

Usage of the MELDEX Digital Music Library

John R. McPherson
Department of Computer Science
University of Waikato
Hamilton
New Zealand
jrm21@cs.waikato.ac.nz

David Bainbridge
Department of Computer Science
University of Waikato
Hamilton
New Zealand
davidb@cs.waikato.ac.nz

ABSTRACT

Online Digital Music Libraries are becoming increasing common and more sophisticated; however, availability of information on how users access and navigate these resources is limited. This type of information is crucial for improving user interface design and for providing users with better supported services.

Here we present an analysis of the logs for our digital music library, Meldex, for a 1 year period to discover patterns of usage.

1. INTRODUCTION

Our Melody Index [1] is part of the New Zealand Digital Library project (nzdl.org). Users can access songs in two ways: they can see the results of a query, or they can browse the song titles alphabetically. Queries can either be melodic or textual. Melodic queries are submitted by either uploading (posting) a short recording of sung or played notes, or by providing a Uniform Resource Locator (URL) to such a recording. Our demonstration page provides some sample recordings. Textual queries are matched against song metadata, such as title or author, and lyrics.

Songs are returned in a variety of different audio formats, such as WAV, MIDI, and Audio Interchange File Format (AIFF). Some collections can also have results returned as an image of the original sheet music. For example, our “Fake Book” collection is built from the results of running optical music recognition over sheet music. Copyright considerations restrict which collections return full-length audio files and images.

Our oldest collection is known as the “Folksong” collection. Based on the Essen and Digital Tradition databases, it consists of 9,354 folk songs which are divided into geographical regions (Chinese, German, Irish and North American). The “Fakebook” collection (mentioned above) consists of 1,235

songs. The “Midi” collection is built from 9,763 MIDI files sourced from the Web, and supports textual and melodic querying. The “MidiMax” collection indexes 17,799 MIDI songs and is more sophisticated, also allowing the indexing and retrieval of motifs. It has been available since October 2000, while the other collections have been online since November 1999.

2. A SELECTION OF STATISTICS

In reviewing prior work for usage analysis, the Variations music library [2] at Indiana University is notable for providing daily statistics online. Given the context of the Variations project, these focus on aggregate performance-oriented statistics such as number of songs retrieved, and maximum, minimum and average retrieval times.

Here we present an analysis of the usage logs of our digital library service for the 12 month period 1 April, 2000 to 31 March, 2001. Most of the results given here that are not for the whole library are for the Folksong collection, as this data set reflects patterns observed across the other collections.

Figure 1 shows the number of daily hits received (the line represents a rolling 7-day average). There is not a noticeable trend here, although there is a drop-off over the Christmas and New Year holidays. There are also several brief periods of server outages.

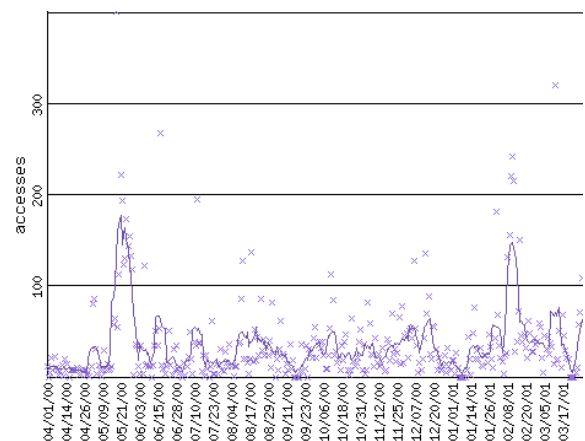


Figure 1: Daily accesses for the folksong collection

Table 1 shows the distribution of visitors to the folksong collection. This is based on all the web pages generated by

Table 1: Top 10 visitor domains for ‘folksong’

Domain	Accesses	%age of total
.net	3,827	29.67
.com	2,128	16.50
Europe	2,102	16.30
unknown	2,090	16.20
.edu	1,001	7.76
Sth. Pacific	661	5.12
Asia	435	3.37
Nth. America	235	1.82
Australia	188	1.46
Sth. America	73	0.57
Totals:	12740	98.77

Table 2: Results pages generated — All collections

Page type	Number
Own audio file with text query	104
Own audio file only	588
Demo audio file with text query	105
Demo audio file only	89
Text query only	1539
Browse titles	1070
Total:	3495

the music library and includes help and query pages, for example, in addition to requests for songs from the collection.

Around 2000 of the hits for the folksong collection are from one site, which appears to have been crawling part of our library website. This accounts for slightly over half of the visits from the .net top-level domain, and also accounts for the two spikes observed in Figure 1. That particular internet address has been filtered from the remaining statistics given here. In addition, the addresses used by Meldex’s principal researchers over this period have been filtered out of all statistics in this report.

Assuming that any accesses from the same IP address with less than five minutes of separation are part of the same “visit”, the average amount of time spent per visit over all Meldex collections was slightly over 2 minutes, and consisted of an average of 3.4 page views. Visits came from just over 1,900 different internet addresses.

Table 2 shows how users get to song listings. Just under 70% of song listings are generated as a result of a query, either audio or textual (or possibly both), with the remainder generated as alphabetical listings of titles.

Table 3: Users’ preferred Audio file format

Audio Format	%age
MIDI	48.5 (834)
WAV	23.8 (410)
Real Audio	17.2 (295)
AIFF	04.0 (69)
Soundblaster VOC	02.6 (45)
Sun u-law	02.3 (39)
Sun AU	01.6 (28)
Total	100 (1720)

Three of the available file formats account for nearly 90% of users’ preference settings, with the MIDI format accounting for nearly half. These settings are listed in Table 3.

Table 4: Folksong Hit Parade - Top 10

Accesses	Name
80	“Auld Lang Syne” [from demo page]
72	“Aéire cinn bó rúin”
62	“The Ash Grove” [from demo page]
52	“Abdul Abulbul Ameer”
49	“Ai erwa”
36	“Three Blind Mice” [from demo page]
30	“A New England Ballad”
25	“Abilene”
25	“A-Beggin’ I Will Go”
22	“Adam and Eve”

Table 4 gives the titles of the 10 most frequently requested songs for the folksong collection. Of the 9,354 songs indexed in this collection, 2,395 (25.6%) have been accessed at least once, and about 1,700 have been accessed exactly once, suggesting that these downloads are the results of users’ individual queries.

3. SUMMARY

Around thirty percent of song lists generated are alphabetical title listings, and most accesses from these lists are for songs that start with the letter ‘A’. We conjecture this is because new users to the library have a strong desire to discover what sort of music is contained in the collection, and accessing songs by titles is the easiest route currently available in the interface.

A result that took us initially by surprise is that forty-four percent of all listings are the result of a text query alone. While it is possible that a wide range of Web users are conditioned to typing queries into a text box, it should not be overlooked that the overhead of entering a music query in our current interface might be too high for many users. Also, analysis of our own group members has shown that for large MIDI collections, a browsing habit that had formed was to enter a text query on some vague topic (for example, “fire”) and see which tunes popped up.

Our Meldex service is available at www.nzdl.org/musiclib.

4. REFERENCES

- [1] Rodger J. McNab, Lloyd A. Smith, David Bainbridge, and Ian H. Witten. The New Zealand Digital Library MELody inDEX. *D-Lib Magazine*, May 1997.
- [2] Jon W. Dunn and Constance A. Mayer. Variations: A digital music library system at Indiana University. In *Proceedings of the Fourth ACM Conference on Digital Libraries*, Berkeley, California, 1999. ACM.