Games and Crowds: Authority Identification

Jakub ŠIMKO*

Slovak University of Technology in Bratislava
Faculty of Informatics and Information Technologies
Ilkovičova 3, 842 16 Bratislava, Slovakia
jsimko@fiit.stuba.sk

Some of the today's computational tasks are still subject to human labor, because computational machinery paradigms are unable to deal with them (esp. in terms of quality). These tasks include metadata acquisition and domain modeling, the two essential processes needed for enabling effective and adaptive hypermedia systems. Hence, a whole field of crowdsourcing-based (and game-based in particular [2]) approaches has emerged to do the job.

However, this field has also its own problems. It is dependent on mass participations of users and meanwhile, it is usually not very effective in using this power as it is based on redundant task solving in order to filter out incorrect task solutions. We believe that the effectiveness of crowdsourcing approaches can be improved through authority identification, i.e. identification of trustworthy contributors with more experience in particular domain, whose problem solutions should be taken with greater weights, assuming their high correctness probability [1]. Authority identification has not been sufficiently addressed yet, and within the domain of games with a purpose (GWAP), it completely absents. We aim to explore the possibilities of authority identification within crowd-based, collaborative and gaming metadata acquisition systems.

In the GWAP domain, our preliminary experiments have demonstrated the increased potential of GWAP solutions that include player domain expertise considerations. During the experiments with the PexAce game – a GWAP for image annotation – we have let the players to work over their own, personal images, rather than general ones. Besides the players were more attracted to the game, they have also been more productive in their annotation efforts. They were able to provide more valuable specific metadata, such as concrete person or event names, not just general descriptions of objects. All this was enabled solely by "smart" game content assignment.

The above approach is, however, bit supervised – the player must provide his own content, or someone else must assign, what content is best for the player to play with. To explore whether there is a way to implement a less supervised approach, we

* Supervisor: Mária Bieliková, Institute of Informatics and Software Engineering

Spring 2012 PeWe Workshop, April 1, 2012, pp. 81–82.

experiment with tracking of the player expertise in different subdomains of the GWAP content. We plan several synthetic experiments over the game logs collected in the *Little Search Game* and *PexAce* GWAPS, in which we want to show, that by measuring the player performance (score gain), we can subsequently build up "player models" and assign later task instances to them with respect to their fields of expertise. The overall hypothesis is, that the convergence to correct problem solutions would become faster, which will spare the human cycles to other tasks.

As for the general crowdsourcing, an analogous experiment is currently underway in the e-learning domain (a software engineering course), where students identify correct and wrong question answers. As an input to the process, we use the last year dataset of question-answer combinations created by students during semestral mini-tests. During the experiment, students are presented with question-answer combinations which they use as learning exercises. Students subjectively evaluate the correctness of answers on the scale 0 to 1. Afterwards, they receive the feedback in the form of average correctness of the answer, computed from other students' evaluations provided so far.

While the primary hypothesis is that such process would bring up correct answer evaluations (which we will validate by the existing real teacher evaluations), we also aim to prove that measurement of the student skills based on his past exercises can improve the crowd-based filtering tasks, if applied during solution voting procedures. In other words, we will measure the performance of students and compute the student rating, which will later serve as a weight of their votes for next question-answer evaluations.

We believe that expertise-aware extensions for the GWAPs and crowd-based applications could significantly improve their performance. An open issue, however, is the possible general (in)applicability of these principles for certain approaches, since the recognition of user models may fragment the crowd to possibly too small groups, where individuals share the same expertise. Then, the approaches that rely heavily on some sort of online collaboration, or cross-user artifact validation (e.g. multiplayer GWAPs) would experience only minor improvements. On the other hand, approaches less dependent on direct collaboration and possessing other means of ensuring output quality (e.g. single player GWAPs) could benefit much more.

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

References

[1] Raykar, V. C., Yu, S., Zhao L. H., Valadez, G. H., Florin, C., Bogoni, L., & Moy, L. (2010). Learning From Crowds. *The Journal of Machine Learning Research*, *11*, 1297-1322.

[2] Šimko, J., Tvarožek, M., & Bieliková, M. (2011). Semantics Discovery via Human Computation Games. *International Journal on Semantic Web and Information Systems (IJSWIS), 7*(3), 23-45.