A Game with a purpose for annotating Greek Folk Music in a Web Content Management System

Emmanouil Giouvanakis, Constantine Kotropoulos, Antonis Theodoridis, Ioannis Pitas

Department of Informatics
Aristotle University of Thessaloniki
Thessaloniki 54124, Greece
Email: egiouvan@csd.auth.gr,costas@aiia.csd.auth.gr,antochat@csd.auth.gr,pitas@aiia.csd.auth.gr

Abstract—A web content management system (Web CMS) has been developed to host a corpus of 405 folk songs for media sharing. It is based on Drupal 7. Exploring various modules, the system offers multiple functionalities, including user registration, layered information about each song in different pages, similar content search, friendship relationships between the registered users, favorite lists, etc. Users can visit song pages, listen to the songs, and provide feedback. In addition to the Web CMS, a game with a purpose (GWAP), called Erasitechnis GWAP, has been designed and deployed on Facebook as an enjoyable way encouraging users to enroll and provide music annotations. Erasitechnis GWAP is a combined approach of existing games with a purpose, which focuses on tagging Greek folk music promoting less frequent tags. By doing so, descriptive tags are collected enabling the creation of a fully annotated dataset, which is exploited to train recommendation or autotagging systems.

Index Terms—Web content management systems; games with a purpose; tagging; computational ethnomusicology;

I. INTRODUCTION

Currently, Music Information Retrieval (MIR) tools have mainly been developed for western popular and classical music. However, the interest for non-Western music continuously grows as is evidenced by the increasing number of papers in recent MIR conferences. Computational methods for automatic classification and topological clustering of large databases of folk music are described in [1]. A platform that extracts and explores pitch annotations in non-Western music, providing musicologically meaningful representations can be found in [2]. Projects like CompMusic, which deals with Turkish, Chinese, and Indian music [3], is an example of a computational ethnomusicology approach.

Greek folk music extends far back in time. It consists of many types of compositions, usually characterized by their place of origin, where they are performed or created (Greek islands, Epirus, Pontus, Thrace, etc.). Apart from territorial criteria, folk songs can also be classified according to their content and the occasion upon which they are performed. Greek folk music covers the whole spectrum of social life including Christmas, carnival, Easter, wedding, and love songs. During joyous occasions, such as on New Year's Day and Sunday before Lent, big dances would take place where people dance and sing, usually accompanied by a band. Historical songs (cleftique) and songs for sad events, are also available.

Many initiatives have been undertaken by various bodies to publish Greek folk music such as the Musical Folklore Archives Melpo Merlie[1], the Research Programme "Thrace and Macedonia"[2], and the various music editions by the Crete University Press[3].

In this paper, we are interested in launching a new initiative by building a web content management system (Web CMS) that enables the information gathering and presentation of each song by distilling existing information sources, such as those mentioned earlier. The Web CMS is based on Drupal 7[4]. For the time being, it hosts a corpus of 405 folk songs for media sharing. The Web CMS enables extracting useful information from the "behavior" of registered users.

Another goal is to collect tags for each song resorting to crowd intelligence, also known as social computation. It has been proved that using webforms on each song page for this purpose is not user-friendly. This led us to design a game with a purpose (GWAP) linked to the Web CMS that makes tagging not only easier, but fun as well. Indeed, using human experts to manually tag hundreds or millions of songs would be a costly and time-consuming process. GWAPs were first introduced by von Ahn et al. [4] as an approach to collect semantic information for multimedia data, such as images, audio, and text. They deal with this problem by giving non-expert users a more entertaining way to contribute to the collection of tags for multimedia data. GWAPs are mainly deployed online, since users can be easily attracted at a low cost. Especially, if they are deployed on a social network, such as Facebook[5], useful metadata can be obtained. For example, content created by users, resource annotations, and friendship relationships between users can be exploited to build a new generation of not only GWAPs, but Web-based applications in general [5]. Erasitechnis GWAP focuses on tagging Greek folk music promoting less frequent tags. Combining various techniques, borrowed from other approaches, a new game has been designed, which enjoys most of the benefits of the existing GWAPs in one.

---

[1]http://www.mla.gr
[2]http://epth.sfm.gr
[3]http://www.cup.gr
[4]http://drupal.org
[5]http://facebook.com

The rest of the paper is organized as follows. In Section II, related work on GWAPs surveyed. In Section III, the Erasitechnis Web CMS is presented, describing the corpus and implementation details. A music annotation game for Greek folk music, called Erasitechnis GWAP, along with implementations details and preliminary results, and a first application to Greek Folk music recommendation can be found in Section IV. Conclusions are drawn in Section V.

II. Related Work on Games With a Purpose

The games with a purpose, released so far, are used for several purposes such as labeling images to improve web searching, adult content filtering, internet accessibility, etc. Specifically for music GWAPs, their goal is either to collect data to tag world's music or to obtain a significant amount of well-tagged songs, which can be used for autotagging all the other songs later on. In the following, a survey of four such games is undertaken.

The ESP Game [6] is an online interactive system in which two players both see the same image and try to guess what their partner is typing. No communication between players is permitted, so the best way for players to move to the next round is to type something relevant to the image they both see. Every time they agree on an image they receive some points and the next round begins. On each round, taboo words (i.e. words that they cannot use), are displayed next to the image. There is a threshold that defines how many partners have to agree on a tag to be appended to the list of taboo words. The game was first launched on August 9, 2003 and in three months time, 13,630 users played the game resulting in 1,217,451 tags for 293,760 images.

Verbosity [7] aims at collecting a database of "common-sense facts". It is meant to be played online by two players. One of the players is chosen as the "Narrator", while the other is the "Guesser". The Narrator describes a secret word, he is given and the Guesser based on the descriptions of his partner tries to guess the word. During the game play, the system collects data from the descriptions and the answers the players give. Over 30,000 unique users have played Verbosity. In fact, it is still among the most popular games in GWAPs site[6].

In TagATune [8], two players are listening to a song and based on the descriptions of their partners they are asked to figure out if they are listening to the same or different song. Points are given only if both players reach on an agreement. The words used by the players to describe the song are assigned as official tags according to their frequency of appearance. Bonus rounds during the game play are available. Based on statistics collected in mid-December 2008, TagATune recorded a total of 49,088 unique games which equals to 439,760 normal rounds. During these rounds, 512,770 tags were collected, 108,558 of them were used by at least two players and 70,908 were unique.

Finally, Herd it [9] is a social game for music annotation and was deployed on Facebook. It does not allow free tagging
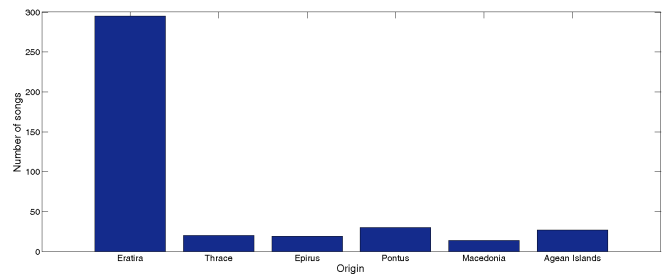

Figure 2. Number of songs for each origin.

but players are asked to complete mini games that require just one click and last 10-20 seconds. For example, players might me asked to choose adjectives contained in bubbles in the game area that best describe the song they are listening to. When the game begins, the player has to choose a genre and automatically is connected in real time with all the other players listening to the same genre. The same song is played to all players. Herd it was released on April 15, 2009 into public beta-testing and 1,049 players were attracted and played 13.831 rounds of the game resulting in 9,941 song/tag pairs.

III. A Web Content Management System For Greek Folk Music

A Web CMS was built, based on Drupal 7, to host our corpus. The URL to the Web CMS is: http://erasitechnis-aiia. web.auth.gr/music. Drupal 7 is a free and open-source content management framework (CMF) written in PHP. Its core contains all the main features of a content management system, such as user account registration and maintenance, system administration, etc. Drupal offers through modules many ways for extending the main functionalities. By exploring various modules, nodes/pages have been created for every song containing information in groups and webforms to collect data are provided to registered users. Furthermore, the system is made social by allowing the users to make friends and store favorite songs in their account.
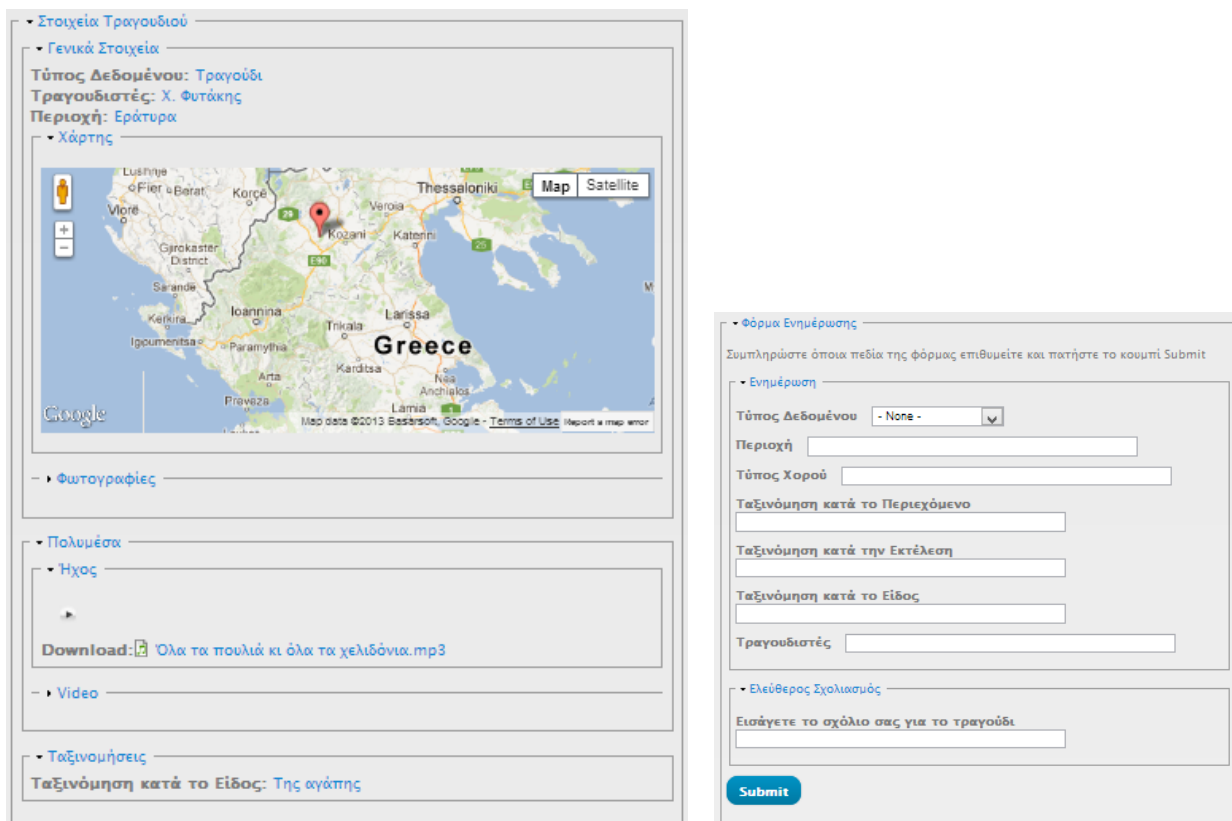
The corpus initially contained about 300 folk songs from Eratyra, Greece[7]. Currently, folk Greek songs from Epirus, Macedonia, Thrace, Pontus, and the Aegean Islands have also been included (see Fig. 2). We foresee that this corpus will be enriched with more folk songs in the future.

Each song has a variety of data fields. Song details are grouped in three categories: *general information*, *multimedia* and *classifications*. In the first group, data are provided about the song *place of origin*, *data type*, *singers*, *dance type* (if the song can be danced), *lyrics*, *geolocation*, and *photos* of the song place of origin, retrieved from Flickr[8]. The second group contains the *audio file* of the song and an embedded *video* from Youtube with a 30 seconds sample of the original song. Finally, classifications by *content*, *type*, and *performance* appear in the third group.

---

[6]http://www.gwap.com/gwap/

[7]http://poseidon.csd.auth.gr/GR/Siniatsikos/index.htm
[8]http://flickr.com

(a) Song information groups. From top to bottom: i) *Γενικά Στοιχεία* (*General Information*): a) *Τύπος δεδομένου* (*Data Type*) *Φωτογραφίες* (*Photos*), ii) *Πολυμέσα* (*Multimedia*): a) *Ήχος* (*Audio*) b) *Video*, iii) *Ταξινομήσεις* (*Classifications*): a) *Ταξινόμηση κατά το Είδος* (*Classification by Type*)

(b) Webform for collecting users' feedback. From top to bottom: i) *Ενημέρωση* (*Update Feedback*): a) *Τύπος Δεδομένου* (*Data Type*) b) *Περιοχή* (*Place of origin*) c) *Τύπος Χορού* (*Dance Type*) d) *Ταξινόμηση κατά το Περιεχόμενο* (*Classification by Content*) e) *Ταξινόμηση κατά την Εκτέλεση* (*Classification by Performance*) f) *Ταξινόμηση κατά το Είδος* (*Classification by Type*) g) *Τραγουδιστές* (*Singers*), ii) *Ελεύθερος Σχολιασμός* (*Free text tagging*)

Figure 1.   Song details and feedback.

After Drupal's installation, a content type has to be defined in order to provide a data structure for each song. This content type, called "Song", holds all the data fields mentioned in the previous subsection. Thus, all songs were stored in the Web CMS in nodes/pages (see Fig. 1(a)) of content type "Song". The following modules were installed and used during implementation.

Drupal 7 supports user account registration, but using the *Social Login*[9] module, login via users' accounts in famous social networks such as *Facebook*, *Twitter*[10], *Google Plus*[11], and *Linkedin*[12] is enabled as well.

Flag[13] module has been used in two ways. Firstly, by creating a favorite flag, users can flag a song as favorite, thus, storing it in their favorite list. Secondly, friendships between the registered users have been developed by creating a friend flag.

Through *Webform*[14] module, forms have been created and attached to every song's page. Webforms contain fields for gathering users' feedback for absent or wrongly registered data in the system and fields, where the users can submit descriptive tags for the songs (see Fig. 1(b)).

The *Views*[15] module has also been used providing a flexible method to control how lists and tables of content, users and other data are presented. This tool is essentially a smart query builder that can build the proper query, execute it, and display the results, given enough information.

Finally, the *Tracker* module, included in Drupal's core by default, has been used to track users behavior and store information about the songs a user has visited.

The result is a new fully functional Web CMS with a number of social features. Users can overview the details of every song by visiting its page, while listening to it and inserting tags.
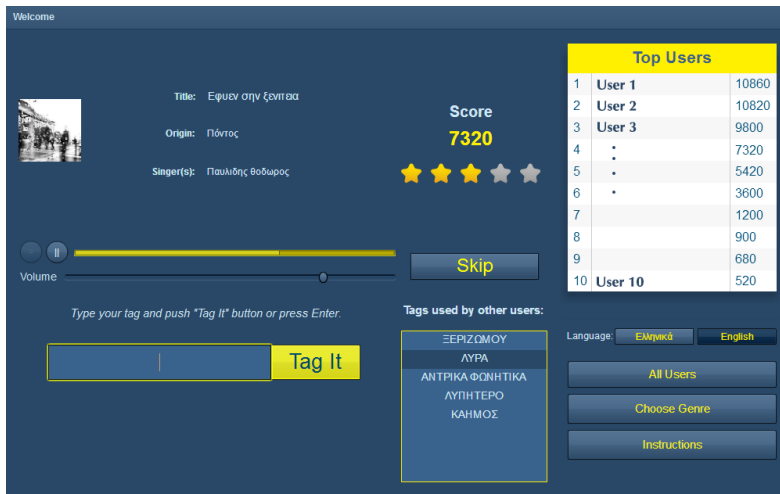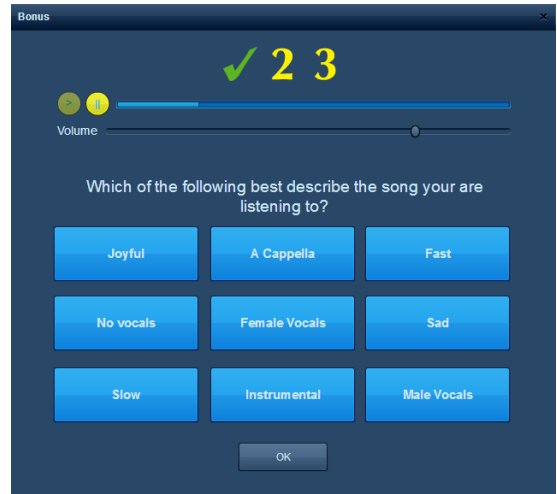
[9]http://drupal.org/project/sociallogin
[10]http://twitter.com
[11]https://plus.google.com
[12]http://linkedin.com
[13]http://drupal.org/project/flag

[14]http://drupal.org/project/webform
[15]http://drupal.org/project/views

(a) Main game.

(b) Bonus round.

Figure 3.   Erasitechnis GWAP interfaces.

IV. Erasitechnis GWAP

Preliminary attempts to massively collect tags through the Erasitechnis Web CMS were rather discouraging. To reverse the situation, the Erasitechnis GWAP has been designed that focuses on the tagging process in a more entertaining way for the users.

Erasitechnis GWAP has been developed on Facebook. Thus, users can login using their personal account and no other registration is required. Unlike the Web CMS, in Erasitechnis GWAP, songs are grouped into categories with respect to their place of origin, which is defined as their genre. The game can be found in the following URLs: http://apps. facebook.com/thaliserasitechnis (Facebook Application), http://erasitechnis-aiia.web.auth.gr/app (Web Application).

A. Game play

Our initial goal was to design a game, which acts as a supplementary tool to collect tags for our Web CMS. The game had to be amusing and easy-to-play in order to attract users asking them to spend time on tagging songs. Thus, the Erasitechnis GWAP had to be: (i) appealing and attractive by designing a simple but complete game interface; (ii) competitive by letting users score points kept in their personal account; (iii) personalized by giving users the ability to choose the genre they would like to listen to before the games starts.

As it was already said, Erasitechnis GWAP is a combined approach that gathers most of the benefits of the so far launched GWAPs in one. By the time one player arrives at the game, he is asked to login and choose what genre he would like to listen to. There are various genres available with respect to the songs' place of origin.

When a genre is chosen, the game begins. On each round, as shown in Fig. 3(a), a song is randomly selected and information about its title, place of origin and singer(s) is displayed. A music player is available to users with pause
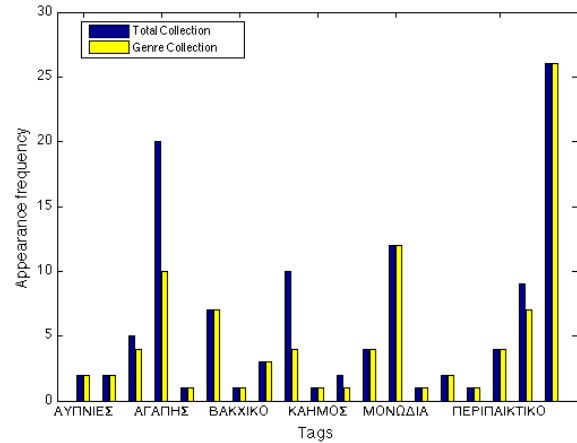

Figure 4.   Appearance frequency of tags for one song.

and play buttons along with a controller to adjust the volume to the desired level.

Each player has his own score and a top-ten-players grid is always visible during the game. To gain points, players have to type tags that best describe the song they are listening to. A list of tags previously submitted for the chosen song is displayed. These tags are chosen among all tags used for this song by firstly calculating the appearance frequency of each tag in the entire corpus and the tag frequency within the collection of songs of the same genre (see Fig. 4) and by keeping only those tags whose their frequencies don't exceed a predefined threshold. Users can double-click on a listed tag and submit it as their own. We think that this list, along with the details provided for each song, are useful hints that help players figure out what type of tags they could possibly type.

Every time a tag is successfully submitted, points are added to the player's score. Points are inversely proportional to the

tag frequency of appearance. Starting with 100 points for completely new tags, points are reduced by 20 for every instance of the same tag proposed by other players. This way, players are urged to insert new tags to gain more points resulting in a much wider collection of tags for every song.

The game uses two criteria to decide if a tag is accepted or not. The first one suggests that one player cannot use the same tag for the same song more than once. The second one aims at maintaining diversity. A threshold is defined for the maximum number of times a tag can be submitted for a particular song from different users. This means that if, for example, the threshold is set to 5, the tag "Fast" for the song "Ikariotiko" can be used from up to 5 users. Either way, upon rejection, a message appears asking players to propose a different tag and no points are added to their score.

Additionally to the points given for every successfully submitted tag, a star is earned. Up to 5 five stars can be collected on every round. Earning the last star results in the beginning of a bonus round.

During the bonus round, which lasts for 3 songs, players listen to a song and have to choose among 9 predefined tags, those that best describe the song they are listening to. The interface of the bonus round is shown in Fig. 3(b). Giving users the ability to select predefined tags rather than typing them makes the tagging process faster and easier. No points are scored for every tag in bonus round, but a significant amount of 1000 points are added to player's score upon bonus round completion. When a bonus round is finished, a new song is loaded and the main game continues. Along with the basic features of the game, a complete list of all users with their scores, instructions of the game, and buttons to change the interface language from Greek to English are available.

Every successful tag submission is stored in game's database in form of [*User*, *Song*, *Tag*]. For example ['*324*', '*45*', '*fast*'] triplet can be read as "User with id '*324*' inserted the tag '*fast*' for the song with id '*45*'". Mapping tags to songs offers great semantic information which can be later on used on music or tag recommendation systems.

B. Results

Erasitechnis GWAP was released on beta-testing on January 1, 2013 and since then we collected 1,500 tags, from a vocabulary of 438 unique stems, for 405 songs. The tags gathered form the Erasitechnis Web CMS (424 tags) are significantly fewer than those collected through the Erasitechnis GWAP (1076 tags). In Table I, the top 10 of the most frequently used tags are listed along with their translation in English. As it was expected, the most frequent tags are generic ones that can best describe a large portion of the corpus. In Table II, informative and interesting tags found in the tail of frequency distribution are mentioned.

C. Exploitation for Greek Folk Music Recommendation

Music recommendation problem is expressed as a hypergraph ranking problem, introducing group sparsity constraints. A hypergraph $G(V, E, w)$ is defined as a set of vertices V and

Table I
TOP 10 MOST FREQUENTLY USED TAGS

Tag	English Translation	Count
αργό	slow	71
χορευτικό	dance song	59
χαρούμενο	joyful	57
ανδρικά φωνητικά	male vocals	44
γυναικεία φωνητικά	female vocals	43
χωρίς φωνητικά	no vocals	36
γρήγορο	fast	31
λυπηρό	sad	28
φωνητικό	vocals	26
αγάπη	love	24

Table II
TAGS IN THE TAIL OF FREQUENCY DISTRIBUTION

Tag	English Translation
δημοφιλές	popular
ευχάριστο	pleasant
θρησκευτικό	religious
καντάδα	serenade
μελαγχολικό	melancholic
μοιρολόι	funeral song
νύφη	bride
παιδικό	nursery song
παράδοση	tradition
βιωματικό	experiential

hyperedges E, to which a weight function $w : E \rightarrow \mathbb{R}$ is assigned [10]. Each hyperedge $e \in E$ contains an arbitrary number of vertices $v \in V$, which is defined as the hyperedge degree $\delta(e) = |e|$, with the set cardinality denoted as $|.|$. Ordinary graphs could be described as hypergraphs with a hyperedge degree equal to 2. Similarly, the degree of a vertex v can be defined as $\delta(v) = \sum_{e \in E|v \in e} w(e)$. Let $\mathbf{H} \in \mathbb{R}^{|V| \times |E|}$ be the vertex to hyperedge incidence matrix, having elements $H(v, e) = 1$ if $v \in e$ and 0 otherwise. The following diagonal matrices are defined: the vertex degree matrix $\mathbf{D}_u$, the hyperedge degree matrix $\mathbf{D}_e$ of size $|V| \times |V|$ and $|E| \times |E|$, respectively as well as the $|E| \times |E|$ matrix $\mathbf{W}$ containing the hyperedge weights. The ℓ_2 norm of a vector is denoted by $\|.\|_2$ and $\mathbf{I}$ is the identity matrix of compatible dimensions.

Let $\mathbf{A} = \mathbf{D}_u^{-1/2} \mathbf{H} \mathbf{W} \mathbf{D}_e^{-1} \mathbf{H}^T \mathbf{D}_u^{-1/2}$ and $\mathbf{L} = \mathbf{I} - \mathbf{A}$ be the positive semi-definite hypergraph Laplacian matrix. For a real valued ranking vector $\mathbf{f} \in \mathbb{R}^{|V|}$, one seeks to minimize $\Omega(\mathbf{f}) = \frac{1}{2} \mathbf{f}^T \mathbf{L} \mathbf{f}$ requiring all vertices with the same value in ranking vector $\mathbf{f}$ to be strongly connected [11]. Vertices participating in many common hyperedges should have similar ranking values. The just mentioned optimization problem was extended by including the ℓ_2 regularization norm between the ranking vector $\mathbf{f}$ and the query vector $\mathbf{y} \in \mathbb{R}^{|V|}$ in [12].

The vertex set V in the hypergraph is made by the concatenation of sets of objects of different kind, such as users, user groups, tracks, or tags. Let each set of objects define

a group. Clearly, each group contributes differently to the ranking procedure. The replacement of the ℓ_2 norm regularization by the Group Lasso term was proposed in [13]. If the hypergraph vertices are grouped into S non-overlapping groups, the ranking recommendation should be optimized by assigning different weights γ_s, $s = 1, 2, \ldots, S$ to each group, yielding the following objective function to be minimized:

$$Q(\mathbf{f}) = \Omega(\mathbf{f}) + \vartheta \sum_{s=1}^{S} \sqrt{\gamma_s \, (\mathbf{f} - \mathbf{y})^T \mathbf{K}_s (\mathbf{f} - \mathbf{y})}. \qquad (1)$$

In (1), ϑ is the regularizing parameter and $\mathbf{K}_s$ is the $|V| \times |V|$ diagonal matrix with elements admitting the value 1 for the vertices, which belong to the s-th group. The recommendation problem can be expressed as finding $\mathbf{f}$, which minimizes $Q(\mathbf{f})$ in (1). By doing so, one can control how the different data groups (i.e., sets of hypergraph vertices) affect the recommendation process.

A very primitive dataset containing 227 tracks, 235 tags, and 41 users has been used for Greek folk music recommendation. Three groups were included: tags, tracks, and users. The 20 mel frequency cepstral coefficients (MFCCs) have been used to encode the timbral properties of the audio signal. They are calculated by employing frames of duration 23ms with a hop size of 11.5 ms and a 42-band filter bank. A Gaussian Mixture Model (GMM) is created for each track with 30 components trained using the Expectation-Maximization (EM) algorithm The distances between the GMM's are computed by using the Earth Movers' Distance, yielding the audio-track (content) similarities. Based on these similarities, the tracks are organized in 10 clusters. At the hypergraph construction phase, a binary incidence matrix whose entries are equal to one for the tracks belonging to the same cluster with each track and a weight matrix is defined containing the distances between the tracks associated to the non-zero entries of the incidence matrix. Preliminary averaged recall-precision curves for 25 users are shown in Fig. 5. The QGS-THALIS(1) curve sets larger weights to tags or tracks than the users. The weights for users, tags, and tracks are set to 0.3, 0.9, and 0.9, respectively. The QGS-THALIS(2) curve weighs more heavily the users than the tags or tracks by setting the respective weights to 0.9, 0.4, and 0.4. The reduced accuracy of the QGS-THALIS(2) curve verifies that the sparse user information does not contribute efficiently to the recommendation process. By adding more hyper-relations to the hypergraph, the recommendation accuracy could be further improved.

V. Conclusion

A Web Content Management System and a music annotation game called Erasitechnis GWAP have been presented. Implementation details have been discussed. The preliminary use of the game by a limited number of users has demonstrated that it contributes significantly to increasing the tag collection. In the future, we shall enrich and expand the corpus with even more songs and evaluate the tags collected, hoping we could use them in Greek folk music and tag recommendation systems. A new version of Erasitechnis GWAP that is based on
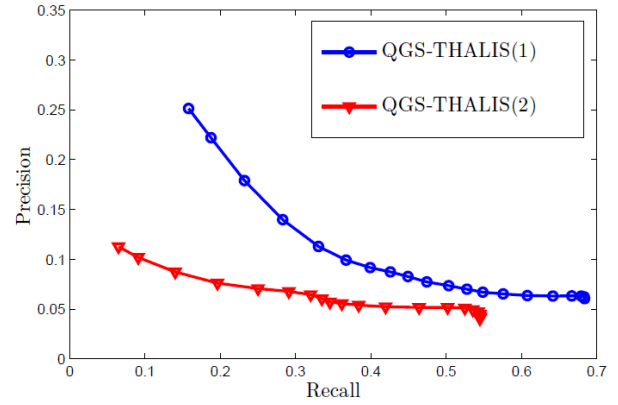

Figure 5.   Averaged Recall-Precision curves for 25 users.

term frequency-inverse document frequency (*tf-idf*) has been released in *http://erasitechnis-aiia.web.auth.gr/appv2*

References

[1] Z. Juhász, "Low dimensional visualization of folk music systems using the self organizing cloud," in *Proc. Int. Conf. Music Information Retrieval (ISMIR 2011)*, 2011, pp. 299–304.

[2] J. Six and O. Cornelis, "Tarsos: a platform to explore pitch scales in non-western and western music," in *Proc. Int. Conf. Music Information Retrieval (ISMIR 2011)*, 2011, pp. 169–174.

[3] X. Serra, "A multicultural approach in music information research," in *Proc. Int. Conf. Music Information Retrieval (ISMIR 2011)*, 2011, pp. 151–156.

[4] L. v. Ahn, "Games with a purpose," *Computer*, vol. 39, no. 6, pp. 92–94, Jun. 2006.

[5] W. Rafelsberger and A. Scharl, "Games with a purpose for social networking platforms," in *Proc. 20th ACM Conf. Hypertext and Hypermedia*, 2009.

[6] L. von Ahn and L. Dabbish, "Labeling images with a computer game," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, 2004, pp. 319–326.

[7] L. von Ahn, M. Kedia, and M. Blum, "Verbosity: a game for collecting common-sense facts," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, 2006, pp. 75–78.

[8] E. Law and L. von Ahn, "Input-agreement: a new mechanism for collecting data using human computation games," in *Proc. SIGCHI Conf. Human Factors in Computing Systems*, 2009, pp. 1197–1206.

[9] L. Barrington, D. O'Malley, D. Turnbull, and G. Lanckriet, "User-centered design of a social game to tag music," in *Proc. ACM SIGKDD Workshop Human Computation*, 2009, pp. 7–10.

[10] C. Berge, *Graphs and Hypergraphs*.   Elsevier Science Ltd, 1985.

[11] S. Agarwal, K. Branson, and S. Belongie, "Higher order learning with graphs," in *Proc. 23rd Int. Conf. Machine Learning*, 2006, pp. 17–24.

[12] J. Bu, S. Tan, C. Chen, C. Wang, H. Wu, Z. Lijun, and X. He, "Music recommendation by unified hypergraph: combining social media information and music content," in *Proc. ACM Conf. Multimedia*, 2010, pp. 391–400.

[13] A. Theodoridis, C. Kotropoulos, and Y. Panagakis, "Music recomendation using hypergraphs and group sparsity," in *2013 IEEE Int. Conf. Audio, Speech, and Signal Processing*, 2013.