

Latent Dirichlet Allocation Model for Raga Identification of Carnatic Music

Rajeswari Sridhar, Manasa Subramanian
B.M. Lavanya, B. Malinidevi and T.V. Geetha
Department of Computer Science and Engineering, Anna University,
Faculty of Information and Communication Engineering, Chennai, India

Abstract: Problem statement: In this study the Raga of South Indian Carnatic music is determined by constructing a model. Raga is a pre-determined arrangement of notes, which is characterized by an Arohana and Avarohana, which is the ascending and descending arrangement of notes and Raga lakshana. **Approach:** In this study a Latent Dirichlet Allocation (LDA) model is constructed to identify the Raga of South Indian Carnatic music. LDA is an unsupervised statistical approach which is being used for document classification to determine the underlying topics in a given document. The construction of LDA is based on the assumption that the notes of a given music piece can be mapped to the words in a topic and the topics in a document can be mapped to the Raga. The identification of notes is very difficult due to the narrow range of frequency and the characteristics of Carnatic music. This inclined us in moving to a probabilistic approach for the identification of Raga. In this study we identify the notes of a given signal and using these notes and Raga lakshana, a probabilistic model in terms of LDA's parameters α and θ are computed and constructed for every Raga by initially assuming a value which is constant for every Raga. This value of α is cultured after determining θ for a given Raga. The θ of a given Raga is computed using the characteristic phrases which is a sequence of notes and is unique for a given Raga. During the Raga identification phase, the value of α and θ are computed and is matched with the constructed LDA model to identify the given Raga. **Results:** Using this model, the Raga identification of Parent Ragas had a lower error rate than that of Child Raga. For parent Raga an average identification rate of 75% was achieved. **Conclusion/Recommendations:** The accuracy of the algorithm can be improved by using more features of Raga lakshana. After identifying the Raga, it can be used as features to be used by a Music Information Retrieval system.

Key words: Fundamental frequency, raga lakshana, raga identification, retrieval system, given raga, carnatic music, input signal, narrow range, descending arrangement

INTRODUCTION

Carnatic music is one of India's traditional systems of music. This system of music differs very much from the Western music (Vijaykrishnan, 2007). The basic differences include even tempered system of music against just tempered system of music, presence and absence of chords, harmony, polyphony against monophony, concept of a pre-defined scale against Raga. In the context of Indian music, North Indian Hindustani music and South Indian music differ in the style of rendering, the depth of pitch and some similarities also exists like Gamakas against meend, hierarchical Raga relationship against time based Raga relationship, etc. In both the systems of music, the main factor governing the music is the Raga which is defined

as the ascending and descending arrangement of notes (Sridhar and Geetha, 2009). A Raga in Carnatic music can be grouped into Parent Raga one that has all the seven notes S, R, G, M, P, D, N which is analogous to C, D, E, F, G, A, B and child Raga which is derived from Parent Raga by removing one or two of the seven swaras either in the ascending, descending or both. In addition to the notes, every Raga is characterized by additional information in terms of Raga lakshana which is the ornamentation that is given to the notes of every Raga. In this study, one of the Raga lakshana characteristics is used for constructing the LDA model which is used for Raga identification. This study is organized as follows. The discussion starts with the basic work that is being done in LDA and Raga identification and is followed by introducing the Raga

Corresponding Author: Rajeswari Sridhar, Department of Computer Science and Engineering, Anna University, 600025, India

lakshana characteristics. We then introduce the algorithm that we have proposed followed by a discussion on the Results. Finally we conclude this study discussing on the future work.

Existing work on raga identification and LDA: Raga identification of Hindustani music has been proposed by many researchers (Belle *et al.*, 2009; Chordia and Rae, 2007; Chordia, 2006). In the study done by Belle *et al.* (2009) Hindustani music Raga identification was performed by using the intonation given to individual swaras. The authors have used features at the swara level, which are extracted from the signal as the peak value of a swara, its mean value, the standard deviation of a swara and distribution of a swara for identifying the swara thereby determining the Raga. The drawback of the system is that the mean value that is computed could not reach a steady state value thereby making it difficult for Raga identification. In addition the study was carried out to determine Ragas that has the same swaras in the Arohana and Avarohana.

In the study done by Chordia and Rae (2007) and Chordia (2006), the authors have performed Raga Classification of Hindustani Raga using Pitch Class and Pitch Class Dyad distributions. In this study the authors have divided the input signal into segments and determined the pitch by using the Harmonic Product Spectrum algorithm (Cuadra *et al.*, 2001). The authors have determined the onset of the input signal to estimate the frequency component at the place of onset. The input is converted to MIDI representation to determine the pitch. Then using the detected pitch, the Pitch Class distribution and Pitch Class Dyad distribution was estimated by determining histogram of the pitch contour. The Pitch Class Distribution and Pitch class Dyad Distribution are used to determine the Raga.

The (Chordia, 2006) have tried the same algorithm which was developed for Hindustani music to Carnatic music. The conversion to MIDI representation led to the loss of Gamaka information and hence the error rate of identification was high. In addition, the identification of notes is very difficult due to the narrow range of frequency and the characteristics of Carnatic music. In Carnatic music every Raga is governed by a Raga lakshana in addition to the Arohana and Avarohana. This Raga lakshana is a non-deterministic value and therefore requires a probabilistic model.

Many such probabilistic models exist for analysing the content of signal, image, information and also in membrane computing (Zhijun and Minghong, 2005). Some of the used probabilistic models include HMM, N-gram model or LDA models (Hu, 2009). After careful consideration and analysing the suitability of the probabilistic model, we moved to LDA model for Raga identification.

MATERIALS AND METHODS

LDA is a document classification algorithm and its use for music has been very minimal. In the study proposed by Hu (2009), the author has analyzed the performance of LDA for text, images and music. The author claims that every document will cater to a set of topics. Every topic will have a set of words specific to it. Hence, these words that describe a topic can be assigned a higher probability. Using the probability distribution in a given document, the number of topics covered by that document can be understood. The authors have used Dirichlet distribution for the distribution of words in a given topic and this distribution is governed by the Dirichlet parameters ' α ' and ' θ '. The parameter α is a K-dimensional parameter that is constant over all documents within a corpus. The parameter θ is the topic weight vector, indicating the amount of each of the K topics in a given document. The initial α is assumed to be a smaller value and this value is re-computed using the value of θ using Baye's theorem. By using the values of α and θ , the distribution in a given document is computed thereby determining the topics in a given document. The author has also explored the determination of these values for images and music.

In the study for images, the image is segmented and each segment is viewed as a bag of words. For each of these segment vectors the Dirichlet parameters are identified to determine the topics in the given image. In the study for music, the author has used LDA for determining the harmonic structure available in a given music piece. The K-dimensional parameter α is computed by assuming the value of 'K' corresponding to the number of notes in major key and minor key which is typically 24. Then the authors have matched document corpus to music corpus, music in the corpus to a document consisting of a distribution of notes. This inspired us to try this study for Carnatic music analysis which has a pre-defined melody called Raga.

In another work done by Bello (2008), the authors have used a modified LDA for determining musical similarity between music files. The authors have modeled the dirichlet parameters using the MFCC feature set which typically conveys the timbral characteristics of a given music.

Raga lakshana characteristics: A Raga lakshana has 13 essential features in addition to the Arohana and Avarohana, as described in literature (Vijaykrishnan, 2007) and comprises of the following:

- Graha: Note at which a raga commences
- Amsa: The note that reveals the melodic entity of the raga-or svarupa-or jiva swara
- Nyasa: The note on which the raga can be concluded
- Mandra: The lowest note that can be played in the raga
- Tara: The highest note that can be played in the raga
- Alpatva: The note used sparingly in the raga
- Bahutva: The note used frequently in the raga
- Apanyasa, same sangati is sung in tara and madhyasthayi, vinayasa, raga sancharas are stopped at a swara - then elaborated in Mandra and tara sthayi
- Sanyasa: Raga is sung and elaborated and finally closed at the adhara shadja swara-s
- Shadava: 6 note sancharas
- Audava: 5 note sancharas
- Antara Marga: Introduction of note or chayya of another Raga

In this study the second Raga lakshana characteristic Amsa is used for training the LDA as it is one characteristic that is more essential in the Raga lakshana characteristics. In addition, the Amsa characteristics can be mapped with LDA better than other characteristics. In LDA, documents are classified into topics based on the frequency of usage of words. Similarly, the characteristic phrase can be mapped to the words of the document and hence this characteristic is used for training the LDA.

Our algorithm for raga identification: In our algorithm, for construction of LDA for Raga identification, the component that is analogous to a word in a document is the characteristic phrase of the Raga. The K-dimensional vector as constructed by the (Hu, 2009) is completely changed here in not using all the notes for determination of dirichlet parameters α and θ but uses a sequence of notes for the determination of these parameters.

The algorithm can be grouped into an online phase and an offline phase. In the online phase the LDA is constructed by initially identifying the characteristic phrase and in the offline phase this characteristic phrase is determined which is compared with the LDA for the identification of Raga. This characteristic phrase consists of a sequence of swaras. Hence to determine the characteristic phrase, it is mandatory to identify the swaras available in the song.

Table 1: Swara and their ratio with S

Swara	Ratio	Swara	Ratio
S	1	M2	27/20
R1	32/31	P	3/2
R2	16/15	D1	128/81
R3	10/9	D2	8/5
G1	32/27	D3	5/3
G2	6/5	N1	16/9
G3	5/4	N2	9/5
M1	4/3	N3	15/8

Swara identification: We use our already proposed segmentation, fundamental frequency estimation and swara identification algorithm to determine the swaras available in the input song. To start with we estimate the fundamental frequency of the input song. The fundamental frequency in a Carnatic music song corresponds to the middle octave Shadja note – S. This frequency is essential since all other notes depend on this note’s frequency. Therefore other fundamental frequency algorithms cannot be used for identifying this middle octave ‘S’. This frequency of the ‘S’ is not constant and varies between singers and between songs. The algorithm is a mutation based technique where a mutating signal is used to mutate the input signal at three positions to determine this frequency of ‘S’. After identifying the fundamental frequency, the input signal is segmented using our segmentation algorithm which is based on the Talam characteristics (Sridhar and Geetha, 2009). A Talam is used along with the song to ensure that the tempo of the song is maintained properly. Therefore each Talam component corresponds to 1, 2, 4, or 8 swaras. This is used as the component for segmentation (Norozi *et al.*, 2010). After segmenting using Talam based segmentation and reducing over segmentation, the understanding is that each segment corresponds to a swara. Therefore we determine the frequency component that is dominant at every segment, by performing Fast Fourier Transform of each segment. Using this frequency component and the fundamental frequency a ratio is computed. This ratio is mapped with the ratio in Table 1 to determine the swara corresponding to every segment. After determining the swara in every segment the sequence of swaras comprising the input song is analyzed, compared with literature and the characteristic phrase is determined. There is more than one characteristic phrase for one Raga. In this situation all the phrases are identified. This is used for the construction of LDA.

LDA construction: To construct the LDA, in our work, we explore the characteristic phrase of notes that is unique for every Raga. This characteristic phrase is mapped onto the heavy weight that LDA gives to the

words that are common for a given topic. As already indicated LDA requires the use of two parameters α and θ , therefore this is our parameter θ which estimates the weight associated with a sequence of notes for a given Raga. The initial value of α is assumed to be uniform and therefore it is computed with all possible combinations and all the phrases are given equal weight. Using the value of θ the value of α is recomputed to determine the actual value for α .

In the training phase the value of the Dirichlet parameters α and θ are estimated and stored which are used for comparison in the testing phase to determine the Raga. Both α and θ are weight vector of probabilities.

α is the generic distribution of patterns in all Ragas. θ is the distribution of swara pattern in one Raga. In our algorithm, during the training phase, we assume initial probability value as being equally distributed for all patterns in α . This value is recomputed by studying the frequency of occurrence of swara patterns in the songs from the training set. Similarly we initialize and re-compute θ by considering songs belonging to a particular raga.

Using the re-computed θ value, the value of α is modified using Baye's theorem. The process of θ computation for other ragas is also performed and the corresponding θ vectors are determined.

In our algorithm the estimated value of α is maintained for a pattern length. If pattern length is 4, all songs with the identifying pattern length as 4 will have one common α and every raga has a θ vector for itself. Using a permutation algorithm, we generate all possible combinations of 4 length patterns that include all 7 swaras. In this study we have not considered the variations that are available within a swara and have considered only its single occurrence. Therefore we have only 7 swaras S, R, G, M, P, D, N as against 17 swaras or 22 intervals (Vijayakrishnan, 2007). We initialize equal probabilities to these patterns with a value as $1/7*7*7*7$.

After initializing the probability value for α , we need to train α with a newer value using songs belonging to all Ragas. In the training process, for every 4 length pattern we encounter in the song, a little weight is added in α for that corresponding pattern. After re-computing the value of α using this process by using a training corpus, the value of θ is determined from α for a Raga.

To construct θ , the system needs to be trained with the songs of one particular Raga. Given a song, for every four length pattern encountered, the weight of that pattern is retrieved from α vector and a little more weight is added and the pattern gets stored in θ

vector. Using this procedure nearly the top 20 patterns for a given Raga is found using training set and is stored as the θ vector. Once again this θ vector is used to refine α vector which is done by adding a small weight to the patterns of a vector that are encountered in the Θ vector. After α gets re-computed for each Raga, the same procedure is repeated to determine these vectors for all Ragas. The pseudo-code of the algorithm is given below in:

```
LDA Construct ()
{
    Determine 4 length pattern combinations and
    assign equal probability
    For every Raga
    {
        Compute  $\alpha$  by choosing songs belonging to all
        Ragas by assigning a little weight if the 4 length
        pattern is encountered in the song
        Compute  $\theta$  by choosing song belonging to one
        Raga and if the 4 length pattern occurs add a little
        weight by choosing from  $\alpha$ 
        Re-compute  $\alpha$  using the computed  $\theta$  vector.
    }
}
```

Raga identification using LDA: In the online testing phase, the input song is given and using α vector, the θ of the song is determined. This θ is compared with θ s of all ragas and for which Raga this θ of the input song is closer to that of the available θ that Raga is chosen as the Raga of the input song. Closeness is determined by the relative positions between the top 10 patterns in the θ 's of all Ragas by computing Euclidean distance.

RESULTS AND DISCUSSION

In our approach for determining the initial value of α , in this algorithm we have used only 4 length phrases of swaras. The challenge was in determining the unique phrase of notes for each Raga. In general a Raga is characterized by two or three of this characteristic phrases, which occur frequently in addition to other characteristics as explained in the Raga lakshana. This was determined from the literature of Carnatic music (Vijayakrishnan, 2007) and also by computing this manually by observing the swara representation of Carnatic songs for each Raga.

The input signal is sampled and the LDA is constructed as explained. Songs sung by Ms. Nithyasree Mahadevan, Ms. Sowmya, Dr. M.S. Subbulakshmi, Dr. M. Balamuralikrishna, belonging to both Parent Raga and Child Ragas are chosen to test the system for efficiency.

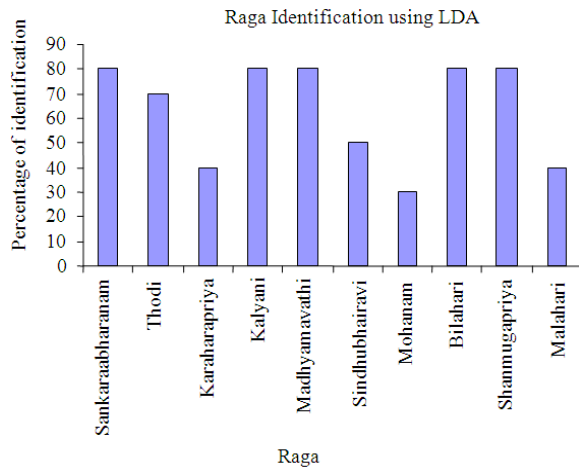


Fig. 1: Raga identification using LDA

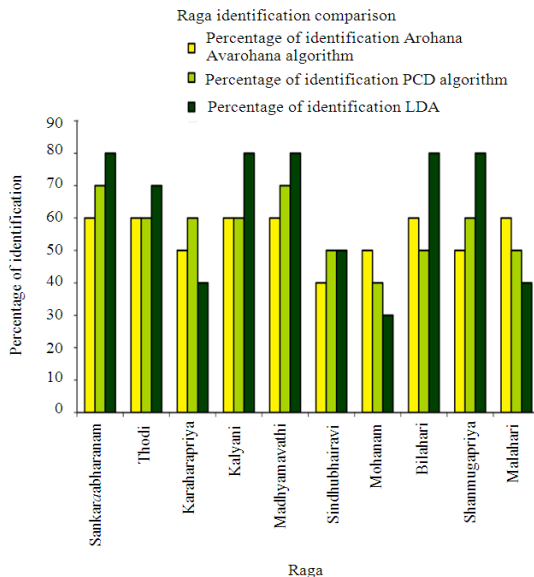


Fig. 2: Comparison of raga identification by different algorithms

The result of the algorithm is given in Fig. 1. From the figure it is observed that the performance of Melakarta Ragas is higher compared to that of Child Ragas. In addition to this, if the characteristics swara phrase of the Ragas is of length 4, the performance is better than that of Ragas whose characteristic phrase is of a different length. In our algorithm to compute a we assumed a 4 length combination of all the 7 swaras. Since all 7 swaras are used for the computation, the computation of a is for the Parent Ragas as child Ragas will not have all the 7 swaras.

However, it is to be observed that the identification rate of Karaharapiya despite being a Parent Raga is low since the characteristic phrase is not of length 4. A comparison between our earlier algorithm for Raga identification, the PCD algorithm and the LDA algorithm is performed and the results are tabulated in Fig. 2.

As can be seen from the figure the LDA based algorithm performed better in the situation where the characteristic phrase was of length 4. In all other situations other algorithms performed better than LDA.

Taking into consideration both Parent Raga and Child Raga, the average identification rate of LDA is 63%, which is comparable to the rate of 57% by the PCD algorithm and 55% by the Arohana Avarohana algorithm.

CONCLUSION

In this study, an algorithm that uses LDA to identify the Raga of a given song is discussed and implemented. This study is based on using a 4-length pattern for the construction of LDA to derive the parameters ∞ and θ . Considering a swara length of 3 or 5 more number of patterns can be generated to determine the initial probability of ∞ . In addition to this, the concept of Child Ragas can be considered by deriving the patterns by omitting one or two swaras. In our algorithm we have not differentiated the variations in the swaras R, G, M, D, N for considering the initial pattern length. This could also be incorporated for determining the swara length pattern, which could be used for identifying the parameters ∞ and θ to increase the efficiency of the system. After identification, the Raga could serve as entities for use in a Music Information Retrieval system (Norozi *et al.*, 2010).

REFERENCES

- Belle, S., R. Joshi and P. Rao, 2009. Raga identification by using swara intonation. J. ITC Sangeet Res. Acad., 23: 1-7.
- Bello, J.P., 2008. Content based music similarity computation using the hierarchical dirichlet process. Proceedings of the 9th International Conference of Music Information Retrieval (ISMIR'08), Drexel University, Philadelphia, pp: 692-692. ISBN: 0615248497
- Chordia, P. and A. Rae, 2007. Raag Recognition using Pitch-class and Pitch-class dyad distributions.
- Chordia, P., 2006. Automatic raag classification of pitch-tracked performances using pitch-class and pitch-class dyad distributions.

- Cuadra, P.D.L., A. Master and C. Sapp, 2001. Efficient pitch detection techniques for interactive music.
- Hu, D.J., 2009. Latent Dirichlet allocation for Text, Images and Music.
- Norozi, A., M.K.A. Ariffin and N. Ismail, 2010. Application of intelligence based genetic algorithm for job sequencing problem on parallel mixed-model assembly line. *Am. J. Eng. Applied Sci.*, 3: 15-24. DOI: 10.3844/ajeassp.2010.15.24
- Sridhar, R. and T.V. Geetha, 2009. Raga identification of carnatic music for music information retrieval. *Int. J. Recent Trends Eng.*, 1: 571-574.
- Vijaykrishnan, K.G., 2007. *The Grammar of Carnatic Music (Phonology and Phonetics)*. 3rd Edn., Mouton De Gruyter Publisher, ISBN-10: 3110183137, pp: 342.
- Zhijun, L.I. and L. Minghong, 2005. EOS: Evolutionary overlay service in peer-to-peer systems. *Am. J. Applied Sci.*, 2: 1401-1406. DOI: 10.3844/ajassp.2005.1401.1406