Improving Community Rating in the Tele-Lecturing Context

Franka Moritz, Maria Siebert, Christoph Meinel
Hasso-Plattner-Institut für Softwaresystemtechnik
Universität Potsdam
Potsdam, Germany
franka.moritz|maria.siebert|christoph.meinel@hpi.uni-potsdam.de

Abstract

A lot of content for tele-lecturing portals was produced in the last decade. But metadata to filter the content was not generated adequately. It therefore takes a long time for users to filter through all the videos available until they find the learning content they are looking for. That is why it is a new challenge to find solutions how to filter the large amount of data with a small base of metadata available. To engage the user community to generate metadata is one option.

This paper compares two rating functionalities as one community feature to enhance search functionalities in tele-teaching portals. The efficiency of the two algorithms is tested against each other using a plug-in-architecture to switch between the two options. One of the algorithms simply calculates a mean rating of all ratings for one content item. The other one applies a rating over several layers of content items and therefore generates more precise results and automatically rates connected content items.

1 Introduction

The learners' main challenges in nowadays society concerning learning are the limited time available and the huge amount of resources and data. Tele-teaching was introduced as one solution were people can learn independent from time and place according to their interests and learning speed. But even in those closed environments the amount of data increases rapidly due to recording technology, like tele-TASK [1], becoming cheaper as well as easier and faster to use. Community rating can be applied to the tele-lecturing context in order to support more precise search and content filtering options for learners and therewith improve the quality and speed of their learning. The usage of rating for e-lectures was motivated and the technical and learning-related issues explained in a previous paper [2].

The utility of the rating algorithm over several layers suggested in the before-mentioned paper will now be compared to a standard rating functionality in terms of the preciseness of the results and speed of the evaluation. The results of the old implementation may be compared to the new version as all results were transformed for the new version. One major problem with the old rating implementation was the lack of participation of users. If only few users rate the content, features that make use of the rate, like search result filtering or recommendation engines, will not work properly because of a lack of data. The new rating algorithm also wants to address this issue.

As sample the rating functionality was implemented at the tele-teaching portal tele-TASK[1] of the Hasso-Plattner-Institut (HPI). As the tele-TASK project includes a recording system as well as a portal for distributing e-lectures, some details of the project will be explained in the next paragraph.

2 Tele-Teaching with tele-TASK

The tele-Teaching Anywhere Solution Kit [1], short tele-TASK, is an e-learning project at the chair Internet-Technologies and -Systems at the HPI. The tele-TASK project was started in 2002 at the university of Trier by developing a hardware system for lecture recording. The goal of the project is the recording and distribution of lectures, seminars, reports and other presentations with as little as possible effort of material and resources.

Therefore an all-in-one solution was developed including hard- and software for lecture recording. Two video steams (a video of the lecturer and screen capturing of his laptop or a smart-board) and one audio stream can be recorded at once. More than 3200 lectures and 7500 podcasts of the tele-TASK archive can be accessed free of charge via web-browser or portable device. The large video archive and the web-platform tele-TASK are the basis for further research and development at the HPI.

A topic that is within the research focus currently is the utilization of community and social web functionalities to

[1] http://www.tele-task.de

enhance the tele-lecturing for students. Rating as one community functionality is introduced and a the application for tele-lecturing explained in the next chapter.

3 Rating in Tele-Lecturing Portals

This chapter will briefly give an introduction into the topic of social web and motivate the combination of tele-lecturing with concepts derived from the social web. A definition of rating, the easiest community feature, is given and the application of rating for the tele-lecturing context explained.

3.1 Community and Social Web Functionalities in Tele-Lecturing Scenarios

Since the beginning of the Web 2.0 [3] era numerous social web portals have evolved and grew very quickly. Their main motivation is fostered around the user participation. A number of social web and community features have been found to be useful to the users. These include blogging, the collaborate creation of wikis, social annotation and tagging, evaluating (eg. rating and commenting), recommending, content sharing and linking of content items [3, 4].

That community functionalities are not only useful for networking, but also for learning context was already found out at the beginning of the e-learning era around 2000 [5, 6]. But only recently research focused on joining tele-lecturing with community functionalities. During the workshop *eLectures 2009* at the conference DeLFI 2009 [7] an approach of integrating tele-lecturing applications into facebook and other social e-learning approaches were shown.

The following paragraph will introduce rating as the fastest to implement and easiest to use communitiy feature.

3.2 Introduction to Collaborative Rating

Rating is “a classification according to order or grade” [8]. In the context of the rating of media items, rating is the quantification of the personally perceived quality of an item. It belongs to community functionalities which originate from Web 2.0 platforms.

Rating is the user-generated enhancement to standard metadata that is easiest for the users. It is usually a small set of integers where the user chooses one of the values. The evaluation of content in this manner is therefore an easy and quick process for the user which he might be more willing to go through than a more time intense process like writing comments or annotations. Facilitating the engagement of users is an important issue in this context as the user participation is usually not very high. A study about the web 2.0 video service YouTube [9] and also experience with the example portal showed this.

Once the rating is accepted by the users, it will facilitate the search in the content as the search results can be ranked according to the ratings. Furthermore it can be used for recommendation systems. If several e-lectures are available as related content to be shown in the recommendations list for a tele-teaching item, the ratings could again be used as one of the factors for ranking items for each related topic.

The idea of the rating over several layers of content items is briefly introduced in this chapter and aspects of the implementation described.

3.3 Calculating the Rating Over Several Layers

In the tele-teaching context there are several content layers where rating can be applied as visualized in figure 1. Usually such a portal consists of lecture recordings that are held by lecturers. The lectures itself are mostly embedded in a larger context, for example the course which runs a whole semester, here called series. Furthermore the lectures are often subdivided into smaller pieces, called scene in this paper. This is done in order to facilitate the usage of mobile players where the content needs to be downloaded, for podcasting and also to simplify a more precise metadata collection and search [10]. As all the three layers include tele-teaching content, all of them should be rateable individually.

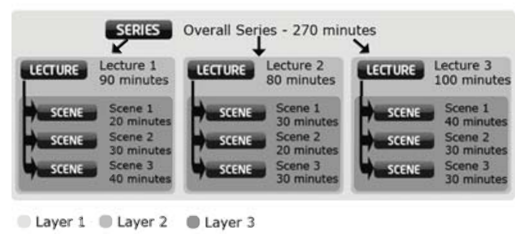

Figure 1. Several content layers in tele-lecturing portals

As the rating across several layers is used, a way of reliably calculating a result that reflects the ratings across the different layers needs to be thought of. The following paragraph shows a calculation that combines the ratings across several layers.

As rating uses a pre-set interval of values that the user can choose and which are used to calculate the mean afterwards, deviant values need not be considered and average calculations like median or truncated mean need not be used to ensure a valid result. As the rating is furthermore not a mean value that is calculated with a factor that includes a relative reference to another unit and no changing rate is required, the arithmetic mean is the mean calculation of choice for ratings. Because the rating shall be calculated across several layers a weighting of the subset ratings is required. The weighted mean (WM) rating of a content item will be calculated by combining and weighting the means (M) of all ratings for the content item and the ratings for its connected content items of the layers underneath and above.

Equation (1) shows how the arithmetic mean of all ratings for one content item is calculated. This equation is the basis for all further calculations of the mean rating that consider a weighting.

$$M_{CSin} = \frac{\sum_{i=1}^{p} R_p}{p} \tag{1}$$

The calculation of the weighted mean for one layer of connected content items (for example all segments that belong to one lecture or all lectures that belong to one series) is shown in equation (2). The factor for weighting the different arithmetic means that were calculated in (1) is the length of the content items. One example: a lecture which is 30 minutes long consists of 3 segments, the first is 5, the second 10 and the third 15 minutes long. The mean rating for the longest segment should have most influence on the weighted mean calculation for the lecture and the other two have lower priority. Equation (2) calculates the combined mean of one layer of content items (as for example the before mentioned three segments) by weighting the means of the single segments with their length.

$$WM_{CLay} = \frac{\sum_{i=1}^{n} L_{CSin_i} \cdot M_{CSin_i}}{\sum_{i=1}^{n} L_{CSin_i}} \tag{2}$$

The overall calculation of the weighted mean for one content item considering all connected layers underneath and on top is shown in equation (3). It follows the same principles as equation (2), but it uses the means of all layers that were calculated with equation (2) and combines them to a weighted mean. The factor for weighting is also a different one now. As one content item has the same length as the sum of all connected items in the layer underneath, the length is no proper weighting factor in this case. The number of ratings is the factor that determines which mean ratings have which prioritization. But as all the segments (which are used for podcasting) together will most certainly receive more ratings than the single lecture they belong to, the ratio of prioritizing only by number of ratings would minimize the effect of the mean rating of the single content item. Therefore the ratio of the number of ratings to the number of content items of the layer is used as weighting factor to combine the means of the different layers.

$$WM_{CSin} = \frac{\sum_{i=1}^{m} \frac{NoR_{CLay_i}}{NoC_{CLay_i}} \cdot WM_{CLay_i}}{\sum_{i=1}^{m} \frac{NoR_{CLay_i}}{NoC_{CLay_i}}} \tag{3}$$

CSin = Single content item
CLay = All content items in one layer
p = Number of ratings per content item
n = Number of content items per layer
m = Number of layers
R = Rating
L = Length of the content item
M = Arithmetic mean of all ratings for one content item
WM = Weighted mean
NoR = Number of ratings
NoC = Number of content items in this layer

3.4 The Rating Algorithm Implementation

The first implementation for rating is a simple one. It focuses on the forms for generating rates by the user and provides only simple evaluation of the results. This implementation uses the average function of the underlying database (see listing 1) to calculate the score of one object. This implementation can even be used inside the search function, but it does not regard the influences between the different data layers.

```
votes = Vote.objects.filter(contentType =
    ctype.id, objectId = data.id).aggregate(
    average=Avg('vote'))
```

Listing 1. Simple rating calculation

Therefore the implementation of the described calculation method was done parallel. We used a plug-in architecture as described in [11] to add the rating functionality. This allowed us to implement the second version on the base of the same database tables as the first one using the data gained so far. Both versions use the same database and can be displayed together on the page. Using the plug-in architecture we can switch the functions on and off, so it is easy to compare both versions and provide the better one in the live version of the portal.

For the calculation of the results of one lecture two nested loops were required. To calculate each included lecture the average of the containing segments has to be calculated as well. The first idea of doing this calculation each time when displaying the results would consume a big amount of calculation time.

So we were forced to think about a better way to implement it. The requirements are less time consumption and ideally the possibility to use the rating results for search result ordering. When thinking of the rating it became obvious, that saving a new rating results happens less often, than displaying a result. Therefore we decided to do the calculation of the results at saving time instead of displaying time.

For saving the results we had to create an extra database table which contained only the calculation results for every object. To reduce the number of calculations, we use these saved intermediate results for calculation as well. This results in the following workflow:

1. If the object which is rated is a segment, then the average rating result of the segment is determined. Af-

terwards the new rating result of the parent lecture is calculated as described in (2). At last the rating result of the corresponding series is calculated as described in (3).

2. If the rated object is a lecture, the average rating results of all segments are retrieved from the database, together with the number of ratings. From the database the metadata of the segments, like the duration, are also fetched. Each rating of a segment is now weighted with the ratio of the duration of this segment and the overall duration of all segments. These values from all segments are added together and represent the average rating of all segments.

   Similar to the calculation of the average of a segment, the average value of the lecture is calculated using the average function of the database. With both values the overall average rating of the lecture is calculated using the number of rates concerning the lecture and the segments.

   Afterwards the average of the series is recalculated like described in (3).

3. If the rated object is a series, then the average rating results of all lectures of this series are retrieved from the database, as well as the number of rates and the lecture metadata. Now the same calculation process starts, which is described in (2), using the data of the lectures like the segment data and the series data like the lecture data.

   Because the data of the segment is included in the rating result of the lecture, it does not have to be calculated again.

Our calculation approach therefore brings the effect, that it minimizes the number of calculations because saved results can be used. It also helps the usage of the rating data for ordering the search results, because the order criteria is directly inside the database and therefore available for the database request itself.

4 Evaluation of the Rating Functionality

Several advantages are expected by using the new rating algorithm. First more exact results and second the automatic calculation of ratings for superordinate content items in the different content layers and therefore more rating results are expected. This is necessary, because although the rating functionality has been online for about six month, only 459 of the about 11400 content items (about 7700 segments, 3300 lectures, 400 series) have been rated and only 25 of those have more than two ratings and can therefore be utilized for rating filtering (average ratings with less than three ratings by different people will not be taken in consideration for fairness reasons). In this paragraph both of these theses are evaluated. Advantages and disadvantages of both calculation methods are compared.

4.1 Advantages of the new rating algorithm

When comparing the number of rating results shown, the algorithm calculating the rating over several layers shows a lot more content items than the standard algorithm. This means that with the help of the calculation over several layers content items that have not been rated by users so far or not enough ratings have been collected to make up an objective results receive a rating calculated via the new algorithm (which can be seen in figure 2). As consequence more rating results will be available as metadata base for further processing within features that use rating to improve the usability and searchability of tele-teaching portals. The last chapter will explain the usage of this metadata base in more detail.

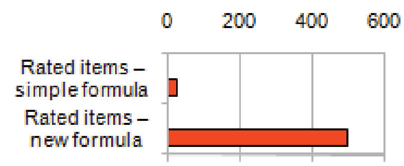

Figure 2. Number of rating results when using the new rating algorithm compared to the old

Second the rating results will be more precise. This is the case because a weighted mean calculation consists of more detailed ratings of the content layers underneath in connection with ratings to that specific content item. An evaluation of the current metadata base did not show a lot of difference in the results of the two rating calculations. Only 42 out of the 459 currently rated objects have differing results with a nearly equal amount of more positive and negative results (see figure 3).

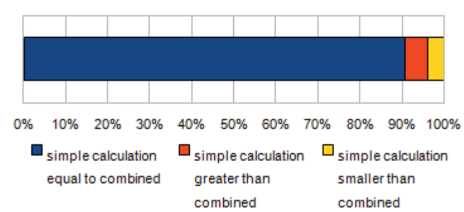

Figure 3. Rating results when using the new rating algorithm compared to the old

This may be the case because of the small interval. When calculating a mean of very deviant values a middle value will be the result and if a lot of people rate in the middle sector the same effect will occur. Therefore most of the mean ratings will be around two to four and the differences between the rating calculations might often only be marginal. Another reason is the ongoing low participation and the therefore small metadata base. There have been 578 ratings for 459 different objects. Only when more connected content items have been rated a proper evaluation on the influence of the new calculation algorithm on the real data can be started.

Until then a sample calculation can show the advantages of the method. Following ratings were inserted for a series with 3 lectures belonging to it:

content item	rating	duration
series	1,1,2,2,1,4,4,3	160
lecture 1	1,1,2,1,1	10
lecture 2	4,4,5,4,5,5,5	90
lecture 3	3,3,3,3,2,4,3,3,2,3	60

Table 1. Sample ratings for connected content items over several layers

In this example the first lecture was only an introduction and therefore not rated very interesting. The last lecture was rated only moderately interesting. But the second one was rated very good. Most of the people who saw the first lecture rated the series as well, later on most students only rated the lectures not the series. Therefore the mean rating of the series itself is not very good (2,25).

Calculation step	result
Simple rating mean calculation	Series = 2,25 Lecture 1 = 1,2 Lecture 2 = 4,57 Lecture 3 = 3
Mean of all lecture ratings weighted by duration	3,77
New series rating calculated over several layers	2,96

Table 2. Sample calculation of rating over several layers for connected content items

Looking at the weighted mean of all the lectures one will get a more positive result of 3,77. When using the calculation over several layers, the final rating result of the series combines the ratings of all lectures and calculates a weighted mean together with the standard mean of the series ratings. As result the series will now be rated more positive and the very positive ratings of the second lecture are reflected in the overall rating of the series as can be seen in figure 2.

4.2 Disadvantages of the new rating algorithm

The major disadvantage of the new algorithm is the increased computing time. At the beginning the new calculation for the rating was performed when opening up each page. This resulted in a huge overhead as a page which included a list of for example 20 content items had to process the rating calculation 20 times and thereby consider all connected content items as well. As solution to reduce the increased computing time, the rating result for the content item and the items connected in the layers above is calculated and stored in the database when a new rating is saved.

4.3 Tradeoff of advantages and disadvantages

When storing the combined rating results for affected content items in the database each time a new rating is entered by a users, the calculation overhead for the new algorithm is minimized. Additionally the results can easily be retrieved each time a content item is shown. This bit of more computing time stands against a large number of more rating results and more precise results. Considering that the user participation for the generation of metadata will not grow exponentially very soon, the required metadata needs to be generated another way. Therefore the usage of the rating calculation over several layers is advisable.

5 Conclusions and Future Work

It was shown that it is senseful to use the suggested rating algorithm across several layers of content items. Now that the decision for one rating algorithm was made, those features that make use of the rating results may be designed and implemented. Some examples of the usage will be addressed in this paragraph.

First of all a content filtering system may utilize the rating results. When a user searches amongst the content of the portal via a keyword, a large amount of content items in all layers of the content hierarchy may be found. The choice which of the search results is the most useful for the student in that moment will be more difficult the more results are found. Therefore it should be possible to filter the results and thereby reduce the total number of them. One filtering option is the rating. The best rated items will be shown first, worst rated items last. As shown by studies with a search engine [12, 13], users only scan through the first few search results and only thoroughly look at the first two to five results. The best rated (and therefore most relevant) results will be on those first pages when using the filtering with the rating results.

Second, a recommendation engine can make use of rating results. When similarities between the current content item and all other items are measured it will often be the case that several relevant items for the same topic or keyword are determined. But neither is there space for a large number of suggestions on the content page nor will the users scan through large numbers of recommended similar content items. When including the rating results into the recommendation algorithm the number of similar content items per related keyword can be reduced by only selecting the best rated item.

In the future the implementation of a group functionality throughout the portal is planned. It was shown that learning in a group is positive for supporting the individuals motivation and eagerness to engage into academic activities [14] and it should therefore be enabled for the digital world too. When the group structure is realized, a lot of features that only worked for either an individual person or all users of a portal will have to be adapted to work for groups as well.

For the rating it will not be possible to have a private rating or a rating for a certain group only. But it will be possible to limit functions that use rating results to the group. One example: a student is member of a group of third semester students. It is very likely that all the members of that group are interested in similar lectures and can make the most use out of similar content items as they study similar courses. Therefore the ratings that another third semester student enters into the portal will most certainly be more relevant for our student than a rating from a sixth semester student, as the student from the same group also has similar previous knowledge to our student whereas the sixth semester student is more advanced.

This means that a recommendation or content filtering system that is limited to influencing factors from within the group will more likely produce more relevant results for all members of that group than one that considers other factors as well. This knowledge can be utilized by offering special recommendation and filtering features for groups as well as special filtering options for search results.

Once the group functionality is available we hope to improve the user participation by supporting more collaboration amongst the students. The more interesting and diverse features can be provided the easier it will hopefully be to attract more users and engage them in an active learning process and an intense usage of the community features. Because only with active participation more metadata will be generated and only then those data will provide real benefits for the users.

References

[1] V. Schillings and C. Meinel, "Tele-TASK – tele-teaching anywhere solution kit," in *Proceedings of ACM SIGUCCS*, Providence, USA, 2002.

[2] F. Moritz, M. Siebert, and C. Meinel, "Community Rating in the Tele-Lecturing Context," in *IAENG International Conference on Internet Computing and Web Services (ICICWS'10)*. Hong Kong: IAENG, 2010.

[3] T. O. Reilly and O. R. Media, "What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software," *Communications & Strategies*, vol. No. 1, pp. 17–37, 2007. [Online]. Available: http://ssrn.com/abstract=1008839

[4] R. Hoegg, R. Martignoni, M. Meckel, K. Stanoevska, and C. Management, "Overview of business models for Web 2.0 communities," in *Proceedings of GeNeMe 2006*, Dresden, 2006, pp. 23–37. [Online]. Available: http://www.alexandria.unisg.ch/Publikationen/31411

[5] R. M. Palloff and K. Pratt, *Building Learning Communities in Cyberspace: Effective Strategies for the Online Classroom*, 1st ed. Jossey-Bass, 1999.

[6] M. J. Rosenberg, *E-learning: strategies for delivering knowledge in the digital age*, R. Narramore, Ed. McGraw-Hill, 2001.

[7] S. Trahasch, S. Linckels, and W. Hürst, "Vorlesungsaufzeichnungen - Anwendungen, Erfahrungen und Forschungsperspektiven. Beobachtungen vom GI-Workshop 'eLectures 2009'," *i-com*, vol. 8, pp. 62–64, 2009. [Online]. Available: http://www.atypon-link.com/OLD/doi/abs/10.1524/icom.2009.0040

[8] "Definition of Rating, accessed on 26/08/2010," http://www.thefreedictionary.com/rating.

[9] M. Cha, H. Kwak, P. Rodriguez, Y.-y. Ahn, and S. Moon, "I Tube, You Tube, Everybody Tubes: Analyzing the World's Largest User Generated Content Video System," in *ICM*, San Diego, California, USA, 2007, pp. 1–13.

[10] A. Groß, B. Baumann, J. Bross, and C. Meinel, "Distribution to multiple platforms based on one video lecture archive," in *SIGUCCS '09: Proceedings of the ACM SIGUCCS fall conference on User services conference*. New York, NY, USA: ACM, 2009, pp. 79–84.

[11] M. Siebert, F. Moritz, and C. Meinel, "Establishing an Expandable Architecture for a Tele-Teaching Portal," in *2010 Ninth IEEE/ACIS International Conference on Computer and Information Science Article*. Yamagata, Japan: IEEE Computer Society, 2010.

[12] L. Granka, H. Hembrooke, B. Pan, T. Joachims, and G. Gay, "In google We Trust: Users' Decision on Rank, Position and Relevancy," *Journal of Computer-Mediated Communication Special Issue on Search Engines 1*, vol. 12, no. 3, pp. 1–36, 2005. [Online]. Available: http://jcmc.indiana.edu/vol12/issue3/pan.html

[13] L. a. Granka, T. Joachims, and G. Gay, "Eye-tracking analysis of user behavior in WWW search," in *Proceedings of the 27th annual international conference on Research and development in information retrieval - SIGIR '04*. Sheffield, United Kingdom: ACM Press, 2004, pp. 478–479. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1008992.1009079

[14] H.-J. So and T. A. Brush, "Student perceptions of collaborative learning, social presence and satisfaction in a blended learning environment: Relationships and critical factors," *Computers & Education*, vol. 51, no. 1, 2008. [Online]. Available: http://portal.acm.org/citation.cfm?id=1371264.1371454