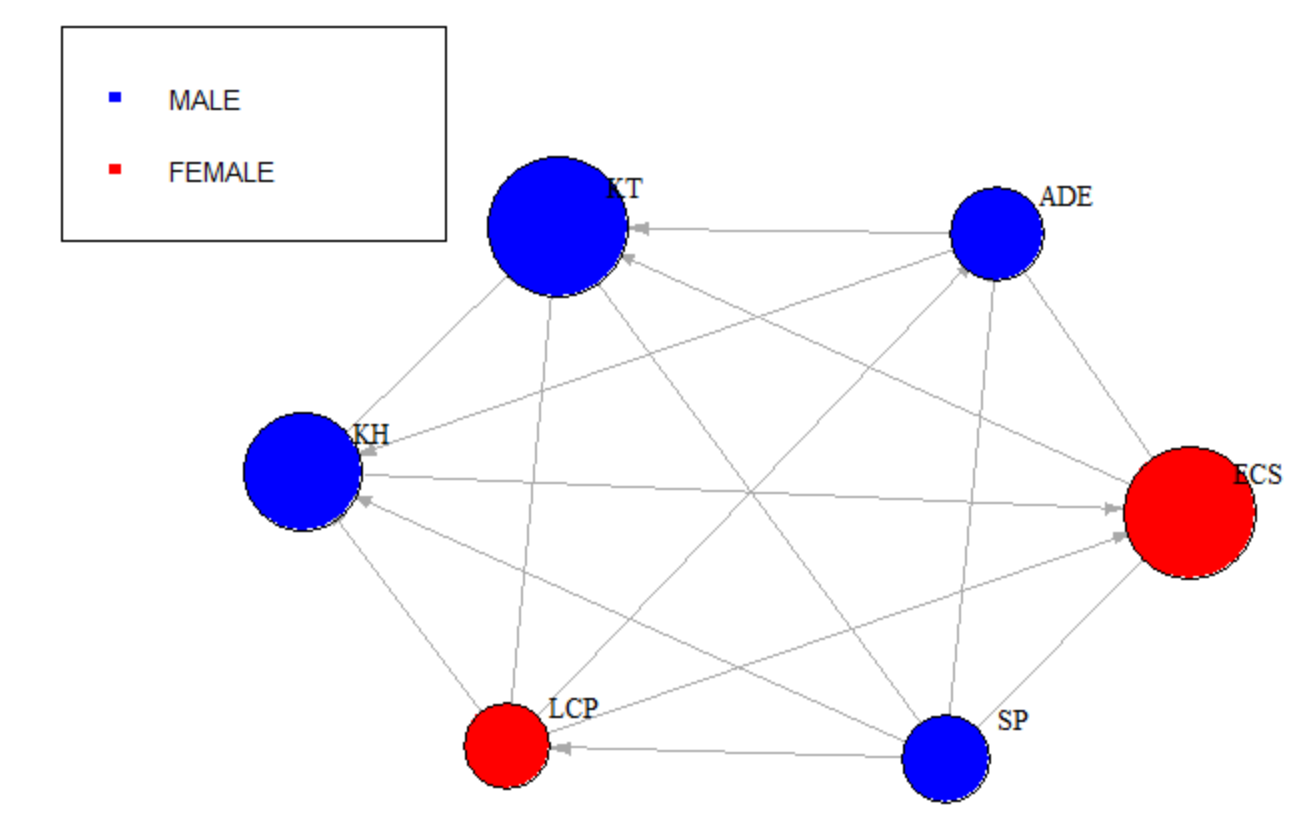
**GROUP M**



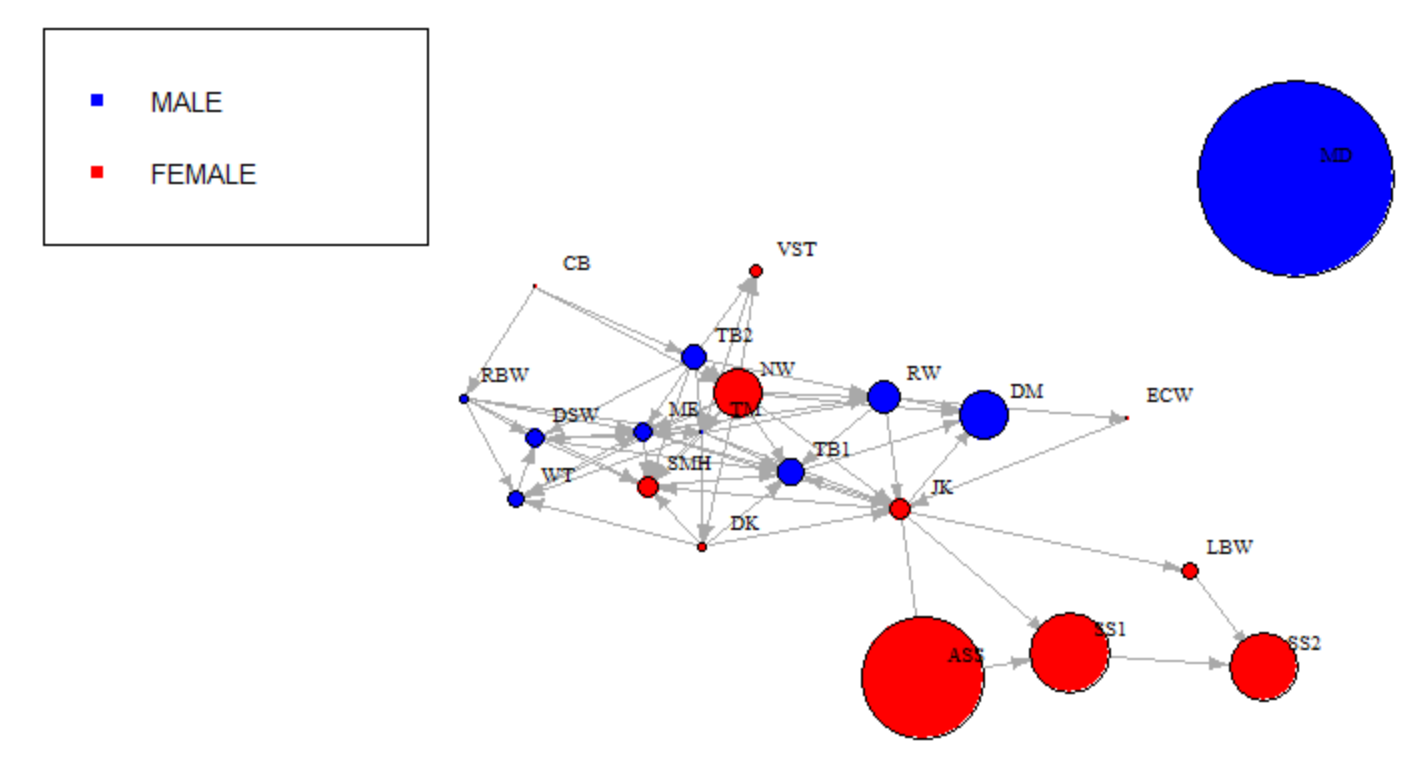
**Observation and explanation**

By visually inspecting the network plot for group M, identify the largest node (the Facebook user with the highest friend count). ***MWB had the highest number of friend count.*** By examining the number of links each node has in this plot, would you agree that this user is the most important node in this network? ***YES, because it has 3 outgoing “to” links and one incoming “from” link. This shows that it’s a hub and an authority in itself given it has many friends also.***

**GROUP S**

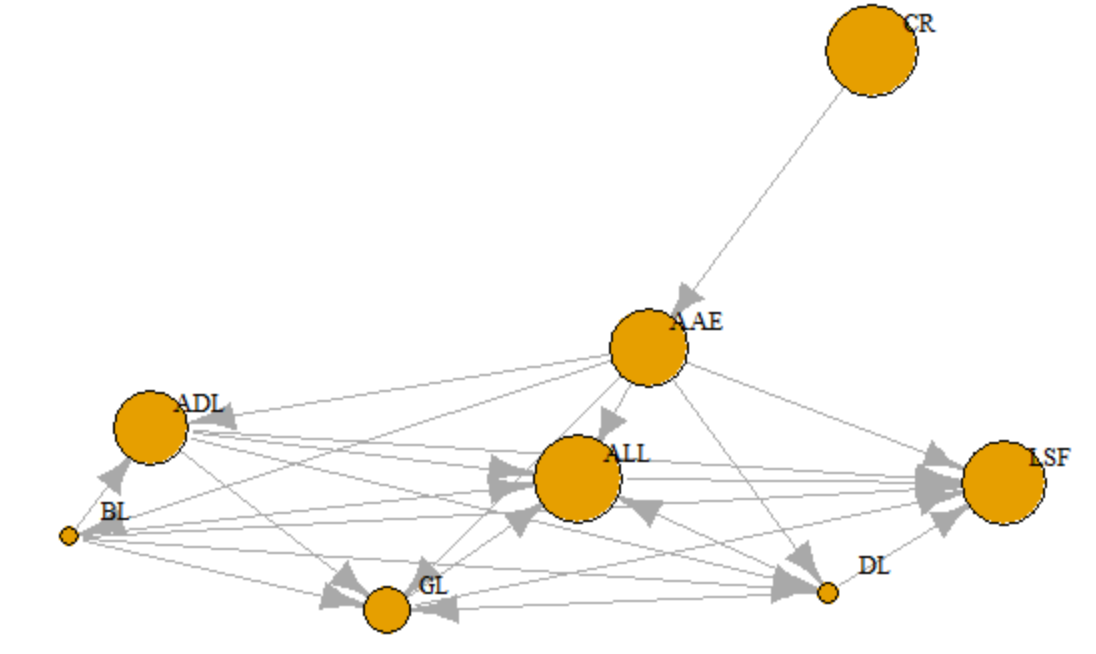


**GROUP W**



1. **Re-plotting the network containing group F Facebook users by gender, as described in task 5.**

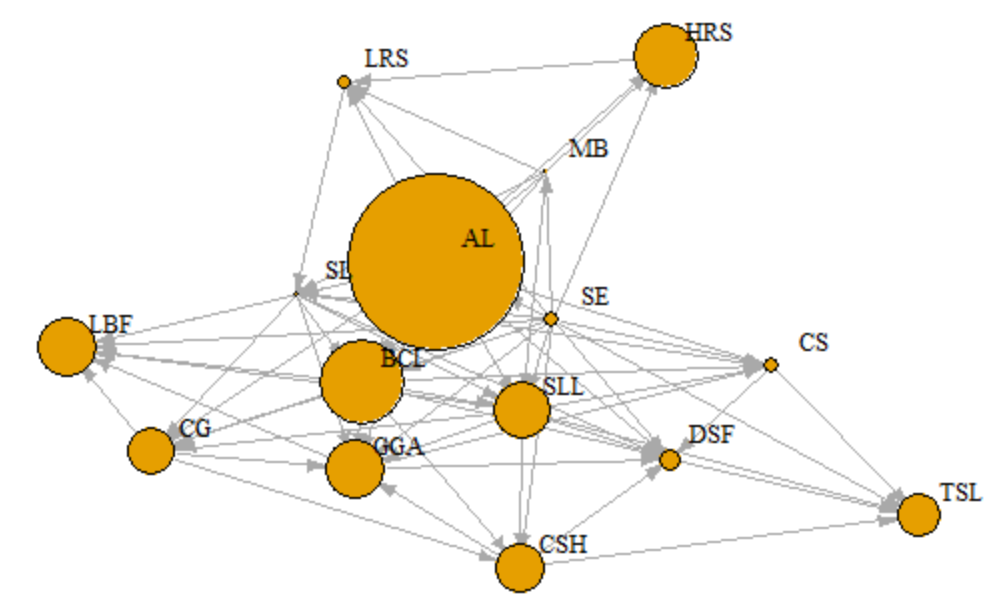
MALE



**Observation and explanation**

Among the male users the most influential based on the friend count as displayed by the size of the nodes and the number of arrows connected to them were; ALL, AAE, and ADL.

FEMALE

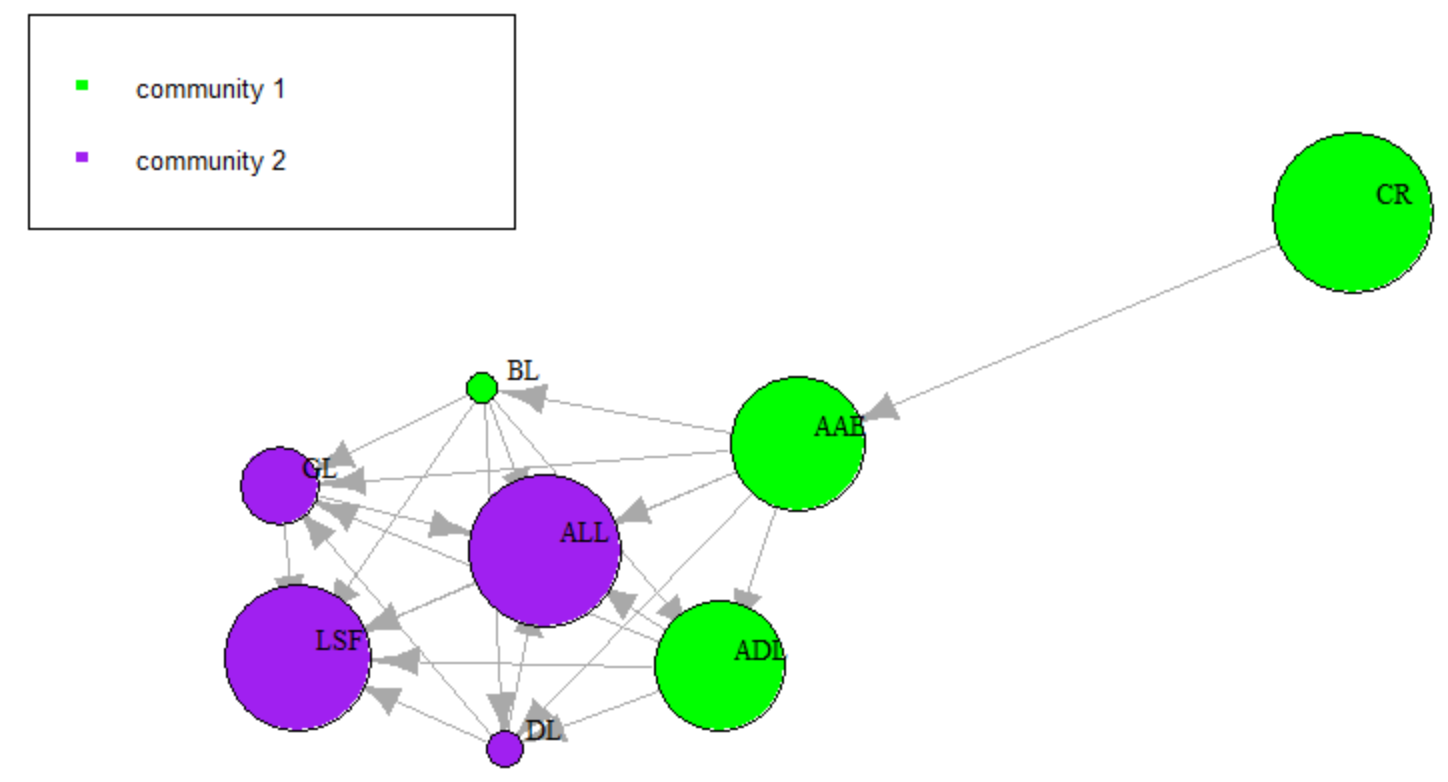


**Observation and explanation**

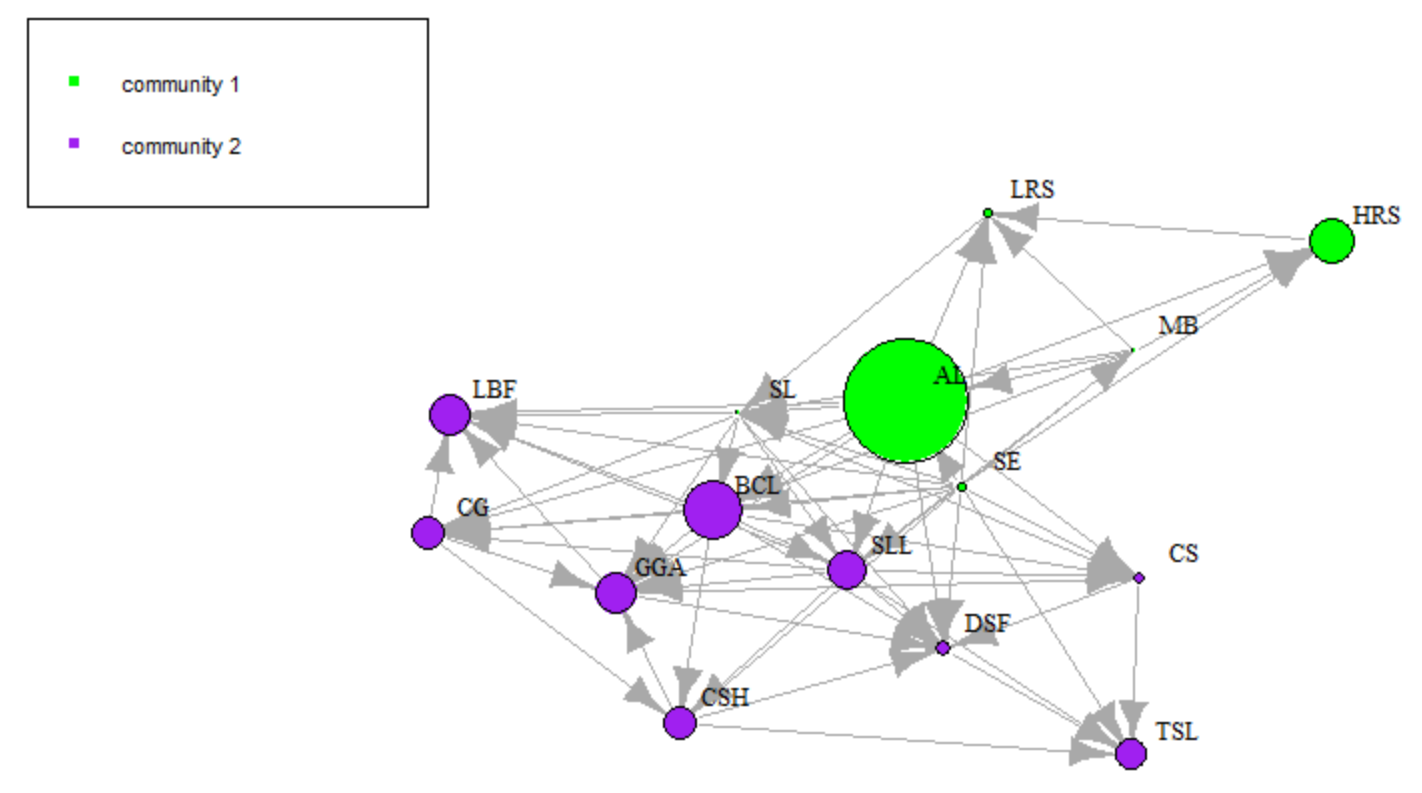
Among the female users the most influential based on the friend count as displayed by the size of the nodes and the number of arrows connected to them were; GGA, SLL, and CG. Then AL has the highest number of friends but is not a good authority since it has only one arrow pointing towards it. Whereas the SL had the lowest number of friends but was quite a good hub having many arrows pointing away from it.

1. **Detection of communities in each plot created in task 5, as described in task 6.**

**Communities among Male users**



**Communities among Female users**

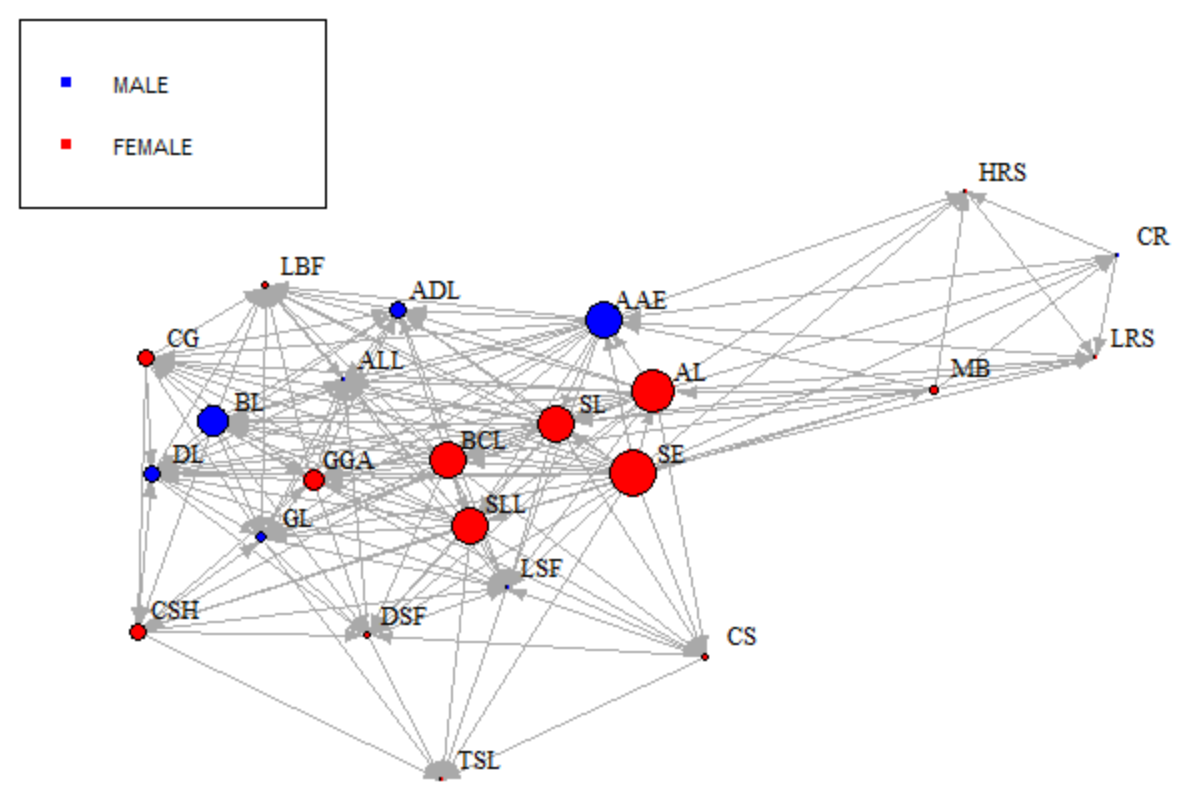


**Observation and explanation**

Male users were equally grouped among the 2 communities. Most female belonged to community 2.

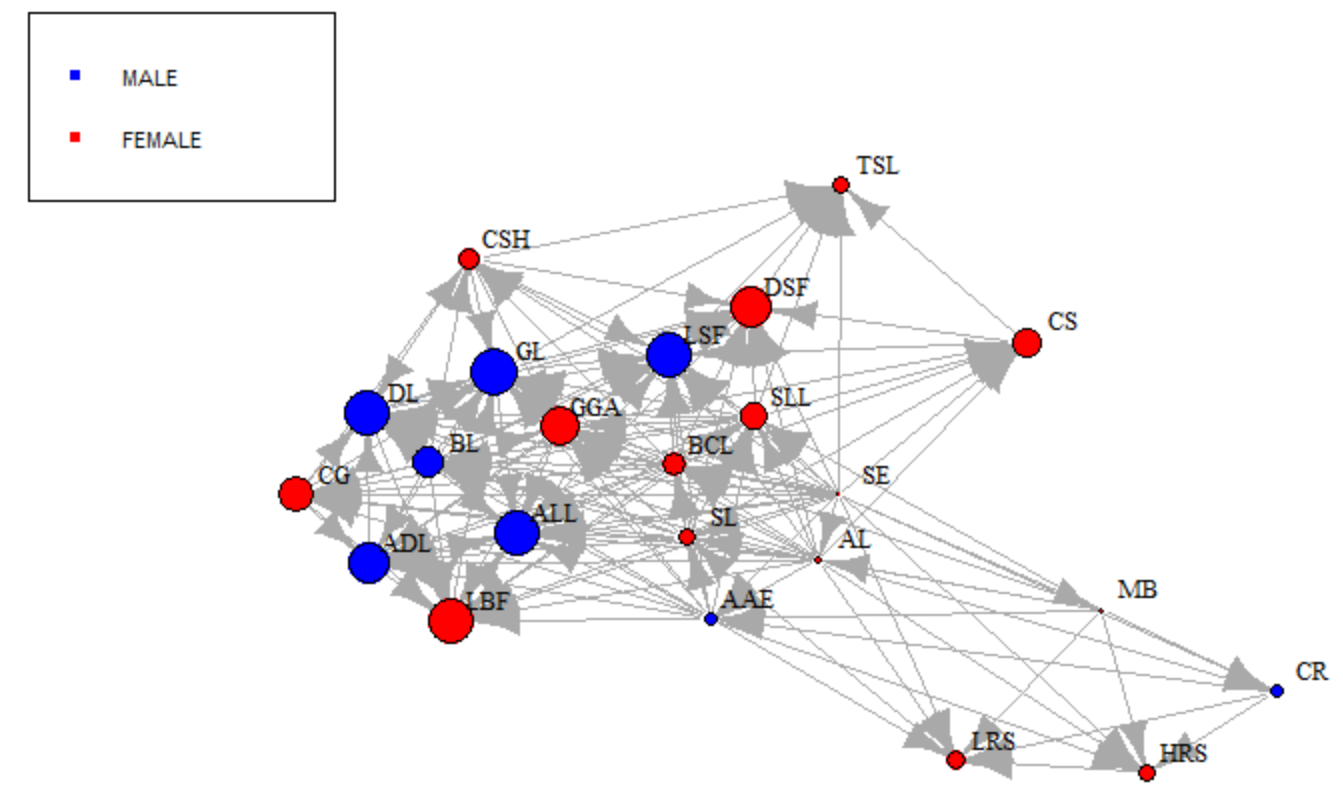
1. **Simplifying the network and calculating hub and authority scores and displaying these details in two networks plots, as described in task 7.**

**The plot based on hubs score multiplied by 10**



**Observation and explanation**

By visually inspecting the network displaying hubs, the 5 most important female Facebook users were; BCL, SL, SLL, AL and SE. These users had large nodes hence high hub scores. A high hub score node points to many good authorities and a high authority score node receives from many good hubs. The hubs that had the most number of arrows pointing out of them to other nodes were the most important.

**Plot based on the authorities scores multiplied by 10** 

**Observation and explanation**

Similarly the above plot on authority scores shows that the larger nodes have many arrows pointing towards them. Hence male users that visually look most important were LSF, GL, ALL, ADL and DL. Whereas their female counterparts were DSF, CS, GGA, LBF and LBF.

1. **Calculating measures of centrality and density, as described in task 8.**

The centrality degrees are as follows;

22, 9, 20, 6, 18, 6, 7, 18, 18, 19, 16, 13, 11, 16, 14, 16, 13, 17, 9, 14, 14, 15, 7

**Conclusion**

Network plotting requires a lot of adjustments to the settings in order to have visually good plot. After the analysis and several plots and breakdowns in terms of sex and groups the Facebook users were analyzed graphically and the following were the findings:

* Most users with many friends had few connections.
* There were some users who had no connection at all e.g. LLB, DW, and LBC.
* Most users with the highest number of friends were male, and the lowest number were female.
* The best hubs (those with many arrows point towards others), were female, while the best authorities (those with arrows pointing from others to them) were female.
* When the users are clustered, 2 communities are detected whereby community 1 consist, mostly of authorities (males) and community 2 consist of hubs (female).