Finding Least Number of Social Media Influencers that can reach to maximum people

Sangeeth Sajan Baby, Siji Rani S

Department of Computer Science, Amrita Vishwa Vidhyapeetham

November 11, 2021

Background

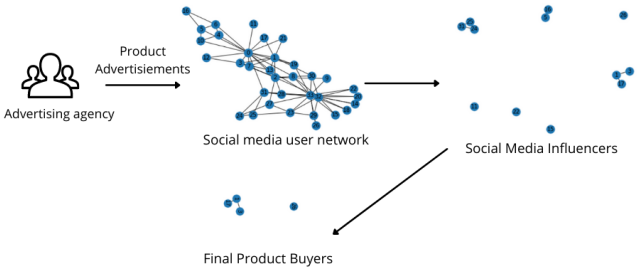

Figure: Advertising agencies are giving advertisements to the social media networks, where only a small amount of people will be highly influential, and a small group of people will be the real buyers.

Our Problem

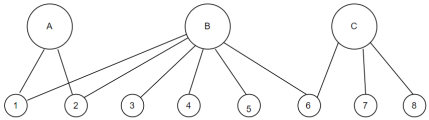

Figure: Network of the Influencers and Followers, A,B,C are influencers and 1-8 are followers

Influencers	Total Followers	Maximum Reach Reach
A	2	2
B	6	6
C	3	3
A,B	8	6
A,C	5	5
B,C	9	8
A,B,C	9	8

Which is the best set?

Figure: Set of Influencers that can reach to maximum followers

Solution Approach

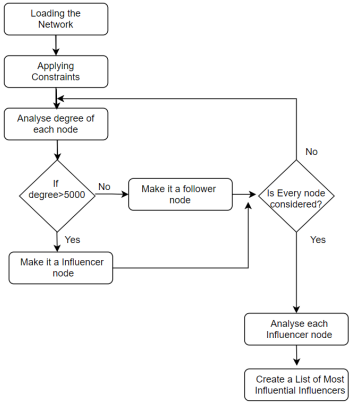

Figure: Major steps in process where we initially categorize the network based on the context and then find the Influencers, and then the least number of Influencers suitable

Solution Approach

UserId	UserName	Language
1	A	Malayalam
2	B	Hindi
3	C	French
4	D	Tamil

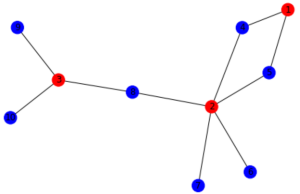

Figure: After applying the constraints, classifying the nodes based on follower count. Nodes in red are influencers and blue are followers

Implementation

```
1  newdf = newdf.select('*').where(col('Language')=='Malayalam')
2  display(newdf)
```

▸ (3) Spark Jobs

	User Node	Language
1	1	Malayalam
2	10	Malayalam
3	11	Malayalam
4	18	Malayalam
5	23	Malayalam
6	27	Malayalam
7	29	Malayalam

Showing all 7 rows.

Figure: Pseudo Code for applying the constraints to the nodes of the network

Implementation

```
1  newdegdic ={}
2  for node in G:
3    newdegdic[node]=G.degree[node]
4  newdegdic
```

```
Out[45]: {1: 0, 10: 0, 11: 0, 18: 0, 23: 2, 27: 1, 29: 1}
```

```
Command took 0.01 seconds -- by amenp2ari20039@am.students.amrita.edu at 11/11/2021, 1:03
```

```
Cmd 20
```

```
1  # Here the mininum number of followers neede are mentioned by the user(
2
3  finalnodelist = list((k) for k,v in newdegdic.items() if v >= 2)
```

Figure: Pseudo Code for categorizing the nodes as Influencers and Followers, the follower count will be assigned by the user.

Implementation

```
list0 = []
list1 = []
for node in G:
    k = G.degree[node]
    if(k>2):
        list0.append(node)
for node in list0:
    k = G.neighbors(node)
    for m in k:
        if m not in list1:
            list1.append(m)
print(list0)
print(list1)
```

Figure: Pseudo Code for finding the least number of Influencers

Observation

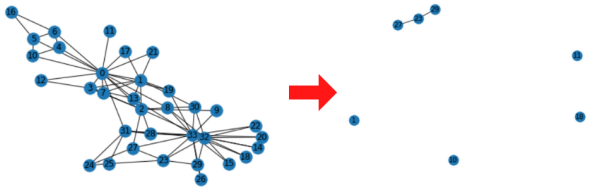

Figure: After applying the constrains, we will get the nodes where our constrains are satisfied.

Observation

```
Out[46]: {(): 0,
 (1,): 2,
 (24,): 2,
 (25,): 2,
 (31,): 2,
 (1, 24): 4,
 (1, 25): 4,
 (1, 31): 4,
 (24, 25): 3,
 (24, 31): 3,
 (25, 31): 3,
 (1, 24, 25): 5,
 (1, 24, 31): 5,
 (1, 25, 31): 5,
 (24, 25, 31): 3,
 (1, 24, 25, 31): 5}
```

Figure: Final output showing the combinations of influencers and their maximum reach.