

Scalar Theory and Knowledge Discovery in the Recording Industry

Rory A. Lewis

University of Colorado at Colorado Springs, 1420 Austin Bluffs Parkway, Colorado Springs, Colorado 80933-7150, USA

Abstract. In the continuing investigation of understanding how music affects human beings from the point of data mining, this paper explores a dimension of Knowledge Discovery in Databases and Music information Retrieval that is not often explored; the music business. In this paper we refer to the music business as those industries comprising the financial and physical networks that first have the means to broadcast songs to the public domain and secondly the ability to sell said music to the masses either online or physically in stores. It is no secret that the music industry has changed substantially in the last twenty years. There was a time that the music industry controlled the masses. Nowadays, it is fair to say that the masses control what they choose to listen to and buy regardless of the discretion of music industry executives. This paper focuses on how KDD, MIR and the music industry are closely related to one another.

1 Introduction

The New York Times recently quoted Aram Sinnreich, a highly respected and powerful media industry statistician and consultant in Los Angeles, said the CD format, introduced in the United States 24 years ago, is in its death throes. "This Christmas (2008) is the last big holiday season for CD sales ... and then everything goes kaput." [12] Between 2000 and 2005, sales of music units in the US have plummeted by 25% to 705 million units [2] [22]. The music industry claimed that file sharing was the culprit and 1) sued thousands of individuals who share files and 2) requested that the Supreme Court deem file sharing as illegal. However, the notion that file sharing is the reason for the recording industry's demise is questioned by analysts and scholars alike [11] [4] [30] [28] [9]. For example, industry analysts are quoted as stating that considering the fact that Apple sold over 1 billion songs in 2007, for an overall total of 5 billion songs sold, [21] when the label's CD sales finally collapse, the record labels will be faced with a tough decision: distributing music nearly exclusively through Apple's iTunes store or rethinking their approach to digital-rights management. [6] This equates to Apple selling over 5 million songs a day, or 58 songs every second. The industry needs an empirical means to determine if a song will be loved, talked about and grow in popularity from a grass roots level. herein is the subject matter of this paper. The author presents a model incorporating statistics, MIR and KDD to present a platform to discover and measure the economic viability of songs

2 The Record Industry's Business Models without KDD

First looking at the 4 prominent statistical methods of analyzing Record Industry metrics but minus KDD and MIR, Oberholzer and Strumpf recently formalized the two most widespread and accepted business models that account for regressing total album sales, including paid digital downloads, on the average number of simultaneous file sharing users in the U.S. They used equation 9 to prove that there is little evidence that growth in the number of on line users has had a statistically or economically effected total on sales.

$$Sales_t = \gamma \times Users_t + v_m + \mu_t \quad (1)$$

where t indicates a month, and v_m are monthly fixed effects which account for seasonality. Using monthly data from August 2002-May 2006 ($N=46$) and defining Sales and Users in millions (with respective sample means of 56.0m and 5.0m), the estimated $\gamma = -0.427$ with a robust standard error of 0.33. Equation 2 provided empirical proof that, for the period 1999 - 2005, download intensity positively influenced long-run sales growth after explicitly measuring for trends in music format popularity.

$$SalesGrowth_g = \alpha + \gamma \times Downloads_g + \lambda \times Listenership_g + e_g \quad (2)$$

where g indicates genre, $SalesGrowth_g$ is the percentage growth in sales over 1999-2005, $Downloads_g$ are measures of genre-specific download intensity, and $Listenership_g$ is the genre-specific listenership growth rate [1] which controls for trends in popularity. [22] In 1985 two statistical tools were developed specifically for the Recording Industry and have been used ever since. These were the Herfindahl-Hirschman Index (HHI) and the CN index. The HHI measures market share in terms of concentration in a particular industry. The HHI, seen in equation 3, is calculated by tracking individual firms U.S. revenues and U.S. market shares in the music industry:

$$HHI = \sum_{i=1}^f S_i^2 \quad (3)$$

Where f is the number of firms participating in an industry, S_i is each firms market share and i represents a firm in a given industry. One can also use the fourth equation (4) , such as the C5 index, to crosscheck the HHI. [5] The C5 index is the combined share of the top five firms in a particular market. It is computed in the following way:

$$C5_j = \sum_{i=1}^5 S_{ij} \quad (4)$$

Where S_{ij} is a firm's i market share of a given industry j , where firms are ordered by size of market share

3 Measuring Scalar Theory in Music

The chordal progressions of songs sold is the indicator or biometric that discerns songs, mathematically, from one another. Measuring chordal progressions can be done manually, or automatically with Blind Source Separation (BSS). This paper’s experiments combines the first four equations into KDD without using BSS processing to separate the sounds, [18] [15] to recognize what instrument is playing what notes to determine the scale and key of the song. Eventually this will be done. In this paper the author first experiments with the viability of the notion that a measuring technique using KDD can be introduced to the statistical models Equations (1 - 4) The field of neuro-endocrinology confirms that music scales evoke signals in the neurotransmitters of humans called emotions [23], [31], [24] MIR accepts this [27] [20] and performs extensive research mining for patterns of emotions in music.[26] [10] [19] The problem was in mining the correct scale and key [29], [32] however Lewis and Raś [25] solved this which leads to the next step which is to measure the tension by synthesizing chordal progressions.

3.1 Chordal Progressions

Of the many attributes in a song that a record executive would consider, apart from the artist’s looks and voice, or the words for the song. When the record executives would listen to a song he would, “hear” chords and determine if that song was going to be a hit or not. Because we calculate the relationships of the fundamental frequency. Combinations of distinct notes on a particular scale and key are called chords. Chords can be played by one instrument such as guitar or piano or as a combination of many instruments each playing particular notes of a chord. When a MIR system analysis a polyphonic wave form it outputs a garbled chaotic stream of notes. For in-depth analysis in referenced areas such as calculating the fundamental frequency of notes (*see* [16]), mining music scalar theory (*see* [13]) in a music database (*see* [33]) and setting a non-Hornbostel hierarchical manner (*see* [14] [17]). The fundamental frequency, f_1 , of a sound wave establishes pitch. The distance of all the other frequencies within an octave, starting from a given note of frequency f_1 and going up in the frequency scale, can be calculated according to the following formula:

$$f_k = f_1 \cdot 2^{k/12} \quad (5)$$

where k is number of semitones separating f_k and f_1 . However, we operate in the non-temporal domain where the fundamental frequency in each frame i is calculated in the form:

$$f(i) = \frac{s}{K_i/n_i} \quad (6)$$

where s is the sample frequency, n_i is the total number of $r(i, k)$ ’s local valleys across zero, where $k \in [1, K_i]$ and K_i is the estimated maximum fundamental period. Here we let $r(i, k)$ be the normalized cross correlation of frame i with lag k . To calculate $r(i, k)$, we look at how it reaches its maximum value

$$r(i, k) = \frac{\sum_{j=1}^{m(i)+n-1} \alpha(j) \alpha(j-k)}{\sqrt{\sum_{j=m(i)}^{m(i)+n-1} \alpha(j-k)^2 \sum_{j=m(i)}^{m(i)+n-1} \alpha(j)^2}} \quad (7)$$

where $k \in [1, K_i], i \in [0, M-1]$ and α is the audio signal, n is the frame size, $m(i) = i * n$, where i is the frame index, M is a number of frames. Our analysis window size is set to 20ms.

In essence, when the record executive would listen to the aforementioned chords he was weighing how the tension, the pull and give, of the music made sense for a particular genre. Tension describes how humans perceive uneasiness or easiness towards a set of tones. Cook's tension function, (equation 8) calculates the difference in interval sizes, is Gaussian in shape where the effects of upper partials calculate the triadic tension [8].

$$t = v \cdot \exp\left(\frac{y-x}{\lambda}\right)^2 \quad (8)$$

where t is the tension, v is the product of the relative amplitudes of the three partials, λ (in our case, $\lambda = 0.60$) is a parameter that determines the steepness of the fall from maximal tension; x and y are the lower and upper of the two intervals in each tone triplet, defined as $x = \log(f_2/f_1)$ and $y = \log(f_3/f_2)$, where the frequencies of the three partials satisfy the relationship: $f_1 < f_2 < f_3$ (in Hertz). The vast majority of scales in contemporary western music consist of 5 to 7 different notes (pitches). To calculate the number of possible scales we assert that the starting and ending notes are fixed and that there are twelve notes in an octave leaving 10 notes between the starting and ending notes [13] where one can see that each octave has one key corresponding to each of the 12 note names. The normal alphabet letters are white keys, the sharps/flats are the black keys. "C" ["C#" "Db"] "D" ["D#" "Eb"] "E" "F" ["F#" "Gb"] "G" ["G#" "Ab"] "A" ["A#" "Bb"] "B".

Chordal Progressions Special Intervals on scales called imperfect fifths, major thirds, minor thirds and the harmonic triads formed 1) the basis of western music until the romantic era, and 2) the basis of the blues and folk music etc. Musicians represent scales in numerous forms all of which are incompatible with knowledge discovery. Scales can be represented using the *spiral array*, where pitch states are associated by coordinates downward along an ascending spiral [7] but its inefficient for discretization. With this in mind the authors chose to represent basic score classification of music not as a music system but rather as an information system $S = (X, A, V)$, called Scale Table, where $A = \{J^I, J^{II}, J^{III}, J^{IV}, J^V, Scale, Region, Genre, Emotion, sma\}$ (see Table 1) [33]. There are N semitones including the tonic t_1 which forms the remaining notes t_2, \dots, t_M in the scale which in turn are distributed over the remaining $N-1$ points. Also, we consider each note by moving from the lowest to the highest note. We cannot repeat a note and this is leaving one possible order, or scale.

X	J^I	J^{II}	J^{III}	J^{IV}	J^V	Scale	Region	Genre	Emotion	sma
X_1	2	2	3	2		Pentatonic Major	Western Blues		melancholy	s
X_2	3	2	1	1	2	Blues Major	Western Blues		depressive	s
X_3	3	2	2	3		Pentatonic Minor	Western Jazz		melancholy	s
X_4	3	2	1	1	3	Blues Minor	Western Blues		dramatic	s
X_5	3	1	3	1	3	Augmented	Western Jazz		feel-good	s
•										
•										
X_{31}	2	1	4	3		Minor 9th	neutral	neutral	not happy	a
X_{32}	2	2	1	2	3	Major 11th	neutral	neutral	happy	a
X_{33}	2	1	2	2	3	Minor 11th	neutral	neutral	not happy	a
X_{34}	4	4				Augmented	neutral	neutral	happy	a
X_{35}	3	3	3			Diminished	neutral	neutral	not happy	a

Table 1. Score Classification Scale Table: In the same way that record executives would judge the viability of the chord progressions we represent jumps between notes in a scale X are represented by $J^I, J^{II}, J^{III}, J^{IV}, J^V$ which correlate to specific scales, regions and genre of music. The values of attribute sma in Table 1 listed as $\{s, m, a\}$ should be read as *scale, mode, arpeggio*

Major chords are formed by the combination of the I, III and V note of the Major Scale For example, the C Major chord is comprised of C, E and G notes [3]. Some music hits only have major chords whereas other songs use tension to divert away from the major chords. The authors represent the distance from the tonic to itself as zero, so we eliminate the redundancy of stating that the distance from the tonic to itself by starting with the distance from the tonic to the 2nd note in the scale.

1	2	3	4	5	6	7
C	D	E	F	G	A	B
0	2	2	1	2	2	2

Table 2. C Major: Header represents the Tonic at 1. For a Major scale in the key of “C”, the tonic is “C”, the 2nd note, 2m is “D”, the 3rd note 3m is and “E” and so forth.

This makes sense to human ears as these major chords sound pleasant. The question is, how come what is pleasant to our ears is so chaotic to machine learning? First a quick look at the physics of chords.

4 Experiment

To determine the relationship between statistical analysis and MIR analysis this paper presents the relationship that does exist between songs that are successful economically and songs that are successful in scalar theory. In short, this paper presents, be it only the beginning (3 notes), of a model that defines what a good record executive would hear that would make him decide a song would

bring in money. The author is not considering attributes such as artists looks, voice, appeal or age. Nor is the author considering melody line, lyrics and instrumental solos. All these items will be addressed in future work. This paper simply demonstrates that a machine, with MIR can mine the music of a song in a similar way that a record executive would listen to a new song. To do this the author has made a double blind domain using songs of the genre blues where the chordal progressions are distinct and the tension is typically steadfast. 180 songs comprising hits and non hits are in the data base with the chordal jumps and progressions. Scores are arranged and determined in according to the author's formula seen in Equation 9. WEKA was used to account for a raw chart score, the MIR score and finally the author's validation score as seen in Equation 9.

$$ChartStrength_t = \int_{\gamma v_m}^{\gamma \mu_t} \frac{\sum_{j=1}^{m(i)+n-1} \alpha(j)\alpha(j-k)}{\sqrt{\sum_{j=m(i)}^{m(i)+n-1} \alpha(j-k)^2 \sum_{j=m(i)}^{m(i)+n-1} \alpha(j)^2}} \quad (9)$$

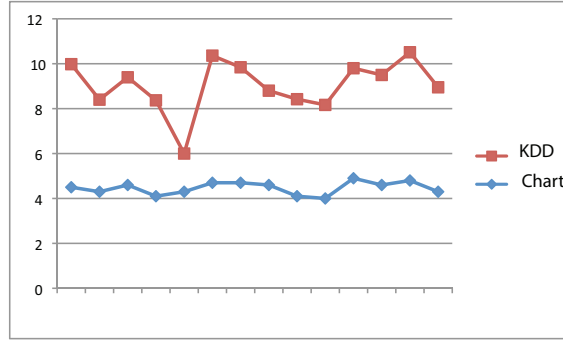


Fig. 1. Fourteen chordal variations of the blues scalar theory based upon blues songs that were successful and those that were not. Each group is amortized to give an average chart score for that type. Integrating the fourteen types into a validating score shows a resemblance to a human's ability to discern what a successful Blues song will be financially

5 Future Work

The ability to show, even in TeX smallest way that KDD methodology can indeed draw a distinction between how a human being can recognize a hit and how KDD can data mine a hit is shown in this paper. Of course the parameters and the domain are highly controlled. This leads to the obvious future work which will be to expand these domains and eventually have KDD provide a means for record executives to help discern a hit from a non hit.

References

1. *Format Trends Report*, 2006.
2. Riaa 2006: The recording industry association of americas 2005 yearend statistics. 2006.
3. A. Shenoy, R. Mohapatra, and W. Yeg. Key determination of acoustic musical signals. *Multimedia and Expo, June 2004. ICME '04. 2004 IEEE International Conference on*, 3(1):1771–1774, 2004.
4. Y. Bakos, E. Brynjolfsson, and D. Lichtman. Shared information goods. *Journal of Law & Economics*, 42(1):117–55, April 1999.
5. t. j. y. v. n. Blanchette, K.
6. E. V. Buskirk. Who’s killing mp3 and itunes? *Wired Magazine*, January(08), 2007.
7. E. Chew. Music information processing a new application for operations research. *Bulletin of AIRO - Associazione Italiana di Ricerca Operativa, Optimization and Decision Sciences*, VII, no.3:9–14, 2002.
8. N. Cook and T. X. Fujisawa. The psychophysics of harmony perception: Harmony is a three-tone phenomenon. *Musicology Review, Parncutt and Richard Publishing*, 1, No. 4:204–209, 2006.
9. M. Fetscherin. Movie piracy on peer-to-peer networks: the case of kazaa. *Telemat. Inf.*, 22(1-2):57–70, 2005.
10. K. Hevner. Experimental studies of the elements of expression in music. *American Journal of Psychology*, bf-48:246–268, 1936.
11. S. Knopper. *Appetite for self-destruction : the spectacular crash of the record industry in the digital age / Steve Knopper*. Free Press, New York :, 2009.
12. J. Leeds. Plunge in cd sales shakes up big labels. *New York Times*, May(28), 2007.
13. R. Lewis and Z. Raś. Rules for processing and manipulating scalar music theory. *Proceedings of the International Conference on Multimedia and Ubiquitous Engineering*, pages 819–824, 2007.
14. R. Lewis and A. Wiczorkowska. Categorization of musical instrument sounds based on numerical parameters. *June 28-30 in Warsaw Poland*, 4585/2007:784–792, 2007.
15. R. Lewis, X. Zhang, and Z. Raś. Blind signal separation of similar pitches and instruments in a noisy polyphonic domain in foundations of intelligent systems lnai no. 4203 springer.
16. R. Lewis, X. Zhang, and Z. Raś. Knowledge discovery based identification of musical pitches and instruments in polyphonic sounds. *in the Special Issue on*.
17. R. Lewis, X. Zhang, and Z. Raś. Mirai: Multi-hierarchical fs-tree based music information retrieval system. *June in Warsaw Poland*.
18. R. Lewis, X. Zhang, and Z. Raś. Knowledge discovery based identification of musical pitches and instruments in polyphonic sounds. *International Journal of Engineering Applications of Artificial Intelligence*, Volume 20 Issue 5 August:637–645, 2007.
19. T. Li and M. Ogihara. Detecting emotion in music. *International Conference on Music Information Retrieval ISMIR, Washington, D.C., and Baltimore, Maryland*, 4th, 2003.
20. R. McClellan. The healing forces of music. *Rockport, MA*, Element Inc., 1966.
21. S. McNulty. Apple tops 5 billion mark in itunes sales. *TUAW*, June(19), 2008.
22. F. Oberholzer and K. Strumpf. The effect of file sharing on record sales: An empirical analysis. *Journal of Political Economy*, 115, 2004.

23. I. Pavel. A hierarchical theory of aesthetic perception - musical scales in theoretical perspectives prometheus. *Leonardo*, Vol. 27 No. 5:417–421, 1994.
24. I. Peretz and K. Hyde. What is specific to music processing? insights from congenital amusia. *Trends Cogn. Sci.* 7, pages 362–67, 2003.
25. W. J. R. Lewis and Z. Raś. Mining scalar representations in a non-tagged music database. in *Foundations of Intelligent Systems*, 2007.
26. N. S. Rickard. Intense emotional responses to music: a test of the physiological arousal hypothesis. *Psychology of Music*, 32, 4th:<http://pom.sagepub.com/cgi/reprint/32/4/371>, 2004.
27. A. Sevgen. The science of musical sound. *New York, NY, ScientificAmerican Books, Inc.*, 1983.
28. C. Shapiro and H. R. Varian. *Information rules: a strategic guide to the network economy*. Harvard Business School Press, Boston, MA, USA, 1998.
29. J. Sloboda. Music and the emotions, british association festival of science. *The Psychology of Music*, page <http://www.ummah.org.uk/forum/showpost.php?p=436890>, 1996.
30. L. Takeyama. The welfare implications of unauthorized reproduction of intellectual property in the presence of demand network externalities. *Journal of Industrial Economics*, 42(1):155–166, April 1994.
31. M. Valentinuzzi and N. Arias. Human psychophysiological perception of musical scales and nontraditional music. *IEEE/Eng Medicine and Biol Mag. March-April 18 (2)*, pages 54–60, 1999.
32. A. Vink. Music and emotion. living apart together: a relationship between music psychology and music therapy. *Nordic Journal of Music Therapy*, 10, No. 2:pages 144–158, <http://www.hisf.no/njmt/artikkelvink102.html>, 2001.
33. A. Wiczorkowska, P. Synak, R. Lewis, and Z. Raś. Creating reliable database for experiments on extracting emotions from music. *Intelligent Information Processing and Web Mining, Advances in Soft Computing*, pages 395–404, 2005.