

Using Game to Enhance Quality of Music Tags

Peter Dulačka*

Faculty of Informatics and Information Technologies
Slovak University of Technology in Bratislava
Ilkovičova 3, 842 16 Bratislava, Slovakia
dulacka@gmail.com

Abstract

Quantity of music metadata on the Web is raising with high speed. It became a real challenge to keep quality of metadata at reasonable level. In this paper we present a game with a purpose called City Lights for music metadata validation, discuss possibilities to achieve more accurate results by using expert recognition and reevaluation.

Categories and Subject Descriptors

H.3 [Information Storage and Retrieval]: Content Analysis and Indexing; K.8 [Personal Computing]: Games

Keywords

game with a purpose, human computing, multimedia, music information retrieval, metadata validation

1. Introduction

The need for multimedia metadata used in categorization and search is today's actual issue. The general groups of approaches to acquire metadata are: (1) automatic acquisition, (2) crowdsourcing or (3) experts involvement. However, even after their combination, the results are not always correct and we need to validate acquired metadata. Our work focus on music and music metadata validation.

The only way to do the validation of gained metadata is to use human labor, which people are not willing to do for free. To motivate people, we propose Game with a

purpose (GWAP) called City Lights¹. The game provides fun as a sufficient motivation for crowds to spend time by validating annotations. GWAPs are games being used to solve problems that machines cannot solve accurately enough, but humans can solve them without a hard effort.

The first games with a purpose were created by Luis von Ahn. ESP Game was a multi-player labeling game for acquiring image annotations [3], Tagatune was a game which obtained music metadata by player's input agreement [1]. Morton et. al. presented Moodswings, which obtained mood of songs [2] by mapping mood into 2D graph. Besides multiplayer games, there have been different approaches too - Šimko et. al. presented a single-player game for creating folksonomy-like network relationships [4]. By analyzing the existing solutions we decided to focus on following aspects of our GWAP: (1) attractive level design and simple gameplay, (2) proper and motivating scoring and (3) building up a player base.

2. Music metadata validation

Our game focuses on validation of existing music tags. Player is given several sets of fetched tags; each set relates to different song. She then hears a part of a song and has to decide, which of the given sets relates to song she is listening to (and gains or loses points respectively). Using her decision, the game is able to determine correctness (*support*) of the presented song-tag relations. Design of the game interface is shown at Figure 1. Unlike other GWAPs, the game does not suffer from cold-start problem [4], because of its single-player design and no need for initial player base.

Each tag starts with initial value of 0 (*support*) and every player action moves it up or down. Tag is validated/removed when it reaches preset support limit. After a player's attempt, there are two possible scenarios. *Incorrect attempt*, when (1) provided tags for the song being played may not be accurate enough or/and (2) tags for different song are accidentally better describing, than provided ones. Player has option to mark tags which persuaded her to select wrong set - they become possible tags for song being played. After an *correct attempt*, player has possibility to mark tags that misled him. Considering bonus as a motivation, we assume that player marked most of misleading tags and tags not marked are correct - we alter support of every tag according to its status. Even when the game works with tag sets, it is able to evaluate tags separately (possible by random tag selection in every game).

*Bachelor degree study programme in field Informatics. Supervisor: Jakub Šimko, Institute of Informatics and Software Engineering, Faculty of Informatics and Information Technologies, STU in Bratislava. Work described in this paper was presented at the 8th Student Research Conference in Informatics and Information Technologies IIT.SRC 2012.

© Copyright 2012. All rights reserved. Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from STU Press, Vazovova 5, 811 07 Bratislava, Slovakia.

Dulačka, P. Using Game to Enhance Quality of Music Tags. Information Sciences and Technologies Bulletin of the ACM Slovakia, Special Section on Student Research in Informatics and Information Technologies, Vol. 4, No. 2 (2012) 30-31

¹accessible at <http://bit.ly/city-lights>

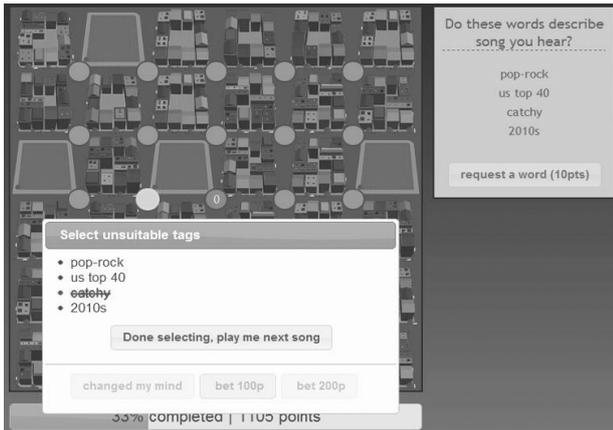


Figure 1: Game interface: music player (left), game board (centre), tags container (right), game log (bottom), window for tag marking (on top)

The power of our game lies in possibility to validate tags without mentioned optional steps, just by considering level of certainty stated by player in attempt and its correctness. However we encourage players not to ignore optional actions by giving them score rewards for completing them.

3. Evaluation

We evaluated our method over 150 songs and 100 annotations for each song. Annotations were fetched from public LastFM database and song previews were played from 7Digital library. We generated five different rounds and introduced game to five test subjects (both male and female in age of 18-30). The experiment contained 20 counted rounds in total. Every set of tags contained four random tags.

We were able to distinguish annotations according to their quality even with scoring rules to be set very strictly. By only 20 player rounds we got rid of annotations such as: *elotmbgmegamixx*, *test*, *nice*, *favorite*, *good lyrics*, *fab*, *etc*. These are either very subjective tags or complete nonsense. On the other hand we validated tags as: *female vocalists*, *love*, *british*, *singer - songwriter*, *pop rock*, *etc*. Some are subjective, but usable at global scale and some are objective and should be validated.

4. Discussion and future work

Our game is able to validate the tags even after couple of rounds, which makes our approach really effective. However, we consider several future work issues to achieve more accurate results.

As first, player expertise should be considered. If we consider genre as an attribute of player, we could prevent a situation where the player does not understand the meaning of genre-specific tag and can therefore hinder the tag validation process. There are two scenarios to achieve this: (1) let the player choose her favourite genre and provide generic genre songs or (2) use data gathered by music social networks and provide songs of player's taste, which could increase her enjoyment and raise quality of obtained results because of familiar tag domain.

Knowing that a player is familiar with certain music domain, we can further measure his level of expertise. The following features could be considered:

- Time needed to make a correct decision. Value can vary because of different quality of tags and should to be decided by comparison other players times.
- Number of tags needed for a correct attempt. An expert should be able to make a correct decision even with the low number of given tags.
- Success rate in choosing correct annotation sets.
- Number of higher (and successful) bets. By placing bets player states confidence about the attempt.

After the separation of expert players from common ones, we can improve our results by retrospective reevaluation of finished games by simulating recorded games with different parameters (prefer recognized experts). With the knowledge of expert actions we can differentiate tags support changes and let actions of expert players have bigger impact on these changes over generic actions. If some players try to manipulate game outputs, we can ignore them (and their actions) completely.

As mentioned, special features have to be presented to motivate players to play the game and to return in the future (some are used in most of GWAPs, some are game-specific and were proposed by players themselves). In our future work, we aim to focus specifically on:

- Separate *per-genre* scoring. By separating genres and high-scores charts, players are motivated to explore unknown songs, which increases their excitement and makes the game usable in multiple genres.
- Different game modes. Variety of game possibilities is necessary in order to maintain the fun effect. Besides classic mode we propose "*Endless game*" (game path length is not predefined constant) and "*Fast game*" (usage of one tag only instead of whole set).
- Tokens for one-time access to special features. Diversity and usage of bonus items helps the game in achieving its goals - game encourages players to fulfill its goals and rewards them for that with tokens. They can be used for in-game song skipping, importing of own songs, nickname change etc.
- Multiplayer mode. Even when high-scores charts are present, players demand direct comparison with others.

Acknowledgements. This work was partially supported by the Slovak Research and Development Agency under the contract No. APVV-0208-10.

References

- [1] E. Law, L. von Ahn, R. Dannenberg, and M. Crawford. Tagatune: a game for music and sound annotation. In *Proceedings of the 8th International Conference on Music Information Retrieval (ISMIR 2007)*, 2007.
- [2] B. G. Morton, J. A. Speck, E. M. Schmidt, and Y. E. Kim. Improving music emotion labeling using human computation. In *Proceedings of the ACM SIGKDD Workshop on Human Computation, HCOMP '10*, pages 45–48, New York, NY, USA, 2010. ACM.
- [3] L. von Ahn and L. Dabbish. Designing games with a purpose. *Communications of the ACM*, 51(8), 2008.
- [4] J. Šimko, M. Tvarožek, and M. Bieliková. Semantics discovery via human computation games. *Int. Journal on Sem. Web and Inf. Systems*, 7(3):23–45, 2011.