Proceedings of the Thirtieth International Florida Artificial Intelligence Research Society Conference

Using Association Rules Mining for Retrieving Genre-Specific Music Files

Louis Rompré,[1,2] Ismaïl Biskri,[2] Jean-Guy Meunier[1]

[1] LANCI – Université du Québec à Montréal, C.P. 8888, Succ. Centre-Ville, Montréal (Qc), H3C 3P8, Canada

[2] LAMIA – Université du Québec à Trois-Rivières, C.P. 500, Trois-Rivières (Qc), G9A 5H7, Canada

rompre.louis@courrier.uqam.ca, ismail.biskri@uqtr.ca, meunier.jean-guy@uqam.ca

Abstract

Retrieving a music file from a large database is a non-trivial task. To support this task, many mechanisms have been developed over the years. However, indexing files remains one of the most popular mechanisms. Several algorithms allow feature extraction from audio signals. Usually, these features are used to describe music content. In this paper, we demonstrate that associations between content-based descriptors can be used as well. We have developed a processing chain which uses association rules mining to find significant relations between content-based descriptors of music files. The significant relations are used to index music files. Experiments conducted demonstrates that the proposed approach can yield interesting results especially with classical music.

Introduction

In recent decades, the digital medium has outpaced the physical medium for storing music. This has led to the creation of large databases. Some of these often contain several millions music files. Considering all available data, searching for specific music files can become a tedious and complex task which may require a considerable amount of effort. Therefore, some titles may remain unexplored leaving potentially interesting files overlooked. To overcome this problem, many mechanisms have been proposed. Among all these mechanisms, indexing files is one of the most employed. Indexes can take different forms. For instance, textual metadata can be used to provide a compact and human readable description of the music content. The entire content of the file needs to be summarized using only a few words. Most of the time, the music genre is used to describe the music content. Non-human readable indexes also exist. Audio fingerprint is an example of this kind of index. An audio fingerprint is a unique content-based signature. It's usually generated following a frequency spectrum analysis of audio signal. When an unknown music file is used as query, its acoustic fingerprint is computed and matched against those stored. Audio fingerprint has allowed the creation of robust recognition tools such as Shazam (Wang, 2003). However, audio fingerprint has its limits because it is not designed to deal with different versions of the same song nor similar songs (Grosche et al., 2012).

We suggest to use association rules mining to find significant relations between content-based descriptors of musical files and use these relations to create indexes. Association rules mining have the capability to extract hidden knowledge in a large data sets. Files data usually contains regular patterns. When these patterns are analyzed, some unexpected properties of data can be discovered. The significant patterns which translate unexpected properties of data are considerate as hidden knowledge. Same hidden knowledge can be observed in similar files. Hence, it can be used to establish relationship between files.

Computer Assisted Files Classification

Classification is a component of many music retrieval systems. It has been proposed to use classification in order to enhance music recommendation (Li et al., 2003). Classification was also applied to enhance automatic audio file annotation (Ness et al., 2008). Classification organizes audio files in such a way that similar files are grouped together. From a computational point of view, classification relies on statistical comparison of content-based descriptors. The whole classification process may be divided in two steps. The first one is the features extraction. Extracted features are used as content-based descriptors. Over the years, many libraries have been developed and shared to facilitate the features extraction (Bray & Tzanetakis, 2005; Lartillot & Toiviainen, 2007; Downie et al., 2005). The main difficulty is not extracting the features; it is choosing the right features. The second step is the classification process itself. Different classifiers such as neural network based classifier can be used to perform this step. One of the limitations of these classifiers is that they are opaque. Therefore, it's difficult to evaluate the contribution of a specific descriptor in the classification result. Both features and classifier have an impact on classifications. Different combinations of features and classifiers should continue to be explored in order to improve the computer assisted classification process.

Using Association as a Descriptor

It is common for the same descriptors to be used in order to describes miscellaneous music files. Even though the same descriptors are used, described files can vary greatly from one to another. In this context, the presence alone of descriptors is not considered as significant information. Therefore, the usefulness of such descriptors as leading parameters in music exploration is limited. However, relationships between descriptors may highlight other particularities of music and it would be appropriate to use the relationships between descriptors in order guide the exploration of audio files. Our assumption is that relations between descriptors are as important as the descriptors themselves.

Association Rules Mining

Association rules mining is a method widely used to discover unexpected relations between data. Unlike opaque methods, association rules are easy to interpret. Interest regarding this method started from the research conducted by Agrawal (Agrawal et al., 1993). Agrawal and his colleagues showed that association rules can be used to extract hidden relations between items contained in a transactional database (Agrawal and R. Srikant, 1994). However, association rules are not limited to transactional database. They can be applied to any field as long as the concept of transaction can be adapted. The following section gives the basic concepts related to association rules.

Let $I = \{i_1, i_2, i_3, \ldots, i_d\}$ be a set of d distinct items then a subset of items from I is called an itemset. It is possible to generate 2^d itemsets from d items. Let $T = \{t_1, t_2, t_3, \ldots, t_n\}$ be a set of n transactions containing items from I, then the support count of an itemset X represents the number of transaction that contain X. The support count is given by Eqn(1).

$$\sigma(\mathrm{X}) = |\{t_i \mid \mathrm{X} \subseteq t_i, t_i \in T\}| \qquad 1$$

The support is a measure similar to support count but the overall number of transactions is considered. An itemset whose support is greater than a minimum threshold is called frequent itemset. The support is given by Eqn (2).

$$\mathrm{S}(\mathrm{X}) = \frac{\sigma(\mathrm{X})}{n} \qquad 2$$

Let X and Y be two itemsets such as $\mathrm{X} \cap \mathrm{Y} = \emptyset$, an association rule is expressed as X→ Y. It represents the co-occurrence regularities between itemsets X and Y. X is called the antecedent and Y is called the consequent. The quality of an association rule is determined by a measure *m* and a threshold. Thus, an association rule X→ Y is considered relevant if $m(\mathrm{X} \rightarrow \mathrm{Y}) \geq$ threshold. The amount of generated rules, their relevance and their utility dependent of both measure and threshold. Evaluation of measures has been the subject of several works (Geng & Hamilton, 2006; Lenca et al., 2008). Even if many others measures exist, the support and the confidence remain the most common measures.

The support of an association rule is the number of transaction that contain both antecedent and consequent. The support is given by Eqn (3). The rules whose support is low may appear only by chance. For this, these rules can be discarded.

$$\mathrm{S}(\mathrm{X} \rightarrow \mathrm{Y}) = \frac{\sigma(\mathrm{X} \cup \mathrm{Y})}{n} \qquad 3$$

The confidence of a rule is the number of transactions that contain both antecedent and consequent among transactions that contain antecedent. It is a precision measure. The confidence is given by Eqn (4).

$$C(\mathrm{X} \rightarrow \mathrm{Y}) = \frac{\sigma(\mathrm{X} \cup \mathrm{Y})}{\sigma(\mathrm{X})} \qquad 4$$

Extraction of Association Rules

Extraction of association rules is divided into two steps which are the detection of frequent itemsets and the generation of rules. Despite their potential, association rules are sensible to combinatorial explosion. The number of possible rules is exponential to the number of items. This number can be estimated using Eqn (5) where *d* is the number of distinct items.

$$R = 3^d - 2^{d+1} + 1 \qquad 5$$

Apriori algorithm can be used to extract association rules (Agrawal & Srikant, 1994). This algorithm uses support and confidence to prune irrelevant associations. However, controlling the number of items to deal with remains the most efficient means to avoid combinatorial explosion.

Association Rules Mining and Music Retrieval

Work has been done on how best to judge the relevance of association rules as well as the quality of their interpretation (Lallich & Teytaud, 2004). Integration of association rules into information retrieval systems (Biskri & Rompré, 2012; Diop & Lo, 2007) and into classification process for text mining (Cherfi & Napoli, 2005) were also studied. Recently, the integration of association rules into music information retrieval system was proposed. They have been used to capture discriminant patterns in symbolic files (Conklin, 2009). They have also been applied to predict the genre of different audio files (Arjannikov & Zhang, 2014). Finally, association rules have been used to enhance music recommendation systems (Swathi & Reddy, 2014).

We propose to used association rule mining to find significant relations between spectral features and use these relations as index for audio files.

Methodology

We have developed a processing chain which use association rule mining to find relations between content-based descriptors. It contains two main operations: transaction creation and finding relevant association rules.

Creating Transactions

An important aspect of the work was to define transactions from audio signal. The retained process contains 5 steps. Figure 1 illustrates how transactions are created from the audio signal.

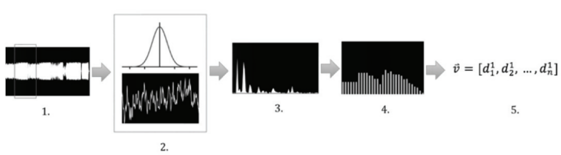

Figure 1. Proposed processing chain to extract significant associations between frequencies.

Splitting Data

The first step that leads to transactions is a segmentation step. The purpose of this step is to represent the signal evolution over time. The signal is subdivided to obtain segments of an approximate duration of 46 milliseconds. To do this, a window containing 2048 samples is slipped into the raw data with a 50% overlap. In the following steps, each segment is then transformed to transaction into transactions.

Application of Window Function

The segmentation process may introduce undesired frequencies. This phenomenon is known as spectral leakage. The second step is merely to apply a window function on each segment to reduce the risk of spectral leakage. The prototype developed allows the use of Hanning, Hamming and Blackman windows which are three window functions widely employed in digital signal processing. Similar results have been obtained using all three.

Getting the Frequency Spectrum

The third step is to obtain the frequency spectrum. Following this, Fourier transform is applied and the magnitude is computed. The output consists of a vector composed of 1024 elements. Mining association rules from a set of transactions which containing 1024 items may be problematic. Worst case scenario, it will cause a memory leak. This entails that additional operations need to be performed in order to reduce the number of items.

Grouping the Frequencies

The fourth step is to group the frequencies. This step is performed in order to limit the number of items inside the transactions and also reduce the risk of combinatorial explosion. Frequencies can be grouped into 9 or 31 spectrum bins. Using 9 bins speeds up the process while using 31 bins gives better resolution. Table 1 shows the 31 bins grouping.

Bin	Min (Hz)	Max (Hz)	Bin	Min (Hz)	Max (Hz)
1	0	21	**17**	625	796
2	22	43	**18**	797	1 012
3	44	64	**19**	1 013	1 248
4	65	86	**20**	1 249	1 614
5	87	107	**21**	1 615	2 002
6	108	129	**22**	2 003	2 497
7	130	150	**23**	2 498	3 143
8	151	172	**24**	3 144	4 005
9	173	193	**25**	4 006	4 996
10	194	214	**26**	4 996	6 309
11	215	236	**27**	6 310	8 010
12	237	258	**28**	8 011	9 991
13	259	322	**29**	9 992	12 510
14	323	409	**30**	12 511	15 999
15	410	516	**31**	16 000	22 050
16	517	624			

Table 1. 31 spectrum bins

Discarding Irrelevant Frequencies

The fifth step is also performed to reduce the risk of combinatorial explosion. Frequencies below a contribution threshold are discarded. This parameter allows removing spectrum bins that are likely to be less informative. Non discarded spectrum bins are taken into account to build the transactions. As a result, each file is represented by a set of transactions containing the most significant spectrum bins.

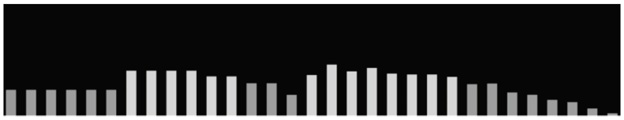

Figure 2. Discarded frequencies below the contribution threshold. The spectrum bins shown in dark gray are not included in the transaction.

Finding and Using Relevant Associations

Unlike many proposed approaches, our approach is not supervised. Therefore, a training stage is not required. It is a straightforward process. The association rules are extracted from the audio files and then used to index files in a database. Association rules mining is performed by using Apriori algorithm (Agrawal & Srikant, 1994). The extracted association rules are sorted and only the *n* top most, that is which with the highest support and confidence, are kept and stored. Finally, association rules are used as query criterion.

Experimentations

To evaluate the retrieval capability of association rules, we have conducted several experiments. The presented results have been obtained using minimal support fixed at 10% and the contribution threshold fixed at 40%. Twenty associations rules have been exploited in order to represent each file.

Dataset

The dataset used for our experimentations is GTZAN. This dataset was initially used by Tzanetakis and Cook (Tzanetakis & Cook, 2002). It consists of 1000 audio tracks. 10 genres are represented: blues, classical, country, disco, hip-hop, jazz, metal, pop, reggae and rock. Each genre is represented by 100 tracks of the same duration. Over the years, GTZAN was widely used by the MIR community (Sturm, 2012). Even if usage of GTZAN may be criticized by some authors (Sturm, 2012), it allows comparison.

Results

The association rules have been extracted for all the audio files. On average, 607 rules were extracted from the classical pieces versus 715 from the rock songs. A first evaluation was performed to determine if the same genre pieces share the same association rules. The compiled results vary significantly from one musical genre to the other. Only five percent (5%) of the rock musical pieces share a significant number of association rules. On the other hand, 90% of the classical musical pieces share a considerable amount of association rules. Table 2 shows the first 20 association rules from two rock pieces. Table 3 shows the first 20 association rules for the two classical ones.

ROCK 25			
X	Y	SUP	CONF
3144;	4006;	0,4	0,8
4006;	3144;	0,4	0,6
2498;	4006;	0,4	0,9
4006;	2498;	0,4	0,5
2003;	4006;	0,4	0,9
4006;	2003;	0,4	0,5
797;	4006;	0,4	0,8
4006;	797;	0,4	0,5
2498;	3144;	0,3	0,8
3144;	2498;	0,3	0,7
1615;	4006;	0,3	0,8
4006;	1615;	0,3	0,5
2003;	3144;	0,3	0,8
3144;	2003;	0,3	0,6
2498;3144;	4006;	0,3	0,9
2498;4006;	3144;	0,3	0,8
3144;4006;	2498;	0,3	0,7
2498;	3144;4006;	0,3	0,7
3144;	2498;4006;	0,3	0,6
4006;	2498;3144;	0,3	0,4

ROCK 65			
X	Y	SUP	CONF
1013;	797;	0,3	0,6
797;	1013;	0,3	0,6
2498;	3144;	0,3	0,7
3144;	2498;	0,3	0,6
2003;	3144;	0,3	0,7
3144;	2003;	0,3	0,6
1615;	797;	0,2	0,6
797;	1615;	0,2	0,5
2498;	2003;	0,2	0,7
2003;	2498;	0,2	0,6
1615;	1013;	0,2	0,6
1013;	1615;	0,2	0,5
1615;	3144;	0,2	0,6
3144;	1013;	0,2	0,5
3144;	1615;	0,2	0,5
1013;	3144;	0,2	0,5
3144;	797;	0,2	0,5
797;	3144;	0,2	0,5
130;	151;	0,2	0,5
151;	130;	0,2	0,5

Table 2. First 20 association rules from two Rock musical pieces.

CLASSICAL 20			
X	Y	SUP	CONF
1013;	517;	0,5	0,8
517;	1013;	0,5	0,6
323;	517;	0,4	0,8
517;	323;	0,4	0,6
410;	517;	0,4	0,8
517;	410;	0,4	0,6
625;	517;	0,4	0,7
517;	625;	0,4	0,6
797;	517;	0,4	0,7
517;	797;	0,4	0,5
1013;	323;	0,3	0,6
323;	1013;	0,3	0,6
323;	625;	0,3	0,6
625;	323;	0,3	0,6
410;	625;	0,3	0,6
625;	410;	0,3	0,6
410;	323;	0,3	0,6
323;	410;	0,3	0,5
797;	410;	0,3	0,6
410;	797;	0,3	0,6

CLASSICAL 25			
X	Y	SUP	CONF
410;	517;	0,6	0,8
517;	410;	0,6	0,7
625;	517;	0,5	0,8
517;	625;	0,5	0,7
625;	410;	0,5	0,8
410;	625;	0,5	0,7
410;625;	517;	0,4	0,8
517;625;	410;	0,4	0,8
410;517;	625;	0,4	0,7
625;	410;517;	0,4	0,6
410;	517;625;	0,4	0,6
517;	410;625;	0,4	0,5
797;	517;	0,4	0,8
517;	797;	0,4	0,5
323;	517;	0,4	0,8
517;	323;	0,4	0,5
1249;	410;	0,3	0,8
410;	1249;	0,3	0,5
797;	410;	0,3	0,8
410;	797;	0,3	0,5

Table 3. First 20 association rules from two Classical musical pieces.

If we consult the data on Table 3 we notice that 12 of the 15 rules are identical for both classical files. The rule '797;' → '517;' can be interpreted as following: when the spectrum bin '797;' describe the audio signal, the spectrum bin '517;' also describe the audio signal (see table 1). This association rule has been used to do research throughout the database. Table 4 indicates the results obtained when rule '797;' → '517;' has been applied as a research criteria. The results show that a total of 115 pieces have been returned. Among these pieces, 72 were classical. As a reminder, the dataset used for the experiment contains 10 genres, each represented by 100 tracks of the same duration. Thus obtaining a recall of 72% and a precision of 63%.

Genre	File	Percentage
BLUES	20	17%
CLASSICAL	**72**	**63%**
COUNTRY	8	7%
DISCO	6	5%
HIP-HOP	1	1%
JAZZ	6	5%
METAL	1	1%
POP	1	1%
REGGAE	0	0%
ROCK	0	0%
Total :	**115**	

Table 4. Results using 1 classical rule.

Table 5 illustrates the results generated when using the association rules '797;' → '517;' , '323;' → '517;' and '625;' → '517;' as research criteria. Adding association rules as research criteria contributes to a direct impact on the results obtained. The recall drops to 56 % but the precision rises to 69%.

Genre	File	Percentage
BLUES	13	16%
CLASSICAL	**56**	**69%**
COUNTRY	3	4%
DISCO	3	4%
HIP-HOP	1	1%
JAZZ	4	5%
METAL	1	1%
POP	0	0%
REGGAE	0	0%
ROCK	0	0%
Total :	**81**	

Table 5. Results using 3 classical rules.

Table 6 shows the results from the use of five association rules as search criteria. The rules used were: '797;' → '517;', '323;'→ '517;', '625;'→ '517;', '410;'→ '517;' and '410;'→ '625;'. These results confirm the differentiation power of the association rules. When using five rules, the recall level dropped slightly to 52%, however, the precision level increased to 75%.

Genre	File	Percentage
BLUES	9	13%
CLASSICAL	**54**	**75%**
COUNTRY	1	1%
DISCO	3	4%
HIP-HOP	1	1%
JAZZ	4	6%
METAL	0	0%
POP	0	0%
REGGAE	0	0%
ROCK	0	0%
Total :	**72**	

Table 6. Results using 5 classical rules.

There is a correlation between the number of used association rules and the recall and precision levels. This correlation is illustrated in Figure 3. The x axis represents the number of rules used while the y axis represents the score obtained.

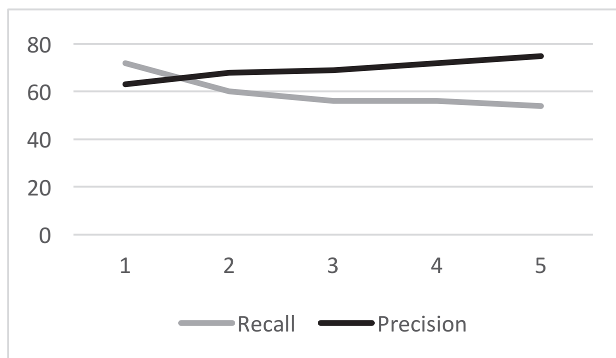

Figure 3. Correlation between the number of association rules, recall and precision.

From our results, we see that the co-occurrence relations shown by the association rules can be informative and discriminant. So, they can be used for retrieving genre-specific music files. Experiments conducted have shown that the proposed method will work best with classical music genre. On the other hand, rock music is the genre with which this method is the least likely to succeed. Similar results have been obtained in (Deshpande et al., 2001; Tzanetakis & Cook, 2002). In fact, the rock genre is often considered as ill-defined category.

The added value of using association rules mining is that it allows the identification of descriptors and the measure of their usefulness. The discovery of discriminant associations offers a multitude of opportunities. The ability to select the right descriptors is one of the main difficulty associated with the development of automated musical data classification systems. The opacity of current methods such as neural network makes it difficult to evaluate the real value of descriptors in the classification process. The proposed method can highlight the importance of these descriptors. Therefore, our method allows the identification of content-based descriptors that should be privileged to index audio files.

Conclusion

In this paper, we suggested to use the association rules to support the exploration of musical data. The proposed methodology promotes the research of musical similarities because it enables us to generate useful descriptors as well as indicate how these descriptors can be used to retrieve genre-specific music file. The conclusion of the experiments demonstrates that association rules can yield interesting results especially for retrieving classical music. Even if this paper is dedicated to music, the proposed methodology can be adjusted to be applied to image, video or text. Thus, we can consider a unique approach applicable to different types of information encoding.

Reference

Agrawal, R., Imielinski T., and Swami, A. 1993. *Minning association rules between sets of items in large databases*, In Proc. of the SIGMOD Conference on Management of Data, pp 207-216.

Agrawal, R., Srikant, R. 1994. *Fast Algorithms for Mining Association Rules*, In Proc. of the 20th International Conference on Very Large Database, pp. 487-499.

Biskri, I., Rompré, L. 2012. *Using association rules for query reformulation*. In Next Generation Search Engines: Advanced Models for Information Retrieval.

Bray, S., Tzanetakis, G. 2005. *Distributed Audio Feature Extraction for Music*, In Proc. ISMIR, pp. 434-437.

Cherfi, H., Napoli, A. 2005. *Deux méthodologies de classification de règles d'association pour la fouille de textes*, Revue des nouvelles technologies de l'information.

Conklin, D. 2009. *Melody classification using patterns*, In Second International Workshop on Machine Learning and Music, pp. 37-41.

Deshpande, H., Singh, R. and Nam, U. 2001. *Classification of Music Signals in The Visual Domain*, *In Proc. the COST G-6 Conference on Digital Audio Effects*.

Diop, C. T., Lo, M. 2007. *Intégration de règles d'association pour améliorer la recherche d'informations XML*. Actes de la Quatrième conférence francophone en Recherche d'Information et Applications.

Downie, J. S, Ehmann, A. and Tcheng, D. 2005. *Music-to-knowledge (M2K): A prototyping and evaluation environment for music information retrieval research*, In Proc. SIGIR Conference on Research and Development in Information Retrieval.

Geng, L., Hamilton, H. J. 2006. *Intertingness Measures for Data Mining: A Survey*, In Journal ACM Computing Surveys, Vol. 38, Issue.

Grosche, P., Müller, M. and Serrà, J. 2012. *Audio Content-Based Music Retrieval*, In Multimodal Music Processing. vol. 3, pp. 157-174.

Lallich, S., Teytaud, O. 2004. *Validation de l'intérêt des règles d'association*. Revue des Nouvelles Technologies de l'Information, Vol. 1, No 2, pp. 193-218.

Lartillot, O., Toiviainen, P. 2007. *A Matlab toolbox for musical feature extraction from audio*, In Proc. of the 10th International conference on Digital Audio Effects.

Lenca, P., Meyer, P., Vaillant, B. and Lallich, S. 2008. *On selecting interestingness measures for association rules: user oriented description and multiple criteria decision aid*", European Journal of Operational Research, p 610–626.

Li, T., Ogihara, M. and Li, Q. 2003. *A comparative study on content-based music genre classification*. In Proceedings of SIGIR conference on Research and development in information retrieval, pp. 282-289.

Ness, S. R., Theocharis, A., Tzanetakis, G. and Martins, L. G. 2009. *Improving automatic music tag annotation using stacked generalization of probabilistic svm outputs*. In Proc. ACM international conference on Multimedia, pp. 705-708.

Sturm, B. L. 2012. *An analysis of the GTZAN music genre dataset*, In Proceedings of the second international ACM workshop on Music information retrieval with user-centered and multimodal strategies, pp. 7-12.

Sturm, B. L. 2012. *A survey of evaluation in music genre recognition*, In Adaptive Multimedia Retrieval: Semantics, Context, and Adaptation, pp. 29-66.

Swathi, V., Reddy, M. S. 2014. *Music Recommendation System Using Association Rules*, In International Journal of Technology enhancements and emerging engineering research, Vol. 2, Issue 7.

Tzanetakis, G., Cook, P. 2002. *Musical Genre Classification of Audio Signals*, In IEEE Transactions on Speech and Audio Processing, Vol. 10, No. 5, pp. 293-302.

Wang, A. 2003, *An Industrial-Strength Audio Algorithm*. In Proc. of ISMIR 2003, pp. 7-13.