

# Motivational Visualization in Peer-to-Peer Systems

Lingling Sun

MADMUC Lab, Department of Computer Science, University of Saskatchewan  
lis102@mail.usask.ca

## ABSTRACT

As peer-to-peer (p2p) applications become popular in online file sharing systems, their drawbacks such as peer free riding, inefficiency of searching, management difficulty, have been brought into attention. Many approaches in solving these problems have been proposed and most of them focused on the improvement of protocols or creating network overlays. However, we believe a sufficient amount of peer participation is essential in a p2p system because without it a p2p system would mostly stay idle and not very useful. So I propose a new approach which is to apply motivational visualization in the interfaces of p2p applications with the goal of encouraging social comparison and inspiring active peer participation in p2p networks.

**Keywords:** community, p2p system, motivational visualization, peer participation, social comparison.

## 1. INTRODUCTION

Since Napster was introduced in 1999, peer-to-peer (p2p) systems have become popular and have been largely used for online file sharing. Compared with traditional client-server networks, a p2p network is highly decentralized, unstructured, and has the advantages of low cost, self-scaling, anonymity, replication and diversity of resources [14]. Users in a p2p network enjoy great freedom of when to come and leave and their privacy is well protected due to the anonymity. But the lack of structure and the variable topologies of p2p systems make the message routing inefficient and the queries suffer from high failure rate.

My previous research discovered three key problems in p2p systems regarding peer participation: free riding, unpredictable routing, and searching delay. Free riding [1, 2, 14] refers to the fact that most of the peers only take resources and services from the network without making contribution to the network. The total contribution compared to the total consumption is negligible. Unpredictable routing is caused by the variable network topology. In a p2p system, there is constantly high frequency of peers joining and leaving the network since they have the freedom to make these decisions at any time. So it is unpredictable what resources and services will be available at what time. When a peer sends out a query, he/she has no idea where the query will be routed since it totally depends on the network topology at that particular point of time and this topology very likely will change soon. This often causes searching delays or lack of results. When a query for a certain type of resource or service is sent out, it is unclear how far the query has to travel in order to find a response. The major cause of the search delay is the message flooding mechanism in Gnutella-like applications [3]. In Gnutella systems, a sender sends its queries to its neighbors who will in turn propagate

the queries to their neighbors. Whether or not the receivers are able to respond to the queries is unknown and not even considered by the sender and the propagators when they send or pass the query. In a network with high frequency of peers joining and leaving, this mechanism is obviously inefficient and will result in a large number of failed queries and unnecessary network traffic.

Many approaches in optimizing p2p systems have been proposed with various focuses such as improving the protocol [11], optimizing the structure [3, 10, 16, 19, 23], or creating network overlays [18, 20]. The evaluations of these approaches showed great increase in the searching speed and the successfulness of queries [18, 20, 23]. However, these evaluations are done in fully or nearly fully utilized networks. Without a critical mass of peer participation, p2p systems will not be fully utilized and so all these improvements will not be able to take effect.

More recently, the lack of cooperative peer participation has been brought into focus [1, 2, 22] but there is not much work being proposed to inspire peer participation. We are currently developing a Gnutella-like file-sharing system, COMTELLA, in our lab. Comtella stands for "Community Gnutella". It motivates active participation from peers through encouraging social comparison and inspiring competition and the feeling of the community.

The rest of the paper is organized in this way: section 2 describes the conceptual design of the Comtella system and presents the problems that we found from the user feedback in the first evaluation. Section 3 presents the related work in social science on motivating user participation and their theories which support my proposed motivational visualization approach. Section 4 is on the actual design of the motivational visualization for Comtella. Section 5 evaluates the visualization design. Section 6 concludes the paper and summarizes the future work for the motivational visualization.

## 2. PREVIOUS WORK – COMTELLA

Comtella is a Gnutella-like application and provides all the basic Gnutella functionalities such as launching searches, displaying searching results, initiating downloads, viewing uploads, and sharing and un-sharing files in the network [11, 13]. The additional features in Comtella are the user modeling, which helps optimize searches and file transfer, and motivational visualization, which will be described in details in section 3. The domain of Comtella is to facilitate sharing academic papers within research a research group or an academic community. Participants in a Comtella network are free to come and leave and it is totally up to them to decide if they share any files or not and what files to share. However, to encourage the sharing of a large number of

useful files from every user, Comtella applies user modeling and rewards the participation with better service.

In a Comtella network each peer creates its neighborhood according to its typical search interests. The peers in the same neighborhood share interest in the same area of research. In order to know each other's interests, a peer needs to learn from experience about other peers' interests so that it gradually builds up **relationships** with others in the interest area. A peer with many interests will have various neighborhoods and will have relationships with others in multiple interest areas, which we call "categories" in Comtella. Relationships are not weighted equally because the strength of each relationship is different. The strength is computed based on the frequency of interactions. An Interaction is a response to a query, a successful download, or a successful upload. The strength is a weighted sum of these three types of interactions. Successful upload has the largest weight. Successful download has a smaller weight but still larger than a response to a query. Each Comtella server constantly relocates its total bandwidth by offering the larger bandwidth to strongly related peers.

The strength of relationships guides the query routing. A Comtella server only sends its queries to peers strongly related in the area of the queries unless there are not enough such peers to ensure a good chance of getting responses to this query. In this case the server will send the query to other related peers. When a server receives a query that it needs to forward, it uses the same method to determine to whom query should be passed. For queries to be successful and efficient, it is important for peers to maintain strong relationships with each other. The assumption here is that peers who have shown the same interest are more likely to share files in that interest area. Under this assumption, queries sent by peers who are strongly related to others in the area will yield responses more successfully and efficiently and the network traffic will also be cut down.

More defects about the user-modeling and the relationship computation are available in [22]. The main problems with the initial version of Comtella is that: 1) the effect of rewarding larger bandwidth to stronger related peers is not noticeable for the users because the network is small and local and the downloading speed is always acceptable; 2) there is no way for peers to see their levels of contribution compared with others.

### 3. RELATED WORK ON USER PERSUASION

We designed a motivational interface in Comtella system to visualize the Comtella communities in different interest areas so that peers can compare their contributions with those of other peers. According to some theories in social psychology, such as the balance theory and the social comparison theory, this can be used to motivate participation and active contribution to the community.

#### 3.1 Theories for Persuasion in Social Psychology

The **Balance Theory** [15] explains how relationships are built in real communities and there are indications that this theory works in virtual communities as well. The theory states that when two people agree in their attitudes to the same issue or social agent, they will have positive feelings to each other and if they disagree in their attitudes, they will have negative feelings to each other (figure 1). If the attitude of one person to the issue changes, this will likely lead either to a change in the feelings between the two people, or to a change in the attitude of the other person to the issue, i.e. the first person will persuade the second person.

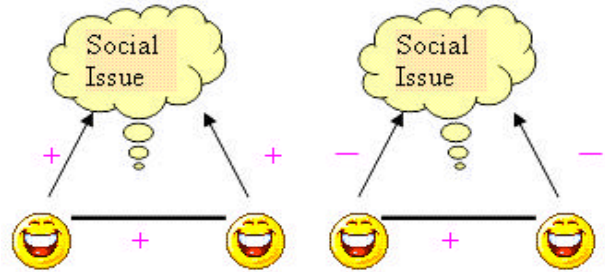


Figure 1 (a)

Figure 1 (b)

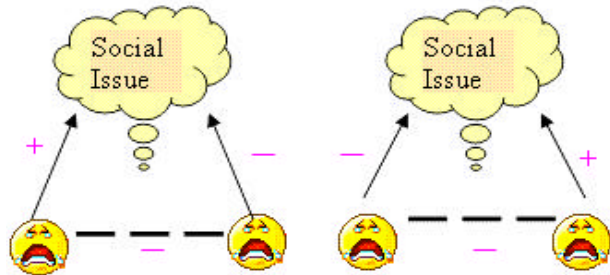


Figure 1 (c)

Figure 1 (d)

The "Social Issue" in figure 1 can be instantiated to anything in a social community such as a person, a topic, an object, a problem, or an event etc. An arrow with a "+" sign represents a positive attitude, an arrow with a "-" sign represents a negative attitude, a line with a "+" sign stands for a positive relationship between two persons, and a dashed line with a "-" sign stands for a negative relationship between two persons in which case the relationship is very likely to be broken. For a triangular structure (figure 1) to be stable, the dot product of the three signs in this triangle has to be positive, which means there has to be either three positive signs (figure 1-a) or two negative signs and one positive sign (figure 1-b,c,d) and could never be one negative sign with two positive signs nor three negative signs. The problem is when the relationship between two persons is negative, there is high possibility for that relationship to break, which will destroy the triangular structure. Thus, when two persons engaged in a positive relationship have different attitudes towards the same important issue, either one of them has to persuade the other to change his/her attitude so that the relationship between these two persons is kept otherwise it is likely that the relationship will break. The balance theory is a good guide in building group collaboration [9] and it can guide us in building user models and grouping users into various communities with different interests.

**Social Comparison.** People commonly have the wish of “fit in with group” [5]. When giving opinion on a subject, one normally tends to compare his/her opinion with the opinions from the peer group and s/he is often influenced by the group. An example given in [5] proves this. There were four groups, A, B, C, and D, of people asked to rate movies. Each group was seated in a separate room so that they would not influence each other. Before they started, group A, B, and C were given different predicted ratings and group D was not. The results showed that a significant proportion of the participants from the group A, B, and C were influenced by the given predictions in their ratings. Another example of the same type of behavior is that a group of people were asked to compare the length of three lines and everyone had to do the task 12 times [5]. They were allowed to discuss with each other but everyone had to come up with his/her own decision. A small portion of the group was the experimenters who deliberately gave wrong opinions and the rest of the group did not know this. The result was that 33% of all the answers were wrong and most subjects gave at least one wrong choice. Both examples explain that social comparison is quite influential in social communities.

**Social recognition** typically goes together with social comparison because the highly recognized members in a social group are usually the targets for the rest of the group to follow and compare with. Generally speaking, people are willing to be positively recognized in their social communities and will work hard to gain this recognition. Once being recognized, people want to maintain their status. Several real-life applications exploiting social comparison and social recognition are given in section 3.2. Another form of social recognition is the appreciation from the community. When a person spends efforts and resources for the benefit of his/her community, if s/he gains great appreciation from the community s/he would be encouraged to do so [9].

### 3.2 Applying Theories from Social Psychology

**Internet Chess Club:** The Internet Chess Club [7] is a successful example of a world-wide system attracting a large number of volunteers inspired by social recognition and social comparison. The Club groups its users into five categories according to their skills. People with good software skills are software contributors and help with designing, improving, and maintaining the system; people good at chess playing join the games and the chess tournaments, compete to become “Titled Chess Players”, and as they play they discover problems with the system and give suggestions and ideas; people interested in socializing work as helpers by answering questions online, or as Administrators who trains players, organizes helpers, advertise forthcoming events and so on; people who are talented in managing organize chess tournaments and resolve conflicts [7]. It has been seen that volunteers generally work hard to maintain or gain reputation in the Club. For example, chess players practice their skills to become “Titled Players” in the tournaments, software contributors compete in inventing new features for the system; helpers and administrators put a lot of effort in answering questions, socializing new comers, and training players etc, so that they are highly recognized in the Club as specialists or experts and may get chances of being promoted to managers. In sum, the social

comparison motivates every volunteer to actively contribute to the Internet Chess Club and the total social capital of the Club constantly increases.

**Slashdot.org** is a widely used online news group and discussion forum. People post information, their own questions as well as answers to others’ questions on Slashdot.org. By providing popular and valuable information or instructive solutions, one can gain recognition from the virtual community of Slashdot.org in terms of votes from others that are accumulated in “Karma”. Depending on the “Karmatic level”, the users can attain moderator-rights for some period of time, which “increases their visibility and power in shaping the forum.

**Orkut** is an online “trusted community”, where one joins only on invitation by a friend and can invite his/her own friends, in which s/he can bring up various comments and issues for discussion. As people discuss and give comments and suggestions, they socialize with each other and find friends with similar interests.

## 4. COMTELLA VISUALIZATION DESIGN

We decided to apply some of these ideas and designed the motivational visualization for the Comtella system to encourage social comparison which will in turn inspire active peer participation.

### 4.1 Techniques

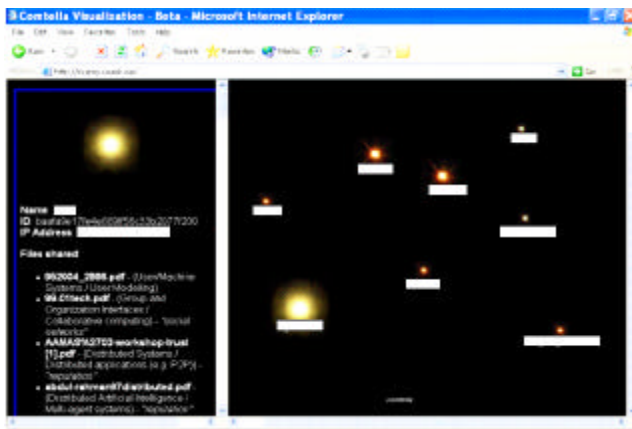
**Hierarchical structure:** A P2P networks can contain many nodes. Using hierarchical structure will facilitate the drawing of this type of graphs and generate organized layout [6, 17, 21]. But there are two difficulties in creating reasonable hierarchies, 1) finding a suitable clustering algorithm, and 2) choosing a proper scale. The clustering algorithm should not only focus on “optimizing the computation and spatial grouping” [21]. It also should make each cluster meaningful and self explanatory. A proper scale should help generate enough levels of hierarchies to map to the levels of significance, hence, should neither be too many levels to confuse users nor too few to blur the difference between each level. I picked up a scale of four to make the visualization cluster nodes into four levels of hierarchies. Each hierarchy represents a certain level of cooperativeness (details in section 3.3.2).

**Visual Attention:** people do not pay attention to everything that they see [12]. Biological studies show that the advantages of graphical visualization over plain text are that 1) the human eyes are more sensitive to colorful pictures than black-and-white texts and 2) the human brains are better at memorizing patterns than words. Effective means of emphasizing something to stimulate visual attention can be one or a combination of the following: highlighting with a particular color, locating at a special place (e.g. center of the picture), using different drawing patterns, etc. These means are aimed at creating contrast to the surroundings. In Comtella visualization, a ME-node (the node of this user) is painted in a special color and the most cooperative peers are placed at the very top of the graph.

**Reusability:** since “richly expressive information visualization is difficult to design and rarely found” [6] it is beneficial that once a useful visualization is created, it can be re-applied in similar situations. The visualization of Comtella separates the data storage from the graph generating so it can be reused to visualize any p2p networks of the similar type and size with proper sets of input data.

## 4.2 Comtella Visualization – Web Version

The first version of the Comtella visualization was designed as a webpage generated on a server separated from the Comtella servent. The reason is that the visualization’s effects on both users and the Comtella system were unclear so if something went wrong with it, the Comtella servent and its main functionality should not be affected. We were not sure whether generating the visualization on the servent would slow down the system too much, and also since the community view was supposed to be the same for all peers, a centralized visualization, generated on a server and viewed as web-page by the servents was a logical choice. The Comtella servent runs independently from the visualization expect it reports necessary data to the database when needed. The database used is the free software MySQL and the visualization generator is written in PHP taking its advantage of creating graphical web pages. To check the Comtella community, a user can either visit the website without running a Comtella servent if s/he knows the web address, or s/he can start a servent and which can help his/her open a browser window with the pre-coded web address.



**Figure 2. Motivational Visualization - web version**

In figure 2 the Comtella community appears as a mid-summer star sky [2]. The size of a star is proportional to the numerical value of its cooperativeness with bigger stars standing for larger cooperativeness values. The sun represents the peer who is most cooperative. The red stars are peers who give more files to others than take files from others, and those who take more than give are visualized by yellow stars. The cooperativeness of a peer is calculated as a weighted sum of four factors: 1) the number of files shared by this peer that were downloaded by others, 2) the number of files that this peer shares, 3) how long this peer stays

online, and 4) the number of files this peer takes from others. Among these factors 1) has the largest unit weight since it indicates the amount of useful contribution a peer makes to the community. If it is not requested by others, a peer’s contribution is not useful to the community. The factor 3) is not active contribution, but as far as a peer stays online, s/he is helping the community by at least forwarding messages. Therefore, a small weight is assigned to this type. Factor 4) is inactive contribution and is assigned a minimum unit weight. It contributes only because the Comtella servents are designed by default to share the files downloaded from others. Thus as a peer downloads more s/he shares more. This creates many replicates of popular files and even if the file’s original contributor is offline, the replicates will continue serve the community. The cooperativeness of a peer is the weighted sum of the factors expressing each type of contribution. There is a small link, “Comtella”, at the very bottom of the right side in figure 2, clicking on which will display the legend of the visualization in the left side. By clicking on a star or the sun in the picture, details will show up in the left side including its alias, its IP address, lists of its shared files, interested areas, and its relationships with its directly connected peers. The locations of stars are randomly generated.

We were not able to find sufficient number of users to do a thorough evaluation of this version of visualization. However we asked some of the graduate students, researchers, and professors in our department to try the system and to give feedback and we received many helpful suggestions. The major problems found by users are:

1. The visualization is not very interactive. A user can click on a star to list the shared files and interests of this peer. But users tend to click on the listed files hoping that they could download or open to read the files. The visualization does not support this function.
2. The graphical location of each star is meaningless. Users are apt to find correlations from the locations of stars but end up with nothing and so being disappointed.
3. The graph is static. Everyone sees the same picture. There is no distinction regarding the different types of contribution (the four factors introduced above) and the different interest areas. Peers who are very interested in one narrow area and are very cooperative in that area may have a smaller star than peers who are not highly cooperative anywhere in particular but contribute something everywhere because the cooperativeness is calculated for the overall contribution in all areas.
4. Only showing the peers who are online does not truly reflect the state of the community. A peer may once see him/her self being the brightest star simply because most of the cooperative peers were offline at the moment. The next time when they were online this peer would suddenly see his/her star shrink down to a tiny one. This is discouraging and makes users lose their interest in social comparison.
5. The visualization shows some unnecessary information, which does not help with enhancing the usability of the system. For example, the IP is private information of the user and should not be shown. The numerical values calculated for the strength of relationships and the cooperativeness are confusing and irrelevant to the users and should not be shown.

6. The visualization should be fairly self-explanatory rather than explained by a legend since users are generally not willing to make the extra effort to read more than they want.

Regarding the problems found in user feedback, I developed the second version of visualization for Comtella.

### 4.3 Comtella Visualization – Current Trial

Instead of generating a webpage on a fixed server, in the new visualization the graph is distributed generated by each Comtella server using data from a central database on a visualization server. But it is still separated from the Comtella server as it is in the previous version and reports data to the database through a separate protocol so that it will not affect the Comtella community. Each Comtella server draws its own picture using Java 2D on its local machine. A peer is represented by a yellow circle on a black background. Even though we prefer the idea of stars to circles, the actual shape of a star is much harder and more time consuming to draw than a circle. The graph shows all the known peers, online or offline. A node is filled if the peer it represents is online, otherwise it is empty. Pink is a preserved color for “Me-node” (the node representing a user him/her self in the Comtella community) in order to make it distinguishable from others. The picture is more interactive than a static webpage. There are several ways for user to interact with the system:

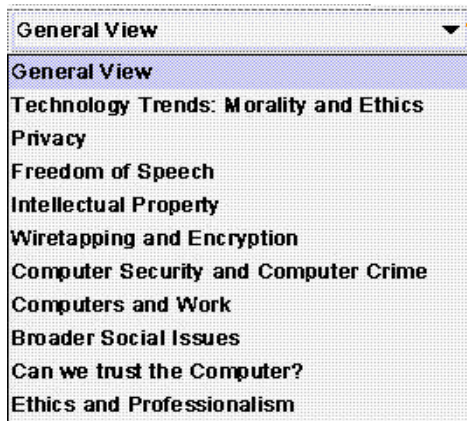


Figure 3. View selection box containing interest areas.

1. A user is able to select an interest area to make the system only visualize the community for this area. Both users and the Comtella system benefit from this feature. Generating graphs on local servers cuts down the workload on the database server which has to handle a lot of database-access requests. Users with different interest are able to see different graphs at any time they want without worrying about affecting others. The community shown in a particular interest area shows the contribution in that area from all the known peers, online or offline. Users are also able to view the overall contribution by selecting the “General View” (figure 3). This feature gives more choices in viewing communities as well as the ability for a user to maintain his/her status in a particular interest area. If a user has a strong interest in one topic and is not interested in others, s/he could still be able to keep his/her superior

status in that area as long as s/he keeps contributing useful files to the community of his/her interest.

2. Users are able to sort the nodes by four criteria, instead of a single cooperativeness value in the old version. The new version of the visualization offers four independent factors to compare the cooperativeness of a peer so users are able to see who is more cooperative in which respect. The factors are the total number of files shared by each peer including both new files and files downloaded from others, the total number of new files shared by each peer, and the usage frequency. The third criterion, usage frequency is the number of times that a peer uses the Comtella system. Figure 4 shows the selectable criteria in the visualization.

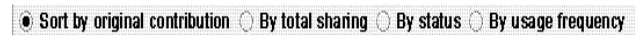


Figure 4. Selectable criteria in the interface

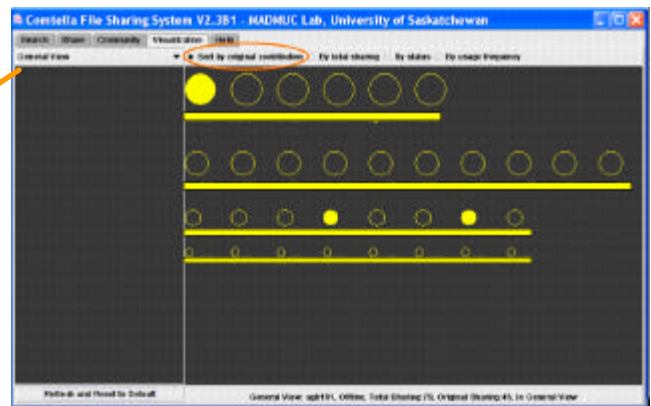


Figure 5. Sort “General View” by the original contribution.

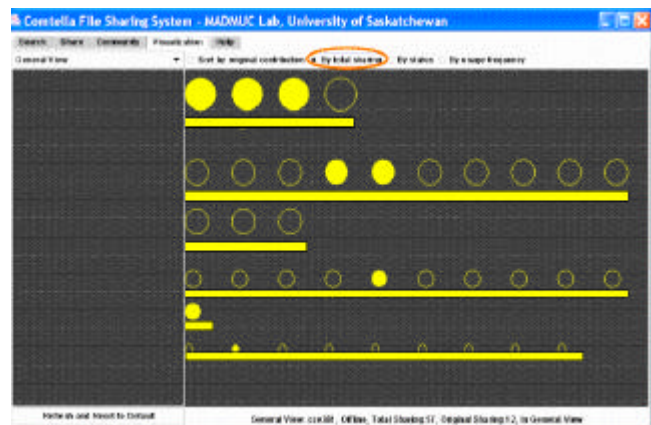


Figure 6. Sort “General View” by the “Total Sharing”.

The criteria selection works in correlation with the view selection. For example, if a user selects the “General View” from the view selection box, the visualization system will generate the overall community showing all peers including the viewer and by default, sorted by the original contribution (the total number of new files) from each peer (figure 5). If this user selects “By total sharing”

without changing the view selection, the system will sort the same nodes in figure 5 by the total number of files shared by each peer and generate a new view as in figure 6.

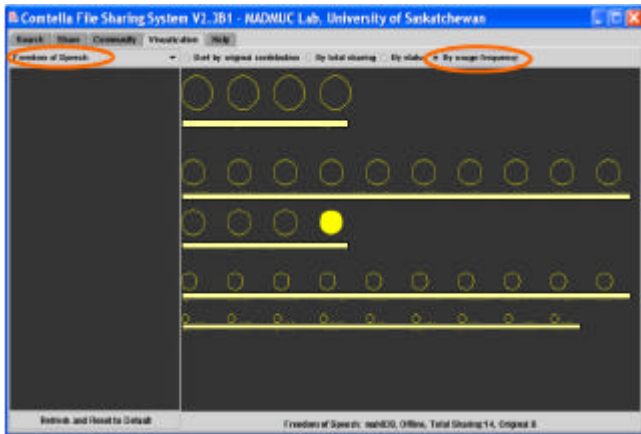


Figure 7. Sort contributes in the topic, “Freedom of Speech” by the “Usage Frequency”.

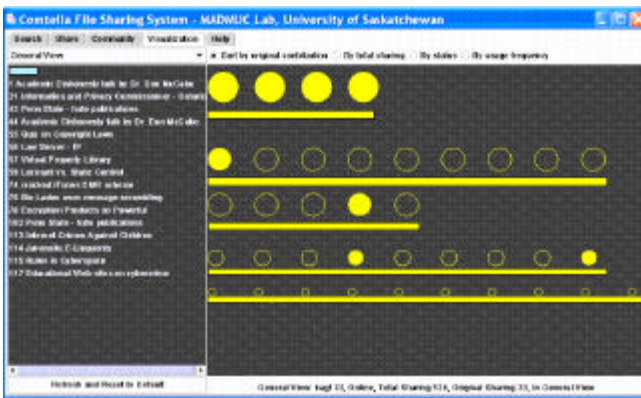


Figure 8. Left side lists the commented files of a node being double clicked.

3. Users are able to click and double-click on any node to get details of that node. At the bottoms of figure 7, there are two bars. On the left is a button to refresh the graph by reconnecting to the database and getting the latest information. On the right is a text bar. When a user moves his/her mouse over a node in the graph, this text bar will display several details of this node as well as the interest area of the current graph. The details include the node’s alias, online or offline status, total number of files shared and the total number of original files contributed by peers in this interest area. A double-click on a node will list on the left side (figure 8) the commented files by this node and a single click on any of the listed files will open up a small window containing the comments given by this node (figure 9).

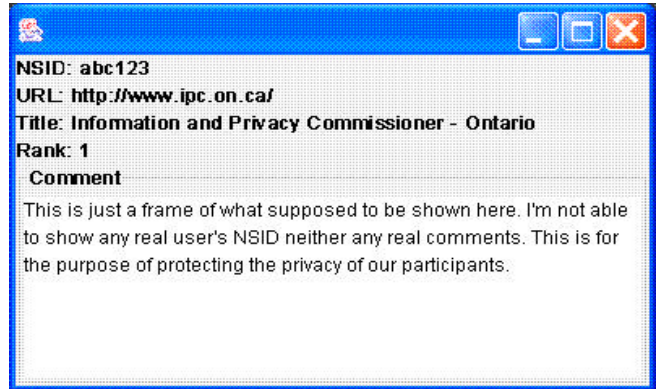


Figure 9. An example of a comment.

It is important to note that in figure 5, 6, and 8 the nodes are not randomly located. Instead they are laid out hierarchically at four levels in the descending order of sizes. Take figure 8 as an example. There is a largest possible size *largest\_size* that a node can have, and there is also a decrement factor *decr\_f*. After sorting the nodes in the descending order according to a user selected criterion, for example “the number of original contributions”, the visualization assigns *largest\_size* to each of the top 10% nodes in the sorted list and draws them at the top-most level in the interface. Among the remaining nodes the visualization takes the top 40% and assigns each of them a size of *2\_size*, where  $2\_size = largest\_size * decr\_f$ , and draws these nodes at the second level in the interface but not necessarily to be the second line from the top since sometimes there too many nodes at the same level to be drawn at the same line. In figure 8, nodes at line 2 and line 3 are all at the second level. The top 50% of the remaining nodes will be drawn at the third level each with a size of  $3\_size = 2\_size * decr\_f$ , and the rest will be drawn at the fourth level with a size of  $3\_size * decr\_f$ . The filled nodes in the figures are peers that happened to be online when I took these screen shoots and the empty nodes are happened to be offline. The trend in these figures is that the cooperativeness of nodes in each criterion decreases from top to bottom vertically and from left to right horizontally.

There are other ways of classifying participants into various hierarchical levels. An obvious alternative is to classify participants into equal-sized groups. Apparently this method will not generate noticeable gaps between the numbers of contributors at different levels therefore may reduce the motivational effect. To be more specific, if there was the same number of peers at the highest level as there was at the lowest level then the best contributors would not be able to feel their rareness and their noticeable positions as much as they would do when they saw themselves being one of the few best at the top-most level, and similarly, the least cooperative peers would not feel so urgent as they would do if they saw themselves being one of the few worst at the bottom-most level. Thus there should not be too many people in either the top-most level or the bottom-most level. However, the top-most level should not appear too high to reach and so its size can not be too small. The size of the lowest level also can not be too small. Otherwise peers at this level may feel too discouraged and shamed and may give up using the system. I personally believe that a top

10% is a reasonable size for the highest level and a similar or slightly larger size is suitable for the lowest level. The vast majority are at the second and the third level. The reason of not putting these two levels together is also for the goal of maximizing the motivational effect. Merging these two levels will create a big-sized middle level which will make the top and the bottom levels appear too small and so create the problems that I just mentioned. Besides, it is always harder to persuade people to contribute more if they see most of the others are just contributing as much as they do (i.e. at the same level as they are).

## 5. EVALUATION

Comtella with the new version of the motivational visualization is currently being used and evaluated by the fourth year computer science students who are taking CMPT 490 (a class about Ethics in Computer Science) in our department. We want to solely experiment on the effect of the motivational visualization and find out if it encourages any social comparison and to what degree it motivates peer participation and contribution. This section describes the details of the experiment setup and a rough analysis of the preliminary results.

### 5.1 Experiment Setup

The experiment started on January 11, 2004 and will finish at the end of April, 2004. At this point of time it has been through ten weeks on eight different topics. In the CMPT 490 class, each week has a specific topic associated with it except week 6 and week 7. Weeks 6 and 7 share the same topic, "computer security and computer crime". The students in this class use the Comtella system throughout the term as a tool helping to find and share up-to-date web-links related to the class material. There are 35 students using the system and 22 of them so far (more will come up later) have consented to give their data for our research purpose, so the evaluation in this paper is made based on the data from these 22 participants.

In order to solely test the effect of the motivational visualization on user participation, we simplified the Comtella system by deleting its other features such as user models and relationships; and in order to make the system better fits the need of the class where all the participants are from, we modified the Comtella server to make it able to share URLs (i.e. bookmarks) instead of actual files. We want to find out if the visualization will encourage any social comparison and to what degree it motivates peer participation and contribution. Students were initially given the Comtella servers without the visualization. The Comtella servers with the visualization were distributed at the beginning of week 8. This section roughly analyzes the current data and a more detailed evaluation will come up after we finish the experiment and obtain more data from users through questionnaire at the end of April.

### 5.2 Evaluation

At this stage we will evaluate the effect of introducing the visualization on four participation factors: 1) the amount of the original contribution, 2) the number of active contributors, 3) the number of given comments, 4) the amount of the totally shared

files including both the new links and the replicates. After the experiment is completed, we will add more factors such as the usage frequency and the time staying online etc.

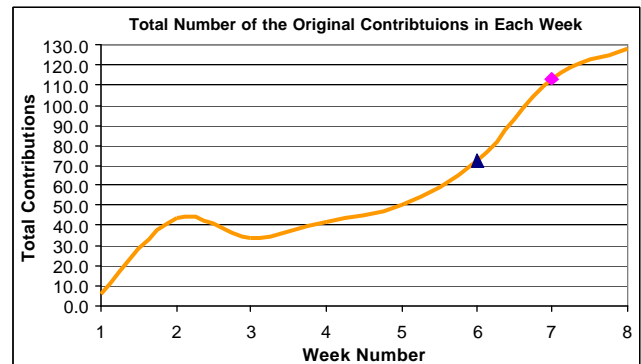
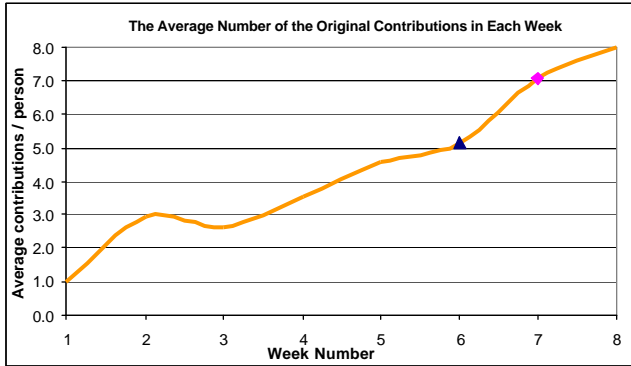


Figure 10. The total number of the original contributions.

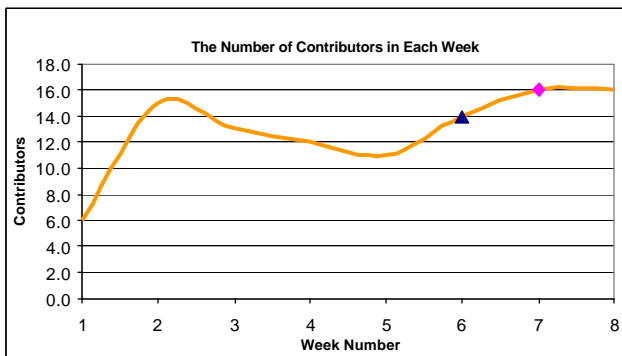
The **original contribution** of a peer is the number of new file links that s/he brings into the Comtella community from the web. Since Comtella servers are designed to share the downloaded links by default, some of the links that a peer shares may not be his/her original contribution but his/her consumption. Increasing the number of useful original contribution is one of the most important goals of the visualization. Figure 10 shows the distribution of the total amount of the original contributions brought by each participant on each topic in the corresponding week(s) over a 10-week period of time. Note weeks 6 and 7 were on the same topic and include in-between the school reading break week so these three weeks are merged together and are represented by one data point in the following figures. Therefore, there are eight data points instead of ten in figures 10 through figure 12 and table 1. The data point 7 corresponds to the calendar week 9, and point 8 to calendar week 10.

Figure 10 clearly shows a large increase in the amount of original contribution at point 7 (week 9), when the motivational visualization has been added into the Comtella servers. This indicates the effectiveness of the visualization in motivating more original contributions. There is another big increase from point 5 to 6 but this is because the data at point 6 is for three weeks. The original contributions at point 7, i.e. after the visualization has been applied to Comtella are twice as much as those in week 5 and almost triple the number in earlier weeks such as week 1, 2, 3, and 4. The extremely low contribution in the first week is caused by the participants' unfamiliarity with both the system and their course material. In the second week, students became familiar with their class material and were able to find more file links to share, meanwhile they also became familiar with the Comtella system and their curiosity also drove them to try out the system. These facts caused the bump at point 2. But after that the curve is normalized at point 3 (week 3) and steadily climbs up by a slight amount for each of the following weeks until the visualization was introduced (in week 9, at point 7). Considering the data at point 6 is the collective for three weeks (the calendar weeks 6, 7 and in between the reading week), the actual amount of original contributions per week in this time period decreased.

This is understandable since most of the students had midterm exams in week 6 and week 7. The sudden jump at point 7 in figure 10 indicates that the application of the motivational visualization in the Comtella a significant increase in the number of the original contributions.



**Figure 11. The average number of the original contributions per participant.**



**Figure 12. The number of contributors in each week.**

Figure 11 shows that the average number of the original contributions brought by each participant matches the pattern in figure 10 except that in figure 10 the slope of the curve from point 3 to point 5 is smaller than it is in figure 11 and in figure 11 the jump at point 7 is not as high as it is in figure 10. These are reasonable if we look at figure 11 and 12 together. Figure 12 shows the number of contributors in each week constantly decreases from point 2 to point 5 and considering the data at point 6 are for three week, the decrement in these three weeks was even greater. So it actually stopped decreasing at point 7 (week 9) instead of point 6. This explains why the slope of the increase from point 3 to 5 is bigger in figure 11 than it is in figure 10. With the Comtella system without the visualization, peers are not able to see their community so social comparison is not possible. As time goes by (from week 2 to 7), participants were losing their interest in the system and some of them even stopped making original contributions. However participants who are still interested in the system continued to bring in new links for sharing. The small increase in the total amount of the original contribution in figure 10 is in fact brought by a decreasing number of contributors thus the average amount of the original contribution in figure 11

increases with a slightly larger slope. The phenomenon of some users becoming very dedicated contributors was observed also later, after introducing the visualization on a great scale.

When the motivational visualization was added into the Comtella servers, participants were able to see their community both from an overall view and from the view on a specific topic. Once participants were able to view their community they were able to compare with each other and this makes the system more appealing and introduced a game-like competitive element. Figure 12 shows that at point 7 (week 9) there is a noticeable increase in the number of the contributors, which indicates that the sudden increase in the total amount of the original contribution (figure 10) is brought by a greater number of contributors (figure 12). Thus, the average contribution per person (point 7 in figure 11), even though still higher than it is in the previous weeks, is not on such a steep curve as the total number of contribution in figure 10.

The **amount of comments** is another important measurement of user participation. It describes the degree of a more active type of participation – giving comments on the shared materials in the community, which would be either one’s own shared materials or the materials brought by the others. This helps with improving the quality of the resources in p2p networks. But it takes users extra time and efforts to do so. For the purpose of their class, participants in our experiment were required to give at least one comment on at least one link per topic i.e. 1 copyright (per week) and it is up to them to give more. Row “Comments” in table 1 shows that participants generally took efforts to give more than one comment on each topic. The data in table 1 fluctuates a lot but still roughly follows the same pattern in figure 10 through 12. The first week has a minimum amount of comments, followed by a big increase, then an obvious decrease, and then smoothly increases until a sudden bump at point 7, which is for week 9 when the visualization was introduced. The same pattern can also be found in the total amount of contribution made by the participants on each topic (rows marked as “Total links shared” in table 1). It seems that the current version of the visualization effectively inspires peer participation through encouraging social comparison. However, after reading through the actual content of most of the contributed links we discovered one major defect of the current visualization. The **quality** of the contributions deteriorated a lot as the number increased after the use of the visualization. A massive number of files contributed after the visualization has been introduced are replicates of the same materials. A lot of similar comments are given on the same material and the comments are generally shorter and less meaningful, i.e. do not provide much guidance to the actual content of the material and sometimes are even irrelevant to the material. Some shared links are not related to the topics that they claim to be about. All of these indicate that the participants did not pay attention on the quality of their contribution and sometimes even did not read their papers before they contributed them, but they simply competed on one dimension – quantity – to exaggerate their nodes in the visualization. Another problem discovered is that, when a peer is reaches the top of the list, there is no particular motivation (unless of course, the fear of losing his/her position) for him/her to continue actively participating and contributing. The short time after introducing the visualization, however did not allow this to happen to many users.



**Table 1. The number of comments and the total number of the shared links.**

Week 1 to 4	1	2	3	4
Comments	31.0	220.0	141.0	75.0
Total links shared	30.0	186.0	122.0	72.0
Week 5 -10	5	6 & 7	9 *	10
Comments	292.0	251.0	395.0	797.0
Total links shared	251.0	240.0	324.0	472.0

\* Week 9 in this table is the first week when the visualization was introduced.

## 6. CONCLUSIONS AND FUTURE WORK

The experiment on the Comtella system with and without the motivational visualization proves that the approach of using motivational visualization to encourage social comparison and inspire active peer participation in p2p communities is feasible. The preliminary experiment results discussed in the previous section already showed that with our current version of the motivational visualization the participants' awareness of their virtual community increases considerably, the amount of the original contributions goes up significantly, and the goal of attracting more participants has also been achieved.

However, in order to gain higher recognition the users tended to contribute a large number of low quality materials. They managed to do so by downloading the materials from others and re-sharing the replicates, by contributing any material that they can found without putting much effort on reading the material and thinking about its content, by copying the others' comments on some materials as their own comments on the same material. The current version of the visualization takes into account only the quantity, not the quality of the contributions and it worked well in stimulating quantity competition but poorly in stimulating quality competition.

Maintaining active participation is another problem. It occurs when a peer is already one of the best contributors. In this case, how to keep his/her continuing actively or even more actively contributing is a question.

The experiment on the current version of the motivational visualization has not been finished yet. There is a substantial amount of data collected and more data will be available by the time the experiment is completed. I am going to further analyze the data and present a more detailed evaluation for this version of the visualization. A questionnaire was developed for interviewing all of the participants in the current experiment to confirm my analysis and explanations as well as get more subjective feedback on how they think of the visualization and their suggestions on the future improvements, whether or not they are emotionally affected by the visualization and which elements were most important in motivating active participation. Although the evaluation in section

5 has shown that the peers in the Comtella community were greatly motivated by the visualization in contributing more materials and giving comments on each other's contribution, the quality of the contributions is still problematic. The future version of the visualization has to find a way to encourage social comparison not only in the quantity of the contribution but also in the quality of the contribution.

Moreover, the current version of the visualization has to be modified to fit in the original Comtella system where users share research papers. As I mentioned that the current visualization is integrated with a simplified version of the Comtella server, which does not have a user model and does not build relationships between users, and is for sharing file links related to a class instead of the actual files, and the students are requested to find class-related links and rewarded with participation marks, which is already a strong source of motivation. These simplifications allowed us to focus on the evaluation of the visualization. But in the future the visualization will be integrated into a fully functioned Comtella server developed previously. More experiments will be carried out to test the overall effects on the peer participation, the amount of the contributions, and more importantly, the quality of the contributions.

## 7. ACKNOWLEDGEMENT

Thank you very much to my supervisor Prof. Julita Vassileva for her instructive suggestions, thank you very much too to all my teammates, especially Christopher Cox and WeiDong Han for their help in the implementation, and most importantly thank you to all the students, researchers, and professors who participated in the experiments by using the Comtella system and gave valuable feedback.

## REFERENCES:

- [1] E. Adar, B.A. Huberman. *Free Riding in Gnutella*. 2002. Available online at: [http://www.firstmonday.dk/issues/issue5\\_10/adar/](http://www.firstmonday.dk/issues/issue5_10/adar/)
- [2] H. Bretzke and J. Vassileva. *Motivating Cooperation in Peer to Peer Networks*. Proceedings User Modeling UM03, Johnstown, PA, June 22-26, Springer Verlag LNCS, 2003.
- [3] Y. Chawathe, S. Rathasamy, L. Breslau, *Making Gnutella-like P2P Systems Scalable*. SIGCOMM'03, August 25-29, 2003, Karlsruhe, Germany.
- [4] D. Cosley, M. Rashid, I. Albert, S. K. Lam, S. M. Mcnee, J. A. Konstan, J. Riedl. *Getting to Know You: Learning New User Preference in Recommender Systems*. IUI'02 January 13-16, 2002, San Francisco, California USA. ACM.
- [5] 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.
- [6] S. G. Eick and G. J. Wills. *Navigating Large Networks with Hierarchies*. IEEE Visualization '93 Conference, pages 204-210.

- [7] M. Ginsburg and S. Weisband. *Social Capital and Volunteerism in Virtual Communities: The Case of the Internet Chess Club*. HICSS-35, Virtual Communities Minitrack, Hawaii, January 2002. IEEE.
- [8] M. C. Humphrey. *Creating Reusable Visualizations with the Relational Visualization Notation*. 11<sup>th</sup> IEEE Visualization 2000 Conference (VIS 2000). October 08 – 13, 2000. Salt Lake City, UT.
- [9] A. Jameson, S. Baldes, and T. Kleinbauer. *Generative Models of Group Members as Support for Group Collaboration*. 8<sup>th</sup> International Conference on User Modeling, Sonthofen, Germany, July 13-17 2001.
- [10] J. Kleinberg. *Small-World Phenomena and the Dynamics of Information*. Advances in Neural Information Processing Systems (NIPS) 14. 2001.
- [11] T. Klingberg and R. Manfredi. *The Gnutella Protocol Specification v0.6 Draft*. June, 2002. Available online at: [http://groups.yahoo.com/group/the\\_gdf/files/Development/GnutellaProtocol-v0.6-200206draft.txt](http://groups.yahoo.com/group/the_gdf/files/Development/GnutellaProtocol-v0.6-200206draft.txt)
- [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*. Available online at: [http://www9.limewire.com/developer/gnutella\\_protocol\\_0.4.pdf](http://www9.limewire.com/developer/gnutella_protocol_0.4.pdf)
- [14] D.S. Milojevic, V. Kalogeraki, R. Ludose, K. Nagaraja K, J. Pruyne, B. Richard, S. Rollins, and Z. Xu, *Peer-to-Peer Computing*, HPL-2002-57, March 8<sup>th</sup>, 2002. Available online at: <http://www.cs.ucr.edu/~vana/kalogeraki-HPL-2002-57.pdf>
- [15] H. Nakanishi, S. Nakazawa, and T. Ishida. K. Takanashi. *Using Balance Theory to Understand Social Agent*. Workshop on the First International Joint Conference on Autonomous Agents & Multi-Agent Systems. July 16 2002 Bologna, Italy.
- [16] W. S. Ng, B. C. Ooi, and K. Tan. *BestPeer: A Self-Configurable Peer-to-Peer System*. 18<sup>th</sup> International Conference on Data Engineering (ICDE'02), p.272, February 26 – March 01, 2002.
- [17] 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.
- [18] A. Rowstron and P. Druschel. *Pastry: Scalable, decentralized object location and routing for large-scale peer-to-peer systems*. Proc. Of the 18<sup>th</sup> IFIP/ACM International Conference on Distributed Systems Platforms (Middleware 2001). Heidelberg, Germany, November 2001.
- [19] Sharman Networks Ltd. KaZaA Media Desktop, 2001. Available online at: <http://www.kazaa.com/>.
- [20] I. Stoica, R. Morris, D. Karger, M. F. Kaashoek, and H. Balakrishnan. *Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications*. SIGCOMM'01, August 27-31, 2001, San Diego, California, USA.
- [21] T. C. Sprenger, R. Brunella, and M. H. Gross. *H-BLOB: A Hierarchical Visual Clustering Method Using Implicit Surfaces*. 11<sup>th</sup> IEEE Visualization 2000 Conference (VIS 2000) October 08 – 13, 2000. Salt Lake City, UT.
- [22] J. Vassileva. *Motivating Participation in Peer to Peer Communities*. Proceedings of the Workshop on ESAW. ESAW'02, Madrid, 16-17 September, 2002. Available online at: <http://www.ai.univie.ac.at/%7Epaolo/conf/esaw02/preproc/E0029.pdf>
- [23] H. Zhang, A. Goel, R. Govindan. *Using the Small-World Model to Improve Freenet Performance*. The 21<sup>st</sup> Joint Conference of the IEEE computer and communications societies, IEEE INFOCOM 2002. Available online at: <http://www.ieee-infocom.org/2002/papers/386.pdf>