Homework 03

MO412 - Network Science

Margarita Lacuaña Apaza
RA: 234924

April 2021

1 Select at random 100 nodes v0, v1, . . . , v99 from this network. Hand in a plain text file (no formatting) with 100 lines of the form

```
import pandas as pd
import numpy as np
import networkx as nx
import matplotlib.pyplot as plt

net = pd.read_csv("net1000-005.tsv",header=None,sep='\t')
graph = nx.from_pandas_edgelist(net, 0, 1)

# ---------- Section a ---------- #

sub_graph = net.to_numpy()
sub_graph = sub_graph[np.random.choice(len(sub_graph),
    size=100, replace=False)]
nodes = []
with open('nodes', 'w') as f:
    for i in range(100):
        f.write(str(i)+" "+str(sub_graph[i][0])+"\n")
        nodes.append(sub_graph[i][0])
#nodes
```

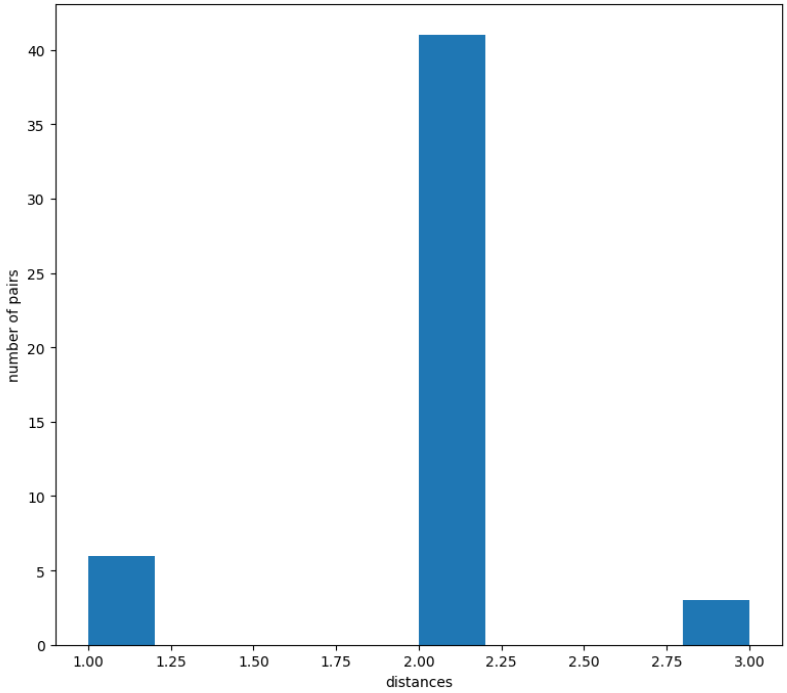
2 Compute shortest paths from v2i to v2i+1, for i ranging from 0 to 49. Hand in a text file with your paths, one per line, with nodes separated by whitespace

```
# ———————————— Section b ———————————— #
distances = []
def shortestPath(i,j):
    path = nx.bidirectional_shortest_path(graph,i,j)
    resp = ""
    distances.append(len(path)-1)
    for i in path:
        resp += str(i)+" "
    return resp

with open('paths', 'w') as f:
    for i in range(50):
        path = shortestPath(nodes[2*i],nodes[2*i+1])
        f.write(path+"\n")
#distances
```

3 Plot a distribution of the distances you found in item 1b. Hand in this plot

```
# ———————————— Section c ———————————— #
distances = np.array(distances)
plt.xlabel('distances')
plt.ylabel('number of pairs')
plt.hist(distances)
plt.show()
```


number of pairs
distances
0
5
10
15
20
25
30
35
40
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00