



Social Networks and Computer Networks

Co-membership of Clients on a Network

J. T. Rigsby and J. L. Solka

`rigsbyjt@nswc.navy.mil; solkajl@nswc.navy.mil`

Naval Surface Warfare Center Dahlgren Division

Agenda

- Social Networks
- The Network
- Client Co-membership
- Wrap-up and conclusions.

Acknowledgments

- This work is supported by the In-house Laboratory Independent Research (ILIR) program.
- We wish to acknowledge helpful discussions with Dr. David Marchette of NSW CDD and Dr. Wendy Martinez of ONR.

Why?

- Short term
 - Prove the network data is or is not random
 - Show proof of concept for anomaly detection
 - Over Time
- Long term
 - Build network topology maps of trust structures
 - Show changes over time
 - Threat deterrence through awareness of changes

Social Network Analysis

- Mathematically describe sociological data.
- Used mainly by sociologists and archeologists.
- Is made of:
 - Actors
 - Events

Modality

- One Mode Network
 - One set of actors or events
 - Internal relationships
- Two Mode Network
 - One set of actors and one set of events
 - Two sets of actors
 - Two sets of events

Social / Computer Network

- Is made of:
 - Computers
 - Clients
 - Servers
- People talk / Computers make connections
- People have commodities / Servers run services
- People consume commodities / Clients use services

Our Network

- 5 users
- 11 computers
- Oct 2002 - February 2003

The Data

- 1899 total servers accessed over 5 months
- Averaged 576 servers accessed per month
- 14 total clients
- Averaged 10 clients

Data Matrices

- October 492 by 10
- November 514 by 9
- December 449 by 8
- January 778 by 10
- February 648 by 11
- Total 1899 by 14

Data Sparseness

- December 449 by 8
- 557 non zero values
- 3035 zero values
- 85% zeros
- Averaged 70 servers per client
- Max 284 Min 3

Clients

- Primary Clients
- Secondary Clients
- Other Clients

Grouping on Commonality

- Commonality value
- Actor Relative Commonality Value
- Actor Pair Relative Commonality Value

December Co-membership

DEC	1	2	3	4	5	6	7	8
1	284	25	6	29	12	18	6	0
2	25	58	3	5	2	5	0	0
3	6	3	12	1	1	1	0	3
4	29	5	1	79	6	9	4	0
5	12	2	1	6	34	8	2	1
6	18	5	1	9	8	67	2	0
7	6	0	0	4	2	2	20	0
8	0	0	3	0	1	0	0	3

December Co-membership Plot

December Comembership

	1	2	3	4	5	6	7	8
1	1							
2		2						
3			3					
4				4				
5					5			
6						6		
7							7	
8								8

Relative Co-membership

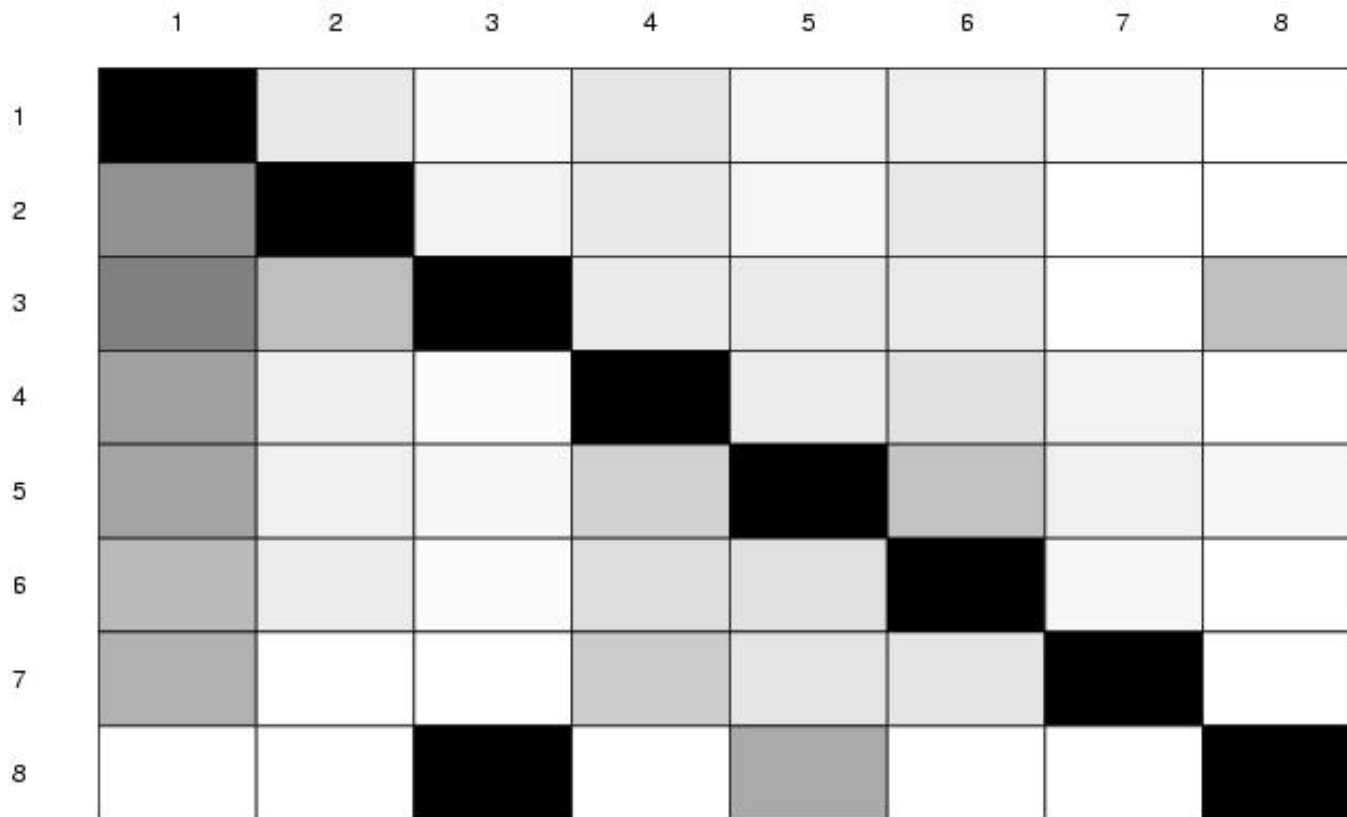
- Normalize data based on Actor
- Divide each matrix value by diagonal value

$$a_{x,y} = \frac{a_{x,y}}{a_{x,x}}$$

- Each row is relative to that actor
- Can see reciprocation of commonality
- Not Symmetric

Relative Dec. Co-membership Plot

December Client Relative Comembership



Pair Relative Co-membership

- Normalize data based on Actor Pairs
- Divide each matrix value by the sum of the 2 associated actor's diagonal values

$$a_{x,y} = \frac{a_{x,y}}{a_{x,x} + a_{y,y}}$$

- Each value is relative to both actors
- Symmetric

Pair Relative Dec. Co-membership

December Client Pair Relative Comembership

	1	2	3	4	5	6	7	8
1	Black	Light Gray	White	Light Gray	Light Gray	Light Gray	Light Gray	White
2	Light Gray	Black	Light Gray	Light Gray	White	Light Gray	White	White
3	White	Light Gray	Black	White	White	White	White	Dark Gray
4	Light Gray	Light Gray	White	Black	Light Gray	Light Gray	Light Gray	White
5	Light Gray	White	White	Light Gray	Black	Light Gray	Light Gray	Light Gray
6	Light Gray	Light Gray	White	Light Gray	Light Gray	Black	Light Gray	White
7	Light Gray	White	White	Light Gray	Light Gray	White	Black	White
8	White	White	Dark Gray	White	Light Gray	White	White	Black

Whole Co-membership Plot

October–February Comembership

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1													
2		2												
3			3											
4				4										
5					5									
6						6								
7							7							
8								8						
9									9					
10										10				
11											11			
12												12		
13													13	
14														14

Whole Relative Co-membership

October–February Client Relative Comembership

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Black	Light	Light	White	Light	Light	White	White	Light	White	White	White	White	White
2	Light	Black	Light	White	Light	White	White	White	Light	White	White	White	White	White
3	Light	Light	Black	White	Light	White	White	White	Light	White	White	White	White	White
4	Light	Light	Light	Black	White	Light	White	Light	Light	White	White	White	White	White
5	Light	Light	Light	White	Black	White	White	White	Light	White	White	White	White	White
6	Light	Light	Light	Light	Light	Black	White	White	White	Light	White	White	White	White
7	Light	Light	Light	Light	Light	White	Black	White	White	White	White	White	White	White
8	Light	Light	Light	Light	Light	Light	Light	Black	Light	Light	Light	White	White	White
9	Light	Light	Light	Light	Light	White	White	Light	Black	White	Light	White	White	White
10	White	Light	Light	Light	White	Light	Light	Light	White	Black	White	White	White	White
11	Light	Light	Light	Light	Light	Light	White	Light	Light	White	Black	White	White	White
12	White	White	White	White	White	White	White	White	White	White	White	Black	Black	Black
13	White	White	White	White	White	White	White	White	White	White	White	Black	Black	Black
14	White	White	White	White	White	White	White	White	White	White	White	Black	Black	Black

Whole Pair Relative Co-membership

October–February Client Pair Relative Comembership

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	Black	Light	Light	White	Light	Light	Light	White	Light	White	Light	White	White	White
2	Light	Black	Light	White	Light	Light	Light	White	Light	White	Light	White	White	White
3	Light	Light	Black	White	Light	Light	Light	White	Light	White	Light	White	White	White
4	White	White	Light	Black	White	Light	Light	White	Light	White	Light	White	White	White
5	Light	Light	Light	White	Black	White	Light	White	Light	White	Light	White	White	White
6	White	White	White	Light	White	Black	Light	White	Light	White	Light	White	White	White
7	White	White	Light	Light	White	Light	Black	Light	White	Light	White	White	White	White
8	White	White	White	Light	White	White	Light	Black	Light	Light	Light	White	White	White
9	Light	Light	Light	Light	Light	White	White	Light	Black	White	Light	White	White	White
10	White	White	White	White	White	White	Light	Light	White	Black	White	White	White	White
11	White	White	Light	Light	Light	Light	White	Light	Light	White	Black	White	White	White
12	White	White	White	White	White	White	White	White	White	White	White	Black	Black	Black
13	White	White	White	White	White	White	White	White	White	White	White	Black	Black	Black
14	White	White	White	White	White	White	White	White	White	White	White	Black	Black	Black

Hypergeometric Distribution

- N objects
- M objects of interest out of the N
 - Duds
 - Defectives
- n items are chosen at random
- X is the number of duds out of the n

Hypergeometric Distribution?

- N objects
 - Total number of servers gone to that month
- M objects or interest out of the N
 - Total number of servers gone to by one client
- n items are chosen at random
 - Total number of servers gone to by another client
- X is the number of duds out of the n
 - Total number of servers gone to by both clients

Distribution

- Probability distribution

$$P(X = x) = h(x; n, M, N) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}}$$

- Expected Value

$$E(X) = n \left(\frac{M}{N} \right)$$

Random Data

- Are the servers people going to randomly associated?
- How does sample size differences affect this?

What to Expect?

- Users that are all employees of the same company
- Users that have multiple clients
- Users that work on similar or dissimilar topics
- Personal Surfing
- Under utilized machines

Whole Co-membership Plot

October–February Comembership

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1													
2		2												
3			3											
4				4										
5					5									
6						6								
7							7							
8								8						
9									9					
10										10				
11											11			
12												12		
13													13	
14														14

Expected Whole Comem. Plot

October–February Expected Hypergeometric Mean

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1														
2		2												
3			3											
4				4										
5					5									
6						6								
7							7							
8								8						
9									9					
10										10				
11											11			
12												12		
13													13	
14														14

Probability Values of Our Data

- One Client went to M sites
- Second Client went to n sites
- Probability of overlap or intersection

Whole Probability Values

October-February Hypergeometric P-value

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	1													
2		2												
3			3											
4				4										
5					5									
6						6								
7							7							
8								8						
9									9					
10										10				
11											11			
12												12		
13													13	
14														14

Detection over Time OCT

October Client Relative Comembership

	1	2	3	4	5	6	7	8	9	10
1	Black	White	Light Gray	White	Light Gray	White	Light Gray	White	Light Gray	White
2	Light Gray	Black	Dark Gray	Dark Gray	Light Gray	Dark Gray	Dark Gray	Dark Gray	White	White
3	Light Gray	White	Black	Light Gray	White	Light Gray	Light Gray	White	White	White
4	Light Gray	Light Gray	Light Gray	Black	Light Gray	White	Light Gray	Light Gray	White	White
5	Dark Gray	White	Light Gray	Light Gray	Black	Light Gray	Light Gray	White	White	Light Gray
6	Light Gray	Light Gray	Light Gray	White	Light Gray	Black	White	White	White	Light Gray
7	Light Gray	Light Gray	Light Gray	Light Gray	White	White	Black	Black	Light Gray	White
8	White	Black	Dark Gray	Dark Gray	White	White	Black	Black	Black	White
9	Dark Gray	White	White	White	White	White	White	White	Black	White
10	White	White	Light Gray	White	Dark Gray	Dark Gray	White	White	White	Black

Detection over Time NOV

November Client Relative Comembership

	1	2	3	4	5	6	7	8	9
1	Black	Light Gray	Light Gray	White	White	White	White	Light Gray	Light Gray
2	Light Gray	Black	Light Gray	White	White	White	White	Light Gray	White
3	Light Gray	Light Gray	Black	White	Light Gray	Light Gray	Light Gray	White	Light Gray
4	White	Light Gray	Dark Gray	Black	Dark Gray	White	Light Gray	White	Dark Gray
5	White	Dark Gray	Dark Gray	Dark Gray	Black	White	Dark Gray	White	White
6	Light Gray	Light Gray	Light Gray	White	White	Black	White	Light Gray	White
7	White	White	Dark Gray	Light Gray	Dark Gray	White	Black	White	White
8	Dark Gray	Light Gray	White	White	White	Light Gray	White	Black	Light Gray
9	Dark Gray	White	Light Gray	Light Gray	White	White	White	Light Gray	Black

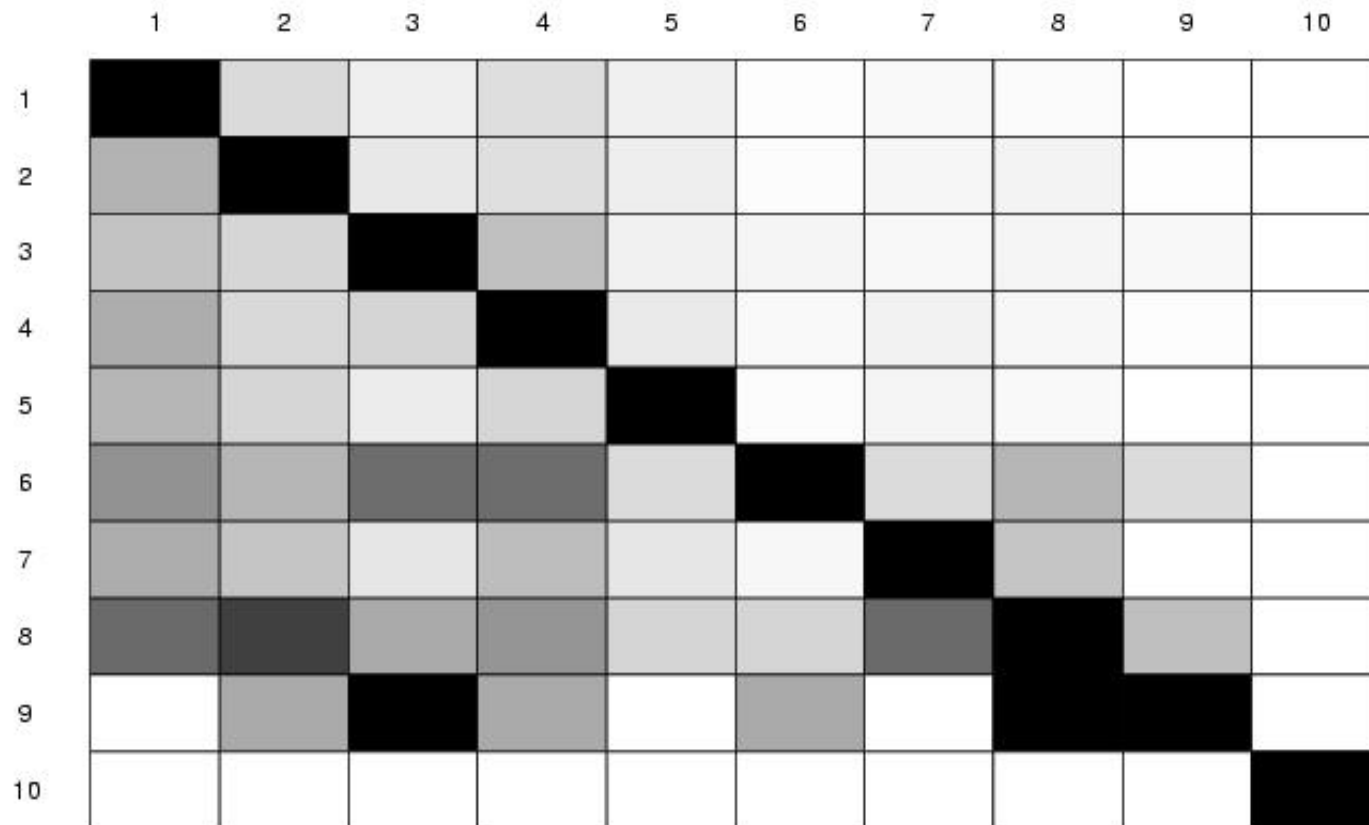
Detection over Time DEC

December Client Relative Comembership

	1	2	3	4	5	6	7	8
1	Black	Light Gray	White	Light Gray	White	Light Gray	White	White
2	Dark Gray	Black	White	Light Gray	White	Light Gray	White	White
3	Dark Gray	Light Gray	Black	Light Gray	White	Light Gray	White	Dark Gray
4	Dark Gray	White	White	Black	Light Gray	Light Gray	Light Gray	White
5	Dark Gray	White	White	Light Gray	Black	Dark Gray	Light Gray	Light Gray
6	Dark Gray	Light Gray	White	Light Gray	Light Gray	Black	White	White
7	Dark Gray	White	White	Dark Gray	Light Gray	Light Gray	Black	White
8	White	White	Black	White	Dark Gray	White	White	Black

Detection over Time JAN

January Client Relative Comembership



Detection over Time FEB

February Client Relative Comembership

	1	2	3	4	5	6	7	8	9	10	11
1	Black	Light Gray	Light Gray	Light Gray	White	Light Gray	Light Gray	Light Gray	White	White	White
2	Light Gray	Black	Light Gray	Light Gray	White	Light Gray	Light Gray	Light Gray	White	White	White
3	Light Gray	Light Gray	Black	Light Gray	White	Light Gray	Light Gray	Light Gray	White	White	White
4	Light Gray	Light Gray	Light Gray	Black	White	Light Gray	Light Gray	Light Gray	White	White	White
5	Light Gray	Light Gray	Light Gray	White	Black	Light Gray	Light Gray	Light Gray	Light Gray	White	White
6	Light Gray	Light Gray	Light Gray	Light Gray	White	Black	Light Gray	Light Gray	White	White	White
7	Light Gray	Light Gray	Light Gray	Light Gray	White	Light Gray	Black	Light Gray	Light Gray	White	White
8	Black	Light Gray	Light Gray	Light Gray	White	Light Gray	Light Gray	Black	Light Gray	White	White
9	White	Light Gray	Light Gray	White	Black	White	White	White	Black	White	White
10	White	White	White	White	White	White	White	White	White	Black	Black
11	White	White	White	White	White	White	White	White	White	Black	Black

Conclusions

- Data is not Random
- Anomaly Detection
 - Over Time

More Work

- Better represent anomaly detects
- Cluster and analyze server co-membership
- Build network infrastructure maps based on trust relationships
- Develop concept of power and commodities