Music Information Retrieval systems: why do individuals use them and what are their needs?

S. Taheri-Panah

Centre for Interactive Systems
Research
Department of Information Science
City University, London

A. MacFarlane

Centre for Interactive Systems
Research
Department of Information Science
City University, London
Email: andym@soi.city.ac.uk*

ABSTRACT

Abstract: To date there has been very little research conducted on the behaviour of music information retrieval (MIR) users, in spite of the immense popularity of free music retrieval systems available on the Internet. In this study we examine the issue of music seeking behaviour through the examination of users life style effect of three different age groups using questionnaires. It was found that lifestyles had a significant impact on users need for music and hence their music seeking behaviour. The importance of social networks in music information seeking was reinforced in this study. An experiment was conducted with three different types of search on the Kazaa MIR system and the participants interviewed in order to collect data. Users found the Kazaa system intuitive and easy to use. Searchers used both song titles and lyrics for finding relevant music items. The insights provided by this study can be of assistance in the development of user focused Internet MIR systems.

Keywords: user music seeking behaviour retrieval

1. INTRODUCTION

More than ever music has become an important and prominent part of our culture with each individual having their own preference for the type of music they listen to, how they retrieve it and where they listen to it. There are many ways in which individuals listen to music; some of these are socially in a bar with friends, whilst exercising at the gym, alone at home etc. The introduction of Internet services such as Napster [1] and Kazaa [2] has given users the ability to gain access to large amounts of digital music previously inaccessible other than by purchase in Music stores/shops. The free availability of music recently means that the music industry has lost its control to some extent over the release and distribution of music. Some of the causes for this growth are evident and self-explanatory e.g. financial reasons. We can identify the availability of free music as one of the main reasons for their success, but there are many more. Advances in the underlying technology and the

infrastructure of both the Internet and computer machinery has made the availability of multi media as prevalent as text documents [3]. The scale of the effect these factors are having on music seeking behaviour is illustrated by the Pew Internet & American Life Project [4]. Through their research they reported that from the internet activities offered to internet users, music retrieval has by far excelled other activities in both its uptake and in its usage e.g. it was reported that in 2000 just over 1.1 million people had used Napster [1], increasing by 506.8% to 6.7 million just six months later. It is this dramatic change in so many people's music retrieval habits which became the driving force for this study, with the focus on the music seeker needs and opinions.

Some studies have already been conducted on music information retrieval on the Internet, but the majority of these have only considered the systems perspective. Although Music Information Retrieval (MIR) have developed considerably over the past few years, there are many aspects of them that require further investigation [5]. An important aspect of this whole phenomenon which has been somewhat neglected is the normal MIR user. These are individuals who are not music experts or professionals: we concentrate on this user group in this paper. The paper has the following structure. Section 2 describes the research aims and objectives. In section 3 we outline the Kazaa system used for data collection purposes. Section 4 describes the research design, while the results from the evaluation are discussed in section 5. A conclusion is given at the end.

2. RESEARCH AIMS AND OBJECTIVES

There is a paucity of research in the study of MIR systems from a user's perspective. With the availability of new music formats and retrieval methods on the Internet, music seekers have many more options available to them. A study addressing the needs of music seekers behaviour is therefore essential. The focus of this study is to identify and understand the needs of the music seeker, through a series of questionnaires and interviews. The aim is to give a better understanding of music seekers and their music retrieval needs. In this study there are two central issues that will be addressed. The first is how, when and why people use music. The second is how well are the needs of music seekers met by these Internet MIR systems.

The aim of this study is to learn what music seekers are looking for and how they perform search.
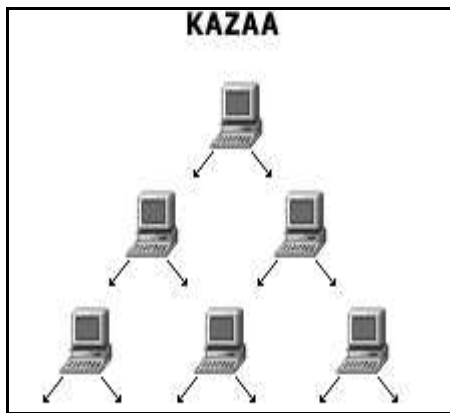
3. THE KAZAA SYSTEM

Figure 1: The Kazaa system Architecture [6]

We use the Kazaa System [2] for our data collection purposes. Kazaa is a file sharing peer to peer network system which is deployed for MIR. A Kazaa server has its own set of files and search facility. The users of Kazaa have access to many types of files including music files, pictures, videos, software and even text documents, thus Kazaa is a multi media retrieval system [2]. Each Kazaa user's computer itself acts as a search engine and locates the file in the computer of other Kazaa users. The system designates some users who have more powerful computers as a supernode. Once a Kazaa user enters a query the request is forwarded to a supernode which firstly conducts an internal search for the requested file in its own archives. If the file is located in the supernode the seeking computer is informed and the item retrieved. If the supernode does not have the file the query is forwarded to other computers connected to the Kazaa system. Once the file has been found the seeking computer is informed through the string of computers the request passed through. The architecture of this system is shown in figure 1. We chose to use this system as Kazaa Media Desktop has been reported to be one of the most popular MIR since Napster was shut down.

4. EXPERIMENTAL DESIGN

We acquired primary data for this study by using questionnaires and interviews. The following gives an overview of the design of these elements for the experiment.

4.1. Questionnaires

The purpose of the questionnaire was to acquire more information about music and associated issues. There were four main issues that were addressed by the questions. These were:

- The methods of music retrieval and collection people use.
- Music usage and its use in conjunction with other activities.
- Music retrieval system (MIR) usage.
- The needs of the users

To ensure the questionnaires were efficient and informative, a pilot questionnaire was conducted on 10 individuals. These questionnaires were given to individuals randomly in a variety of locations. The results from the pilot questionnaires were used to improve and produce a final version of the questionnaire to use in the study. This research was conducted in July and August of 2003. The revised questionnaires were given to 45 individuals to complete in a number of locations including universities, libraries and coffee shops. To produce a fair and accurate study the numbers in each group and the ratio of males to females was kept as balanced as possible. A summary of the individuals who participated in this study is given in table 1.

Group	Age Range	Total No	No of females	No of males
A	19 and under	17	8	9
B	20 – 39	15	8	7
C	40 and over	13	7	6

Table 1: The age groups used in the questionnaire and the number of individuals in each group.

4.2. Interviews

Pre interview searches were conducted in the researchers working environment, with access to Kazaa. The researcher remained present to assist the user and help the participants with Kazaa. This research was conducted to provide a more qualitative set of information for the research. In the experiment, 10 participants were asked to carry out three tasks on Kazaa. Each participant was given a worksheet to record their results. Three tasks were given to the user:

- Known item (or song) searches.
- Search of new and old songs from the same music genre.
- Search of songs from different genres.

The researcher went through each task on the worksheet before the observation session started. The participants were given the option to look for artist's names and song titles using Internet search engines if they could not think of any themselves. Once the search tasks were complete the participants were interviewed about their experience with the Kazaa MIR.

5. EXPERIMENTAL RESULTS

5.1. Analysis of the user group
The data in table 1 above has been combined with the information regarding the amount of free time available for listening to music (see figure 2).
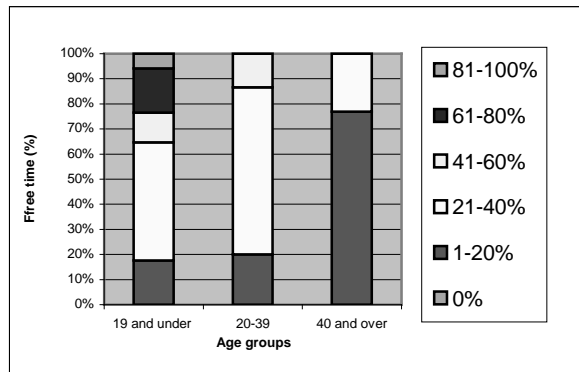

Figure 2: Relationship between an individual's age and their free time listening to music.

A number of interesting observations about this group can be made from this data. Firstly as age increases less time is spent listening to music. Interestingly in the 19 and under age group we see a wide spread of data. This suggests that the individuals within the older age groups have similar habits, particularly with regards to music. Looking more generally at the music usage across all ages we can infer that the younger music users have more flexibility with their time. For the older participants on the other hand, time is restricted because of the extra responsibilities they have. This shows that the younger participants in our study have more time available to them and can therefore spend more of it on activities such as listening to music. Listening to music as an activity was also compared to other activities. Here it was found that in the 19 and under age group, listening to music was preferred in all of the other categories except for watching TV and socialising with friends. This changed in the 20-39 age group, where the only preferred category to music was watching TV.

For the 40 and over group we see a slight shift to watching TV and socialising with friends being preferred over music. There is also a fondness and preference for reading in this group: music was consistently rated at the third or fourth most popular Internet activity across all the age groups. Interestingly the 19 and under and the 40 and over group both rated communications as their top Internet activity. In contrast purchasing goods or services was the most popular Internet activity for the 20-39 age group. From this we can only speculate that at this point in life, the developing careers and lifestyle of this age group contributes to a more hectic schedule than the other two age groups. It is possible that more of these individuals are likely to purchase goods using the Internet since they may not have a lot of spare time.

Considering the life styles of the other two groups, we would expect them to be more settled. Consequently they have more time to dedicate to less important Internet activities as in this case for communication purposes. The Internet activities of these individuals may perhaps be connected to the level of computer knowledge and experience of each age group (see figure 3). This clearly shows how the participants of the age groups are distributed in accordance of their IT literacy. The majority of the beginners were the 19 and under group followed by the 40 and over. This is plausible since many 20-39 year olds use computers on a daily basis for work. This would give them the ability to perform more complex activities using their computers and the Internet.
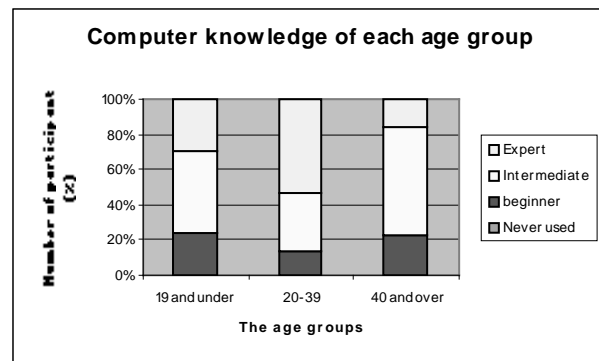

Figure 3: The level of computer knowledge across the age groups.

5.2. Analysis of the participant's musical requirements
Our results illustrated the relevance and importance of music in all age groups. We found that a high percentage of the participants in this study pay particular attention to the lyrical contents of songs. The majority of people across the age groups used music to fulfil two common categories; music that makes you feel happy and music that masks problems. However the majority of the younger individuals, including the 19 and under and the 20-39 age groups also listened to music to make them feel energetic whilst their older counterparts used it for relaxation. This is credible if we consider the difference in the social activities of the two younger groups: the majority of these participants said that they mainly listen to music outside of their homes. The explanation suggested for this lies with the social activities of these age groups, where going to the gym, bars and clubs is more common.

It was found that an influential factor on people's music choice is in the form of TV and radio. Friends were also found to be an influence since music is used for many social activities for both young and old: the majority of our participants said that they listen to music in the company of others. The influence friends have on each others music taste is considerable: overwhelming 96% of the participants said that they discuss music with their friends. With these findings we

can assert that people inform each other about music that they purchase, listen to and possibly the music retrieval methods they use. In order to examine this issue in more detail, the participants were asked to estimate the number of their friends who they share their music taste with. The majority of the participants across all of the age groups, in particular the 19 and under group, shared their music taste with most of their friends.

5.3. Analysis of the experiment and interviews

The majority of the individuals who participated in the observations and completed the mini-questionnaire had used MIR before. Altogether around 80% of these individuals claimed to have used MIRs, with Kazaa being the most popular. This is important since the effective use of these systems depends on who the expected users are and their level of knowledge [7]. After having used the Kazaa MIR in this observation most people said they would speak of it favourably, irrespective of prior usage. Users were able to use Kazaa with little or no training. The new users of Kazaa were inquisitive and did enquire about some of the more unique features offered by this system, such as the theatre function. Although when the user becomes aware of the availability of images and videos, this turns out to be self-explanatory. The participants were asked to rate some of the features of the system from very important to very unimportant (with the values of 5 to 1 respectively). From this a score was then calculated in accordance to how all of the individuals rated the features. With this score it is easier to determine the most important feature for all of the participants. The evidence from calculated scores showed that the speed of MIR along with the quality of the music are the most important features to the participants. Unfortunately neither one of these features is controlled by the Kazaa system. The speed depends primarily on the type of Internet connection the Kazaa users are using. The quality of music depends on, what type of files the users share on the system through their computer. A good search facility was also considered to be important by the participants; this is logical since a good search engine is necessary in locating files on the system. This is a feature that can be controlled by the system administrators and they can make it as flexible or as precise as they want. There are benefits to both of these; a more precise search engine will return fewer but more accurate files. In comparison a more flexible search engine will be able to locate slightly different variations of the search term, giving the MIR user the opportunity to find items that have been spelt incorrectly. In all these instances a rigid search engines would intentionally block out items user may be interested in.

It was found that the most preferred and used type of query by these individuals was by using the titles of songs, which was followed by using the artist's name. Surprisingly some individuals said that they used lyrics from songs to find a specific item. This is important to explore in further studies since occasionally individuals may not know or remember the details of a particular song [8]. This suggests that those who share items and those who look for items, pick up on and enter the same parts of a song instead of the title. A particular feature that was favoured by the participants was that this system gave its users the ability to look through the items of the users connected to the network. The explanation for this is that if an individual has one music item that a music seeker likes and wants, then it is likely that they have a few more (that are not being directly searched for by the music seeker). This brings back the social issues in music sharing. Most of the participants mentioned that they thought the results of the advanced search were not as good or as accurate as they had expected. The normal quick search was found to be more efficient. When asked about the provision of additional features, the participants did not think that supplying information on these systems is necessary for the most part. However, the participants did suggest that information on artists would be useful.

6. CONCLUSIONS AND FURTHER RESEARCH

By identifying the importance of music to users, the issues surrounding music retrieval have been explored. From this research, it was found that different people used it for different reasons: for personal enjoyment in a variety of circumstances and with friends. Through analysing the interviews it was found that these music retrieval systems have been designed with a variety of music seekers in mind. That is they are aimed at users with novice to expert knowledge of computers and the Internet. The system used text queries to locate and retrieve items, which is suitable for almost all music seekers. Although the more experienced music seekers and the music professionals would benefit from conducting their searches more accurately using music pitch and rhythm, the ordinary user would find it difficult to use such evidence in their searches [11]. However, from this research it appears that for the most part, the needs of music seekers are currently being met by MIR systems. It should be noted that this issue is causing significant copyright problems [12,13], which could potentially have a significant impact on the usage of MIR systems such as Kazaa.

Further research in this area is essential, particularly as the music seeking behaviour of users will change as they become older – for example it is clear that the behaviour of users in the over 40 group will change when those with more advanced IT and Internet skills move into that group. Therefore while we are sure that our sample is representative in the current time frame, it may not be in the next five or ten years due to changes in the makeup of the groups. It is therefore important to keep monitoring the behaviour of users to track any changes as they occur. A further issue is understanding precisely what the 'information problem' is in music retrieval. Textual information retrieval has some understanding of this issue [14], but it is unclear

what such a concept means in music retrieval, if indeed it has any meaning at all. User studies are useful, but some kind of cognitive framework is required if we are to better understand the music seeking behaviour of MIR users.

7. ACKNOWLEDGEMENTS

We are grateful to Tim Crawford for his advise in preparing this paper.

8. REFERENCES

[1] Smith, T. (2000) Napster has huge number of users shock! [online] Available on: http://www.theregister.co.uk/content/6/13789.html [visited 22nd April 2004]

[2] Kazaa web site, http://www.kazaa.com [visited 22nd April 2004]

[3] Tang, N. (1999) Multimedia Information Retrieval Systems: An Overview, [online] Available on: http://www.cs.ucla.edu/~tang/papers/multimedia-IR.pdf [visited 22nd April 2004]

[4] Graziano, M. and Rainie, L. (2001) The music downloading deluge: 37 million American adults and youths have retrieved music files on the Internet. Available on: www.pewinternet.org/reports/pdfs/PIP_More_Music_Report.pdf [visited 22nd April 2004]

[5] Byrd, D. and Crawford, T. (2002) Problems of music information retrieval in the real world, *Information processing and management*,38 (2), pp 249-272.

[6] Veltman, M. (2002) The Dutch solution to peer-to-peer software: the latest ruling in the Kazaa-case, [online] Available on: http://www.kent.ac.uk/law/undergraduate/modules/ip/handouts/2002_3/Kazaa_essay.doc [visited 16 June 2003]

[7] Lippincott,A.(2001) Issues in content-based music information retrieval, *Journal of Information Science*, 28 (2) 2002, pp. 137–142.

[8] Cunningham, S. (2002) User studies: A First Step in Designing an MIR Testbed [online] Available on: http://music-ir.org/evaluation/wp1/wp1_cunningham.pdf [visited 22nd April 2004]

[9] Futrelle, J. (2002) Three Criteria for the Evaluation of Music Information, Retrieval Techniques against Collections of Musical Material [online] Available on: http://music-ir.org/evaluation/wp1/wp1_futrelle.pdf [visited 22nd April 2004]

[10] Chevallet, J. and Nigay, L. Characteristics of users' needs and activities: A design space for interactive information retrieval systems. Available on: http://iihm.imag.fr/publs/1996/HCIIRChavelletNigay96.pdf [visited 22nd April 2004]

[11] Uitdenbogerd, A. and Zobel, J. (2000) Music Ranking Techniques Evaluated, In International Symposium on Music Information Retrieval, [online] Available on: http://goanna.cs.rmit.edu.au/~jz/fulltext/acsc02uz.pdf [visited 22nd April 2004]

[12] Beverley, P. (1999). Protecting Copyright, Increasing Choice and Generating Revenue from the Internet, *in* Fox, M (2002)Technological and social drivers of change in the online music industry, *First Monday*, [online] 7 (2) Available on: http://www.firstmonday.dk/issues/issue7_2/fox/index.html [visited 22nd April 2004]

[13] Austin, S. (2004). Web warning to pop pirates, Metro, Friday 26th March, p1.

[14] Ingwersen, P. (1992). Information Retrieval Interaction, Taylor Graham, Available on: www.db.dk/pi/iri [visited 22nd April 2004]