Four Chords Go a Long Way: Measuring Chord Progression Similarity in Chinese Popular Music

Jiaying Li[1] and Nat Condit-Schultz[1]
[1]Georgia Institute of Technology, Atlanta, GA, USA

Georgia Tech | Center for Music Technology

INTRODUCTION

Music Similarity has great significance for music searching and classification, copyright maintenance, and music criticism, as well as artistic implications.

Chord Progression Similarity Index (CPSI) based on Markov Model matches music theoretic predictions well and shows high compatibility in types and numbers of chords in chord progressions.

CHINESE POP MUSIC DATABASE

o Chinese Music Market Survey:
- Users of music APPs under Tencent Music and Entertainment (TME) account for 94% of music streaming in China;
- QQ Music, Kugou Music, and Kuwo Music include 72.8% of users (TME's annual financial report)

o Sampling Method:
- Music Average Ranking: $r_{ij} = \frac{1}{3}\sum_{k=1}^{3} r_{ij}(k)$
  - song j
  - year $i \in \{2012, 2013, ..., 2021\}$
  - music APP $k \in \{1, 2, 3\}$, where k = 1 represents QQ Music, k = 2 represents Kugou Music, k = 3 represents Kuwo Music.

o Chord Progression Notations:
- We accessed existing chord-transcriptions of 200 sampled songs from the Echangwang website:
  - We encoded chord progressions from the verses and choruses of each track;
  - We transferred the key of the songs into major keys during data sampling.
- E.g: F->G->Am->C

CPSI MEASURING ALGORITHM BASED ON MARKOV MODEL

o We split every chord in a chord progression into notes. Each chord progression forms a **12 × 12** transition matrix.

o Root-based Measuring Model:
- For a chord progression K with chords $\{C_i\}$, where i represents its location index in K, we simplify it as a directed Markov chain based on its roots $\{R_i\}$. Denote the transition probability from root R_i to R_j as P_{ij}.

0	1	2	3	4	5	6	7	8	9	10	11
C	C#	D	D#	E	F	F#	G	G#	A	A#	B
or C	Db	D	Eb	E	F	Gb	G	Ab	A	Bb	B

The relationship between notes and numbers

R_0: F $\xrightarrow{1}$ R_1: G $\xrightarrow{1}$ R_2: A $\xrightarrow{1}$ R_3: C

Example of chord progression "F->G->Am->C"

o Construction of Chord Progression Transition Matrix:
- For notes N_{ij} in chord C_i, their transition probabilities to the notes N_{i+1j} in the next chord C_{i+1} are the same;
- The transition probability p_{ij} from note N_{ij} to N_{i+1j} is defined as: $p_{ij} = \frac{P_{ij}}{\max j}$

o CPSI:
- For chord progression K_1 and chord progression K_2, define their similarity index as the Euclidean distance of their note transition matrix Π_1 and Π_2:

$$\rho_{1,2} = \sqrt{\sum_{i,j}[\Pi_1(i,j) - \Pi_2(i,j)]^2}$$

o Metric Weighting:
- Metric and hypermetric position are extremely important to musical organization and to music perception, especially the first note;
- Assume that the practical human perception weights w_i for chord C_i, the transition matrix P is modified to:

$$P_{ij\,modified} = \frac{w_i P_{ij}}{\sum_{i,j} w_i P_{ij}}$$

CPSI EXAMPLES

| Chord Progression 1	Chord Progression 2	w1	w2	CPSI
C->Am->F->G	C->Am->F->G	10%	10%	0.0000
C->Am->F->G	Am->F->G->C	10%	10%	0.0552
C->Am->F->G	C->Am->F->C	10%	10%	1.6382
C->Am->F->G	C->G->Am->F	10%	10%	1.8555
C->Am->F->G	Cmaj7->Em->G->Dm	10%	10%	2.1538

Examples of chord progression pairs and corresponding CPSI

CPSI VERSIONS

o We use the CPSI to compute more fine-grained measures of similarity in the dataset and compare the following twelve versions:
- Version 1.x.x: No weight for chords;
- Version 2.x.x: Add weight to leading chord;
- Version x.1.x: Non-simplified chords;
- Version x.2.x: Simplified chords into triads;
- Version x.3.x: Root-based model;
- Version x.x.1: Progression group as basic analysis unit;
- Version x.x.2: Combining all chord progressions together as basic analysis unit.

RESULTS

o Chord Progression Exact Reuse

Non-simplified version   Simplified version   Root version
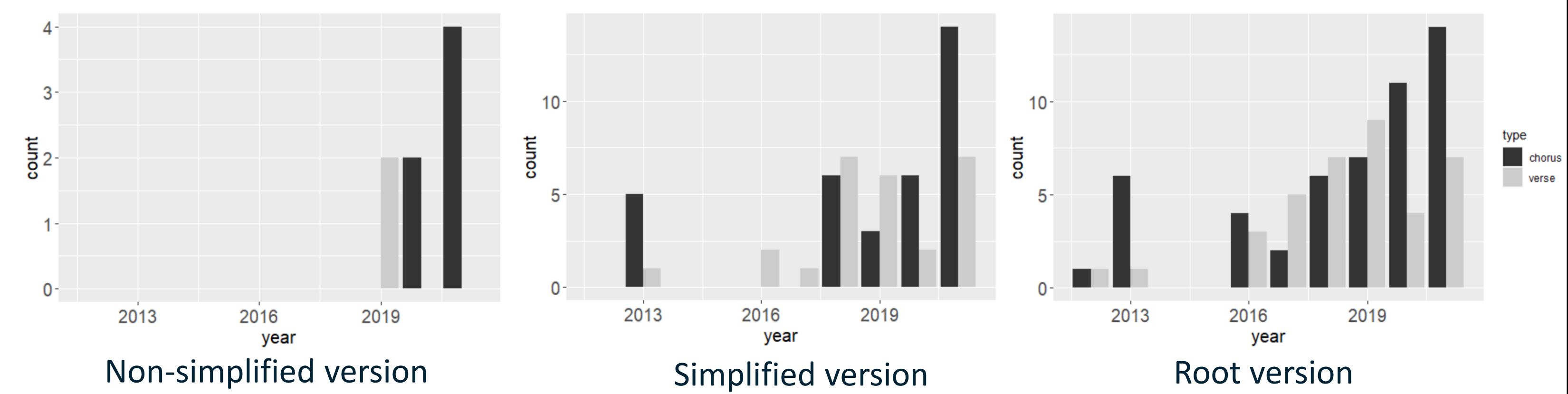
o CPSI Analyses
- In order to select the appropriate versions, we applied one-way ANOVA test to each version. The null hypothesis (H0) of the ANOVA is no difference in means for CPSI values grouped by years.

Version	p for Verse CPSI	p for Chorus CPSI	p for Total CPSI
1.1.1	0.00282 **	0.0141 *	0.00205 **
1.1.2	0.0557 .	0.0128 *	0.014 *
1.2.1	0.00298 **	0.0575 .	0.00702 **
1.2.2	0.0145 *	0.204	0.0384 *
1.3.1	0.00186 **	0.873	0.0662 .
1.3.2	0.467	0.519	0.441
2.1.1	0.00327 **	0.0122 *	0.00191 **
2.1.2	0.0562 .	0.0113 *	0.0131 *
2.2.1	0.00271 **	0.24	0.0338 *
2.2.2	0.0147 *	0.19	0.0365 *
2.3.1	0.00219 **	0.922	0.0775 .
2.3.2	0.565	0.47	0.463

ANOVA test for each version, where the significance codes '**' represents p < 0.001, '*' represents p < 0.01, and '.' represents p < 0.05.

- Only four versions are qualified versions: 1.1.1, 1.1.2, 2.1.1 and 2.1.2

CPSI Data Analysis by Mean   CPSI Data Analysis by Median
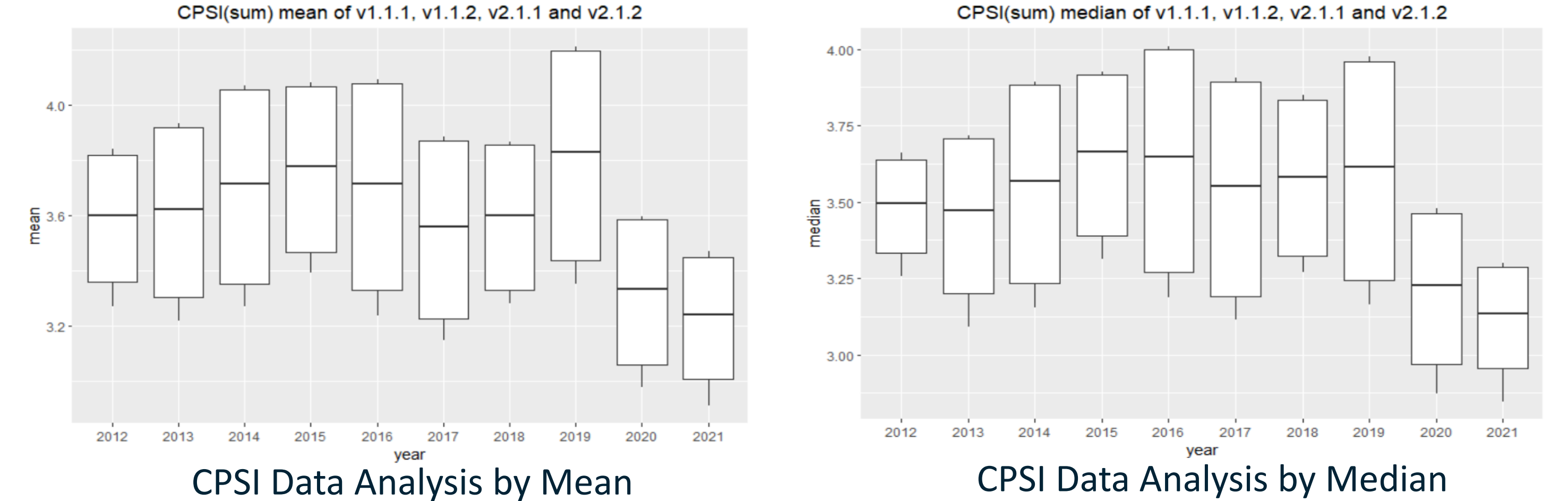
o Conclusions
- Mann-Kendall test for monotonic trend analysis shows that the decreasing CPSI is significant in the past five years.
- Comparing to the verse, chorus part more likely appears the same chord progressions.

CONTACT INFORMATION

Jiaying Li, Computational & Cognitive Musicology Group, Georgia Tech Center for Music Technology
jli3269@gatech.edu
Dataset: https://github.com/JiayingLi0803/harmonicSimilarityProject_2022