Exploring the Network Dynamics of Hip-Hop Music Collaboration Among Artists Associated with Oakland, California

Gregory M. Tozzi
gtozzi@gmu.edu

INTRODUCTION

Hip-hop culture is typically thought of as being represented in four urban art forms: emceeing or rapping, turntablism or deejaying, B-boying or breakdancing, and graffiti artistry. Two cultural elements – *knowledge* and *overstanding* – are sometimes also added to these performance elements to create a more complete definition of hip-hop culture. As can be expected of a movement in which there is not an accepted taxonomy, additional elements are sometimes added to these six [Alim 2002]. While important from the standpoint of understanding the broader context of hip-hop as a social movement, the cultural elements and two of the performance elements – B-boying and graffiti – have largely been eclipsed by the musical elements, and, in particular, by rap [Alim 2002].

Emerging in New York's South Bronx in the 1970s as a purposeful counterpoint to rampant gang violence [Chang and Watkins 2007], the musical elements of hip hop have expanded well beyond their urban African American origins to capture the interest and imagination of suburban white America [Kitwana 2006] and indeed, a good part of the wider world [Androutsopoulos and Scholz 2003].

Through its evolution, hip-hop's musical disciplines, those of the MC and the DJ, demonstrate a degree of collaboration – including musical collaboration extending outside of the hip-hop community [Light 2004] – substantially greater than that seen in other musical genres [Smith 2006]. Indeed, as Smith noted, collaboration among hip-hop artists frequently extends beyond the usual boundaries of record labels and geographic areas. This study explores techniques to build temporal graphs resulting from hip-hop collaboration from data available from crowd-sourced internet sources, to explore the structure and dynamics of the resulting temporal graph, and to test popular notions of the evolution of hip-hop and a novel technique for assigning communities among the graph's nodes using collaborative dynamics.

Smith showed that the hip-hop network was vast in 2006, and it is reasonable to assume that it has grown considerably since that time. To keep this study tractable but contextually meaningful, the study focuses on the network of hip-hop artists bearing a tie, as identified by the crowd, to Oakland, California.

DATA ACQUISITION AND GRAPH CREATION

Data sources and acquisition

This study relied on the crowd-sourced MusicBrainz database for information related to artists, their recordings, and their collaborations. MusicBrainz's database is extremely large and extends well beyond hip-hop [MetaBrainz 2017]. To define the network more narrowly, the Wikipedia List of hip-hop musicians page [Wikipedia 2017] was employed as an authoritative list of hip-hop artists. Bespoke R functions were developed to simplify access to the MusicBrainz API. MusicBrainz assigns each name under which an artist records a unique hexadecimal identifier.

Graph construction

As stated above, initial explorations were limited to a graph constructed from artists hailing from or being closely associated with Oakland, California as identified by contributors to the MusicBrainz database. While not as central to hip-hop as Los Angeles and New York, Oakland – and the San Francisco Bay Area more generally – plays an important role in the evolving the culture [Amoeba 2013]. Emphasizing the importance of the Bay Area, rapper E-40 stated that the, "rap game without the Bay Area is like old folks without bingo," [Jones 2006].

Among the metadata provided by MusicBrainz is an artist's city of origin or activity. MusicBrainz indicated that 119 hip-hop artists had a defensible connection to Oakland, and these artists were used as the vertices in the Oakland graph. Of the original 119, 89 were retained as having formed intra-Oakland connections during their careers. Temporal connection data for each artist were scraped from MusicBrainz after a careful review of the HTML structure used to report the site's entries.

Three issues rose to the fore during the scraping and graph creation process. The first was the impact of reliance on crowd-sourced data to determine an artists' geographic affiliation. The Oakland hip-hop collaboration graph does not include the famous Bay Area rapper E-40 because his geographic association in MusicBrainz is with Vallejo. Likewise, rap legend 2Pac Shakur's period in Oakland as a member of the rap collective Digital Underground do not qualify him as an Oakland rapper per the MusicBrainz data [MetaBrainz 2017]. This issue and those that follow highlight the difficulty in deciding the domain knowledge that a research should apply when curating a data set. In this case, no domain knowledge was applied to adjust the crowd's reported sense of geographic affiliation.

The second issue was that of posthumous collaboration. Mortality among hip-hop artists is disturbingly high, and several nodes in the Oakland hip-hop collaboration graph represent deceased artists [Fernandez 2015]. Often, deceased hip-hop artists continue to be assigned collaborations as their previous works are sampled and additional recordings are released [Geslani 2017]. There may be some utility associated with collaborations of this sort, so nodes were left active after artists' deaths.

The third issue was name disambiguation. Hip-hop artists frequently employ pseudonyms, and MusicBrainz issues a unique identifier to every name an artist uses without directly performing disambiguation. This task can be performed using additional data available from MusicBrainz; this task was automated in the current study with purpose-built R code.

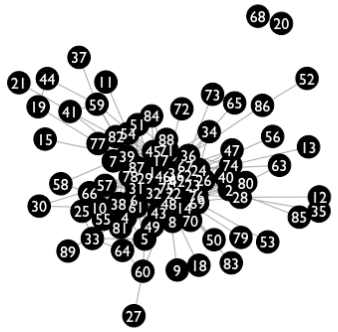

FIGURE 1 - THE OAKLAND HIP-HOP COLLABORATION GRAPH AS OF 2016

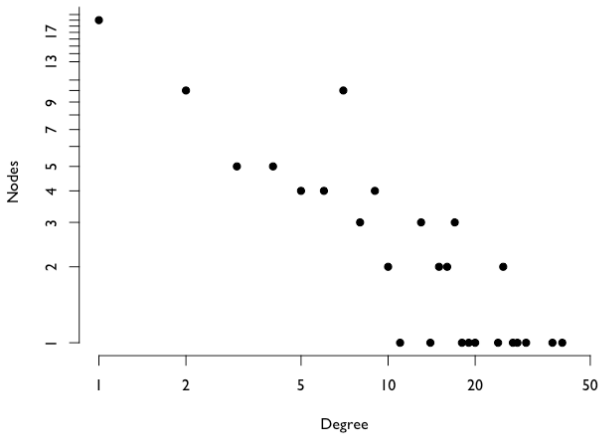

FIGURE 2 - DEGREE DISTRIBUTION IN THE OAKLAND HIP-HOP COLLABORATION GRAPH AS OF 2016

The graph was finally constructed in yearly temporal steps – the smallest time step common to all of the data scraped from MusicBrainz – between the first recorded intra-Oakland hip-hop

collaboration in 1992 and the most recent complete year, 2016. The resulting graph's structure in 2016 is shown as Figure 1.

Smith noted that the degree distribution in the comprehensive graph of hip-hop collaboration he produced indicated the possibility of a power law relationship [Smith 2006]. Plotting the distribution of degree for the Oakland hip-hop collaboration graph in 2016 yields Figure 2 and suggests a possible power law relationship. Care should be exercised before accepting this conclusion, however, because of the relatively small number of nodes in the Oakland graph. Preferential attachment is a plausible mechanism for the formation of the Oakland hip-hop collaboration graph. One would expect that new rappers would want to collaborate with established artists in order to establish their careers and benefit from network effects. If this is the case, a power law distribution of degree is likely [Newman 2001].

DYNAMIC INFLECTION AND THE HYPHY MOVEMENT

Santoro et al. [2011] suggest that inflection points in certain measures compared over time indicates that the mechanisms through which the graph is being created have changed. Dynamic analysis of this sort may have forensic benefit in that it highlights for researchers certain time periods for further investigation. To that end, the evolution of density and clustering in the Oakland hip-hop collaboration graph were computed.

Density and clustering dynamics

The Oakland hip-hop collaboration graph described above undergoes a pronounced reduction in density from 1992 to 2003 as shown in Figure 3. This effect was noted in a study of the social network resulting from scientific publishing [Santoro et al. 2011] and is the result of vertex initiation outpacing the development of new edges. Beginning in 2004, the graph's density recovers somewhat indicating that edge creation is proceeding at a rate quickly enough to overcome the rate that new nodes are being added to the graph.

The evolution of the clustering coefficient of the Oakland graph is presented in Figure 4. The graph suggests a steady reduction in the density of sub-communities in the network through roughly 2003. From 2004 forward a characteristic of the graph is the development of somewhat denser sub-communities. As with the reported density dynamics, this regime shift in clustering indicates that the mechanisms of graph creation have changed during the period of inflection.

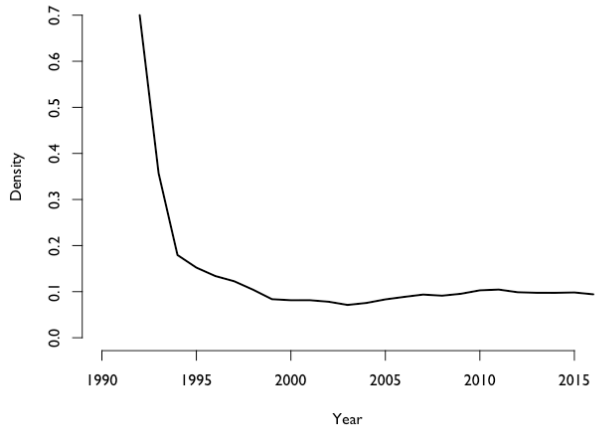

FIGURE 3 - DENSITY DYNAMICS

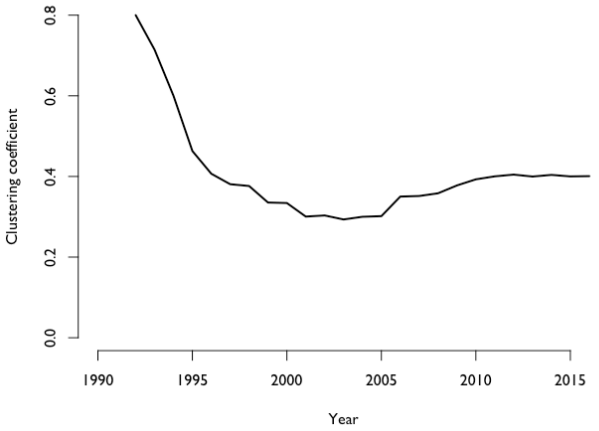

FIGURE 4 - CLUSTERING DYNAMICS

Explaining the inflection – the Hyphy movement

The regime shift evident in Figures 3 and 4 was caused by a surge of artists entering into new collaborations that began in 2004. Figure 5 shows the trend new intra-Oakland collaborations. Examining the data reveals that rapper Mac Dre,

founder of Thizz Entertainment, and the Thizz-associated group the Delinquents were responsible for the bulk of the new intra-Oakland hip-hop collaborations in 2004. Mac Dre is commonly viewed as a leading figure in the development of the Bay Area's unique *hyphy* style of hip-hop [Nguyen 2014]. In 2005, following Mac Dre's murder, the artists entering into the greatest number of intra-Oakland hip-hop collaborations were Mistah F.A.B., Bavgate, and Keak da Sneak. Mistah F.A.B. and Keak da Sneak are both strongly associated with the hyphy style [Horowitz 2016].

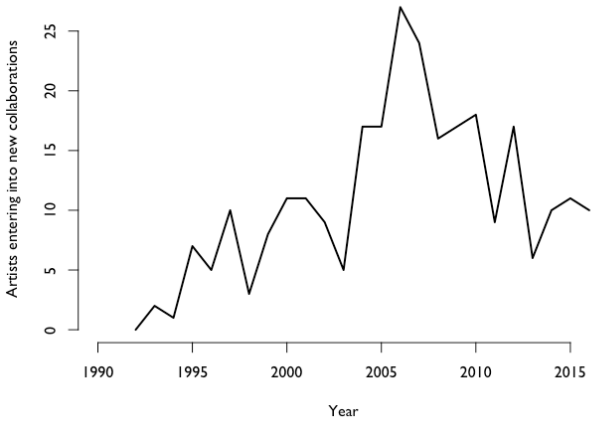

FIGURE 5 - EVOLUTION OF INTRA-OAKLAND HIP-HOP COLLABORATION

The hyphy movement was characterized by up-tempo, bass heavy hip-hop music extolling the virtues of hyphy-specific cultural elements like impromptu automotive-themed street parties referred to as *sideshows* at which cars would occasionally be allowed to run driverless in an act called *ghostriding the whip*, the consumption of the drug ecstasy or *thizz*, and fashion elements like large oversized sunglasses referred to as *stunna shades*. Hyphy music was borne of collaborations between Bay Area affiliated artists and those with ties to the South – or, commonly, the *Dirty South* – including Lil Jon, a pioneer of the South's *crunk* hip-hop style [Ibid].

Is there a link between the hyphy hip-hop subculture and the change in the dynamics governing the evolution of the Oakland hip-hop collaboration graph? Downloading the relatively frequency of searches for the term "hyphy" from Google Trends from the start of Google's trend tracking in 2004 through 2016, generating annual averages for Google's monthly data, and then regressing the number of artists entering into new intra-Oakland collaborations against those averages yields the statistically significant relationship $(R^2 = 0.591, p = 7.94\times10^{-4})$ presented in Figure 6 but also shows signs of a non-normal residual distribution. Testing further using Kendall's rank correlation indicates a statistically significant correlation $(p = 3.40\times10^{-2})$.

The data strongly suggest, but do not prove, a link between the evident changes to the mechanisms underpinning the development of the Oakland hip-hop collaboration graph and the emergence of the hyphy style of hip-hop music in roughly 2004. If a link existed, a plausible narrative could likely be built around the change in Oakland's cachet resulting from the arrival of a popular hip-hop genre with a strong regional affiliation. Capitalizing on the new style would provide strong incentives for increased intra-Oakland (and intra-Bay Area) collaboration among artists.

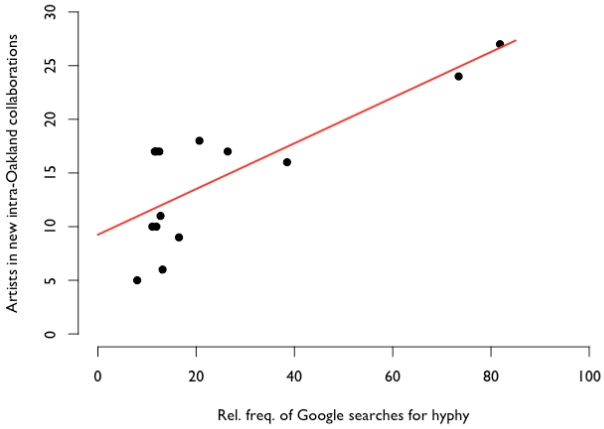

FIGURE 6 - CORRELATION OF GOOGLE SEARCHES FOR "HYPHY" AND ARTISTS ENTERING INTO NEW INTRA-OAKLAND HIP-HOP COLLABORATIONS

BROKERAGE DYNAMICS AND INCENTIVES

Burt's arguments suggest a system wherein structural holes provide a potential for value capture that is then realized by closing those holes through brokerage [Burt 2009]. Hypothesizing an

ambitious actor who understands these mechanisms, that actor should divide effort between seeking out extra-clique connections that form structural holes and brokering connections between potential relationships. Temporal graphs like the Oakland hip-hop collaboration graph provide opportunities to find evidence of this behavior.

Identifying potential brokerage

Testing notions around brokerage requires a means of assigning potential brokerage actions to actors. A potential action, rather than a true action, is the best that one can hope to identify without using structured interviews or some other technique outside the immediate scope of social network analysis.

Define a *potential brokerage event* as occurring if two nodes in an actor's first-degree ego network at time period *t* that do not share an edge develop an edge in time period $t+1$. Consider the case of rapper San Quinn in 2004. San Quinn's first-degree ego network in 2004 included connections to rappers Mac Dre, Keak da Sneak, and Bavgate. Neither Mac Dre nor Keak da Sneak were connected to Bavgate in 2004. In 2005, Bavgate formed connections with both Mac Dre and Keak da Sneak. It is impossible to state conclusively that these connections formed because of actions taken by San Quinn, but the possibility certainly exists. Stated another way, the evolution of connections described above are necessary, but insufficient, to prove that San Quinn brokered the connections described. For purposes of simply counting potential brokerage events assignable to node *N* that are realized in period $t+1$, it is sufficient to begin with *N*'s ego network in period *t*, and a network constructed from the vertices of *N*'s ego network in period *t* and the edges of *N*'s ego network in period $t+1$. Call the first network N_t with number of edges E_t and the second network $N_{t,t+1}$ with edge count $E_{t,t+1}$. The number of potential brokerage events occurring in period $t+1$ is then

$$1. \quad PBE_{t+1} = E_{t,t+1} - E_t$$

This way of defining potential brokerage events is not robust to events that occur within the temporal graph's time scale. For instance, the dates of collaboration events contained in the Oakland hip-hop collaboration graph are accurate only to the year. A hypothetical artist entering into initial individual collaborations with two other artists in February 2006 who subsequently brokers a collaboration between those two artists in May 2006 is not awarded credit for a potential brokerage event because the proposed method disregards new nodes. The method is robust, however, to incorrect assignment of potential brokerage events due to a node entering into initial collaborations in a given year with two artists who are already connected through collaboration.

Results

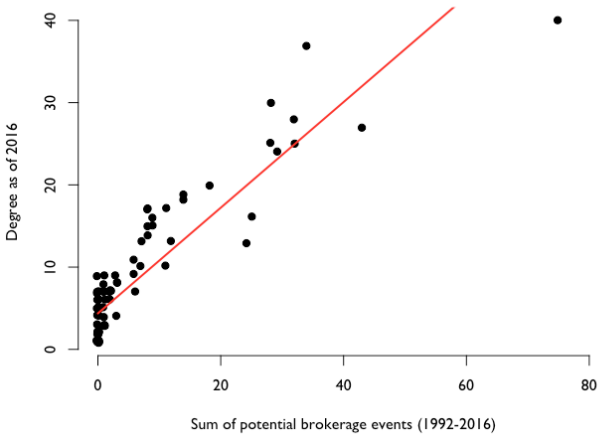

FIGURE 7 - INDIVIDUAL DEGREE AS A FUNCTION OF BROKERAGE

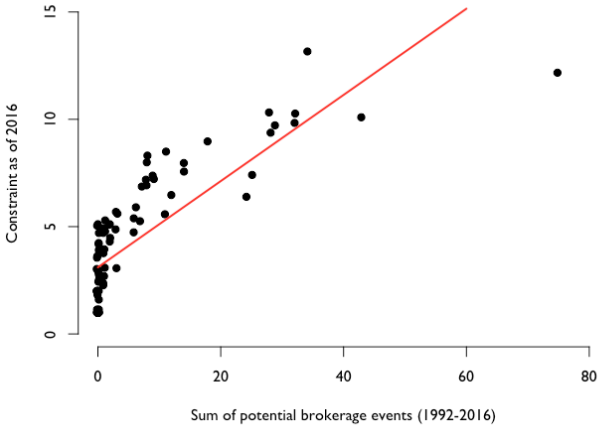

FIGURE 8 - INDIVIDUAL CONSTRAINT AS A FUNCTION OF BROKERAGE

Applying Equation 1 to all nodes in the Oakland hip-hop collaboration graph from 1992 to 2016 produces a sum of potential brokerage events for each artist represented in the graph. Figure 7 shows the degree of each node in the graph in 2016 plotted against the artists' sums of career potential brokerage events. The linear model indicated on the figure suggests a strong correlation ($R^2 = 0.813, p < 2.2\times10^{-16}$) between an artist's degree in 2016 and the career sum of the artist's potential brokerage.

Plotting the inverse of each artist's Burt's constraint in 2016 against the career sum of potential brokerage events produces FIGURE **8**. The linear model indicated on the figure suggests a correlation ($R^2 = 0.639, p < 2.2\times10^{-16}$) between an artist's degree in 2016 and the career sum of the artist's potential brokerage.

Results in context

A regular theme in hip-hop lyrics is the notion of an artist aggressively overcoming barriers to financial success against a backdrop of profoundly challenging economic and social hardships. Frequently, success is attributed to *hustling*, a nuanced term that encompasses selling and deal making in the informal and illicit economies. [Charnas 2011]. In *Hope I Don't Go Back*, prominent Bay Area rapper E-40 reflects on the travails of his past life in Vallejo, California selling *yayo*, a slang term for cocaine, to make *mail*, a slang term for money. The song's chorus, "I hope I don't go back to slangin' yayo, slangin' yayo, to get my mail," is simultaneously a reflection on drug dealing as a stepping stone and a commentary on the perceived tenuousness of success in hip-hop [E-40 1998].

The hustle dynamic in hip-hop may give rise to the brokerage relationships described in the preceding section. Qualitative evidence suggests that both highly connected artists and up-and-coming artists seek one another out based on the perception that a relationship will yield career advancement. Rapper Lil Dicky sought out a collaboration with legendary artist Snoop Dogg and funded it with a campaign on the crowd-funding platform Kickstarter. The collaboration resulted in the recording and video *Professional Rapper* likely launched Lil Dicky to the number 2 position on the Billboard Top R&B/Hip-Hop Albums chart [Robehmed 2016]. Rapper and producer Eminem saw potential in newcomer 50 Cent, eventually signing the latter to the Shady Aftermath record label and accelerating a career that saw 50 Cent reaching a net worth in excess of $150 million [Markman 2014].

This context provides support for a narrative of the dynamics generating the results presented in FIGURE **7** and FIGURE **8**. Hip-hop's biggest names and new talent eagerly seek one another out to shore up and advance their successes. Those artists who demonstrate the willingness and ability to broker subsequent relationships become preferred partners, sought out by others in the space and able to approach new talent successfully. Because this dynamic is perceived to be important, successful brokers seek out connections that create structural holes that can eventually be filled with a portion of the resulting value accruing to the broker. This narrative can be intuited from the dynamics of the Oakland hip-hop collaboration graph. It should be viewed merely as plausible, however, until tested using the techniques of classical social science.

DYNAMIC COMMUNITY DETECTION

Methodology

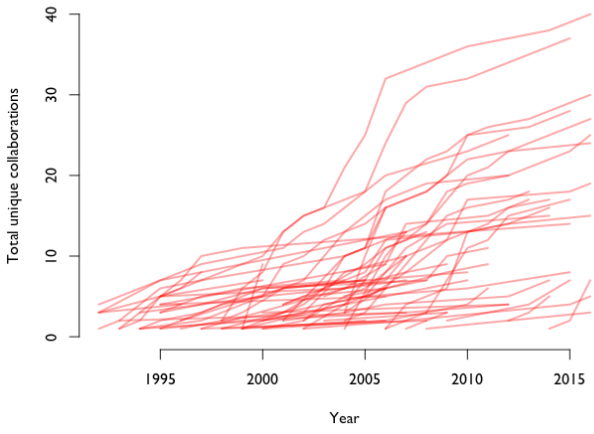

FIGURE 9 - TRACES OF CONNECTION ACQUISITION FOR THE 89 MEMBERS OF THE OAKLAND HIP-HOP COLLABORATION GRAPH

Conventional community detection methods rely on the static structure of networks to identify

groups of significance [Fortunato 2010]. Efforts to glean community affiliations from temporal graphs have been made relatively recently. More recent attempts have been made to determine community affiliation based on adjacency tensor factorization. The adjacency tensor this method relies on is essentially a temporal stack of static adjacency matrices. [Gauvin et al 2014].

Rather than relying on a direct understanding of the temporal progression of connections in a network, perhaps community affiliations of practical interest can be detected simply from patterns of edge acquisition in a manner agnostic to the identity of the nodes with which the acquired edges form connections. Put another way, it may be possible to group nodes into communities based on how those nodes behave with respect to making connections.

The new connection acquisition behavior of the 89 artists in the Oakland hip-hop collaboration graph is plotted as Figure 9. The time series plotted show the cumulative number of novel collaborations than an artist has entered into. From these traces, it is easy to form a time series for each artist that gives the count of novel collaborations by year. Clustering of time series can be accomplished through hierarchical clustering, a method that has been applied to the task of optimizing portfolios of assets by identifying and avoiding correlations [Tola et al. 2008]. Borrowing from this technique, clustering of time series of novel connection acquisition can be accomplished by:

1. Computing a time series consisting of the number of novel connections an actor forms in a given time period.
2. Computing the variance-covariance matrix of the resulting time series.
3. Applying a hierarchical clustering algorithm to the difference of the resulting matrix. This can be accomplished, for example, through application of the `hclust` function in R. This function implements a number of algorithms. An algorithm implementing Ward's criterion was selected because it was found to produce the best modularities in this instance [Murtagh et al. 2014]. This selection should not be considered a general best practice, however, as the choice of clustering algorithm may be specific to the domain.
4. Determining a horizontal cut point – or *tree height* – for the resulting hierarchical tree. The cut point in this study was determined through computation of modularity of the resulting community assignments for a range of potential tree heights.

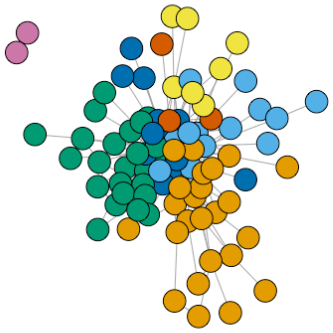

FIGURE 10 - COMMUNITIES RETURNED USING CONVENTIONAL STATIC CLUSTERING

Applying a traditional static community detection algorithm to the Oakland hip-hop collaboration graph in 2016 provides a baseline against which to compare the proposed method. Accomplishing this task using the default settings in iGraph yielded the community structure indicated in Figure 10 [Csardi G, Nepusz 2006]. The spatial relevance of the community apportionment is obvious from inspection, but the structure of that apportionment is made a bit more clear by examining the grouped time series of novel connection acquisition presented in Figure 11. The orange, green, and dark blue communities are dominated by one primary node (although an argument could be made in support of the idea that there are two dominant nodes in the case of the orange group). The dominant artists in these three groups are Too $hort, Keak da Sneak, and

San Quinn, all artists noted for their intra-Oakland collaboration [Paine 2013].

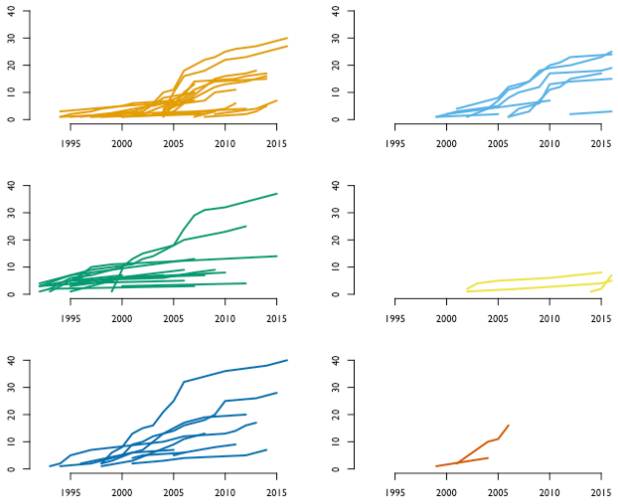

FIGURE 11 - CLUSTERED TRACES OF NOVEL CONNECTION ACQUISITION UNDER STATIC COMMUNITY DETECTION (GROUP 7 NOT SHOWN)

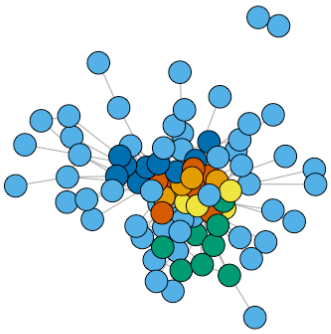

FIGURE 12 - CLUSTERING UNDER THE DYNAMIC COMMUNITY DETECTION ALGORITHM

Applying the dynamic community detection algorithm produces the tree with optimal cut point shown in Figure 12. The resulting communities, shown in Figure 13 are spatially relevant for those nodes that are the most central members of the graph. Nodes on the periphery are grouped into a single community based on their relatively low rate of connection acquisition; these are the least active intra-Oakland collaborators.

The analogous figure to Figure 11 is presented as Figure 14. Here we see that the algorithm has grouped the time series in a seemingly sensible manner. More interesting, the memberships of the individual groups have contextual value that is immediately apparent to one with domain knowledge. The yellow group represents iconic members of the Oakland hip-hop community who achieved immediate success in the early years of hip hop. The orange group includes those high-profile artists who were most strongly associated with Oakland's transition from the *Mobb Music* style in the 1990s to the hyphy style in the mid-2000s. The dark blue community are the hyphy early adopters. The red community are primarily late adopters of hyphy. The green traces are deceased and defunct artists. While the relevance of these groupings is immediately apparent with the benefit of domain knowledge, there are as yet no measures that marry contextual value quantitatively with community assignment. This is, therefore, clearly a work in progress.

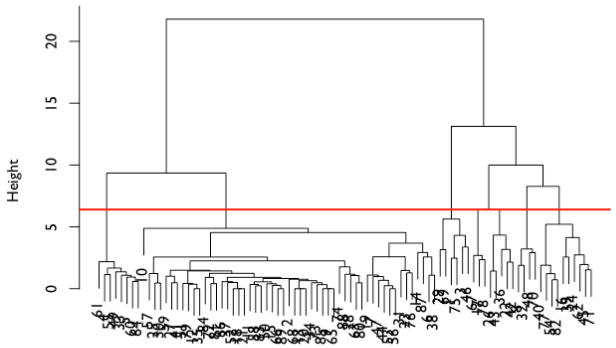

FIGURE 13 - TREE SHOWING CUT POINT IN RED

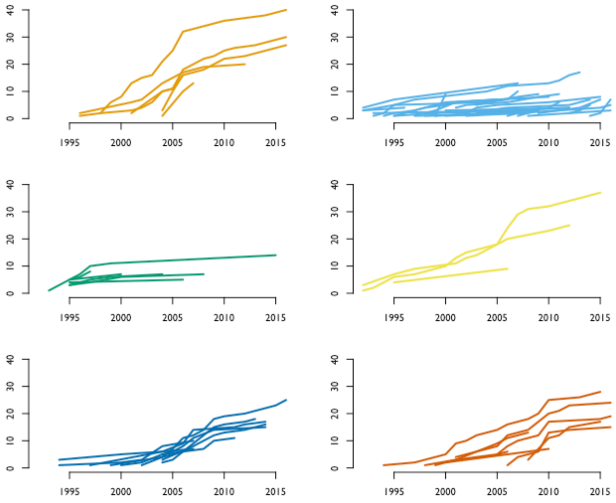

FIGURE 14 - TRACES OF CONNECTION ACQUISITION UNDER DYNAMIC COMMUNITY DETECTION

REFERENCES

Alim, H. Samy. “Street-conscious copula variation in the hip hop nation.” *American Speech* 77, no. 3 (2002): 288-304.

Amoeba Records. “Hip-Hop History Tesdays: The Roots of Oakland Rap and the Birth of Bay Area Hip-Hop in the 1980s.” *amoeba.com*. Web. Accessed April 2, 2017.

Androutsopoulos, Jannis and Arno Scholz. “Spaghetti funk: appropriations of hip-hop culture and rap music in Europe.” *Popular Music and Society* 26, no. 4 (2003): 463-479.

Burt, Ronald S. *Structural holes: The Social Structure of Competition*. Cambridge: Harvard University Press, 2009. 18-29.

Chang, Jeff, and S. Craig Watkins. “It’s a Hip-hop World.” *Foreign Policy* 163 (2007): 58.

Charnas, Dan. *The Big Payback: The History of the Business of Hip-Hop*. Penguin, 2011. Chap 1.

Csardi G, Nepusz T: The igraph software package for complex network research, InterJournal, Complex Systems 1695. 2006. http://igraph.org.

E-40. “Hope I don’t go back.” The Element of Surprise. Sick Wid It Records. November 12, 1998, Digital.

Fernandez, Lisa. “Rapper ‘The Jacka’ Killed While Making Music: Friend.

Fortunato, Santo. "Community detection in graphs." *Physics Reports* 486, no. 3 (2010): 75-174.

Gauvin, Laetitia, André Panisson, and Ciro Cattuto. "Detecting the community structure and activity patterns of temporal networks: a non-negative tensor factorization approach." *PloS One* 9, no. 1 (2014): e86028.

Geslani, Michelle. “Faith Evans details posthumous Notorious B.I.G. duets album, shares two new songs – listen.” COS. February 3, 2017. Web.

Horowitz, Steven J. “An Oral History of Hyphy.” *Complex Magaizne*. June 13, 2016.

Jones, Steve. “Hyphy pulls a Bay Area Breakout.” *USA Today*. April 13, 2006.

Kitwana, Bakari. *Why White Kids Love Hip Hop: Wankstas, Wiggers, Wannabes, and the New Reality of Race in America*. Basic Civitas Books, 2006.

Light, Alan. “About a Salary or Reality?” *That’s the Joint: A Hip Hop Studies Reader.* Ed. Murray Forman and Mark Anthony Neal. New York: Routledge (2004): 137-145.

Markman, Rob, “Eminem and 50 Cent Remember When They First Met in This Video.” mtv.com. MTV. December 23, 2014. Web.

Murtagh, Fionn, and Pierre Legendre. "Ward’s hierarchical agglomerative clustering method: which algorithms implement ward’s criterion?" *Journal of Classification* 31, no. 3 (2014): 274-295.

MetaBrainz Foundation. *MusicBrainz*. musicbrainz.org. Web. Accessed March 13, 2017.

Newman, Mark E.J. "Clustering and preferential attachment in growing networks." *Physical Review* E 64, no. 2 (2001): 025102.

Nguyen, Dean Van, “Vallejo rapper Mac Dre pioneered the hyphy movement.” *WaxPoetics*. June 2, 2014.

Paine, Jake. “San Quinn speaks openly about he and Richie Rich jumping Tupac in 1992.” hiphopdx.com. Web. August 22, 2013.

Robehmed, Natalie, “Professional Rapper? How Lil Dicky Mixed Comedy and Hip-Hop to Top the Charts.” *Forbes*. February 9, 2016.

Santoro, Nicola et al. “Time-varying graphs and social network analysis: Temporal indicators and metrics.” *3rd AISB Social Networks and Multiagent Systems Symposium* (May 2011): 32-38

Smith, Reginald D. “The network of collaboration among rappers and its community structure.” *Journal of Statistical Mechanics: Theory and Experimentation* 2006, no. 02 (2006): P02006.

Tola, Vincenzo, Fabrizio Lillo, Mauro Gallegati, and Rosario N. Mantegna. "Cluster analysis for portfolio optimization." *Journal of Economic Dynamics and Control* 32, no. 1 (2008): 235-258.

“List of Hip Hop Musicians.” *Wikipedia.* Wikimedia Foundation. Web. Accessed March 10, 2017.