

Musical Group Membership Network Analysis

Social Network Analysis – University of Michigan/Coursera

Peer Assessment - Option 1

Author - bobisthedataguy

Introduction

Modern musical groups (bands) are composed of individual members often bringing unique talents. Those talents include songwriting skills, vocal talents, instrumental skills, and performance abilities. The unique combination of members in a musical group can result in the creation of entertainment artifacts (recording, performances, etc.) greater than that which could be produced by the individual members of the group.

One interesting social aspect of musical groups is the change that occurs in group composition. Band members leave one group only to join another existing group or to start a new group composed of former members of other existing groups. An individual may end up being connected to many other musicians as a result of their membership in different musical groups over time. The intent of this project is to study a musical group social network of individual in the network based upon the musical groups they have been a member of.

The premise of this study is to investigate the connection of musicians based upon their common membership in musical groups. It is understood that this is not the only measure of relationship between these individuals. In addition, there are additional variables that would indicate the strength of the connection between individuals. For example, just because individuals are in the same group, they might not have been in the group at the same time. This aspect has not been considered in this study, but is definitely something to consider in building the edges between nodes in any future studies. This has been ignored in this analysis as this detail was not available.

Data Sources

Obtained a tab separated file of musical group membership (group_membership.tsv) from freebase. The extract date of this file is May 6, 2013. The file contains 106,274 rows. This file contains linkages between a musical group and a group member. Below is the description of the file attributes:

Musical Group Membership Properties

Property	Description
Member	The musician who is part of the group
Group	The group (such as a band or orchestra) to which the member belongs
Roles	The primary instrument(s) or vocal role(s) which the member played in the group. Note that this should be the name of the instrument ("guitar") or voice/role ("tenor", "lead vocals"), rather than the related profession ("guitarist" or "lead vocalist").

Property	Description
Period (start)	When this member joined or founded this group
Period (end)	When this member left this group

Below is a sample of the dataset:

name	id	member	group	role	start	end
	/m/062vhlh	Hank Williams III	Antiseen			
	/m/0654bxy	John Cooper	Skillet	Lead vocalist,Acoustic guitar,Bass		1996
	/m/05kgt58	Christoph Kohli	Span	Bass guitar,Vocals		
	/m/05147hs	Rodrigo Aravena	Men at Work	Bass guitar,backing vocals		2000
	/m/05cm2fy	Vincent Kenis	The Honeymoon Killers			
	/m/05nqgx6	Derek Lee Rock	Suburban Legends			
	/m/05nqdz6	Emil Johnson	Black Flag			
	/m/05nnk5v	Jeff Plate	Savatage			
	/m/05nn7vv	Rowland Charles Gould	Level 42			

This file was loaded into a relational database table to facilitate ease of manipulation. Some minimal data cleansing was performed on this data.

Initial analysis attempts focused on analyzing the entire network of musicians in the source dataset. This proved to be overwhelming. As a result, it was determined that a subset of musical groups would be analyzed in order to facilitate meaningful analysis. It was decided that this subset would be composed of certain related musical genres that were prevalent in the 1970s and 1980s. The following genres/sub-genres were included due to the author's hypothesis that 'cross-pollination' of musicians occurred between these musical genres:

- Proto-punk
- Post-punk
- New Wave
- No Wave
- Power Pop
- New Romantics
- Synthpop
- Sophistipop

This is a very subjective selection, and it is clear that opinions may differ greatly regarding the relatedness of these categories.

Lists of musical groups were extracted from Wikipedia for each category indicated above. This data was cleansed prior to further processing.

All data was loaded into a sql database tables for ease of manipulation and processing. A comprehensive subset of groups was created using the following database view:

```
CREATE VIEW "selectgroupview" AS select grp from
newromantics
union
select grp from newwave
union
select grp from nowave
union
select grp from postpunk
union
select grp from powerpop
union
select grp from protopunk
union
select grp from sophistipop
union
select grp from synthpop;
```

The above view is used to filter out any group membership data for groups that do not fall within the above genre classifications.

Please note – “grp” is used for the musical group column names due to “group” being a SQL reserved word.

Methodology

The following extract files were generate using SQL queries:

File	Description
nodes.csv	Comma separate file containing all individual members of musical groups within the subset genres that have 1 or more connection with other individuals based upon shared musical group membership.
edges.csv	Comma separate file containing Source-Target relationship connections between individuals contained in the nodes.csv file.

Relationship edges are created based solely upon shared membership in the same group. In other words, If 2 individuals were in the same musical group, a relationship edge is established.

SQL used to create nodes.csv:

```
select distinct a.member as Id,  
              a.member as Label  
from  
  membergroup a,  
  membergroup b,  
  selectgroupview c  
where  
  a.grp = c.grp  
  and a.grp = b.grp  
  and a.member < b.member  
  
union  
  
select distinct b.member as Id,  
              b.member as Label,  
from  
  membergroup a,  
  membergroup b,  
  selectgroupview c  
where  
  a.grp = c.grp  
  and a.grp = b.grp  
  and a.member < b.member;
```

SQL used to create edges.csv:

```
select a.member as Source,  
       b.member as Target,  
       a.grp as Label  
from  
  membergroup a,  
  membergroup b,  
  selectgroupview c  
where  
  a.grp = c.grp  
  and a.grp = b.grp  
  and a.member < b.member;
```

A sample of the nodes data:

```
"Id";"Label"  
"A.C. Newman";"A.C. Newman"  
"Aaron A Train Smith";"Aaron A Train Smith"  
"Aaron Escolopio";"Aaron Escolopio"  
"Aaron Flora";"Aaron Flora"  
"Aaron Rossi";"Aaron Rossi"  
"Aaron Smith";"Aaron Smith"  
"Abbo";"Abbo"  
"Adam Clayton";"Adam Clayton"  
"Adam Cox";"Adam Cox"
```

And as sample of the edges data:

```
"Source";"Target";"Label"  
"A.C. Newman";"Blaine Thurier";"The New Pornographers"  
"A.C. Newman";"Daniel Bejar";"The New Pornographers"  
"A.C. Newman";"John Collins";"The New Pornographers"  
"A.C. Newman";"John Collins";"The New Pornographers"  
"A.C. Newman";"Kathryn Calder";"The New Pornographers"  
"A.C. Newman";"Kurt Dahle";"The New Pornographers"  
"A.C. Newman";"Neko Case";"The New Pornographers"  
"A.C. Newman";"Todd Fancey";"The New Pornographers"  
"Aaron A Train Smith";"Aaron Smith";"Romeo Void"
```

These input files were imported into Gephi for network analysis. The edges were imported as unweighted.

Results

The resulting graph had 2653 nodes with 12115 undirected edges between these nodes. The average degree for the graph is 4.567, indicating that this is the average number of connections between members within the graph. The overall network diameter is 17, which represents the maximum distance between any pair of individual represented in the graph. There are 308 connected components in this large graph. The average cluster coefficient is 0.957 and the average path length is 6885. The graph density is small at 0.003.

The giant component is composed of 38.6% of the nodes (1024 nodes) and 53.9% of the edges (6539 edges). The average degree in the giant component is higher than the larger graph at 6.386. The resulting graph density is still small in the giant component at 0.012. The average clustering coefficient is 0.913 and the average path length is 6954 in the giant component.

The highest degree individuals in the graph are in the table that follows. This would possibly indicate the possibility of “cross-pollination” with many other individual musicians. The high degree means they have been in the musical groups with a large number of other individual musicians. Possibly either

they were in a large number of bands, or more likely, they were in bands that experienced high turnover in the individual members.

Id	Degree
Micky Billingham	48
Mark Brzezicki	47
Jim Sclavunos	43
Craig Adams	43
Johnny Marr	43
Jay dee Daugherty	42
Barry Adamson	40
Andy Anderson	39
Kid Congo Powers	39
Jools Holland	38
Paul Mulreany	38
Simon Wolstencroft	38
Spencer Birtwistle	38
Steve Hanley	38
Steven Trafford	38
Tom Head	38
Tommy Crooks	38
Tony Friel	38
Una Baines	38
Yvonne Pawlett	38
Adam Helal	38
Adrian Flanagan	38
Ben Pritchard	38
Brian Fanning	38
Brix Smith	38
Craig Scanlon	38
Dave Bush	38
Dave Milner	38
David Spurr	38
Ed Blaney	38
Elenor Smith	38
Jim Watts	38
Jonnie Brown	38
Julia Nagle	38
Alex Lee	38
Joe Allen	38
Pat Fish	38
Karen Leatham	38
Karl Burns	38
Kate Themen	38
Kay Carroll	38
Keiron Melling	38
Kenny Brady	38

Id	Degree
Marc Riley	38
Marcia Schofield	38
Mark E. Smith	38
Martin Bramah	38
Mike Leigh	38
Neville Wilding	38
Paul Hanley	38
Pete Greenway	38
Simon Archer	38
Simon Rogers	38

Below is a graph filtered for these top degree individuals:

Member	Group	Degree
Mark Brzezicki	Big Country	47
Mark Brzezicki	Casbah Club	47
Mark Brzezicki	Procol Harum	47
Mark Brzezicki	The Cult	47
Mark Brzezicki	Ultravox	47
Jim Slavunos	8-Eyed Spy	43
Jim Slavunos	Grinderman	43
Jim Slavunos	Nick Cave and the Bad Seeds	43
Jim Slavunos	Sonic Youth	43
Jim Slavunos	Teenage Jesus and The Jerks	43
Jim Slavunos	The Cramps	43
Jim Slavunos	The Vanity Set	43
Craig Adams	Spear of Destiny	43
Craig Adams	The Cult	43
Craig Adams	The Mission	43
Craig Adams	The Sisters of Mercy	43
Johnny Marr	7 Worlds Collide	43
Johnny Marr	Electronic	43
Johnny Marr	Johnny Marr + the Healers	43
Johnny Marr	Johnny Marr and Billy Duffy	43
Johnny Marr	Modest Mouse	43
Johnny Marr	The Cribs	43
Johnny Marr	The Pretenders	43
Johnny Marr	The Smiths	43
Johnny Marr	The The	43
Jay dee Daugherty	Patti Smith Group	42
Jay dee Daugherty	The Beat	42
Jay dee Daugherty	The Church	42
Jay dee Daugherty	The Waterboys	42
	Barry Adamson, Anita Lane and the Thought System of	
Barry Adamson	Love	40
Barry Adamson	Buzzcocks	40
Barry Adamson	Magazine	40
Barry Adamson	Nick Cave and the Bad Seeds	40
Barry Adamson	The Birthday Party	40
Barry Adamson	Visage	40
Andy Anderson	A & M	39
Andy Anderson	Hawkwind	39
Andy Anderson	The Cure	39
Andy Anderson	The Glove	39
Kid Congo Powers	Congo Norvell	39
Kid Congo Powers	Knoxville Girls	39
Kid Congo Powers	Nick Cave and the Bad Seeds	39
Kid Congo Powers	The Cramps	39
Kid Congo Powers	The Gun Club	39
Jools Holland	Billy Preston & Jools Holland	38

Member	Group	Degree
Jools Holland	Jools Holland & His Rhythm & Blues Orchestra	38
Jools Holland	Jools Holland and Jamiroquai	38
Jools Holland	Squeeze	38
Jools Holland	The The	38
Jools Holland	Tom Jones & Jools Holland	38
Paul Mulreany	The Blue Aeroplanes	38
Paul Mulreany	The Jazz Butcher	38
Simon Wolstencroft	The Fall	38
Simon Wolstencroft	The Weeds	38
Spencer Birtwistle	The Fall	38
Steve Hanley	The Fall	38
Steve Hanley	Tom Hingley and the Lovers	38
Steven Trafford	The Fall	38
Tom Head	The Fall	38
Tommy Crooks	The Fall	38
Tony Friel	The Fall	38
Tony Friel	The Passage	38
Tony Friel	The Teardrops	38
Una Baines	The Fall	38
Yvonne Pawlett	The Fall	38
Adam Helal	The Fall	38
Adrian Flanagan	The Fall	38
Ben Pritchard	The Fall	38
Brian Fanning	The Fall	38
Brix Smith	The Adult Net	38
Brix Smith	The Fall	38
Craig Scanlon	The Adult Net	38
Craig Scanlon	The Fall	38
Dave Bush	Elastica	38
Dave Bush	The Fall	38
Dave Milner	The Fall	38
David Spurr	The Fall	38
Ed Blaney	The Fall	38
Elenor Smith	The Fall	38
Jim Watts	The Fall	38
Jonnie Brown	The Fall	38
Julia Nagle	The Fall	38
Alex Lee	Placebo	38
Alex Lee	Strangelove	38
Alex Lee	Suede	38
Alex Lee	The Blue Aeroplanes	38
Alex Lee	The Jazz Butcher	38
Alex Lee	Warm Jets	38
Joe Allen	Strangelove	38

Member	Group	Degree
Joe Allen	The Blue Aeroplanes	38
Joe Allen	The Jazz Butcher	38
Pat Fish	J.B.C.	38
Pat Fish	The Blue Aeroplanes	38
Pat Fish	The Jazz Butcher	38
Karen Leatham	The Fall	38
Karl Burns	Public Image Ltd.	38
Karl Burns	The Adult Net	38
Karl Burns	The Fall	38
Karl Burns	The Teardrops	38
Kate Themen	The Fall	38
Kay Carroll	The Fall	38
Keiron Melling	The Fall	38
Kenny Brady	The Fall	38
Marc Riley	Marc Riley with The Creepers	38
Marc Riley	Shirehorses	38
Marc Riley	The Fall	38
Marcia Schofield	The Adult Net	38
Marcia Schofield	The Fall	38
Mark E. Smith	The Fall	38
Mark E. Smith	Von Südenfed	38
Martin Bramah	The Fall	38
Martin Bramah	The Teardrops	38
Mike Leigh	The Fall	38
Neville Wilding	The Fall	38
Paul Hanley	The Fall	38
Paul Hanley	Tom Hingley and the Lovers	38
Pete Greenway	The Fall	38
Simon Archer	AAAK	38
Simon Archer	The Fall	38
Simon Rogers	E-Zee Possee	38
Simon Rogers	Head Honcho	38
Simon Rogers	Incantation	38
Simon Rogers	Ramp	38
Simon Rogers	Slacker	38
Simon Rogers	The Adult Net	38
Simon Rogers	The Fall	38

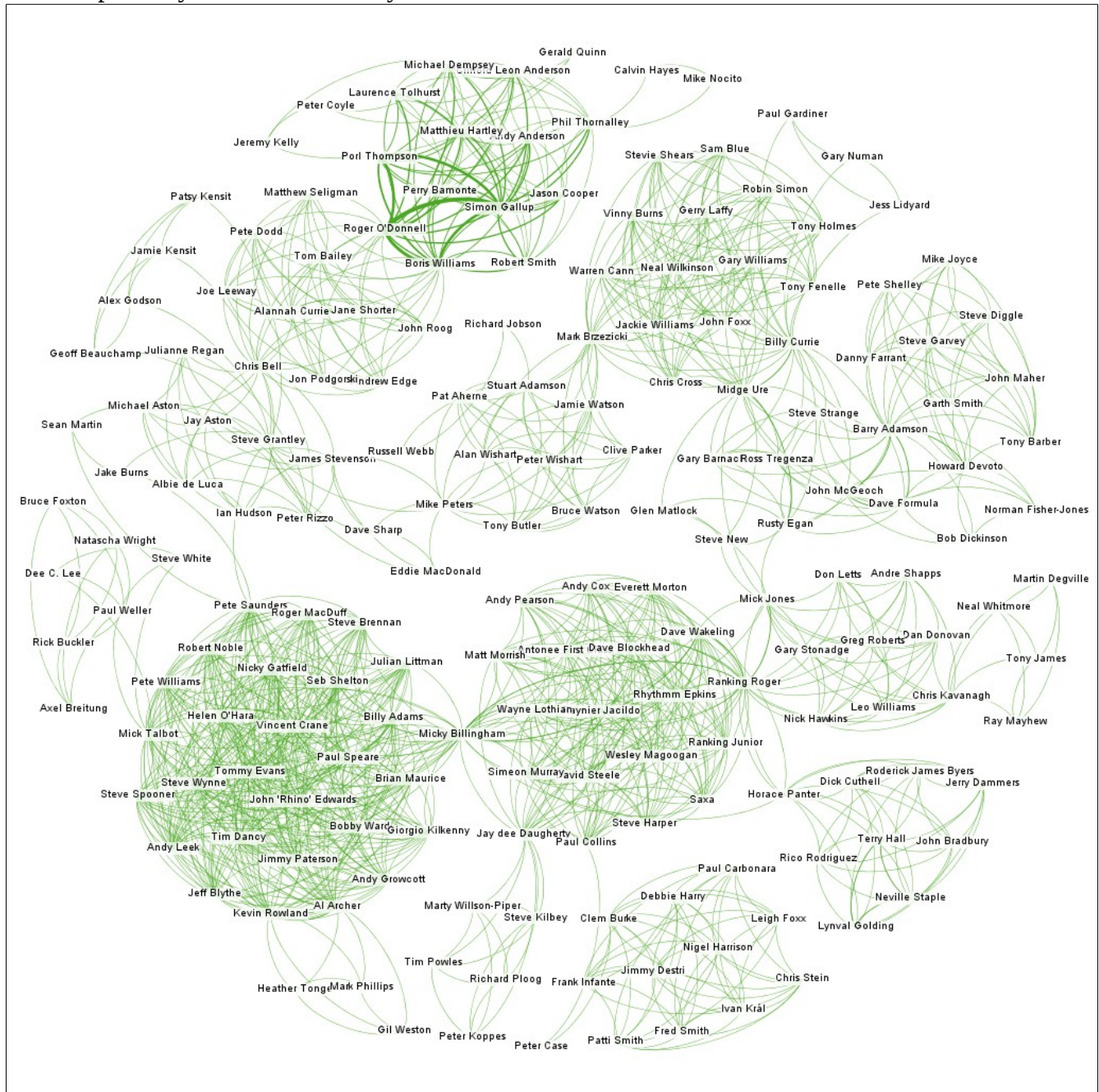
One interesting thing that jumps out of this data is that the musical group “The Fall” appears to have had quite a few individuals in the top degree list as members.

Future Analysis and Data Considerations

The dataset used can benefit from further enhancement. First, there will be benefit from determining a way to weight the edges. Possibly this can be done by obtaining group membership dates. In addition,

membership dates may be useful in filtering out any memberships that did not overlap between individuals. Additionally, shared membership is not the only determining factor in relationships between musicians. Various other factors (e.g. named influences) might be useful in future analysis.

Additionally, further detailed analysis is probably beneficial and in order. Sunsetting the data further into individual genres will probably lead to more meaningful analysis. For example, below is a graph of the memberships within the “new wave” genre. This produces a graph that is easier to deal with and this will probably lead to richer analytical results.



Data Source References

Freebase, Musical Group Membership, http://www.freebase.com/music/group_membership,
December 12, 2006

Wikipedia, Listings of Musical Groups, <https://www.wikipedia.org/>, Various