

# SOCIAL NETWORK ANALYSIS

## COMMUNITY DETECTION ALGORITHM

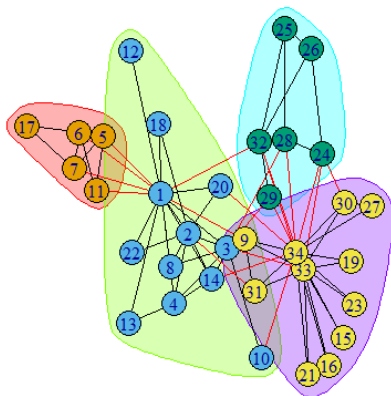
1. Different community algorithms applied on the karate.gml file.

Multiplelevel.community

```
> karate <- read.graph("karate.gml",format= "gml")
> wc <- multilevel.community(karate)
> modularity(wc)
[1] 0.4188034
> membership(wc)
[1] 2 2 2 2 1 1 1 2 4 2 1 2 2 2 4 4 1 2 4 2 4 2 4 3 3 3 4 3 3 4 4 3 4 4
> compare(wc,V(karate)$value,method="nmi")
[1] 0.5866348
> plot(wc,karate)
> plot(karate, vertex.color=membership(wc))
```

Modularity=0.4188034

Membership= 34 (2 2 2 2 1 1 1 2 4 2 1 2 2 2 4 4 1 2 4 2 4 2 4 3 3 3 4 3 3  
4 4 3 4 4)  
Qcs= 0.5866348



edge.betweenness.community

```
> wc <- edge.betweenness.community(karate)
> modularity(wc)
[1] 0.4012985
> membership(wc)
[1] 1 1 2 1 3 3 3 1 4 5 3 1 1 1 4 4 3 1 4 1 4 1 4 2 2 4 2 2 4 4 2 4 4
```

```
> compare(wc,v(karate)$value,method="nmi")
[1] 0.5798278
> plot(wc,karate)
> plot(karate, vertex.color=membership(wc))
```

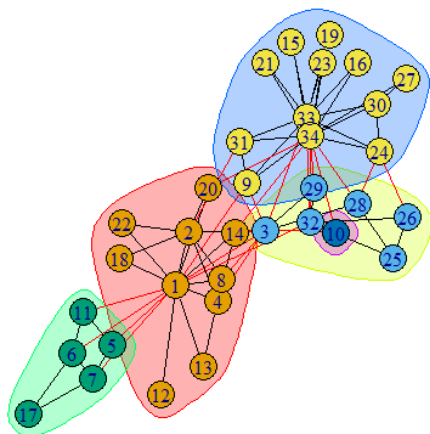
Modularity = 0.4012985

Membership = 34 (1 1 2 1 3 3 3 1 4 5 3 1 1 1 4 4 3 1 4 1 4 1 4 4 2 2 4 2 2 4 4 2 4 4)

The quality of obtained community structure compared to the ground-truth

Community structure is:

$Q_{cs} = 0.5798278$

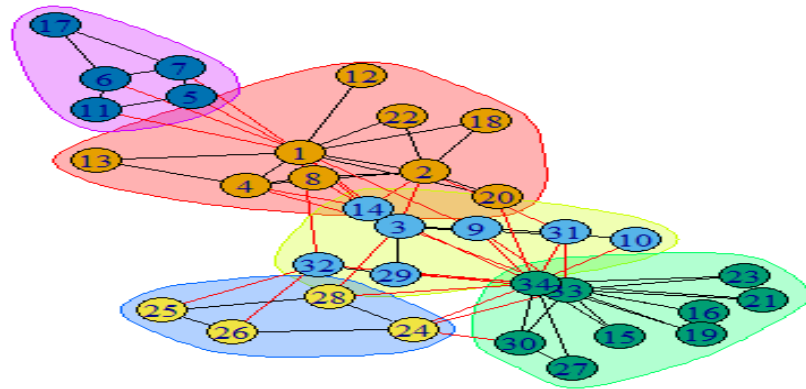


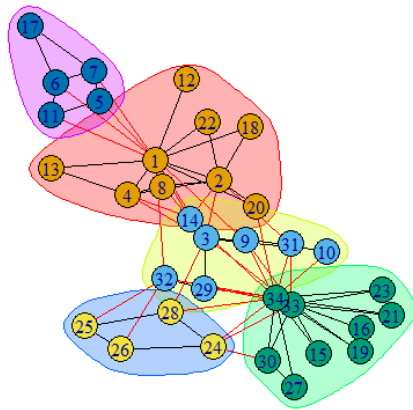
Walktrap.community

```
> wc <- walktrap.community(karate)
> modularity(wc)
[1] 0.3532216
> membership(wc)
[1] 1 1 2 1 5 5 5 1 2 2 5 1 1 2 3 3 5 1 3 1 3 1 3 4 4 4 3 4 2 3 2 2 3 3
> compare(wc,v(karate)$value,method="nmi")
[1] 0.504178
> plot(wc,karate)
> plot(karate, vertex.color=membership(wc))
```

Modularity = 0.3532216

Membership = 34 (1 1 2 1 5 5 5 1 2 2 5 1 1 2 3 3 5 1 3 1 3 1 3 4 4 4 3 4 2 3 2 2 3 3)

$$Q_{cs} = 0.504178$$

$$Q_{cs} = 0.6994882$$



## Conclusion

Base on the different results on the karate file, I can conclude that the Infomap.community algorithm with  $Q_{cs} = 0.6994882$  is the best community detection algorithm on the karate file.

2. Different community algorithms applied on the dolphins.gml file.

```
dolphins <- read.graph("dolphins.gml",format= "gml")
> wc <- multilevel.community(dolphins)
> modularity(wc)
[1] 0.5185317
> membership(wc)
[1] 1 2 1 5 5 4 4 2 5 4 1 5 3 4 3 5 3 4 5 2 3 5 4 5 5 2 2 2 2 5 2 4 4 3 3
5 3 3 3 3 3 4 1 3 3 5 3 1 4 3 3 5 3 1 4 5 4 4 3 5 4 1
> compare(wc,v(dolphins)$value,method="nmi")
[1] 0.5108534
> plot(wc,dolphins)
> plot(dolphins, vertex.color=membership(wc))
```

Multilevel.community

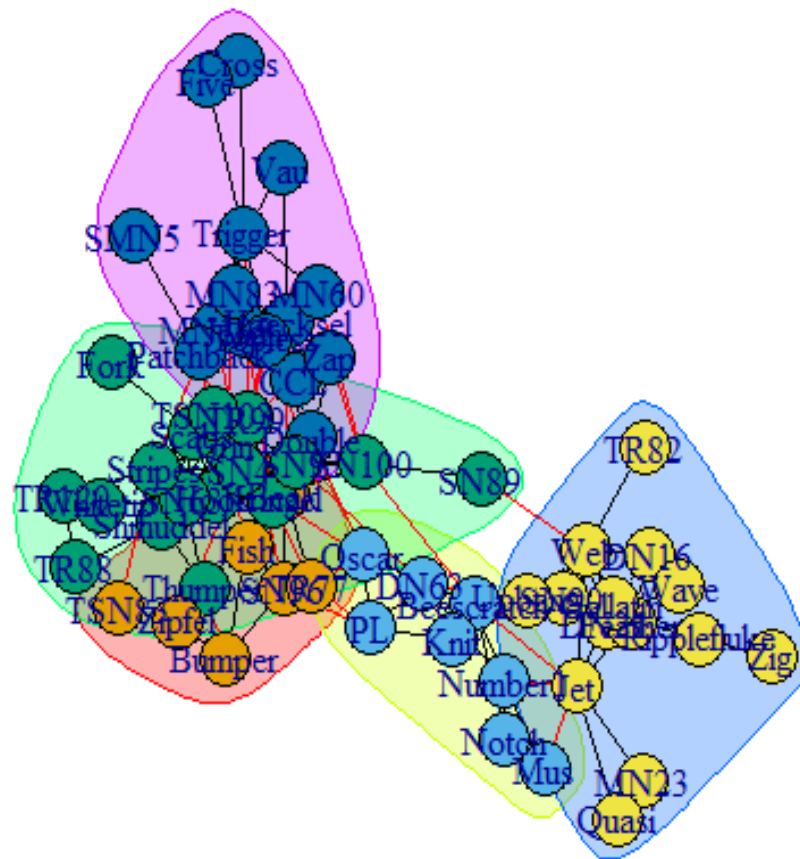
Modularity = 0.5185317

Membership = 60 (1 2 1 3 4 2 2 3 2 1 4 3 2 3 4 3 2 4 2 3 4 2 4 4 2 2 2 1 4 1 2 2 3 3 4 3 3 3 2  
3 2 1  
3 3 4 3 1 2 3 3 4 3 5 2 4 2 2 3 3 2 5)

The quality of obtained community structure compared to the ground-truth

Community structure is:

$Q_{cs} = 0.5108534$



```
edge.betweenness.community
```

```
> wc <- edge.betweenness.community(dolphins)
> modularity(wc)
[1] 0.5193821
> membership(wc)
[1] 1 2 1 3 4 2 2 2 3 2 1 4 3 2 3 4 3 2 4 2 3 4 2 4 4 2 2 2 1 4 1 2 2 3 3
4 3 3 3 2 3 2 1 3 3 4 3 1 2 3 3 4 3 5 2 4 2 2 3 3 2 5
> compare(wc,v(dolphins)$value,method="nmi")
[1] 0.5541605
> plot(wc,dolphins)
> plot(dolphins, vertex.color=membership(wc))
```

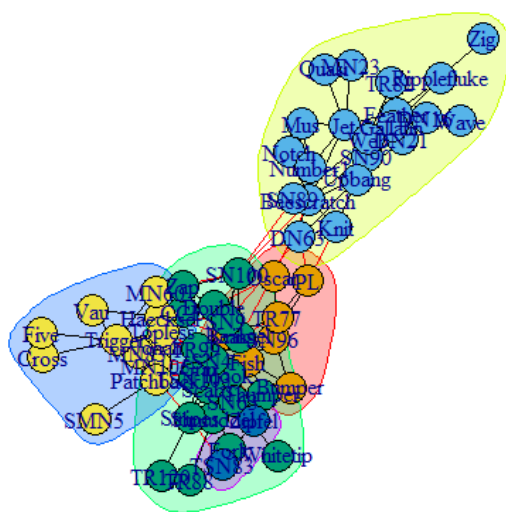
Modularity = 0.5193821

Membership = 60 ( 1 2 1 3 4 2 2 2 3 2 1 4 3 2 3 4 3 2 4 2 3 4 2 4 4 2 2 2 1 4 1 2 2 3 3 4 3 3 3  
 2 3 2 1  
 3 3 4 3 1 2 3 3 4 3 5 2 4 2 2 3 3 2 5)

The quality of obtained community structure compared to the ground-truth

Community structure is :

$Q_{cs} = 0.5541605$



Walktrap.community

```
> wc <- walktrap.community(dolphins)
> modularity(wc)
[1] 0.4888454
> membership(wc)
[1] 2 1 2 3 3 1 1 1 3 1 2 3 2 1 2 3 2 1 3 1 2 3 1 3 3 1 1 1 1 3 1 1 4 2 2
3 3 2 2 1 2 1 2 2 2 3 2 2 1 2 2 3 2 2 1 3 1 1 2 3 4 2
> compare(wc,v(dolphins)$value,method="nmi")
[1] 0.53725
> plot(wc,dolphins)
> plot(dolphins, vertex.color=membership(wc))
```

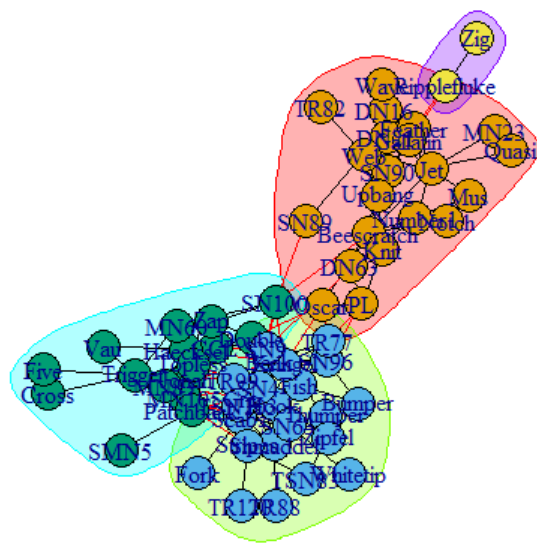
Modularity = 0.4888454

Membership = 60 (2 1 2 3 3 1 1 1 3 1 2 3 2 1 2 3 2 1 3 1 2 3 1 3 3 1 1 1 1 3 1 1 4 2 2 3 3 2 2 1  
 2 1 2  
 2 2 3 2 2 1 2 2 3 2 2 1 3 1 1 2 3 4)

The quality of obtained community structure compared to the ground-truth

Community structure is:

$Q_{cs} = 0.53725$



Infomap.community

```
> wc <- infomap.community(dolphins)
> modularity(wc)
[1] 0.518947
> membership(wc)
[1] 5 4 5 6 3 2 2 4 6 2 5 3 1 2 1 3 1 2 3 4 1 3 2 3 3 4 4 4 4 3 4 2 2 1 1
3 6 1 1 6 1 2 5 1 1 3 1 5 2 1 1 3 1 1 2 3 2 2 1 6 2 1
> compare(wc,v(dolphins)$value,method="nmi")
[1] 0.4811711
> plot(wc,dolphins)
> plot(dolphins, vertex.color=membership(wc))
```

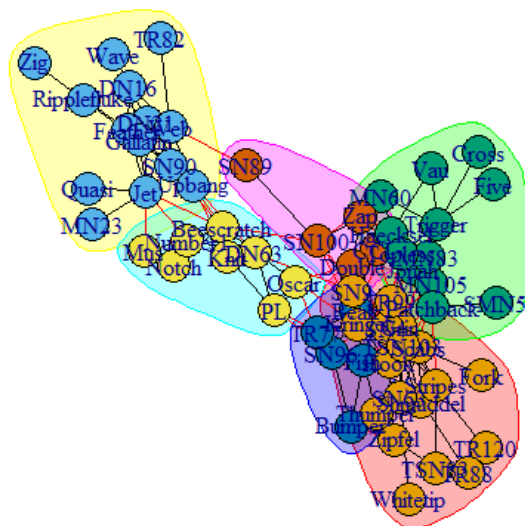
Modularity = 0.518947

Membership = 60 (1 2 1 3 4 2 2 2 3 2 1 4 3 2 3 4 3 2 4 2 3 4 2 4 4 2 2 2 1 4 1 2 2 3 3 4 3 3 3 2  
 3 2 1  
 3 3 4 3 1 2 3 3 4 3 5 2 4 2 2 3 3 2 5)

The quality of obtained community structure compared to the ground-truth

Community structure is:

$Q_{cs} = 0.4811711$



## Conclusion

Base on the different results on the dolphins file, I can conclude that the edge.betweenness.community algorithm with  $Q_{cs} = 0.5541605$  is the best community detection algorithm on the dolphins file.

3. Different community algorithms applied on the football.gml file.

Multilevel.community

```
> football <- read.graph("football.gml", format= "gml")
> wc <- multilevel.community(football)
> modularity(wc)
[1] 0.6020821
```



```

> membership(wc)
[1] 9 2 1 4 9 4 1 9 9 9 4 4 8 1 8 1 9 6 8 5 6 9 9 9 4 2 8 6 4 5 5 8 1 2
8 5 6 2 8 1 4 9 8 8 7 2 3 1 7 3 4 9 4 3 8 5 6 7 6 6 1 8 6 6 1 6 7 3 9 4 6
8 4 3 4 7
[77] 6 9 9 5 5 4 5 3 4 8 7 6 3 2 4 7 7 9 5 6 6 6 6 4 8 1 5 4 2 9 2 1 4 9 2
3 9 7 6 3
> compare(wc,v(football)$value,method="nmi")
[1] 0.8549734
> plot(wc,football)
> plot(football, vertex.color=membership(wc))

```

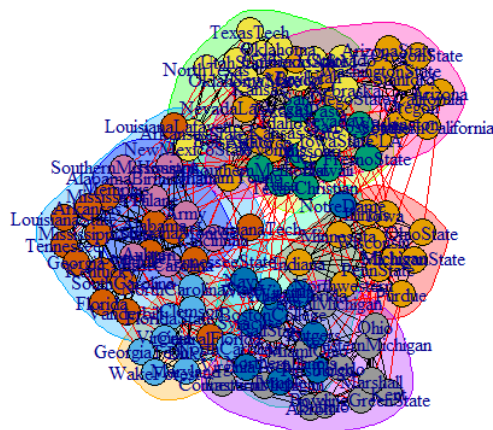
Modularity = 0.6020821

Membership = (9 2 1 4 9 4 1 9 9 9 4 4 8 1 8 1 9 6 8 5 6 9 9 9 4 2 8 6 4 5 5  
8 1 2 8 5 6 2 8 1 4 9 8 8 7 2 3 1 7 3 4 9 4 3 8 5 6 7 6 6 1 8 6 6 1 6 7 3  
9 4 6 8 4 3 4 7 6 9 9 5 5 4 5 3 4 8 7 6 3 2 4 7 7 9 5 6 6 6 4 8 1 5 4 2 9  
2 1 4 9 2 3 9 7 6 3)

The quality of obtained community structure compared to the ground-truth

Community structure is:

Qcs = 0.8549734



edge.betweenness.community

```

> wc <- edge.betweenness.community(football)
> modularity(wc)
[1] 0.6005129
> membership(wc)

```



```

> membership(wc)
  [1] 1 10 9 7 1 7 9 8 8 1 7 1 4 9 4 9 1 5 4 3 5 8 8
1 1 10 4 5 1 3 3 4 9 10 4 3 2 10 4 9 7 1 4 4 2 10 6 9
2 6 1
 [52] 8 7 6 4 3 5 2 2 2 9 4 5 2 9 5 2 6 8 1 5 4 7 6
7 2 5 8 8 3 3 7 3 6 7 4 2 5 6 10 1 2 2 1 3 5 5 2 7
4 9 3
[103] 7 10 1 10 9 7 8 10 6 8 2 5 6
> compare(wc,v(football)$value,method="nmi")
[1] 0.8873604
> plot(wc,football)
> plot(football, vertex.color=membership(wc))

```

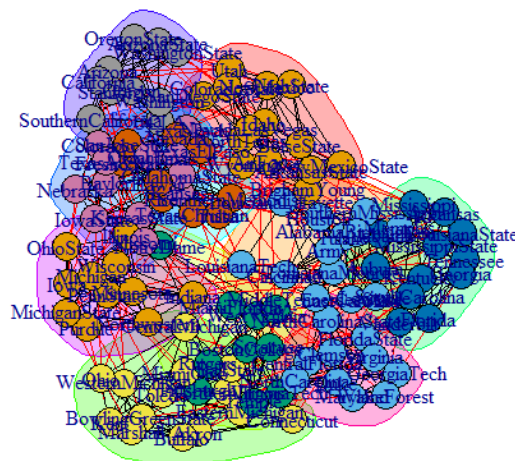
Modularity = 0.6038112

Membership = 60 (2 1 2 3 3 1 1 1 3 1 2 3 2 1 2 3 2 1 3 1 2 3 1 3 3 1 1 1 1 3 1 1 4 2 2 3 3 2 2 1  
2 1 2  
2 2 3 2 2 1 2 2 3 2 2 1 3 1 1 2 3 4)

The quality of obtained community structure compared to the ground-truth

Community structure is:

Qcs = 0.8873604



Infomap.community

```

> wc <- infomap.community(football)
> modularity(wc)

```

```

[1] 0.5777064
> membership(wc)
[1] 10 8 3 12 10 4 3 5 5 10 4 11 1 3 1 3 10 2 1 6 2 5 5
10 11 8 1 2 11 6 6 1 3 8 1 6 1 8 1 3 4 10 1 1 9 8 7 3
9 7 11
[52] 5 4 7 1 6 2 9 12 12 3 1 2 12 3 2 9 7 5 11 2 1 4 7
4 9 2 5 5 6 6 4 6 7 4 1 9 2 7 8 11 9 9 10 6 2 2 12 4
13 3 6
[103] 4 8 10 8 3 4 5 8 7 5 9 2 7
> compare(wc,V(football)$value,method="nmi")
[1] 0.911059
> plot(wc,football)
> plot(football, vertex.color=membership(wc))

```

Modularity = 0.5777064

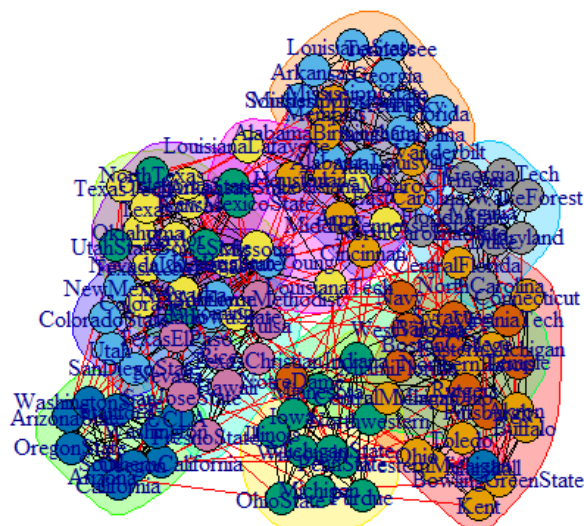
Membership = 60 (1 2 1 3 4 2 2 3 2 1 4 3 2 3 4 3 2 4 2 3 4 2 4 4 2 2 2 1 4 1 2 2 3 3 4 3 3 3 2  
3 2 1

3 3 4 3 1 2 3 3 4 3 5 2 4 2 2 3 3 2 5)

The quality of obtained community structure compared to the ground-truth

Community structure is:

Qcs = 0.911059



Conclusion

Based on the different results on the football file, I can conclude that edge.betweeness

community algorithm with  $Q_{cs} = 0.8788884$  is the best community detection algorithm on the footballs file.

2. Download the following file wikipedia.gml from <http://lipn.fr/~kanawati/ars>. This is a snapshot of the wikipedia network. Nodes have a label attribute that gives the title of the wikipedia page. Apply different community detection algorithms on this network and evaluate the outcome using the label attribute

For the second question it gives a longer time on even the community detection algorithm

KODAVATI SAI NARAYANA PHANINDRA

EISTI, ADEO2