

Assignment 1

Jacky(Hsiu-Hsuan) Yeh

Table of contents

1. Generate the network (with 12 nodes) with the given edge set, complete the corresponding analysis and visualization tasks.	1
1) Report the adjacency matrix, and plot this network with labels.	3
2) Find the node(s) with maximum betweenness, what's the maximum betweenness value?	4
2. Import the friendship network in a class represented by an adjacency matrix "class" into R:	5
1) What is the highest page rank value among nodes?	5
2) What is the size of the largest component in this network?	5
3) What is the global clustering coefficient and assortivity of outdegree in this network?	6

```
import cudgraph, os, cudf
import graphistry
import networkx as nx

import warnings
warnings.filterwarnings("ignore")
graphistry.register(
    personal_key_id = 'R7THUYM8LJ',
    personal_key_secret = os.getenv('GRAPHISTRY_API_KEY')
)
```

1. Generate the network (with 12 nodes) with the given edge set, complete the corresponding analysis and visualization tasks.

Edge set: - 1 and 2, 1 and 3, 1 and 4, 1 and 8, 1 and 10, 1 and 11, 1 and 12 - 2 and 3, 2 and 4, 2 and 5, 2 and 6 - 3 and 5, 3 and 6 - 4 and 7, 4 and 9, 4 and 10, 4 and 12 - 5 and 9, 5 and 11 - 6 and 12 - 7 and 11 - 8 and 12 - 9 and 10 - 11 and 12

```

source = [1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 3, 3, 4, 4, 4, 4, 5, 5, 6, 7, 8, 9, 11]
dest = [2, 3, 4, 8, 10, 11, 12, 3, 4, 5, 6, 5, 6, 7, 9, 10, 12, 9, 11, 12, 11, 12, 10, 12]

edge_list = cudf.DataFrame({
    'source': source,
    'destination': dest,
})

edge_list['weights'] = 1

edge_list

```

	source	destination	weights
0	1	2	1
1	1	3	1
2	1	4	1
3	1	8	1
4	1	10	1
5	1	11	1
6	1	12	1
7	2	3	1
8	2	4	1
9	2	5	1
10	2	6	1
11	3	5	1
12	3	6	1
13	4	7	1
14	4	9	1
15	4	10	1
16	4	12	1
17	5	9	1
18	5	11	1
19	6	12	1
20	7	11	1
21	8	12	1
22	9	10	1
23	11	12	1

1) Report the adjacency matrix, and plot this network with labels.

```
G = cudagraph.Graph(directed=True)
G.from_cudf_edgelist(
    edge_list,
    source = 'source',
    destination = 'destination',
    edge_attr = 'weights',
    renumber = False,
)

G.to_pandas_adjacency()
```

	9	2	4	1	11	7	8	3	5	6	10	12
9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0
2	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0	1.0	0.0	0.0
4	1.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	1.0	1.0
1	0.0	1.0	1.0	0.0	1.0	0.0	1.0	1.0	0.0	0.0	1.0	1.0
11	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
7	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	1.0	0.0	0.0
5	1.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
10	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
12	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

```
(
    graphistry
    .bind(source='source', destination='destination', edge_title="weights")
    .edges(edge_list)
    .plot()
)
```

<IPython.core.display.HTML object>

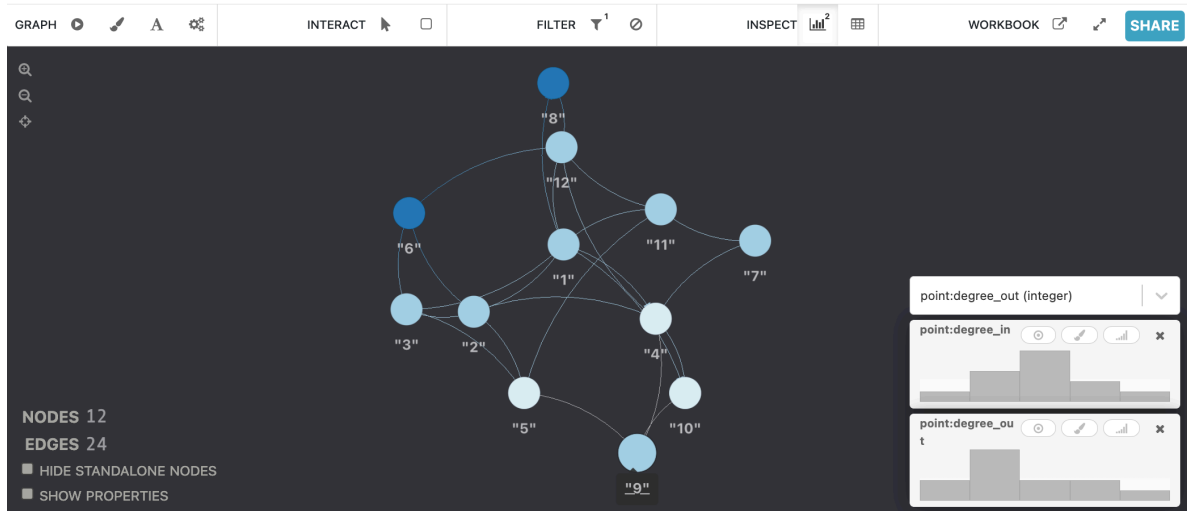


Figure 1: network_with_label

2) Find the node(s) with maximum betweenness, what's the maximum betweenness value?

the maximum betweenness value is 0.037879

```
(
  cugraph
    .betweenness centrality(G)
    .sort_values('betweenness centrality', ascending=False)
)
```

	vertex	betweenness centrality
4	4	0.037879
5	5	0.034091
9	9	0.015152
11	11	0.015152
6	6	0.011364
2	2	0.007576
3	3	0.007576
7	7	0.007576
0	0	0.000000
1	1	0.000000
8	8	0.000000

	vertex	betweenness centrality
10	10	0.000000
12	12	0.000000

2. Import the friendship network in a class represented by an adjacency matrix “class” into R:

```
df = cudf.read_csv('class.csv', header=None)
G = cudgraph.Graph(directed=True)
G.from_numpy_array(df.to_numpy())
```

1) What is the highest page rank value among nodes?

the highest page rank value is 0.003985

```
cudgraph.pagerank(G).sort_values('pagerank', ascending=False)
```

	vertex	pagerank
36	541	0.003985
2	810	0.003886
80	1360	0.003732
9	581	0.003705
56	875	0.003429
...
1568	1560	0.000152
1569	1564	0.000152
1570	1567	0.000152
1571	1580	0.000152
1572	1581	0.000152

2) What is the size of the largest component in this network?

the size of the largest component is 1214

```
components = cudgraph.connected_components(G, connection='strong')
components.groupby('labels').size().sort_values(ascending=False)
```

```

labels
0      1214
337      3
644      3
909      2
418      2
...
8        1
868      1
1407     1
1298     1
1186     1
Length: 348, dtype: int64

```

3) What is the global clustering coefficient and assortivity of outdegree in this network?

the global clustering coefficient is 0.13869634654329321 and the assortivity of outdegree is 0.19815374096688262

```

G = nx.from_numpy_array(df.to_numpy())
nx.transitivity(G)

```

0.13869634654329321

```

nx.degree_assortativity_coefficient(G, x='out', y='out')

```

0.19815374096688262