Emergent Groups

MB 109

Agenda

- Definition
- A theory of subgroup formation
- Detecting subgroups analytically
  - methods
  - Data
    - Camp data
    - Zachary Karate data
    - Beach data

What's an emergent group?

- An emergent group emerges from pairwise interactions
  - Is a global structure that comes from local interactions
- As opposed to
  - Formal groups like teams, organizations
  - Classes, like "whites" or "people who live in the Mods"
- Emergent groups don't have clear boundaries, clear membership
  - At least at first

Evolution of Emergent Groups

- Pairwise interactions among a population
- Coalescing into subgroups
- Identification with group
- Development of group culture
  - Norms of behavior
  - Practices
  - Common symbols, references
  - World view

How does the coalescing happen?

- Similarity based transitivity
  - If friendship comes from similarity (homophily) then because similarity is transitive, so will friendship
- Time based transitivity
  - If A interacts a lot with B and also a lot with C, some of that time will have to be with B and C together, or else you run out of hours
- Reduction of cognitive dissonance
  - Suppose I like you and you like Bill. If I don't like Bill, something's got to give ..
- Cultural schemas / identification
  - Friend of a friend is supposed to be a friend; enemy of a friend is an enemy, etc.
- Categorizing processes / perception of groups

Detecting emergent groups

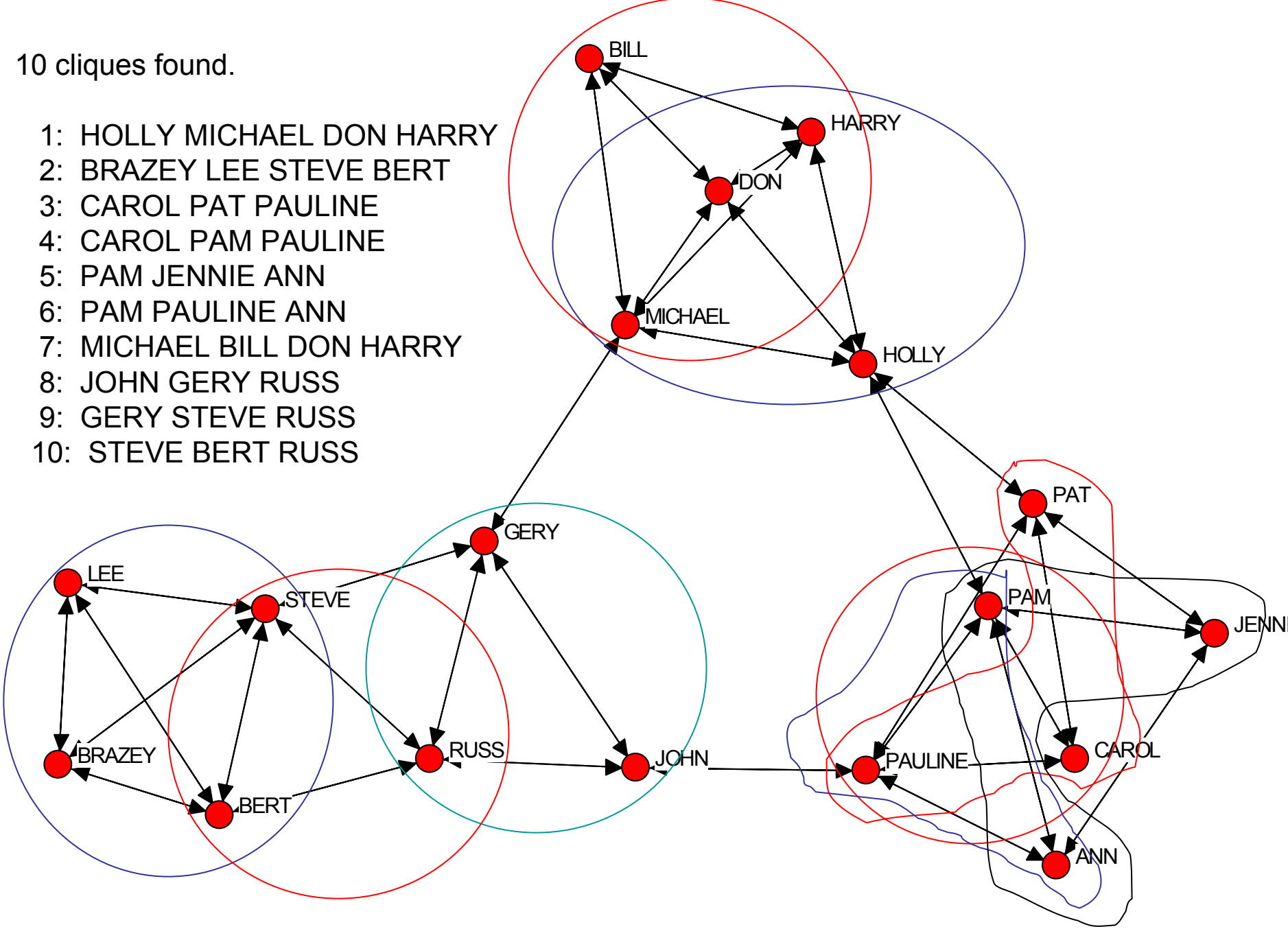
- Run clustering algorithms that search for tell-tale signs in the patterns of interaction
  - Luce & Perry Cliques
  - Factions
  - Hierarchical clustering

Luce & Perry Cliques

- Sets of nodes in which every member of the set has direct tie to every other member of the set
- Issues
  - Can be overlapping
  - What if there is one pair in the group that doesn't get along?

10 cliques found.

1: HOLLY MICHAEL DON HARRY
2: BRAZEY LEE STEVE BERT
3: CAROL PAT PAULINE
4: CAROL PAM PAULINE
5: PAM JENNIE ANN
6: PAM PAULINE ANN
7: MICHAEL BILL DON HARRY
8: JOHN GERY RUSS
9: GERY STEVE RUSS
10: STEVE BERT RUSS



Factions

- Mutually exclusive sets of nodes that are more connected to each other than to outsiders

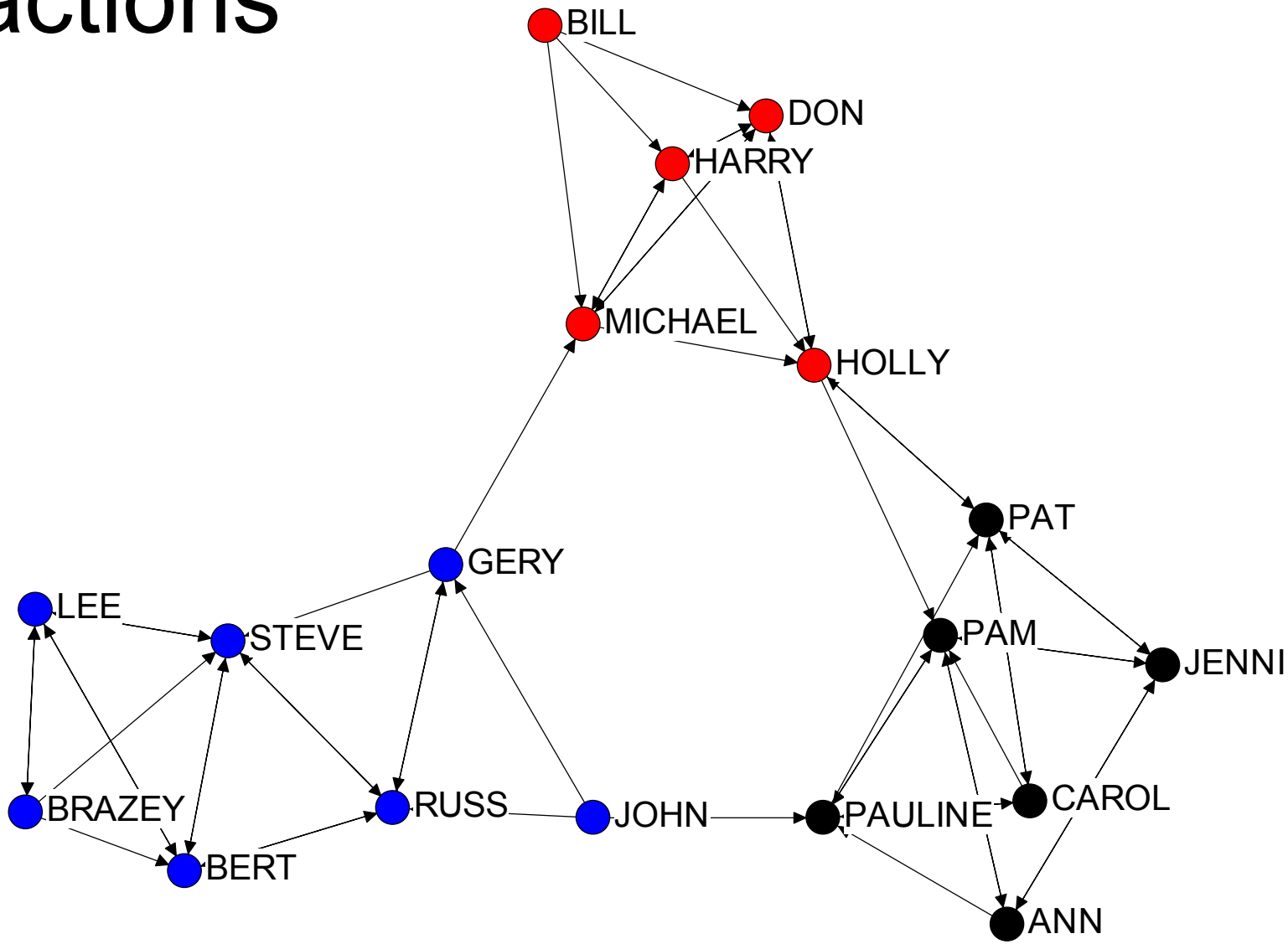
Idealized Factions

		5 P	6 J	3 C	4 P	7 P	8 A	1 8 R	2 B	1 3 J	1 6 S	1 1 L	1 7 B	1 5 G	1 0 B	1 4 H	1 2 D	1 H	9 M
5	PAT	-	1	1	1	1	1												
6	JENNIE	1	-	1	1	1	1												
3	CAROL	1	1	-	1	1	1												
4	PAM	1	1	1	-	1	1												
7	PAULINE	1	1	1	1	-	1												
8	ANN	1	1	1	1	1	-												
18	RUSS							-	1	1	1	1	1	1					
2	BRAZEY							1	-	1	1	1	1	1					
13	JOHN							1	1	-	1	1	1	1					
16	STEVE							1	1	1	-	1	1	1					
11	LEE							1	1	1	1	-	1	1					
17	BERT							1	1	1	1	1	-	1					
15	GERY							1	1	1	1	1	1	-					
10	BILL														-	1	1	1	1
14	HARRY														1	-	1	1	1
12	DON														1	1	-	1	1
1	HOLLY														1	1	1	-	1
9	MICHAEL														1	1	1	1	-

Factions

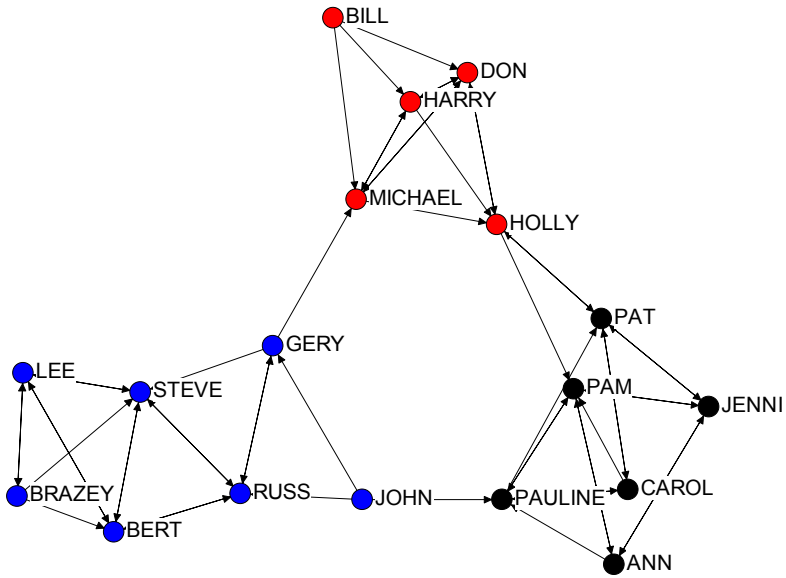
		5 6 3 4 7 8	1 1 1 1 1 1	1 1 1	
		5 6 3 4 7 8	8 2 3 6 1 7 5	0 4 2 1 9	
		P J C P P A	R B J S L B G	B H D H M	
5	PAT	- 1 1 1		1	
6	JENNIE	1 - 1 1			
3	CAROL	1 - 1 1			
4	PAM	1 1 - 1 1		1	
7	PAULINE	1 1 1 - 1	1		
8	ANN	1 1 1 -			
18	RUSS		- 1 1 1 1		
2	BRAZEY		- 1 1 1		
13	JOHN	1	1 - 1		
16	STEVE		1 1 - 1 1 1		
11	LEE		1 1 - 1		
17	BERT		1 1 1 1 -		
15	GERY		1 1 1 -	1	
10	BILL			- 1 1 1	
14	HARRY			1 - 1 1 1	
12	DON			1 1 - 1 1	
1	HOLLY	1 1		1 1 - 1	
9	MICHAEL		1	1 1 1 1 -	

Factions

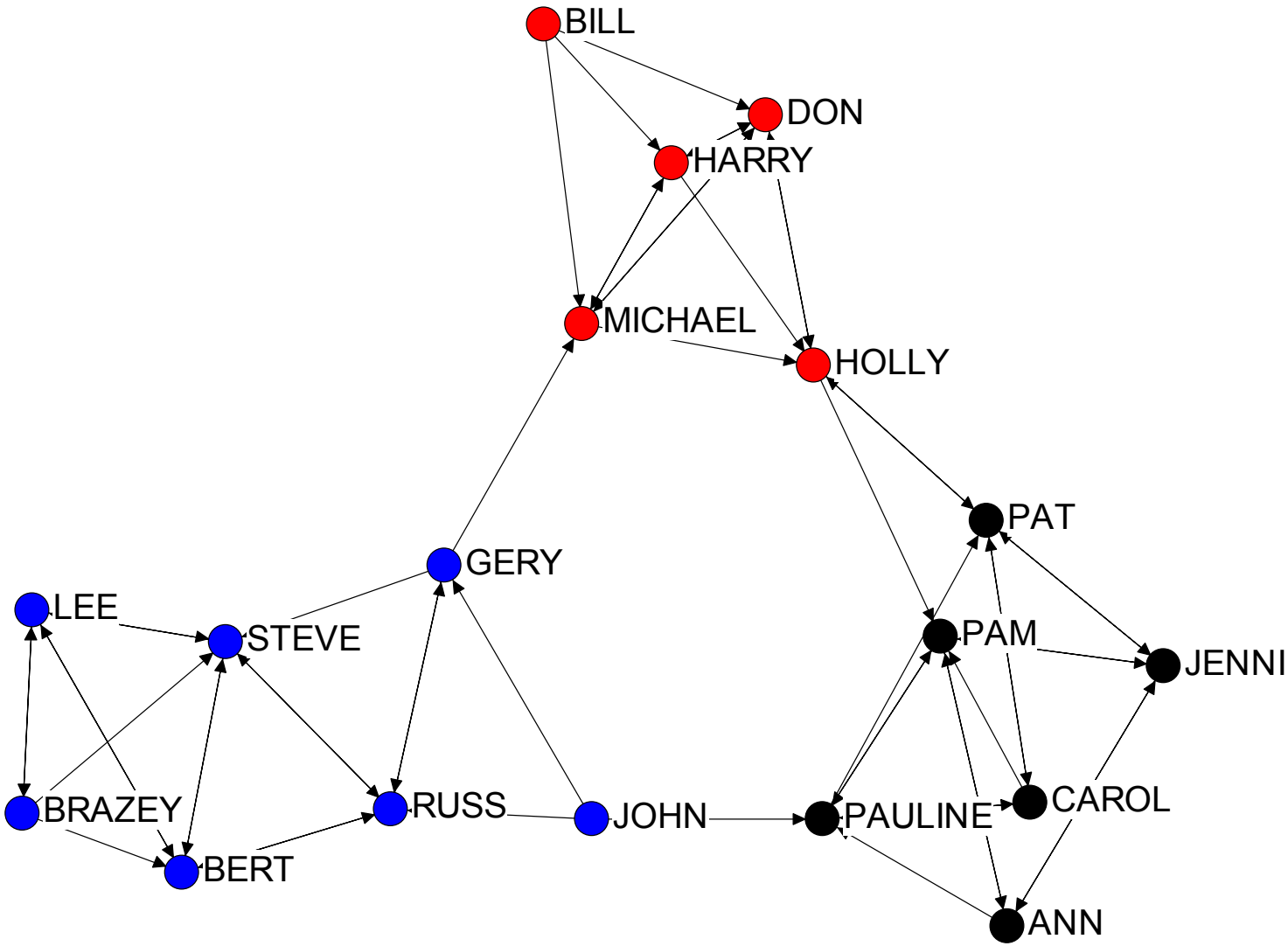


Newman-Girvan

- Find the most critical tie in a network, and delete it.
  - If the network breaks into pieces, great
  - If not, delete the 2nd most critical tie, and so on
- Recalculate critical ties, and repeat

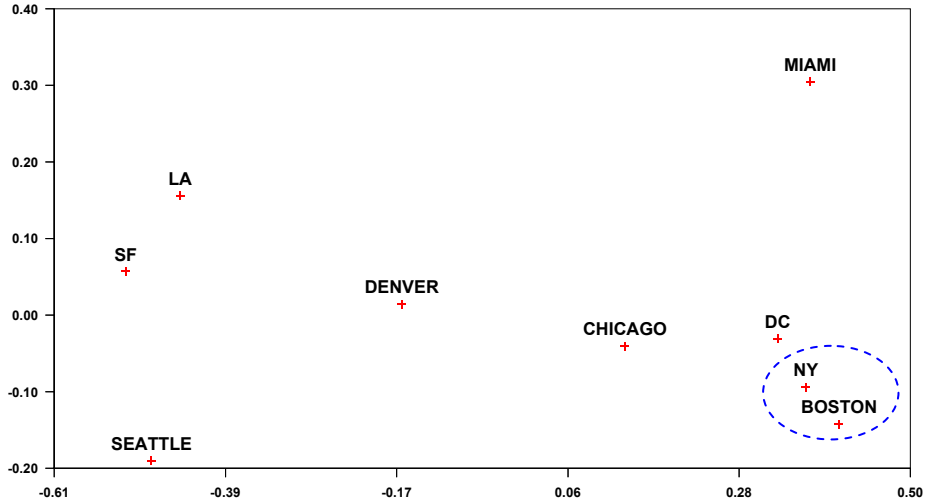


Newman-Girvan



Johnson's Hierarchical Clustering

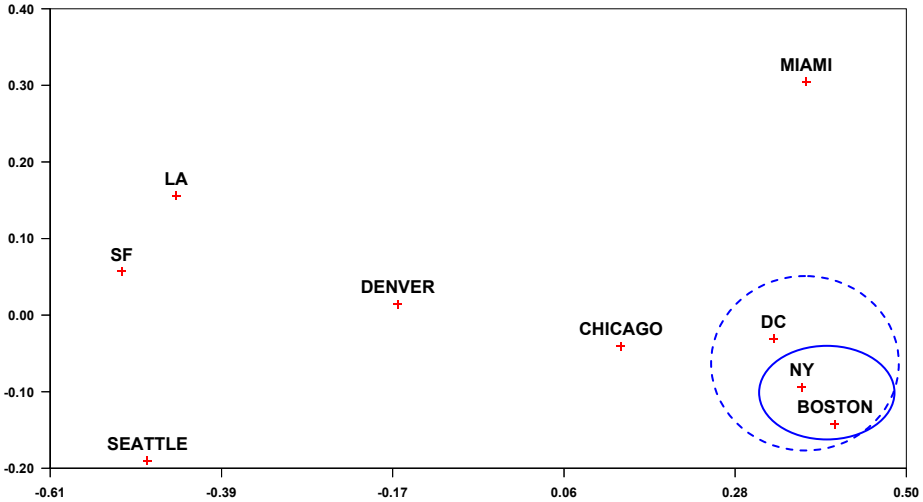
- Finds hierarchically nested groupings
- Divides nodes into mutually exclusive groups, but then joins groups together into larger, less cohesive groups




```
                    M  S        B        C  D
                    I  E  S  L  O  N  D  H  E
                    A  A  F  A  S  Y  C  I  N
Level               4  6  7  8  1  2  3  5  9
-----               -  -  -  -  -  -  -  -  -
  206               .  .  .  .  XXX   .  .  .
  233               .  .  .  .  XXXXX .  .  .
  379               .  .  XXX  XXXXX  .  .  .
  671               .  .  XXX  XXXXXXX  .  .
  000               .  XXXXXX  XXXXXXXX  .  .
  996               .  XXXXX  XXXXXXXXXX
 1059               .  XXXXXXXXXXXXXXXXX
 1075               XXXXXXXXXXXXXXXXXXX
```

	BOS	NY	DC	MIA	CHI	SEA	SF	LA	DEN
BOS	0	206	429	1504	963	2976	3095	2979	1949
NY	206	0	233	1308	802	2815	2934	2786	1771
DC	429	233	0	1075	671	2684	2799	2631	1616
MIA	1504	1308	1075	0	1329	3273	3053	2687	2037
CHI	963	802	671	1329	0	2013	2142	2054	996
SEA	2976	2815	2684	3273	2013	0	808	1131	1307
SF	3095	2934	2799	3053	2142	808	0	379	1235
LA	2979	2786	2631	2687	2054	1131	379	0	1059
DEN	1949	1771	1616	2037	996	1307	1235	1059	0

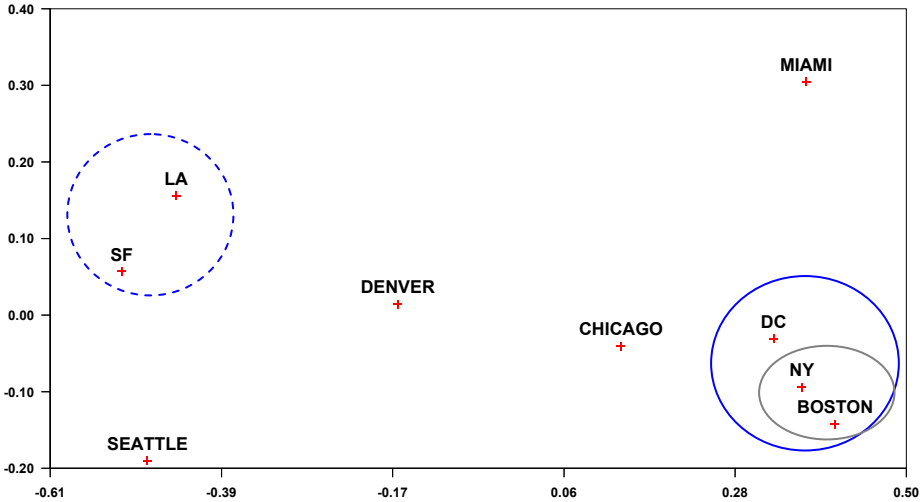
Closest distance is NY-BOS = 206, so merge these.




```
                  M  S        B        C  D
                  I  E  S  L  O  N  D  H  E
                  A  A  F  A  S  Y  C  I  N
Level             4  6  7  8  1  2  3  5  9
-----             -  -  -  -  -  -  -  -  -
  206             .  .  .  .  XXX   .  .  .
  233             .  .  .  .  XXXXX .  .  .
  379             .  .  XXX  XXXXX  .  .  .
  671             .  .  XXX  XXXXXXX   .  .
  000             .  XXXXXX  XXXXXXXXX   .
  996             .  XXXXX   XXXXXXXXXXX
 1059             .  XXXXXXXXXXXXXXXXXX
 1075             XXXXXXXXXXXXXXXXXXX
```

	BOS NY	DC	MIA	CHI	SEA	SF	LA	DEN
BOS/ NY	0	223	1308	802	2815	2934	2786	1771
DC	223	0	1075	671	2684	2799	2631	1616
MIA	1308	1075	0	1329	3273	3053	2687	2037
CHI	802	671	1329	0	2013	2142	2054	996
SEA	2815	2684	3273	2013	0	808	1131	1307
SF	2934	2799	3053	2142	808	0	379	1235
LA	2786	2631	2687	2054	1131	379	0	1059
DEN	1771	1616	2037	996	1307	1235	1059	0

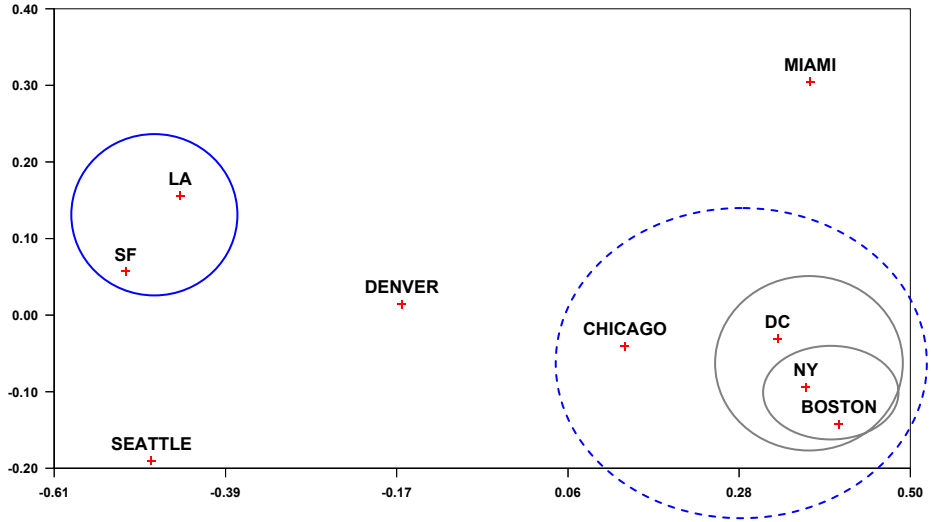
Closest pair is DC to BOSNY combo @ 223. So merge these.

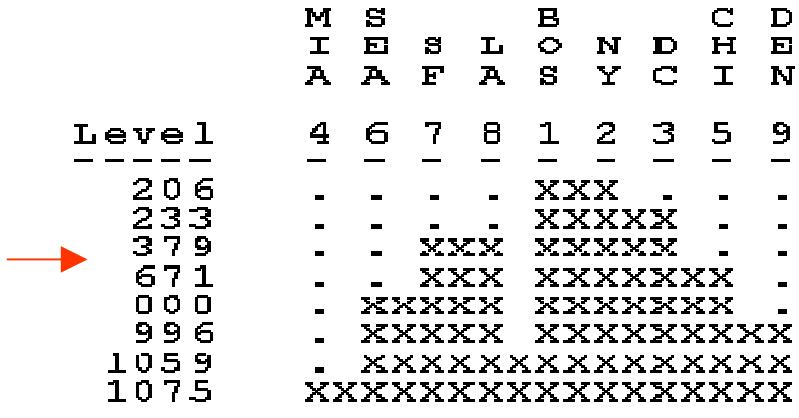



```
Level
-----
  206
  233
  379
  671
  000
  996
 1059
 1075
```

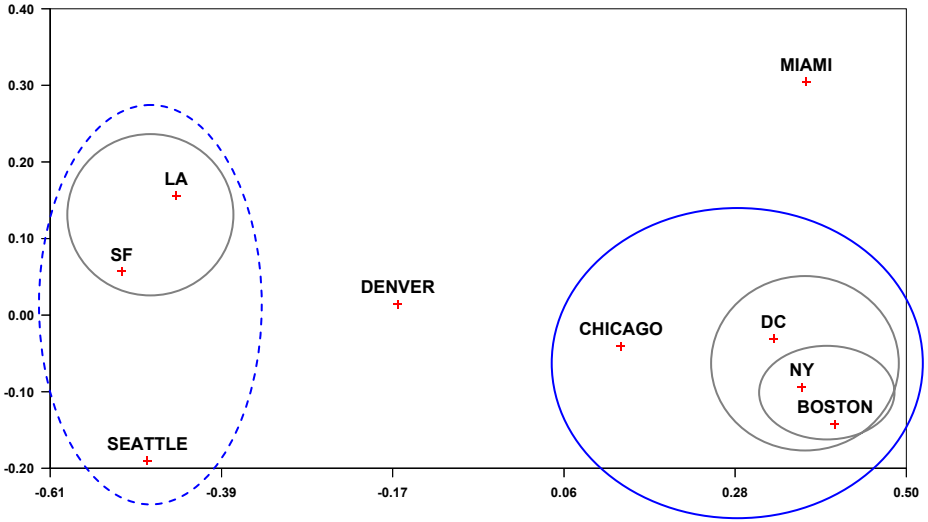
```
M  S        B        C  D
I  E  S  L  O  N  D  H  E
A  A  F  A  S  Y  C  I  N
-  -  -  -  -  -  -  -  -
4  6  7  8  1  2  3  5  9
-  -  -  -  -  -  -  -  -
.  .  .  .  XXX  .  .  .
.  .  .  .  XXXXX  .  .
.  .  XXX  XXXXX  .  .
.  .  XXX  XXXXXXX  .
.  XXXXXX  XXXXXXXXX  .
.  XXXXX  XXXXXXXXXXX
.  XXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXX
```

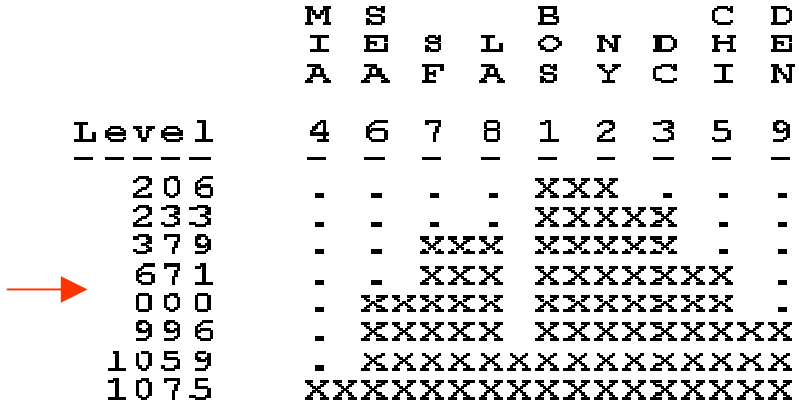
	BOS/ NY/ DC	MIA	CHI	SEA	SF	LA	DEN
BOS/NY DC	0	1075	671	2684	2799	2631	1616
MIA	1075	0	1329	3273	3053	2687	2037
CHI	671	1329	0	2013	2142	2054	996
SEA	2684	3273	2013	0	808	1131	1307
SF	2799	3053	2142	808	0	379	1235
LA	2631	2687	2054	1131	379	0	1059
DEN	1616	2037	996	1307	1235	1059	0



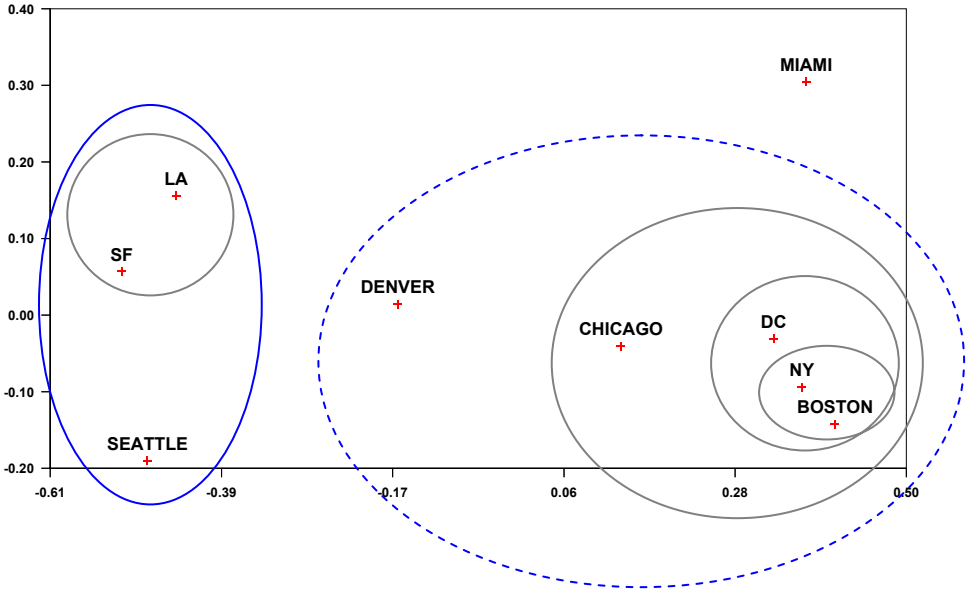


	BOS/ NY/DC	MIA	CHI	SEA	SF/LA	DEN
BOS/NY/DC	0	1075	671	2684	2631	1616
MIA	1075	0	1329	3273	2687	2037
CHI	671	1329	0	2013	2054	996
SEA	2684	3273	2013	0	808	1307
SF/LA	2631	2687	2054	808	0	1059
DEN	1616	2037	996	1307	1059	0



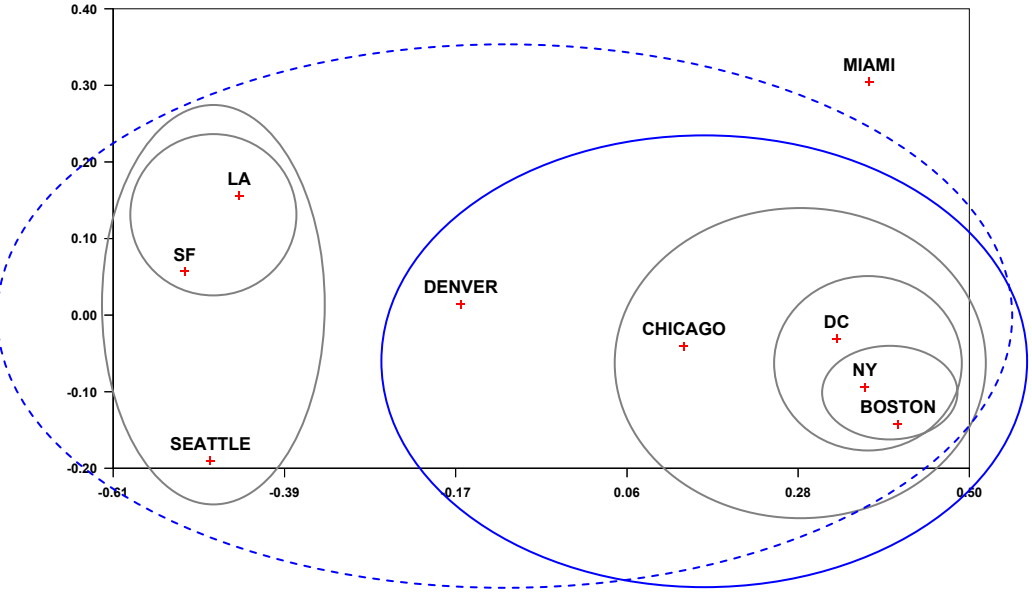


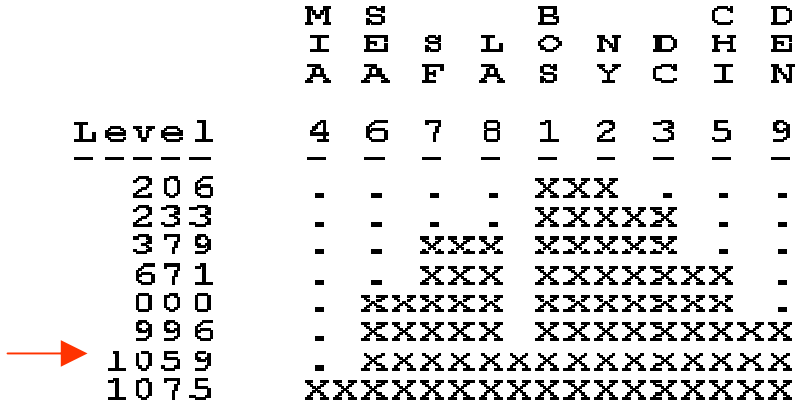
	BOS/ NY/D C/ CHI	MIA	SEA	SF/L A	DEN
BOS/NY/DC/C HI	0	1075	2013	2054	996
MIA	1075	0	3273	2687	2037
SEA	2013	3273	0	808	1307
SF/LA	2054	2687	808	0	1059
DEN	996	2037	1307	1059	0



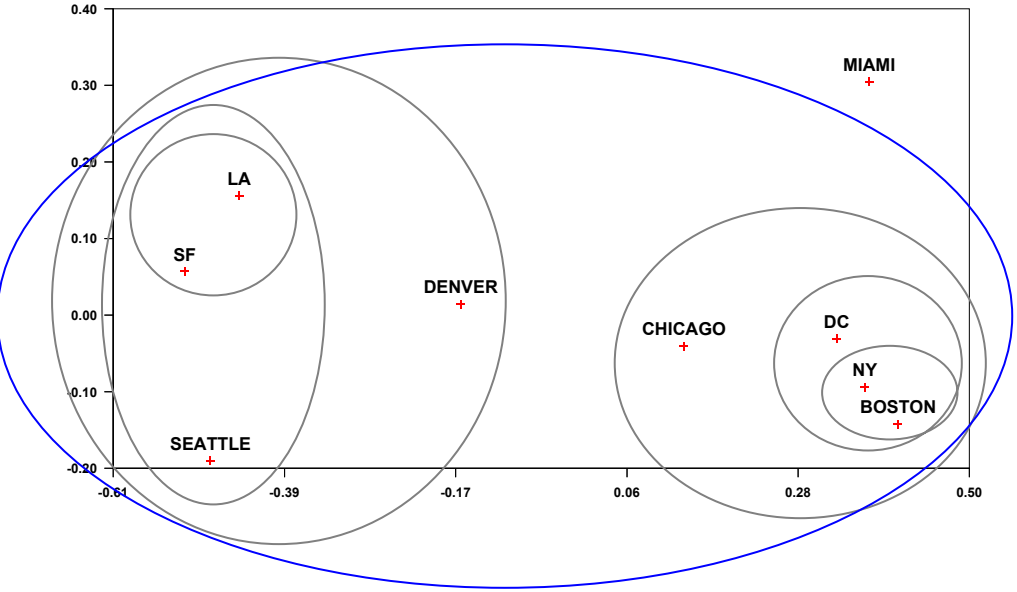

```
                M  S        B        C  D
                I  E  S  L  O  N  D  H  E
                A  A  F  A  S  Y  C  I  N
Level           4  6  7  8  1  2  3  5  9
-----           -  -  -  -  -  -  -  -  -
  206           .  .  .  .  XXX   .  .  .
  233           .  .  .  .  XXXXX .  .  .
  379           .  .  XXX  XXXXX  .  .  .
  671           .  .  XXX  XXXXXXX   .  .
  000           .  XXXXXX  XXXXXXXX  .  .
  996           .  XXXXX   XXXXXXXXXX
 1059           .  XXXXXXXXXXXXXXXXX
 1075           XXXXXXXXXXXXXXXXXXX
```

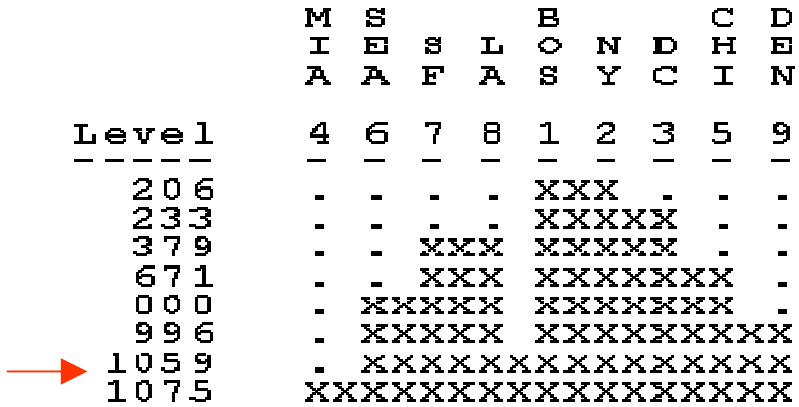
	BOS/ NY/D C/C HI	MIA	SF/L A/SE A	DEN
BOS/NY/DC/ CHI	0	1075	2013	996
MIA	1075	0	2687	2037
SF/LA/SEA	2054	2687	0	1059
DEN	996	2037	1059	0





	BOS/ NY/D C/CHI /DEN	MIA	SF/LA /SEA
BOS/NY/DC/ CHI/DEN	0	1075	1059
MIA	1075	0	2687
SF/LA/SEA	1059	2687	0





	BOS/NY/DC/CHI/DEN/SF/LA/SEA	MIA
BOS/NY/DC/CHI/DEN/SF/LA/SEA	0	1075
MIA	1075	0

Geodesic Distances

		1 H	2 B	3 C	4 P	5 P	6 J	7 P	8 A	9 M	10 B	11 L	12 D	13 J	14 H	15 G	16 S	17 B	18 R
1	HOLLY	0	4	2	1	1	2	2	2	1	2	4	1	3	1	2	3	4	3
2	BRAZEY	4	0	5	5	5	6	4	5	3	4	1	4	3	4	2	1	1	2
3	CAROL	2	5	0	1	1	2	1	2	3	4	5	3	2	3	3	4	4	3
4	PAM	1	5	1	0	2	1	1	1	2	3	5	2	2	2	3	4	4	3
5	PAT	1	5	1	2	0	1	1	2	2	3	5	2	2	2	3	4	4	3
6	JENNIE	2	6	2	1	1	0	2	1	3	4	6	3	3	3	4	5	5	4
7	PAULINE	2	4	1	1	1	2	0	1	3	4	4	3	1	3	2	3	3	2
8	ANN	2	5	2	1	2	1	1	0	3	4	5	3	2	3	3	4	4	3
9	MICHAEL	1	3	3	2	2	3	3	3	0	1	3	1	2	1	1	2	3	2
10	BILL	2	4	4	3	3	4	4	4	1	0	4	1	3	1	2	3	4	3
11	LEE	4	1	5	5	5	6	4	5	3	4	0	4	3	4	2	1	1	2
12	DON	1	4	3	2	2	3	3	3	1	1	4	0	3	1	2	3	4	3
13	JOHN	3	3	2	2	2	3	1	2	2	3	3	3	0	3	1	2	2	1
14	HARRY	1	4	3	2	2	3	3	3	1	1	4	1	3	0	2	3	4	3
15	GERY	2	2	3	3	3	4	2	3	1	2	2	2	1	2	0	1	2	1
16	STEVE	3	1	4	4	4	5	3	4	2	3	1	3	2	3	1	0	1	1
17	BERT	4	1	4	4	4	5	3	4	3	4	1	4	2	4	2	1	0	1
18	RUSS	3	2	3	3	3	4	2	3	2	3	2	3	1	3	1	1	1	0

Hierarchical Clustering

```
                P           M
                A     J     I           B
              C U   H E     C H         R S
              A L   O N   B H A     B A T G J R
            P R I P L N A I A R D L E Z E E O U
            A O N A L I N L E R O E R E V R H S
            T L E M Y E N L L Y N E T Y E Y N S

                              1   1 1 1 1   1 1 1 1
Level       5 3 7 4 1 6 8 0 9 4 2 1 7 2 6 5 3 8
-----       - - - - - - - - - - - - - - - - - - -
1.000       XXXXX XXX XXX XXXXXXX XXXXXXX XXXXX
1.333       XXXXX XXXXXXX XXXXXXX XXXXXXX XXXXX
1.457       XXXXX XXXXXXX XXXXXXX XXXXXXXXXXXXX
1.481       XXXXXXXXXXXXX XXXXXXX XXXXXXXXXXXXX
2.723       XXXXXXXXXXXXXXXXXXXXX XXXXXXXXXXXXX
3.142       XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
```