

# Social Visualization Encouraging Participation in Online Communities

**Abstract.** In order to encourage users to participate more actively and bring more contributions to online communities, we propose to create a motivational community visualization based on the social comparison theory from social psychology. This paper describes the design of a static and a dynamic version of this visualization developed in our lab, explains the advantages and the disadvantages of the static version and the reason why we decided to develop the dynamic version. Results from the evaluation on the dynamic version are presented.

**Keywords:** stimulating participation, online communities, social visualization

## 1 Introduction

The “cold-start” problem is well-known in most online communities, e.g. peer-to-peer (P2P) file-sharing networks, discussion forums, IRC systems, social networking or blogging systems. While some web-based online communities manage to attract users and grow enormously, others never reach the “critical mass” of active users needed to sustain the community and ensure that there are enough new things happening in the community, for example, new shared resources, posts, or blogs which can attract users to revisit the community. Often online communities created to serve a specific role in a certain geographic or institutional context remain unused. It seems like a “chicken and egg” problem: A community is only interesting if many people are participating and contributing; but to get users to contribute, you have to provide an interesting community first.

This paper proposes to motivate users to participate by visualizing the community and the levels of participation of all community members, hoping that the social visibility will enact social norms, stimulating users to engage in responsible and reciprocal behaviour. The paper is organized as follows: in the next section, we give an overview of some of the work on motivation from the area of social psychology as well as other approaches to using visualization in online communities. In section 3, we describe our first approach to design the motivational visualization and the lessons learned from its brief deployment, as well as a second, improved version of the visualization. Section 4 presents the results of evaluation of the second version. These results and the implications as well as directions for future work are discussed in section 5.

## 2 Related research

Most studies on human motivation have been done in the area of social psychology and for real communities. Leon Festinger found in his famous 1954 paper on social comparison [9], that humans tend to compare their achievements and actions with people who they think are similar to them in some way. For example, when a student

wants to know if she is good at math, typically, she compares herself with the other students taking the same math class, rather than with her professor. However, when there is no suitable peer group, people will compare with almost anyone [9]. On the other side, when one knows that others will compare with him/her, one acts more responsibly. People normally want to be positively recognized in their community and are willing to make an effort to gain social reputation, providing the effort is affordable and worthwhile compared with the potential benefit of the reputation.

Asch's conformity study [1] had proved that people generally want to "fit in groups". The fitting behaviour at an interpersonal level happens, for example, when someone sees a friend doing or believing in something and starts believing or doing the same thing. Fitting, at a collective level happens, for example, when one sees trends in the behaviour of others, e.g. the style in dressing, and changes his/her own style of dressing to fit in, even if only for a particular occasion.

It seems that social conformity is a motivator for users in online communities too. Cosley et al.'s experiment [9] in the context of an online movie rating systems also proved that people generally want to "fit in" their group. In this experiment subjects were divided into four groups, A, B, C, and D. Each group was seated in a separate room so that they would not influence each other and were told to rate movies that they have seen before. The users in the first three groups A, B, and C were shown system-predicted ratings for each film they were rating and group D was not. The results showed that a significant proportion of the ratings given by the first three groups of participants correlated with the system-predicted values they were shown, which means that these users tried to conform to the predictions [4]. This experiment indicates that the designers of the software infrastructure of the online community can exploit the phenomenon of social comparison to influence user's behaviour. For example, if they are aware that most of the other users, similar to themselves contribute actively to a community, and that their contributions are lacking, they may be stimulated contribute more to the community [3]. For social comparison to take place, however, users have to be made aware of the behaviour of other users as well as of their own behaviour. Visualization has been used in online communities to create awareness about the other users and the things happening in the community. VisitorVille [16] visualizes websites as cities and visitors as passengers in busses. VIUM [19] is designed to stimulate user reflection on their learning and displays learning concepts as a graph of texts in different fonts, colours and brightness to represent how much the user knows about each concept.

Social visualization approaches using different metaphors have been proposed to stimulate the activation of social norms in the online community. For example, the Babble System [8] visualizes a chat room as a pie with moving dots on the pie representing users to show which users are actively participating in the conversation (those, whose dots are close to the center) and which users are mostly listening (on the periphery). The Chat Circles [21] uses circles filled with different colors and texts representing the conversations. The Task Proxy [7] visualizes task groups in a company as differently-coloured cells in a honeycomb creating conditions for social comparison and therefore some social embarrassment on those groups who are lagging behind. This embarrassment would create social pressure and motivation to perform on par with others.

Erickson proposed a number of guidelines for designing social visualizations for online communities [6-8]. He makes an important distinction between “translucence” and “transparency”, emphasizing that the information showed in the visualization does not necessarily have to be very detailed and exact. In most of the cases, it is better just to give the user a general idea, and even in some cases a certain level of misrepresentation may be beneficial. Also, customizations should be avoided; all users should see the same thing so that they can feel responsible for their actions, since they know that others see the same things as they and are aware of what they do [6].

There are also general design guidelines developed in the visualization community, which are aimed at reducing cognitive overload, using properties of human vision, and usability guidelines. The choice of metaphor is very important, since an appropriate metaphor is intuitive to use and doesn't require a complex legend for interpretation. Applying hierarchical structure [5, 18] and using composable layout and visual sets [15] are always helpful, when designing information-compact visualizations for large networks. Proper use of location and color contraction of visual components will successfully attract attention [12]. “Richly expressive information visualization is difficult to design and rarely found” [11], so it is always beneficial to make the visualization easily reusable in similar situations.

In the next section, we describe two versions of community visualization designed according to the guidelines mentioned above. The goal of the visualization is to stimulate social comparison and to motivate users to contribute resources to an academic paper-sharing online community called “Comtella”.

### **3 Two Visualization Designs for a File-Sharing Community**

Comtella is a paper sharing online community developed in the Madmuc Lab at the University of Saskatchewan. It started in 2002 as a peer to peer, Gnutella-based file-sharing system (originating the name “Comtella”, which stands for “Community Gnutella”) [20]. It evolved through several versions, using different technologies, including a version in 2003-2004 where users shared paper URLs, not files and all servers resided on a server to ensure the simultaneous presence of nodes needed to allow sharing in a small community. Later (in 2004-2005), it was re-implemented as an entirely centralized, web-based system for sharing URLs with a data-base on the server storing contributed URLs. Comtella was used to support students sharing academic papers in the MADMUC research lab and in several (5) undergraduate classes at the Computer Science Department at the University of Saskatchewan. The common feature in all the versions is that the community is relatively small (up to 40 users), closed, and that users share and search academic articles by the “areas of interest” or topics in the class (according to the weekly class schedule). The community visualization for Comtella was included from the very first Comtella implementation [3] and has evolved through two versions, a static and dynamic one. The next sections present the design of these two versions.

### 3.1 Static Design

The first version of the visualization was designed for the peer-to-peer version of Comtella. In this version, there were two important components of participation that were desirable for ensuring a sustainable community:

- sharing a lot of files (new or downloaded from others), which ensured redundancy and availability of resources necessary for the P2P search to work, and
- contributing a lot of new files, to increase the diversity of resources in available in the community.

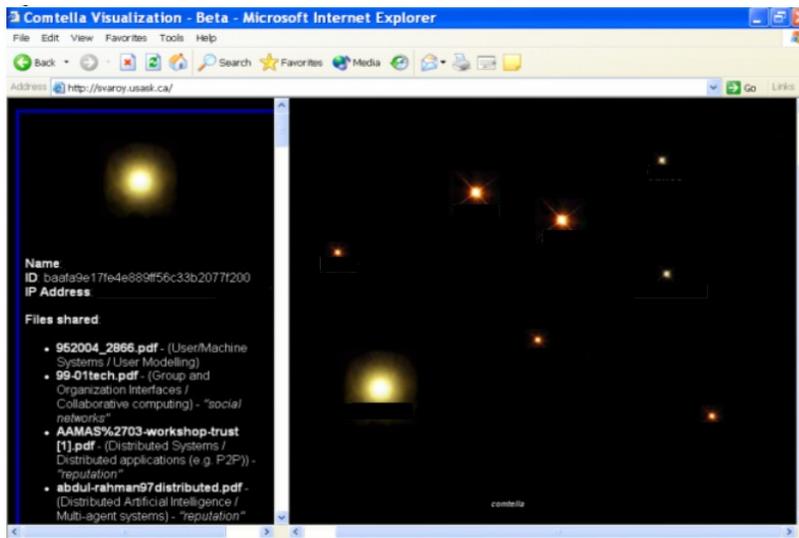


Fig. 1. Motivational visualization - static version (user names and IP address are blacked to preserve the privacy of the participants).

Therefore the visualization showed these two aspects of participation to encourage social comparison among users. It was designed a static webpage, which showed all users online at the moment. The same view was shown to every user, as recommended by Erickson [6]. The visualization (shown in Figure 1), used the metaphor of a mid-summer night sky, first introduced in [2]. The size of a star was determined by the user's number of contributions (files shared in the Comtella community). Users who shared more than the average number of files per user in the community had larger stars, users who shared roughly the same as the average had medium-sized stars, and users who shared less than the average had small stars. The color of a star was either red or yellow. Red stars represented users who shared more new files than the number of files they downloaded from other users in the community, and yellow stars represented users who downloaded from others more than the number of new files they shared with the community. In this way people who enriched the shared resources pool with new files were recognized. The "sun", which

displayed as a big yellow star, represents the “best user” among those who were currently online. The “best user” was the one that shared more than everyone else and has contributed new things to the community, rather than downloading from others

We did not get a chance to carry out a systematic evaluation for this visualization. There were 16 users (graduate students in the MADMUC lab, as well as several computer science professors from the department and several other students). Approximately 10 of them were using it for more than one month. Most of the users accessed the visualization webpage at least once in each log-in session. About 200 unique papers were shared in the Comtella community, but mostly by one user. Through interviewing the users we obtained feedback about the visualization that revealed the following problems:

1. The visualization did not provide enough interactivity.
2. The graphical location of each star on the screen was random and meaningless, but users were trying to attach meaning to it.
3. The motivational effect was not very strong, though everyone found it a good idea in principle.
4. The users mostly had specialized interest in one topic, e.g. “distributed systems” and were forced to compare themselves with users with interests in different topics, e.g. “human computer interaction”, whom they didn’t necessarily know or want to compare with. It would have been better to see oneself as a “bright star” in the sub-community interested in a given topic.
5. The visualization was also misleading since it showed only the users who were online at the moment and calibrated the sizes of stars accordingly. In this way the best user at the moment, even if with relatively few contributions would appear as the Sun. However, this lead to inconsistency in time, since for example, a user who saw herself as large star one day, found that her star has shrunk when he logged in the next time, even though she had made several contributions in the meantime, simply because a bigger contributor happened to be logged at this time. This inconsistency was reported by users as unfair and actually discouraging contributions.
6. Because of visualizing only the users who were on-line at the moment, often users saw themselves as the only star on the night sky, which amplified a feeling of isolation and lack of community, instead of creating a feeling of co-presence.
7. The visualization was not self-explanatory. It needed a legend to explain the meaning of each size and colour of a star.

There was no evidence that the static visualization encouraged participation in any way. However, the users generally liked the idea and indicated in their feedback that an enhanced version of this visualization would be useful in quickly discovering what their colleagues were working on, easily finding out what the hot topics were, and raising the users’ awareness of the online community. We re-designed the visualization by taking the user feedback into consideration and the resulting, dynamic design is presented in the next section

### 3.2 Dynamic Design

The dynamic version of the visualization allows users to specify how they want to view the community by selecting different criteria, and the visualization is generated upon request with the latest data from the database. In this way, the users are provided with a simpler graphical language based only on the size and colour of the “stars”, but these two dimensions get a different meaning depending on the criterion selected by the user. The following viewing criteria were introduced: view by topic, view by number of original contributions (default), view by number of shared files (also called total contributions), view by login-frequency, and view by status which can be “bronze”, “silver” or “gold”, depending on a summative measure of the total participation of the user in the system. Every user can use the same set of selection criteria to view the community, so everyone has an equal opportunity to see each possible customized view (somewhat coherent in theory with Erickson’s guideline [6]). This allows social comparison to happen in different dimensions, depending on how the users may wish to define their “peers”. The visualization is implemented as a graph-generating application written in Java embedded in the Comtella client interface.

The metaphor used in this design is the same (night sky with stars), but it is highly stylized – instead of stars, the visualization shows a group of nodes (i.e. circles) on a black background (see Figure 2). The static version of the visualization, taking the advantage of the web browser and HTML, was able to display pictures of real stars saved as JPG images. However, in the dynamic design the system has to draw every component of the visualization, depending on the selected view. It is easier and faster to draw primitive geometry elements such as a circle and a dot

Unlike the static version, all users are shown, regardless on whether they are on- or off-line at the moment. This alleviates the problem with the inconsistent size of stars in the static version and creates a stronger feeling of “co-presence” in the system. Each user is represented by node, which is either filled or empty. A filled node represents a user who is currently online and an empty node an offline user.

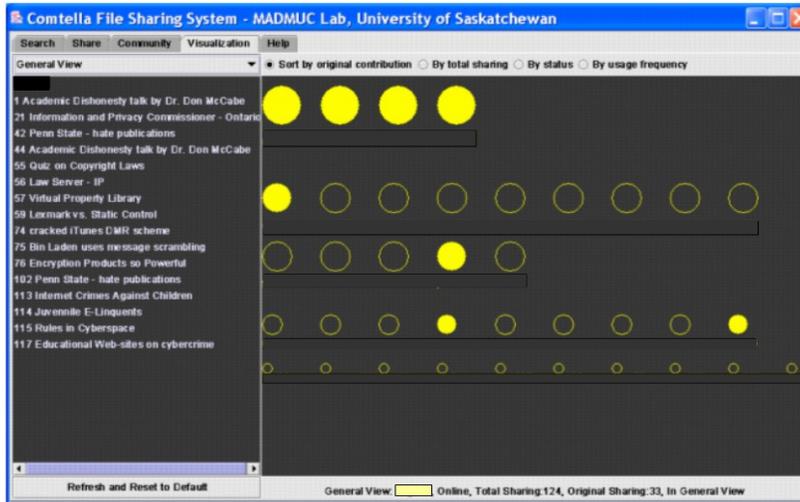


Fig. 2. The dynamic design of the visualization (usernames hidden for privacy).

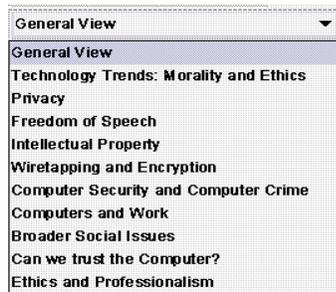


Fig. 3. Topic selection box

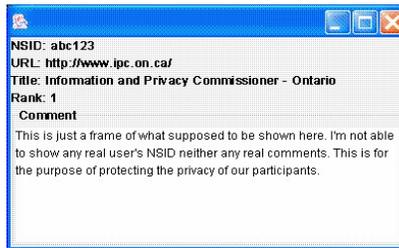


Fig. 4. A comment window

The size of a node is defined by the contribution of the represented user in a selected topic under the selected criterion. Different views can be generated depending on the topic selected by the user from the pre-defined topic list (see Figure 2, the pull down menu on the top left corner shown enlarged in Figure 3). If the user does not make any selection the default topic is the “General View”, showing the nodes with sizes reflecting and sorted by their original contributions for all the topics. The users are able to double-click on a node to see the list of papers shared by this user (the left black part of the window in Figure 2); a single click on any of these files opens up a comment-window (see Figure 4). When a user moves his/her mouse over a node in the visualization, a brief summary of the contributions made by the represented user appear in the bottom bar of the window, on the right side, below the visualization

Unlike the static version, here the nodes are not randomly located. Instead they are laid out hierarchically at four levels in the descending order of sizes. According to the guidelines [6, 8] for designing community visualizations, it is not necessary to visualize the exact difference in the numbers of articles shared by two different users. Instead, a certain amount of approximation is preferable, since it is harder for the users to discover small differences in the sizes visually. So instead of determining the size of a node by the exact number of contributions, we classified the contributions into four levels and users classified in the same level are represented by nodes of the same size. The following is the algorithm according to which the classification is done. If the total number of users is  $N$ :

$X = 10\% * N \rightarrow$  the number of nodes at Level One,

$Y = 40\% * (N - X) = 36\% * N \rightarrow$  the number of nodes at Level Two,

$Z = 50\% * (N - X - Y) = 27\% * N \rightarrow$  the number of nodes at Level Three,

$R = 27\% N \rightarrow$  the number of nodes at Level Four.

An obvious alternative to this approach is to have all four classes of equal size, i.e. to divide the total number of users into equally-sized groups. This method will not generate noticeable gaps between different levels and therefore may reduce the motivational effect. To be more specific, if there is the same number of users at the highest level as there is at other levels, then the best contributors will not be able to feel the exclusiveness of their high social status as much as they would, if they see themselves as some of the few at the exclusive top level. Similarly, the users who barely make any contributions to the community may not feel so urged to participate more if the majority of users do not contribute. There would be a stronger motivation to comply, if there are not many people at the bottom level and they are among the few. For this reason, there should not be too many people at either the top level or the bottom level. However, the top level should not appear too hard or impossible to reach, therefore its size cannot be too small; otherwise, it will be limited to only one or two users. Similarly, the size of the lowest level also can not be too small; otherwise, users at this level may feel too discouraged and may give up using the system. By not combining the two middle levels we also hope for maximizing the motivational effect; it would be harder to persuade people to improve, if they see that most of the others are just like them. Moreover, merging the second and the third levels will create a super-sized middle level which will make the top and the bottom levels appear too small and exclusive.

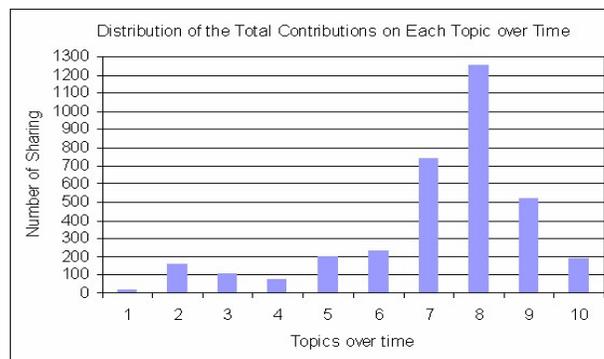
In the next section we present the evaluation of the social visualization which was carried out in a undergraduate computer science class.

#### **4 Evaluation of the Dynamic Visualization in Comtella**

Comtella was used to support students in CMPT 490 to share web links to class-related articles and do their course-work. CMPT 490 (consequently renumbered to CMPT 409) was a 4<sup>th</sup> year computer science class on social impact of information technology offered by the Computer Science Department at the University of Saskatchewan. The list of the areas of interest corresponded to the weekly topics discussed in the class, as shown in Figure 3. The experiment lasted three months, or

thirteen weeks, and covered ten different topics. Each topic was discussed for one week except the sixth topic, “computer security and computer crime” which was discussed for two weeks plus an extra week in between (the readingweek break); thus this topic ran over three weeks. The last week of the class was entirely dedicated to team-project presentations, and there was no particular topic in Comtella for this week.

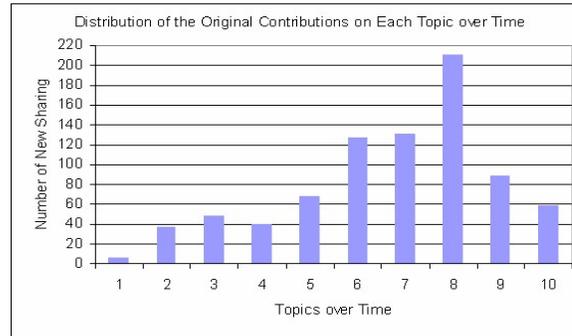
The students were first given a Comtella client with interface excluding the visualization. In the middle of the term, after the sixth topic (i.e. after the reading week break) they were given a new client with interface including the visualization. System usage data was collected from each participant and compared before and after the introduction of the visualization. The effect of the visualization was evaluated by four participation metrics: (1) the total number of shared articles for each topic, (2) the number of original (new) shared articles on each topic, (3) the number of comments given on the shared articles, and (4) the number of ratings given on the shared articles. Also a qualitative study of the user experience was conducted through an online questionnaire which the students voluntarily filled after the end of the class. The questions were trying to find out if the students felt that they belonged to a community and if the visualization had encouraged them to compare themselves with others, to compete, and to what degree it motivated them to participate.



**Figure 5.** Distribution of the Total Contributions on Each Topic

Figure 5 shows the distribution of the total contributions made by all users over the 10 topics. It shows that the number of total contributions after introducing the visualization is significantly higher than the number contributed before the visualization was introduced. In particular, the contributions for topic 7 are about 3 times higher than for each of the previous topics. Similarly, Figure 6 shows the distribution of the original contributions made by all users on each topic throughout the experiment. topic 7 represents the contribution in the first week after applying the new interface including the social visualization. Even though the number of contributions for topic 7 is not much larger than for those for Topic 6, one should keep in mind that topic 6 was “Computer Crime and Security”, which traditionally

attracts most interest in the class and is therefore discussed over two weeks (plus one week of break between them), so the students had plenty of time and intrinsic motivation to bring new contributions.



**Figure 6.** Distribution of the Original Contributions on Each Topic

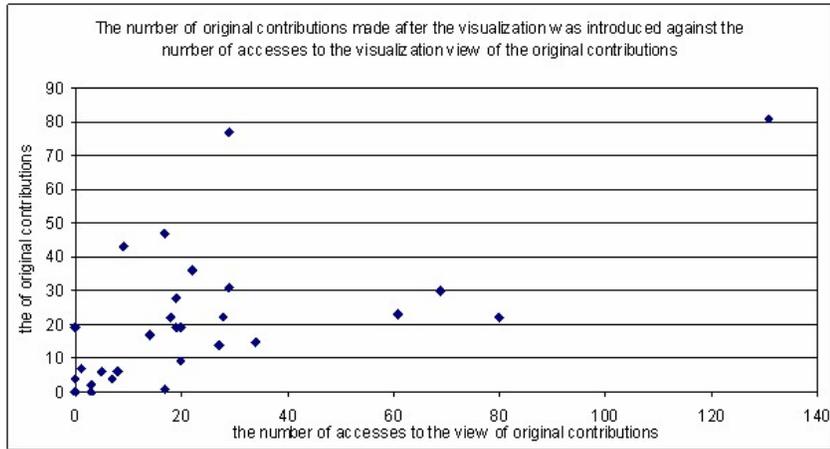
The marked increase in contributions (also in ratings and comments) in the second half of the term is obvious from Figure 7. Row marked “before 7” in this figure refers to the sum of the data from topics one to six, and “after 7” refers to the sum of the data in all topics after and including topic 7. Figure 7 shows that the visualization had a strong positive effect on the students’ sharing behaviour. Yet, towards the end of the term, the number of contributions dropped down close to the levels before introducing the visualization. We are not sure what the reason for this decline in participation was: there are many factors that play a role in a real classroom experiment, such as the inherent interests of students in certain topics, the amount of the coursework (most class projects are due by the end of the term, including the class project for CMPT 490, which limits the amount of time the students have to spend searching and contributing papers to Comtella). Also the professor’s promotion of the system in the class, may have played a role. The novelty effect is well known in the area of Human-Computer Interaction and may account for the initial interest in the students to use the system with the new interface, which died off as users got familiar with the visualization.

	Total Contribution		Original Contribution		Comments		Ratings	
	number	%	number	%	number	%	number	%
All	3526	100%	821	100%	888	100%	578	100%
before 7	803	22.77%	331	40.32%	176	19.82%	73	12.63%
topic 7	745	21.13%	131	15.96%	162	18.24%	112	19.38%
after 7	2723	77.23%	490	59.68%	712	80.18%	505	87.37%

**Figure 7.** Participation Data

Did those users that accessed the visualization more often contribute more than those who didn’t access the visualization? Figure 8 maps the original contributions made by

each user against the number of times s/he accessed the visualization (the original contributions view). Each point represents a user, where the x-coordinate of the point shows the number of times she accessed the visualization and the y-coordinate shows the number of original contributions made by this user. We can see that most points are scattered around the diagonal but there are also 5 outliers.



tests on the data shown in Figures 8 and 9. The correlation coefficient for the data in Figure 8 is 0.66, while the correlation coefficient of data in Figure 9 is 0.34. This means that the visualization had a stronger effect on encouraging original contributions.

The usage data also showed that users did not want to make the effort to do extra selection on the sorting criteria (i.e. to select different views). They usually left the sorting criteria as default (“sorted by original contribution”) and this indicates why the visualization encouraged more original contributions than anything else. None of the users ever selected on “sorted by usage frequency”, which indicates this criterion was irrelevant for the users.

For the qualitative study, all students were invited to fill in a questionnaire and received a 3% bonus towards their total mark in the class as a reward for the effort. To prevent any influence on the evaluations that the students would give, the students received the same benefit regardless if they participated in the experiment (signed the consent form) or if they used the system. The students who didn’t sign consent exited the questionnaire immediately and no data was stored in the system about them apart from the fact that they have logged in to fill it, so that they could get the 3% bonus. The class instructor had no access to the students’ data about the experiment and didn’t know who used the system and who did not, she wouldn’t know who answered the questionnaire and who did not. The answers were anonymized immediately, as well as all the participation data from the database, to mitigate eventual influence on the answers. The questions regarding the visualization together with the results (what percentage of the 35 students chose each answer) are listed below:

1. What would your reaction be if you saw yourself as one of the smallest nodes in the visualization?
  - a. Take immediate action: share more links to make your node larger. (35% chose this option)
  - b. Think of sharing more links, but later. (19%)
  - c. Feel unhappy, but do nothing. (10%)
  - d. Feel that the system is unfair, so it doesn't make sense to contribute. (10%)
  - e. Do not care, so will do nothing. (16%)
  - f. Other, please specify: (10%)

In the optional comments given for this question, a couple of students expressed concern about the quality of the shared articles as they noticed that not all the users with bigger nodes were sharing useful resources.

2. If you saw yourself as one of the largest nodes in general, what would you do?
  - a. Feel proud of your status and try to contribute even more. (42% chose this option)
  - b. Feel proud, but at the same time, in some sense 'exploited'. The others are not bringing in so much, so I will stop or decrease my contributions. (6%)
  - c. Feel worried, you may be raising the bar too high and others may hate you or you may be perceived as an 'overachiever' by the others. (6%)
  - d. Feel nothing, since it is not important to me. (19%)
  - e. Other, please specify: (26%)

In the specifications given, some students said that if they were already one of the biggest stars, they would only share more links if they had time and in this case they would focus on the quality of their contributions. One student stated that if she saw him/her self as one of the biggest stars she would feel it was too easy to reach a high contribution level and the evaluation system needed to be improved.

3. If their answers to the above two questions are affected by whether the topics of interest to them?

1. a. Yes (55%)
2. b. No (32%)
3. c. I don't know. (13%)

This suggests that the users' competitive attitude doesn't depend much on whether a topic is of interest or not. It is important to notice that users replied in the context of a class where topics are changed based on a weekly schedule and are pre-selected. Perhaps the answers would have been different in a interest-based community e.g. a research group (as the system in with the static visualization described in Section 3.1) where users have long term personal interests.

4. What would you like to know about other users?

	<b>-2</b>	<b>-1</b>	<b>0</b>	<b>1</b>	<b>2</b>
who is online	13%	16%	26%	23%	22%
how much others contributed	3%	6%	13%	35%	43%
who downloaded from me	3.23%	9.68%	12.90%	35.48%	38.71%
who I downloaded from	3.23%	19.35%	19.35%	38.71%	19.35%
who gave similar ratings as me	3.23%	9.68%	22.58%	41.94%	22.58%
am I a freeloader or an contributor?	16.13%	16.13%	12.90%	35.48%	19.35%

To sum up, about 77% of the subjects wanted to know how much others contributed and 55% were interested in knowing if others see them as freeloaders or as active contributors. This indicated there is great potential for motivating subjects to contribute more through this prototype visualization

## 5 Discussion

The experimental results and the user feedback discussed previously showed that the motivational visualization effectively increased the participants' awareness of their Comtella community and encouraged social comparison as a result both the total contributions and the original contributions went up significantly, and participants gave more comments and ratings. The effect is more obvious on the original contribution than it is on the total contribution as the t-test indicates.

It should be acknowledged that it is hard to control the “noise” in real world experiment during a deployment of a system in class, such as the users’ inherent interests in different topics and the currency and relevancy of the topics in the real world. It can be argued that the increase in participations may be caused by the subjects being more interested in topics discussed in the latter half of the term when the visualization was introduced. It can also be argued that the topics discussed latter in the term were “hotter” and it was easier to find relevant articles on the web. However, according to the class instructor the opposite is true: the most interesting topics for students and richest of materials on the web are “Privacy”, “Freedom of Speech”, “Intellectual Property”, “Wiretapping and Encryption” and “Computer Security and Computer Crime”, which were discussed in the earlier half of the class. The topics discussed latter after the new interface was introduced, “Computers and Work”, “Broader Social Issues” are too broad and often fail to attract student’s interest. For the last two topics “Can we Trust the Computer”, and “Ethics and Professionalism”, it is harder to find online materials, since they are of more professional interest than general.

One clear conclusion was that user-customizable views are hardly needed, since users tend not to use them; most users only used the default view. This seems to support the first guideline in design visualization by Erickson (2003). Our main motivation for providing the different views was to simplify the complex graphical language (expressed in the legend) necessary to decode the meaning of the visualization, based on our experience from the first version of the visualization. In retrospect, we could have not visualized unnecessary information and still provided a single view.

One general observation that was made in the experiment was that as the quantity of contributions increased, their quality somewhat deteriorated. This may be because 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. Several users found ways to game the system and exaggerate their nodes in order to gain higher status and visibility. Motivating social comparison in the quality of the contributions, comments, and ratings is an important future direction of research. There exists a lot of research on reputation mechanisms that can be helpful in this endeavor.

Motivating active users to continue their contributions or even increase them is another problem. In the questionnaire, one student indicated that if one was already the best contributor and was visualized as the largest node, there was no motivation for him/her to continue contributing. Another one indicated that she would then try to submit higher quality contributions. Probably the visualization has to take into account both the quality and quantity of user contributions and change dynamically the view to stimulate social comparison depending on what is needed most by the community – just any papers (if there is still very small number of shared resources) or high quality papers (if there is an abundance of resources).

From a technical point of view, a major issue in the design of the visualization is the clustering 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.

One should not forget that the ultimate purpose of the Comtella system (apart from being an experimental tool for testing motivational approaches for user participation) was to facilitate students in finding and reading fresh materials related to the topics of their studies or research. One student commented that she didn't like the competition for bringing resources, since people were caring more about finding resources than about reading them, and *this* should be the main goal in the system. It seems that reading articles found by others is also an important way of participation, even if it is "invisible" from the viewpoint of the community, since it does not contribute to enlarging the pool of shared resources. It is important to remember that the resources are only valuable for the community, if they are *read* by the community members. Otherwise, participation becomes a game with no higher, in this case educational, purpose. Therefore, we should try to find ways of encouraging people to participate even as a "silent audience", if not as active contributors.

**Acknowledgement:** This work is supported by the second author's NSERC Discovery Grant.

## References

- [1] S. E. Asch. (1951) Effects of Group Pressure upon the Modification and Distortion of Judgments. In *Groups, Leadership, and Men*, pages 177-190.
- [2] C. Alexander, S. Ishikawa, M. Silverstein, with M. Jacobson, I. F.-King, and S. Angel. (1977). *A Pattern Language: Towns, Buildings, Constructions*. New York: Oxford University Press.
- [3] H. Bretzke and J. Vassileva. (2003). Motivating Cooperation on Peer to Peer Networks. *Proceedings User Modeling UM '03, LNAI 2702*, pp.218-227.
- [4] D. Cosley, S. K. Lam, I. Albert, J. A. Konstan, J. Riedl. *Is Seeing Believing? How Recommender Interfaces Affect Users Opinions*. CHI 2003, April 5 – 10, 2003, Ft. Lauderdale, Florida, USA.
- [5] S. G. Eick and G. J. Wills. (1993). Navigating Large Networks with Hierarchies. *IEEE Visualization '93 Conference*, pages 204-210.
- [6] T. Erickson. *Designing Visualizations of Social Activity: Six Claims*. ACM CHI April 5 – 10, 2003 Ft. Lauderdale Florida, USA.
- [7] T. Erickson, W. Huang, C. Danis, and W. A. Kellogg. *A Social Proxy for Distributed Tasks: Design and Evaluation of a Working Prototype*. ACM CHI 2004, April 24-29, 2004, Vienna, Austria.
- [8] T. Erickson and W. A. Kellogg. *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.

- [9] L. Festinger. A Theory of Social Comparison Processes. (1954). Human Relations, 7, 117-140.
- [10] Free Peers, Inc. (<http://www.freepeers.com/>), BearShare (<http://www.bearshare.com/>).
- [11] M. C. Humphrey. Creating Reusable Visualizations with the Relational Visualization Notation. 11th IEEE Visualization 2000 Conference (VIS 2000). October 08 – 13, 2000. Salt Lake City, UT.
- [12] Victor A. F. Lamme. Why Visual Attention and Awareness are Different. TRENDS in Cognitive Sciences Vol. 7 No. 1 January 2003.
- [13] LimeWire. *The Gnutella Protocol Specification v0.4 Document Revision 1.2.*: [http://www9.limewire.com/developer/gnutella\\_protocol\\_0.4.pdf](http://www9.limewire.com/developer/gnutella_protocol_0.4.pdf)
- [14] MetaMachine. *eDonkey2000 and Overnet* (<http://www.edonkey2000.com/>).
- [15] T. Pattison, R. Vernik, and M. Phillips. Information Visualization Using Composable Layouts and Visual Sets. Conferences in Research and Practice in Information Technology, Vol. 9. Commonwealth of Australia, 2001.
- [16] R. Savage <http://www.visitorville.com/meet-the-mayor.html>, VisitorVille <http://www.visitorville.com>
- [17] Sharman Networks Ltd. KaZaA (<http://www.kazaa.com/>) Media Desktop, 2001.
- [18] T. C. Sprenger, R. Brunella, and M. H. Gross. H-BLOB: A Hierarchical Visual Clustering Method Using Implicit Surfaces. 11th IEEE Visualization 2000 Conference (VIS 2000) October 08– 13, 2000. Salt Lake City, UT.
- [19] J. Uther and J. Kay. (2003). VIUM, a Web-Based Visualization of Large User Models. Proceedings User Modeling, UM'03, LNAI 2702, pp. 198-202.
- [20] J. Vassileva. Motivating Participation in Peer to Peer Communities. Proceedings of the Workshop on ESAW. ESAW'02, Madrid, 16-17 September, 2002. <http://www.ai.univie.ac.at/%7Epaolo/conf/esaw02/preproc/E0029.pdf>.
- [21] F. B. Viegas and J. S. Donath. Chat Circles. Proceedings of the SIG CHI conference on Human factors in computing systems: the CHI is the limit, 1999.