An Improved Design and a Case Study of a Social Visualization Encouraging Participation in Online Communities

Julita Vassileva[1⊕], Lingling Sun[2]

[1] Computer Science Department, University of Saskatchewan, Canada
[2] Solutions AB TELUS Business Transformation, Edmonton, Canada.
jiv@cs.usask.ca, gloriasun@hotmail.com

Abstract. The paper describes a further development of the design of a motivational visualization encouraging participation in an online community. The new design overcomes shortcomings in previous designs, by using more attractive appearance of the graphic elements in the visualization, by giving up the largely unused in the previous design user customization options. The visualization integrates more information in one view, and uses an improved user clustering approach for representing graphically their different levels of contribution. A case study of the new design with a group of 32 students taking a class on Ethics and Computer Science is presented. The results show that the visualization had a significantly effect participation and with respect to two activities (logging into the community and rating resources).

Keywords: participation, online communities, social visualization, evaluation.

1 Introduction

Social visualization approaches using different metaphors have been proposed to stimulate the activation of social norms in groupware and online communities [5, 8]. We proposed [2, 7] a motivational visualization aimed at encouraging participation in an online sharing community. Grounded on the theories of social conformity [1] and social comparison [6], this approach was evaluated in an online community, supporting undergraduate computer science students in a class of Ethics and IT. The experimental results and the user feedback [7] showed that the motivational visualization effectively increased the students' awareness of their community and encouraged social comparison. As a result the contributions of shared resources in the community increased significantly, and participants gave more comments and ratings.

We found that user-customizable views were hardly needed, since most users checked only the default view which displayed users ranked according to the number of resources they shared in the community and didn't explore the other views which showed ranking according to status, login frequency, or number of downloaded resources. This supports Erickson's first guideline [4] for design of community visualization.

⊕ author for correspondence

From a technical point of view, a major issue in the design of the visualization is the algorithm used to classify users into four contribution levels based. Ideally, the algorithm should find a significant difference between the marginal cases on both sides of a boundary between two clusters. More specifically, if the list of nodes is sorted by the original contributions, there could be the case that the last node at the top level only shares one or two files more than the first node at the second level, which may share 10 more files than the second node at this level. A better algorithm should find reasonable gaps between contributions of users to classify these users into different levels. A compromise between desired sizes and sharper boundaries would be a direction to explore in the future.
We observed that as the quantity of contributions increased, their quality somewhat deteriorated. Several users found ways to game the system and exaggerate their nodes in order to gain higher status and visibility. A possible reason for this behavior was that the visualization showed only the quantity of the articles shared by each user regardless of the quality i.e. the visualization did not encourage comparison among the users with respect to the quality of their contributions. Motivating social comparison in the quality of the contributions, comments, and ratings was set as an important future direction of research. Motivating active users to continue their contributions or even increase them was another problem. We concluded that the visualization had to take into account both the quality and quantity of user contributions.
The paper describes a new version of the visualization which takes into account the requirements derived from the previous designs. It provides a minimal user interaction – it requires only the user to select a topic (an area of interest), instead of selecting both a topic and a sorting criterion. By default, the topic is set to the one discussed in the current week of the class, according to the course curriculum, so if the user wants to see the current view of the community, he/she doesn't need to select anything. The semantics of the different sorting criteria used in the previous version are represented into one picture with more complexity and dimensions, which generate visually more attractive and consistent view of the community. The visualization applies a better algorithm to smooth the classification of users into different levels of contributions. The visualization has been redesigned as a web-based application, supported by Apache Tomcat web server v4.0 and mySQL database server.

2 Design

The new design of the visualization had a new, more complex visual language, more attractive stars, a new clustering algorithm. The visualization uses again the nigh-sky metaphor. However, we use a more complex visual language. Unlike the previous design which used just two dimensions: the size (four possible sizes) and color (yellow filled circle or black empty one), in the current design we use four dimensions: size, colour, level of brightness and shade.

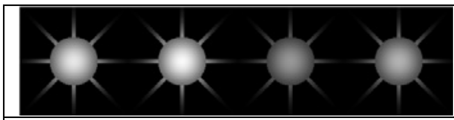
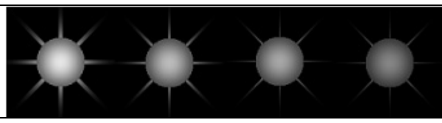
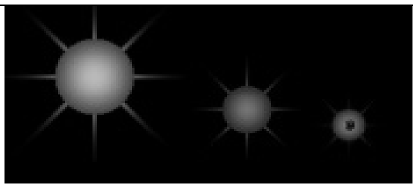
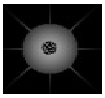
	
a. Different colours denote different memberships (status).	b. Different brightness denotes different quality of contributions.
	
c. Different sizes denote different number of contributions	d. Shaded stars denote users who are offline at the moment.

Fig. 1. The four dimensions of the graphical language

Each star also has a particular color and a certain level of brightness of that color. The color of a star indicates the membership level of the represented user (Fig.1a). The membership (status) is a combined measure of the user's participation which depends on the number and quality of the user's contributions (new links and ratings). A yellow star represents a user who holds Gold membership, a white star represents a user who has a Silver membership, a red star, a Bronze membership, and a Green star, the lowest "plastic" level of the membership (the initial membership level for everyone when s/he first starts to use the system).

Each star has a certain level of color density, which visually appears as the brightness to represent the reputation level of a user (Fig. 1b). There are also four levels of reputation. Brighter stars represent users with higher levels of reputation.

The size of a star indicates the number of links shared by the represented user (Fig.1c). There are four possible levels of contribution: the users who contribute the most links are at Level 1, and the users who contribute no links – at Level 4.

If the center of a star is covered a by a black shade, this indicates that the represented user is currently offline, otherwise, s/he is online (Fig. 1d). In this design, "a user is offline" means that the user has not been active in the past ten minutes in the Comtella community. A user may have a combination of any contribution level, membership level, reputation level, and be either online or offline.

The arrangement of the stars in representing the users in the visualization is fixed (see Fig. 2), while in the previous design it was a result of dynamic sorting according to the criterion chosen by the user. In this way the user can easily locate him/her self as s/he gets familiar with using the system. Each user can create his/her alias, under which s/he is known in the community. The users can see their alias name and the aliases of their peers by moving the mouse on top of a star.

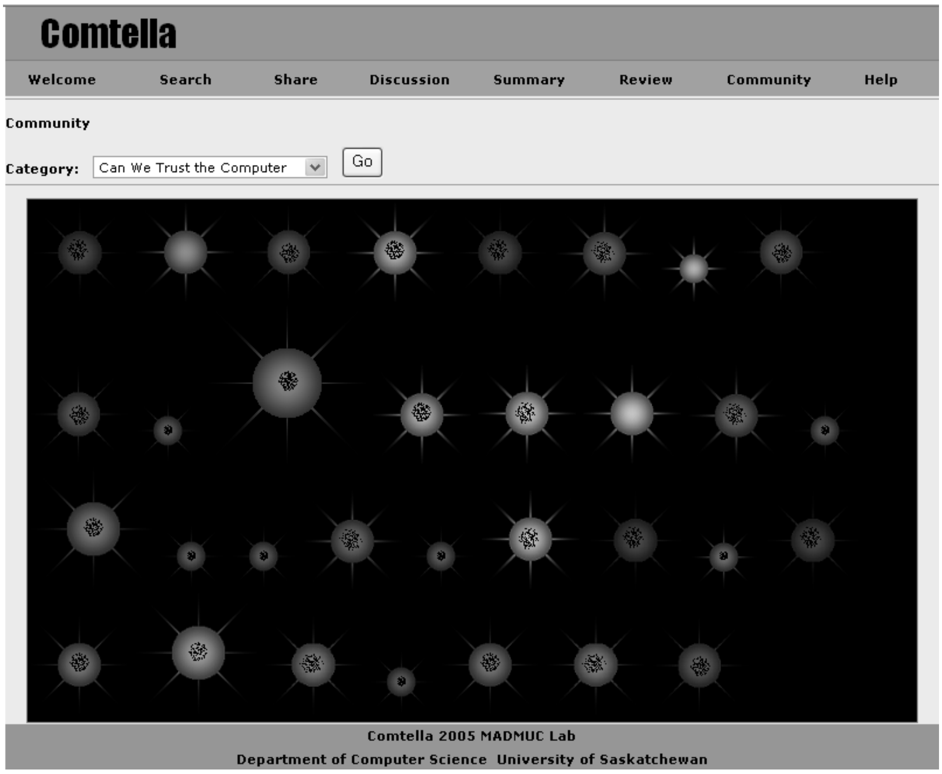


Fig. 2. A screen shot showing the new design of the visualization.

Our hypothesis was that this design will continue to motivate users when they have already become good contributors, which was one of the problems in the previous design. It is almost impossible for a user to achieve the highest levels in all criteria, e.g. having the first contribution level, gold membership, and highest reputation. In this way, there will always be a way of improvement for the user, or a factor that motivates a user to contribute. As Fig. 2 shows the largest star does not necessarily have to be a gold member or the brightest star. This is because the size of a star is solely determined by the number of contributions (shared URLs) by the represented user, while the membership is calculated based on other criteria (the quality of these contributions, as well as by the number and quality of ratings given by the user for the contributions of other users). Some users may feel satisfaction from becoming the brightest small green star, by contributing only a few but highly rated papers.

The images used to represent users in this visualization design are cartoon versions of stars on a black background. In this design we gave up the idea of generating the stars on user request for the goal to having more realistic/beautiful stars. However, unlike the very first design of the Comtella visualization [2, 7], we did not use JPEG images of real stars since they could not be manipulated consistently in terms of colors and brightness to achieve the variety of sizes, colours and different levels of brightness that we wanted to have in the new version. The pictures we used in this design are pre-generated by a program written in OpenGL, and saved as .PNG files after being processed by Microsoft photo editor.

This algorithm was designed to solve the problem of insignificant boundaries between two consecutive clusters of users when classifying these users into different contribution levels. This problem was obvious in the previous design and created feelings of unfairness in the students who were close to the margins of different contribution classes. With the old algorithm, the first contribution level should always contain the top three users, and according to the contributions for topic 1 (i.e. the first column in Fig. 3), for example, some of the zero-contributors will be classified into the first contribution level, while the rest will be classified into the other levels, which is obviously not fair. The new algorithm prevents this unfairness. It is illustrated in Fig. 3, and works as follows:

case 1: Sort in a list (L) the users in descending order of their contributions for a given topic.

case 2: Set everyone who shares nothing with a contribution level = 4.

case 3: If everyone shares something, but they all share the number of files i.e. make the same contribution, set their contribution level = 3.

case 4: Else:

Find the biggest gap in contributions among the top 20% of the users in L and mark it gap_1. For example, if the biggest gap in this range falls between user A and user B, where A is in front of B in L (i.e. A shares more than B), then gap_1 = the index of A in L. Set users before gap_1 with a contribution level = 1.

Find gap_2 which is the biggest gap after gap_1 among the top 50% of the users in L, and set contribution level = 2 to all the users after gap_1 but before gap_2.

Find gap_3, the biggest gap among the rest of the users, and set everyone between gap_2 and gap_3 with a contribution level = 3, and those after gap_3 with a contribution level = 4. However, if there are some users who have not contributed anything, then gap_3 will be the index in L of the last non-zero contributor.

The brightness level is computed using the average reputation of the user's shared URLs (referred to as "paper-reputation" in the following context) defined in [3]. If the highest paper-reputation of all the users, either online or offline, is H then everyone whose paper-reputation is H will have the brightest star (i.e. the highest reputation level). If a user's paper-reputation is less than H, for example r, then another value R is computed as $R = r/H$. If $R > 0.9$ with an allowable margin of 0.05 (i.e. $R > 0.85$) then this user will also have the brightest star; otherwise, if $R > 0.55$ then this user will have a second brightest star (i.e. reputation level 2); otherwise, if $R > 0.35$ then this user will have a dark star (i.e. reputation level 3); and if $R <= 0.35$ then this user will have the darkest star (i.e. lowest reputation level), which makes it almost fade into the background.

topic 1	topic 2	topic 3	topic 4	topic 5	topic 6	topic 7	topic 8	topic 9	topic 10	overall
3	5	14	14	24	28	23	15	20	20	124
0	5	8	10	19	28	15	15	13	10	111
0	4	7	10	16	21	11	10	11	8	102
0	2	6	9	12	15	10	10	8	6	77
0	1	5	9	10	12	8	10	7	6	69
0	1	5	7	8	11	7	6	7	6	58
0	1	5	7	6	11	6	6	6	6	49
0	1	5	7	6	11	6	6	6	6	48
0	0	5	6	6	11	6	6	6	5	47
0	0	5	6	6	9	5	5	5	5	45
0	0	5	6	6	9	5	5	5	5	43
0	0	4	6	5	7	5	5	5	4	40
0	0	4	5	5	6	5	5	5	4	39
0	0	4	5	5	6	5	4	4	3	38
0	0	4	4	5	6	4	4	4	3	35
0	0	4	4	4	5	4	4	4	3	34
0	0	4	4	4	5	4	3	4	3	34
0	0	3	4	4	5	4	3	4	2	33
0	0	3	4	4	4	3	3	3	2	32
0	0	3	4	3	4	3	3	3	2	29
0	0	3	3	3	4	2	2	2	0	28
0	0	3	3	3	3	2	2	2	0	25
0	0	2	2	3	3	1	0	2	0	18
0	0	1	2	3	3	0	0	0	0	15
0	0	1	2	2	2	0	0	0	0	13
0	0	0	1	2	2	0	0	0	0	11
0	0	0	0	0	0	0	0	0	0	4
0	0	0	0	0	0	0	0	0	0	3
0	0	0	0	0	0	0	0	0	0	3
0	0	0	0	0	0	0	0	0	0	0

Fig. 3. An example output of the classification algorithm.

3 Case Study

This design of the visualization was evaluated in a case study with a group of 32 forth-year computer science students taking CMPT 408, a class on Ethics in Computer Science, offered by the Department of Computer Science from January 17 to April 8, 2005, a total of 12 weeks. The first 10 weeks were dedicated to the experiment and the last two weeks were for the online questionnaire survey. The list of categories for sharing URLs corresponds to the topics discussed in the class. Each topic was discussed in one week following the class curriculum, except for "Computer Crime and Security" in the middle of the term which was discussed for two weeks with the reading-week break in between, so this topic ran over weeks 4, 5, and 6.

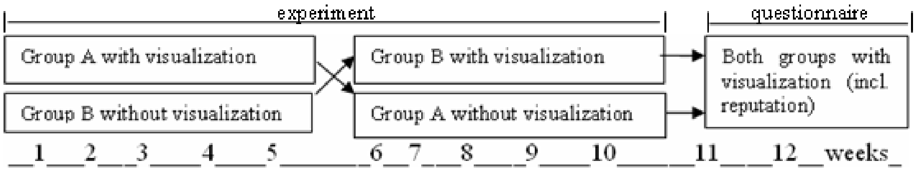


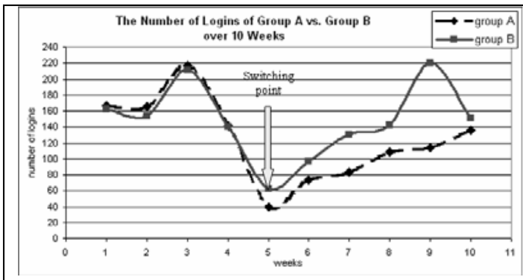
Fig. 4. Experiment Time Schedule

The experimental subjects were divided randomly into two groups of equal size, 16 students in each group, and the experiment duration was split into two equal parts as well, 5 weeks in each part (see Fig. 4). The midnight on Sunday February 20, 2005 was the “switching point” — at this point the two groups were switched so that Group A, who had access to the visualization during the first five weeks, was not able to use the visualization any more, and Group B, who was not able to use the visualization in the first five weeks, gained access to the visualization. The reason for switching the two groups in this way was to reduce as much as possible the ordering effects and the effect of novelty. However, the novelty effect could not be entirely eliminated. In the case of this experiment, it was stronger on Group A than it was on Group B because Group A was the first group who had access to the visualization, so for them both the system and the visualization were new. The subjects in Group A had no knowledge about the visualization when they were exposed to it, but the subjects in Group B had at least heard about the visualization from their colleagues, with whom they shared classes, and worked on the class project. So the visualization was not as new to Group B as it was to Group A.

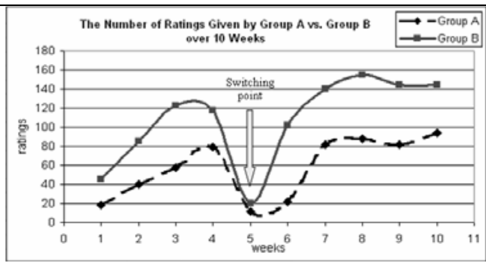
The hypothesis was that the visualization would motivate the subjects to contribute more papers and ratings and to participate more actively in the Comtella online community by logging in more frequently and reading more papers.

The quantitative results about the participation of the two groups are shown in Fig. 5. The dark dashed line in each chart represents the performance of the Group A and the lighter solid line represents the performance of Group B with respect to each activity. The X-axis shows the time duration of the experiment in terms of weeks, starting at Week 1 and ending at Week 10. The Y-axis shows the number of times subjects logged in to the Comtella system. Each data point represents the total number of activities of a given type for all students in the corresponding group and week. The groups were switched at midnight on the last day of Week 5 i.e. the beginning of Week 6. Weeks 4, 5 and 6 were dedicated to the same topic and the students, shared most of their URLs on this topic in Week 4 and almost nothing in weeks 5 and 6. Moreover, Week 5 was the reading-week break. This explains the big drop in Week 5 in each of the figures 5a-d.

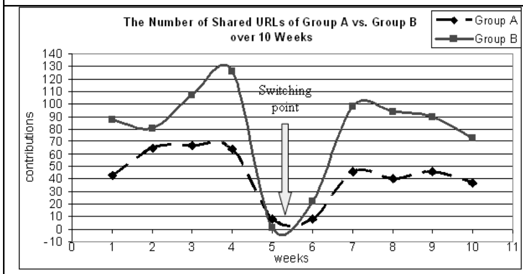
Figure 5a represents the total number of logins made by the subjects each group on a weekly basis and Figure 5b represents the number of ratings given by subjects in each group on a weekly basis. Giving ratings is a major type of activity in the Comtella community. It takes effort to read and evaluate the material, and the rating constitutes a valuable contribution to the community since reasonable ratings will guide users to find good articles. Another important type of contribution is sharing papers (URLs). Figure 5c compares the number of URLs shared by the subjects of Group A with the number of URLs shared by the subjects of Group B on a weekly basis. Figure 5d represents how many times subjects from each group read a paper shared by others in the community (as a read we count just opening the URL of the paper).



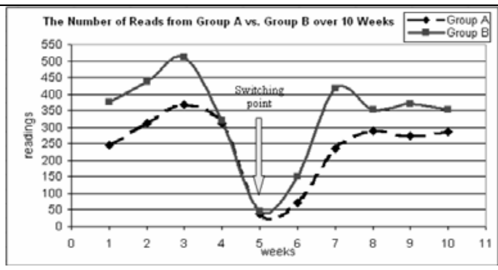
a. The Number of Logins of students from Group A vs. Group B



b. The Number of Ratings Given by Group A vs. Group B



c. The Number of Papers (Shared URLs) contributed by Group A vs. Group B



d. The Number of Read Papers (clicked URLs) by Group A vs. by Group B

Fig. 5. The different types of participation by the students in Group A compared to those of Group B.

It is clear from Fig. 5 that there is a difference between the participation of Group A and Group B. Group B participated more actively than Group A in all activities. Since the experimental subjects were assigned randomly into groups, it happened so that one group contained more active members than the other. In this case, we need to adapt our hypothesis to correspond to this unintended bias. The modified hypothesis, that takes into account the fact that one of the groups (B) is more active is based on the original hypothesis: that the visualization would motivate both groups to participate more. This means that it is expected that the difference between the participation levels of the two groups would be smaller when the less active group has access to the visualization and the more active group does not have access to the visualization (which is the case during the first period of the experiment, before the switch). On the contrary, the difference between the participation levels of the two groups would be larger when the more active group has access to the visualization and the less active group does not have access to the visualization (the case during the second period of the experiment). Figure 6 illustrates the effect of the growing difference between the participation level (we will call it “performance” for brevity) of the two groups in the two periods of the experiment according to the modified hypothesis.

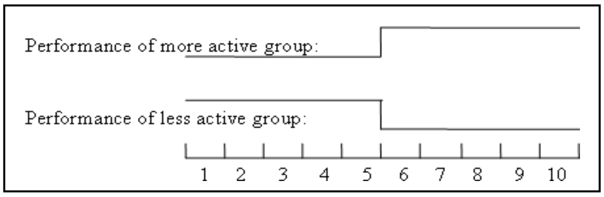


Fig. 6. Modified Experimental Hypothesis.

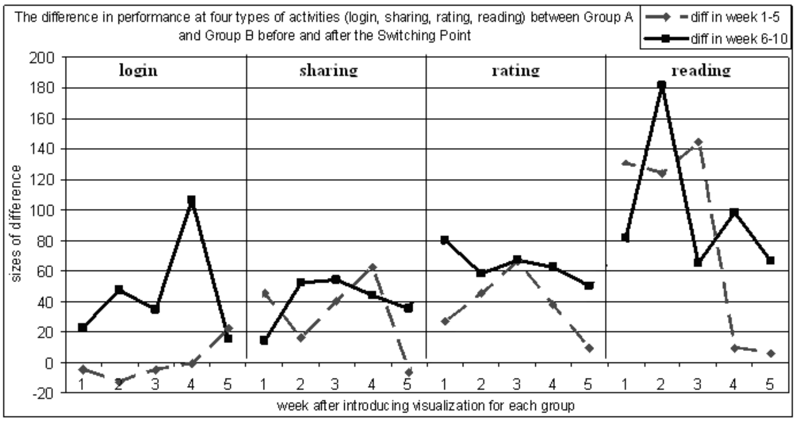


Fig. 7. Differences between the contributions of the two groups in the four activities.

Fig. 7 shows the differences between the contributions made by Group A – Group B for each week and for each type of activity. The X-axis is divided into four sections, each representing the difference in the performances in a particular type of activity (login, sharing URLs, rating URLs, or reading). Each section along the X-axis contains five segments, from 1 to 5, each segment representing a pair of weeks (Week 1 paired with Week 6 as marked by 1, Week 2 paired with Week 7 as marked by 2, Week 3 paired with Week 8 as marked by 3, Week 4 paired with Week 9 as marked by 4, and Week 5 paired with Week 10 as marked by 5). Thus, each point on the darker solid line is comparable to the point on the lighter dashed line. For example, the first point on the solid line represents the first week when Group B had access to the visualization and Group A did not, and the first point on the dashed line represents the first week when Group A had access to the visualization and Group B did not, and so on for the rest of the points. The solid line is mostly above the dashed line which indicates that most of the times the difference in the performances of the two groups after the switching point is bigger than it is before the switching point. This seems to confirm the modified hypothesis.

We performed two tests for statistical significance on the differences in each activity: t-test and the Wilcoxon's Matched-Pairs Signed-Rank Test. We found a significant difference in the performances of the two groups regarding the *login* and the *rating* activities. According to both the t-test and the Wilcoxon's test, the significance for logins is greater than 95%: the significance for rating is 97.5% according to the t-test, and 95% according to the Wilcoxon's test. However, the results for sharing and reading activities are weak. The t-test shows the probability of the difference in sharing activity being random is 29%, and the probability of the difference in reading activity being random is 33%.

The visualization had stronger motivational effect on Group B, the active group, than it had on Group A, the less active group. The reason why the active group was motivated more effectively is not clear. One possible explanation is that if a group is generally more active then the students in this group probably care more about their contribution levels and care if other users see them as good users or freeloaders.

The users provided qualitative feedback in the last two weeks of the term by filling in a questionnaire for which they received a bonus participation mark of 2% towards their final grade. A summary of the user answers to each question related to the visualization are presented below.

1. Please rank the following reasons for which you used the visualization
(1: most important; 5: least important):

	1	2	3	4	5
appears interesting	15%	20%	30%	20%	15%
find articles	10%	0%	20%	20%	50%
compare contributions	30%	35%	15%	5%	15%
check who contributed what	5%	20%	5%	35%	35%
find top contributors	15%	25%	30%	10%	20%

2. Please rank the following (from -2: "very poor" to +2: very good):

	-2	-1	0	1	2
overall	9%	0%	23%	59%	9%
support tool for the class cmpt408	9%	4%	13%	35%	39%
usability	11%	21%	21%	42%	5%
reliability (crashes etc.)	10%	19%	14%	43%	14%
visualization attractive	10%	5%	35%	30%	20%
visualization useful	10%	5%	35%	40%	10%
visualization intuitive	10%	15%	35%	25%	15%
visualization effective	25%	15%	40%	20%	0%
quality of shared links	20%	0%	25%	45%	10%

fairness	10%	0%	10%	65%	15%

Half of the subjects ranked “visualization attractiveness” as positive compared with the first design where only 34% of the subjects ranked this criterion positively and only 18% thought the first design “appeared interesting”. 40% of the subjects thought the visualization was intuitive and 20% thought it was effective; however, the first design gave a slightly better result on the intuitiveness (48% positive ranking) than the final design. These numbers indicate that this version of the visualization is more successful in general compared with the first version.

3. What would be your reaction if you saw yourself as one of the smallest stars (regardless of its color and brightness) in the visualization?

a. Take immediate action: share more links to make your star larger (20%)

b. Think that you should probably share more links, but later (45%)

c. Feel unhappy but do nothing (0%)

d. Feel that the system is unfair, so it doesn't make sense to contribute (0%)

e. Do not care, so will do nothing (20%)

f. Other - please specify: (15%)

The data indicates that 65% of the users were motivated to contribute more if they saw their stars were not big enough in the visualization.

4. If you saw yourself as one of the largest stars (regardless of its color and brightness), would you:

a. Feel proud of your status and try to contribute even more. (40%)

b. Feel proud, but also in some sense "exploited", stop bringing more links. (10%)

c. Feel worried, you may be raising the bar too high and the others may hate you or you may be perceived as "overachiever" by the others. (10%)

d. Feel nothing, since it is not important for me. (35%)

f. Other - please specify: (5%)

Most (55%) of the users were not motivated to contribute more once their stars are big enough in the visualization, and there is some discouraging factor as option b indicates.

5. Please rank the following factors according to how strongly they motivated you to contribute (1: strongest; 6: weakest):

	1	2	3	4	5	6
community visualization	8%	19%	19%	11%	8%	35%

earn higher membership	22%	19%	18%	15%	15%	11%
earn higher marks	43%	21%	4%	14%	11%	7%
bringing good papers	23%	12%	35%	11%	11%	8%
being best user	11%	28%	14%	18%	11%	18%
having best papers	11%	25%	18%	18%	18%	10%

The results show that a significant source of motivation is the social comparison, stimulated by the visualization. Forty-six (46%) of the subjects ranked the community visualization as strong motivator (1 to 3), 53% - being the best user, 54% - bringing the best papers, 59% - the status. Yet, there were other, stronger motivators - the expectation of receiving good marks was ranked as a strong motivator by 68% (these students are obviously extrinsically motivated), and bringing good papers was ranked as strong motivators (1 to 3) by 69% - students who seem to be intrinsically motivated.

6. Did you find the final visualization represents fairly your overall level of contribution in the class?

a. Yes (60%)

b. No (25%)

c. If No, why?

The justification given by the students who selected “b” above was based mainly on the dichotomy between quantity vs. quality of contribution. They thought there should be more emphasis given on the quality of the shared materials. 15% of the subjects were not sure about the overall fairness, and from the given justifications, we found this uncertainty was caused by unfamiliarity with the system.

4 Discussion

Comparing with the feedback from the case study of the previous design of the motivational community visualization [7], these results show that users generally preferred the new design and found it more attractive. The feedback from the case study shows that a higher percentage of users (compared with the case study of the previous design) used the visualization to check who contributes how much and who the top contributors are. The new design of the visualization effectively motivated user contributions in each of the two groups A and B under the test condition. The experimental results confirm the hypothesis that the visualization helps shrinking the difference in the performances of the two groups when the less active group had access to the visualization and the more active group did not, and the visualization amplifies the difference in the performances of the two groups when the more active group had access to the visualization and the less active group had no access to the visualization. The motivational effect is more obvious on the active group than it is on the inactive group. The *t*-test and the Wilcoxon's Matched-Pairs Single Rank Sum test show that the difference in the performances of the two groups before and after the switching point is significant for login and rating activities but not for sharing and reading activities. To conclude on each specific type of activity separately, the experiment needs to be run for a longer period of time, or we need to double the size of the experimental subjects so that we could run the experiment with two groups, one group with the visualization and the other group without, in parallel. Due to the limited class duration (12-13 weeks) the first option is not feasible, but increasing the number of subjects or running experiments in the same class under the same conditions for two consecutive years could be a direction of research.

The users generally found the new visualization design useful and interesting. The effect of the community visualization on motivating contributions and more active participation was shown in both case studies, but the significance of the effect is different depending on what is visualized, how it is visualized (i.e. what graphical representation is used), if it is easy enough for users to read and understand the visual representation (i.e. how intuitive the pictures are), and so on. The results indicate that the visualization is more effective on people who are naturally competitive and care about others' opinions and views on themselves. For people who are not competitive, sociable, or do not care about others' opinions on themselves, the visualization is not an effective motivator, since it was designed to facilitate social comparison. Competition is a form of upward social comparison in which one compares and tries to "fit in" with the elite, top-performing sub-group [8].

One important conclusion is that the simpler the visualization is, the more predictable the effect is. As it was observed in the case study with the first design, users usually do not select any sorting criterion and rely on the default view, i.e. sorted by original contributions, so the node representing each user was only different in size. The nodes remained the same in color and there was no difference in brightness; even if users selected another sorting criterion, the visualization still visualized only one (the selected) criterion at a time. Therefore, the first design was one-dimensional visualization with the dimension determined by user's selection of the sorting criteria, which most of the times was "by original contribution" (the default view). The users related the size of their star with the number of their original contributions and this representation provided a clear direction for social comparison and improvement. That is why the first design was more effective in motivating original contributions, as the experimental results from the first study showed.

In comparison, the second design appears to be less effective than the first design in motivating user contributions in terms of original contributions (new shared papers / URLs). However, the second design was good in motivating diverse contributions, which is probably more desirable than one-dimensional contributions (just in one activity, sharing new papers). A complex visualization showing several dimensions at a time (e.g. size representing contribution level, color representing membership level, and brightness representing reputation level) is interpreted differently by users. Users can focus on different dimensions to compete, rather than one particular area of competition such as the number of contributions, so the motivational effect is dispersed to a variety of activities. If a longer time was available for the experiment and more data for analysis, perhaps we would have seen a significant effect of the visualization on other user activities, i.e. sharing papers, and reading papers.

One clear conclusion for the designer is that when there is a clear goal about which type of contributions or participation is needed for the community, the visualization should represent just the user performance according to this type of participation or contribution. During the lifetime of an online community, different needs arise and different activities should be encouraged at different times [3] and so the default community view should be adapted to represent the activity that has to be encouraged at the moment.

5 Conclusions and Future Directions

Most visualizations discussed in the literature have been created with the purpose of informing the users about activities in the online community, since they allow a quick grasp of complex information. To our best knowledge, there are no other visualizations specifically targeted at motivating user participation in the online communities.

This paper proposed a new, improved design of the motivational community visualization targeted on encouraging participations in an experimental sharing community. The whole experience of developing this prototype visualization tells that it is not straightforward to create a motivational visualization. Apart from the great amount of information that needs to be represented, it has to be easy to operate, intuitive, attractive, and powerful enough to represent different semantics.

There are other interesting directions for further research, including the following:

1. Dynamic adaptation of particular dimensions (e.g. different sorting criteria) that are visualized depending on what is needed mostly in the community (e.g. need more shared files, need better quality shared files, or need more people to rate or comment on the shared files).

2. If decided to present more than one dimension (e.g. size, color, brightness) in the design, it would be better to experiment on one dimension at a time, instead of testing all the dimensions at the same time. The experience from the two major experiments described in this thesis indicates that one-dimensional visualization is easier to be predicted and controlled because of less noise.

3. Exploring the impact on user participation of incorporating more semantics through new dimensions of the star metaphor that haven't been used so far: such as the distance between stars based on, for example, the similarity in taste or ratings given by users.

4. Representing likeness between users, e.g. who reads whose contributions most often, who rates whose contributions most often, who normally rates whose contributions high and rates whose contributions low etc.

5. Investigating the effects of different graphical representations of an online community. The proposed prototype of the visualization in this paper chose a specific metaphor, a staring sky, but there are alternatives, from simple representations such as dots, circles, beehives, tables with numbers, charts, graphs, to complex metaphors such as cities, gardens, or combinations of any of the above. Which particular

representation works best depends on the purpose of the visualization and the online community that it serves (e.g. the age of the members, their attitudes to computers, etc.) Investigating the effect of different metaphors for presenting community information is worthwhile.

6. Creating a more advanced graphical representation, e.g. allowing the navigation in the cosmos, like in a 3-D game. For example, in the second design of the proposed prototype of the visualization interface, it might be possible to group users with similar interests into subgroups and visualize it by a galaxy, clicking on which will cause the expansion of this galaxy and displaying the inside view of this galaxy; or clicking on a star will navigate users to the group of friends of this star (based on what criterion to define a user is a friend of another user could be an interesting research topic), etc.

Acknowledgement: This work is supported by the NSERC Discovery Grant Program.

References

1 Asch S. E. "Effects of Group Pressure upon the Modification and Distortion of Judgments". In Groups, Leadership, and Men, 1951, pp. 177-190.
2 Bretzke H., Vassileva J. "Motivating Cooperation in Peer to Peer Networks", Proceedings User Modeling UM03, Johnstown, PA, Springer Verlag LNCS 2702, 2003, 218-227.
3 Cheng R., Vassileva J. "Design and Evaluation of an Adaptive Incentive Mechanism for Sustained Educational Online Communities". User Modelling and User-Adapted Interaction, special issue on User Modelling Supporting Collaboration and Online Communities, 16 (2/3), 2006, 321-348.
4 Erickson T. "Designing Visualizations of Social Activity: Six Claims." ACM CHI'2003 Proceedings April 5 – 10, 2003 Ft. Lauderdale Florida, USA.
5 Erickson T. and Kellogg W. A. "Social Translucence: Using Minimalist Visualizations of Social Activity to Support Collective Interaction". In Designing Information Spaces: The Social Navigation Approach (eds. K. Hook, D. Benyon, A. Munroe), Springer-Verlag: London, 2003, pp. 17-41.
6 Festinger L. "A Theory of Social Comparison Processes." Human Relations,7, 1954, 117-140.
7 Sun, L., Vassileva J. "Social Visualization Encouraging Participation in Online Communities", In Groupware: Design, Implementation, and Use, Proceedigns of CRIWG'2006, Springer LNCS 4154, 349-363.
8 Garcia S., Tor A. "Rankings and Competition: Social Comparison in the Shadow of Standards", Social Science Research Network Library, 2005. (available at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=880505#PaperDownload).