

Evaluation form – Progress Reports

V1.06 Feb. 2004

Project ID	2003252	Project Title	An Interactive Recommender System	
Student Name	Andrew Yeung		Supervisor	Professor Rich Zemel, Dept of CS
Section #	6	Coordinator	D. Beresford	

Presentation	Copy graded: <input type="checkbox"/> electronic <input type="checkbox"/> paper	<input type="checkbox"/> comment summary in paper
Coordinators Signature:		Grade /10 (each report worth 2.5%)

Technical Evaluation	Total Mark	Suggested Mark	Notes
Progress	10		<ul style="list-style-type: none"> following game plan? (milestones accomplished; milestones missed & why)
Organization	10		<ul style="list-style-type: none"> use of human resources? use of non-human resources? efficiency of efforts? preparation & foresight?
Method	15		<ul style="list-style-type: none"> are the modules / steps sufficiently tested when 'done'? do these tests have good structure? have the interaction specifications of the parts been properly developed?
Decision making	10		<ul style="list-style-type: none"> suitable response to difficulties? appropriate change of course when encounter obstacle or new development? modification of process as necessary?
Creativity / Complexity / Effort	35	C/C _____ Effort x _____ x 35	Creativity and Complexity from charts. Scores for creativity, complexity and effort are multiplied by maximum score / 1000 to get the grade.
TOTAL	80		(report #1 worth 7.5 & report #2 worth 12.5% of the student's final grade)
Supervisor	<input type="checkbox"/> Accept suggested technical mark <input type="checkbox"/> Change technical mark (reasoning attached in separate sheet, marks in blank column above)		
Supervisor's Signature:			

Note to students: There is a design award, the Aloha award, that you might wish to apply for. Please check the course website for instructions on what to do to be considered for this award.

**The Edward S. Rogers Sr. Department of Electrical and Computer Engineering
University of Toronto**

**ECE496Y Design Project Course
Individual Progress Report**

Title: An Interactive Recommender System

Project I.D.#: 2003252

Prepared by: Andrew Yeung andrew.yeung@utoronto.ca

Supervisor: Professor Rich Zemel, Dept. of Computer Sciences

Section #: 6

Section Administrator: D. Beresford

Date: February 23, 2004

TABLE OF CONTENTS

EXECUTIVE SUMMARY:	5
INTRODUCTION:.....	6
PROGRESS:.....	6
CONCLUSION:	13
REFERENCES:.....	15
APPENDIX A - Unix Installations And Environment Setup:	17
APPENDIX B - Song Sample Format:	19
APPENDIX C - User Profile Information:	20
APPENDIX D - MySQL Database Structure:	21
APPENDIX E - Testing Procedures:	22
APPENDIX F - Oral Presentation Outline:	23
APPENDIX G - Milestones Version 1 (Original):.....	24
APPENDIX H - Milestones Version 2:.....	26

EXECUTIVE SUMMARY:

The pace of development in the design of an Active Recommender System has been going well, without any major problems that have caused our design project to have to undergo significant revision in terms of goals and requirements. Prior to this report our group had been working on separate modules of the system, and this report details the integration process and preparation for testing and research. Milestones have had amendments made to start and end dates, and these alterations have been made as a result of our group gaining a better understanding of the time requirements to be spent on each milestone, minor delays from debugging and problem-solving, and previously unknown conflicts between milestones and our other academic courses. There has been addition and removal of milestones, as well as a changing of responsibility of a milestone. The result of this has helped our team divide up work more efficiently and to define a clearer agenda, ultimately giving us the ability to be better prepared as we enter each successive milestone.

Each member of the Active Recommender System has completed their milestones to date, with minor delays due to problems including coding errors and integration troubles between modules in the project. The milestones that Andrew has been responsible for has gathered essential user information, tested the entire system to ensure proper operation, defined formats and methodologies for future project stages, and uncovered areas of improvement. Problems found have been resolved through a variety of methods including consultation with members of the team and academic staff, online research, and thorough testing emphasizing both module and system-wide testing.

INTRODUCTION:

Milestones under the responsibility of Andrew focused on the gathering of user profile and music rating data through an online survey, system-testing of the Active Recommender System with all three modules incorporated, the coordination and defining of the research phase, and the oral presentation format. A significant amendment to the milestones was the removal of one milestone after it was decided that the milestone would not impact the success of our project and would only add unneeded complications. Additionally, two milestones, original scheduled for Andrew and Bernard, were placed under the responsibility of Bernard because of easier integration by one person, and three additional milestones were added for Andrew to allow for easier progress accountability.

PROGRESS:

Milestone #1 [New] - Creation of Website to gather User Ratings: Website created to gather ~75 user ratings of music using PHP, MySQL and HTML, as well as profile information such as age, sex, mood, favorite genre [See Appendix B, C]. Due January 17, 2004.

Responsibility: Andrew Yeung

Status at start of reporting period: Website plans designed for integration with MySQL database using PHP.

Status at end of reporting period: Website, database completed Jan. 5, 2004.

Actions: Due to the lack of publicly available resources for music rating data, a webpage (~4000 lines of code) was created of 126 MP3 songs in 10 separate genres, with a minimum of 10 songs per genre in the format shown in Figure 1.

Webpage was linked to MySQL database of information [See Appendix D]. Each song can be rated on a scale from 1[least liked] to 5 [most liked] and user can also select their sex, mood, favorite genre and age. Approximately 75 users have gone through the survey. Website: [1], PHP/MySQL Resources [2, 3].

ID:	GENRE:	TITLE:	ARTIST:	RATING:	MP3:
1	Blues	Blues Brothers	Mini the Mocher	<input checked="" type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	Sample
2	Classical	Beethoven	3rd Movement	<input type="radio"/> 1 <input type="radio"/> 2 <input type="radio"/> 3 <input type="radio"/> 4 <input checked="" type="radio"/> 5	Sample
3	Reggae	Wayne Wonder	No Letting Go	<input type="radio"/> 1 <input type="radio"/> 2 <input checked="" type="radio"/> 3 <input type="radio"/> 4 <input type="radio"/> 5	Sample
:	:	:	:	:	:
:	:	:	:	:	:

Figure 1: Sample table to gather user ratings

Decisions: Because of the difficulty in expecting users to rate all 126 songs, a minimum requirement of 5 songs per genre was made, and the user was also asked to rate songs that they did not know. This was done to expedite survey completion time, and to allow all songs to have a similar number of ratings. We decided to use mainstream artists when possible to make it easier for users to have familiarity with the music, and also because it was simpler to obtain the music, although a newspaper article providing links to independent music was found, which could be used to obtain music for future use [4]. Completion of website was achieved before the original due date, but time was extended to gather user data. Development was done on supplied Computer Science AI Group Linux machines and installation procedures are shown in Appendix A.

Testing & Verification of Progress

1. Website Link Verification: All website links, mp3 song samples tested to ensure that they were not broken and played specified song sample.

2. General Test: User Rating System tested and verified locally in Com. Sci. AI Lab(where developed) by checking MySQL database for results.
3. Public Test: User Rating System tested and verified using public terminals to ensure that website was not blocked by any U of T Computer Science firewalls (where website is hosted) and that communications port was open and available.

Results: Song samples were changed to play in a separate window instead of opening up on the same page to allow the user to maintain focus on the main window. Website color-coded to differentiate genres and give a more professional look. Minor corrections to HTML / PHP code to correct broken links and coding errors. Met with CS administrators to acquire privileges to host site.

Milestone #2 [removed] - XML structure definition: With coordination between the three main modules to ensure optimal compatibility, a Document Type Definition structure will be created for use in the transfer of music information between the web server, application and database, and ultimately to the user. Due Nov. 17, 2003.

Responsibility: Andrew Yeung

Status at start of reporting period: Research for XML structure finished and structure created [5].

Status at end of reporting period: Milestone removed due to unnecessary requirements for project.

Actions: Our design group discussed the need for XML integration into the project, and after a comparison between benefits and disadvantages, as well as the

amount of integration work required, it was decided that the project would continue without the use of XML.

Decisions: The original decision to use XML was made because it allowed for an application independent medium of transport that could be used by other systems which had the proper XML structure definition. Because our project is only a small scale development, having information that is easily usable by others is not a vital requirement, and would only be a long-term benefit if the recommender system was used by other research groups. By removing the milestone, time was saved in having to investigate how to parse and read XML data using Java and MySQL.

Testing & Verification of Progress: DTD structure verified to conform to XML defined standards and follow example structures [5].

Milestone #3 [Responsibility Change] - Integration of web server and

database: Web server and database communication processes designed such that data will remain accessible for future sessions by a user.

Responsibility: Bernard Ma

Status at start of reporting period: Not started

Status at end of reporting period: Completed January 24, 2004.

Actions: Researched methods to allow communication between web server and database (Java and MySQL) components, and found an official Java Database Connectivity (JDBC) driver - MySQL Connector/J [6], provided to Bernard M.

Decisions: The change in responsibility to have Bernard M. do both the 'Integration of web server and database' milestone as well as the 'Integration of web server and application' was decided because Bernard has the most experience working with the web server and it is the base component in both of the

milestones, and the majority of the programming required for the task involved modification on the web server side. The JDBC-3.0 API used to connect the web server to the database was chosen because of its stability, speed and compatibility with the Tomcat server and MySQL database.

Testing & Verification of Progress: Bernard's responsibility in the milestone included all testing and verification of correct operation.

Milestone #4 [New] - General Testing of entire web server / application /

database functionality: Exhaustive system tests done to ensure that components communicate properly, and data is being recorded and used accurately. Due Feb13, 2004. Website: [7]

Responsibility: Andrew Yeung

Status at start of reporting period: Not started

Status at end of reporting period: General Testing completed Feb. 11, 2004.

Actions: Recommender System was tested to ensure that all 3 modules (GUI, algorithm, and database) worked correctly with each other and that information was being correctly handled as a fully functioning system. Account creation and logins were tested, ratings of all the songs were done, and suggestions on web site from a visual standpoint were made.

Decisions: Discussions were made regarding the best way to present information clearly and cleanly to the client when they are using the system. We decided on an interface that had a constant familiar style, and simplicity, with few graphics to allow audio content to load quicker, allowed for more efficient loading of pages and easy to understand content.

Testing & Verification of Progress

1. Exhaustive Testing: Rating of all songs in database individually verified to check that system does not stall when there are no more songs available to rate. Check that information written to database.
2. Visual Test: Making sure that important aspects of webpage are highlighted or made to stand out. Uniform tables used to display information, and current generated list of recommendations always available for viewing. Personal and public opinions used to decide on best presentation format.
3. Navigation Test: Opinions from group and public users used to decide on layout of web pages, and how sitemap would be structured. Browser compatibility on Microsoft and Mozilla web pages verified. Usability tests following recommendations found at [8].
4. Account Creation Test: Making sure that duplicate accounts cannot be created, and that new accounts could be immediately used to log into website. Password verification incorporated. MySQL database tables checked to verify correct information entered.

Results: System stalled during exhaustive testing because there was no handling of the case when all the songs had been rated. Error fixed by looping the list of songs. Additional buttons added to site to allow user to quickly return to main menus with minimal (1-2) mouse button clicks. Suggestions made and implemented for modifications to web page code to fix formatting incompatibilities when displaying pages on Mozilla web browser.

Milestone #5 - Coordination of Research Schedule: Discussion with graduate student Ben Marlin regarding ACF testing methodologies, and two methods of testing will be decided upon, numbers of test subjects to be used, and types of measuring scales to be used to determine accuracy and usefulness of recommendations. Due February 13, 2004.

Responsibility: Andrew Yeung

Status at start of reporting period: Not started

Status at end of reporting period: Testing processes finalized Feb. 7, 2004.

Actions: A variety of online resources were consulted to research on methods that people used to test their collaborative filtering methodologies [9, 10, 11].

Decisions: The goals in coming up with testing procedures were to have plans that would provide non-simulated results, so we decided to go with a process that required public participants to utilize the system. Another goal was to minimize the time required for the participant to run through the tests. The use of two different testing methodologies allows for a variety of result comparisons.

Testing & Verification of Progress:

Two tests, one which requires user to rate all songs and the other which requires user to rate only x number of songs [See Appendix E]. Analysis of test results is a separate milestone under the responsibility of Cavan Y.

Milestone #6 - Oral Presentation: Presentation format decided upon and roles in presentation assigned. Decisions regarding appropriate diagrams and results to display will be considered and created. Rehearsal of presentation and analysis of strengths and weaknesses. Due March 26, 2004.

Responsibility: Andrew Yeung

Status at start of reporting period: Preliminary ideas discussed with team on method of presentation.

Status at end of reporting period: Rough outline of oral presentation made.

Actions: Attended design project presentations and considered the variety of formats used and methods to convey information such as speaker rotations, style of dress, visual aids, amount of information detail, and audience questions.

Decisions: Our initial oral presentation outline will have each of the members talk about their particular module of the project that they worked on, and then a discussion of test results and their significance will follow. The introduction of the presentation will provide definitions, focus, goals and objectives, and our conclusion will have thoughts on futures expansions to the project and impact of the research. Oral Presentation deadline was changed once our group was able to determine the exact presentation date, so that time would be properly budgeted.

Testing & Verification of Progress: N/A. Presentation Outline: Appendix F.

CONCLUSION:

Our design project group currently has a working recommender system in place, and we are now beginning the research portion of the project, to compare Active Collaborative Filtering versus Collaborative Filtering. The changes between our original and current list of milestones can be seen in Appendix G, H, and they have occurred as a result of the following considerations:

- A) Taking into account previously unconsidered requirements
- B) Defining a more thorough testing schedule

C) Allowing for more preparation for successive milestones by working and deciding on key aspects earlier

The next phase for the project is focusing on research, to compare the effectiveness of the Active Collaborative Filtering technique with the currently used Collaborative Filtering methodology. We believe that the ACF technique, which uses an approach to gather the most useful information at every instance instead of randomly querying user ratings, will prove itself to be more truthful at predicting music that a user will like, and the analysis, oral and design project presentations, and the concluding final report will summarize these results.

REFERENCES:

- [1] [Online] Available <http://giraffe.ai.utoronto.ca:30491/musicrating.php> 'User Rating Gathering Site' [cited 2003 Dec 28]
- [2] [Online] Available <http://phpclasses.iplexx.at/browse.html/package/782.html> 'PHP/MySQL Tutorial' [cited 2003 Dec 5]
- [3] [Online] Available <http://www.freewebmasterhelp.com/tutorials/phpMySQL> 'PHP/MySQL Tutorial' [cited 2003 Dec 5]
- [4] Gill, Alexandra. Available HTTP: <http://www.globetechnology.com/servlet/story/RTGAM.20040119.gtcbc19BNStory/Technology/> "Indie music and beyond" [cited 2004 Jan 19]
- [5] [Online] Available HTTP: <http://www.xmlfiles.com/dtd/> "XML DTD Tutorial" [cited 2003 Oct 25]
- [6] [Online] Available <http://www.mysql.com/downloads/api-jdbc-stable.html> 'Java / MySQL driver' [cited 2004 Jan 2]
- [7] [Online] Available <http://giraffe.ai.utoronto.ca:30492/ACF> [cited 2004 Jan 15]
- [8] [Online] 'Testing Guide -Things to keep in mind when doing usability tests ' Available <http://filebox.vt.edu/users/sakella/web/Evaluation.htm> [cited 2004 Jan 18]
- [9] O'Connor, Mark & Jon Herlocker. "Clustering Items for Collaborative Filtering." Available http://www.csee.umbc.edu/~ian/sigir99-rec/papers/oconner_m.pdf [cited 2004 Jan 13]
- [10] [Online] 'CF testing techniques' Available http://www.multiagent.com/Topics/Recommender_Systems-Collaborative_Filtering/ [cited 2004 Jan 9]
- [11] [Online] 'REFEREE: Recommender System Tester ' Available <http://www.neci.nec.com/~lawrence/papers/referee-vldb02/referee-vldb02.pdf> [cited 2004 Jan 17]
- [12] [Online] Available <http://mp3splt.sourceforge.net> [cited 2003 Oct 3]
- [13] [Online] Available <http://users.otenet.gr/~jtcliper/tgf/> "The GodFather" [cited 2003 Oct 4]
- [14] [Online] Available <http://www.grouplens.org/data/million/readme.txt> [cited 2003 Dec 6]

- [15] [Online] 'CF Recommender Algorithms' Available www.cs.tau.ac.il/~fiat/dmsem03/ItemBasedCollaborativeFilteringRecommendationAlgorithms.ppt [cited 2004 Jan 9]
- [16] Breese, John S., David Heckerman and Carl Kadie. "Empirical Analysis of Predictive Algorithms for Collaborative Filtering." *In Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence, Madison, WI, July 1998. Morgan Kaufmann.*
<http://research.microsoft.com/users/breese/cfalgs.html> [cited 2004 Jan 13]

APPENDIX A - Unix Installations And Environment Setup:

- for installations, type the command `./configure --help` for configuration options

Setting up Java / Tomcat Environment:

- 1) go to pwd (`cd ~`)
- 2) `pico .cshrc`
- 3) Add the line:
'set path =
(\$path/direct/pkgs/jdk1.4.0/linux/j2sdk1.4.0_01/bin)' under
the #set path... line
- 4) 'source .cshrc' to reload file and setup environment
- 5) The 'source' command doesn't really work, so after you edit the .cshrc file, log out of the computer and log back in again to ensure that environment has been reloaded
- 2) Add 'setenv JAVA_HOME /pkgs/jdk1.4.0/linux/j2sdk1.4.0_01'
to the .cshrc file to be able to run Tomcat application

Installation of Apache 2.0.48 (<http://www.apache.org>)

Download Apache tar ball (`httpd-2.0.48.tar.gz`)

1. `tar xvzf httpd-2.0.48.tar.gz`
2. `cd httpd-2.0.48`
3. `./configure --prefix=/homedirectory/apache
--enable-module=so`
4. `make`
5. `make install`

Now you need to configure `httpd.conf` to make apache use port 30491.

Edit the following lines in:

`/homedirectory/apache/conf/httpd.conf`

```
ServerName giraffe.ai.toronto.edu  
Listen 30491
```

Finally start apache:

```
/homedirectory/apache/bin/apachectl start
```

Connect to apache using: <http://nameofmachine:30491>

Installation of MySQL 4.3.4 (<http://www.MySQL.org>)

Download MySQL tar ball (MySQL-2_0_48.tar.gz)

1. tar xvzf MySQL-4_0_17.tar.gz
2. cd MySQL-2_0_48
3. ./configure --prefix=/homedirectory/MySQL
4. make
5. make install

Installation of PHP 4.3.4 (<http://www.php.net>)

Download PHP tar ball (php-4.3.4.tar.gz)

1. tar xvzf php-4.3.4.tar.gz
2. cd php-4.3.4
3. ./configure --prefix=/homedirectory/php434 --with-MySQL=/homedirectory/MySQL --enable-track-vars --with-xml --with-apxs2=../apache/bin/apxs
4. make
5. make install

Installation of phpESP (<http://phpesp.sourceforge.net>)

Download phpESP tar ball (phpESP-1.6.1.tar)

1. tar xvf phpESP-1.6.1.tar
2. mv phpESP-1.6.1 phpESP
3. cp ~phpESP/scripts /homedirectory/MySQL/bin/
4. /homedirectory/MySQL/bin/MySQLadmin -u root password [password]
5. /homedirectory/MySQL/bin/MySQL -uroot -p[password] < scripts/MySQL_create.sql

APPENDIX B - Song Sample Format:

Genre	# of songs	Size of Song	Recording Specifications	Song ID3 Information
Blues	11	~1MB per song	128Kbps 44Hz	Artist Song Title Genre
Classical	12			
Country	13			
Folk	10			
Jazz	10			
Misc	11			
New Age	15			
Reggae	10			
Rock	22			
Soundtrack	12			

Figure B.1: Standardized Sample Music Breakdown

- MP3SPLT: Software tool used to retrieve 1 minute song samples [12]
- The GodFather: Software tool used to supply song information [13]

APPENDIX C - User Profile Information:

Age : < 15
15 – 30
30 – 45
45 – 60
> 60

Current Favourite Genre : Blues
Rock
Classical
Country
Folk
Jazz
Miscellaneous
Newage
Reggae
Soundtrack

Current Mood : Bored
Confused
Happy
Mad
Neutral
Sad
Tired

Sex : M (Male)
F (Female)

- Profile information based on Grouplens format [14]

APPENDIX D - MySQL Database Structure:

Hosted on : giraffe.ai.utoronto.ca

Database Name : musicratings

Database Tables : a) music
b) ratings
c) users

Table Breakdown :

a) music

Field	Type	Null	Key	Default	Extra
Id	int(11)		PRI	NULL	auto_increment
Genre	varchar(15)				
Artist	varchar(35)				
Title	varchar(45)				

b) ratings

Field	Type	Null	Key	Default	Extra
Id	int(11)		PRI	NULL	auto_increment
_1	varchar(4)	YES		NULL	
_2	varchar(4)	YES		NULL	
.
.
.
_125	varchar(4)	YES		NULL	
_126	varchar(4)	YES		NULL	

c) users

Field	Type	Null	Key	Default	Extra
Id	int(11)		PRI	NULL	auto_increment
Age	varchar(15)				
Fav_Genre	varchar(15)				
Mood	varchar(15)				
Sex	varchar(4)				

APPENDIX E - Testing Procedures:

- Statistical Accuracy Metrics used as a measure for evaluating the quality of the recommender system: user ratings compared to numerical recommendation scores in the test data set [15]

- Mean Absolute Error : provides deviation value of recommendations from the actual supplied user values

$$MAE = \frac{\sum_{i=1}^N |p_i - q_i|}{N}$$

- Testing will be done using a popular methodology called 'cross-validation'
- Cross Validation / Leave One Out evaluation: a rating is predicted by using *all the data except the rating itself* [reference Breese, John S]
- Of the 75 user records submitted, 10 user ratings, all who have rated the entire set of 126 songs, will be randomly selected and used in the following two tests:

Method 1) - all songs in the system will be rated
- match recommended list with the list of ratings
- calculate average percentage error between two lists

Method 2) - rate x number of songs in the system
- match recommended list with the list of ratings
- calculate average percentage error between two lists

- Results will be performed using both Collaborative Filtering and Active Collaborative Filtering methodologies. This is accomplished by modifying the recommender system to ask user to randomly rate songs (CF method) versus asking the user to rate songs that will provide the most information to the system when making recommendations (ACF method)
- Currently only user ratings will be used for testing, although additional information gathered such as age, sex, mood, favourite genre may also be used

APPENDIX F - Oral Presentation Outline:

- Thursday April 8 7-8pm BA1240
- MS PowerPoint Format (18pt Arial Font)
- Work Casual Dress (ie Collared Shirt and Khaki)
- ~ 30 seconds per slide (28 slides / 14 minutes)

INTRODUCTION (~5 slides / 2.5 min) Andrew Yeung:

- Title slide
- Purpose of project
- Definitions
- Goals
- Overview of Recommender System

MODULE 1: Java Tomcat Web Server Interface - Bernard Ma (~4 slides / 2 min):

- Go over format of web server
- Advantages / Disadvantages of using Java Tomcat language

MODULE 2: MySQL Database Andrew Yeung (~4 slides / 2 min):

- Describe how new user information is stored in database
- Advantages / disadvantages of MySQL
- Provide information on website used to gather ratings for algorithm

MODULE 3: Active Collaborative Filtering Java Algorithm - Cavan Yie (~4 slides / 2 min):

- Describe ACF algorithm in detail using diagrams
- Advantages / disadvantages of Java

TEST RESULTS - Bernard Ma (~6 slides / 3 min):

- Testing Methodologies
- Charts showing results between Active Collaborative Filtering and Collaborative Filtering techniques

PROGRESS, FUTURE WORK AND CONCLUSION - Cavan Yie (~5 slides / 2.5 min):

- Summarize work accomplished, significance of results
- Possible extensions to project and future work to be done

APPENDIX G - Milestones Version 1 (Original):

List each major milestone in chronological order. Assign ONE key team member that has ultimate responsibility to each milestone. Use up to two pages if necessary.				
#	Description	Assigned to	Start Date	End Date
1.	<p>Clear definition of project objectives, methodologies and software and hardware components that will be used:</p> <p>Collection and research of resources and papers that will be used for project. Obtain computer resources from Professor Zemel and obtain access privileges to Artificial Intelligence labs.</p>	Andrew	09/01/2003	10/17/2003
2.	<p>Implement Active Collaborative Filtering Methodology in Java.</p> <p>The back-end of the system will be coded in Java. This includes the calculations of EVOI for queries, model fitting for the probability model, and computing rating predictions. Its main purpose is to return a set of recommendations to a user given a database of user ratings.</p>	Cavan	10/17/2003	11/21/2003
3.	<p>Database Implementation in MySQL:</p> <p>MySQL database will be created to store music information. Decisions on the music information to be displayed, types and categories of music to be used, and music sample format and size will be made based on research of related sites and considerations such as transfer speed.</p>	Andrew	10/17/2003	11/07/2003
4.	<p>Webserver / GUI Setup using Jakarta Tomcat, Java Server Pages, Apache Server.</p> <p>A user friendly and informative interface will be designed and implemented. The GUI will provide features such as information about the music to be rated (mp3 sound clips, artist information etc.).</p>	Bernard	10/17/2003	11/14/2003
5.	<p>XML structure definition for music, users and ratings:</p> <p>A Document Type Definition structure will be created for use in the transfer of music information between the web server, application and database modules, and ultimately to the user. This step requires an understanding and coordination with the three main modules to ensure optimal compatibility.</p>	Andrew	10/31/2003	11/17/2003
6.	<p>Integration of web server and database:</p> <p>Interface for modular database component will have a focus on easy and fast access with webserver. Testing of speed of data retrieval will be looked at and optimizations and modifications to both components will be considered if necessary.</p>	Andrew	11/14/2003	12/19/2003

7.	<p>Integration of web server and application.</p> <p>In order to maximize the effectiveness of the system, the Web Server must interact quickly with the application. The speed of use will be tested, and modifications to both components will be made if necessary.</p>	Bernard	11/21/2003	12/19/2003
8.	<p>Integration of application and database. (2 weeks)</p> <p>Enabling the Java code to successfully retrieve and update data to the database. Ensure that communication between the two components is smooth.</p>	Cavan	01/04/2004	01/25/2004
9.	<p>Testing, comparison and implementation of alternative recommender algorithms. (2 weeks)</p> <p>Comparison of "active" versus "non-active" approaches to collaborative filtering. Analyze their performances and accuracy to real life usage.</p>	Cavan	01/26/2004	02/12/2004
10.	<p>Oral Presentation:</p> <p>Summarization of progress will be condensed into presentation format. Visuals and aids will be considered and created to provide an effective emphasis and attract audience. Rehearsal of presentation and analysis of strengths and weaknesses.</p>	Andrew	02/13/2004	02/27/2004
11.	<p>Design Fair Poster Presentation</p> <p>The poster will attempt to give an overall view of our design project while keeping in mind the audience will mostly be comprised of 3rd year ECE students. The results of our research will be displayed through the use of charts and graphs. A computer will also be available to provide a hands on demonstration of our system.</p>	Bernard	02/27/2004	03/16/2004
12.	<p>Completion of group final report</p> <p>Integration of the documentation of all components which make up the design project including diagrams, figures, references.</p>	Cavan	03/16/2004	04/08/2004

APPENDIX H - Milestones Version 2:

(February 23, 2004)

List each major milestone in chronological order. Assign ONE key team member that has ultimate responsibility to each milestone. Use up to two pages if necessary.				
	Description	Assigned to	Start Date	End Date
1.	<p>Clear definition of project objectives, methodologies and software and hardware components that will be used:</p> <p>Collection and research of resources and papers that will be used for project. Obtain computer resources from Professor Zemel and obtain access privileges to Artificial Intelligence labs.</p>	Andrew	09/01/2003	10/17/2003
2.	<p>Implement Active Collaborative Filtering Methodology in Java.</p> <p>The back-end of the system will be coded in Java. This includes the calculations of EVOI for queries, model fitting for the probability model, and computing rating predictions. Its main purpose is to return a set of recommendations to a user given a database of user ratings.</p>	Cavan	10/17/2003	01/21/2004
3.	<p>Database Implementation in MySQL:</p> <p>MySQL database will be created to store music information. Decisions on the music information to be displayed, types and categories of music to be used, and music sample format and size will be made based on research of related sites and considerations such as transfer speed.</p>	Andrew	10/17/2003	11/07/2003
4.	<p>Web server / GUI Setup using Apache/Jakarta Tomcat Server and Java Servlets</p> <p>A user friendly and informative interface will be designed and implemented. The GUI will provide features such as information about the music to be rated (mp3 sound clips, artist information etc.).</p>	Bernard	10/17/2003	12/13/2003
5.	<p>Creation of Website to gather User Ratings using PHP, MySQL and HTML</p> <p>Website created and tested to gather approximately 75 user ratings and other information such as age, sex and mood for songs in the database. Information will be gathered and stored in multiple MySQL database tables to be used by the recommender system</p>	Andrew	11/15/2003	01/17/2003
6.	<p>Integration of web server and database.</p> <p>In order to maximize the effectiveness of the system, the Web Server must interact quickly with the database. Communication process must be designed such that data will remain accessible for future sessions by a user.</p>	Bernard	12/14/2003	01/24/2004

7.	<p>Integration of web server and application</p> <p>In order to maximize the effectiveness of the system, the Web Server must interact quickly with the application. Integration will focus on the speed of communication between the modules. Extensive testing must be done to demonstrate the correctness of the data exchanged. Modifications to the modules will be made if necessary.</p>	Bernard	12/14/2003	01/24/2004
8.	<p>General Testing of web server / application / database functionality:</p> <p>As integration of web server, application and database progresses, general tests of the system will be conducted to determine boundary exceptions, broken links, correctness of algorithm implementation, and ease of communication and speed between modules.</p>	Andrew	01/17/2004	02/13/2004
9.	<p>Integration and testing of application and database.</p> <p>Enabling the Java code to successfully retrieve and update data to the database. Ensure that communication between the two components is smooth.</p>	Cavan	01/04/2004	02/13/2004
10.	<p>Coordination of Research Schedule.</p> <p>Decision on two methods of testing will be decided, numbers of test subjects to be used, and types of measuring scales to be used to determine accuracy and usefulness of recommendations.</p>	Andrew	01/30/2004	02/13/2004
11.	<p>Research, comparison and implementation of alternative recommender algorithms.</p> <p>Comparison of "active" versus "non-active" approaches to collaborative filtering. Analyze their performances and accuracy to real life usage.</p>	Cavan	02/13/2004	03/05/2004
12.	<p>Oral Presentation:</p> <p>Summarization of progress will be condensed into presentation format. Visuals and aids will be considered and created to provide an effective emphasis and attract audience. Rehearsal of presentation and analysis of strengths and weaknesses. Results of testing will be added in as tests are completed.</p>	Andrew	02/13/2004	03/26/2004
13.	<p>Design Fair Poster Presentation</p> <p>The poster will attempt to give an overall view of our design project while keeping in mind the audience will mostly be comprised of 3rd year ECE students. The results of our research will be displayed through the use of charts and graphs. A computer will also be available to provide a hands-on demonstration of our system.</p>	Bernard	02/27/2004	03/16/2004
14.	<p>Completion of group final report</p> <p>Integration of the documentation of all components which make up the design project including diagrams, figures, references.</p>	Cavan	03/16/2004	04/08/2004