

Mining the performance history of the New York Philharmonic, 1842–2015

Programming trends and performer networks

Eamonn Bell

Department of Music
Columbia University
www.columbia.edu/~epb2125
<epb2125@columbia.edu>

Introduction

The recent release into the public domain of historical concert program data in XML format by the New York Philharmonic affords the opportunity of diachronic analysis of musical practice in the city, as reflected in the selected pieces performed and records of the participating performers themselves. By borrowing techniques from the domains of information extraction and time-series analysis, we can explore trends in concert programming over the history of the institution.

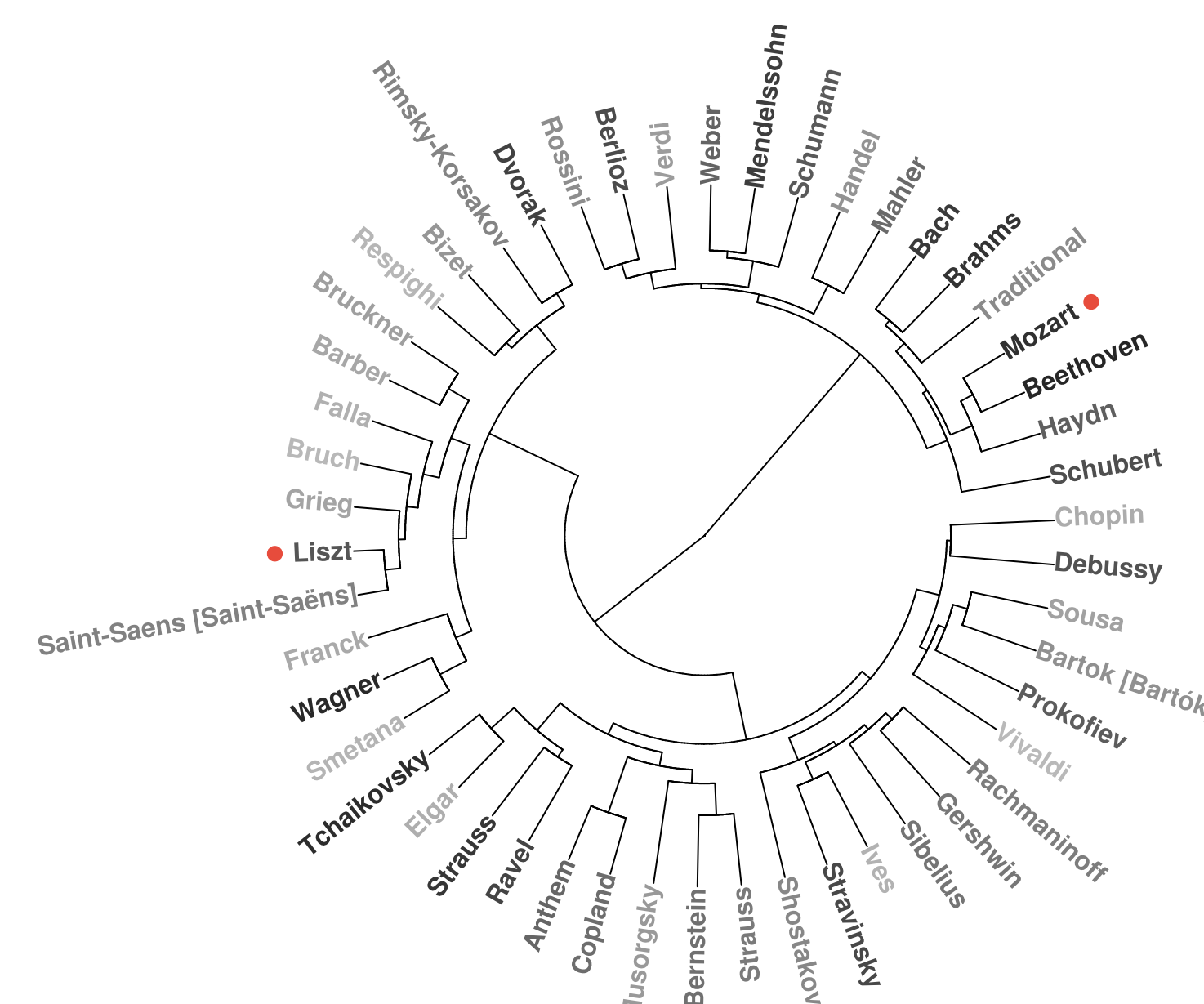
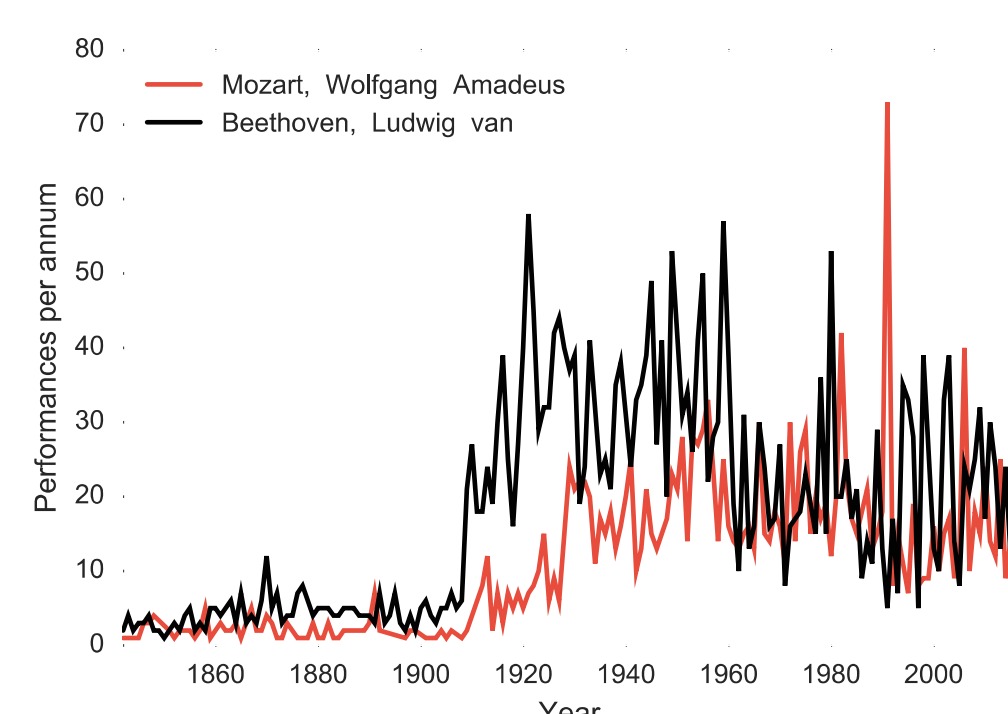
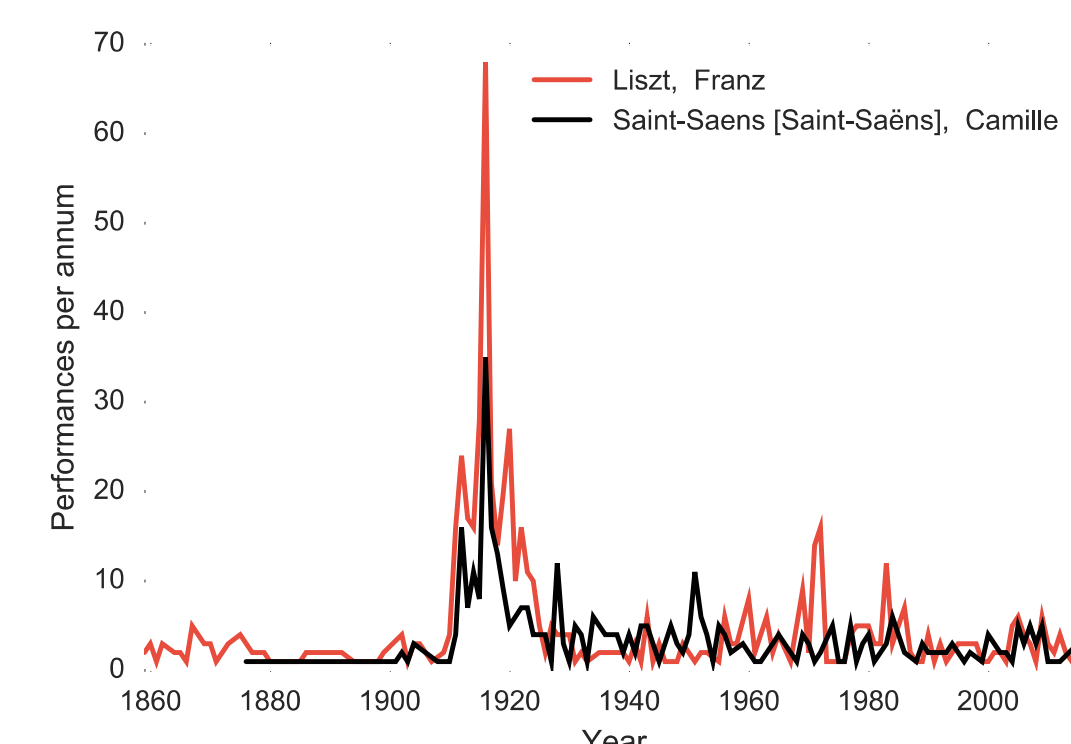
Concert programs

- Every concert consists in a selection of works by some composers
- **Who are the most ‘characteristic’ composers in a generation?**
- Represent the performances in a decade as ‘bag-of-composers’ (composer/performance-count pairs)
- Apply **tf-idf** (term-frequency/interdocument-frequency) transformation
- Rank each decade’s composers by both raw and transformed performance counts
- **Raw** counts reflect canon: concert “warhorses”
- **Transformed** rankings reflect characteristic composers e.g. emergence of musical theater in 50s/60s under conductor Bernstein
- Same can be done for pieces.
- Applications to: cultural history, playlist curation

Decade	Top composer (raw counts)	Top composer (tf-idf)
1920	Wagner	Hadley, Henry Kimball
1930	Wagner	Debussy, Claude
1940	Beethoven	Tchaikovsky, Pyotr Ilyich
1950	Beethoven	Rodgers, Richard
1960	Tchaikovsky	Bernstein, Leonard
1970	Mozart	Stravinsky, Igor
1980	Beethoven	Prokofiev, Sergei
1990	Beethoven	Copland, Aaron
2000	Beethoven	Bernstein, Leonard
2010	Tchaikovsky	Bernstein, Leonard

Performance time-series clustering (number of performances per year)

A time-series of the number of performances **per composer** over time was constructed for each of the top 50 composers (by number of performances) in the dataset. The results of hierarchical clustering on these time-series using the Shape-Based Distance metric [1] are shown in the fan dendrogram below. The line charts show two examples of discovered correspondences between composer performance volumes over time.



Composers’ names at the leaves of this dendrogram are shaded according to the midpoint of their birth and death dates. Darker is earlier; lighter is later.

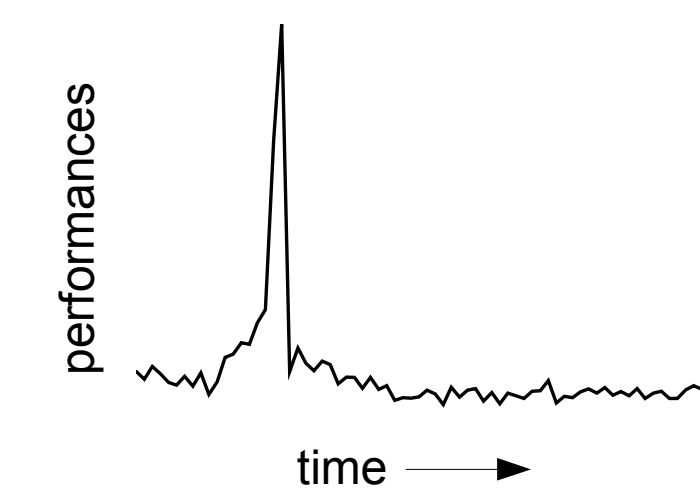
Social networks of soloists

Network of soloists constructed by adding an edge between soloists who collaborate on the same performance. Degree (inbound and outbound edges) indicates number of collaborations; eigenvector centrality indicates influence.

Soloist	Degree	Details
Stacy, Thomas	49	Cor anglais soloist (ret. 2010)
Sylar, Sherry	37	Associate Principal Oboe (since 1984)
Levinson, Eugene	34	Former Principal Double Bass
LeClair, Judith	33	Principal Bassoon (since 1981)
Drucker, Stanley	33	Clarinet; ret. 2009 after 60 years

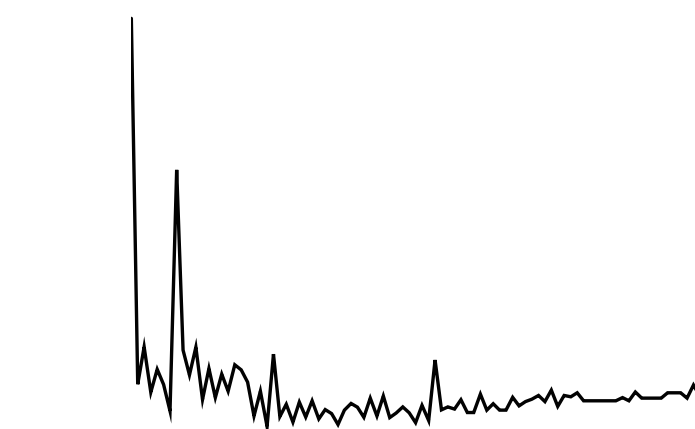
Soloist	Eigenvector Centrality	Details
Tritle, Kent	1	New York choral conductor (b. 1960)
Ralske, Erik	0.81	Met Opera Orchestra since 2010
Myers, Philip	0.72	Principal Horn since 1980
Carabella, John	0.68	Second Horn 1960-1968?
Alessi, Joseph	0.67	Principal Trombone since 1985

A similar time-series **per work** over time was constructed for each of a sample of 250 works in the dataset. The centers of learned k-Shape clusters [1] are shown in the line charts below.



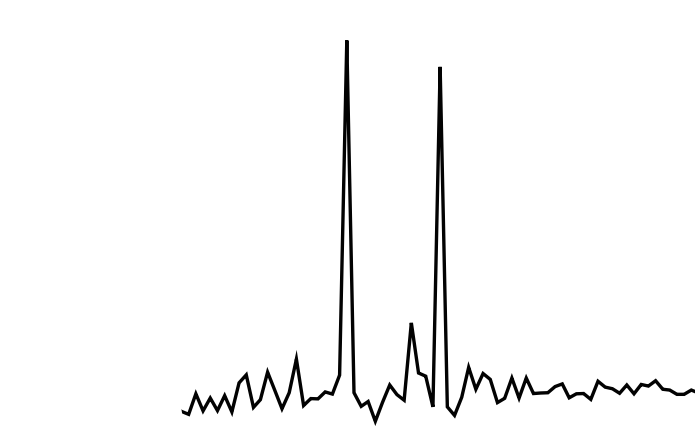
Cluster 1 (n = 48)

- Schumann, Robert OVERTURE, SCHERZO, AND FINALE, OP.52
- Siczynski, Rudolf WIEN, WIEN (VIENNA, CITY OF MY DREAMS)
- Haydn, Franz Joseph SYMPHONY NO. 101, D MAJOR (THE CLOCK)



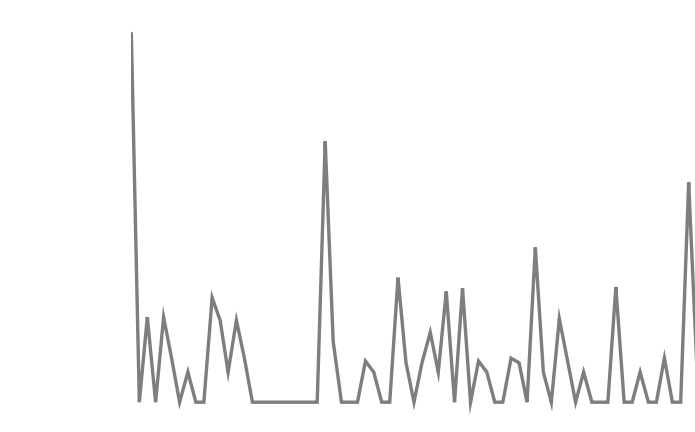
Cluster 2 (n = 22)

- Rodgers, Richard CAROUSEL
- Verdi, Giuseppe TROVATORE, IL
- Mozart, Wolfgang Amadeus CONCERTO, VIOLIN NO. 5, A MAJOR, K.219



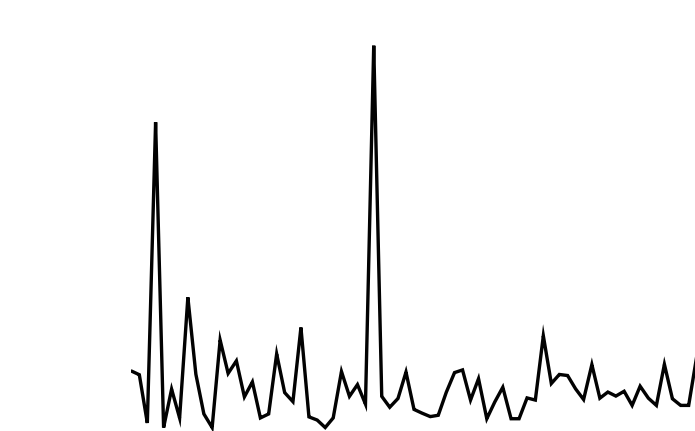
Cluster 3 (n = 24)

- Ravel, Maurice PAVANE POUR UNE INFANTE DEFUNTE
- Schuman, William NEW ENGLAND TRIPTYCH
- Strauss, Johann, Jr. & Josef PIZZICATO POLKA



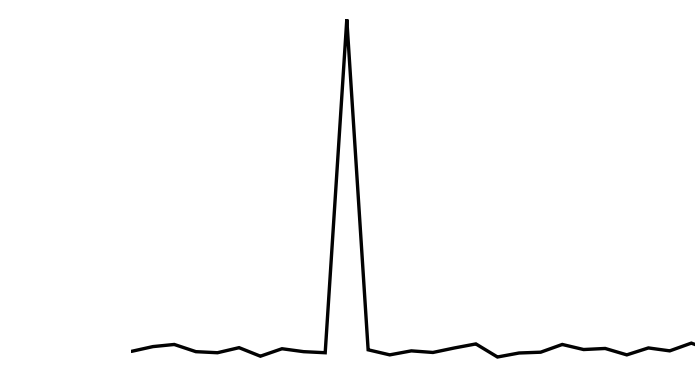
Cluster 4 (n = 4)

- Massenet, Jules MANON
- Mozart, Wolfgang Amadeus CONCERTO, PIANO NO. 22, E-FLAT MAJOR, K.482
- Ravel, Maurice INTRODUCTION AND ALLEGRO FOR HARP WITH FL/CL/STRINGS



Cluster 5 (n = 13)

- Bach, Johann Sebastian CON., VIOLIN, E MAJOR
- Goldmark, Karl SAKUNTALA OVERTURE
- Mozart, Wolfgang Amadeus CONCERTO, PIANO NO. 24, C MINOR, K.491



Cluster 6 (n = 140)

- Ponchielli, Amilcare GIOCONDA, LA
- Khachaturian, Aram GAYANE, SUITE NO. 1
- Haydn, Franz Joseph SYMPHONY NO. 95

References

- **Dataset:** New York Philharmonic Performance History, (released under CC0 at <https://github.com/nyphilarchive/PerformanceHistory>)
- **Replicate:** <https://github.com/eamonnbell/nypposter>

[1] Paparrizos, John, and Luis Gravano. "k-Shape: Efficient and Accurate Clustering of Time Series." Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data. ACM, 2015