# Quantifying the Salience of Musical Characteristics From Unstructured Text

Gopala Krishna Koduri and Xavier Serra

Music Technology Group, Universitat Pompeu Fabra, Barcelona, Spain
gopala.koduri@upf.edu

**Abstract.** Music is a discerning window to the rich diversity of the world. We hypothesize that identifying the differences between music from different cultures will lead to richer information models representative of them. Using five music styles, this paper presents a novel approach to bring out the saliences of a given music by rank-ordering its characteristics by relevance using a natural language text corpus. The results agree with the cultural reality reflecting the diverse nature of the music styles. Further, to gather insights into the usefulness of this knowledge, an extrinsic comparative evaluation is performed. Similarities between entities in each music style are computed based on a salience-aware semantic distance proposed using the knowledge acquired. These are compared with the similarities computed using an existing linked-data based distance measure. A sizable overlap accompanied by an analysis of users' preferences over the non-overlapping portions indicate that the knowledge acquired using our approach is indeed musically meaningful and is further complementary in nature to the existing structured information.

## 1 Introduction

Music traditions from around the world share a few common characteristics. Yet, they differ substantially when viewed within their geographical and cultural context [1]. Even among the seemingly usual characteristics, such as the musical concepts (melody, rhythm, ...) and the people involved in making the music (performers, composers, ...), their relevance and role vary from music to music. Consider the role of dance in different music styles. In Flamenco, it becomes an integral part of the music and is therefore seen as an important aspect of the music itself. Whereas in Jazz, it is not as closely associated.

Most commercial music platforms are agnostic to such differing characteristics of music, which inhibits them from scaling their recommendation services to meet the cultural diversity. To a certain extent, collaborative filtering techniques [2] and context-based recommendation systems [3] implicitly avail such information latent in listener activities and the community provided data such as tags. However, to our knowledge, there are no known approaches that explicitly incorporate the relevance of different musical characteristics.

We formally define the problem of quantifying the relevance or salience of characteristics of a given music as follows. $E$ is a set of entities that make up the music, which includes its entire vocabulary. $C$ is a set of its characteristics. Any given entity can posses more than a characteristic. $C_k$ is a set of entities that share a characteristic, $c_k$. Entities include names of scales, chords, raagas, rhythm cycles, people and so on. An example for a characteristic is composing ($c_k$). All the entities who possess this characteristic constitute a set, $C_k$.

$$\begin{aligned} E &= \{e_i \mid e_i \text{ is an entity}\} \\ C &= \{c_i \mid c_i \text{ is a characteristic}\} \\ C_k &= \{e_i \mid e_i \text{ has a characteristic } c_k\} \end{aligned} \tag{1}$$

The first part of this paper presents our system, called *Vichakshana*[1], for quantifying the salience of the characteristics ($C$) of a given music and rank-ordering them, thus bringing out the most defining aspects of each music. Using the scores of $C$, we then propose a salience-aware semantic distance ($SASD$) to discover the related entities of a query entity. In the second part of the paper, We use an evaluation methodology to compare the results of a recommendation system[2] using $SASD$ with a linked data based recommendation system [4]. Our primary intention is to understand the common and complementary aspects between the knowledge available as linked open data and the information our approach extracts.

The remainder of the paper is organized as follows. In sec. 2 & 3, we discuss the related work and describe the data we work with, respectively. Secs. 4 & 5 present our approach with details of its application on different music styles, and the $SASD$. In sec. 6, we present the evaluation methodology and an extrinsic comparative analysis of the recommendation system built using $SASD$. In sec. 7, we conclude with a summary of the paper, the current work in progress and possible future directions.

## 2 Related work

Within the context-based MIR, there are broadly two classes of approaches based on the characteristics of the semantic information they obtain from the data. One class of approaches take advantage of the knowledge in such data in an *implicit* manner for applications ranging from playlist generation and auto tagging, to music recommendation, search engines and interfaces (see [3] for a recent review of the related work).

The other class of approaches can be understood as belonging to the larger linked open data and semantic web movements [5]. There has been a considerable effort in publishing structured music information such as DBTune[3], LinkedBrainz[4] and Music Ontology [6] besides others. As of today, music related information makes up for a sizeable portion of the available linked open data. Freebase[5], the open-licensed data source which powers Google's knowledge graph, contains close to 200 million triples

[1] *Vichakshana*, in Telugu language, means wise discretion.
[2] In this paper, we use the term *recommendation system* loosely to mean any system that can be used in retrieving related entities.
[3] http://dbtune.org/
[4] http://linkedbrainz.org/
[5] http://freebase.com/

about music, making it the largest topic so far. Together with ontologies, linked open data cloud forms a rich source of explicit knowledge about music.

As a result, there is a growing interest in both enriching and using linked open data for developing new approaches for context-based MIR. Foafing-the-music [7] uses multiple data sources such as contextual-data from RSS feeds, content-based features and user profiles, and combines them using ontologies. This information is further used for music recommendation. Jacobson *et al* [8] show that the network structure of artists on Myspace correspond to the genres. This data is also published using music ontology [6] to be used in other research. DBrec system [4] uses a linked-data semantic distance (LDSD) to relate musical entities, based on the link structure of the music-related resources on DBpedia, resulting in explanatory recommendations. Our approach borders between these two classes. It aims to extract information from unstructured contextual-data that can be exposed to be used by music applications.

## 3 Data

The natural language descriptions are a rich source of data about a given music. The web is voluminous in this sense, but also very noisy: with varying spellings, scholarly value etc. As the impact of such noise on the results is difficult to keep track of, we chose to present the results of our approach using text corpus extracted from the Wikipedia. Further, for our work, we need to acquire the characteristics of a given music. Automatically detecting them is part of the research on ontologization at the intersection of information extraction and knowledge engineering domains, which is a challenge in itself. These characteristics often directly correspond to the subsumption hierarchies and the class memberships in ontologies [9]. In this paper, we address the issue of rank-ordering the characteristics based on their salience. Therefore, in order to avoid digression from the problem being addressed, we rely on Wikipedia for obtaining the characteristics which roughly correspond to the categories each page is associated with. We keep only the plain text from the pages removing other structured information such as hyperlinks, info-boxes and tables.

| **Music** | **Pages** ($E$) | **Categories** ($C$) | **Words** |
|---|---|---|---|
| Baroque | 2439 | 2476 | 901243 |
| Carnatic | 618 | 631 | 251533 |
| Flamenco | 322 | 1113 | 100854 |
| Hindustani | 697 | 492 | 317241 |
| Jazz | 21566 | 14500 | 5797726 |

**Table 1.** Details of the text-corpus taken from Wikipedia.

We have selected five different music styles to work with: two Indian art music traditions (Carnatic and Hindustani), Baroque, Flamenco and Jazz, which together constitute a diverse set of music styles. Table. 1 shows the number of pages, categories (which correspond to $E$, $C$ respectively), and the cumulative number of words across all the pages for each music style.

## 4 Vichakshana

A given entity in a music can be characterized by the references to other related entities in its description. In a way, such references can be understood to *explain* the given entity. Analysis of the structure of a network of references combined with the characteristics of each entity would yield us certain insight into the nature of the music. This is the intuition that our approach, *Vichakshana*, builds upon.

The process broadly consists of three steps: entity linking, entity ranking and salience computation. The first step involves identifying the references to other entities from the content of a given page. This is performed using the DBpedia spotlight[6], which uses a combination of language-dependent and -independent approaches to contextual phrase spotting and disambiguation [10]. A weighted directed graph ($G$) is created with the entities as nodes and the references as edges. The weight of an edge ($w_{e_i,e_j}$) is defined as follows:

$$w_{e_i,e_j} = \frac{n_{e_i,e_j}}{\Sigma_k n_{e_i,e_k}} \tag{2}$$

where $n_{e_i,e_j}$ is the number of references from $e_i$ to $e_j$. We have observed that the link structure in the graphs thus obtained is very sparse. Therefore, the references to entities which are outside the set of $E$ are eliminated. Table. 2 shows topology of all the graphs.

| Graph | Nodes | Edges | Density | Avg. Clust. | Avg. Deg. |
|---|---|---|---|---|---|
| Baroque (I) | 14278 | 44809 | 0.0002 | 0.002 | 3.14 |
| Baroque | 2059 | 7118 | 0.0017 | 0.018 | 3.46 |
| Carnatic (I) | 4524 | 12952 | 0.0006 | 0.003 | 2.86 |
| Carnatic | 602 | 3291 | 0.0091 | 0.03 | 5.47 |
| Flamenco (I) | 2671 | 5459 | 0.0008 | 0.004 | 2.04 |
| Flamenco | 312 | 846 | 0.0087 | 0.027 | 2.71 |
| Hindustani (I) | 7011 | 17754 | 0.0004 | 0.002 | 2.53 |
| Hindustani | 681 | 3774 | 0.0081 | 0.027 | 5.54 |
| Jazz (I) | 87918 | 381514 | 0.0 | 0.004 | 4.34 |
| Jazz | 17650 | 119107 | 0.0004 | 0.019 | 6.75 |

**Table 2.** Topology of the graphs obtained on entity linking, before and after the references to entities outside $E$ are eliminated. Rows with `(I)' denote the former.

In order to compute the salience score for a given $C_i$, we require a measure for the relevance of the constituting entities in the given music. Pagerank is a widely used

[6] We use a locally deployed version of DBpedia spotlight with the statistical backend, available openly at https://github.com/dbpedia-spotlight/dbpedia-spotlight

algorithm to compute the relevance of a node in a hyperlink graph [11]. Intuitively, it is an iterative algorithm in which nodes acquire relevance from their incoming edges. A reference from a node with a high pagerank to another node contributes positively to the relevance score of the latter. In this sense, it can also be understood as a variant of the eigenvector centrality [12]. We use a slightly modified version of the original pagerank algorithm to use edge weights in propagating the score of a given node to its neighbors. Eq. 3 describes the corresponding computations.

$$\begin{aligned} A_{e_i,e_j} &= w_{e_i,e_j} \\ D_{e_i,e_i} &= max(e_i^{out}, 1) \\ P &= D(D - \alpha A)^{-1}\beta \end{aligned} \tag{3}$$

where $A$ is the adjacency matrix corresponding to the graph $G$, $D$ is the diagonal matrix with the diagonal elements set to the out degree of the corresponding node ($e_i^{out}$), $P$ is the resulting pagerank values of all the nodes. $\alpha$ is an activation constant set to 0.85 in our analysis, and $\beta$ is an array of additive constants which are all set to 1. For more explanation on pagerank and the constants, we refer the reader to [12, 11].

Given a $C_k$, a naive and simple salience score can be the mean of pagerank scores of all the constituting entities. Remember that an entity can have multiple characteristics. A simple scoring method, such as this one, would imply that the pagerank score of a given entity equally contributes to the salience score of every $C_k$ it belongs to. However, it is desirable that an entity contributes more to those characteristics which are more specific to it. As our data does not contain this information, we hypothesize that the fewer the number of other entities which share a characteristic with the given entity, the more specific it is. Formally, if each entity ($e_i$) represents a document with the characteristics ($c_i$) as the terms, the inverse document frequency of a $c_k$ with respect to $E$ would yield us a measure that can be used to weigh the pagerank score of a given entity in computing the saliences of all $C_k$ it is associated with. Eq. 4 describes this process in detail along with the steps for computing the salience score of a $C_k$ (given by $S_{C_k}$) from the pagerank values of its entities.

$$\begin{aligned} \text{idf}(c_k) &= \log \frac{|E|}{1 + |\{e_i \in C_k\}|} \\ S_{C_k} &= \frac{1}{|C_k|} \sum_{e_i \in C_k} P(e_i) \times \text{idf}(c_k) \end{aligned} \tag{4}$$

This gives us a list of characteristics of a music ordered by their salience. We have observed that several characteristics have a considerable overlap between them. For instance, the characteristics *Music festivals in India* and *Carnatic music festivals in India* have more or less the same set of entities, with respect to Carnatic music. We consider them redundant even though semantically one is a more specific form of the other. As we will see in sec. 5, this is undesirable for applications using these salience scores. We handle such cases by *merging* them and assigning a common rank to each such group. This is performed using an undirected weighted graph constructed with the characteristics ($C$) as nodes. The weights of edges correspond to the cosine similarity between the corresponding sets of entities, $C_i$ and $C_j$. Those edges with a weight less than 0.5 are

| Baroque | Carnatic | Flamenco | Hindustani | Jazz |
|---|---|---|---|---|
| Anglican saints, Organ improvisers | Carnatic music terminology | Spanish dances, Spanish folk music | Carnatic music | African-American music, Western swing ... |
| Music in Leipzig, Thomaskantors | People from Tiruvarur district | People from Algeciras, Spanish people of Portuguese descent | Formal sections in music analysis | Burials at Woodlawn Cemetery (Bronx) |
| Composers for cello | Indian classical music | Andalusian music | Indian classical music | Rec labels- established in 1916, disestablished in 1940 |
| Harpsichord, Keyboard instruments | Ragas | Andalusian music, Flamenco styles, Spanish music, Vocal music | String instruments | Presidential Medal of Freedom recipients |
| Music catalogues | Chennai culture | Latin jazz musicians, Spanish guitarists | Hand drums | Bass (sound) |
| Baroque instruments, Necked bowl lutes ... | Carnatic music | 1950 births | Ragas | American jazz |
| Collected editions of classical composers, Johann Sebastian Bach | Indian Vaishnavites, Kannada people | Romani guitarists | Bangladeshi-, Hindustani-, Pakistani-musical instru. | ABC Records artists |
| 1685 births | Dvaita, Indian philosophers ... | Spanish musicians | Culture of- Bihar, Uttar Pradesh ... | Amplified instruments, Bass guitars ... |
| People from Halle (Saale) | Carnatic Ragas | People from Córdoba, Andalusia | Necked bowl lutes, String instruments with sympathetic strings ... | Jazz ensembles |
| German Lutherans | Hand drums, Pitched percussion ... | Male ballet dancers, Spanish dancers | Hindustani music | American Buddhists, Converts to Buddhism ... |
| English people of German descent ... | Telugu people | Cancer deaths in Spain | Sitars | EMI |
| Cantatas by Johann Sebastian Bach | Hindu monarchs, Maharajas of Travancore ... | 1958 births | Carnatic music terminology | Jazz instruments |
| Medieval music | Bhakti movement | People from Cadiz | Music schools in India, Vocal gharanas | Jazz genres |
| Composers for violin | 1680 deaths, Hindu poets, History of Andhra Pradesh, Telugu poets | 2004 deaths | Carnatic Ragas | Pablo Records artists |
| 1750 deaths | People from Thanjavur district | Flamenco groups | Dark Horse Records artists, Grammy Award-winning artists | Companies based in California, Labels distributed by UMG ... |

**Table 3.** Top 15 characteristics ordered by their salience to different music styles. Note that as Carnatic and Hindustani share a large portion of musical terminology which are categorized into Carnatic music on Wikipedia, we see many Carnatic music characteristics for Hindustani music.

filtered out, and then the closely related communities are identified using the Louvain method [13]. Each such community represents a group of characteristics which have a great overlap between the corresponding entities, and is assigned a common rank based on the new salience score recomputed using eq. 4 considering each community as a single characteristic.

It is also observed that the weights from Eq. 4 inadvertently resulted in a high rank for characteristics that are relevant to a musician, but not to the given music in general. For instance, if a very popular musician also happens to be a politician, the political characteristics are ranked high even though they are irrelevant to the music. However, if there is a certain regularity to such associations (eg: more musicians are also politicians), it is dsizeesirable to incorporate and rank those characteristics. Towards this extent, we constrain the ranked characteristics to a set of those which have at least a minimum number of entities linked to them. A high threshold includes the risk of discarding meaningful characteristics, while not employing such threshold would result in spurious ranking. Merging overlapping characteristics minimizes the impact of irrelevant ones provided they are not associated to a musical entity by chance. We have empirically chosen the value for the threshold to be three.

Table. 3 shows the top few characteristics of each music style, ordered by their salience to the music. The similarities and contrasts between the music styles are quite apparent. In Baroque, Carnatic and Flamenco, various groups of people identified by their region/language occupy a prominent place, while in Hindustani and Jazz, such groups are relatively less prominent. This might be due to the fact that the latter two are spread out over larger regions than the former three. In Carnatic and Hindustani, the terminology and musical concepts turn out to be more relevant than in other music styles. In Jazz and Flamenco, the salience of record labels is quite high whereas in other music styles, it is almost non-existent or very less. This can be due to the fact that the primary medium by which Carnatic/Hindustani music reaches people is a concert. In the case of Baroque, this is because it is no longer an active music tradition. The results also highlight the distinct features of each music. The prominence of religion in Carnatic music, dance in Flamenco, and Gharanas/schools in Hindustani music is noticeable and each contrasts with other music styles.

A direct objective evaluation of these results in not feasible as it is impractical to obtain a consensus on a highly subjective notion such as the relevance or salience of something/ somebody in a music. Therefore, we present an extrinsic evaluation of the results using the task of music recommendation. We use the salience scores of the characteristics of a given music to relate the entities using a distance measure, and compare the results with a recommendation system that feeds on linked-data.

## 5 Salience-aware semantic distance

Using the graph ($G$) and salience scores ($S$), we propose a salience-aware semantic distance (SASD). It is a weighted average of three parts. The first and prominent part is a function of salience scores. The second part is a function of length of the shortest path between the two nodes in $G$, while the third part is a function of their cocitation index, which is the number of other nodes in $G$ that point to both the given nodes. Eq. 5

formally defines the three parts and the sum. The values of all the parts of the distance and the weighted sum range between 0 (nearest) and 1 (farthest).

$$
\begin{aligned}
S_{e_i,e_j} &= \{S_{C_k} \mid e_i \in C_k \text{ and } e_j \in C_k\} \\
A'_{e_i,e_j} &= \begin{cases} 1 & \text{if } A_{e_i,e_j} > 0 \\ 0 & \text{otherwise .} \end{cases} \\
D1 &= \frac{1}{1 + |S_{e_i,e_j}| + mean(S_{e_i,e_j})} \\
D2 &= \frac{p^2_{e_i,e_j}}{1 + p^2_{e_i,e_j}} \\
D3 &= \frac{\Sigma_k A'_{e_k,e_i} A'_{e_k,e_j}}{\sqrt{\left|\Sigma A'_{e_k,e_i}\right| \times \left|\Sigma A'_{e_k,e_j}\right|}} \\
SASD_{e_i,e_j} &= \left(\frac{1}{2}\right) D1 + \left(\frac{1}{4}\right)(D2 + D3)
\end{aligned}
\tag{5}
$$

where $S_{e_i,e_j}$ corresponds to the salience scores of the common characteristics of $e_i$ and $e_j$, $A'$ corresponds to the adjacency matrix of the unweighted graph equivalent of $G$, and $p_{e_i,e_j}$ is the length of the shortest path between $e_i$ and $e_j$ in $G$. The first part of the distance is weighed more making the role of knowledge extracted using *Vichakshana* more pronounced in relating the entities. The role of the other two parts of the distance is often limited to further sort the list of related entities obtained using the first part.

# 6 Evaluation

## 6.1 Methodology

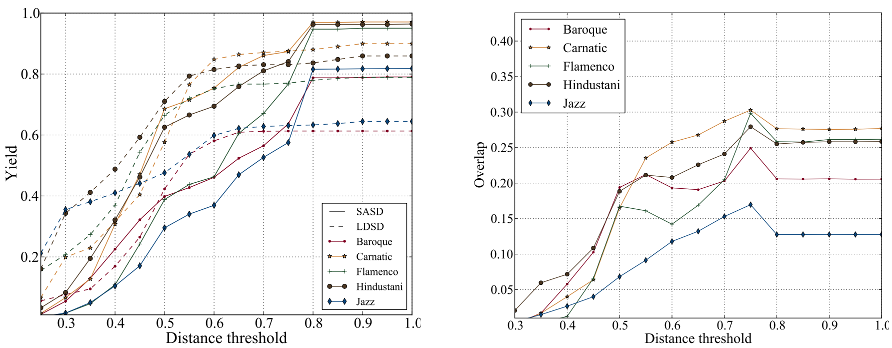
Our evaluation methodology is primarily intended to streamline the two stages of the objective and the subjective forms of evaluations for comparing the recommendation systems. The first stage corresponds to an objective comparison of the results over three measures: yield ($Y$), overlap ($O$) and rank-correlation ($RC$) within the overlap. The former two measures are defined as follows:

$$
\begin{aligned}
Y^I &= \frac{\left|\{e_k \mid \left|R^I_{e_k}\right| > 0\}\right|}{|E|} \\
O_{e_k} &= \frac{|R^I_{e_k} \cap R^J_{e_k}|}{\max(|R^I_{e_k}|, |R^J_{e_k}|)}
\end{aligned}
\tag{6}
$$

where $R^I_{e_k}$ denotes an ordered-list of recommendations for a given entity $e_k$ generated using an approach $I$. Therefore, $Y^I$ is the proportion of entities which have non-empty set of recommendations using approach $I$. $O_{e_k}$ is the proportion of the common set of entities in $R^I_{e_k}$ and $R^J_{e_k}$. For measuring rank-correlation, we use Kendall's Tau, which is preferred over other standard alternatives as it is known to be robust for smaller sample sizes. We use the Tau-b variant which accounts for tied pairs [14].

The set $E$ is divided into three different sets based on our analysis in this stage. The first one ($E_1$) is the set of entities for which $O \geq \frac{1}{3}$ and $RC$ is greater than the median of all values. This set corresponds to those entities where both approaches broadly agree with each other. The second set ($E_2$) consists of those entities where $O \geq \frac{1}{3}$ and $RC$ is less than the median of all values. The last set of entities ($E_3$) is where $O < \frac{1}{3}$.

In the second stage, which is a subjective form of evaluation, the music experts (mostly practicing musicians) record their feedback for questionnaires based on the latter two sets of entities. The one based on $E_2$ has, for each query entity, two rank-ordered lists with exactly the same set of entities (i.e., the overlapping portion). The experts are asked to pick the better one. The motive behind this is to understand whether one system is preferable to the other in ranking the entities. The questionnaire based on $E_3$ also has two lists for each query entity, but this time without an emphasis on the order of the items. The experts are asked to pick the entities in each list that are the most relevant to the query entity, and also the overall better list. Evidently, the motive here is to understand which of the approaches produces more appropriate recommendations. An analysis of their opinions would let us know whether a particular approach is preferred over the other, and can further be used to investigate why.




(a) The yield ($Y$) for both the systems shown for different music styles.

(b) The thick-lines show the mean overlap between $R^{sasd}$ and $R^{ldsd}$ for different music styles. The dotted lines shows the standard deviation.

**Fig. 1.** Results for the analysis of overlap between the two recommendation systems. X-axis in both the figures denote the distance threshold beyond which two entities are considered unrelated.

### 6.2 Results & discussion

In order to ensure that the system we compare our results with, has access to the same data source for a fair comparison, we chose DBrec system [4], which is based on DBpedia[7]. As it is shown to perform comparably well with other context-based approaches

[7] DBpedia collects the structured content from Wikipedia.

that build on diverse sources of data, it helps us to put the results of our system in perspective with both the linked-data based and the other context-based systems. However, note that our system uses only the salience scores and the entity references, but not the structured data from Wikipedia.

For all the experiments hence forth, the size of the recommendations corresponds to ten[8]. Fig. 1(a) shows $Y^{sasd}$ and $Y^{ldsd}$, and fig. 1(b) shows the mean overlap between $R^{sasd}$ and $R^{ldsd}$ for different distance thresholds. $Y^{ldsd}$ steeply rises until 0.6 and saturates, indicating that the practical limit for LDSD between two entities is 0.6, where as it is 0.8 for SASD. In line with this, the mean overlap in fig. 1(b) rises until a distance threshold of 0.75, where the overlap for all the music styles between the two systems is the maximum. Following that, it slightly drops, which must be a consequence of the gain in $Y^{sasd}$ compared to $Y^{ldsd}$ as shown in fig. 1(a). We can deduce that there is a sizable overlap between the recommendations of the two systems. However, which of the non-overlapping recommendations are more meaningful is an issue we will have to address in the subjective evaluation.

| **Mus** | **Pos** | **Mean** | **Std** | **Neg** | **Mean** | **Std** |
|---|---|---|---|---|---|---|
| Baroque | 59% | 0.47 | 0.25 | 41% | -0.32 | 0.3 |
| Carnatic | 56% | 0.47 | 0.26 | 44% | -0.29 | 0.26 |
| Flamenco | 69% | 0.41 | 0.22 | 31% | -0.2 | 0.29 |
| Hindustani | 65% | 0.49 | 0.26 | 35% | -0.23 | 0.24 |
| Jazz | 55% | 0.47 | 0.25 | 45% | -0.27 | 0.28 |
| **Avg** | 60% | 0.46 | 0.24 | 39% | -0.26 | 0.27 |

**Table 4.** Results for rank-correlation between the two approaches, showing the % of entities with positive and negative rank-correlation, along with their mean and standard deviation.

The rank-correlation ($RC$) is analyzed between those $R^{sasd}_{e_k}$ and $R^{ldsd}_{e_k}$ which have an overlap of 0.3, with a distance threshold of 0.75 (which is roughly the same as the configuration corresponding to the highest overlap from fig. 1(b)). Note that $RC$ ranges from -1 (complete disagreement) to 1 (perfect agreement). We consider a value of zero to be a negative correlation. Table. 4 shows the results. We observe consistently more positive correlations across all styles of music. Further, the mean of the positive correlations indicates a strong agreement between the recommendation systems. Based on the analysis so far, we divide the entities into three sets as discussed in sec. 6. Table. 5 shows the proportions of the three sets for different music styles.

For the subjective evaluation with music experts to further understand $E_2$ and $E_3$, we randomly sampled 20 entities from Carnatic music[9] ensuring that there is equal coverage of popular and less popular entities, as well as $E_2$ and $E_3$. Note that these entities

[8] Results for other sizes of the recommendations show a similar behavior. For the sake of brevity, we skip the rest.

[9] Carnatic music was chosen for the subjective evaluation as the authors have better access to its community compared to other music styles.

| Music | $E_1$ | $E_2$ | $E_3$ |
|---|---|---|---|
| Baroque | 17% | 12% | 31% |
| Carnatic | 27% | 21% | 39% |
| Flamenco | 31% | 14% | 29% |
| Hindustani | 28% | 14% | 39% |
| Jazz | 11% | 8% | 34% |

**Table 5.** Proportions of the $E_1$, $E_2$ and $E_3$ across all the music styles.

comprise of not just artists and songs, but any musical entity (eg: a place). The measure of popularity of an entity is its PageRank value in the graph $G$.

A total of 424 responses were recorded from 10 Carnatic music experts[10], all of whom are practicing musicians with a median age of 25. Table. 6 shows the aggregate results for questionnaires based on $E_2$ and $E_3$. The overall results do not seem to indicate a strong preference to one system or the other. However, it is evident that the responses concerning different entities are very divided. There are certain interesting observations from the responses over $E_3$. The number of cases in $E_3$ where DBrec system is preferred seems slightly higher. Yet, the number of cases where more entities are specifically marked relevant to the given query entity is higher for $SASD$.

In order to further understand this phenomenon, we have gone through the recommendations from the two systems and the responses recorded for each entity. Consider the case of *Kshetrayya*, a composer in $E_3$. The list of recommendations using $SASD$ is dominated by other composers sharing some characteristics (like geographic location, language etc). Those from DBrec system ranks the performers who often sing his compositions higher than the fellow composers. This resulted in more experts preferring DBrec system. However, the number of recommendations explicitly marked as relevant are marginally higher for $SASD$. Another example is the case of *M. Balamuralikrishna*, a performer. The recommendations from $SASD$ do not include his teacher whereas those from DBrec system do. This is a result of a low recall in entity linking using DBpedia Spotlight. While the ratio of experts preferring DBrec system to SASD is 7:2, the corresponding ratio of absolute number of entities selected as relevant is 7:6. There are several other cases like these. This trend clearly indicates that though our system missed few important relations between entities (such as those between different types of entities), the recommendations made are still very relevant, often times even more than the recommendations from DBrec system.

More formally, for a given query entity, $SASD$ is inherently biased to select and rank higher other entities which are of the same type (Eg: Composer). It is also highly sensitive to recall in entity linking. On the other hand, DBrec system has access to a richer link structure that spans different entity types.

When we view the results again in the light of this stark difference in the nature of information both the systems had access to, the results seem to clearly indicate that the information extracted using *Vichakshana* is meaningful in itself. Further, it is comple-

[10] By experts, we mean those who are thoroughly well-versed with the domain and the surrounding activity.

mentary to the existing structured content on Wikipedia which DBrec depends on, and is useful for improving the music recommendations, both in terms of better ranking and more importantly, finding appropriate content.

| | $E_2$ | | | | $E_3$ | | | |
|---|---|---|---|---|---|---|---|---|
| | SA. | LD. | Both | None | SA. | LD. | Both | None |
| Overall preference | 30% | 30% | 10% | 30% | 40% | 50% | 10% | 0% |
| Entities specifically marked as relevant | n/a | n/a | n/a | n/a | 50% | 30% | 20% | 0% |

**Table 6.** Results of the subjective evaluation of the two recommendation systems. The first row of results show the % of query entities where a particular recommender system is more favored. The second row shows the % of query entities where more number of entities in the corresponding recommendation list are marked as specifically relevant to the query entity.

## 7 Conclusions

We have presented and formally defined the idea of quantifying the salience of characteristics of a music, and how it leads us to extracting culture-specific information about a music using text documents. We have shown that the performance of a recommendation system built using the information extracted is comparable to that of the linked-data based recommendations. The main contributions of the paper are as follows:

- A novel approach that quantifies the salience of characteristics of a music.
- A salience-aware semantic distance that builds upon the knowledge extracted.
- An evaluation methodology that allows for a streamlined objective and subjective comparison of recommendation systems.
- Open parallelized python implementations of *Vichakshana* and SASD[11] for reuse by research community.

Our work in progress and the future plans include:

- Evaluation on other music styles.
- Scaling the approach to work with web data.
- Publishing and integrating the results with the linked open data.
- Adapting it to other thematic domains such as movies and games.
- Conducting large-scale user (non-expert) evaluation.

## 8 Acknowledgements

We are thankful to Dr. Mohamed Sordo for his early inputs, Vignesh Ishwar and the Carnatic music community for their help in evaluation. This research was partly funded

[11] Available at https://github.com/gopalkoduri/vichakshana-public

by the European Research Council under the European Union's Seventh Framework Program, as part of the CompMusic project (ERC grant agreement 267583).

## Bibliography

[1] Serra, X.: A Multicultural Approach in Music Information Research. In: ISMIR. (2011) 151--156
[2] Sarwar, B., Karypis, G., Konstan, J., Riedl, J.: Item-based collaborative filtering recommendation algorithms. In: World Wide Web, ACM (2001) 285--295
[3] Knees, P., Schedl, M.: A Survey of Music Similarity and Recommendation from Music Context Data. ACM Trans. Multimedia Comput. Commun. Appl. **10**(1) (December 2013) 2:1--2:21
[4] Passant, A., Decker, S.: Hey! ho! let's go! explanatory music recommendations with dbrec. The Semantic Web: Research and Applications **1380**(2) (2010) 411--415
[5] Bizer, C., Heath, T., Berners-Lee, T.: Linked data-the story so far. International Journal on Semantic Web and Information Systems **5**(3) (2009) 1--22
[6] Raimond, Y.: A Distributed Music Information System. PhD thesis, University of London (2008)
[7] Celma, O.: Foafing the Music Bridging the semantic gap in music recommendation. In: International Semantic Web Conference, Athens, GA, USA (2006) 927--934
[8] Jacobson, K., Sandler, M., Fields, B.: Using Audio Analysis and Network Structure to Identify Communities in On-Line Social Networks of Artists. In: ISMIR. (2008) 269--274
[9] Petasis, G., Karkaletsis, V.: Ontology population and enrichment: State of the art. In: Knowledge-driven multimedia information extraction and ontology evolution. Springer-Verlag (2011) 134--166
[10] Daiber, J., Jakob, M., Hokamp, C., Mendes, P.: Improving efficiency and accuracy in multilingual entity extraction. In: International Conference on Semantic Systems. (2013) 3--6
[11] Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank citation ranking: bringing order to the web. Technical report, Stanford InfoLab (1999)
[12] Newman, M.: Networks. Oxford University Press (March 2010)
[13] Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment **2008**(10) (October 2008) P10008
[14] Kendall, M.G., Gibbons, J.D.: Rank Correlation Methods. 5th edn. E. Arnold (1990)