國立交通大學

資訊科學與工程研究所

碩 士 論 文

個人化音樂情緒反應預測系統之建造

Building a Personalized Music Emotion Prediction System

研 究 生：葉展彰

指導教授：曾憲雄　教授

中 華 民 國　九 十 五　年 六 月

個人化音樂情緒反應預測系統之建造

Building a Personalized Music Emotion Prediction System

研 究 生：葉展彰　　Student：Chan-Chang Yeh

指導教授：曾憲雄 博士　　Advisor：Dr. Shian-Shyong Tseng

國 立 交 通 大 學

資 訊 科 學 與 工 程 研 究 所

碩 士 論 文

A Thesis

Submitted to Institute of Computer Science and Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

June 2006

Hsinchu, Taiwan, Republic of China

個人化音樂情緒反應預測系統之建造

研究生：葉展彰　　　　　　　　　　　　　　指導教授：曾憲雄博士

國 立 交 通 大 學 資 訊 科 學 研 究 所

摘要

隨著多媒體技術的進步，數位化的音樂已經廣為散佈，傳統的音樂研究集中在音樂的分類、推薦以及分析，現今的研究則試著要找出音樂與人類情緒反應的關係，第一種作法，找出音樂屬性與音樂情緒反應的關係，建立一個音樂情緒的反應模型，預測所有人的音樂心情反應，此種做法最主要的缺點就是沒有考慮到個人的差異性情緒反應；另一種作法，針對每個人訓練出音樂情緒反應模型，再利用這個模型來預測音樂情緒反應，此法雖然有考慮到個人間的差異性，但是主要的缺點則在於訓練模型時的時間浪費，事實上，雖然個人的反應存在差異性，但依然具有群組性的行為模式。

本論文主要改進以上缺點，提出一個個人化的音樂情緒反應預測分析系統，考慮使用者背景的差異性，來預測使用者的音樂情緒反應，其分析過程總共包括五個階段：1)資料前處理；2)使用者情緒反應群體分群；3)使用者情緒反應群體分類；4)音樂情緒預測；5)個人化音樂情緒反應規則整合。經過以上五個階段，就可以產生個人化的音樂情緒反應預測規則，並將之建於個人化音樂情緒反應預測系統內，用來預測音樂情緒。在使用的過程，只要預先知道使用者的背景相關資料，輸入某首音樂，便可依照音樂的屬性值，預測此人的音樂情緒反應。

本論文最後邀請二十四個人聆聽二十首音樂來做實驗，利用此訓練資料產生出個人化情緒反應預測規則，並邀請十個人聆聽四首音樂來做測試，測試結果可達百分之七十的準確度。

關鍵字：個人化音樂情緒反應預測、資料探勘、分群法、分類法

Building a Personalized Music Emotion Prediction System

Student: Chan-Chang Yeh Advisor: Dr. Shian-Shyong Tseng

Department of Computer and Information Science
National Chiao Tung University

Abstract

With the development of multimedia technology, digital music has been wide spread and research on music is getting more and more popular. Traditional researches focus on music analysis, classification and recommendation. Nowadays researchers try to study the relationship between music and listeners' emotion. Some researchers use single model to predict people's emotional response to music. However, the effect of music emotion response prediction is limited without taking listeners' differences into consideration. Some researches try to build a model for each listener, and use the model to predict each one's emotional response. But without using the similarity of users' emotion response, the training time of model is an overhead.

Therefore, we propose a Personalized Music Emotion Prediction (P-MEP) system to assist predicting users' music emotion concerning with users' differences. To analyze listeners' emotion response to music, a series of data mining techniques are applied. Our approach includes the following phases: 1) Data Preprocessing Phase, 2) User Emotion Group Clustering Phase, 3) User Group Classification Phase, 4) Music Emotion Classification Phase, 5) Personalized Music Emotion Prediction Rules Integration. The personalized music emotion prediction rules will be generated and embedded in the personalized music emotion prediction system.

During the application procedure, the P-MEP system will predict the listener's emotion response to music. Finally, 24 listeners are invited to annotate the emotion of 20 pieces of music to be the training data and another 10 listeners are invited to annotate the emotion of 4 pieces of music to be the testing data. The result of the experiment shows that the generated personalized music emotion prediction rules can be used to predict emotional response to music concerning with listeners' differences and the accuracy is up to 70%.

Keywords: Personalized Music Emotion Prediction, Data Mining, Classification, Clustering

誌　謝

本篇論文的完成，必須感謝很多人的支持與協助。首先要感謝我的指導教授曾憲雄博士，由於老師細心的指導與勉勵，讓我的論文能夠順利的完成，而他治學嚴謹的態度，讓我在研究上觀念的釐清受益匪淺，待人處世方面，也有很大的啟發。同時，也必須感謝我的口試委員，洪宗貝教授、葉耀明教授以及黃志方教授，他們對此篇論文提出許多寶貴的建議，讓此篇論文更臻完善。

此外，也要感謝佩瑾學姊以及瑞峰學長，他們對於此研究提供了不少的建議與協助，在與他們討論的過程中，有相當大的收穫。另外，慶堯學長、俊銘學長、順傑學長、衍旭學長、威勳學長以及哲青學長，你們的意見與指教，也豐富了論文的內容。

實驗室的同窗好友，昇翰、珍妮、嘉文、永彧、晉璿、仁杰、南極、喚宇，謝謝你們陪我度過這既忙碌又充實的碩士生涯，我永遠不會忘記一起挑燈夜戰的時光。

最後，我要感謝我的家人與朋友，你們的支持是我向前邁進最大的動力，有你們的陪伴，我對未來將無所恐懼。要感謝的人相當的多，在此僅向所有幫助過我的朋友，致上我最深的謝意。

Table of Contents

List of Tables

Lists of Figures

List of Algorithms

Chapter 1 Introduction

Emotion recognition is very important in studying music. Composers in the past era wrote the score to express their emotion. Performers such as conductors, pianists and violinists try to simulate the composer's emotion and add their self-emotion when playing music. People who listen to the music try to feel the composer's or performer's emotion when they wrote or played the music. Music has very significant influences to listeners' emotion no matter how the music makes the listeners relaxing, happy, anxious or sad. Therefore emotion expression is the most charming part of music.

Generally speaking, music with rapid tempo and major key will induce the happy emotion for a listener. While music with slow tempo and written in minor key will have a feeling of sadness, and music with a rapid tempo combined with dissonance will induce fear [1]. But how can we predict the listener's emotion when he or she listens to music?

With the development of multimedia technology, music has been digitalized in the form of .WAV or .MIDI and has now been in widespread use. To find out some useful information or knowledge from such huge amount of music files has become an important issue. In current researches, most of them focused on music structure analysis [10][16][17][19][20], music classification [3][14][18] or implementing music recommendation system [2][8][11]. However, few researches studied perceived emotion when a person listens to music.

In fact, a music emotion prediction system or music recommendation system by music emotion is needed. It may be useful in choosing the background music in restaurant, bookstore or department store to relax guests' mood or stimulate the willingness to shopping. In addition, choosing the music in movie or music therapy are also the applications of the system. Some researches [11][18] focus on study the relationship between music patterns and emotion and their systems could be used to predict the emotion of listeners. However, the listeners will get the same emotion in the system even though they have different backgrounds. In fact, listeners' backgrounds will affect the emotion to music.

In this research, we proposed a Personalized Music Emotion Prediction (P-MEP) System based on the user profiles and the music attributes to predict a listener's emotion when he or she listens to music. There are personalized music emotion prediction rules which are obtained from a series of data mining approach named personalized music emotion prediction analysis procedure and then used to predict the music emotion in this system.

Our approach includes the following five phases: **1. Data Preprocessing Phase:** In this phase, we collect three kinds of data to be the training case including user profiles, music attributes and the music emotion. **2. User Emotion Group Clustering Phase:** The clustering technique is applied to divide listeners into several user emotion groups. Listeners with the similar music emotion are clustered into the same user emotion group. **3. User Group Classification Phase:** After the user emotion group constructed, every created cluster is tagged with a cluster label. Then the classification technique is applied to find out the relationship between the user profiles and the user emotion group. **4. Music Emotion Classification Phase:** In this phase, the classification technique is applied to find out the relationship between music attributes and music emotion for each user emotion group. **5. Personalized Music Emotion Prediction Rules Integration:** Finally, a rule integration method is applied to combine these two parts of rules generated in phase 3 and 4. Then the personalized music emotion prediction rules are generated.

During the applying procedure, the P-MEP system firstly classifies the new listener into a specified user emotion group by the user profiles and then predicts the emotion of the listener with the music attributes extracted from the new music.

The rest of the thesis is arranged as follows. Section 2 gives a review on related music researches. Section 3 and 4 introduces the personalized music emotion prediction analysis procedure in detail. The system will be described in Section 5. Experiments design and the results of the experiments will be analyzed in Section 6. Finally, Section 7 concludes the thesis.

Chapter 2 Related Work

2.1 Traditional Computer Music Researches

With the development of digital music technology, many researches are based on the content based music information retrieval. The music data they used is either MIDI files or the audio files.

Chai and Veroce [3] in MIT Media Lab used the Hidden Markov Models to do automatic classification about Irish, German and Austrian folk music. Shan et al. [24] extracted the melody form polyphonic MIDI files and did the music classification by chord sets. Li and Ogihara [14] used genre taxonomy to do the classification about music genre including Blues, Classical, Country, Disco, Hip-Hop, Jazz, Metal, Pop, Reggae and Rock. All of the above researches are based on the music classification.

For the music recommendation system, there are two approaches named the content-based filtering approach and the collaborative filtering approach. Chen and Chen [4] proposed a music recommendation system based on users' preference by clustering the music MIDI files in the database in advanced. Cano et al. [2] proposed a music audio recommendation method based on the instrumentation, rhythm and harmony from music audio signals. Iwahama et al. [8] proposed a content-based filtering system for music data. In the system, a decision tree is generated to classify whether the user likes or dislikes the music. Kuo and Shan [13] proposed a personalized music filtering system based on melody style classification about music MIDI files including Chinese folk song, Enya, Chopin and Bach. The melody style was described by chord sets.

There are some researches focused on music structure. Kuo and Shan [12] proposed a music retrieval system by melody style. The chord sets were also used to describe the melody style. Maddage [19][20] analyzed audio music by rhythm extraction, chord detection and vocal / instrument boundary detection for popular music. Chen et al. [5] proposed a method to do music segmentation by rhythmic features and melodic shapes. The melodic shapes included convex, descending, ascending and concave which describe the difference between adjacent pitch notes.

There are still some researches focused on mining sequential patterns in music. The technique is mainly applied in finding out the music motive like the famous sol-sol-sol-mi in Beethoven's Fifth Symphony [10][16]. And the repeating patterns were then used in query and classification [15].

Traditional researches studied music classification, music recommendation and music analysis over digital music. However they didn't discuss the issue of emotion.

2.2 Music Emotion Analysis

Here we review three papers that are related to music emotion.

In [11], Kuo et al. proposed an emotion-based music recommendation system based on film music. Totally 107 film music midi files from 20 animated films including the productions of Disney, Studio Ghibli and Dream Works were selected to do the experiments. The music attributes included melody, rhythm and tempo and the emotion model that was proposed by Reilly [22] was used to describe the music emotion. There were 30 kinds of emotion that had been grouped into 15 groups. The relationship of music features and emotions of the training data was discovered by a graph-based approach named Mixed Media Graph (MMG) and the MMG method was further refined to be the Music Affinity Graph that was suitable to use in the music recommendation system. During the application time, user queried a kind of music, then the system tried to rank all the music stored in the music database with their music features which were related to this kind of emotion. Music with top ten score was recommended to the user. The average score of recommendation results achieved 85 %.

In [18], Lu et al. focused on the acoustic music data and three features sets including intensity, timbre and rhythm were extracted to represent the characteristics of a music clip using some digital signal processing technique. There were totally 800 representative music clips of 20 seconds selected from 250 pieces of music and the Thayer's two-dimensional model of mood including Exuberance, Anxious/Frantic, Contentment and Depression was adopted to describe the music emotion. Three experts joined the experiment and annotated music in the database. Initially, the Gaussian mixture model (GMM) was utilized to model each feature set. A classification model was obtained by classifying music into 4 emotions and then used to detect the music mood in an entire music clip to distinguish every sub-music clip with different music mood. The precision of resultd in the experiment is about 81.5%.

As mentioned above, only a music classification or recommendation model could be obtained in their systems. It means that everybody should have the same emotional response to music. Accordingly, these models are not suitable to represent the situation that listeners with different background may have different emotional response to music.

In [25], 18 music features including pitch, interval, tempo, loudness, note density, tonality, etc. were used to represent the music midi files and a modified Thayer's emotion model with six kinds of emotions including Joyous, Robust, Restless, Lyrical, Sober, Gloomy were used. Support Vector Machine (SVM) was used to do the classification over the music data. In experiment, 20 different listeners with and without music backgrounds were joined to label the music data. Wang considered the phenomenon that listeners with different backgrounds may have different emotional response to music. So they trained the music classification model for all the 20 different listeners. Then these 20 sets of music emotion classification rules were used to predict the listener's emotion to music.

In this research, they predict the emotion in a user-adaptive way. But everyone should have a trained prediction model before prediction. It will be an overhead and the method is difficult to be applied for everybody. Although listeners with different backgrounds may have different emotional response to music but those with similar backgrounds may have similar emotional response to music. The prediction rules and the knowledge could be reused to predict new listeners' emotion.

In order to improved the drawbacks in the previous researches, the fact that listeners with different backgrounds may have different emotional response to music should take into consideration to model the phenomenon that people' music emotional response has a grouped effect. We propose an analysis procedure that consists of data mining techniques and will be described in detail in the following sections.

Chapter 3

Personalized Music Emotion Prediction (P-MEP)

For the same music, listeners with different user profiles may have different emotion to music but some may have similar emotion to music. Generally speaking, music with different music patterns may result in different emotion to listeners. In order to model the characteristics of the grouped emotional response to music of listeners and find out the relationship between music patterns and music emotion of listeners, an analysis procedure of the Personalized Music Emotion Prediction (P-MEP) is proposed.

3.1 Problem Definition

To analyze the listeners' music emotion with music patterns and their backgrounds, we have two assumptions: Firstly, for the same music, listeners with different user profiles may have different emotions to music. Secondly, music with different music patterns or different music attribute values may have different emotion to listeners.

During the construction of the predicting model, there are two terminologies: 1) User Emotion Group (UEG): User emotion group is a group consisting of several listeners who have the similar emotional response to music. According to experts' suggestions the total number of user emotion group could be three to five. 2) Representative Music: Representative music is a kind of music that makes the emotional response of the listeners different and is selected by expert. It results in that listeners would be separated into different user emotion groups according to their emotion to representative music.

Inputs of the problem are user profiles, music attributes and music emotion. The problem is to analyze music emotion concerning with listeners' differences and music attributes. Output is the personalized music emotion prediction rules.

3.2 Feature Selection

In this section, three kinds of input data including user profiles, music attributes and music emotion will be described in detail.

(1) User Profiles:

The user profiles including 'Gender', 'Age', 'Education Status', 'Job', Geographical Location', 'Constellation' and some 'Personality' [28] index, etc. are defined as user profiles vector in Definition 3.1 to represent a listener's backgrounds. Feature selection about user profiles is based on expert's suggestion and [28] and may have effects on music emotion.

Definition 3.1: User Profiles Vector

Listener: L (Background, Personality)$_i$ denotes the i-th listener, and it has 2 categories containing 24 tuples.

- **Background = <GN, AG, ES, CS, GL> denotes the background of listener**
- **Personality = <A, B, C, …, R, S> denotes the personality of listener**

The value of each user profile attribute is listed in Table 3.1.

Table 3.1: The User Profiles of a Listener

Category	Attribute	Value	
Background	**Gender (GN)**	**M:** *Male,* **F:** *Female*	
	Age (AG)	**L:** [l ,l+Δ) M: [l+Δ, l+2Δ) H: [l+2Δ, μ]	
	Education Status (ES)	**E:** *Elementar,;* **J:** *Junior,* **S:** *Senior* **U:** *Undergraduate,* **M:** *Master,* **D:** *Doctor*	
	Constellation (CS)	**Aqu:** *Aquarius*	**Ari:** *Aries*
		Can: *Cancer*	**Cap:** *Capricornus*
		Gem: *Gemini*	**Leo:** *Leo*
		Lib: *Libra*	**Pis:** *Pisces*
		Sag: *Sagittarius*	**Sco:** *Scorpio*
		Tau: *Taurus*	**Vir:** *Virgo*
	Geographical Location (GL)	**N:** *North,* **C:** *Center,* **S:** *South,* **E:** *East*	
Personality	**A**	**1:** Placid	**2:** Vivid
	B	**1:** Like to stay alone	**2:** Like to socialize
	C	**1:** Like physical exercise	**2:** Like brainstorming
	D	**1:** Expansive	**2:** Shy
	E	**1:** Unambitious	**2:** Ambitious
	F	**1:** Confident	**2:** Unconfident
	G	**1:** Self-centered	**2:** Sensitive to others
	H	**1:** Pedantry	**2:** Creative
	I	**1:** Patient	**2:** Inpatient
	J	**1:** Despotic	**2:** Like to be led
	K	**1:** Amiable	**2:** Severe
	L	**1:** Mild	**2:** Combative
	M	**1:** Unpredictable tempered	**2:** Stable tempered
	N	**1:** Serious	**2:** Humorous
	O	**1:** Cold	**2:** Enthusiastic
	P	**1:** Responsible	**2:** Irresponsible
	Q	**1:** Moderate	**2:** Irritable
	R	**1:** Obstinate	**2:** Like to be guided
	S	**1:** Self-restraint	**2:** Easy-going

(2) Music Attributes:

Music has often been divided into three categories including monophonic, homophonic and polyphonic based on the amount of concurrency present. Monophonic music is a kind of music that only one note sounds at a time, homophonic music is music in which multiple notes may sound at once, but all notes start and finish at the same time, and polyphonic music is the most general form of music, in which multiple notes may sound independently. We are concerned with polyphonic music and choose the track with the highest pitch density to be the main track [4], because composer uses the most distinct pitch notes in the main track. The pitch density is defined in Equation 3.1.

Equation 3.1: Pitch Density of Music

$$Pitch\ density = \frac{NP}{AP} \quad [4]$$

- *NP is the number of distinct pitches in the track*
- *AP is the number of all distinct pitches in MIDI standard, i.e. 128* [27]

In [9], some music attributes including tempo, mode, loudness, pitch, intervals, melody, harmony, tonality, rhythm, tempo, articulation, amplitude envelope and musical form were proved to have the effect in listener's emotional response to music, and may interact with each other. In this thesis, three important sets of music attributes are extracted from the music midi file data which are listed in Table 3.2 and are defined as the music attributes vector in Definition 3.2 to describe a piece of music.

Definition 3.2: Music Attributes Vector

Music: M (Pitch, Rhythm, Velocity, Timber, Mode)$_i$ denotes i-th music and has 5 categories containing 12 tuples.

- **Pitch = <PM, PS, IM, IS, PE, PD> denotes the pitch information of music**
- **Rhythm = <TD> denotes the rhythm information of music**
- **Velocity = <LM, LS> denotes the velocity information of music**
- **Timber = < TB > denotes the timber of music**
- **Mode = < MD, TN > denotes the mode and tonality of music**

For the music attributes, PM, PS, IM, IS, PE, PD, TD, LM, LS are numerical attributes, while TB, MD, TD are categorical attributes.

Table 3.2: Music Attributes

Category	Attributes	Description
Pitch	**Pitch Mean (PM)**	Mean of pitch value of the main track.
	Pitch Standard Deviation (PS)	Standard deviation of pitch value of the main track.
	Interval Mean (IM)	Mean of the difference of adjacent two note pitch in the main track.
	Interval Standard Deviation (IS)	Standard deviation of the difference of adjacent two note pitch in the main track.
	Pitch Entropy (PE)	$PE = -\sum_{j=1}^{NP} P_j \log P_j$,where $P_j = \frac{N_j}{T}$, Nj is the total number of notes with the corresponding pitch in the main track and T is the total number of notes in the main track. [4]
	Pitch Density (PD)	Defined in Equation 3.1
Rhythm	**Tempo Degree (TD)**	Ratio of the number of fast measures to the number of measures in the main track. A measure is a fast measure if the average note duration in the measure is shorter than one. [4]
Velocity	**Loudness Mean (LM)**	Average value of the note velocities
	Loudness Standard Deviation (LS)	Standard deviation of the note velocities
Timber	**Timber (TB)**	Timber of the main track
Mode	**Mode (MD)**	Mode of the music (Major, Minor)
	Tonality (TN)	Tonality of the music (C, C#, D, D#, E, F, F#, G, G#, A, A#, B)

(3) Music Emotion:

The music emotion vector tagged by m-th listener for n pieces of music is defined in Definition 3.3.

Definition 3.3: Music Emotion Vector

Music Emotion: ME $(E_1, E_2, E_3, \ldots, E_i, \ldots, E_n)_m$ denotes music emotion of m-th listener.

- **E_i represent the emotion of i-th music**

A music emotion concept hierarchy based upon two emotion model proposed by Thayer [25] and Reilly [22] is proposed to describe the listeners' emotional response to music. In Thayer's two dimensional model of mood, there are Exuberance, Anxious/Frantic, Contentment and Depression emotions. Among them, Exuberance and Contentment are positive emotions while Anxious/Frantic and Depression are negative emotions. In addition, Exuberance and Anxious/Frantic are mapped to high intensity while Contentment and Depression are mapped to low intensity [18].

In Reilly's emotion model, there are 30 kinds of emotions grouped into 15 groups in [11], where the emotions in the same group assumingly have the same emotion. In our thesis, we adopted different grouped policy. We map the Reilly's emotion model to Thayer's mood model considered whether they are positive or negative emotion and whether the intensity is high or low.

The music emotion concept hierarchy is shown in Figure 3.1. The first layer has one element named ***Emotion*** in the first layer named **Root**. There are two elements named ***Positive (A1)*** and ***Negative (A2)*** in the second layer named **Layer A**. In the third layer named **Layer B**, ***Exuberance (B1)*** and ***Contentment (B2)*** have the AKO relation to ***Positive (A1)***, while ***Anxious/Frantic (B3)*** and ***Depression (B4)*** have the AKO relation to ***Negative (A2)***. This layer is referred to Thayer's two dimensional model of mood. Finally, in the fourth layer named **Layer C**, there 30 elements marked ***C1***, ***C2***, …, ***C30***. They also have the AKO relation to ***B1***, ***B2***, ***B3*** and ***B4***. This layer is referred to Reilly's emotion model.

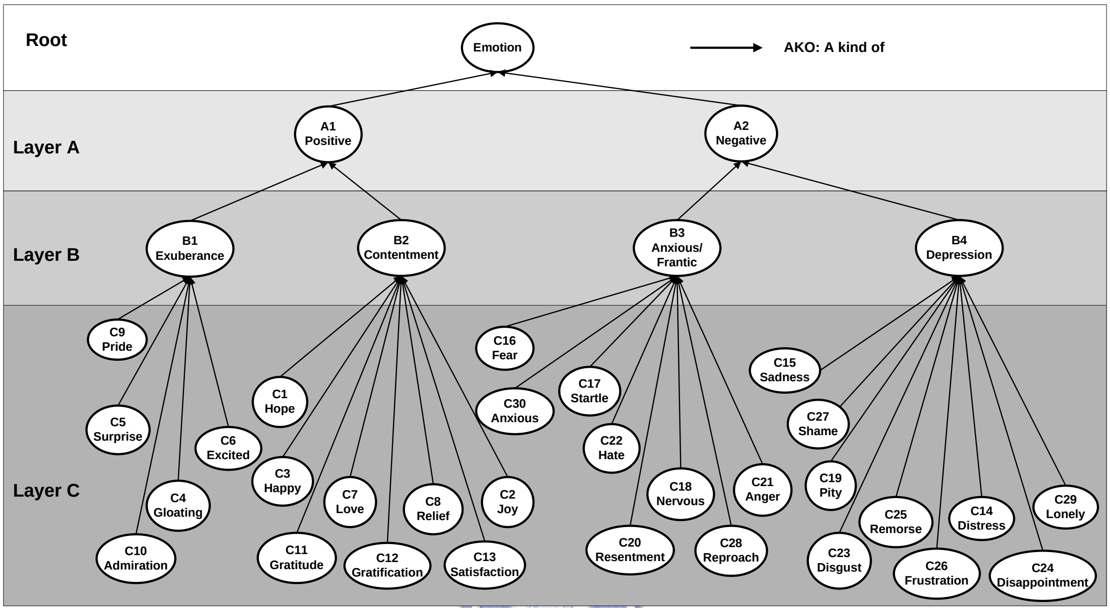

Figure 3.1 : Music Emotion Concept Hierarchy

3.3 Analysis Procedure

In order to predict the music emotion concerning with listeners' differences, we propose a Personalized Music Emotion Prediction (P-MEP) Analysis Procedure which consists of a series of data mining process to extract the personalized music emotion prediction rules and then adopt them in the prediction system.

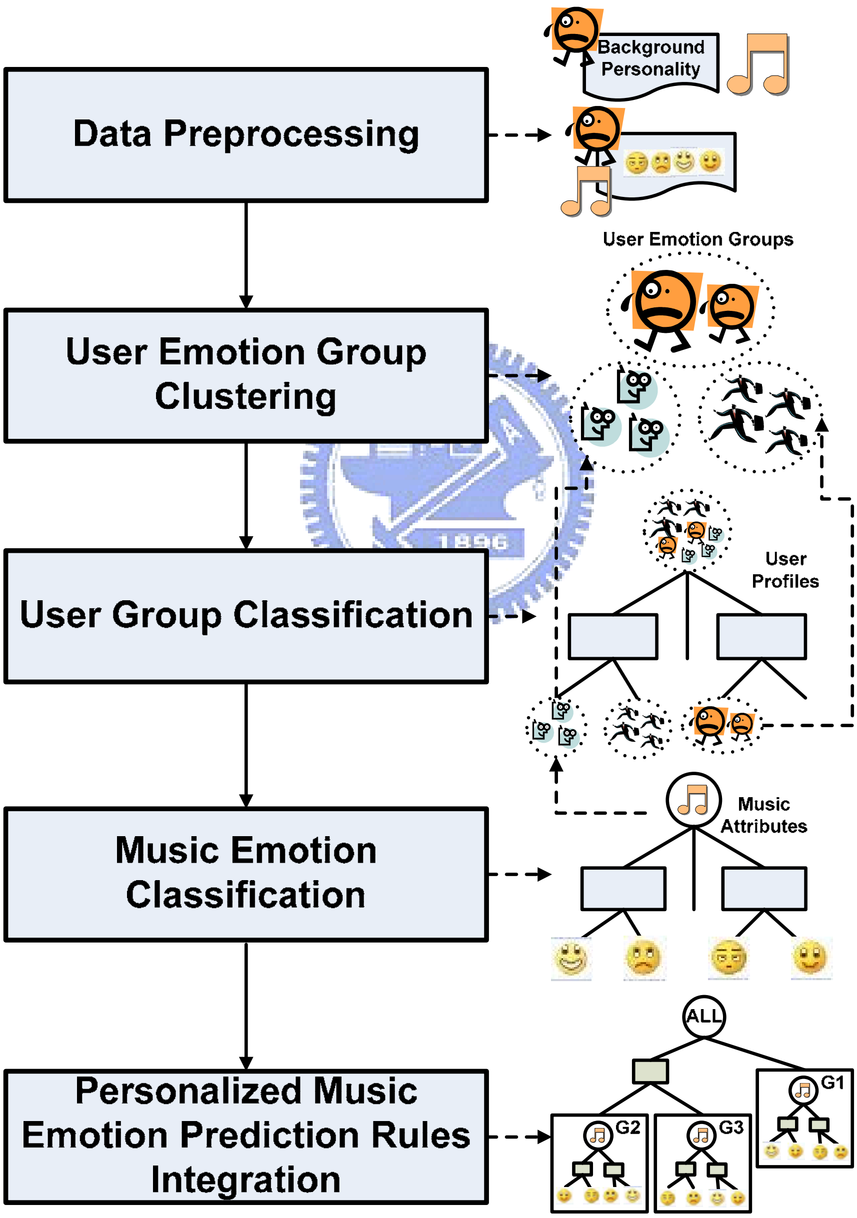

Figure 3.2 : Analysis Procedure of Personalized Music Emotion Prediction

As shown in Figure 3.2, the analysis procedure including five phases is described as follows:

(1) **Data Preprocessing Phase:** In this phase, the user profiles, music attributes and music emotion are collected. Firstly, all the listeners fill up a designed questionnaire which includes all the user profiles on it. Then listeners write down the emotion they perceived after representative music play. Finally, music attributes are obtained by parsing midi file. All these three kinds of data are preprocessed and then saved into a database.

(2) **User Emotion Group Clustering Phase:** Music emotion tagged by listeners to music is transformed into emotion vectors. Then the Robust Clustering Algorithm for Categorical Attributes (ROCK) method is applied to cluster listeners into several user emotion groups.

(3) **User Group Classification Phase:** In this phase, ID3 classification algorithm is applied to train the relationship between user profiles and user emotion group number. All the listeners are classified into the user emotion group according to their user profiles. The output of this phase is the user group classification rules.

(4) **Music Emotion Classification Phase:** For each user emotion group, we have the representative music emotion vector of m music to represent all the emotion of every listener in a user emotion group. Then C4.5 classification algorithm is applied to the music attributes and music emotion to train the music emotion prediction rules for each user emotion group.

(5) **Personalized Music Emotion Prediction Rule Integration Phase:** The rule integration method is used to combine user group classification rules and music emotion prediction rules to be the personalized music emotion prediction rules.

Chapter 4 Rule Generation of P-MEP

4.1 Data Preprocessing

The purpose of this phase is to collect three kinds of data, including user profiles, music attributes and music emotion. The output is these three kinds of dada. Suppose there are m listeners and n pieces of representative music and the listeners annotate the emotion tag after they listen to music. Three kinds of the vectors including user profiles vector, music attributes vector and music emotion vector are the input of the P-MEP analysis procedure.

Example 4.1: User profiles, music attributes and music emotion vector

In the following section, an example contained 12 listeners and 16 pieces of music will be used to describe the analysis procedure of P-MEP. The user profiles and music attributes are simplified in this example. Table 4.1 shows twelve listeners with their user profile value, Table 4.2 shows sixteen pieces of music with their music attribute value, and Table 4.3 shows the music emotion of 16 pieces of music tagged by these listeners. The emotion model used is the Reilly's emotion model.

Table 4.1: Simplified User Profiles of 12 Listeners

Listener	Gender	Major in Music
L1	Male	Yes
L2	Male	Yes
L3	Male	Yes
L4	Male	Yes
L5	Male	No
L6	Male	No
L7	Female	No
L8	Female	No
L9	Female	Yes
L10	Female	Yes
L11	Female	Yes
L12	Female	Yes

Table 4.2: Simplified Music Attributes of 16 Pieces of Music

Music	Pitch Mean	Tonality	Mode
M1	60	C	Major
M2	72	C	Major
M3	61	C	Major
M4	86	D	Major
M5	52	C	Major
M6	43	C	Major
M7	56	D	Major
M8	49	D	Major
M9	23	F	Minor
M10	25	G#	Minor
M11	29	G#	Minor
M12	65	F	Minor
M13	37	F	Minor
M14	56	F	Minor
M15	21	F	Minor
M16	34	F	Minor

Table 4.3: Emotion of 16 Pieces of Music Tagged by 12 Listeners

	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M 10	M 11	M 12	M 13	M 14	M 15	M 16
L1	1	2	7	11	6	9	10	4	16	17	30	21	26	14	27	15
L2	1	13	7	11	5	10	10	4	16	20	30	21	26	19	27	15
L3	1	2	11	11	5	9	10	4	16	17	18	21	26	14	23	15
L4	1	2	7	3	5	9	6	4	16	17	30	28	26	14	27	29
L5	16	17	30	21	26	14	27	15	1	2	7	11	6	9	10	4
L6	16	20	30	21	26	19	27	15	1	13	7	11	5	10	10	4
L7	16	17	18	21	26	14	23	15	1	2	11	11	5	9	10	4
L8	16	17	30	28	26	14	27	29	1	2	7	3	5	9	6	4
L9	26	14	27	15	16	17	30	21	6	9	10	4	1	2	7	11
L10	26	19	27	15	16	20	30	21	5	10	10	4	1	13	7	11
L11	26	14	23	15	16	17	18	21	5	9	10	4	1	2	11	11
L12	26	14	27	29	16	17	30	28	5	9	6	4	1	2	7	3

4.2 User Emotion Group Clustering

The first task of the analysis procedure would be to find out several user emotion groups by the music emotion that listeners tagged to the chosen representative music. Here we apply the clustering method named ROCK to do clustering.

ROCK is the abbreviation of A Robust Clustering Algorithm for Categorical Attributes that was proposed by Guha, Rastogi and Shim in 2000 [6]. Clustering is a useful technique for grouping data that points within a single group/cluster have similar characteristics. Traditional clustering methods such as K-Means and K-Medoids, etc. are only suitable for clustering data point with numerical values because the way they calculate the similarity of data point is to calculate the numerical distance between data point in the numerical space and it could be only used in numerical data point. But for the categorical and boolean attributes like our emotion model, this kind of method will not be suitable to perform such a clustering task. So here we choose the ROCK method and modify the way of similarity calculation based on our emotion concept hierarchy.

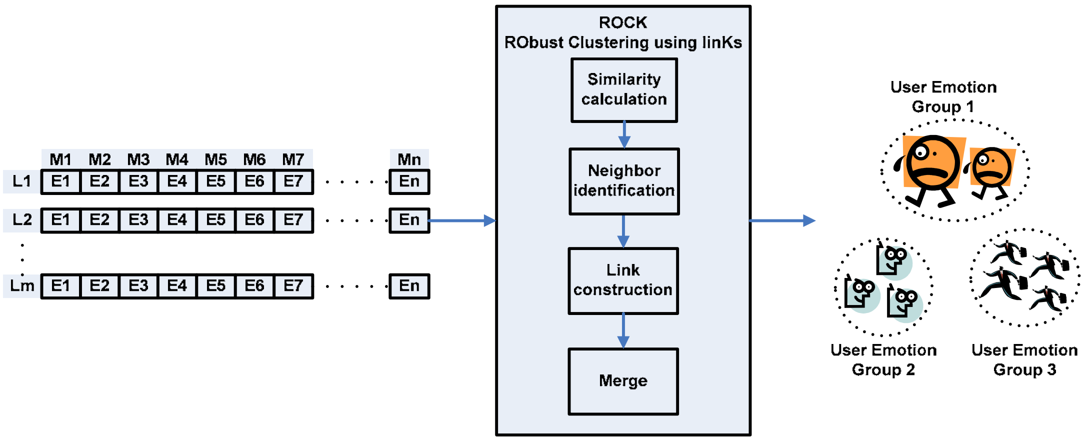

Figure 4.1: User Emotion Group Clustering

The procedure of this phase including several steps is shown in Figure 4.1. The input is music emotion vector of m listeners and the similarity of each data point will be calculated by the proposed music emotion concept hierarchy. Then neighboring points are identified if the similarity value of a pair of point is larger than the threshold. Accumulate the number of links between each pair of points if they have the same neighbor. Every cluster is maintained by a heap and initially it contains only a point. Starting merging the pair of clusters with largest goodness value is performed until there are sufficient numbers of desired clusters or there is no more links between each pair of clusters. Finally, a hierarchical clustering tree would be constructed. According to the expert's suggestion, the number of user emotion groups would be three to five. So we cut in the appropriate place in the hierarchical clustering tree to get 3 to 5 user emotion groups. The detailed algorithm is shown in Algorithm 4.1.

Algorithm 4.1: User Emotion Group Clustering (Modified ROCK Algorithm)

Purpose: To cluster all the listeners into several user emotion groups according to their emotion to representative music.

Symbol Definition:

C_i: The ith Cluster

L_i: The ith listener

Input: Emotion of n representative music of m listeners (ME), Similarity Threshold (θ), Desired Number of Clusters (DNC)

Output: User Emotion Group Label of m listeners (UEGL)

Step1: Calculate the similarity among each ME, $\forall$ i, j, ($i \neq j$), if similarity of L_i and $L_j > \theta$, then L_j and L_j are neighbors.

Step2: Calculate link between neighbors

Step3: Repeatedly execute this step until number of clusters < DNC or there is no more links between each pair of clusters

3.1 Choose the largest goodness measure of C_i and C_j, where the goodness measure $g(C_i, C_j) = \dfrac{link[C_i, C_j]}{(n_i + n_j)^{1+2f(\theta)} - n_i^{1+2f(\theta)} - n_j^{1+2f(\theta)}}$, where $f(\theta) = \dfrac{1+\theta}{1-\theta}$

3.2 Merge C_i and C_j

3.3 Recalculate the link among new clusters

Step4: Output the complete hierarchical clustering tree and UEGL

The similarity calculation of two listeners is based on the equation below.

Similarity of listener L_x and L_y $S_{xy} = \sum_{i=1}^{n} S_i$ /n, *where S_i=similarity of ith music.*

To determine the S_i value, let's go back to the emotion concept hierarchy proposed in previous section. In Layer C (Reilly's model), if emotion of ith music of L_x and L_y are the same, then S_i=1. If emotion of ith music of L_x and L_y are different, then roll up to Layer B (Thayer's model), the S_i value is listed in Table 4.4. For the value 0.8, because in Layer C they are different, but when roll up to Layer B, they are in the same group, so their similarity is set to be 0.8. For the value 0.6, in Layer B, they are either in positive group or negative group, so their similarity is set to 0.6. For value 0.3, although they are not simultaneously in positive or negative group, but they are in either high intensity group or low intensity group, so their similarity is set to 0.3. Finally, for those are not in positive or negative group or high or low intensity group, their similarity is set to 0.

Example 4.2: User Emotion Group Clustering

Following Example 4.1, in ROCK, the similarity of each pair of point will be first calculated. In the initial stage, each listener represents a data point. Table 4.4 shows the similarity value among twelve listeners. Then referred to Step1 in Algorithm 4.1, the neighboring points will be identified if the similarity value of each pair of points is larger than the threshold θ which is 0.6. Table 4.5 shows the number of links among each data point referred to Step2 in Algorithm 4.1. Then referred to Step3, the algorithm starts to merge each cluster based on the goodness value. Finally the hierarchical clustering tree is generated and shown in Figure 4.2. There are three user emotion groups. The result is listed in Table 4.6.

Table 4.4: Similarity Among 12 Listeners

	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12
L1	0.969	0.975	0.969	0	0	0	0	0.2	0.2	0.2	0.2
L2		0.956	0.95	0	0	0	0	0.2	0.2	0.2	0.2
L3			0.956	0	0	0	0	0.2	0.2	0.2	0.2
L4				0	0	0	0	0.2	0.2	0.2	0.2
L5					0.969	0.975	0.969	0.4	0.4	0.4	0.4
L6						0.956	0.95	0.4	0.4	0.4	0.4
L7		**Symmetric**					0.956	0.4	0.4	0.4	0.4
L8								0.4	0.4	0.4	0.4
L9									0.969	0.975	0.969
L10										0.956	0.95
L11											0.956

Table 4.5: The Numbers of Links Among 12 Listeners (data points)

	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12
L1	2	2	2	0	0	0	0	0	0	0	0
L2		2	2	0	0	0	0	0	0	0	0
L3			2	0	0	0	0	0	0	0	0
L4				0	0	0	0	0	0	0	0
L5					2	2	2	0	0	0	0
L6						2	2	0	0	0	0
L7							2	0	0	0	0
L8			Symmetric					0	0	0	0
L9									2	2	2
L10										2	2
L11											2

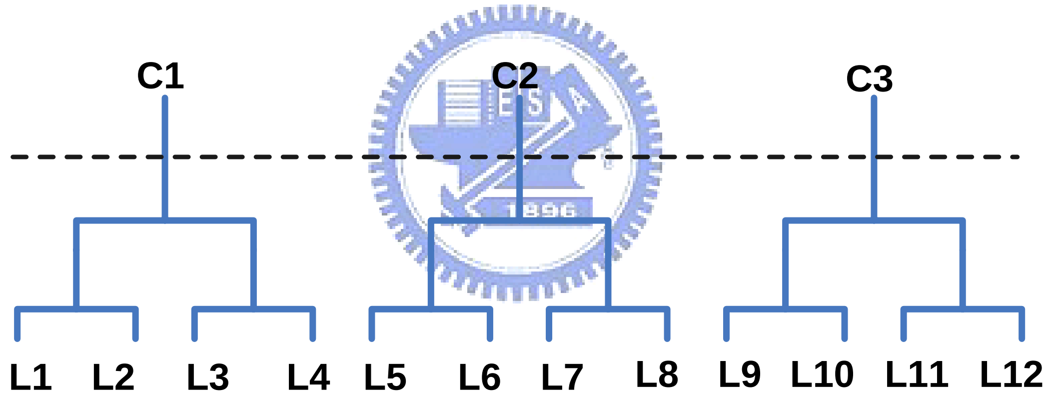

Figure 4.2: Hierarchical Clustering Tree of 12 Listeners

Table 4.6: 12 Listeners and Their User Emotion Group

Listener	L1	L2	L3	L4	L5	L6	L7	L8	L9	L10	L11	L12
Cluster	C1	C1	C1	C1	C2	C2	C2	C2	C3	C3	C3	C3

4.3 User Group Classification

After the representative user emotion groups were constructed, the next step would be finding out the relationship between user profiles of the listeners and the label of the user emotion groups. That is to say we want to know listeners with what kinds of the user profiles would be classify into which user emotion group.

The well-known decision tree algorithm named ID3 [7] is applied. The input in this step is the user profiles and the number of the user emotion group of m listeners. The procedure is shown in the Figure 4.3 and the algorithm is shown in Algorithm 4.2.

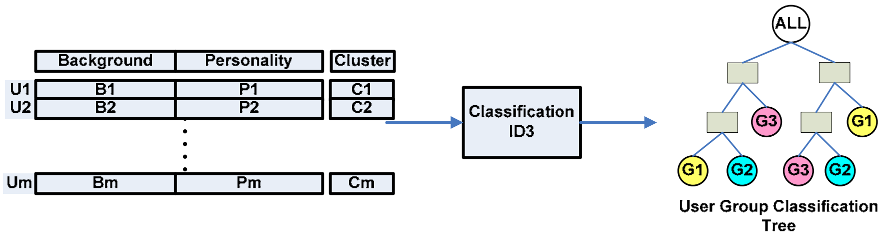

Figure 4.3: User Group Classification

Algorithm 4.2: User Group Classification Algorithm (ID3 Algorithm)

Purpose: Find the relationship between user profiles of the listeners and the label of user emotion groups.

Input: User Profiles of m listeners (UP), User Emotion Group Label of m listeners (UEGL)

Output: User Group Classification Rules (UGCR)

Step1: UP and UEGL are the set of training instances and let them be T

Step2: Choose an attribute that best differentiates the instances contained in T. That is to say to choose an attribute with the largest gain ratio.

Step3: Create a tree node whose value is the chosen attribute. Create child links from this node where each link represents a unique value for the chosen attribute. Use the child link values to further subdivide the instances into subclasses.

Step4: For each subclass created in step 3:

4.1 If the instances in the subclass satisfy predefined criteria or if the set of remaining attribute choices for this path of the tree is null, specify the classification for new instances following this decision path.

4.2 If the subclass does not satisfy the predefined criteria and there is at least one attribute to further subdivide the path of the tree, let T be the current set of subclass instances and return to step 2.

Step5: Output the UGCR.

Example 4.3: User Group Classification

Following Example 4.2, our input now is the user profiles with the user emotion group number which is shown in Table 4.7. After we apply the ID3 classification method, we will get a decision tree that is shown in Figure 4.4. And the rules are listed in Figure 4.6.

Table 4.7: User Profiles and User Emotion Group of 12 Listeners

Listener	Gender	Major in Music	User Emotion Group
L1	Male	Yes	C1
L2	Male	Yes	C1
L3	Male	Yes	C1
L4	Male	Yes	C1
L5	Male	No	C2
L6	Male	No	C2
L7	Female	No	C2
L8	Female	No	C2
L9	Female	Yes	C3
L10	Female	Yes	C3
L11	Female	Yes	C3
L12	Female	Yes	C3

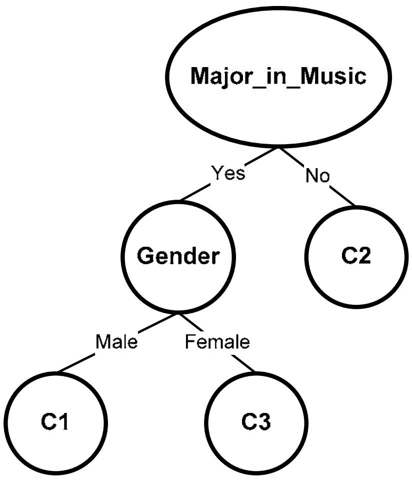

Figure 4.4: User Group Classification Tree

4.4 Music Emotion Classification

After the user emotion groups are constructed, at the same time, the task of music emotion classification will be performed. That is to say, we want to find out music with which kind of music attributes values or music patterns would results in which kind of emotion. Here we apply the C4.5 method to perform this task.

C4.5 was proposed by Quinlan [21]. He is also the inventor of ID3. Unlike ID3, C4.5 is suitable to classify the data with numerical, categorical and boolean attributes. But for ID3, only the categorical and boolean attributes could be input into the classifier. If there is numerical attribute, it should firstly be divided into several segments. Then the ID3 could be applied. But here for the music attributes such as pitch mean, pitch standard deviation, interval mean, etc., we don't want to initially divide the range into several segments because no one knows where the rational boundaries are. So we apply the C4.5 algorithm to perform the music prediction.

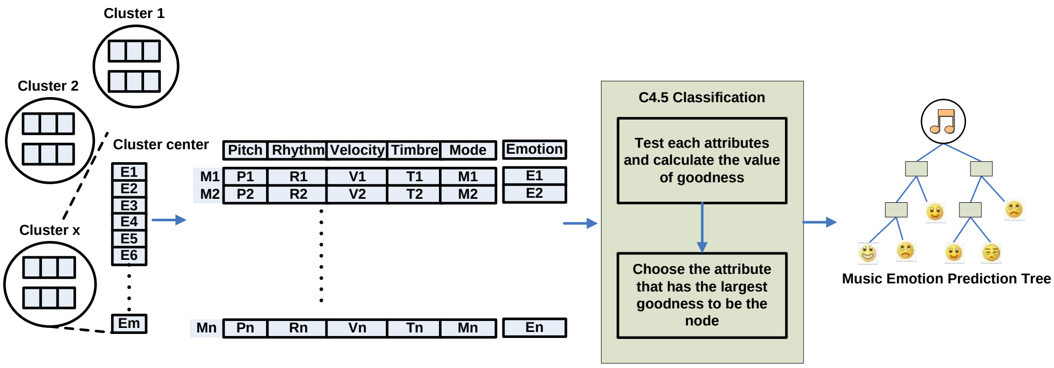

Figure 4.5: Music Emotion Classification

Figure 4.5 shows the step of music emotion classification. In section 4.2, we have gotten the number of cluster of these m listeners by clustering. Before applying C4.5 to find out the music emotion prediction rules for each user emotion group, we should firstly get the emotion center of each user emotion group. When determining the emotion center, we could roll up in the emotion concept hierarchy layer. For example, we could transform the emotion in layer C (Reilly's model) to layer B (Thayer's model) or directly to layer A (Positive or Negative). The emotion center of each user emotion group is determined as the majority of all listeners' music emotions in every user emotion group. For example, if there are four listeners in user emotion group 1 who are L1, L2, L3 and L4. The total number of music is five. They are M1, M2, M3, M4 and M5. For M2, the emotions of these four listeners are E1, E2, E1, E2 and E1 respectively. The emotion center of user emotion group for M2 is E1. So the input in this phase would be music attributes and emotion center of each user emotion group. We then apply the C4.5 algorithm to each user emotion group to generate the emotion prediction rules for each user emotion group. In C4.5, there is also a goodness index to determine the classified node. Attribute with the largest goodness value would be selected as the node in every step. The algorithm of music emotion prediction is shown in Algorithm 4.3.

Algorithm 4.3: Music Emotion Classification Algorithm (C4.5 Algorithm)

Purpose: Find out the relationship between music attributes and music emotion

Input: Music Attributes of n representative music (MA), Emotion of n representative music of m listeners (ME)

Output: Music Emotion Prediction Rules (MEPR)

Step1: MA and ME is the set of training instances and let it be T

Step2: Choose an attribute that best differentiates the instances contained in T. That is to say to choose an attribute with the largest gain ratio.

Step3: Create a tree node whose value is the chosen attribute. Create child links from this node where each link represents a unique value for the chosen attribute. Use the child link values to further subdivide the instances into subclasses.

Step4: For each subclass created in step 3:

4.1 If the instances in the subclass satisfy predefined criteria or if the set of remaining attribute choices for this path of the tree is null, specify the classification for new instances following this decision path.

4.2 If the subclass does not satisfy the predefined criteria and there is at least one attribute to further subdivide the path of the tree, let T be the current set of subclass instances and return to step 2.

Step5: Output the MEPR.

Example 4.4: Music Emotion Classification

Following Example 4.3, we have gotten the number of cluster of these 12 listeners by clustering. And the emotion centers of user emotion group C1, C2 and C3 are listed in Table 4.8. Here we select the emotion of layer B in the proposed emotion concept hierarchy. That is to say we roll up the emotion model from layer C to layer B. Then we apply the C4.5 algorithm to each user emotion group. We could get the music emotion prediction tree for each user emotion group which is shown in Figure 4.6.

Table 4.8: Emotion Center of Each User Emotion Group

	M1	M2	M3	M4	M5	M6	M7	M8	M9	M10	M11	M12	M13	M14	M15	M16
L1	2	2	2	2	1	1	1	1	3	3	3	3	4	4	4	4
L2	2	2	2	2	1	1	1	1	3	3	3	3	1	4	4	4
L3	2	2	1	4	1	1	4	1	3	3	3	4	4	4	4	4
L4	2	1	2	2	1	1	1	1	3	3	3	3	4	2	4	4
C1	2	2	2	2	1	1	1	1	3	3	3	3	4	4	4	4
L5	3	3	3	3	4	4	4	4	2	2	2	2	1	1	1	1
L6	3	3	3	3	4	4	4	4	2	2	2	2	1	1	1	1
L7	3	3	3	3	4	4	4	4	2	2	2	2	1	1	1	1
L8	3	3	3	3	4	4	4	4	2	2	2	2	1	1	1	1
C2	3	3	3	3	4	4	4	4	2	2	2	2	1	1	1	1
L9	4	4	4	4	3	3	3	3	1	1	1	1	2	2	2	2
L10	4	4	4	4	3	3	3	3	1	1	1	1	2	2	2	2
L11	4	4	4	4	3	3	3	3	1	1	1	1	2	2	2	2
L12	4	4	4	4	3	3	3	3	1	1	1	1	2	2	2	2
C3	4	4	4	4	3	3	3	3	1	1	1	1	2	2	2	2

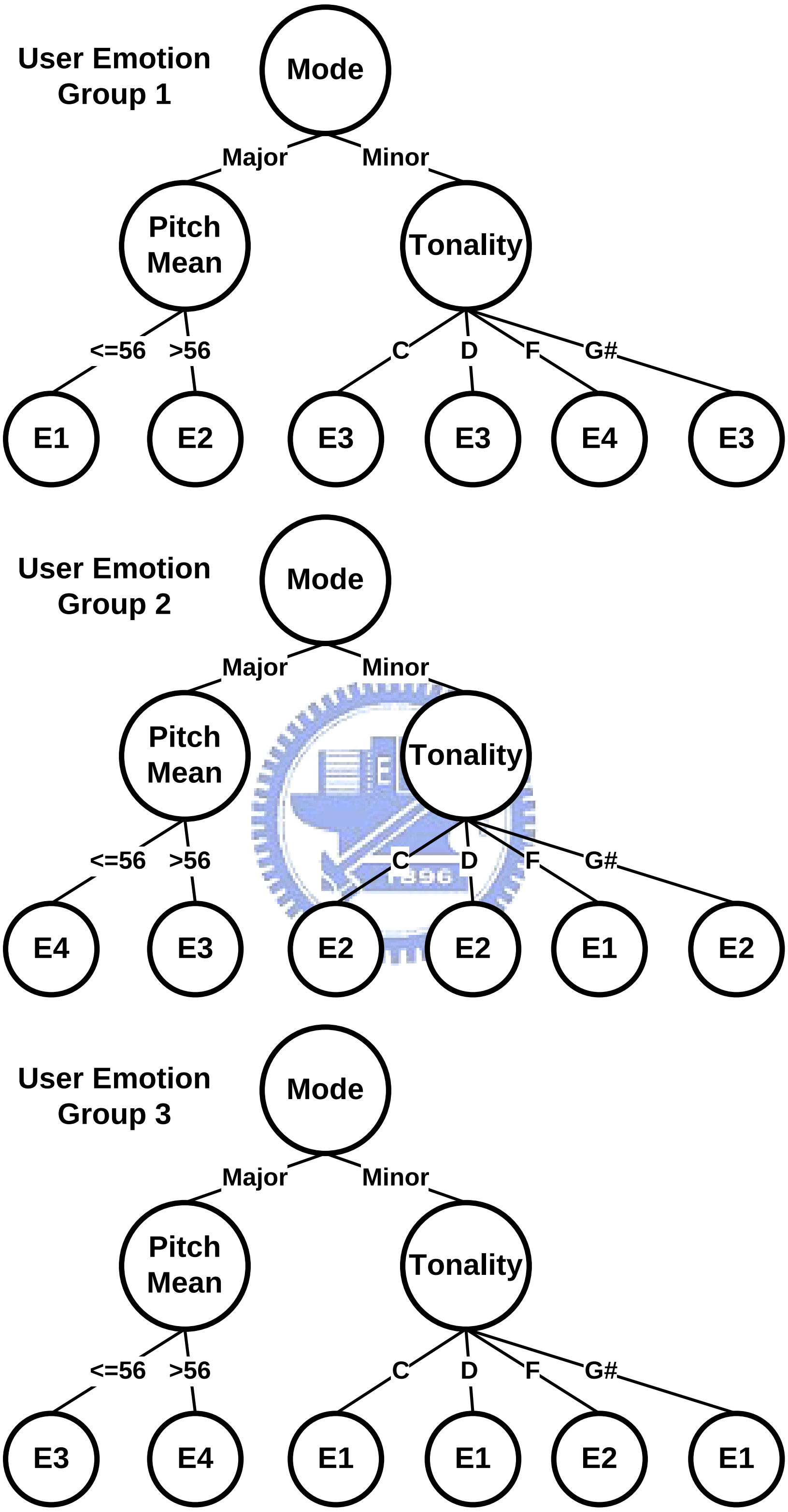

Figure 4.6: Music Emotion Prediction Tree

4.5 Personalized Music Emotion Prediction Rules Integration

After the previous phases finished, we got one decision tree for user group classification and decision tree for music emotion prediction for each user emotion group. Finally, we apply the rules integration method to integrate these two kinds of rules to become the personalized music emotion prediction rules. Then the rules would be combined with an inference engine to perform the music emotion prediction concerning with user differences. The procedure is shown in Figure 4.7 and the algorithm is shown in Algorithm 4.4.

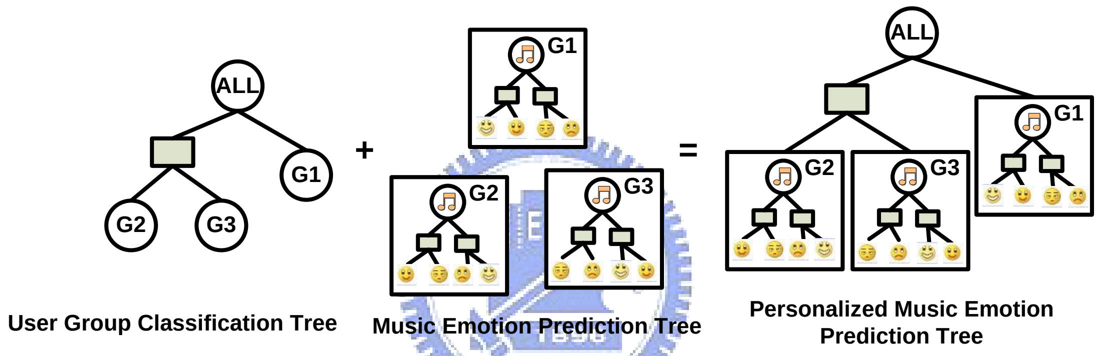

Figure 4.7: Personalized Music Emotion Prediction Rules Integration

Algorithm 4.4: Personalized Music Emotion Prediction Rules Integration Algorithm
Purpose: Integrate the rules that are generated in user group classification phase and music emotion prediction phase.
Input: User Group Classification Rules (UGCR) and Music Emotion Prediction Rules (MEPR)
Output: Personalized Music Emotion Prediction Rules (P-MEPR)
Step1: For each user emotion group
Merge both rules
Step2: Output the P-MEPR

Example 4.5: Personalized Music Emotion Prediction Rules Integration

Following Example 4.4, after we merge the user group classification tree and the music emotion prediction tree in Figure 4.4 and 4.6. The outcome is the personalized music emotion prediction tree that is shown in Figure 4.8. And the rules are listed in Figure 4.9.

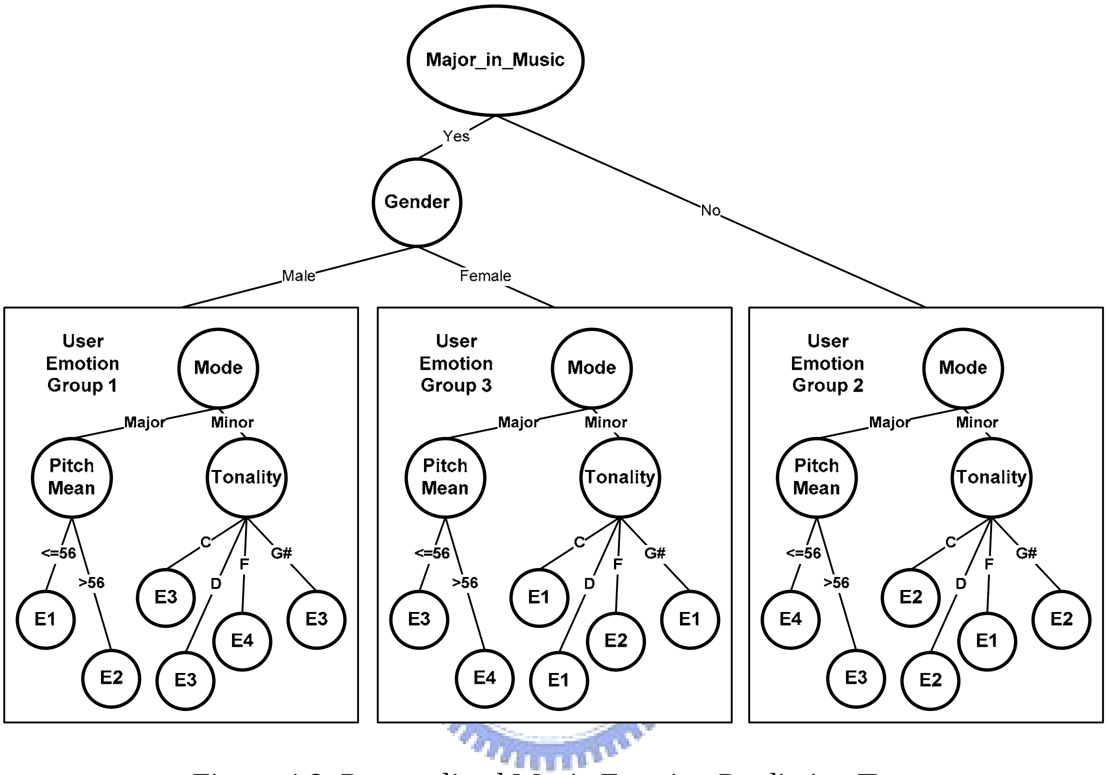

Figure 4.8: Personalized Music Emotion Prediction Tree

Rule 1: If *Major_in_Music* = *yes* and *Gender* = *male*, then *cluster* = *C1*

Rule 1_1: If *Mode* = *major* and *Pitch_Mean* <= *56*, then *emotion* = *E1*

Rule 1_2: If *Mode* = *major* and *Pitch_Mean* > *56*, then *emotion* = *E2*

Rule 1_3: If *Mode* = *minor* and *Tonality* = *C*, then *emotion* = *E3*

Rule 1_4: If *Mode* = *minor* and *Tonality* = *D*, then *emotion* = *E3*

Rule 1_5: If *Mode* = *minor* and *Tonality* = *F*, then *emotion* = *E4*

Rule 1_6: If *Mode* = *minor* and *Tonality* = *G#*, then *emotion* = *E3*

Rule 2: If *Major_in_Music* = *no*, then *cluster* = *C2*

Rule 2_1: If *Mode* = *major* and *Pitch_Mean* <= *56*, then *emotion* = *E4*

Rule 2_2: If *Mode* = *major* and *Pitch_Mean* > *56*, then *emotion* = *E3*

Rule 2_3: If *Mode* = *minor* and *Tonality* = *C*, then *emotion* = *E2*

Rule 2_4: If *Mode* = *minor* and *Tonality* = *D*, then *emotion* = *E2*

Rule 2_5: If *Mode* = *minor* and *Tonality* = *G#*, then *emotion* = *E1*

Rule 2_6: If *Mode* = *minor* and *Tonality* = *F*, then *emotion* = *E2*

Rule 3: If *Major_in_Music* = *yes* and *Gender* = *female*, then *cluster* = *C3*

Rule 3_1: If *Mode* = *major* and *Pitch_Mean* <= *56*, then *emotion* = *E3*

Rule 3_2: If *Mode* = *major* and *Pitch_Mean* > *56*, then *emotion* = *E4*

Rule 3_3: If *Mode* = *minor* and *Tonality* = *C*, then *emotion* = *E1*

Rule 3_4: If *Mode* = *minor* and *Tonality* = *D*, then *emotion* = *E1*

Rule 3_5: If *Mode* = *minor* and *Tonality* = *G#*, then *emotion* = *E2*

Rule 3_6: If *Mode* = *minor* and *Tonality* = *F*, then *emotion* = *E1*

Figure 4.9: Personalized Music Emotion Prediction Rules

Through Sections 4.1 to 4.5, we have applied a series of data mining technique including ROCK, ID3 and C4.5 to generate the personalized music emotion prediction rules that is shown in Figure 4.9. According to the rules, listeners in different user emotion group have different emotional response to music. For example, if a listener major in music and is female, then she is clustered into C3 according to Rule 3. Now she listens to music in C minor, she will feel Exuberance (E1) according to Rule 3_3. During the application phase, these rules are used to predict music emotion of listeners. Firstly, the listener is clustered into their user emotion group according to Rule 1, 2 or 3. Then the sub rules are used to predict their emotional response according to the value of music.

Chapter 5 System Implementation and Experiment

The personalized music emotion prediction rules are generated from a series of data mining procedure that have been described in previous section in detail. Now the task would be building the personalized music emotion prediction system. The personalized music emotion prediction rules will be embedded in the system to inference the emotional response to music of a listener with his or her user profiles.

5.1 The P-MEP System

The main page of the system is shown in Figure 5.1. During the application procedure, a new user comes. The user should firstly register an account in the system which is shown in Figure 5.2. Then the system will show a questionnaire for the user to fill in the user profiles which is shown in Figure 5.3. After the user submitting the questionnaire, the system will inference the user emotion group which the user belongs to according to the user profiles. The rules are the set of user group classification rules which are generated in user group classification phase. The user now has finished the register procedure.

After the user finished the register procedure, they could use the system to predict the emotional response to music. The interface of inputting the music midi file is shown in Figure 5.4. Then the system will predict the emotion according to the music prediction rules which are generated in music emotion prediction phase. Figure 5.5 shows the result of emotion prediction.

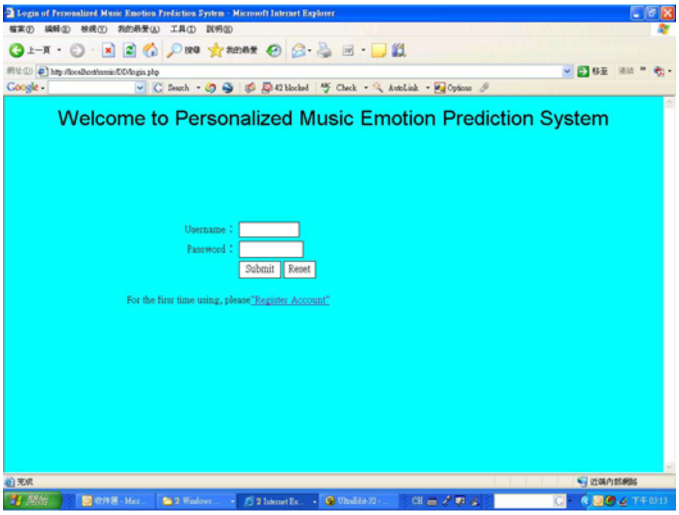

Figure 5.1: The Main Page of P-MEP System

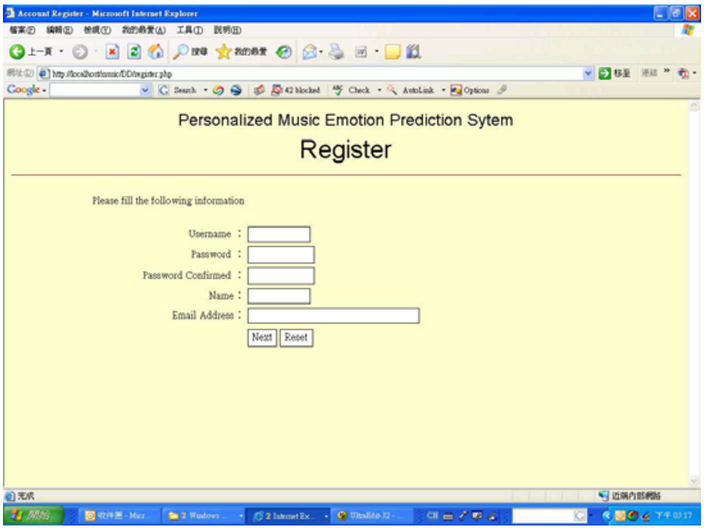

Figure 5.2: The Registration Form of P-MEP System

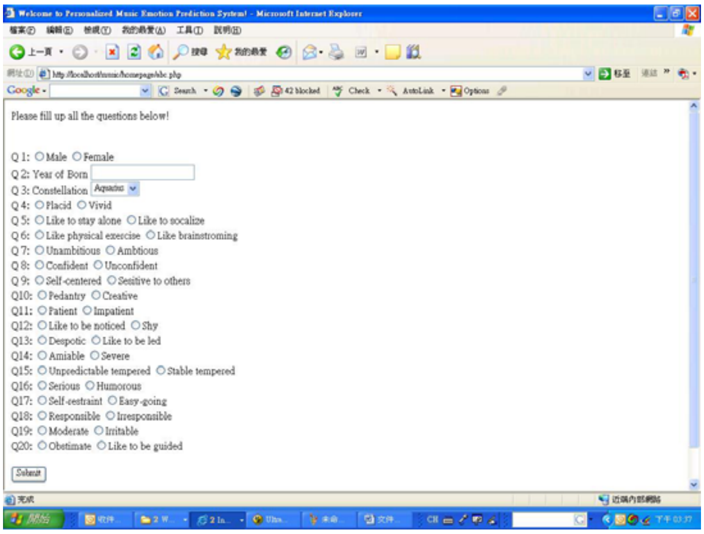

Figure 5.3: User Profiles to be Filled

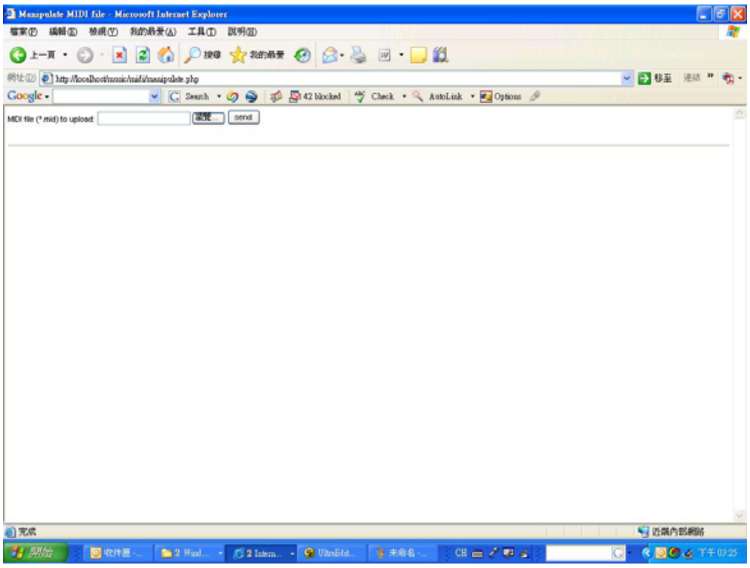

Figure 5.4: The Form to Load Midi File

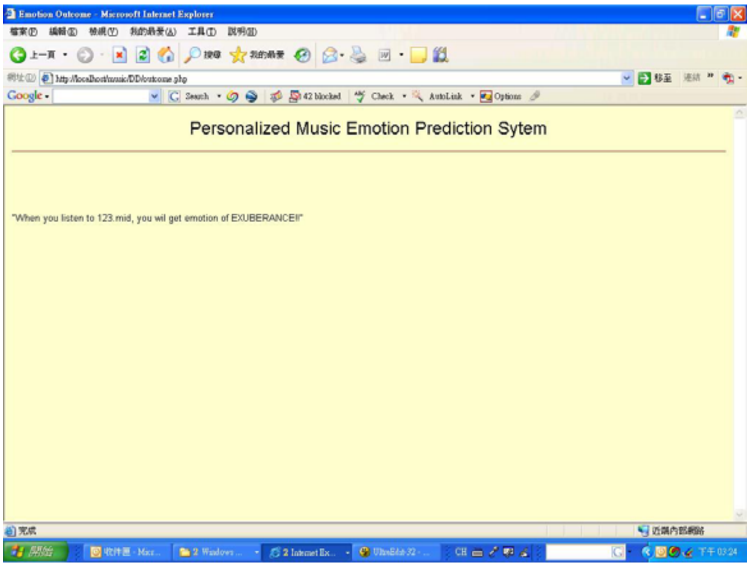

Figure 5.5: The Predicted Music Emotion

5.2 Experiment

In the experiment, 20 pieces of representative music midi files are selected by the expert. All the music is cut into short clips about 1 minute. 24 listeners with different backgrounds join this experiment to annotate the music emotion in Reilly's model to each music clip. In these twenty-four listeners, 15 are male and 9 are female. These 20 music attributes vectors, 24 music emotion vectors and 24 user profiles vectors are the training data. The threshold of similarity is set to be 0.6. The results of the experiments are shown in Figure 5.6.

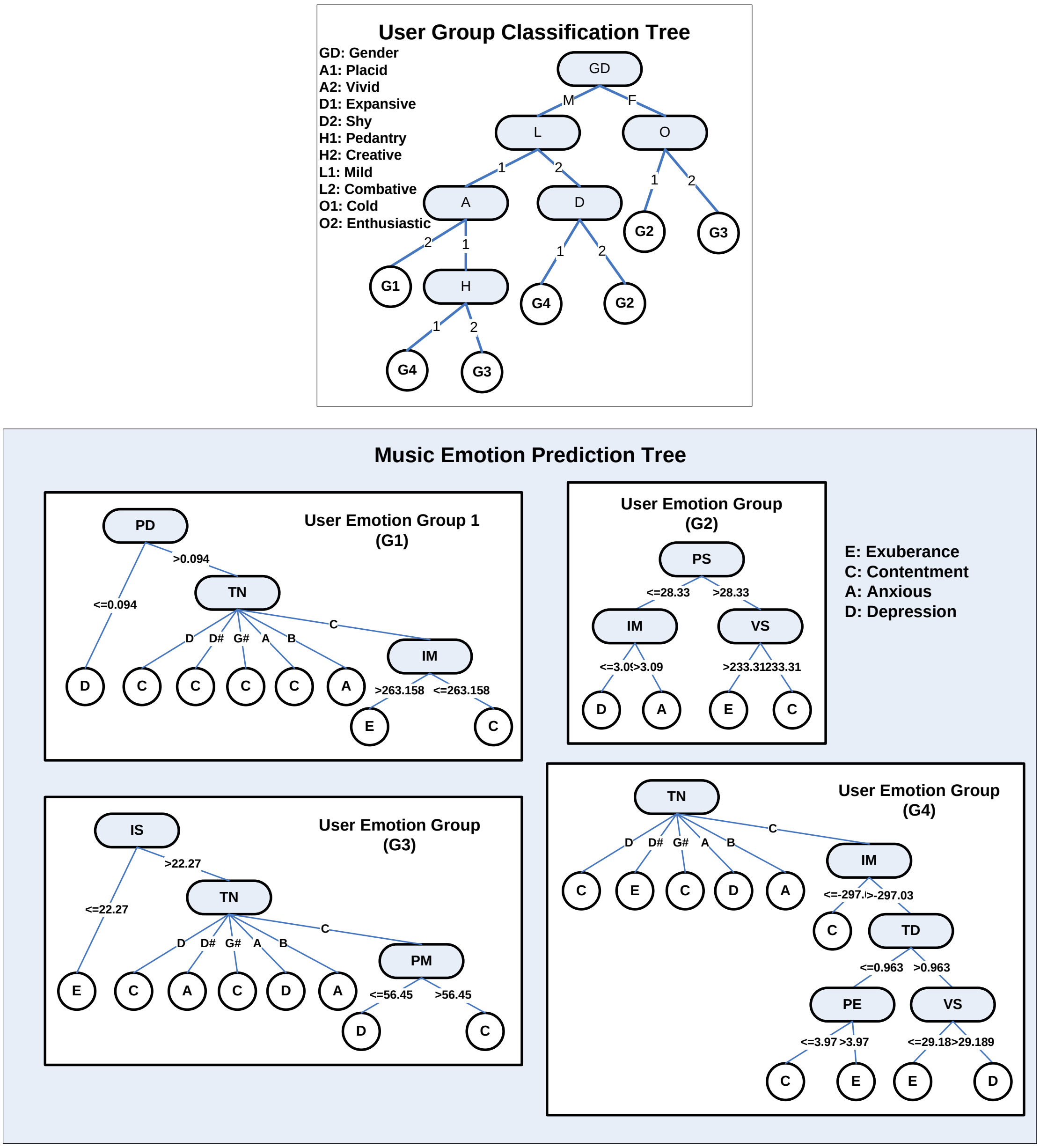

Figure 5.6: The Results of the Experiment

Then another 10 listeners including 5 male and 5 female listeners are invited to annotate 4 different music clips. They are the testing data set. And the accuracy of the experiment is 70%.

In Figure 5.6, there are four listeners in user emotion group 1 (G1) and they are all male. There are five listeners in user emotion group 2 (G2), while 2 are male and 3 are female. There are seven listeners in user emotion group 3 (G3), while 2 are male and 5 are female. Finally, there are eight listeners in user emotion group 4 (G4), while 1 is female and the rest are male. Listeners in G1 have the personality that is vivid, listeners in G2 have the personality that is shy and cold, listeners in G3 have the personality that is enthusiastic if the listener is female and creative if the listener is male, and finally listeners in G4 have the personality that is pedantry or straightforward.

The listeners in G1 have positive emotion to music when the pitch density of the music is greater than 0.094 and the tonality is C, D, D#, G# and A, and have negative emotion when the pitch density of the music is less that 0.094 or the pitch density is greater that 0.094 and the tonality is B. The listeners in G2 have positive emotion when the standard deviation of pitch of music is greater than 28.33 and have negative emotion when it is less than 28.33. The listeners in G3 have positive emotion when the standard deviation of interval is less than 22.27 or when it is greater than 22.27 and the tonality is D, G# or B and have negative emotion in the rest cases. The listeners in G4 have negative emotion when the tonality of music is A or B and when mean of interval is greater than 297, tempo degree is greater than 0.96 and velocity standard deviation is greater than 29.18 and have positive emotion for the rest cases.

Chapter 6 Concluding Remarks

With the rapid development of multimedia technology, researchers now try to study the relationship between music and emotion and the music emotion prediction system is needed. In this thesis, the proposed P-MEP analysis procedure consisting of data mining techniques are used to analyze the listeners' music emotion concerning to their background differences and the personalized music emotion prediction rules are generated. In addition, the P-MEP system is built based on the rules. The result of the experiment shows that our personalized music emotion prediction system can be used to predict music emotion of listeners concerning with their different backgrounds.

In the near future, we will add more attributes to user profiles and music attributes such as culture related attributes to get more information about listeners and music. In addition, collecting more data such as more listeners and different kinds of music such as percussion music will improve the accuracy of the generated rules. Furthermore, the music audio files could be used to improve the research by extracting the music attributes that described in previous section and the music emotion concept hierarchy could be modified to match the emotion which composers used in writing music scores. The research can further focus on the difference of perceived emotion between people who major in music or not. Finally, in the P-MEP system, the mechanism of user feedback could be added to improve the system and the incremental mining technique should be used. In application, the system can be used to predict patients' emotions when applying in music therapy or to prompt some products to customers by broadcasting the music that will arouse the willing to shop.

References

[1] "The biology of music," *The Economist,* Feb 12-18. 2000.

[2] P. Cano, M. Koppenberger and N. Wack, "Content-based music audio recommendation," in *Proceedings of the 13th Annual ACM International Conference on Multimedia,* 2005, pp. 211-212.

[3] W. Chai and B. Vercoe, "Folk music classification using hidden markov models," in *Proceedings of International Conference on Artificial Intelligence,* 2001,

[4] H. C. Chen and A. L. P. Chen, "A music recommendation system based on music and user grouping," *Journal of Intelligent Information Systems,* vol. 24, pp. 113-132, 2005.

[5] H. C. Chen, C. H. Lin and A. L. P. Chen, "Music segmentation by rhythmic features and melodic shapes," in *Proceedings of the 2004 IEEE International Conference on Multimedia and Expo,* 2004, pp. 1643-1646.

[6] S. Guha, R. Rastogi and K. Shim, "ROCK: A robust clustering algorithm for categorical attributes," *Information Systems,* vol. 25, pp. 345-366, 2000.

[7] J. Han and M. Kamber, *Data Mining :Concepts and Techniques.* San Francisco: Morgan Kaufmann Publishers, 2001, pp. 550.

[8] K. Iwahama, Y. Hijikata and S. Nishida, "Content-based filtering system for music data," in *Proceedings of the 2004 International Symposium on Applications and the Internet Workshops,* 2004, pp. 480-487.

[9] P. N. Juslin and J. A. Sloboda, *Music and Emotion :Theory and Research.* Oxford ;New York: Oxford University Press, 2001, pp. 487.

[10] J. L. Koh and D. C. Yu, "Efficient feature mining in music objects," in *Lecture Notes in Computer Science* , vol. 2113, H. C. Mayr, J. Lazansky, G. Quirchmayr and P. Vogel, Eds. Berlin / Heidelberg: Springer, 2001, pp. 221-231.

[11] F. F. Kuo, M. F. Chiang, M. K. Shan and S. Y. Lee, "Emotion-based music recommendation by association discovery from film music," in *Proceedings of the 13th Annual ACM International Conference on Multimedia,* 2005, pp. 507-510.

[12] F. F. Kuo and M. K. Shan, "Looking for new, not known music only: Music retrieval by melody style," in *Proceedings of the 2004 Joint ACM/IEEE Conference on Digital Libraries,* 2004, pp. 243-251.

[13] F. F. Kuo and M. K. Shan, "A personalized music filtering system based on melody style classification," in *Proceedings of the 2002 IEEE International Conference on Data Mining,* 2002, pp. 649-652.

[14] T. Li and M. Ogihara, "Music genre classification with taxonomy," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing,* 2005, pp. 197-200.

[15] C. R. Lin, N. H. Liu, Y. H. Wu and L. P. Chen, "Music classification using significant repeating patterns," in *Lecture Notes in Computer Science* , vol. 2973, Y. J. Lee, J. Z. Li, K. Y. Whang and D. Lee, Eds. Berlin / Heidelberg: Springer, 2004, pp. 506-518.

[16] N. H. Liu, Y. H. Wu and L. P. Chen, "An efficient approach to extracting approximate repeating patterns in music databases," in *Lecture Notes in Computer Science* , vol. 3453, L. Zhou, B. C. Ooi and X. F. Meng, Eds. Berlin / Heidelberg: Springer, 2005, pp. 240-251.

[17] N. H. Liu, Y. H. Wu and L. P. Chen, "Efficient K-NN Search in Polyphonic Music Databases Using a Lower Bounding Mechanism," *Multimedia Systems,* vol. 10, pp. 513-528, 2005.

[18] L. Lu, D. Liu and H. J. Zhang, "Automatic mood detection and tracking of music audio signals," *IEEE Transactions on Audio, Speech and Language Processing,* vol. 14, pp. 5-18, Jan. 2006.

[19] N. C. Maddage, "Automatic structure detection for popular music," *Multimedia, IEEE,* vol. 13, pp. 65-77, 2006.

[20] N. C. Maddage, C. Xu, M. S. Kankanhalli and X. Shao, "Content-based music structure analysis with applications to music semantics understanding," in *Proceedings of the 12th Annual ACM International Conference on Multimedia,* 2004, pp. 112-119.

[21] J. R. Quinlan, *C4.5 :Programs for Machine Learning.* San Mateo, Calif.: Morgan Kaufmann Publishers, 1993, pp. 302.

[22] Reilly, W. S. N., "Believable Social and Emotion Agents," 1996.

[23] P. Y. Rolland and J. G. Ganascia, "Pattern detection and discovery: The case of music data mining," in , vol. 2447, D. J. Hand, N. M. Adams and R. J. Bolton, Eds. Berlin / Heidelberg: Springer, 2002, pp. 190-198.

[24] M. K. Shan, F. F. Kuo and M. F. Chen, "Music style mining and classification by melody," in *Proceedings of the 2002 IEEE International Conference on Multimedia and Expo,* 2002, pp. 97-100.

[25] R. E. Thayer, *The Biopsychology of Mood and Arousal.* Oxford University Press, 1989,

[26] M. Wang, N. Zhang and H. Zhu, "User-adaptive music emotion recognition," in *Proceedings of the 2004 7th International Conference on Signal Processing,* 2004, pp. 1352-1355.

[27] P. Winsor, *Automated Music Composition.* Denton, Tex.: University of North Texas Press, 1989, pp. 312.

[28] A. Ben-Tovim 與 D. Boyd 著，發掘兒童音樂潛能：如何選擇樂器，陳軍譯，世界文物，台北市，初版，民國 86 年。