Stimulating User Participation in a File-Sharing P2P System Supporting University Classes

Julita Vassileva, Ran Cheng, Lingling Sun, Weidong Han

University of Saskatchewan, Computer Science Department, Saskatoon S7N 5A9 Canada

jiv@cs.usask.ca

Abstract We have implemented a small-scale peer-to-peer based environment called Comtella, which allows faculty, graduate and undergraduate students to contribute and share URLs of class-related resources, like academic and popular articles. The system has been deployed for three months in a university course on ethics and information technology with thirty five undergraduate students. We deployed two motivational strategies to encourage users to contribute resources. The evaluation results showed a significant increase in participation and the number of contributions. Future work will focus on ensuring mechanisms for collective quality control of the contributions.

1. Introduction

Typically, most peer to peer (P2P) file sharing applications are large-scale, involving thousands of peers. They have been used for sharing music, video-clips, movies and games and due to copyright issues have been typically in the grey zone between legal and illegal. There have been few attempts to build smaller scale P2P applications to serve a particular organization. There are specific problems faced in the development of small-scale P2P applications, and ensuring presence, participation and contribution is the most crucial one. A small-scale system can not rely on the redundancy of resources ensured in a large-scale system. Ensuring that there are enough peers on-line is crucial, since otherwise the files shared by off-line peers are not available and the infrastructure for propagating queries suffers. In addition to presence, it is important to ensure constant flux of new shared resources. Otherwise the system quickly reaches a state where everyone has downloaded all resources available which one is interested in, and there is no reason to search in the system, since there is nothing new. In a large system, like most successful existing P2P filesharing systems (KaZaA, eDonkey, BitTorrent etc.), the huge numbers of users participating and sharing resources ensures that there are enough resources available for search at any moment of time, and enough new resources are added. However, for small scale systems, it is not trivial to ensure these things.

This paper presents our experience in developing a successful small-scale P2P system. We have developed a system called Comtella that allows sharing papers downloaded from the web, or web links to papers for the purpose of a research group or a class of students. It uses a form of physical centralization to ensure presence and motivational strategies to persuade users to bring in new resources. The paper is organized as follows: the next section presents the Comtella project. Section three discusses the first experiences and problems encountered. Section 4 reviews approaches for ensuring incentives for contributions in P2P systems and on—line communities. Section 5 presents an implementation of Comtella to support a university undergraduate class. Section 6 presents the strategies adopted in Comtella for motivating user participation and section 7 – the evaluation results. Section 8 presents the main conclusions and directions for future work.

2. The Comtella Project

The Comtella system was developed at the Multi-Agent, Decentralized, Mobile and Ubiquitous Computing (MADMUC) lab at the Computer Science Department to support graduate students in the laboratory to share

research papers found on-line [14]. Comtella uses an extension of the Gnutella 0.6 protocol [8] and is based on a JTella client [11]. Each user needs to download a application (called "servent") which allows sharing new papers with the community (typically, pdf files) and searching for papers shared by oneself and by the other users. The shared papers need to be annotated with respect to their content as belonging to a subject category.

We did not want to include a heavy-weight full-text search techniques on board of the client and decided to use a shared hierarchical organization of categories among the peers, because it allows both indexing new documents and search in form of browsing through a hierarchy which is a familiar technique for users, similar to the way they store and locate files on their hard disks. In retrospect was not a wise decision, first, because it is hard to make users agree on a unified taxonomy of subject categories and second, because it implies a form of "semantic" centralization for the P2P system. We used a list of categories that we believed was well-accepted – an extended subset of the ACM subject category index, containing the areas of interest of the department. However this category index turned out to be outdated, large and complex, so we had to develop techniques around this to allow users to delete and suggest new categories and a community based-process for modifying the hierarchy categories. The problem of pushing this list back to the clients was solved with periodic release of new versions, but there were compatibility problems since not all users upgraded their clients regularly. We are currently looking for better ways of achieving a decentralized semantic organization that allows a convenient way of indexing new papers and convenient search options.

Currently a hierarchical category list is used to classify the shared papers. The user searches by specifying a category and receives a list of all own papers and papers shared by others related to this category. From the list of results, the user can download the desired papers and view them in a browser. Thus the basic functionality of Comtella is similar to several recent systems which have attempted to harvest the efforts of independent libraries, web-catalogues and individual users for sharing bibliographic information and papers, most prominently LOCKSS (http://lockss.stanford.edu/), Bibster (http://bibster.semanticweb.org/) and DAD [7].

3. First Experiences and Problems

Comtella was used for three months in our department on an experimental basis using our campus network (10Mb/sec) which includes various platforms: Windows, Linux, Unix, Solaris, NetBSD and MacOS. There were logistics issues, related to the fact that the system was fully distributed. The necessity to have one servent on each computer and the impossibility to relate more than one servent to the same user meant that the users who wanted to use the system at home and in the office had to always leave their servents running on both machines, so that they could access from work their own papers shared at home and in reverse. The user's servents on the home computer and on the work computer are just like servents of two different users, with different ids, lists of shared papers etc. In order to access the papers shared by another user, that user has to have his/her servent running and in this case, to be able to access one's own files on the other computer, one had to keep that computer (with a Comtella servent on it) running. This proved to be a problem, because users typically switch off their home computers when they are at work. The users tended to start their servents only when they wanted to search for papers and to quit them afterwards. Due to the relatively small size of the user group (25 users), this lead to very few servents being on-line simultaneously, so typically there were very few (if any) results to a query.

The problem of participation is well known also for large scale systems. It is very important to ensure a critical mass of on-line servents to maintain an infrastructure that guarantees successful searches and attracts more users. Bit Torrent for example, uses a protocol that finds and uses several different sources to download a single file to ensure redundancy (if some of them shut down) and better speed of downloads. But to achieve this redundancy Bit Torrent too needs a critical mass of users who keep their servents running.

Finally, we observed that even when most of the users keep their servents running for a certain time, this is not enough to ensure a useful system. If there are no new resources injected regularly into the system (by users bringing in and sharing new papers), very soon it makes no sense for a user to search in his/her main area of interest since there is nothing new in the system that the user has not seen or downloaded before. Ultimately, the system reaches equilibrium where everyone has all papers that everyone else has. The sharing system is then useless. In order to achieve a dynamic and useful system, the users have to share new papers regularly. Motivating users to contribute new resources is an important problem that can not be solved on a protocol level. The user needs to be involved and motivated to feel as a part of a community and incentives need to be built in to stimulate active participation.

4. Work on Incentives for User Participation

Studies on P2P networks [1] show that only few of the peers contribute resources. Many existing p2p systems have developed their own strategies to motivate users to make contributions to their virtual communities. Only several more characteristic examples are listed below.

Simple solutions to this problem have been used in the popular file-sharing systems like KaZaA and LimeWire. Typically, the interface makes terminating the servent particularly hard for the user: even when the user clicks the “close window” box, the application is not shut down, but runs in the background, passing messages and serving requests.

"Direct Connect" (www.neo-modus.com) and Limewire deploy a simple strategy to ensure user contributions – they force users to share a minimum amount of resources. If a user fails to meet the required quota, his or her access will be limited or completely denied. Although this method encourages dedicated users to contribute, there is no data about the numbers of users who decided not to participate at all because of this constraint. Anecdotal data suggests that many people tend to not join such exclusive communities since they are forced to contribute before receiving any benefits.

Many authors have proposed theoretical models for economic incentives in terms of micro-payments for uploads and downloads [4,9]. Mojo Nation (www.mojonation.net) attempted to introduce an electronic currency and micro-payments to provide economic incentives for sharing resources. The users needed to pay for each download and the users who share resources got paid. However, the approach was not successful and Shirky [12] pointed the reason for the failure in the cognitive cost for each transaction. The act of buying anything, even if the price is very small, creates mental transaction costs, that is, the energy required to decide whether something is worth buying or not. People would rather pay a flat rate than think about the cost of every small purchase.

Another way to promote users to participate and contribute is by rewarding the active users with better quality of services (faster downloads), as Kazaa Lite and eMule do. The systems record the actions of users and maintain numeric participation levels for each user. The speed of downloads the user can get is determined by the user’s participation value, which seems to be calculated as a function of the difference between how much resources (in MB) have

been downloaded from the by other users and how much resources the user has downloaded from others.

The main problem with this strategy, as well as with other economic-based strategies is that they do not reward users who share files that are not of wide common interest. Such users feel frustrated and treated unfairly, they withdraw and the variety of files offered decreases.

Similarly to P2P systems, some server-oriented online communities use various mechanisms to motivate people to join and contribute high quality resources (typically, posts). Although these online systems are different from P2P systems in terms of architecture, the ideas of their motivation strategies can be applied to P2P systems. The best example in this category is Slashdot (slashdot.org) which measures the users' contributions to the community in a unit called "karma" [10]. If the user's posts are highly rated by the moderators, the user earns karma in the system related to some special privileges. For example, the user's subsequent posts begin life at a higher rating than usual. The users with high karma are more likely to be chosen as moderators in the future. Users who are moderators can rate other users' posts, consuming their own karma. That means the users with high karma have more ratings to give away and are therefore more influential in the community. This strategy also stimulates the members to submit high quality posts.

We chose to study approaches that are not based on payment and instead looked into the area of social psychology on research about how people can be motivated to do particular things. We found two persuasion strategies, deployed in advertisement and customer relationship management (CRM) which we believe can be useful for our purpose. These include visualization of the community [3] and introducing a concept of hierarchical community memberships, based on the user's contribution to the community.

We applied these two strategies in a new motivational version of Comtella which was deployed in an undergraduate class on Ethics and Information Technology. In order to enable Comtella to serve as a tool to support a class some modifications from the original design were required, which are presented in the next section. A detailed description of the motivational strategies and the experiment results are presented in sections 4 and 5.

5. Comtella in the Ethics Class

"Ethics and Information Technology" is a fourth year undergraduate computer science class at the University of Saskatchewan discussing topics related to privacy, freedom of speech, intellectual property, deskilling, workplace issues, outsourcing, as well as professionalism and ethical decision making. The class uses an excellent textbook [2]. However, it presents mainly US-related context (legislation and example cases). The nature of the material demands that there is a lot of discussion, since there are many "hot" issues (for example, job outsourcing, software patenting) and legal case-developments (legislation related to intellectual property and file sharing for example).

Comtella was introduced in this class as a tool for students to find and post on-line articles and materials related to the weekly themes of the class. In previous class offerings the students had to find and post class-related web-links on their personal websites dedicated to the class. More details and comparison of the two approaches of sharing class materials can be found in [15].

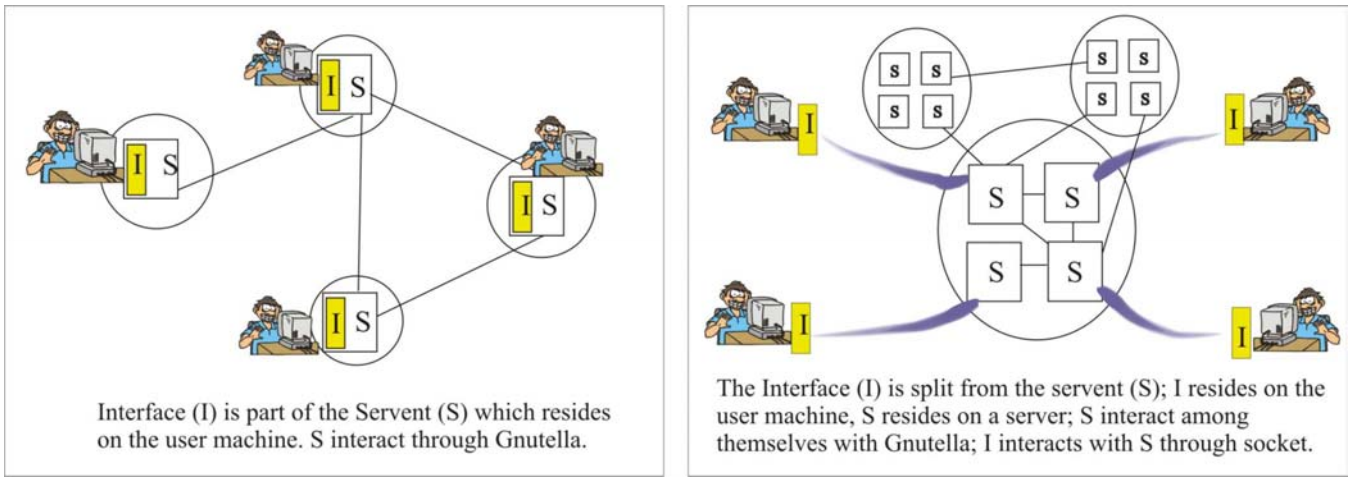

Figure 1. Modified architecture to ensure that servents run constantly and shared files are always available.

Three modifications of the Comtella system were necessary to ensure the basic level of participation and thus – the reliability needed for a course-support system.

First, we separated the user interface from the basic servent's functionality and moved all servents on a server machine (see Figure 1, right). We provided for the users a downloadable Java-Swing interface to allow them to log into their servents on the server. In this way we could also control the access to the system and limit it to include only the registered students in the class. The shared resources by each servent, together with the servent resided on the server machine and thus were always available independently of whether the user was currently logged in or not; since the servents ran all the time. In this way, physically, we had a server-based system. Logically, however, the system was decentralized, since all the servents even though residing on the same machine used the Gnutella protocol to search for documents (weblinks in our case). At any point they could be moved away from the server to a different server, or could be distributed among various servers and even moved back to the users' machines, as standard Gnutella clients (Figure 1, left).

Second, instead of sharing the entire files containing the articles, as in the original Comtella, the users of this version shared only web-links, i.e. the URLs of articles that they found on-line, just like the sharing of bookmarks in DAD [7]. The interface allowed viewing of the article in a browser directly by clicking on the list of search results, without the need of downloading and sharing the link (see Figure 2).

The categories used to annotate the bookmarks corresponded to the main themes of the class curriculum and textbook chapters. There was no hierarchy; annotating of shared web links and searching by theme was straightforward. Each theme was covered in one week, with the exception of theme 6 (Computer Crime and Security), on which three weeks were devoted with one week of spring-break in-between.

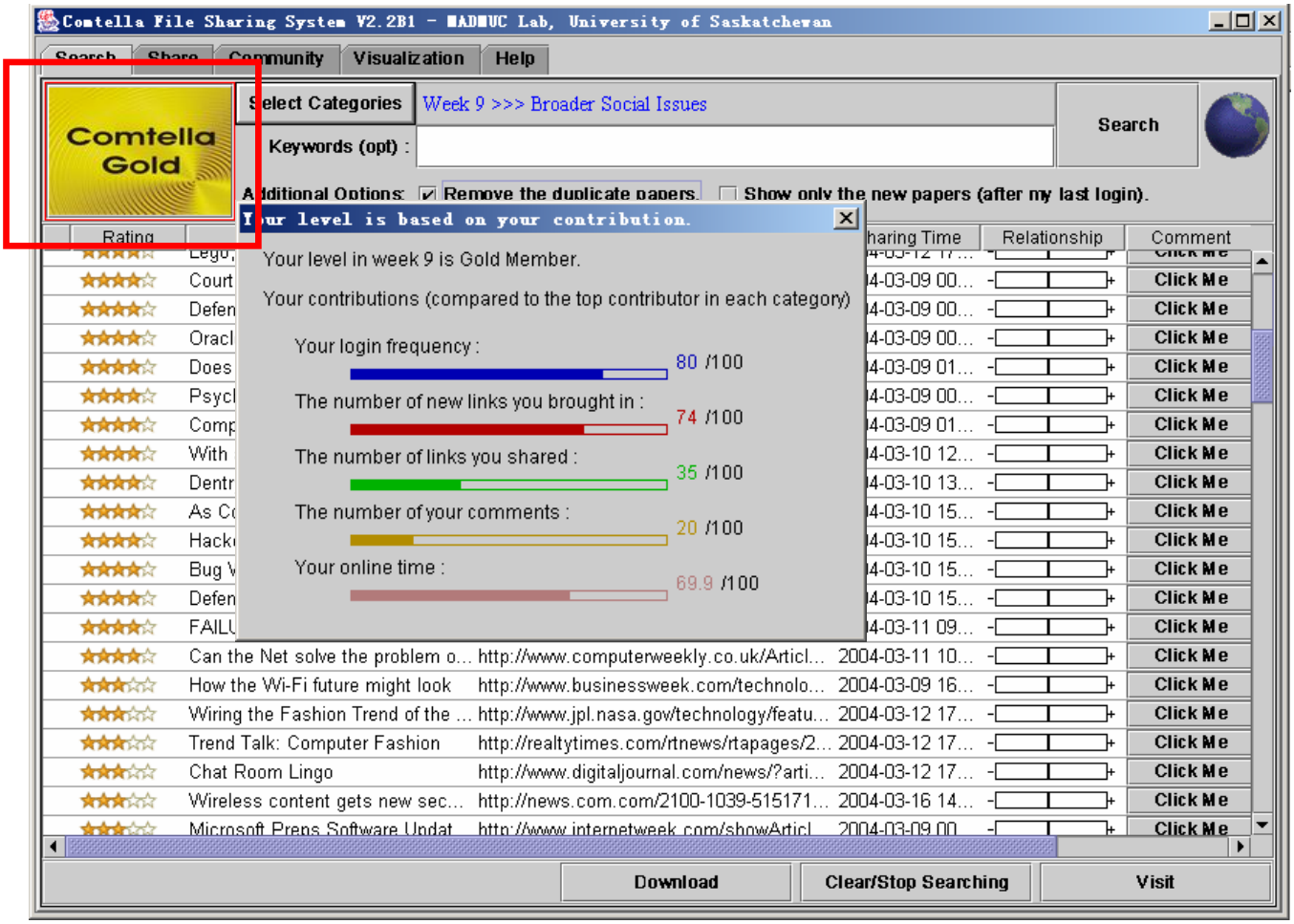

Figure 2: Search in Comtella and explanation of the user's membership level.

Third, the Comtella interface was modified to allow users to rate and comment the articles that they shared. The students were encouraged to download and share web links that they found from others only if they have viewed and liked the article and wanted to comment on it. The intention was to minimize the duplication of resources among the peers. In this new version of Comtella, duplication of resources was no longer needed, since all servents were active all the time on the server. In this way a resource (web link) shared by a user was always available independently if the user was connected or not to the servent.

6. Motivating User Participation in Comtella

To ensure that there is a constant flux of new resources contributed to the system we applied two principle strategies to motivate user participation:

- Introducing hierarchical membership in the community of users, and
- Visualization of the community and the users' activities.

The main idea of the first strategy is to stratify the user community into different classes based on their participation and give different privileges to the different classes, e.g. different

interface options and a better visibility in the community. This strategy is used in Customer Relationship Management (CRM), for example, by airlines or supermarket chains that reward loyal customers with points accumulated on their club membership cards and which can be exchanged in free groceries or travel. Users who have achieved a higher membership level fear that they may loose it together with the privileges associated with it, and act so as to avoid this, e.g. tend to purchase their groceries in the particular store, or fly with particular airline [13]. In Comtella users accumulate participation points and are rewarded in terms of better quality of service (better search options).

The metric for computing user participation takes into account the following contribution categories: the number of resources (web links) that are newly introduced in the system by the user, the number of shared resources by the user (also including those copied from other users), the time the user spent on-line (i.e. logged into the servent), the number of ratings and comments given by the user. The participation score is computed as a weighted sum of these factors (normalized with respect to the total contribution by the whole class in each contribution category).

Three levels of membership are introduced: bronze, silver and gold. The membership level of each user is re-calculated on a weekly basis, based on the level of the user's participation score from the previous week. The user can see his/her level visualized by a virtual "card" in the upper-left corner of the search screen (see Figure 2). If the user clicks on the card, he/she can see a graphical representation of his/her contribution level (several bars comparing each aspect of participation to the current top contributor in this aspect). This representation suggests the activities in which the user should engage to improve his/her membership level (the open window in Figure 2). Users with higher membership level are rewarded with extended search functions, e.g. filtering out duplicate results or articles that the user has already seen and sorting the results alphabetically, by provider, by date of sharing.

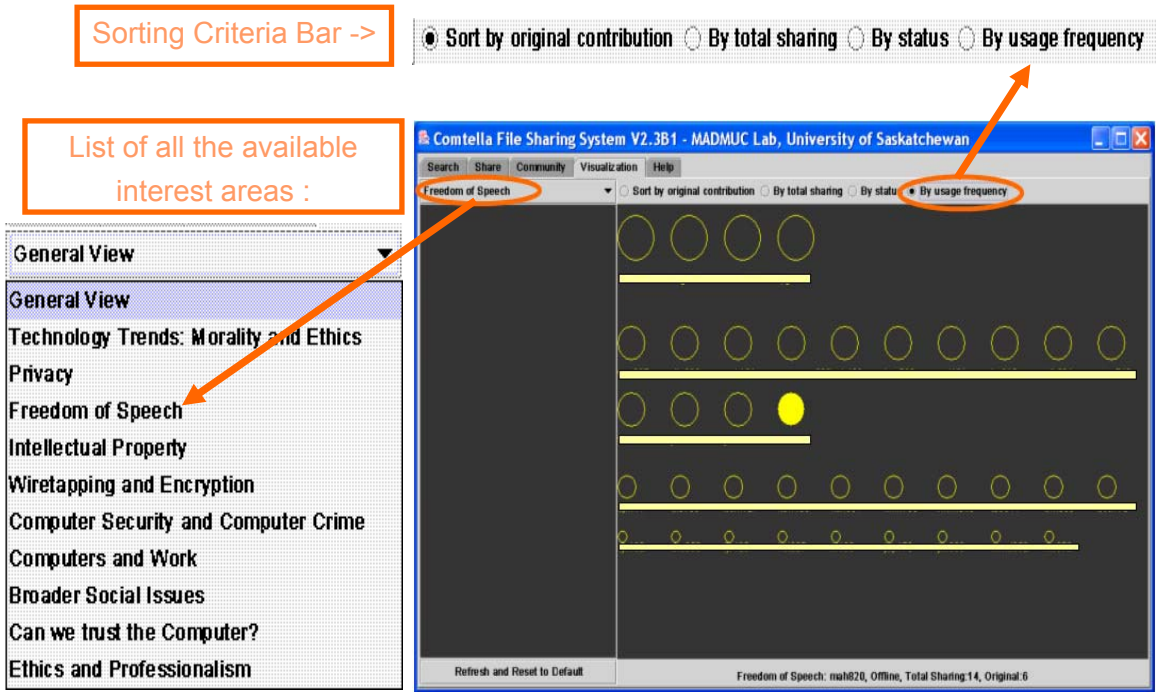

Figure 3: Visualization of the community with respect to the current membership level of the users (the user ids below each node are blanked to preserve user anonymity)

The second motivational strategy deployed in Comtella uses a visualization of the community. Such visualization creates social awareness and can stimulate social comparison and competition among users [5], and reward highly contributing users with a feeling of significance in the community.

The visualization represents the users as nodes, whose radius shows the level of contribution of the user (see Figure 3). Users who participate more actively in the system are rewarded with better visibility: their nodes are larger.

The colour of the node shows the membership level (gold, silver, bronze); the node is filled if the user is currently logged into his/her servent and empty if the user is not logged in.

Clicking on a node shows the web links shared by the corresponding user. The user can choose different views of the community: a weekly view that shows the contributions of the users for a given week (category), and particular forms of contribution (see Figure 3). In this way, the user can view the current state of the community.

7. Experimental Results

In order to evaluate the motivational influence of the proposed motivational strategies, we compared the level of contributions of the same group of students using two versions of the Comtella system: first a version without and then a version with the motivational strategies. The experiment took place during the "Ethics and Information Technology" class offered in the second term (January to April) of 2003/2004 by the Computer Science Department with thirty five (35) registered students. We introduced the version implementing the motivational strategies in the middle of the class (after the three weeks dedicated to theme 6). The results showed that the number of new shared web links increased significantly after the introduction of the motivational version The median contribution number for the period before introduction of the motivational version was 51, while in the four weeks after introducing the motivational version it increased to 154, i.e. 3 times. The peak of contributions happened two weeks after introducing the new version. After that peak, there was a decline with the level of new contributions going close to that before the new version. Obviously, the motivational strategies were able to stimulate user participation, at least in the short term.

The quality of shared web links decreased after introducing the motivational version. While before introducing the new version there were virtually no shared articles that were unrelated to the current theme, after introducing the motivational version there were 25 unrelated links shared in the first two weeks and close to 50 in the third week. Obviously, some users were sharing links just to boost their contribution numbers. Four of the 35 users showed this behaviour. One extreme case was the top contributor of the class, who shared 131 (13%) of the 973 links submitted by all students, of which 40 (~30%) were irrelevant.

We observed that users clicked on the membership card often to check their participation level, which showed that they cared about it. The additional functions available to the Gold and Silver members were used by approximately 65% of the users who had access to these functions. The visualization was also used often, though not in its full functionality. The most frequently used view was the default one which showed the size of nodes based on the number of the new papers they brought in. Maybe due to this default setting some of the users concluded that this is the most important form of participation (versus for example, commenting or rating papers, or being on-line), and tried to maximize it by "cheating".

After the last week of the experiment, the students were asked to answer a detailed questionnaire about their experience with Comtella. Thirty-one (31) students responded to the questionnaire. Seventy percent of the users had positive reaction (+1 or +2 on a scale from minus 2 to plus 2) and sixty-seven percent said that they would recommend the system to be used in other classes.

When explaining this increase of contributions we can not ignore two facts: first, that the motivational system was used only for a relatively short time (four weeks) and second that there were external factors in the experiment, since it was taking place in a real class situation where students are exposed to various pressures, motivations and time limitations. The effect of novelty is well documented in human-computer interaction research, where significant effects in user performance can be observed after introducing a new interface item, which become insignificant in a longitudinal study due to the fact that the users become accustomed to the interface. Nevertheless we believe that there is strong evidence that in our experiment the motivational strategies deployed in Comtella played the intended role. This can be seen from the increase in the amount of attempted "cheats", the usage of the interface functions related to the rank and from the student questionnaires.

8. Conclusions

We see two main contributions in our work. First our experience shows that P2P content sharing systems can be applied successfully in small-scale applications, for example, in a research group or to support a class that requires student involvement in locating relevant resources and discussion. However, certain modifications and compromises are necessary with the purist distributed P2P architecture in order to ensure a critical mass of users necessary for the functioning of the system. These modifications in our case involved a physical centralization (moving the servents on one dedicated machine) and applying incentive mechanisms to motivate users to contribute resources.

Our experiment shows that applying persuasion strategies is a promising way to stimulate participation and active contributions by students. More studies and studies with longer duration are needed to gain insight in the dynamic process of social motivation and there will be specifics depending on the complex interplay of factors from the context of a particular application.

Generally, care should be taken, since every effective incentive mechanism can motivate some people to cheat. To prevent this, appropriate measures must be taken. In the case of Comtella, it is important that not only quantity, but quality of contribution is rewarded.

We are currently working on a mechanism allowing users to peer-review and rate each others' contributions, on motivating users to rate contributions, and on measures to police cheaters, i.e. invalid ratings. The Slashdot strategy would be very appropriate to apply in Comtella in combination with the existing two strategies of hierarchical memberships and community visualization. It will help to motivate users to rate the web-links provided by other users and in this way to deal with the problem of some users trying to cheat their way to a higher membership levels by submitting irrelevant web-links. It will also provide an attractive reward for high-membership users in terms of more actual power (ability to rate up or down weblinks contributed by other users) in the community. We are currently testing a new Comtella version deploying this strategy.

Of course, the introduction of ratings may not solve the problem, since users may find other ways to increase their standing, e.g. by submitting false ratings, or by forming cliques. But if there aren't any attempts to "cheat" the

system, this most likely shows that the persuasion mechanism is not really working. We believe that a "healthy" level of cheating should be expected, though of course, it should be discouraged, to ensure fairness and not compromise the attractiveness of the rewards for honest users.

More information about Comtella, as well as a downloadable version of the servent is available on the MADMUC Lab homepage at: http://bistrica.usask.ca/madmuc/peer-motivation.htm

About the Authors

Julita Vassileva is an associate professor in Computer Science at the University of Saskatchewan. Her interests are in user modeling, advanced learning technologies, and user issues in decentralized peer-to-peer applications.

Ran Cheng, Lingling Sun and Weidong Han are Masters Students at the Computer Science Department studying use persuasion, community awareness and visualization in peer-to-peer networks.

References

[1] Adar, E. & Huberman, B.A., (2000). Free Riding on Gnutella. Available online at http://www.firstmonday.dk/issues/issue5_10/adar/

[2] Baase, S. (2003) A Gift of Fire: Social, Legal, & Ethical Issues for Computers & the Internet. Prentice Hall.

[3] Bretzke H., Vassileva J. (2003) Motivating Cooperation in Peer to Peer Networks, Proceedings User Modeling UM03, Johnstown, PA, June 22-26, Springer Verlag LNCS 2702, 2003, 218-227.

[4] Buragohain, C., Agrawal, D., Suri S. (2003) A Game Theoretic Framework for Incentives in P2P Systems. Proc. of the 3rd International Conference on P2P Computing (P2P2003), Linkoping Sweden.

[5] Cialdini R. (2001) Influence: The Science of Persuasion. Scientific American, February, 2001, 76-81.

[6] Comtella (2002) available to download from http://bistrica.usask.ca/madmuc/peer-motivation.htm

[7] Cordasco G., Scarano V., Vitolo C. (2004) A P2P Distributed Adaptive Directory. Proceedings Adaptive Hypermedia 2004, Springer LNCS.

[8] Gnutella (2000) available (accessed Aug 17, 2004) at http://www.gnutella.com

[9] Golle P., Leyton-Brown K., Mironov I. (2001). Incentives for Sharing in Peer-to-Peer Networks, *Proceedings EC'01*, October 12-17, 2001, Tampa, Florida, ACM press, 264-267.

[10] Johnson S (2002) Emergence: The Connected Lives of Ants, Cities and People.

[11] JTella (2000) available on-line at http://www.javaworld.com/javaworld/jw-10-2000/jw-1006-fileshare-p1.html

[12] Shirky C (2003). Fame vs Fortune: Micropayments and Free Content, First published September 5, 2003 on the "*Networks, Economics, and Culture*" mailing list, available online at: http://shirky.com/writings/fame_vs_fortune.html

[13] Nabi R. (2002) Discrete Emotions and Persuasion, The Persuasion Handbook: Developments in Theory and Practice, In J. Dillard & M. Pfau, Sage Publications, 289-308.

[14] Vassileva, J. (2002) Motivating Participation in Peer to Peer Communities, Proceedings of the Workshop on Emergent Societies in the Agent World, ESAW'02, Madrid, 16-17 September, 2002, Springer Verlag LNAI 2577, 141-155.

[15] Vassileva J (2004) Harnessing P2P Power in the Classroom. Proceedings Intelligent Tutoring Systems, ITS'2004, Lester, J.; Vicari, R.; Paraguacu, F. (Eds.). Lecture Notes in Computer Science No. 3220 305-314.