

User Motivational Mechanism for Building Sustained Online Communities

A Thesis Submitted to the College of
Graduate Studies and Research
in Partial Fulfillment of the Requirements
For the Degree of Master of Science
in the Department of Computer Science
University of Saskatchewan
Saskatoon, Saskatchewan

by

Ran Cheng

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

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

Abstract

The proliferation of online communities on the Internet nowadays may lead people to the conclusion that the development of custom-made communities for particular purpose is straightforward. Unfortunately, this is not the case. Although software providing basic community infrastructure is readily available, it is not enough to ensure that the online community will “take off” and become sustained. Most online communities suffer from the scarcity of the user participation in their initial phase. To address the problem, this thesis proposes a motivational mechanism to encourage user participation. The main idea is to introduce a set of hierarchical memberships into online communities and reward active users with better quality of services. The mechanism has been applied in a small-scale online community called Comtella and evaluated. The results showed that, although the mechanism was able to motivate users to participate more actively and make more contributions, it led to a deteriorating quality of user contributions, catalyzed “information overload” in the community and resulted in a decrease in user participation towards the end of the study.

Therefore, to regulate the quality and the quantity of user contributions and ensure a sustainable level of user participation in the online community, the proposed mechanism was improved so that it was able to adapt the rewards for particular forms of participation for individual users depending on their reputation and the current need of the community, thereby influencing their actions of contributing. The improved mechanism was also implemented and evaluated in the Comtella system. The results of evaluation showed that the mechanism can guarantee stable and active user participation and lower the level of information overload in the online community and therefore it can enhance the sustainability of the community.

Acknowledgements

First of all, I would like to sincerely thank my supervisor, Dr. Julita Vassileva, for her persistent support, guidance, help, and encourage during the whole process of my study and my thesis.

I also want to express my thankfulness to my advisory committee: Dr. John Cooke, Dr. Ralph Deters and Dr. Valery Chirkov for their valuable suggestions and comments.

In addition, I would like to thank my dear friends, Weidong Han and Lingling Sun, for the time and effort they put into the Comtella project and also thank all the students of the Department of Computer Science who participated in the evaluation studies on the Comtella system. Without their active participation, I would never accomplish this work.

Finally, I would like to thank my parents for their support and love all the time. I want to dedicate my thesis to them and tell them they are the best parents in the world.

Table of Contents

Permission to Use	i
Abstract	ii
Acknowledgements.....	iii
Table of Contents.....	iv
List of Figures.....	vi
List of Tables	vii
Chapter 1 Introduction.....	1
Chapter 2 Related Work on User Motivation.....	3
2.1 Strategies for motivating user participation in client-server systems	3
2.2 Strategies for motivating user participation in peer-to-peer systems	5
2.3 Discussion	7
2.4 Summary	8
Chapter 3 The Mechanism of Hierarchical Memberships	9
3.1 Overview of the solution.....	9
3.2 Measuring the users' contribution.....	10
3.3 Determining the users' memberships	10
3.4 Updating the user's memberships	12
3.5 Rewards for active users	13
3.6 Summary	13
Chapter 4 Evaluation of the Hierarchical Membership Mechanism	14
4.1 Overview of Comtella.....	14
4.2 Implementing the mechanism in Comtella	16
4.3 Evaluation	20
4.4 Discussion.....	25
4.5 Summary.....	26
Chapter 5 The Improved Incentive Mechanism.....	28
5.1 Approaches to decentralized quality control.....	28
5.2 Collaborative rating mechanism	30
5.3 Adaptive reward mechanism.....	31
5.3.1 Community model.....	32
5.3.2 Individual model.....	33
5.3.3 How the mechanism works	35
5.4 Summary	37
Chapter 6 Evaluation of the Improved Incentive Mechanism.....	39
6.1 Re-implementing the Comtella system.....	39
6.2 Implementing the mechanism in the new Comtella.....	41
6.3 Evaluation	45

6.4 Summary	58
Chapter 7 Conclusions and Future Work.....	59
References	63
Appendix A: The Socio-psychological Foundation of the Hierarchical Membership Mechanism	66
Appendix B: The Questionnaire Used in the Study in 2003-2004 Winter Session and the Results.....	68
Appendix C: The Process of the Wilcoxon Signed-Rank Test.....	77
Appendix D: Technical Details of the Web-based Comtella System.....	80
Appendix E: The Consent Form Used in the Study in 2004-2005 Winter Session.	82
Appendix F: The Questionnaire Used in the Study in 2004-2005 Winter Session and the Results.....	84
Appendix G: The Process of the Mann-Whitney Test	95

List of Figures

Figure 1. The user interface for the gold members.....	16
Figure 2. The window showing user’s contributions.....	17
Figure 3. The visualization panel.....	19
Figure 4. The total number of articles shared in each week.....	22
Figure 5. Total logon times of all users in each week.....	22
Figure 6. A segment of a search result list.....	31
Figure 7. An overview of adaptive reward motivation mechanism.....	32
Figure 8. The changing of the community reward factor (F_C).....	33
Figure 9. The changing of the individual reward factor (F_I).....	35
Figure 10. The relation between the adaptive weights for sharing (W_S) and rating (W_R).....	37
Figure 11. The architecture of the Web-based Comtella system.....	40
Figure 12. Welcome page of Web-based Comtella.....	42
Figure 13. The interface for sharing resources in Comtella.....	43
Figure 14. Different interfaces for users in different membership levels.....	44
Figure 15. The number of ratings contributed in the two groups in each week...	48
Figure 16. The number of articles shared for each topic in the two studies.....	52
Figure 17. The average times of reading articles for the users in both groups....	55
Figure 18. The average times of logon for the users in both groups.....	56
Figure 19. The component-level architecture of the Comtella system.....	80

List of Tables

Table 1. Different weights for five cooperative activities.....	17
Table 2. Rewards for the users with different memberships.....	19
Table 3. The use of additional functions.....	24
Table 4. Differences between the two systems for the two groups.....	45
Table 5. The relation between the articles' summative ratings and the times they were summarized by the users.....	49
Table 6. The differences between the users' real contributions and their expected contributions.	53
Table 7. The differences between the actual number and the expected number in the test group for the eight topics.....	54
Table 8. The users' estimations of the percentages of the articles with high, medium and low quality in both groups.....	55
Table 9. The number of ratings, the times of reading and number of logon actions in each week and the correlation coefficients for both groups.	57
Table 10. X_A , X_B , X_A-X_B , the absolute value and the sign of X_A-X_B for each user.....	78
Table 11. Level of statistical significance.....	79
Table 12. The numbers, the ranks, the sum and the average of the ranks in both groups.....	96

Chapter 1

Introduction

The Internet connects people to an ever-increasing source of information. It allows them to share and exchange resources and ideas through various online communities although they may be dispersed geographically. Online communities are social networks of users who share similar interests, practices or purposes, communicate regularly over a common medium, and are governed by some policies and norms. The proliferation of online communities nowadays may lead people to the conclusion that the development of custom-made communities for particular purposes is straightforward. Unfortunately, this is not the case. Although software providing basic community infrastructure is readily available, it is not enough to ensure that the online community will “take off” and become self-sustainable.

Most online communities suffer from scarcity of user participation and contributions in their initial phase. This is because a system with a small number of active participants is less useful. Most users tend to free-ride in the beginning until they are convinced in the benefit of sharing, which reduces the probability of finding useful resources in the system. When users realize the system is not useful, they are discouraged from participating further, and the community becomes stagnant and ultimately winds down to non-existence. This situation matches the findings by Hiltz and Turoff (1978). According to their research, a “critical mass” (or a certain number of active users) must be reached for an online community to be self-sustainable.

On the contrary, if there are more participants who are willing to contribute resources, the probability that a user finds what she wants becomes much higher. Users who benefit from the system perceive it as useful and tend to log on more often and make more contributions. Therefore, user participation, as well as the value of the system increases and shows the “network effect” (Reed 1999).

Thus, there is a feedback loop in online communities (Vassileva 2002). User participation determines the level of usefulness of the community and the usefulness of the community in reverse influences user participation. In order to have this feedback loop work in a positive way, incentives are needed to attract the users to join and contribute to the community.

Therefore, I propose a motivational strategy to encourage user participation in the online community. The basic idea is to introduce a set of hierarchical memberships into the online community and assign different memberships to the users depending on their levels of participation in the system. Higher memberships are associated with certain rewards: higher visibility in the community, more power, or better quality of service. The underlying hypothesis is that such rewards would motivate users to actively participate and contribute to the community. The proposed strategy was implemented and evaluated in an online community, based on the Comtella system (Vassileva 2002). Developed at the MADMUC Lab at the University of Saskatchewan, Comtella is a class-support system with a peer-to-peer infrastructure. It enables a community of researchers and students to share and exchange resources, typically Web-articles and research papers. Like other small-scale online communities, Comtella faces the challenge of ensuring user participation and contributions.

The thesis is organized as follows. Chapter 2 presents an overview of related work on motivating user participation in online communities. Chapter 3 describes the proposed mechanism of hierarchical memberships. Chapter 4 presents the evaluation of the hierarchical membership mechanism and discusses the findings. In Chapter 5, how the proposed mechanism was improved is described. Chapter 6 explains how the evaluation of the improved mechanism was setup and performed and discusses the results of the evaluation. Chapter 7 summarizes the conclusions of this work and proposes some directions for future research.

Chapter 2

Related Work on User Motivation

The research on encouraging user participation in online communities is still in its infancy. This chapter reviews the incentive mechanisms in existing online communities. The infrastructure of an online community has implications on the range of possible incentive mechanisms. Depending on the infrastructure, two categories of communities exist - client-server online systems and peer-to-peer systems.

2.1 Strategies for motivating user participation in client-server systems

Compared to systems with fully decentralized infrastructure, centralized systems allow more control over the policies in the communities they support and over the behavior of their users. Therefore, motivational mechanisms can be deployed more easily and efficiently.

“Slashdot” (slashdot.org) set a milestone in the development of strategies for motivating user participation in online communities. To measure the users’ contributions to the community, the founder of the system, Rob Malda, coined a term called “karma”. If the user’s posts are highly rated by the moderators, the user earns karma, which is associated with 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. After a user becomes a moderator, she spends her own karma rating other users’ posts. In this way, the users with high karma have more karma to rate posts and therefore are more influential in the community. Slashdot’s moderation mechanism has two merits. Firstly, it distributes the task of evaluating posts among the

large user pool, thereby making achievable a job that would otherwise have been overwhelming. Besides, the final ratings of posts are more unbiased since they are computed based on ratings from many users. Slashdot's strategy stimulates the members to submit high quality posts. Furthermore, "it set up an environment where community leaders could naturally rise to the surface" (Johnson 2001).

Sharing can be considered a socially desirable activity. Real human communities have evolved norms that reward such activity. However, in virtual communities, these norms do not always emerge. Erickson and Kellogg (2000) pointed out that there was a lack of social cues in online communities and to compensate for this, the community members and their activities should be visualized to enhance their mutual awareness and responsibility. Based on this idea, an online chat system named Babble was designed. In the system, a visualization component showed in real time the presence and activities (e.g. talking, listening) of the users engaged in conversations. The visualization allows the participants to get social cues (e.g. who is currently active, who is drifting away) and to coordinate their interactions better. This approach serves small to medium-sized groups very well, but it is hard to apply in online communities with a large population, where it would be hard for the users to identify who is represented in the visualization.

Jenny Preece (2001) considered that a sense of responsibility was very important for community members to be cooperative. Therefore she proposed three possible approaches to enhance users' responsibility: using registration procedure to prevent users from jumping between communities; reducing the anonymity to make users identifiable; recording users' past behaviors to build their reputations. Although all these approaches are helpful in building up the users' responsibility in the community, they have some shortcomings. As Preece realized, a registration procedure may deter some users from participating in the community. Identifying users in online systems is always a controversial issue. People are generally not willing to provide their real name or other personal information while participating in online social activities, since they have concerns for their privacy and safety. If they are forced to reveal their identity, most of the users may choose not to join a community. Building user reputation for behaving cooperatively in an online community is a very promising idea because good reputation is a social reward for cooperative behaviors in real human communities. However, not

all users could be motivated by social recognition and glory; to motivate more users, physical benefits (e.g. more power or authority, better quality of service) remain necessary.

Millen and Patterson (2002) suggested that an online community, although it may comprise many discussion forums, should have one common entry, where new messages and calendar items within the community are highlighted. This common space is supposed to encourage members to observe the activities in the community, thereby enhancing the sense of community. Besides, a notification service is recommended, through which members are informed when new information is posted on the community forums. Millen and Patterson found that the probability to visit the site doubles for both active and inactive users who get notifications. To some extent, this community entry space and notification strategy is similar to Erickson's community visualization because both of them try to promote social awareness of community members to stimulate their engagement.

Motivating user contributions through notification service was also attempted in the well-known movie recommender system, MovieLens (Cosley et al. 2003). In MovieLens, the contributions are the users' ratings for movies. An approach to stimulate users to rate movies through sending them email-invitations (Beenen et al. 2004) was proposed. The authors experimented with different texts in the email-invitations. The evaluation showed that users were influenced more by personalized messages emphasizing the uniqueness of their contributions and by the messages that state a clear goal (e.g. number of movies the user should rate). This approach, however, is questionable as a long-term solution because the effect of receiving email invitations will likely wear off with the time passing by. Yet, it is interesting that personalization in the messages is important and the messages setting specific goals are more persuasive than general appeals.

2.2 Strategies for motivating user participation in peer-to-peer systems

Peer-to-peer online communities are built on decentralized infrastructure. They need more user participation than client-server online systems because in peer-to-peer systems

there is no centralized server to ensure permanent availability of resources. To ensure quality of service, peer-to-peer systems rely on the redundancy of resources shared by the users in the system and this redundancy is ensured when people tend to share the files that they have downloaded from others and when they tend to stay online (Vassileva 2002). Therefore, many peer-to-peer systems have their own strategies to motivate users to share both original and downloaded resources.

"Direct Connect" (www.neo-modus.com) and Limewire (www.limewire.com) are very similar in their user motivational strategies, which are in fact policies enforced on the users. The users are forced to share a minimum number of resources. If a user fails to meet the requirement, her access to the resources of the systems will be limited or completely denied. This policy created a lot of frustrations in the users and attempts to hack the systems. Most people were not willing to join the community since they were forced to contribute before receiving any benefits.

Mojo Nation (www.mojonation.net) attempted to introduce an electronic currency and micro-payments (Golle et al. 2001) to provide economic incentives for sharing resources. The users had to pay a small amount for each download and the resource providers received a micro-payment for each download. The assumption behind this approach is that if the payments are miniscule, the users will not notice that tiny amounts of money are extracted from their virtual or real accounts, while these payments would add up to something significant for the provider of the resource. However, the users do notice, since they are being asked to buy something. The act of buying anything, even if the price is very small, creates mental transaction costs (Shirky 2003), that is, the energy required to decide whether something is worth buying or not. Mental transaction costs create a level of inconvenience that cannot be removed simply by lowering the prices of goods.

KaZaA promotes users to participate and contribute by rewarding the active users with better quality of services. The system records the actions of users and maintains a numeric participation level for each user, which determines the speed of downloads the user can get. How the participation level of a user is computed is not revealed but it seems to be a function of the difference between how much data (in MegaBytes) other users have downloaded from her and how much she has downloaded from others.

Actually, this strategy is based on reciprocation. Since the user's benefit (or privilege) is associated with the amount of resource other users download from her, this strategy can motivate users to stay online and make more contributions and even more, compel them to be concerned about the quality and the potential demand of the resources they share because obviously there is no reward for sharing resources that nobody downloads. However, the strategy can result in the following situation. A user who is sharing many files that are not of common interest may receive a low participation level and a low-quality service, since relatively few people, or maybe no one at all will download these files. Thus users who share rare files for a specific narrow interest area are disadvantaged. They may withdraw and the diversity of resources shared in the system will decrease.

2.3 Discussion

Many existing motivational mechanisms are based on the idea of rewarding users' participation and contributions with certain benefits. For example, KaZaA provides better services for active users, Slashdot reward users with some privileges, and Mojo Nation uses economic incentives to motivate users to contribute. The basic idea of rewarding user contributions is quite promising however the ways these systems measure the users' contributions have shortcomings. In Mojo Nation and KaZaA, the value of user contributions is measured depending on the times these contributions are downloaded. The contributions that are seldom downloaded will not bring benefit to the contributors. This may stimulate all users to contribute popular resources of common interest, which will decrease the variety of the resources in the community. A fairer approach to measure user contributions should take into account contributions that are not in big demand but ensure diversity in the resource pool of the community. The Slashdot mechanism measures user participation by the quality of contributed postings measured by the number of user ratings. This means that users can receive rewards only if the resources (postings) they shared are rated. However, even good resources may not attract enough attention to be rated, especially if they are shared relatively late. If user rewards are based exclusively on ratings, the mechanism may be not fair enough. In

order to use effectively rewards to motivate users to contribute, the approach to measure the users' contributions needs to be more comprehensive and considerate so that all contributions are fairly rewarded.

The idea of building user reputation in the online community (Preece 2001, Johnson 2001) is an effective way to enhance users' responsibility. Users usually use aliases in the online community and are hard to identify. Keeping a representation of their reputation can make them more concerned about their behaviors in the community and encourage them to contribute. Of course, reputation alone can not motivate everyone to contribute. For many people good reputation should be linked to some real-world rewards. Yet there is evidence that reputation-ranking can be a strong incentive for many people, since it allows for social comparison and competition. Erickson's suggestion of visualizing users and their activities is a good idea to increase users' awareness of other users and the whole community, especially for medium or small scaled systems. Besides, if all the users are visualized differently in a common place depending on their individual participation and contribution levels, it is possible to arouse social comparison and competition between the users. In this way, not only the user will be aware of the presence of other users, but also they will be triggered to compete with others by participating and contributing more actively.

2.4 Summary

This chapter reviews related research on motivating user participation in online communities. It can be seen that most of the existing motivational mechanisms have some limitations. But some ideas are quite valuable, for example, KaZaA's strategy of rewarding active users with better services and Slashdot's mechanism of building user reputation for contributing high quality resources. In next chapter, a new incentive mechanism is proposed, based on hierarchical memberships, which is inspired by some of the ideas discussed in this chapter.

Chapter 3

The Mechanism of Hierarchical Memberships

In this section, an incentive mechanism is proposed, which is different from the ones used in the existing online communities. Although this mechanism is devised for one specific system, Comtella, it can be applied to other online communities. In this chapter I discuss the general mechanism; the next chapter discusses the Comtella system and the implementation and evaluation of the mechanism.

3.1 Overview of the solution

The basic idea is to introduce a set of hierarchical memberships into the community. For example, users are given different memberships, such as “gold”, “silver”, “bronze”, etc., depending on their contributions to the system. The more the user’s contributions, the higher is her membership level. The users with higher-level membership receive better services and enjoy some privileges or special rights. Rewarding active users is the substance of the mechanism and it is coherent with two persuasion theories from social psychology¹. Several questions need to be decided in order to implement the mechanism: how to measure the user contributions, how to classify users in different membership classes, how to update users’ memberships and how to reward users of a given membership class. The way these questions are decided would influence the resulting motivational effect. These questions are discussed in the following sections.

¹ A discussion of the socio-psychological foundations of the hierarchical membership mechanism can be found in Appendix A.

3.2 Measuring the users' contribution

It is desirable that the users in an online community engage in certain cooperative activities that are beneficial for the community, for example, contributing original resources, evaluating the resources in the system, staying online (for peer-to-peer communities), etc. For different online communities, the desired activities may be different.

When the user is evaluated in terms of her contributions, all her activities that are beneficial for the community should be taken into account. These activities will be called “cooperative” activities from now on. Therefore, the approach is the following. First, a numeric value should be maintained for each cooperative activity ($V_i | i=1,2\dots N$) to represent the user's performance on the activity. For example, if staying online is considered a cooperative activity, the value may account for how long the user stays online per time unit (e.g. one week). Since the importance of the cooperative activities to the community may be different, a different weight is introduced for each cooperative activity ($W_i | i=1,2\dots N$). Then the overall evaluation of the user's contributions “ V_{oe} ” can be calculated through Formula (1).

$$V_{oe} = \sum_{i=1}^N W_i * V_i \quad (1)$$

3.3 Determining the users' memberships

First of all, it has to be decided how many membership levels are desirable for the community. If the number of the levels is too small, the users with different participation levels would not be differentiated well. When the user realizes that people who hold the same membership did not contribute as much as she did, she will feel treated unfairly and will be discouraged from participating. On the other hand, too many membership levels would complicate the hierarchy, hereby bringing difficulties for the users in figuring out which level is higher and more desirable. The exact number of the memberships could be different in particular systems, depending on the size (or expected size) of the community and the differentiation of rewards or privileges that can be

assigned to the members at different levels. Aeroplan®, Air Canada's frequent flyer branch devised four different memberships to distinguish their customers: regular, prestige, elite, and super elite. The Indian caste system (Callaham and Pavich 1998) has five levels. Generally, to ensure that people will not be confused by the hierarchy of the memberships, the number of the membership levels should not be greater than six; to distinguish the users with different participation levels, the number should not be less than three. In Comtella, the users are ranked into three levels of membership.

An important question is how large each membership group should be. In general, the number of the users in the highest membership group should be relatively small, because the users in this group, representing the highest participation level in the community, is not easy to be stimulated further. According to the theory of discrete emotions, their motivation would be to try to maintain their memberships.

Most of the users should be classified into the middle levels. The users in these levels have already progressed to the higher level and have reaped some rewards. On one side they may want more rewards since they have already seen that it is achievable to gain a higher membership and know how to do it. Also, according to the theory of discrete emotions, these users will have fear of losing their gained higher memberships and the associated rewards. Therefore, they are motivated from two sides. Both of the two possibilities could become their motivation to make more contributions to the community. In addition, according to the social validation theory, the fact that most of the users in the community are in or above the middle levels would bring pressure and stimulation for the inactive users with low membership.

One way to classify the users into different levels is to fix the proportion of users in each level. This is the approach I used in Comtella. Let's assume that there are three levels of membership, e.g. gold, silver and bronze, and the percentages of the users in these levels are 10%, 60% and 30% respectively. First, the users are sorted by V_{oe} (the evaluation of contributions) in decreasing order. Then the top 10% of them are put on the first level; the middle 60% and the bottom 30% of the users are put in the second and third level respectively. In this way, the proportion of the users in each level is easily controlled to accord with the guidelines. But the user's membership depends not only on her participation level but also on the participation levels of others. It is possible that a

user who contributes more than before will not obtain the higher membership level she desires just because the evaluation of contributions (V_{oe}) of many other users has also increased. This may frustrate the user. So the proportions need to be adjusted in terms of the participation level of the whole community. For example, if most users' V_{oe} increase, then the proportions of the high-level members should become larger.

Another possible way to classify the users is through setting a fixed threshold for each membership level. The user can obtain the membership as long as her participation level reaches the corresponding threshold. The merit of this approach is that it has a strong motivational effect before the users achieve the highest membership levels because upgrading memberships are achieved only through contributing more and participating more actively. However, it is hard to control the number of the users in each membership level. If most of the users in the community are in the highest level, the motivational effect would become very weak. The readjustment of the fixed thresholds should be performed very cautiously. If the users' memberships are degraded due to the readjustment of the thresholds, they would become very upset and begin to question the fairness of the mechanism.

3.4 Updating the user's memberships

Basically, there are two ways to update the user's memberships. One is to check whether the user qualifies for the next higher membership whenever she performs any of the cooperative actions and to upgrade her membership once she meets the requirement. The other way is to update all users' memberships at regular intervals, for example every week or every month. Obviously the first way is applicable only if the membership levels are decided through fixed thresholds, that is, the user's membership only depends on her own contributions and participation. The second way can work with both of the two membership determining mechanisms but the users can not see the change in their membership levels immediately, which may weaken the motivational effect. The frequency of membership updating should depend on how active the users' participation is or is expected to be. For example, if the users are not expected to participate actively, we can update the memberships monthly or every few months. In educational context,

the users are expected to be involved and contribute every week since usually a different topic is discussed every week according to the curriculum. In this case, weekly updating is a better choice.

3.5 Rewards for active users

When a user has managed to upgrade her membership, it is very important to have her realize that the system does reward her participation and contributions. Otherwise, the user would feel that obtaining a high level in the membership hierarchy is useless and may stop participating and contributing. What should be the reward for the cooperative users? First, the membership itself is a sort of reward, since users' memberships are public in the community. It is desirable to show the membership of each user in community visualization (Bretzke & Vassileva 2003). A special way of appearance (e.g. different colors) can serve as a kind of status recognition from the community, and high level members would gain a status of celebrities of the community. However, offering only this reward is not enough since not all users are motivated by social comparison. Better services need to be provided for active users. For example, in Comtella, some additional functions that facilitate the search for resources are used as the rewards for the users with high level memberships. The definition of a "better service" may be different in different systems. Generally, what is offered as better services should be what users really need while using the system. The reward should deserve or outweigh the users' effort to upgrade their memberships.

3.6 Summary

This chapter describes the proposed mechanism of hierarchical memberships, including the answers to several key questions about the detail design of the mechanism. In the next chapter, an overview of the Comtella system is given. Then, how the proposed mechanism was implemented and evaluated in Comtella is presented and the results of the evaluation are discussed.

Chapter 4

Evaluation of the Hierarchical Membership Mechanism

To setup an evaluation environment, the mechanism of hierarchical memberships was implemented in the Comtella system. A ten-week study to evaluate its effectiveness was carried out while the system was used in the course on Computer Ethics and Information Technology offered by the Department of Computer Science in 2003-2004 winter session. This section first gives an overview of the Comtella system and presents the modifications needed to make it support the class. Then it describes how the proposed mechanism is implemented based on Comtella and how the evaluation was carried out. Finally, the results of the evaluation are presented and discussed.

4.1 Overview of Comtella

Developed at the MADMUC Lab at the University of Saskatchewan, Comtella is an online system that enables a community of researchers and students to share and exchange Web-based articles or learning resources. Due to the copyright issue, the users are allowed to share only the URL of Web resources instead of the actual files. When they share a resource, they need to specify a topic that the resource is most related to. The topics in the system are defined in advance. The users can search for the resources they want by topic and keyword. Besides, they can rate and add text comments to the resources that they shared or download from others.

Originally, the system had a pure peer-to-peer infrastructure that was based on the Gnutella Protocol v6.0 (Klingberg & Manfredi 2002). To use Comtella, the users had to download a small client application. The client was based on an open-source Gnutella client in Java, called Jtella. The interface of the client was implemented using the Java Swing Graphic User Interface library and the communications between the distributed

components of the system were based on Java Socket technology.

However, as shown in a previous experiment with Comtella, the entirely distributed peer-to-peer infrastructure was not suited to support a small-scale community, since the availability of resources depends critically on whether there are enough users staying online. University students can not keep their computers and clients running all the time and therefore for most of the time the shared resources were unavailable to others. To solve this problem and allow a reliable service supporting required coursework, a modified version of the system was set up, in which the part of the user client program that forwarded and answered queries to other clients and kept the catalogue of all shared resources, as well as the resources themselves, was moved from the user's local machine to a central server where it was kept running all the time. The users kept on their own machine only the part of the client that provided the user interface and at login connected to the part of the client residing on the server. In this way, even though the user may be not online, her resources were always available for others in the community. So Comtella was no longer a pure peer-to-peer system but a hybrid of a peer-to-peer and a client-server system.

Another modification to Comtella was introduced to follow the dynamics of the course focus. According to the curriculum of most university classes, there is a different topic discussed every week. The students are expected to share Web-based articles or learning resources related to the topic discussed in each week. To support this rhythm of activities, Comtella should stimulate timeliness for user contributions, so that in any week the users are motivated to contribute on the current topic. This is different from general online communities where contributions in any topic/category are appreciated at any time.

Despite this, Comtella bears the common features of general online communities. It needs user participation to provide good service and has the feedback loop of user participation and system usefulness. A motivational mechanism is also needed in Comtella to ensure a thriving and sustainable community.

4.2 Implementing the mechanism in Comtella

To apply the proposed motivational mechanism to the Comtella system, a set of three memberships are introduced into the community: gold, silver and bronze. All users are grouped into these three levels according to their contributions to the community. Different but analogous graphic user interfaces are provided to the users with different memberships. Figure 1 shows the interface for the gold members of the community.

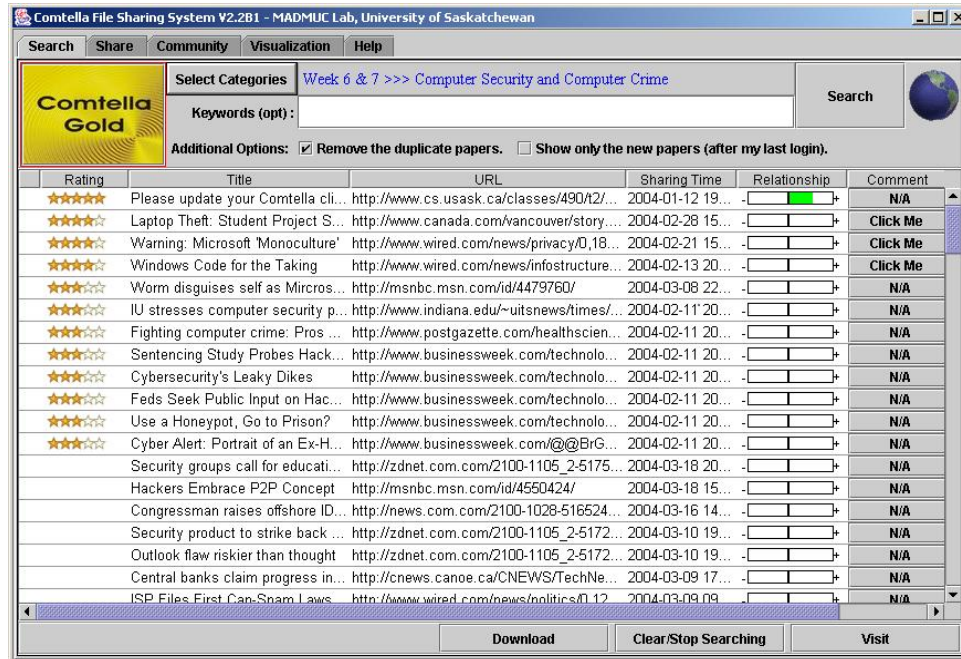


Figure 1. The user interface for the gold members

On the search panel of the interface (default panel when a user logs on the system) a symbolic membership card is displayed (see upper left corner in Figure 1), which shows the user's current membership level. If the user clicks on the card, a new window would pop up and show the user's participation and contributions, based on which her current membership level is computed (Figure 2). The window shows the user's contribution as a proportion of the contribution of the top user in each cooperative activity. I chose to show a proportion instead of the absolute value of the contribution to facilitate further social comparison and to help the user understand which activity she needs to emphasize.

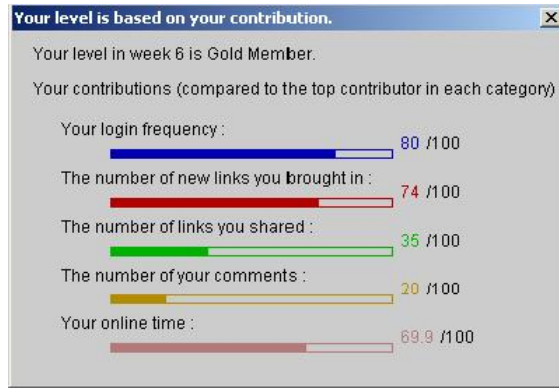


Figure 2. The window showing the user's contributions

In Comtella, it is desirable that the users engage in the following five cooperative activities:

1. Stay online;
2. Log on the system frequently;
3. Download resources and share them with others;
4. Bring new resources into the system;
5. Comment on the resources they have experienced.

Therefore, when the user's contributions are evaluated, her performance in all these activities is taken into account. The user's comprehensive evaluation is calculated using Formula (1) in Section 3.2. According to the relative importance of the five cooperative activities in the class context, I set the values of the five weights ($W_i | i=1,2,3,4,5$) as shown in Table 1.

Table 1. Different weights for five cooperative activities

i	Cooperative activities	W_i
1	Stay online	1
2	Log on the system	1
3	Download and re-share resources	0.5
4	Bring new resources	5.5
5	Comment on resources	2

After the introduction of a central server, the availability of the shared resources is guaranteed even though the providers are offline. Therefore, staying online has become less significant for the Comtella community than it used to be in the pure peer-to-peer version of the system and is given relatively small weight. Similarly, downloading

resources and re-sharing the copies is not so important, because increasing the number of duplicates of the resources does not increase the availability of the resources; on the contrary, it requires more space on the server. Logging on the system is a basic level of participation, so it has low weight. Sharing new resources is given high weight because it is very important to have new articles for Comtella to provide good service. Besides, the users were encouraged to comment on the articles they have read, so that these comments may give other users some recommendation for choosing articles. Other cooperative activities, like accessing/reading resources or rating resources were not considered as important and were not included in the comprehensive evaluation of the user contributions.

The users' memberships were updated weekly based on their participation and contributions in the pervious week. Since the way to classify users into different levels is based on their relative places instead of absolute values of their participation and contributions, their memberships have to be updated simultaneously on a fixed period of time. Weekly-based updating is a good choice because the users are expected to be involved and contribute every week.

In Comtella, the users' memberships are public inside the community. If the user switches to the visualization panel, the system would show a hierarchical representation of all the users' IDs together with their memberships (Figure 3). This representation is supposed to trigger social comparison and thus stimulate the user to contribute more to measure up with her peers.

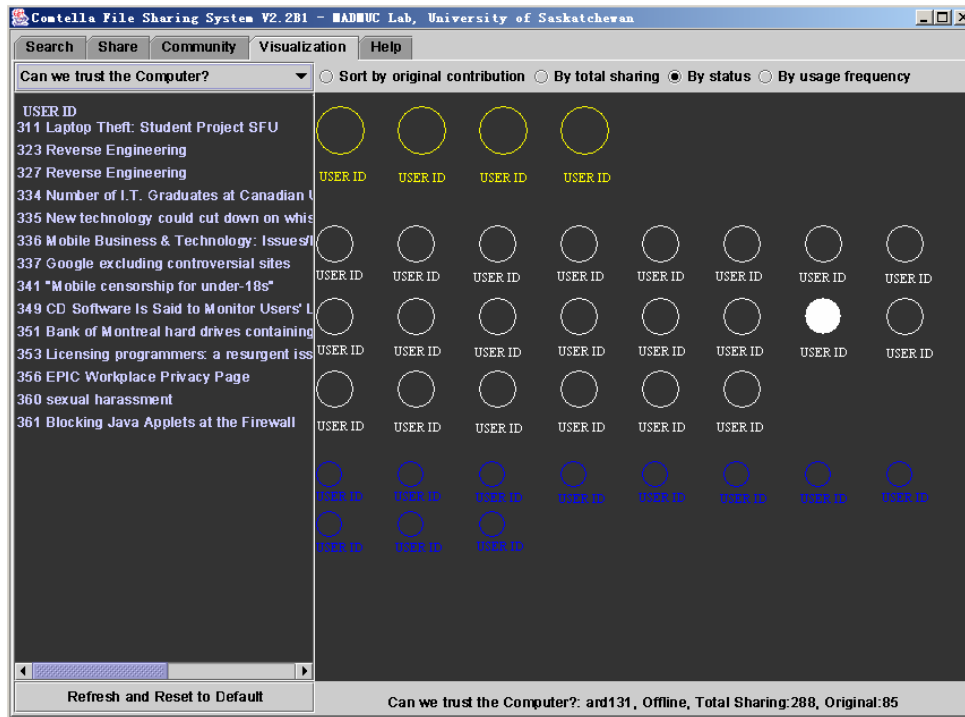


Figure 3. The visualization panel (User identification is removed to preserve privacy; a filled circle indicates the user is currently online.)

In the previous version of Comtella, it was difficult for users to find the target resources because the search results are usually numerous and there are many duplicate resources which are returned from different peers. Therefore, some useful extra-functions are provided for the users in each level to facilitate their search for resources. And the users with higher-level membership are rewarded with more of these extra-functions (Table 2).

Table 2. Rewards for the users with different memberships

Users	Addition Functions
Bronze member	<ul style="list-style-type: none"> ▪ Sort search results by resource title
Silver member	<ul style="list-style-type: none"> ▪ Sort search results by resource title ▪ Remove duplicate results
Gold member	<ul style="list-style-type: none"> ▪ Sort search results by resource title ▪ Remove duplicate results ▪ Show only new resources, shared after the user's last logon ▪ Sort search results by rating, share time, or provider

4.3 Evaluation

- **Introduction**

After the Comtella system was updated, a case study was launched in a course on Computer Ethics and Information Technology offered in 2003-2004 winter session. The students were encouraged to use Comtella to share URLs to the web articles related to the topics of the course. According to the curriculum, a new class topic was introduced into the community at the beginning of each week and during this week the articles shared by the students should be related to the topic. 5% of the final mark of the course was based on the number of URLs the student shared in Comtella. The goal of the study is to evaluate the effect of hierarchical membership mechanism on motivating users to participate and contribute. The study aimed to answer the following questions.

- Did the strategy succeed in stimulating the users to do the five cooperative activities?
- To what extent were the users stimulated?
- Did the users really care about their membership levels?
- Were the additional (reward) functions really useful? How often were they used?

- **Participants**

Thirty-five (35) fourth-year students of the Department of Computer Science were invited to use the Comtella system while taking the course.

- **Methods**

The study took ten weeks. To isolate the effect of the mechanism, in the first six weeks, the proposed motivational mechanism was not applied, so that the users used only the basic Comtella functionality. This provided us with a baseline for comparison of the contribution levels. At the beginning of the 7th week the proposed mechanism with hierarchical memberships was introduced. The users' memberships were updated weekly, depending on their participation and contribution level in the previous week.

To collect the data to answer the above four questions, the client and the server of Comtella were programmed to trace the users' actions in the system and report the data to a central database. The following data were recorded during the study:

- Whenever the user logged on the system, the logon time, logout time, and the user's identification were recorded.
- Whenever the user shared or downloaded a resource, the action, the resource and the user's identification as well as the time were recorded.
- Whenever the user clicked on her membership card, the action and the user's identification as well as the time were recorded.
- Whenever the user used the additional functions, the time, the function being used and the user's identification were recorded.
- All the users' actions on the visualization panel were recorded.

In addition, I used a post-experiment questionnaire² to collect more information from the users at the end of the study.

- **Results**

Through analyzing the raw data, I obtained the following results.

- 1. Users' contributions and participation increased.***

According to the statistics, the users' contributions and participation increased from the 7th week of the study, the week when the hierarchical memberships were introduced into the community. Figure 4 shows the sum of all the articles shared by all the users for each week. It is evident that the number of new articles contributed in Week 7, 8 and 9 is greater than those in the first six weeks. Although there is a decrease in Week 10, the average of the numbers in the last four weeks is greater than that in the first six weeks (158.8 vs. 55.3). I also calculated for each user the average number of resources she shared in the first six weeks and in the last four weeks and used Wilcoxon Signed-Rank test (Lowry 1999) to verify the hypothesis that the users contributed more articles in the

² The questionnaire and the results can be found in Appendix B

last four weeks than they did in the first six weeks³. The results of the test showed that the z-ratio was equal to 3.954, which confirmed the hypothesis with statistical significance beyond the 0.0005 level. Besides, the number of users' comments increased as well from the 7th week. On average the users contributed 168.3 comments in each of the first six weeks but they contributed 372.0 comments in each of the last four weeks.

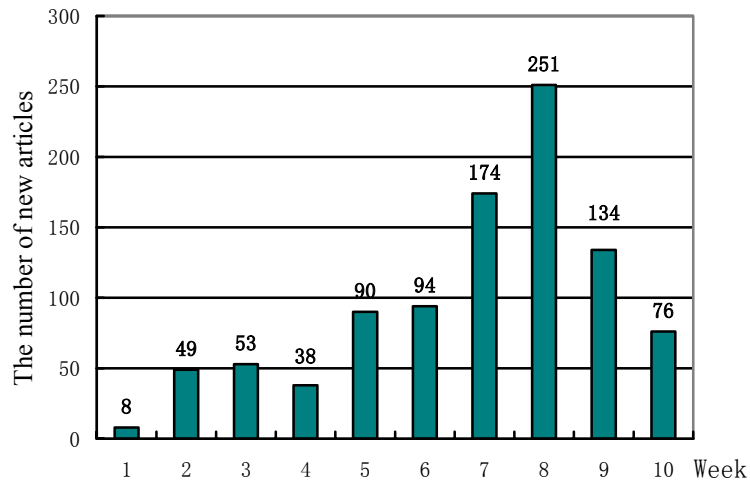


Figure 4. The total number of articles shared in each week

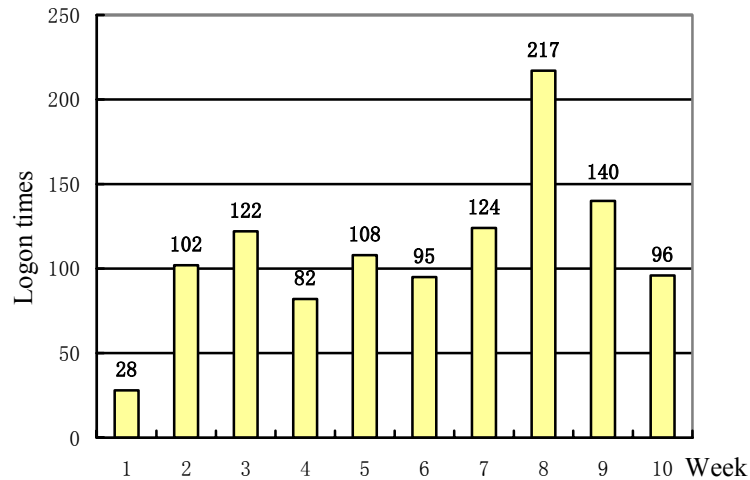


Figure 5. Total logon times of all users in each week

The same effect was observed with the logon actions. As can be seen in Figure 5, the

³ The detailed process of the Wilcoxon Signed-Rank test is described in Appendix C

total number of times of the users logging on the system per week increased from the 7th week. The z-ratio of Wilcoxon Signed-Rank test based on each user's average number of times of logon in first six weeks and in last four weeks was equal to 2.812, which proved that users logged on the system more frequently in the last four weeks than they did in the first six weeks with statistical significance beyond the 0.005 level.

Nevertheless, both the number of user contributions and logon times decreased in the last week of the study. One reason for this may be the heavier workload of the students at the end of the term, which didn't leave them enough time to dedicate to Comtella. It is also possible that the increase in participation is just the effect of the novelty of the mechanism. This question will be discussed further in the next section.

2. Nearly half of the users checked weekly the evaluation of their contributions.

According to the records from the database, 77% of the users (27 users) clicked their membership cards at least once to check the details about their levels of participation in comparison with the top contributor in each aspect. On average the membership card was clicked 10.8 times per user since it was introduced in Week 7. Besides, the records showed 15 users (43%) clicked the cards every week to check their participation and contribution levels. Although the users might click their cards by mistake, the probability is small. Most of those users who clicked their cards did this to check the evaluation of their contributions.

3. The users who paid more attention to their membership status were more inclined to upgrade their membership intentionally.

This pattern was found by analyzing the results of the questionnaires and the records from the database. Of all the participants, 57% (20 users) indicated that they have tried to upgrade their memberships. However, of those who had clicked their membership card to check their participation levels in each of the four weeks (15 users), 93% (14 users) said that they tried to upgrade their memberships. I also noticed that the users who never clicked their membership cards (8 users) did not attempt to earn higher-level membership. These users were not motivated by the proposed strategy. Obviously, there was a correlation between the efforts the users made to upgrade their memberships and the regularity with which they checked their participation measures.

4. Some of the extra (reward) functions were used frequently, others were not.

Six extra functions were provided for the different membership levels to reward active users with better services. From the data collected in the last four weeks, I found that some of these functions were often used by the users, but some were not. For each of the functions, I calculated the following metrics to evaluate its usefulness:

- The number of the users who were eligible to use the function (NE);
- The number of the users who used the function (NU);
- The percentage of the users who used the function ($P=NU / NE$);
- The total number of times the function was used (T);
- The average number of times the function was used per user ($A=T / NE$).

The data about how the extra functions were used are listed in Table 3. It can be seen that sorting search results by title and removing duplicate results were the functions used most by the participants. Sorting by provider and sorting by rank were used as well and some participants used them frequently. However, the other two functions were used seldom. The function of showing only new resources was used less than once by each user on average. These functions are probably not what users really need. From the results, I concluded that the reward functions provided for gold and silver members are not attractive enough. More useful functions are needed to stimulate users to upgrade their membership levels.

Table 3. The use of additional functions

Addition functions	The members who can use it	NE	NU	T	P	A
Remove Duplicate Resources	Gold; Silver.	32	20	167	62.5%	5.22
Show only new resources	Gold.	13	5	12	38.5%	0.92
Sort by rank	Gold.	13	4	37	30.8%	2.85
Sort by title	Gold; Silver, Bronze	35	15	343	42.9%	9.80
Sort by share time	Gold.	13	3	13	23.1%	1.00
Sort by provider	Gold.	13	8	34	61.5%	2.62

5. The quality of users' contributions declined.

With the increase in the quantity of users' contributions, the quality of the

contributions decreased. There were four users who shared a lot of articles that were not related to the topics of the class. They were obviously trying to raise their participation scores and gain a higher membership level or maintain their gold level. Besides, since the moment the hierarchical memberships were introduced, the quantity of the users' comments increased, but the length of the comments became shorter on average. Some students even plagiarized (by copying and pasting) others' comments. The increased number and lower quantity of the contributions was noticed by the users, which led to some complains in the questionnaire that some of their colleagues are trying to cheat the system.

6. *Different users have different contribution patterns.*

I found that different users contributed in different ways. Some contributed many, but average or poor-quality resources, while some contributed few, but very good ones. While it is unknown whether a particular contribution pattern is associated with some personal characteristic or understanding based on previous experience with online communities, it is interesting to see this result. In order to improve the overall quality of the contributions in the community, it seems better to have an adaptive incentive mechanism that can discourage the contributions from the users who do not regard quality while encouraging the contributions from the users that tend to share high quality articles.

4.4 Discussion

The results of evaluation demonstrate that the proposed strategy is capable of motivating users to participate and make contributions in an online community. After the hierarchical memberships were introduced in the Comtella community, the users began participating in the system more actively and contributing more resources than before. The questionnaires showed that 57% of the users were successfully stimulated to participate and contribute. Besides, most of the users were concerned about their membership levels. In the process of the study, some students even requested that the mechanism used to determine the memberships was made public, so that they can optimize their strategies to earn higher level memberships.

Nevertheless, the results also show that there was a decrease in several forms of user participation in the last week of the study. There are three possible reasons for this. Firstly, it may be because the students had heavier workload in the last week of the class due to project deadlines and incoming final exams and did not have as much time as before to participate in the system. Another possible reason is that the increase in contributions and participation since the 7th week is due to the novelty effect of the mechanism and after about three weeks, this effect faded away so that the users' contributions and participation returned to the levels before. Therefore, a longer term evaluation is necessary to test whether the proposed mechanism has persistent motivational effect on users. The third possible reason is that user participation was discouraged by the deteriorating quality and the large quantity of the contributions in the system towards the end of the study.

The results have shown that after the mechanism was applied in Comtella, the users' contributions increased in quantity but somewhat declined in quality. This made it hard for the users to find good articles in the system, resulting in their disappointment reflected in the negative comments in the post-experiment questionnaire. A similar phenomenon called "information overload" (Shenk 1997) has been observed in other online communities. It is typically a result of an influx of many new contributions with medium or low quality, which makes the users feel swamped by a mass of unwanted information. Jones and Rafaeli (1999) found that the users' most common response to it is to reduce or end their participation in the community, both as contributors and as consumers. Therefore, the decreased level of participation in the last week could be due to the information overload emerging in the system although the two other possible reasons can not be excluded yet. In order to guarantee the active and stable user participation in the online community, it is necessary to avoid the information overload by controlling the quantity of user contributions in the system, motivate users to contribute high-quality resources and inhibit the inferior contributions.

4.5 Summary

This chapter presents the evaluation of the hierarchical membership mechanism. The

results of the evaluation are presented and discussed. The study showed that the mechanism was able to motivate more than half of the users to participate but it resulted in a deteriorating quality of user contributions and information overload in the system. This impaired the sustainability of the online community. Therefore, the motivational mechanism needs to be improved so that it encourages only high-quality user contributions and inhibit inferior ones.

Chapter 5

The Improved Incentive Mechanism

5.1 Approaches to decentralized quality control

To ensure active and sustained user participation and avoid information overload in the online community, the incentive mechanism has to take into account the quality of user contributions, i.e. to reward the contributions with high quality, inhibit inferior ones and restrict the total number of contributions.

It is not easy to measure the quality of each contribution impartially and accurately because quality measures are mostly subjective. Centralized moderation is feasible only for small and narrowly focused communities, where members have very similar evaluation criteria. For medium or large online communities (e.g. the ones with more than 100 users), a decentralized moderation for quality measurement is necessary. A real world example of decentralized moderation is the impact factor which measures the quality of journals or papers by counting the times they were cited. Although it is somewhat controversial whether the impact factor is able to represent fairly the quality of research papers (e.g. Merton & Zuckerman 1968), it indicates the extent to which the paper is used by other scholars. In a similar way, one can measure the quality of a posting in an online community by counting the times it was viewed (clicked). However, this method is based on the assumption that people who view a resource hold a positive attitude to its quality, which is not always the case.

Another way of evaluating the quality of resources or comments is through explicit user ratings (e.g. the peer-reviewing process in academia, online communities like Slashdot, etc.). Since the final evaluations of resources are computed based on ratings from many users, they are more unbiased. However, a study of the Slashdot rating mechanism showed that some deserving comments may receive insufficient attention

and end up with an unfair score, especially the ones with lower initial rating and those contributed late in the discussion (Lampe & Resnick 2004). The same effect appears in the majority of decentralized quality measurement systems and it is sometimes referred to the “rich get richer” or “Matthew effect” (Merton & Zukerman, 1968). It is impossible to prevent this effect and ensure a totally fair mechanism in a decentralized rating system since the postings with high ratings become more visible and tend to be read and rated more often. However, knowledge of this effect can help an individual to develop strategies to achieve a better standing. Obviously, the timeliness of contributing resources is important and users should be encouraged to contribute early because late contributions are unable to receive enough attentions and less useful for the community. This is especially relevant in a class-support system like Comtella or I-Help (Greer et al. 2001) since the topics typically change on a weekly basis and late contributions tend to be neglected. Apart from playing a role in the likelihood that a contribution will be read and rated, the timeliness of the contributions reflects the needs of the community. Early in the discussion period, when there are not many contributions, it is important for the community to bring new contributions, so that there are enough materials to be read. Later, when the number of contributions is already high, it is more important for the community that the users rate the contributions, since in this way they provide guideline for finding the good contributions among the many. Therefore the rewards for different types of cooperative activities (e.g. sharing new resources; giving ratings) should be different depending on the time after introducing the new topic.

A challenge in the systems that rely on decentralized moderation is to ensure that there are enough user ratings. MovieLens’ approach of sending users email-invitations to encourage them to rate movies (Beenen et al. 2004) may be effective for some time. However, it is questionable as a long-term solution since the effect of receiving email-invitations will likely wear off with the time passing by. To stimulate users to rate resources constantly, persistent incentives are necessary.

The evaluation of the motivational mechanism showed that different users had different contribution patterns. Some contributed many, but poor-quality resources, while some contributed few, but high-quality ones. Therefore, the incentive mechanism should be adaptive to the patterns of contributions of different users. The mechanism

should stimulate the users contributing few high-quality resources to contribute more and it should inhibit contributions from users who contribute many low-quality resources unless they improve the quality of their contributions.

In addition, the mechanism should adapt to the total number of resources desirable in the online community. For example, the rewards for contributing new resources should be decreased if there are already too many resources in the community that could cause information overload. However, how to decide what amount of resources is desirable is an open question. In a context of a class-support community, the instructor may be able to decide about a desirable and possible number of contributions.

Based on the discussion above, an improved incentive mechanism was implemented in the Comtella system to encourage users to rate the resources. Depending on the quality evaluation from user ratings, the mechanism is able to adapt the rewards of different forms of participation for individual users depending on their current reputations and the current needs of the community. The goal is to influence the users' actions of contributing in terms of both sharing resources and rating resources so that they benefit the system most. The mechanism consists of two components: the collaborative rating mechanism and the adaptive reward mechanism.

5.2 Collaborative rating mechanism

The collaborative rating mechanism is inspired by the Slashdot moderation system. In order to have a broader source of ratings, all users can rate others' contributions. They can award any contribution +1 or -1 point as they like or dislike it, consuming one of their own points (Figure 6). However, the users with higher membership levels receive more points to give out, which means they are more influential in the community. To ensure that contributions initially have equal chance to be read and rated, the initial rating for every new contribution is zero regardless of its providers' membership level or the quality of her previous contributions (This is different from Slashdot, where the postings of the users with higher karma start at a higher rating.) In the end, the final rating for the contribution is the sum of all the ratings it has obtained. The summative

rating for each contribution is displayed in the list of search results, which can be sorted by the user and viewed as a “top 10” list of articles for any topic (Figure 6).

As a persistent incentive for users to rate contributions, a virtual currency is introduced, called “*c-point*”. Initially, each user only has a limited number of *c-points*. Whenever the user rates an article, she is awarded a certain number of *c-points*, depending on her reputation of giving high-quality ratings (discussed in Section 5.3). The users can use the earned *c-points* to increase the initial visibility of their postings in the search result list. Most users desire that their contributions appear in salient positions, e.g. in the first place or among the top 10, because in those positions they will have a better chance to be read and rated. The Comtella search facility displays all the contributions matching a query in a sorted list according to the number of *c-points* allocated by the contributors (Figure 6). Unlike Slashdot, the proposed mechanism allows the user the flexibility to invest any number of their *c-points* in a particular posting to increase its visibility. The *c-points* are rewarded immediately after users rate articles, bringing instant gratification, which could be a powerful incentive for some users. Since the users with higher membership levels can rate more articles, they are potentially able to earn more *c-points*.

Result:		<<Previous Next>>		Total: 5 Page		
Cpoint	Paper Title	Earned Ratings	My Rating	View Times	Fake?	Fak Cou
40+	PORNOGRAPHY: SOCIAL EXPRESSION OR SOCIAL DISEASE?	1	<input type="text" value=""/> Rate	7	Fake	0
30+	Google ? the only archive we'll ever need?	2	<input type="text" value=""/> Rate	8	Fake	0
20+	Technology & Happiness	4	<input type="text" value="1"/> Rate	12	Fake	0
20+	Video Games, Not TV, Linked to Obesity in Kids	4	<input type="text" value="-1"/> Rate	13	Fake	0
10+	Alzheimer's patients to trial MS labs life-blog gadget	3	<input type="text" value=""/> Rate	4	Fake	0
10+	Special Issues for Teens	2	<input type="text" value=""/> Rate	8	Fake	0

Figure 6. A segment of a search results list

5.3 Adaptive reward mechanism

The adaptive reward mechanism is introduced as another improvement of the mechanism of hierarchical memberships. The basic idea is to adapt the rewards of particular forms of participation for individual users and to display personalized

motivational messages for them depending on their current reputation and the current need of the community, thereby influencing and directing the users' behaviors of contributing. Figure 7 presents an overview of the mechanism. It uses two models: community model and individual model.

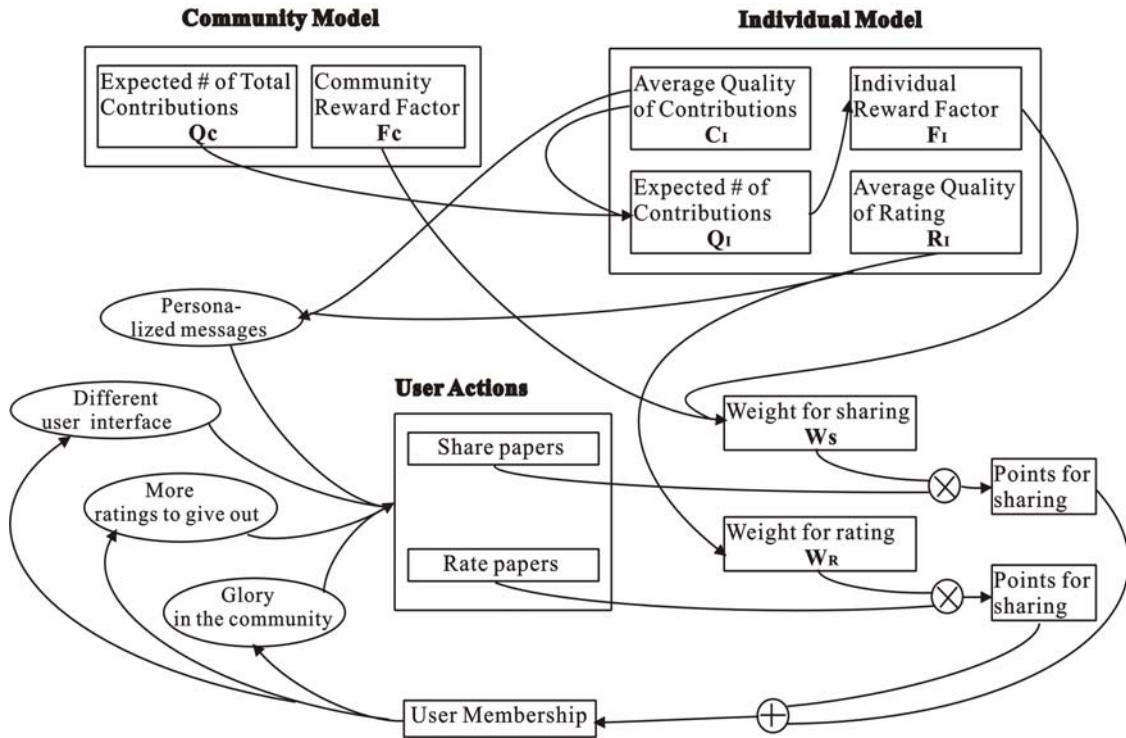


Figure 7. An overview of adaptive reward mechanism

5.3.1 Community model

The community model is used to describe the current phase of the whole community. It includes the expected sum of resources desired for current topic (Q_C) and the community reward factor (F_C). Q_C is used to restrict the overall number of resources in the system. For each week, when a new class topic is introduced, Q_C is set by the community administrator (e.g. the instructor of the course) for the new topic, depending on her knowledge of certain features of the topic (e.g. how interesting it is expected to be for the users, the relative amount of materials available online that could be shared) and the users' situation (e.g. how much time and energy they can devote, depending on their coursework, exams, etc.). F_C reflects the extent to which newly contributed resources are useful for the whole community. Generally, new resources are needed as

soon as possible after a topic has been announced or opened and those contributed late in the period are less useful than the ones contributed early. Therefore, F_C has its maximum value when a new topic is introduced and decreases gradually with the time. After the middle of the discussion period, it decreases faster (Figure 8).

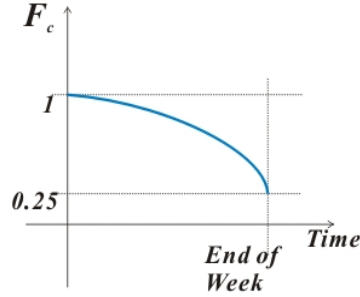


Figure 8. The changing of the community reward factor (F_C)

5.3.2 Individual model

Each user has an individual model that keeps her reputation of contributing high-quality resources and giving high-quality ratings and contains the data describing her current state. Since the summative rating of a resource denotes its quality, a user's reputation of contributing resources (C_I) is defined in a straightforward way as the average summative rating of all the resources she has shared so far.

However, the quality of ratings can not be defined so easily, since they are by nature subjective. Although in educational online system, instructors or teaching assistants may have the ability to evaluate the quality of ratings, the workload of evaluating all user ratings is overwhelming for one or even several persons. Moreover, in general online communities there are no such arbiters available. I chose to measure the quality of each rating by the difference between it and the average of all the ratings the resource gets. The smaller the difference, the higher the quality of the rating. Accordingly, a user's reputation of giving high-quality ratings (R_I) is defined as the reciprocal of the average difference between all ratings she has made and the respective average ratings. Formula 2 shows how R_I is calculated. Here $r_i (i=1,2,3 \dots N)$ are the ratings the users has made; and $\bar{r}_i (i=1,2,3 \dots N)$ are respective average ratings.

$$R_I = \frac{N}{\sum_{i=1}^N |r_i - \bar{r}|} \quad (2)$$

The approach is based on the assumption that the average rating of a resource can reflect the opinion of the majority and is less biased. Since this metric can be easily skewed if users intentionally rate close to the average rating of the resource, the average rating should not be shown directly to the user.

The expected number of resources contributed by the user (Q_I) is a fraction of Q_C . The users with higher C_I will get a larger Q_I , which means the users with a good reputation of contributing high quality resources are expected to share more while the expected total number of contributions from all users is fixed. If details are ignored, Formula (3) can demonstrate how Q_C is distributed among users.

$$Q_I \approx Q_C \cdot \frac{C_I}{\sum C_I} \quad (3)$$

Q_I : the number of resources expected from the user

Q_C : the expected sum of resources desired from all users

C_I : the user's reputation of contributing resources

The individual reward factor (F_I) defines the extent to which the resources contributed by the user are being rewarded. Generally, the contributions beyond the user's expected number (Q_I) should not be encouraged. So F_I is a function that is a constant value as long as the number of the user's contributions is less than or equal to her Q_I . When the number exceeds the expectation, F_I drops to one fourth of the constant value instantaneously and keeps decreasing with the increment of the users' contributions (Figure 9).

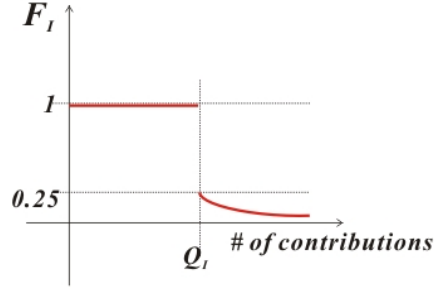


Figure 9. The changing of the individual reward factor (F_I)

5.3.3 How the mechanism works

In the adaptive reward mechanism, varying weights $W_i(t)$ for particular forms of contribution are applied to evaluate users' comprehensive participation and determine their membership levels, which are associated with different rewards and privileges. If we represent with $t=(1,2,3 \dots T_i)$ the sequence of the contributions in each kind (e.g. sharing resources, giving ratings), the overall evaluation of a user's contributions (V_{oe}) is calculated through Formula (4).

$$V_{oe} = \sum_{i=1}^n \left[\sum_{t=1}^{T_i} W_i(t) \right] \quad (4)$$

The weights are adaptable to the states of the user's individual model and the community model at the current time. The current values of the weights are shown to the user at logon time so that she can see what rewards she will receive for each type of the contributions. So are the personalized motivational messages, which inform the user of the number of resources expected from her for the current topic (Q_I) and, if the user's C_I (the reputation for sharing) or R_I (the reputation for rating) is lower than certain limit, remind her to pay attention to the quality of the resources or the ratings that she will contribute.

The adaptive weight for sharing resources (W_S) is calculated through Formula (5). Here W_{S0} is a constant, which is the initial value of the weight.

$$W_S = W_{S0} \cdot F_C \cdot F_I \quad (5)$$

F_C : the community reward factor

F_I : the individual reward factor

W_S is equal to W_{S0} , its initial constant value, when a new topic begins and the number of the user's contributions has not reached their expected value Q_I . After that, it decreases gently with time as F_C does (Figure 8). Whenever the number of the user's contributions goes beyond her Q_I , W_S sharply decreases to one fourth of its original value (as F_I does in Figure 9) and continues to decrease with the accumulation of the user's contributions and time.

It can be seen that W_S inherits the features of both reward factors, F_C and F_I . In this way, a user who shares many resources but does not care about their quality gets a low C_I (the user's reputation of contributing resources) and a small Q_I (the number of resources expected from her) and therefore, little reward for her subsequent contributions. Thus the personalized message to the user would be to contribute less new resources and try to improve their quality. This situation continues until the user finally improves her reputation in sharing. On the other hand, if a user tends to share a small number of good resources, she obtains a high C_I and a large Q_I , which encourages her to share more resources. And she will earn more rewards by sharing new resources unless she starts compromising the quality. For both kinds of users, early contributed resources always earn more points. Hence, the adaptive weight W_S is able to restrict the quantity of user contributions, elicit the contributions from the users with a good reputation, discourage the contributions from the users who ignore quality and stimulate users to share early in the discussion period.

The adaptive weight for giving ratings (W_R) is proportional to the user's reputation of giving high-quality ratings (R_I) (see Formula 6). The users who have gained a good rating reputation get higher weight for their subsequent ratings, which should stimulate them to rate more resources. However, those with low R_I will not get much reward for rating. They have to rate less and improve the quality of their ratings to win their reputation back and this would be the suggestion of the personalized message.

$$W_R = K \cdot R_I \quad (K \text{ is a constant}) \quad (6)$$

The adaptive weights are designed not only to influence users' patterns on particular forms of contributions but also to direct them to choose among several rewarded activities when they want to contribute to the community. For example, the difference between W_S and W_R may stimulate the user to make the type of contributions that are more valuable to the system at the time when she logs in. If the number of resources expected from her (Q_I) and her reputation of giving ratings (R_I) are not taken into account, Figure 10 shows the relation between W_S and W_R in one discussion period. In the first half of the period, there is a strong demand for new resources in the community since the topic is newly introduced. Therefore, sharing resources is rewarded higher at that time. In the second half of the period, when there are already many resources in the system, W_R becomes higher than W_S because more ratings are needed to help users to filter out low-quality resources.

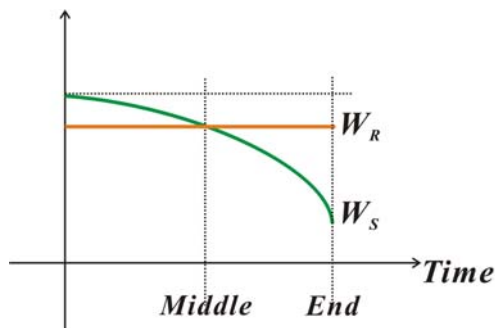


Figure 10. The relation between the adaptive weights for sharing (W_S) and rating (W_R)

5.4 Summary

This chapter presents the design of an improved incentive mechanism, the goal of which is to regulate the quality and the quantity of user contributions, decrease the level of information overload in the online community and make it more sustainable. The mechanism has the following properties:

- It provides a persistent incentive for the users to rate resources (the c-points);
- It adapts the rewards of particular forms of contribution for the users based on:
 - their individual reputation for these forms of contributions

- the needs of the community at the moment;

- It takes into account the quality of both contributed resources and ratings;
- It directs users' actions of contributing by adaptive rewards and personalized motivational messages.

To evaluate the effectiveness of the mechanism, a case study was launched. Thirty-one fourth-year students used the modified Comtella system again in the course on Ethics and Information Technology for nine weeks in its next offering in the spring of 2005. The next chapter presents the process and the results of the study.

Chapter 6

Evaluation of the Improved Incentive Mechanism

To test whether the mechanism discussed in the previous chapter can achieve the goals of regulating the quality and quantity of contributions and ensuring sustained participation in online communities, a second case study based on the Comtella system was carried out. Before the study, Comtella was redesigned and re-implemented to improve the system performance and to have all the new features of the mechanism. This section first introduces the reimplementation of the Comtella system. Then the process of the case study is presented and the results are discussed.

6.1 Re-implementing the Comtella system

The previous version of Comtella has some shortcomings with respect to reliability, performance and usability. Since the communications between the clients and the central server were based on Java Socket, the server could not handle a high volume of user requests. Some requests were simply ignored when the server got very busy. Access from off-campus computers was usually blocked by the network firewalls. This leads to complaints about some resources disappearing after being shared/uploaded. Besides, the users seemed not satisfied with the interface of the system. Their comments on the usability were mostly negative.

To improve the reliability, usability and scalability of the system, Comtella was redesigned and re-implemented as a Web-based system with the architecture shown in Figure 11. The new Comtella system has a pure client-server infrastructure. It runs on Apache Tomcat Server 5.0 and MySQL Database 4.0. The users are provided with a friendlier Web interface, so they do not need any client-side application to access the system. The communications between the clients and the Web server are based on the

standard HTTP protocol so that firewalls do not obstruct them. All these technologies are very mature and widely applied, so the performance and the reliability of the system are guaranteed.

The new Comtella was built using a Java Web-application development framework – WebWork (OpenSymphony 2004). It is a Java third-party library that can facilitate the development of Java Web applications. Following a mature framework to develop a Web application can enhance the readability of the code and the extensibility of the system. Among several existing frameworks, WebWork was chosen because of its clear modularized structure and its simplicity⁴.

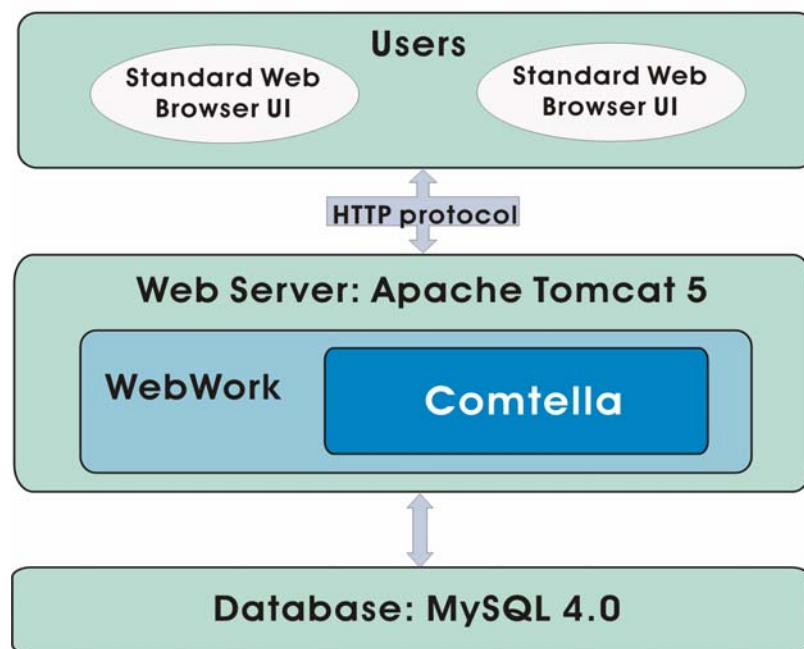


Figure 11. The architecture of the Web-based Comtella system

The new Comtella system had more powerful functionality than the previous one. In addition to the basic functions such as sharing articles, searching for articles, rating articles and commenting on articles, the new system allows the users to upload their summaries (e.g. Word or PDF files) for particular articles and to review and evaluate each other's summaries. Besides, a discussion forum is provided, which allows users to post opinions related to each of the topics of the class (one for each week). Although

⁴ More technical details about the Web-based Comtella system can be found in Appendix D.

these features are not relevant to the proposed incentive mechanism, they are useful to support community activities in educational context.

6.2 Implementing the mechanism in the new Comtella

The improved incentive mechanism discussed in Chapter 5 was implemented in the new Comtella system, which was used as a test environment for the evaluation of the mechanism. Figure 12 shows the new “welcome page” (the front page) of the system. On this page, the following information and features are provided for the user.

- The user’s contribution levels in the pervious week and in the current week. Similar to the old system, the different forms of user contributions (e.g. sharing, rating) are evaluated separately. The comprehensive evaluation is computed based on the user’s performance in each form of contribution and the weight for it. To enhance the motivational effect, the contribution levels in the current week are updated immediately after the contributions are made, so the user will see how her level increases after she makes a contribution.
- Weights for different forms of contributions. As discussed in Section 5.3, the weights for contributing new links and for giving ratings are updated constantly according to the user’s reputation and the needs of the community. The weights for the other activities are constant.
- The user’s current membership level. It is updated weekly, based on the comprehensive evaluation of the user’s contributions in the previous week.
- Personalized messages for the user. Personalized messages inform the user of the number of articles she should share for the current topic and also remind the user to pay attention to the quality of her contributions and ratings when necessary.
- Community news. To enhance the users’ awareness of the community and other members, the system administrator inform users about new events in community and provide course information such as deadlines for submitting assignments, time and place of examinations etc. Besides, gold members are also allowed to release community news, as one of their privileges.

- Top users and best papers of last week. The top users and the best papers of the previous week are announced in the welcome page, as another stimulation to contribute. The top users are those three or four users with the highest comprehensive level of contributions in the system. The best papers are the articles that received the most ratings and are shown in a list of five to seven together with the users who contributed them.

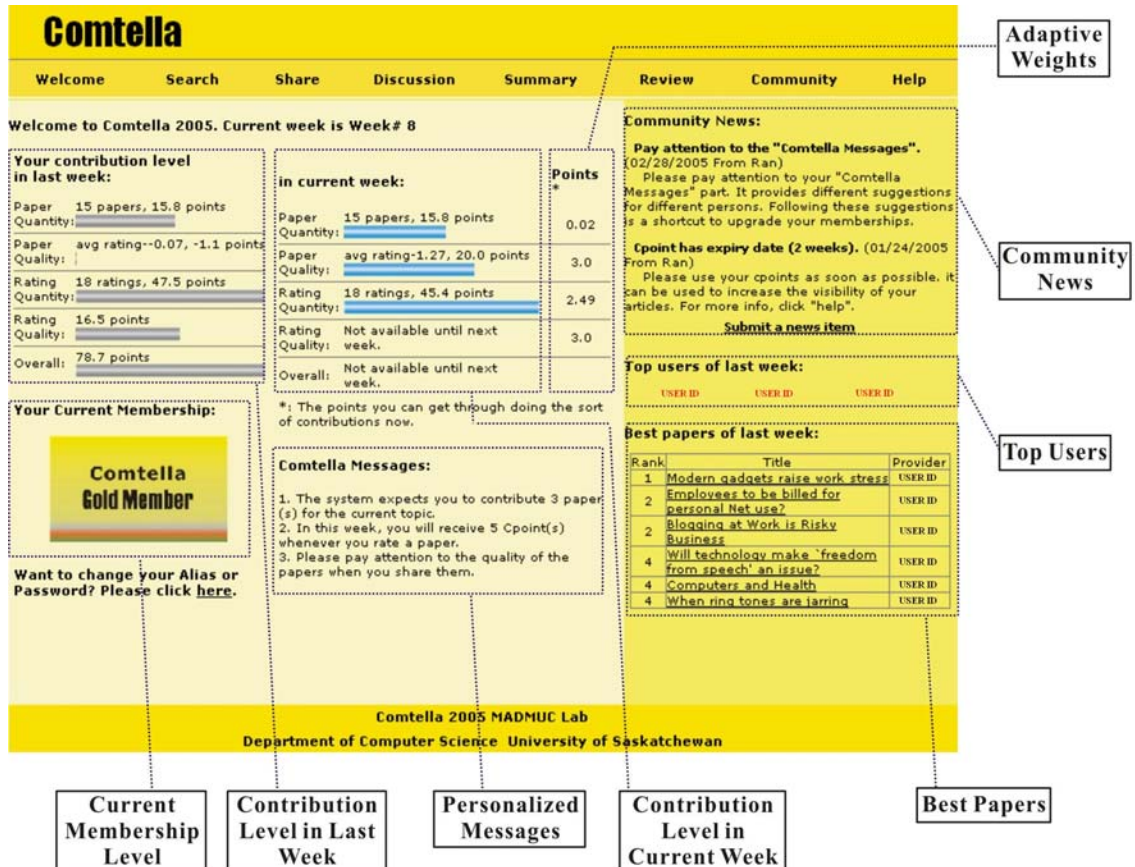
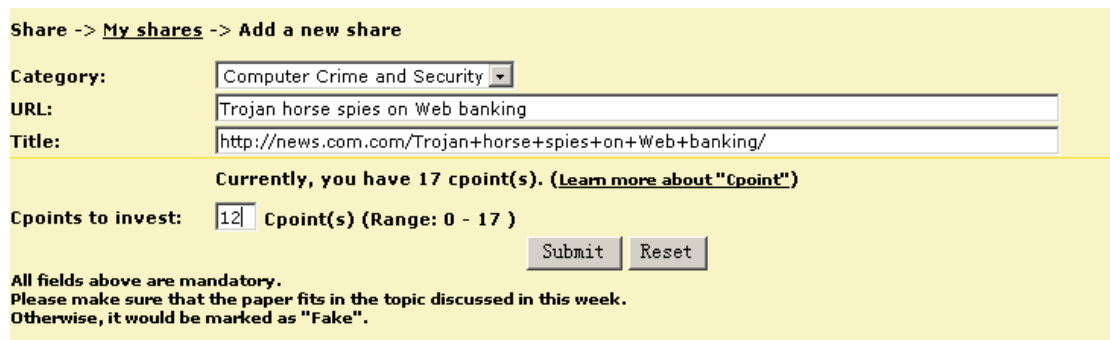


Figure 12. Welcome page of Web-based Comtella

The rating interface is embedded in the “search page” of the system (Figure 6), through which the user can rate the articles immediately after searching and viewing them. The current total of earned ratings and the times the article was viewed are shown for each article allowing the user to select for reading the articles that have been viewed often by others. By default, the articles are sorted in the list by the number of *c-points* and the sharing time. The users can re-sort them by article title, number of earned ratings or view-times.

Whenever the user rates an article, she will obtain a certain number of *c-points*. The number depends on the user's reputation of giving ratings (evaluated with Formula 2 in Section 5.3.2). To stimulate users to consume their *c-points* sooner and rate more articles, the *c-points* are effective only within two weeks after they are obtained. When sharing a resource, the user has the option to invest a certain number of *c-points* in the resources (Figure 13). The user can decide how many *c-points* to attach to the article, depending on how interesting, readable and relevant they think it is. However, to avoid the unbounded competition for the largest number of *c-points*, there is a limit to the number of *c-points* that can be invested in one article (50 *c-points*).



Share -> [My shares](#) -> Add a new share

Category:

URL:

Title:

Currently, you have 17 cpoint(s). ([Learn more about "Cpoint"](#))

Cpoints to invest: Cpoint(s) (Range: 0 - 17)

All fields above are mandatory.
Please make sure that the paper fits in the topic discussed in this week.
Otherwise, it would be marked as "Fake".

Figure 13. The interface for sharing resources in Comtella

The users' memberships are updated at the beginning of each week, based on their contribution levels in the previous week. The comprehensive evaluation of the user's contributions is computed through Formula 4 in Section 5.3.3. To limit the variable space so that it is possible to evaluate the improved mechanism, in this version of Comtella, only sharing articles and giving ratings are considered to be cooperative activities. However, not only the times they are performed, but also the quality of articles shared and the ratings made are taken into account. The weights for the number of articles and the number of ratings depend on the user's reputation in each of these activities (contributing high quality papers and making high quality ratings) and the needs of the community at the moment, while the weights for the quality of them are invariable.

There are four hierarchical membership levels in the new Comtella system – Gold member, Silver member, Bronze member and Plain member. The user interfaces

provided for the members with different memberships are different in appearance and color but same in terms of functionality (Figure 14). Although higher level members are not offered additional functions anymore, they still have some privileges in the community. For example, they receive more ratings to give out, which means they are more influential in the community. Also, Gold members are enabled to publish the news about themselves or the community at the front page of the system.

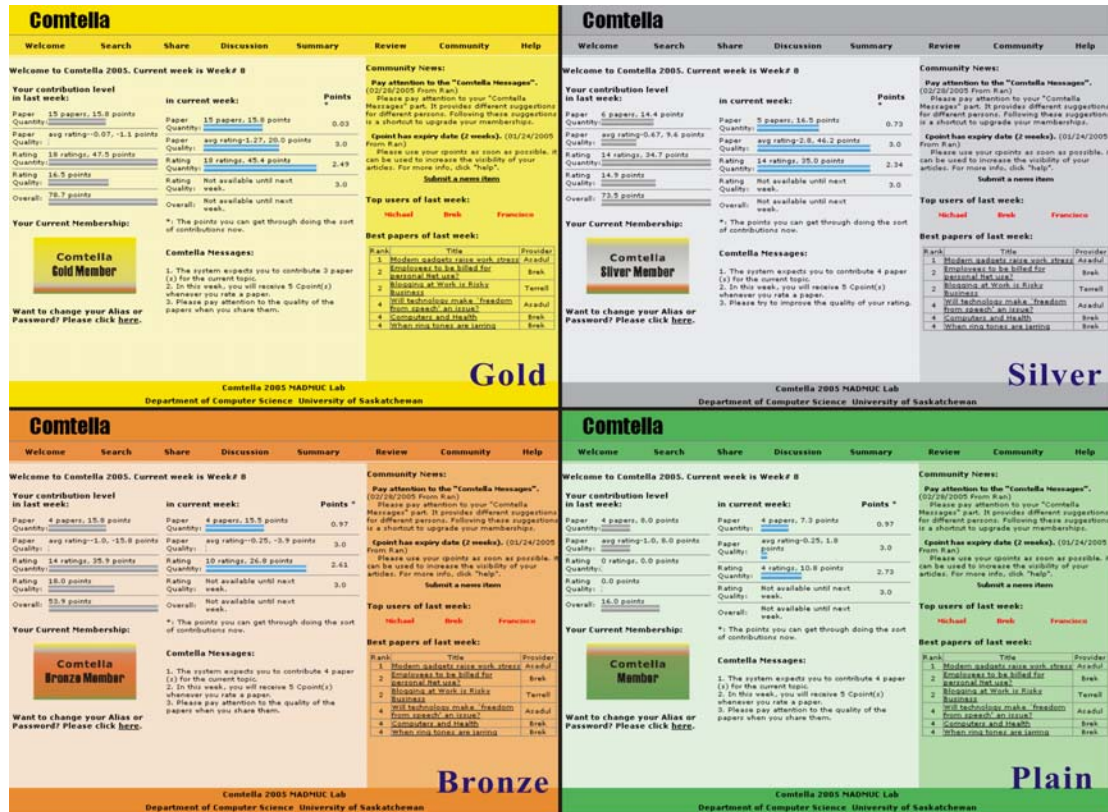


Figure 14. Different interfaces for users with different membership levels

According to the evaluation plan, the users were divided into two groups: the test group and the control group. Two different systems were created to serve the two groups which were identical except for the extra features (see Table 4) that were offered in the system for the test group.

Table 4. Differences between the two systems for the two groups

Feature	System for Test Group	System for Control Group
Hierarchical Memberships	✓	✓
Showing Contribution Levels in Previous and Current Week	✓	✓
Interface for Rating Articles	✓	✓
<i>Cpoints</i> as Reward for Rating	✓	✗
Adaptive Weights for Sharing and Rating	✓	✗
Personalized Messages	✓	✗

6.3 Evaluation

- **Introduction**

After two systems were set up, a second case study based on Comtella was launched in the same course on Computer Ethics and Information Technology offered in the 2004-2005 winter session to evaluate the effectiveness of the improved incentive mechanism. The study was carried out for 9 weeks. The participants were encouraged to use the system to share Web-articles related to the topic of each week. According to the curriculum, one different topic was introduced in each week except that the same topic was discussed in the 4th week and the 5th week. In total, eight topics were used during the study.

5% of the final mark of the course was based on the number of articles the student shared during the study. Another 5% was based on the student's participation in discussions. They could choose to participate either in the classroom discussion or in the online discussion forum in Comtella.

- **Participants**

Thirty-one (31) fourth-year undergraduate students who took the course were the participants. They were assigned into two groups: a test group of 15 users and a control group of 16 users. Since there might be some cultural and gender-based differences in the users' initial predisposition for participation, the assignment of users to groups was

based on having equal proportion of Canadian to foreign and male to female students in each group. Before the study, all the participants voluntarily signed consent forms⁵, which granted me permission to use their data.

- **Methods**

To evaluate the effect of the improved incentive mechanisms, I compared the behaviors of the test group, which used the system with all the features of the mechanism and the control group, which used the system where some functions were blocked, as shown in Table 4. To avoid the effects that the contribution patterns of one group can have impact on the behaviors of the other group, the two groups of users formed completely separated online communities, but followed the same class schedule, shared lectures and classroom and had the same coursework throughout the study.

After the evaluation, a post-experiment questionnaire⁶ was distributed to the participants to collect feedback and information about their experiences in their online communities. Comtella was programmed to record the times and the types of various user actions and contributions in the online communities, including logging on the system, sharing, reading and rating articles, commenting on articles, uploading summaries, etc. The data from the questionnaire and the system-logs were analyzed and compared to answer the following questions:

- Did the users in the test group rate articles more actively?
- How well did the summative ratings reflect the real quality of the articles in both groups? If the users in the test group gave more ratings than those in the control group, did the summative ratings in test group reflect the real quality of the articles more accurately?
- Did the users in the test group tend to share resources earlier in the week?
- Was there a significant difference with respect to the total number of articles between the test group and the control group? Did the users in the test group share the number of articles that was expected from them?

⁵ The consent form can be found in Appendix E

⁶ The questionnaire and the results can be found in Appendix F

▪ Did the test group as a whole produce a higher percentage of high-quality articles? What was the users' perception with respect to information overload in each group?

- **Results and discussion**

The data from the system-logs and the post-experiment questionnaire showed the following results and user contribution patterns.

1. ***The users in the test group were more active in rating articles.***

The system records showed that the number of ratings given in the test group was consistently higher than that in the control group in each week (Figure 15). Throughout the nine weeks, the difference between the total numbers of ratings in both groups was significant - 1065 in the test group and 593 in the control group. The percentage of the articles rated was also higher in the test group than in the control group (78.3% in the test group vs. 53.3% in the control group). I also counted the ratings contributed by each individual user and applied the Mann-Whitney test (Lowry 1999) to verify the hypothesis that the users in the test group rated articles more actively than those in the control group did during the study⁷. The results of the test showed that the z-ratio was equal to 1.727, which confirmed the hypothesis with statistical significance beyond the 0.05 level. This clearly shows that the incentive mechanism with *c-points* and the associated rewards showed sustained effectiveness in stimulating users to rate articles during the experiment.

Besides, the users' attitudes towards the *c-points* were mostly positive. Among the 15 users in the test group, nine users (60%) indicated in the questionnaires that the *c-points* were useful to make their articles more visible to others; eleven of them (73.3%) used more than 40% of all the *c-points* they earned to make their contributed articles more visible. Of all the *c-points* awarded to the users, 62.8% were attached to articles.

⁷ The detailed process of the Mann-Whitney test is described in Appendix G

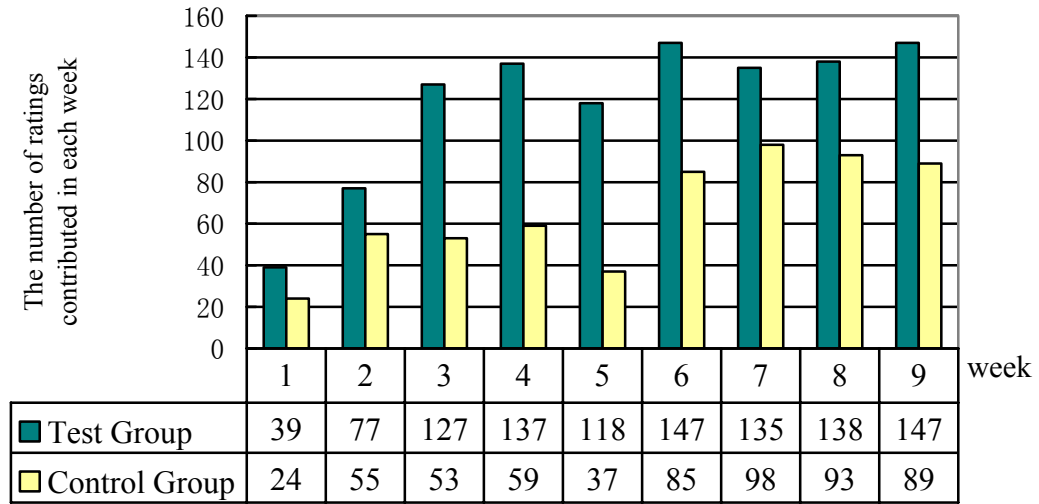


Figure 15. The number of ratings contributed in the two groups in each week

2. *The articles with higher ratings were more likely to be chosen by users to summarize.*

I found this result by analyzing users’ activity-logs in both systems. I observed that the articles with higher summative ratings were chosen by the students for summarization⁸ more times on average. We can safely assume that the students tended to choose interesting or good-quality articles to summarize because they desired good marks for their summaries. The attention and the effort they paid to the articles they chose also indicated the articles were valuable. So the number of times an article was chosen can reflect its quality. Then the correlation between articles’ ratings and the number of the times they were chosen can be explained by two hypotheses:

- i. The articles’ summative ratings also reflect the quality of the articles so that they showed a correlation with the number of times the articles were summarized.
- ii. The students tend to choose articles by looking at their summative ratings. If this hypothesis is true, the correlation does not mean that the ratings correctly reflect the quality of the articles.

In the exit questionnaire the users were asked whether they chose articles to summarize according to the summative ratings or the times the articles were viewed or

⁸ Writing summaries was a weekly-based assignment of the course; students could freely choose one article shared in the system to summarize.

none of these two indicators. Seven users answered that their selection was guided by the summative ratings; two users answered that their choice was guided by the number of times the articles were viewed; the remaining 22 users (71.0%) indicated that they chose articles based on none of these two indicators.

Since there is a connection between the article’s summative rating and the number of times it was viewed, I removed from consideration the data from the seven users who looked at the ratings and the two users who looked at the number of view-times, and took into account only the data from the 22 users who stated that they didn’t choose articles based on any of these indicators. The results still showed a strong correlation between the articles’ rating and the number of times they were summarized by the 22 users. Table 5 shows that the articles that earned higher ratings were on average chosen and summarized more often.

Since these 22 users selected articles according to neither the summative ratings nor the number of view-times and the correlation between the articles’ summative ratings and the number of times they were chosen by these users is statistically evident, I conclude that the articles’ summative ratings are able to reflect their quality.

Table 5. The relation between the articles’ summative ratings and the times they were summarized by the users

R: Summative rating N: The number of the articles that had the rating T: The total times these articles were chosen for summarization (by the 22 users) N/T: The number of times these article were chosen on average			
R	N	T	N/T
Greater than 4	7	5	0.714
4	26	15	0.577
3	89	31	0.348
2	167	34	0.204
1	303	39	0.129
0	47	7	0.149
Less than 0	154	6	0.039
Not rated	407	18	0.044

An interesting result is that although there was no limit on the type or the length of the articles that could be shared, most of the articles shared by the users in the two groups were magazine articles, which were very popular and easy to read. Some

academic papers were also shared during the study but their number was small. The articles with high summative ratings were usually the ones that were easy to read and understand and were very interesting. Academic papers that were long and profound were usually ignored by the users. So the quality reflected by summative ratings indicated the interestingness and readability of the articles more rather than the technical or academic quality of the articles.

3. The users in the test group were more satisfied with the summative ratings for their articles.

I found that the users in the test group were more satisfied with the summative ratings for their contributions. The data from the questionnaire showed that 53% of the users (eight users) in the test group thought that the final ratings for the articles they shared reflected fairly the quality, while in the control group only 31% (five users) thought so. This is partly because the articles in the control group were less rated than the ones in the test group. The ratio of the articles to the ratings in the test group is 1: 1.74 and that in the control group is 1: 1.01. Around half (46.7%) of the articles in the control group were not rated at all while in the test group only 21.7% were unrated. More deserving articles did not receive ratings in the control group than in the test group. Therefore, the quality evaluation based on collaborative rating requires a critical number of user ratings. Before this number is reached, increasing the number of user ratings in the system through incentives can improve the accuracy of the quality evaluation of the shared resources.

4. The users in the test group tended to share resources earlier in the week.

The system-logs showed that the users in both systems shared more articles in the first half of each week than they did in the second half. 66.1% of all the articles in both systems were contributed in the first three days of the week. Yet, the users in the test group shared a higher percentage of their contributions in the first three days (71.3%) than the users in the control group (60.6%) and the difference between the two groups was significant in each week (ranging between 7% and 14%).

In Comtella, one topic was active in the system for only one week and the resources shared for the topic could be rated only in that week. So the resources posted earlier in

the week gained more time and chance to attract attention and to collect ratings, which was realized by 17 users (54.8%) across the two groups (according to the data from the questionnaire). This explained why the users in both groups tended to contribute articles in the first half of the week.

The higher percentage of early contributions in the test group proved that the adaptive reward mechanism was effective in motivating users to share resources early. The reason for this conclusion is that the adaptive reward mechanism was applied only in the system used by the test group and it was the only difference between the two systems that was related to the timeliness of making contributions. As discussed in Section 5.3, for the users of test group, the adaptive weight for sharing resources decreased with the time in one week.

5. There was no significant difference in the total number of articles between the two groups. Compared to corresponding data from the pervious study, the numbers of articles shared by two groups for the topics show less fluctuation in this study.

The difference between the total numbers of articles shared over nine weeks in the two groups was not significant (613 articles shared in the test group vs. 587 in the control group). The number of articles contributed by each user ranged from 3 to 111 in the test group and from 0 to 124 in the control group. The standard deviation in the test group was slightly smaller than that in the control group (29.4 vs. 32.1).

Since the number of participants in the two studies was nearly the same (35 in the pervious study, 31 in this study) and the eight topics in this study are same as the last eight topics in the pervious study, it is not surprising that the numbers of articles shared for these topics in the two studies were comparable. Figure 16 shows that the numbers of articles shared for these topics in the second study fluctuate less than the corresponding numbers in the first study and consistently fall into a higher range. More articles were shared for Topic 4 in the second study because this topic was kept in both systems for two weeks while others were kept for only one week. Except for that, the numbers of articles shared for the topics are basically in the same range. It is reasonable to expect some differences between the numbers since the interestingness of the topics is hardly

the same. The average number of articles shared for all the topics in the second study (150.0) is close to that for the last four topics in the first study (158.8), when the motivational mechanism had been introduced and it is higher than the average number of articles shared for the first four topics in the first study (68.8). These results show that the improved incentive mechanism persistently stimulated users to contribute more resources throughout the study. This answers the question raised by the decline of contributions in the end of the previous study regarding the possibility that the increase of contributions was short-lived due to the novelty effect.

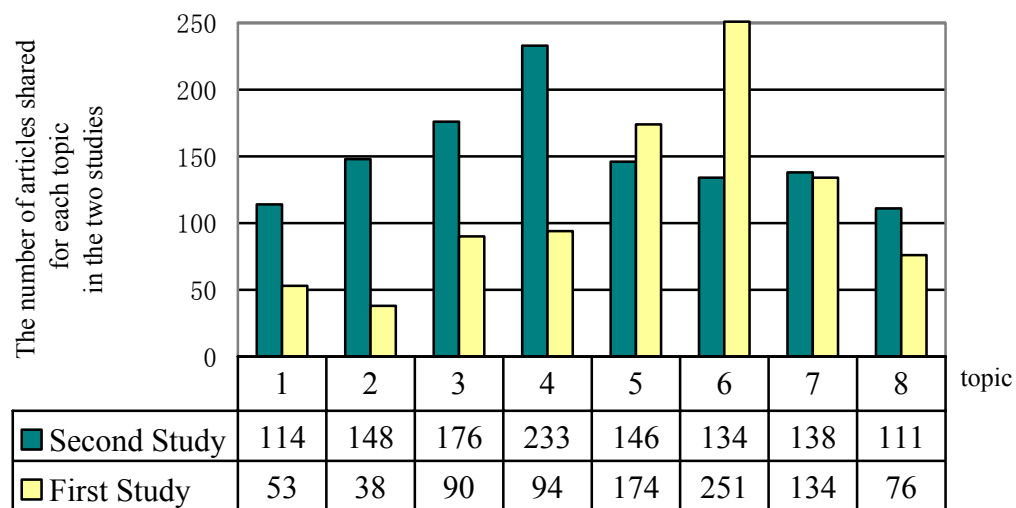


Figure 16. The number of articles shared for each topic in the two studies

6. In the test group, more than half of the users tried to share the number of articles that was expected from them. The overall number of articles in the test group was not excessive.

In the questionnaire, the users of the test group were asked whether they followed the system suggestion about the number of articles they were expected to contribute. Eight users (53% out of 15) said that they tended to share the number of resources that was expected from them. I calculated for each user the average difference between the actual shared number and the expected number over nine weeks and found that for eight of the users (53%) the average differences were less than 2, which means these eight users almost contributed the number of articles that was expected from them. Interestingly, the two groups of eight users did not entirely overlap. Table 6 shows each user's answer to this

question and the average of the differences between the actual number of articles she shared and the number expected from her in the nine weeks. These results indicate that more than half of the users in the test group were persuaded to share resources in or close to the number that was expected from them.

Table 6. The differences between the users' real contributions and their expected contributions

User #	Average Difference	Answer to the question
04	0.75	No
07	0.88	Yes
12	0.88	Yes
13	1.00	Yes
03	1.13	Yes
10	1.13	Yes
01	1.38	Yes
15	1.63	More
14	2.00	No
11	2.25	Yes
08	2.75	Less
05	3.38	Yes
09	5.25	Less
06	8.38	No
02	9.00	More

The question
Did you pay attention to the number of articles the system expects you to contribute?
 Four options:
 Yes - Yes, I tried to share in the number the system expected.
 More - I always shared more than the number.
 Less - I always shared less than the number.
 No - I did not care about that number; I shared as many as I wanted.

In the test group, the total number of contributions did not have a big discrepancy from the overall expected number for most of the topics since about half of the users tried to share the number expected from them and the extra contributions made by users who tended to share more were compensated roughly in the same number by the contributions of users who tended to share less. Table 7 shows the differences between the total number of shared articles and the overall expected number for the eight topics. It can be seen that the differences for all the topics were less than 20% except for Topic 3, 5 and 8. The students over-contributed for Topic 3 possibly because the topic

(“wiretapping and encryption”) happened to be of highest interest for them. They under-contributed in the last week possibly because of their heavy coursework load at the end of the term. But for the difference of Topic 5, I have not found any explanation. The overall difference for the eight topics was equal to 12.3% of the total of the expected numbers. This shows that the number of articles in the test group was not excessive and the approach of setting specific goal for each user to contribute was helpful to control the overall quantity of resources in the system.

Table 7. The differences between the actual number and the expected number in the test group for the eight topics

E: The expected number of articles shared for the topic. A: The actual number of articles shared. D: The difference between A and E, equal to (A-E). P: The percentage of the difference, equal to (A-E) / E.									
Topic	1	2	3	4	5	6	7	8	Sum
Week	1	2	3	4 & 5	6	7	8	9	
E	60	60	61	121	62	62	60	60	546
A	57	66	98	121	83	72	69	47	613
D	-3	6	37	0	21	10	9	-13	67
P	-5.0%	10.0%	60.7%	0.0%	33.9%	16.1%	15.0%	-21.7%	12.3%

Note: The same topic was discussed in week 4 and 5. So there was only one expected number for both of the weeks.

7. In both groups, the users’ attitudes towards the quality of the articles were generally neutral. The users in the test group were more active in terms of logging on the system and reading articles.

As for the quality of the articles in both systems, the questionnaire asked the users in both the control and test group to give the rough estimate of the percentages of articles with high, medium and low quality in their respective systems. The data in Table 8 shows the averages of users’ estimations, which indicates that their attitude towards the quality of the articles in their own communities was basically neutral. However, it is hard to compare the quality of the articles in the two groups based on this data because the users in any group had experience only in one system and they might have different criteria of quality evaluation.

Table 8. The users' estimations of the percentages of the articles with high, medium and low quality in both groups

Group	High	Medium	Low
Test Group	23.5%	47.1%	29.4%
Control Group	27.3%	41.5%	31.1%

The analysis based on the system-logs showed that the users in the test group participated in the system more actively than the users in control group did. The number of times of reading articles and logging on the system were computed for both groups of users over nine weeks. In each week, on average, the users in the test group consistently read more articles and logged on the system more times than the users in the control group (except that in the first week the control group had more logon times). Figures 17 and 18 show the average number of times of reading articles and logon in the nine weeks in both groups. Throughout the study, the total number of times of reading in the test group and in the control group were 3407 and 2373 respectively; the total number of logon times were 1714 and 982. This clearly shows that the activities of reading articles and logging on the system in the test group were more frequent than those in the control group, which proves that the improved mechanism could ensure more active and sustained user participation. A noteworthy fact is that these activities were rewarded in neither of the systems for the two groups. Therefore, the more active user participation can not be directly attributed to a particular incentive in the mechanism.

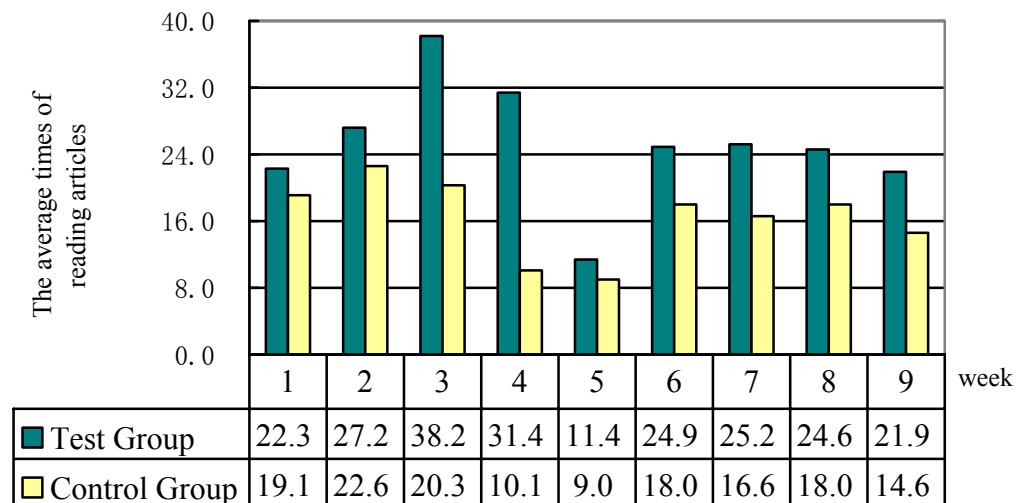


Figure 17. The average number of times of reading articles for the users in both groups

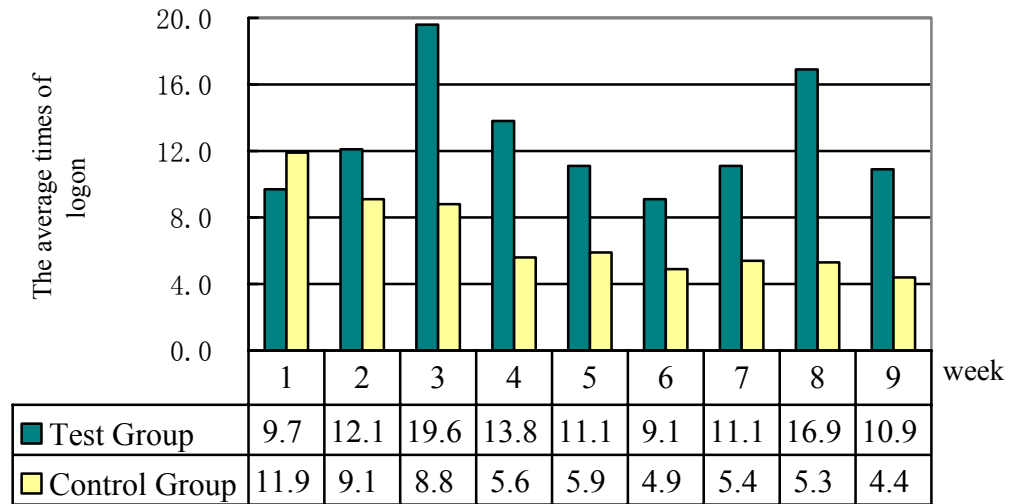


Figure 18. The average number of times of logon for the users in both groups

One possible explanation is that the test group users' more active participation in reading articles and logon is a by-product of stimulating them to rate articles. However, this is not the case. I have looked over the relevant data and found that there was no correlation between the number of ratings contributed and the times of reading articles or the times of logon in both systems during the nine weeks. Table 9 listed the number of ratings contributed, the number of times the users read articles and the number of times they logged on the system in each of the nine weeks for both groups. Obviously, more ratings in a particular week are not related to more readings or more logons in that week (e.g. for the test group, Week 9 and for the control group, Week 7). I also computed for both systems the correlation coefficients of the number of ratings and the times of reading or logon over the nine weeks (see Table 9). It can be seen that none of them is greater than 0.25, which clearly shows that the activities of reading and logon are not correlated with rating in both systems.

Actually, for all the users in both groups, the numbers of the articles they read are far more than the numbers of articles they rated. The system-logs showed that the users rated only about 30% of the articles they read on average. The users of the test group rated a higher percentage than those of the control group (34.6% vs. 25.7%). Obviously, the users could rate more articles that they had read without any extra effort. If their goal was just to rate articles, they definitely did not need to read new articles since there were

plenty of articles that they had read but not rated yet. Therefore, the incentive mechanism was able to increase the proportion of the articles the users rated from the articles they read, but unable to stimulate users to read more articles. This explains why more ratings in particular weeks did not guarantee more times of reading articles in those weeks during the study.

Rating articles is more independent of the action of logging on the system. The users logged on the system for various purposes and they could rate as many articles as they want through one logon. So encouraging users to rate articles had no influence on their activity of logging on, which is also showed in Table 9.

Table 9. The number of ratings, the times of reading and number of logon actions in each week and the correlation coefficients for both groups

Week	Test Group			Control Group		
	Rating	Reading	Logon	Rating	Reading	Logon
1	39	335	145	24	306	191
2	77	408	181	55	361	146
3	127	573	294	53	325	140
4	137	471	207	59	162	89
5	118	171	166	37	144	94
6	147	374	137	85	288	79
7	135	378	167	98	265	87
8	138	369	253	93	288	85
9	147	328	164	89	234	71
Correlation coefficient with the # of ratings		0.124	0.246		0.067	-0.757

In addition to the difference in the times of reading and logon between the two groups, I also observed that in the control group the times of logon decreased during the nine weeks and it showed a negative correlation with the number of ratings because ratings increased during the study (see Table 9). However, in the test group no such decrease was observed. This gives some hint that the levels of information overload are different in the two groups. Jones and Rafaeli (1999)'s research indicated that information overload results in a decrease or end of the users' participation in online communities. Our results from the previous evaluation also showed that users' contributions and participation decreased after the system was flooded with shared articles. Evidently, user participation is roughly in inverse proportion to the level of

information overload in the online community provided that there are no motivational factors working on the users. Since reading articles and logging on system are not the activities that were encouraged by the improved mechanism in any group, the difference between the users' levels of engaging in these activities in the two groups can show that the information overload in the test group was less serious than that in the control group.

6.4 Summary

This chapter presents how the Comtella system was re-implemented to improve its performance and to include the new features of the improved incentive mechanism and how the case study was setup and carried out to evaluate the effectiveness of the mechanism. After the study, some important results were found through analyzing the system-logs and the users' feedbacks.

Firstly, the users of the test group contributed more ratings than the users of the control group. They also contributed their articles earlier in each discussion period. These mean that the improved mechanism encouraged users to rate articles and to contribute articles when they were most needed by the community.

I also found that the articles with higher ratings were more likely to be picked for summarization even by the users who did not use the article rating as a guide for choosing articles to summarize. This means that the articles with high ratings were really more interesting and desirable and therefore of higher quality.

The evaluation was unable to yield clear positive results about the effect of mechanism on encouraging users to contribute high-quality articles. Still, I found that the users of the test group participated in the system more actively in terms of reading articles and logging on the system, which implies the improved mechanism lowered the level of information overload in that group.

In the next chapter, these results are discussed further and some conclusions are generated. And as well, some possible directions for future work are presented.

Chapter 7

Conclusions and Future Work

To have an online community survive and become self-sustained, a critical mass of user participation has to be reached. The goal of the hierarchical membership mechanism proposed in Chapter 3 is to motivate users' contributions and participation in the online community, thereby achieving its thriving and sustaining. The results of the first case study clearly showed that the mechanism was able to stimulate users to make more contributions and participate more actively in the system. However, the fact that there was a fixed reward for contributions made some of the users try to game the system by sharing many resources of lower quality. This catalyzed information overload in the community and then led to users' disappointment and a decrease in their participation toward the end of the study.

To achieve a sustainable level of user participation in online communities, it is important to control the quality and the quantity of users' contributions and avoid information overload or degrade its level. Therefore, the mechanism of hierarchical memberships was augmented with two mechanisms described in Chapter 5 and the improved mechanism was reevaluated in another case study. The first mechanism that was added, the collaborative rating mechanism ensures a decentralized way of measuring the quality of contributions by encouraging the users to rate each other's contributions. Based on this quality measurement, the adaptive reward mechanism encourages users' contributions differently, taking into account the users' individual reputation and the current needs of the community.

The results from the second case study showed that the improved mechanism worked well in the online community, and was able to achieve most of the goals as desired. Firstly, the data showed that throughout the study, the users in the test group consistently contributed more ratings than the users in the control group, which proves

that the collaborative rating mechanism can persistently stimulate the users to rate more articles. Besides, a quite evident correlation between the numbers of times the articles were chosen for summarization and the articles' summative ratings was found, which shows that the summative ratings can reflect the quality of the articles. The users in the test group seemed more satisfied with the ratings earned by their contributions because their contributions were rated more than those in the control group. Apparently, a critical number of user ratings have to be reached for peer-evaluation based quality evaluation systems to work accurately.

The data showed that the improved mechanism successfully motivated the users in the test group to contribute their articles *earlier* in each discussion period. The overall number of the contributions in the test group didn't exceed the expected number in most of the weeks. More than half of the users in that group were persuaded to contribute articles in or close to the number that was expected from them. These results show that the mechanism is helpful to control the overall number of resources in the system. Besides, it was found that throughout the study, the users in the test group were consistently more active in reading articles and logging on the system, the activities that were not rewarded by the mechanisms in the two groups. This indicates the improved mechanism is able to ensure more active and sustained user participation in the online community. According to the results from the previous evaluation on Comtella and the research by others (Jones and Rafaeli 1999), users' participation level decreases with the aggravation of information overload. Therefore, the users' different participation levels in the two groups indicate the level of information overload in the test group was lower than that in the control group. However, whether information overload was entirely avoided in the test group is still unknown.

Although the results of the case studies of the proposed mechanisms are quite positive and exciting, there are still some questions that have not been answered by the study and also some very interesting directions that deserve further research.

Firstly, whether the improved incentive mechanism is able to improve the quality of the resources in online community is still unknown. The results from the questionnaire about the quality of shared articles are unable to answer the question about which group as a whole produced higher-quality contributions. I suppose that the articles in the test

group are of higher quality since they were read more frequently by users. However, more work is needed to prove this hypothesis. One possible solution is to invite a new group of students to evaluate the articles shared in both systems. Because none of them have contributed any of the articles, their opinions would be more objective.

Secondly, the performance of the collaborative rating system can be improved. It has been found that a higher number of user ratings could make the quality evaluation more accurate. However, when the number of ratings reaches the critical mass level, the summative ratings will converge towards the community measure of quality of the resources, and more ratings will not improve the accuracy of the evaluation. It would be valuable to find out the critical mass of user ratings for the peer-based rating systems to work properly. When this threshold is known, the rewards for rating can be adapted so that rating articles is rewarded more before the threshold is reached. In addition, the problem that some good articles may be unread and end up with no ratings remains unsolved. To encourage users to rate unrated articles, it is possible to offer the first rater of the article some extra benefit, especially if the article ends up with high rating.

Thirdly, a recommender system can be applied to help the users to overcome the problem of information overload. Although the adaptive reward mechanism has proved to be able to lower the level of information overload in the online community, the problem can not be completely eliminated, especially for large-scale online systems with hundreds of users or more. It would become much harder to control the overall quantity of the contributions in the system if the user population keeps increasing. Therefore, we could recommend for the users some articles they might be interested in or put these articles on the top of the result list. This may help them to filter out the information they do not want. The collaborative filtering algorithm (Herlocker et al. 2000) could be applied and based on the ratings of the articles. Besides, it is also possible to model relations between the users based on the ratings they gave. In a way similar to collaborative filtering systems, we can find out the users who always rate the same articles similarly or closely and can recommend their contributions to each other, since they may have similar taste or interests.

Finally, some changes in the adaptive reward mechanism will be needed, if it is applied into general online communities. The adaptive reward mechanism was designed

in the context of the educational system Comtella. The main characteristic of Comtella is that the topic for sharing resources in the system is changed weekly and at any given time there is only one active topic being discussed. Therefore, the evaluation of users' contributions and reputation and the updating of the users' memberships follow a weekly rhythm. However, in general interest-based online communities, like discussion forums, blog-systems, collaborative filtering systems, etc., usually the contributions are going on simultaneously in many forums/categories and there is no time limit to the participation in each forum/category. In spite of these differences, it is still possible to motivate or control users' contributions based on the quality data of their previous contributions. We could build the user's reputation of making contributions on all the topics or on each topic separately, depending on different system requirements. The discussion on some categories is ephemeral. For these categories, it is also necessary to encourage users to contribute early because later contributions may not get enough notice. Every online community has its own characteristics. There are no mechanisms that are perfect for all systems. However, the basic ideas of rewarding desirable activities, assigning different status and service to users to arouse comparison and adapting users' reward to influence their contributions can be applied widely. These ideas are coherent with three persuasion theories from social psychology and have proved effective in building a thriving and sustained online community.

References

- Beenen, G., Ling, K., Wang, X., Chang, K., Frankowski, D., Resnick, P., & Kraut, R. E. (2004). Using Social Psychology to Motivate Contributions to Online Communities. *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, Chicago, Illinois, 2004, 212-221.
- Bretzke H., Vassileva J. (2003). Motivating Cooperation in Peer to Peer Networks. *Proceedings of the International Conference on User Modeling*, Johnstown, PA, June 22-26, 2003, Springer Verlag LNCS 2702, 218-227.
- Callaham, T., & Pavich, R. (1998). Indian Caste System. Available online at: <http://www.csuchico.edu/~cheinz/syllabi/asst001/spring98/india.htm>.
- Cialdini, R. B. (2001). The Science of Persuasion. *Scientific American*, 2001, 76-81.
- Cosley, D., Lam, K. S., Albert, I., Konstan, A. J., & Riedl, J. (2003). Is seeing believing?: how recommender system interfaces affect users' opinions. In: Cockton, Gilbert, Korhonen, Panu (ed.): *Proceedings of the ACM CHI 2003 Human Factors in Computing Systems Conference*. Ft. Lauderdale, Florida, USA, 2003, 585-592.
- Erickson, T., & Kellogg, W. A. (2000). Social Translucence: An Approach to Designing Systems that Mesh with Social Processes. *ACM Transactions on Computer-Human Interaction* 7, 1, 2000, 59-83.
- Golle, P., Leyton-Brown, K., & Mironov, I. (2001). Incentives for Sharing in Peer-to-Peer Networks. *Proceedings of the 3rd ACM conference on Electronic Commerce*, Tampa, Florida, 2001, 75-87.
- Greer, J., McCalla, G., Vassileva, J., Deters, R., Bull, S., & Kettel, L. (2001). Lessons Learned in Deploying a Multi-Agent Learning Support System: The I-Help Experience. *Proceedings of AIED*, San Antonio, 2001, 410-421.
- Herlocker, J. L., Konstan, A. J., & Riedl, J. (2000). Explaining collaborative filtering recommendations, *Proceedings of the 2000 ACM conference on Computer*

- supported cooperative work*, 2000, 241-250.
- Hiltz, S. R., & Turoff, M. (1978). *The network nation: Human communication via computer*. Addison-Wesley Publishing Company, Inc., London, 1978.
- Johnson, S. (2001). *Emergence: The Connected Lives of Ants, Brains, Cities, and Software*. Publisher: Scribner, 2001, 152-162.
- Jones, Q., & Rafaeli, S. (1999). User Population and User Contributions to Virtual Publics: A Systems Model. *Proceedings of the international ACM SIGGROUP conference on Supporting group work*, Phoenix, Arizona, 1999, 239-248.
- Klingberg, T., & Manfredi, R. (2002). *The Gnutella Protocol Specification v0.6 Draft*. Available online at <http://rfc-gnutella.sourceforge.net/draft.txt>.
- Lampe, C., & Resnick, P. (2004). Slash(dot) and Burn: Distributed Moderation in a Large Online Conversation Space. *Proceedings of the SIGCHI conference on Human factors in computing systems*, Vienna, Austria, 2004, 543-550.
- Lowry, R. (1999). Subchapter 11a: The Mann-Whitney Test & Subchapter 12a: The Wilcoxon Signed-Rank Test, *Concepts and Applications of Inferential Statistics*, 1999. Available online at <http://faculty.vassar.edu/lowry/webtext.html>.
- Merton, R., & Zuckerman, H. A. (1968). The Matthew Effect in Science: the Reward and Communication Systems of Science are Considered. 1968, *Science* 199, 3810, 55-63.
- Millen, D. R., & Patterson, J. F. (2002). Stimulating social engagement in a community network. *Proceedings of the 2002 ACM conference on Computer Supported Cooperative Work*, New Orleans, 2002, 306-313.
- Nabi, R. L. (2002). Discrete Emotions and Persuasion. *The Persuasion Handbook: Developments in Theory and Practice*, Sage Publications, 2002, 289-308.
- OpenSymphony (2004). *WebWork Online Documentation*. Available online at: <http://www.opensymphony.com/webwork/wikidocs/Documentation.html>
- Preece, J. (2001). *Online communities: Usability, Sociability, Theory and Methods*. In Earnshaw, R., Guedj, R., Dam, van A., & Vince, T. (Eds) *Frontiers of Human-Centred Computing, Online Communities and Virtual Environments*. Springer

Verlag: Amsterdam, 2001, 263-277.

Reed, D. P. (1999). That Sneaky Exponential—Beyond Metcalfe's Law to the Power of Community Building. *Context magazine*, spring 1999. Available online at: <http://www.contextmag.com/archives/199903/DigitalStrategyReedsLaw.asp>.

Shenk, D. (1997). *Data smog: Surviving the information glut*. HarperCollins, New York, 1997.

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

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

Vassileva, J., Cheng, R., Sun, L., & Han, W. (2004). Stimulating User Participation in a File-Sharing P2P System Supporting University Classes. *P2P Journal*, July 2004 issue, 14-23. Available online at: http://www.p2pjournal.com/issues/Jul_04.pdf.

Appendix A: The Socio-psychological Foundation of the Hierarchical Membership Mechanism

Although online communities are virtual communities in cyberspace, they share many features with real social communities. Therefore, some theories of social psychology could be used to guide the design of the mechanisms to motivate user participation and contributions in online communities. The mechanism of hierarchical memberships is supported by two persuasion theories from social psychology, that is social validation theory and the theory of discrete emotions.

- **Social validation**

Social validation theory implies that one of the main ways that people decide what to do in a particular situation is to look at what others are doing or have done; if many individuals have decided in favor of a particular idea, more people would tend to follow this way (Cialdini 2001). According to this theory, users' memberships are made public in the online community in order to arouse the comparison and the competition between the users. Moreover, the users are visualized in different colors depending on their memberships and in different size depending on the quantity of their contributions (Vassileva et al. 2004). In this way, they will try to compare and compete with each other by contributing more and participating more actively.

- **Theories of discrete emotions**

Discrete emotions are “the emotions that have unique appraisal patterns, motivational functions, and behavioral associations” (Nabi 2002). Common emotions such as fear, anger, sadness, joy, etc. are discrete emotions. There are theories about each discrete emotion. The theory of fear is most relevant to the proposed mechanism. According to the theory, people will feel fear when they perceive some threat to themselves or their properties. This makes the incoming messages, especially those about how to avoid the threat more persuasive to them (Nabi 2002). Therefore, if a user

who holds relatively high level membership and enjoys better services stops participating in the community or contributes less than before, the system will show that the amount of her participation is decreasing, which may arouse her fear that her membership is going to be degraded. At this time, a message related to the actions that the user can take in order to avoid demotion may provide effective persuasion (e.g. pointing to one or more of the cooperative activities that the user has been neglecting).

Appendix B: The Questionnaire Used in the Study in 2003-2004

Winter Session and the Results

1. Did you pay attention to your own membership level?
 - a. Yes, every one of the 4 weeks when the option was available. 48%
 - b. In 3 of the weeks. 19%
 - c. In 1 or 2 of the weeks. 12%
 - d. Never. 19%

2. Do you care about who were the Gold members in each week?
 - a. Yes, all 4 of the weeks when the option was available 29%
 - b. In 3 of the weeks. 0%
 - c. In 1 or 2 of the weeks. 29%
 - d. Never. 41%

3. Did you ever try to upgrade your membership by making contributions to the system?
 - a. Yes 57%
 - b. No 43%If yes, did you succeed?
 - c. Yes 35%
 - d. No 22%

4. Did you think the mechanism for determining the user's membership levels was fair? Please rate:

-2,	-1,	0,	1,	2
9%	25%	16%	32%	16%

Would you like to comment on the mechanism? _____

 - Changing the rules after I find the holes, that's not fair. :-)
 - The system (from my understanding) based your status mostly on how much you shared. So many people would contribute poor or unsuitable papers, and then share everything else other students shared just to improve their rankings. I felt that my time was better spent searching for higher quality page.
 - too many irrelevant articles people used to jack-up their ratings
 - How can I answer this question when I have no idea of the membership algorithm being used.
 - I think it was poor. Every week I contributed 4 to 5 articles. This was good enough for me to maintain a silver level. In the last week I was bumped back down to bronze because students registered like 100 links. This

provides no motivation to place any links at all if the system can be manipulated

- I have to admit, I did not contribute much, so I didn't get a good feel for how the mechanism works.
- Some people took advantage of the earlier bug of reporting original links when they were really sharing other people's original files.
- Students dump large amounts of links without even reading them. What use is finding links if they don't read it.
- I don't know the mechanism. I would like to know if there were set threshold levels, percentages at each level, or what. It appeals to our base competitive nature.
- Was too easy to gain a higher membership just by leaving Comtella connected, or logging on and off repeatedly
- too many factors
- Downloading others materials is not a good measurement as classmates just did this during class time.
- It seemed like irrelevant things would get the upgrade, and important things would not.
- I didn't use the Gold/Silver/Bronze release
- I believe that even though I had a high weekly contribution ranking and original contribution ranking I should not have appears silver in most cases. The comparison should be made to the number of users in the system, not the top users.
- had no idea what was needed to upgrade.
- It is certainly the best approach of ranking individual user by their original contribution. I believe number of contribution is not very important but reading articles and making comment is more important. Because, we have to read and learn the problem and solution. This is the main objective.
- good.
- I thought it was fair but the feedback provided was in weekly intervals which I thought was too late by then.
- If the mechanism is based on number of contributions and not the quality of contributions it will invite people to submit crappy links that are only a nuisance to people who are trying to read articles. I ended up reading more or less only articles that had been graded, even if they were only rated
- Seemed fair to me. I really noticed when I got busy at the end of the term, my membership card reflected it.... I fell from silver to bronze!
- I suppose, but I think most people just submitted links that they either saw were already submitted and were rated high, or, they just submitted anything they could find from the usual web sites (wired, slate, etc.) without much thought to the actual topic.
- It is hard to evaluate the quality of contributions

5. Did you think the information provided when you click on the membership card explains why you have this membership level?

- a. Yes 48%

- b. No 19%
- c. I didn't know there was an explanation 32%

6. If yes, do you think it suggests appropriately what you can do to improve your membership level?

- a. Yes 41%
- b. No 3%

7. Please list some ways to 'cheat' the system in computing your status, if you can think of some.

- Leaving the client on over night, logging in and logging off frequently, adding one letter to the comments, rating papers randomly
- The 1000 pound gorilla tactic I used seemed to work pretty well. That the system counted very old links as new, and mine, wasn't right t
- Contribute papers not related to the subject of the week. Share every single paper contributed by other students.
- contribute junk, click on articles to share but not read them, stay online for a long time
- Downloading a whole bunch of links of other people to boost your overall contribution.
- Form a group of students, submit large numbers of papers, and highly rate each other.
- I am unaware of any was besides dumping in tons of poor links.
- I believe that if you share a link that you downloaded from someone else, it can count as an original contribution from you.
- As mentioned previously in #10. And to leave computers running Comtella while they went out for coffee.
- Download all the articles, and share any article is not relevant
- Just dumped lots of links, even irrelevant ones.
- Staying logged on for days. Logging on and off for fun during class. Linking every article you can find on some tech websites which is vaguely relate
- Just by leaving Comtella connected, or logging on and off repeatedly
- 'contribute' a lot of garbage
- Stay online until Comtella times out. Download everyone's links.
- Leaving Comtella logged on for hours. Downloading and sharing as many links as possible. Sharing irrelevant links.
- Add a massive amount of links that you personally have not checked.
- To share many papers and to simply login frequently to Comtella or leave it running even when not being used.
- There was one I know of, download a bunch of links that others contributed
- I have seen lots of students log on the system all the time. This should be stopped right now. One thing should be ensure - everyone have to read arti-post "dummy" comments -leave client logged in -share URLs that may not be very good (ie that you would not share except to increase your status)

- Randomly contribute papers even ones that are not related. -Contribute papers you have not read just for the purpose of contributing more links.
- Sharing a lot of links maybe? Submitting totally links, whether they are relevant or irrelevant junk?
- leave yourself logged in all day without really using Comtella - log in and out over and over - contribute many articles found on google without a
 - Submit off-topic or low quality papers.
 - Just submit links that others have already submitted and were rated highly.
 - the number of contributions; online time calculation

Were you tempted to try some of them?

Yes – 20% No – 80%

8. What did you think of the current 'gold' members?
- a. They are volunteers who did a good job for the community. 12%
 - b. They spent a lot of effort and are rewarded. 22%
 - c. They contributed a lot, but some of the articles were not good. 38%
 - d. They probably cheated their way to gain 'gold' status. 12%
 - e. They raised the standard and wasted everybody's time. 12%

9. If you were ever a silver or gold member of the community, please rate the usefulness of the following additional search interface options:

Remove the duplicate papers	(-2,	-1,	0,	1,	2)
	9%	3%	19%	25%	41%

Show only the new papers (after my last login)	(-2,	-1,	0,	1,	2)
	12%	3%	45%	22%	16%

Sort the search results by rating	(-2,	-1,	0,	1,	2)
	9%	6%	29%	35%	19%

Sort the search results by title	(-2,	-1,	0,	1,	2)
	6%	6%	25%	29%	32%

Sort the search results by time of sharing	(-2,	-1,	0,	1,	2)
	9%	19%	48%	9%	12%

Sort the search results by provider time	(-2,	-1,	0,	1,	2)
	19%	12%	45%	9%	12%

10. Did you prefer to use Comtella in the beginning of the term when there were fewer links available, or later, when there were more?

- a. When there were fewer links 35%
- b. When there were more links 25%
- c. The number of links didn't matter 38%

11. Did the extra functionality in the Silver and Gold interface help you to more easily find the papers you were interested in?

- a. Yes 45%
- b. No 32%

12. In your opinion, what is the most serious usability problem of Comtella?

- Much too slow from off campus. It took several minutes to find all the links for the week. Plus if my client ever crashed at any time during a search, I would not be able to login until they were all restarted.
- The connection. The login fails with no feedback; even after the bridge comes up it can hang for hours. Even after you login you can lose contact with anything outside your own data at any time.
- Users can share a link that has already been shared by another user - just call it something different. This confused me - especially in the initial versions where it was unclear who was the original contributor.
- confusing what the options are. still not sure of what articles I personally contributed. too many articles then it is very slow to load, maybe limit the preferred # you wish to see. not intuitive as to how to share or comment on an article -- shouldn't have to go to the file properties
- unable to share links using osx
- Not being able to log in half of the time.
- Student-designed interface.
- Ugly gui. Cannot sort the table by clicking on the table heading. The visualizations were poor: all information should be displayed in a better object. Clicking on the various different categories was a bad design decision.
- I think my biggest issue was that by the time I found a good link, that week was already over. I found about 15-20 links, but they were always for a previous week. Frustrating.
- I cannot visit the link provided by others by double clicking it. I also didn't like that I had to wait for such a long time for the visualization screen/canvas to refresh.
- When I did my summary, it would time out or something and then I could not log in for a few days
- Lack of functionality. No menus, toolbars and etc. like most applications.
- This is a massive tool with functions, options, views, etc that I still don't understand. I don't know if this is a problem because each person is going to use it to the level he or she is comfortable with or has time for.
- That I couldn't open links in my browser.
- centralized server, articles with inferior quality
- Crashing. If keywords could automatically be generated from the url or description
- Double clicking links downloads them. It should just view the article.
- Searching is difficult and results are difficult to sort. Especially when looking for papers that we can write a weekly summary on.
- it is very slow. It takes 30 seconds to start up the Comtella.
- I cannot think of any serious usability problems. It has a good idea to it. I just don't like the idea of membership levels.

- It is too slow and unreliable. Java makes searching slow and it takes forever to gather results.
- login doesn't always work, too many problems not logging in.
- Comtella was down lots of time. I e-mailed lots of time to our Professor for fixing the system. Please, make sure - Comtella to work most of the time except any serious problem. Honestly, I really like Comtella.
- crashing clients 2. Mac version seemed buggy (especially the visualization)
- unstable. sometimes, the link i click will not pop out the article.
- There needs to more than just simple features to encourage a higher rank. I don't know what it is but something that is rewarding but doesn't divide the gold and silver users to have an active advantage in using the system over the bronze users. Perhaps a link to a computer sci
- The number of links. The interface is not so easy to understand either. There were a number of thinks that I did not understand how to use or why I should use them. For instance I never submitted my summaries through Comtella, could not figure out how. I remember the first time I was about to use it
- There are "hidden" functions. It is not apparent what you can click and what you cannot. From one of the questions in this form, I got the hint that you can click on the circles that represent users. I never would have guessed that you can click on the circles!
- Crashing servants preventing logins
- The interface is quite bad, but I think Swing has more to do with this than poor design. Also, it rarely finds links the first time it is used. Perhaps a centralized approach would be better.
- Cannot control the quality of contributions. Sometime, it is not reliable

13. Would you like to give any comments or suggestions about visualization or about the membership levels?

- It was kinda bland. By this I mean it wasn't too stimulating and it took alot of time to find the location to roll over the node. Double click also never worked.
- Number of logins should not count, but time online should.
- The visualizations could be "prettied up" a little bit :). I found the membership levels a little pointless. Having limited usability for new users isn't very encouraging in my opinion.
- Like the membership levels but the determination of the level should be used to enhance the system and not encourage people to cheat the system. Had a hard time finding a good article for week 11 since there was so much junk.
- Visualizations were nice.
- Membership levels did not motivate me at all.
- It motivates people to submit more links, but the links got worse and worse. Different criteria needs to be developed for membership level.
- I think the visualization is a good idea, and the membership levels are a fun way of making someone want to contribute.

- More sort criteria
- membership level are unfair and inaccurate.
- I think you should explicitly state how they are come up with. I realize that 100% is the great person for that area during last week, but what are the cut-off levels of percentages of people who are put in gold bronze and silver.
- I like the idea of the competition but the quality of information is also important. If there was a way to use the ratings for membership levels??
- It's really annoying not to be able to sort by something because some people cheated the system.
- visualization is very intuitive and impressive.
- I didn't download the latest release so I wouldn't be tempted to fall into the role of a competitive article submitter. I would rather see quality articles then lots of articles.
- Membership levels need to be computer in a better way. I'm not exactly sure how but they are not completely accurate. Also, Comtella is a ugly and needs some more "spunk".
- nope
- Membership status must be determined by original contribution and comments made by the users. Visualization is perfect - I think.
- if the contribution will count to the final mark, i will like to see other's contribution for the link.
- Visualization was not as attractive as the lecture from the beginning of the year where each node was in fact a star instead of a circle.
- It was kind if nive to find out how much others contributed. It also made me a bit mad at certain individuals how obviously where only trying to get better marks by posting a lot of links.
- I think it is very exciting to have. I wish that you could click on one of the circles to get information about all of that user's contributed articles and their comments. (Maybe this feature is already in Comtella but I didn't know about it!!) I also think that relationships c
- No comment.

14. Would you like to give any comments, criticisms or suggestions about Comtella in general?

- The option to restart a client would have been nice so that if it wasn't working toward the end of the week that I could restart it without having the to write an email and wait for it to come back on.
- I think it has potential, but it still needs work. The connection has to be fixed. And there should be an easier way to comment articles. Maybe a properties button on the search screen?
- I think it has a lot of potential. It still needs a few issues worked out, but there was significant improvement over the term. I think it would have been really neat if it was fully decentralized and it was a requirement for 490 students to leave it running between 60-80% of the time (or something
- Overall, it was useful for 490 as far as ease of sharing. Lots of junk articles

though -- quality should be more important than quantity. I'm still not sure what all the components mean though -- it isn't clear what articles I submitted are.

- The default java look and feel is quite ugly! Computer Scientist never seem to care about looks, that is a mistake in my opinion
- Just that it down much of the time.
- I didn't like being forced to use Comtella for this class.
- Ok, needs to be refined.
- Just about it's use in the class. Part of the problem is that if you don't find a link right in the week when that topic is being covered, it seems useless.
- It was a good tool to motivate us to search through articles online and stay up to date with things.
- Maybe have a rating system for bad links, maybe gold members can delete bad links
- The GUI could look better or more polished. It looks very unprofessional.
- Even amongst us comp. sci. people who live by technology, we didn't trust Comtella. I talked to many classmates who used E-handin because they didn't trust handing in summaries on Comtella. Stop the version barrage.
- I think it was a very beneficial tool for this class. I cannot imagine doing this without Comtella. it would not be near as effective.
- It is a good idea, but needs to be made prettier and faster and more reliable. it is too slow and I'm impatient. The search results need to be similar to speed as kaza for example and should not slow down my computer. The servants should also be local and create a decentralized system
- after a few iterations, classes, to tweak it. It will be better
- I want to show my respect to my Cmpt 490.3 professor and the team worked behind Comtella to support us for learning. I really like this participating approach provided by Comtella. Finally, I wish you all the best and happier days to come. Thanks again.
- may want to duplicate articles immediately, when we add links to the Comtella.
- It was a good interface but the connection was poor when the system was congested with users. The system was not consistent when I used Comtella from home at times. I would only get my local files and not files/links shared by others. But overall its a good system to use. It needs better colors
- The idea is nice but not perfect (yet). I think people would be less willing to accept the system if they did not know that it was a research project and under development. However compared to visiting peoples personal webpages it is very good. I wish there was some kind of rating system for all ar
- A very good system and I think it can be made even better! I look forward to seeing future versions.
- Good start. Definitely better than web pages of links.
- I originally used the class web site to find readings and e-handin to submit my summaries because Comtella was such a pain to use. I only resorted to

- using it when I couldn't get updated weekly links from the class web site.
- very helpful- but need to improve the reliability and efficiency

Appendix C: The Process of the Wilcoxon Signed-Rank Test

The Wilcoxon Signed-Rank test was used in the first study to verify whether the users contributed resources more actively and logged on the system more frequently in the last four weeks than they did in the first six weeks. First, I made sure that both of the cases can meet all the preconditions of the test. The three preconditions that the Wilcoxon Signed-Rank test can be applied are listed below:

1. The paired values of X_A and X_B are randomly and independently drawn;
2. The dependent variable (in this case is the average number of resources shared or the average number of logon times) is intrinsically continuous, capable in principle, if not in practice, of producing measures carried out to the n^{th} decimal place;
3. The measures of X_A and X_B have the properties of at least an ordinal scale of measurement, so that it is meaningful to speak of "greater than," "less than," and "equal to".

Obviously, all the three preconditions were met in the two cases. The data from all the users in the study were taken into account. So they should be counted as randomly and independently drawn. The average numbers of resources the users shared in the first six weeks and in the last four were intrinsically continuous and comparable, so are the average numbers of time they logged on the system. After making sure that the conditions were met, I followed the Wilcoxon Signed-Rank procedure to perform the test. Here I take the average numbers of resources users shared as an example.

Let us assume that X_A is the average number of resources the user shared in the last four weeks and X_B is the average number of resources the user shared in the first six weeks. First, the difference between X_A and X_B is calculated for each user. Then, the signs and the absolute values of the differences need to be separated. The differences that are equal to 0 are eliminated from consideration at this point, since they have no

impact on the result of the test. And then all the users are ranked depending on the absolute values of the difference from the lowest to the highest.

Table 10. X_A , X_B , X_A-X_B , the absolute value and the sign of X_A-X_B for each user

Subj.	X_A	X_B	X_A-X_B	Sign of X_A-X_B	Absolute value of X_A-X_B	Rank	Signed Rank
1	0	0	0	-	0	-	-
2	0	0	0	-	0	-	-
3	0	0	0	-	0	-	-
4	0	0	0	-	0	-	-
5	0	0	0	-	0	-	-
6	1.167	1.000	-0.167	-1	0.167	1	-1
7	2.000	1.750	-0.250	-1	0.250	2	-2
8	1.333	1.000	-0.333	-1	0.333	4	-4
9	0.333	0.000	-0.333	-1	0.333	4	-4
10	0.833	0.500	-0.333	-1	0.333	4	-4
11	1.333	1.750	0.417	1	0.417	6	6
12	1.500	2.000	0.500	1	0.500	7	7
13	0.167	0.750	0.583	1	0.583	8	8
14	0.833	1.500	0.667	1	0.667	9	9
15	4.333	5.250	0.917	1	0.917	10	10
16	1.167	0.000	-1.167	-1	1.167	11	-11
17	0.000	1.250	1.250	1	1.250	12	12
18	2.833	4.250	1.417	1	1.417	13	13
19	1.500	0.000	-1.500	-1	1.500	14	-14
20	1.667	4.000	2.333	1	2.333	15	15
21	2.000	4.750	2.750	1	2.750	16	16
22	1.167	4.000	2.833	1	2.833	17	17
23	2.667	5.750	3.083	1	3.083	18	18
24	2.500	6.000	3.500	1	3.500	19	19
25	2.167	5.750	3.583	1	3.583	20	20
26	1.167	5.000	3.833	1	3.833	21	21
27	2.333	7.250	4.917	1	4.917	22	22
28	2.667	8.000	5.333	1	5.333	23	23
29	0.000	5.750	5.750	1	5.750	24	24
30	3.833	9.750	5.917	1	5.917	25	25
31	1.833	8.750	6.917	1	6.917	26	26
32	2.333	9.750	7.417	1	7.417	27	27
33	1.500	11.500	10.000	1	10.000	28	28
34	7.333	22.000	14.667	1	14.667	29	29
35	0.833	19.750	18.917	1	18.917	30	30
Sum of the Signed Rank							385

Table 10 shows the X_A , X_B , $X_A - X_B$, the absolute value and the sign of $X_A - X_B$ for all the users after they are ranked. Tied ranks have to be included where appropriate. For example, if two numbers with the same value ranked #5 and #6, they are given the average of these two ranks, which is 5.5.

Then, the sum of the signed ranks (W) is computed, which is equal 385 in this case. If we assume that users did not contribute more resources in the last four weeks, the expected value of W (μ_w) should be equal 0. The number of signed rank (N) is 30. For any particular value of N , the standard deviation of the sampling distribution of W (σ) is calculated through Formula 7. In this case σ was equal to 97.237.

$$\sigma = \sqrt{\frac{N(N+1)(2N+1)}{6}} \quad (7)$$

z -ratio is calculated through Formula 8. The correction for continuity is -0.5 when W is greater than μ_w and +0.5 when W is less than μ_w . In this case σ was equal to 3.954.

$$z - ratio = \frac{(W - \mu_w) \pm 0.5}{\sigma} \quad (8)$$

Finally, z -ratio is referred to the normal distribution table to get the corresponding level of statistical significance (Table 11). It can be seen that the result falls into the level of 0.0005, which shows that the hypothesis that the users contributed resources more actively in the last four weeks than they did in the first six weeks is true with statistical significance beyond 0.0005 level.

Table 11. Level of statistical significance

Level of statistical significance				
0.05	0.025	0.01	0.005	0.0005
z-ratio				
1.645	1.960	2.326	2.576	3.291

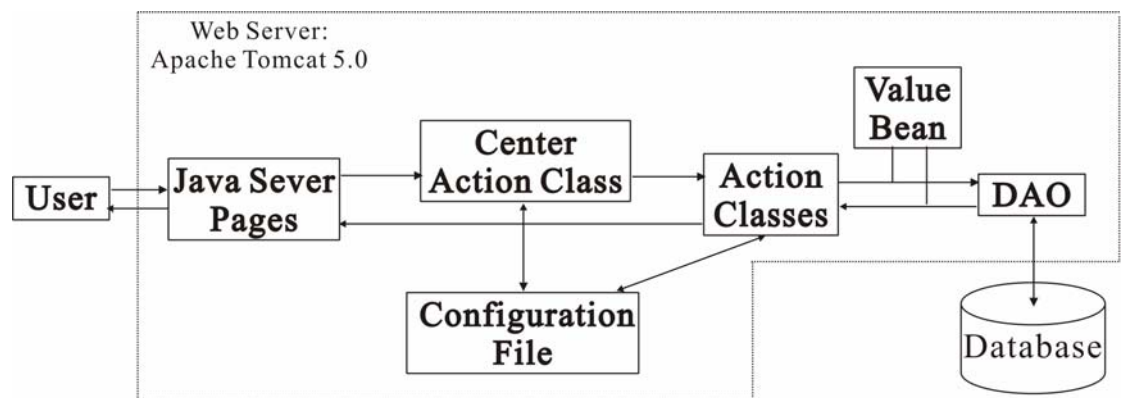
For more information about the Wilcoxon Signed-Rank test, please refer to Subchapter 12a of Concepts and Applications of Inferential Statistics, authored by Richard Lowry (1999). The book is available online at

<http://faculty.vassar.edu/lowry/webtext.html>.

Appendix D: Technical Details of the Web-based Comtella System

The Web-based Comtella system runs on Apache Tomcat Server 5.0 and MySQL Database 4.0. It was developed based on a Java Web-application development framework, WebWork (OpenSymphony 2004). Figure 19 shows the architecture of the whole system in component-level.

Figure 19. The component-level architecture of the Comtella system



The whole process of the user request being responded is as followed. The user interacts with Java Server Pages when she accesses the system. She enters on the pages the information needed for certain transaction she wants. Then the request from the user is forwarded to the central Action Class.

The central Action Class checks the user's identification and status and looks up the configuration file to dispatch the user's request to one of the Action Classes depending on what kind of transactions the user requests. Different Action Classes are responsible for different requests. The configuration file defines the mapping between each kind of user request and the Action Class that handles the request. It also defines which Java Server Page is used to display certain kind of results of the database operations (e.g. whether the database operation succeeds or not).

The Action Class retrieves the information needed for the transaction from the request, packs it into one single Value Bean and calls certain DAO class with the Value Bean as parameters. The DAO class performs certain database operation (e.g. searching, inserting, or updating) and returns the result back to the Action Class. Based on the result, the Action Class selects corresponding Java Server Page to display the result for the user by referring to the configuration file, which finishes one user request.

The Comtella system includes 60 Java Server Pages, 39 Action Classes, 17 DAO classes, 30 Value Beans and one configuration file. The database that supports the system includes 21 tables.

Appendix E: The Consent Form Used in the Study in 2004-2005 Winter Session

CONSENT FORM

Title of the Study:

Evaluation of the Effect of a Ratings-system and Community Visualization on ensuring a sustainable level of contributions in the Resource Sharing Environment COMTELLA

Researchers:

Julita Vassileva, Associate Professor, Computer Science Department; 966-2073

Ran Cheng, M.Sc. Student, Computer Science Department

Lingling Sun, M.Sc. Student, Computer Science Department

The aim of this study is to investigate the effect of a new ratings-system and a community visualization on ensuring a sustainable level of contributions and participation rates in a resource-sharing environment. Comtella is a web-based system developed at the MADMUC lab of the Computer Science Department, which allows a limited number of users to share links to articles that they have found with other users of the network. This particular experimental version of the system has, in addition to the main functionality, a virtual economy of ratings which results in different user status and an informative community visualization which is hoped to help motivate sustained user contributions.

Potential Benefits:

The possible benefit to the participants will be a more convenient access to class resources (papers, found on the Web). It will allow the users to make use of the search results of their peers, which can lead to a synergy in the class efforts to stay current in their knowledge of the literature related to the class topics.

Potential Risks:

It is hard to envisage any risks or side effects of the usage of the system. The papers that are shared are publicly available on the web, and we don't envisage copyright issues. If we become aware of any negative effects during the study, we will inform immediately the participants or interrupt the study. We may discontinue a participant's involvement in the study, if they use the environment to communicate links offensive, copyrighted or inappropriate materials. In the event of a participant withdrawing from the study, his / her data will be deleted and destroyed insofar as possible.

Collection and Storage of Data:

During the study, data will be collected about the users' actions related to sharing, rating,

and commenting links, as well as accessing the community visualization, for example the view of who is currently on-line, or the view of who shares papers in a particular area, or the view of who has contributed most new articles so far. This data will be correlated in anonymous form with user-participation data, in terms of number of new links found, links downloaded from other peers, number of rated and summarized papers. All data will be stored anonymously and will be available only to the investigators involved in the study. All data about the users will be stored securely for a minimum of five years, on a password-protected computer system, and any additional on-paper data will be kept securely locked in Dr. Vassileva's office.

Confidentiality:

The anonymity of the collected data and the identity the subjects will be completely protected. Only aggregate data will be reported in publications; the names and identities of the participants will not be published in any form.

Right to Withdraw:

Each participant is free to withdraw from the study at any time; this will not affect the participant's academic status or success in the class. In the event of a participant withdrawing from the study, his / her data related to the experiment will be deleted. The participants will be advised of any new information that may have a bearing on the participants' decision to continue in the study.

Questions:

If you have any questions concerning the study, please feel free to ask at any point; you are also free to contact the researchers at the numbers provided above if you have questions at a later time. Any questions regarding your rights as a participant may be addressed to the Behavioral Research Ethics Board through the Office of Research Services (966-2084), or through Julita Vassileva (966-2073). If you wish to acquire information on the results of the research once the study is completed, send a request to Julita Vassileva at jiv@cs.usask.ca.

Consent to Participate:

I have read and understood the description provided above; I have been provided with an opportunity to ask questions and my questions have been answered satisfactorily. I consent to participate in the study described above, understanding that I may withdraw this consent at any time. A copy of this consent form has been given to me for my records.

Signature:

Date

(Name of Participant)

Signature of Researcher

Appendix F: The Questionnaire Used in the Study in 2004-2005

Winter Session and the Results

The questionnaire for the test group and the results:

1. Why did you rate articles?
 - a. I enjoy rating others' articles. 20%
 - b. It is related to membership and I want to upgrade my membership. 26.7%
 - c. It allows earning Cpoints that I need to invest in my articles. 33.3%
 - d. Other reasons: _____ 20%
 - It's not much effort to rate, so why not?
 - I did not rate articles
 - All of the above, It can also prevent someone from reading a poor article or encourage someone to read a good article.

2. Did you try to earn and use Cpoints?
 - a. Yes 73.3%
 - b. No 26.7%

3. If your answer to the above question is "yes", what do you think about Cpoints?
 - a. They are useful to make my articles more visible to others 60%
 - b. They are not useful. Even though I invested some Cpoints in my articles, they did not attract much attention from others anyway. 13.3%
 - c. Other comments: 0%

4. If your answer to question 3 is "no", then the reason you did not use Cpoint is:
 - a. The Cpoint mechanism is too complicated to understand or use. 0%
 - b. I did not like to rate others' articles, so I did not get Cpoints. 0%
 - c. I do not care whether my articles can attract others' attention or not, so I usually did not invest may Cpoints in my posts. 0%
 - d. Other reasons: 26.7%
 - too complicated
 - While I agree with c, I find this a very distracting system of evaluating contributions. While the 'fake' mechanism is necessary with the degree of over lap or cross referencing available on the internet I do not think that having students rate other articles is beneficial as there are many biases which can affect this interaction.
 - I do not care whether my articles can attract others' attention

- do not like to rate others
5. Do you think the final earned rating of your posts (the sum of others' ratings for your posts) can fairly reflect the quality of your posts?
 - a. Yes 53.3%
 - b. No 46.7%
 6. Do you believe that you could manage to make your posts earn more ratings from others?
 - a. No 73.3%
 - b. Yes 26.7%
 7. If yes, how do you think one can succeed best in doing it?
 - a. By sharing really good articles 0%
 - b. By sharing really good articles and investing Cpoints in them 53.3%
 - c. By sharing papers early in the week. 13.3%
 - d. By viewing my own paper many times to increase its views counter
 8. Did you try to use any of these strategies?
 - a. No, I did not try to do that. 46.7%
 - b. Yes, I used: (please, list all that you used in order of frequency): 53.3%
 - By sharing really good articles By sharing really good articles and investing Cpoints in them and By sharing papers early in the week
 - Sometimes, I used to search good articles for the next week beforehand so that I can post those at the beginning. It is observed that at the start articles generally earn points. But the danger is if any article initially earns any negative points then it has sure chances to earn more lots of negative points.
 - share th article earlier and invest more Cpoints on the article
 - I tried to share good articles early in the week. I shared them early in the week so that I could share the articles I found. I would then add cpoints to articles that I thought were really good or that had started to receive positive ratings.
 - C-Points & Good Articles. Also a catchy title tends to bring people in.
 - Cpoint investment, sharing early, and view clicking
 - By sharing papers early in the week
 - Investing Cpoints
 9. When you tried to find an article to read or to write a summary, which kind of sorting functions in the "Search" window did you find was most useful?
 - a. The default sequence based on Cpoints and share time is good enough. 26.7%

- b. Sorting by paper title. 6.7%
- c. Sorting by earned rating. 13.3%
- d. Sorting by view times. 6.7%
- e. I did not use any of these sorting facilities. 46.7%

10. Which one in the following do you think is most important to the thriving of the Comtella system:

- a. Quality of the articles shared. 33.3%
- b. Quantity of the articles shared. 6.7%
- c. Quality and quantity are equally important to the system. 60%

11. How many articles did you read per week in average in Comtella?

Max – 25; Min – 2; Average – 11.47.

12. Do you think it is hard to find the interesting articles in the system?

- a. There were many links in the system and I had to spend much time filtering out the ones I was not interested in. 66.7%
- b. It is okay. 26.7%
- c. Easy to find good articles. 6.7%

13. Do you think Comtella’s mechanism to determine the users’ membership grades is fair or not? (please rank it in a 1 to 5 scale, 5 is excellent and 1 is very poor)

1,	2,	3,	4,	5
0%	26.7%	20%	33.3%	20%

14. How would you classify the typical weekly shared papers in quality categories?

High-quality	_____%
Medium-quality	_____%
Low-quality	_____%

The result of this question is listed in Table 8.

15. Did you sense the following connections in the Comtella system? (you can select more than one answer)

- a. When your posts were highly rated by others, the system would expect more articles from you in the next week. 46.7%
- b. If your posts did not get many ratings from others, the system would expect fewer articles from you in the next week. 13.3%
- c. If your ratings were usually different from the ratings given by others, the points and Cpoints you can get for subsequent rating would decrease. 0%
- d. If your ratings were usually similar to other ratings, the points and Cpoints you can earn for subsequent rating would increase. 13.3%

Did you sense any other connections: (please, explain)

- I actually didn't notice any of those connections and believe they

would be a confusing system to be used by a class. In an online community or solely online course I would expect the fewer/lower ratings = fewer articles expected to be a boon.

- When I turned into gold member I had to rate more articles, as well as contribute more.
- I really didn't pay attention to what the system was telling me. I'd see the indicators on the first page, the amount of ratings/Cpoints I had left and what not but I didn't pay particular attention to how it changed over the weeks.

16. Would you like to give any comments, criticism, or suggestions

a. About the community visualization?

- I think community visualization should not display our name in any form. Individual member can know his current status by email or any other intelligent way. This may change or influence the rating.
- The visualization can motivate the student to contribute more papers, but it also let the other students know who contribute the paper, so some of the student may rate the paper based on how well they know that guy rather than the quality of the paper
- I don't think that you should be able to see who submitted which articles until that week is over. This will help to eliminate persons from helping their friends out by rating their articles positively.
- Perhaps you should try to use something like trees where everyone begins on the same level, but height and breadth of the tree increase as your participation increases.
- Interface was very hard to use
- I think if you show the community visualization you may as well show who posted what articles.
- Fun to look at but what's the purpose?
- One thing I don't understand is why the names of the stars pop-up as alt text instead of being simply labeled underneath their respective star? Well, at least they sometimes come up in the alt text. Why not do it that way instead of making me hunt around? The color scale should also be made explicit ... is yellow better than red or is it better than gray?

b. About the membership mechanism?

- Honestly, I really don't like existing system. I have contributed the maximum number of articles and earned lots of good ratings and selected the best contributor but my membership status does not display this. At the final I got very good marks. After looking at my final marks I am happy but if I look back to my membership I am not happy.
- should make the evaluation algorithm known by all students
- The membership mechanism is good; it can motivate the student to contribute more paper by trying to get a higher membership.

- The current membership mechanism is designed to keep those who are currently on top on top. As long as you submit 3 or 4 decent articles and do all of your ratings you are unlikely to fall in membership.
 - I find the color levels interesting, but too inflexible as they reset each week. Something which reflects where you stood last week would have been appreciated. In order to self-track your progress.
 - I think it is a good weight of participation. However I feel there should be more emphasis on quality, and way less on quantity
 - What's a membership mechanism?
 - Maybe the community was too small to properly motivate me to improve my membership, so I don't think I'm qualified to comment on the mechanism as I never really experienced it.
- c. About Comtella in general?
- The intention to use Comtella for this class is very good and novel idea. I believe Comtella will turn out one of the best tool to share ideas and views in future. I also believe Comtella is still under its development stage - it will overcome all the initial difficulties in time. Hope for the best.
 - there's the potential for Comtella to be very useful, however, maybe not in an academic environment. seems as if it's more about the marks, if anything.
 - The Comtella in general is good, easy for the user to use.
 - The user name and password should not be displayed on login (in the address bar). The process id's should also be hidden since their display allows persons to hack their way into the community visualization when they don't have access to it.
 - I think that the reviews should have a scale of 1-10. Personally I struggled with assigning marks sometimes because the summary wasn't worth 100% but it was worth more than 80%. An alternative to this would be having two scales. One for summarizing the information accurately and one for the quality of the writing (grammar, flow, word use, etc.)
 - I think that the expectations for summaries and reviews need to be made clearer. I would also like to see the definition of article clarified since many members of the community seemed to have different opinions on what constituted an article.
 - While an interesting system I believe that the search functions need further refining and that there should be a chat relating to the top post of each previous week in the following so each person who reads the article can contribute to a discussion rather than relying on summaries and reviews of those summaries.
 - This system is not ready, we spent too much time discussing problems with Comtella rather than the topics of the class.

- used Comtella as little as possible and basically forfeit my participation marks because I disliked the system so much
- I think there are a few bugs in the system which should have been fixed especially when they were visible.
- Kind of difficult to use. Often times confusing. It could have been used for discussion topics in class a little more to make it feel a little more useful. As well there should be mechanism for relating membership to Comtella discussion, to encourage online debate.
- "The Comtella" will only ever be as useful as the members that use it.
- Okay, I have a couple things. First, I despise those little pop-up windows that give some extra information or let me upload a summary, etc. I understand the appeal to use them, but they are a horrible nuisance for the user. Second, there are too many 'modes' in Comtella. The principle artifact (i.e. articles) never changes so why must I go to a different tab to upload a summary for an article I just read? A different tab to review? Can this not all be aggregated under one page? It becomes such a chore to do anything and I think that really curtailed my enthusiasm for using the system.

The questionnaire for the control group and the results:

1. Why did you rate articles?

- a. I enjoy rating others' articles. 18.8%
- b. It is related to membership and I want to upgrade my membership. 31.3%
- c. Other reasons: _____ 50%
 - If ratings did not affect your level of status, then watch EVERYONE stop rating. This is very skewed this way. It tells me nothing.
 - it was suggested by the instructor although it is not required, rewarded or increase my membership
 - like to rate and membership
 - It not only matters to my own membership but it is also fair for others how contribute really good papers.
 - I felt that I ought to do it but given limited time it seemed more important to find articles to share than to rate other peoples' articles even though rating gives more points for less work. Others must have felt much the same way I did because few papers received any ratings, let alone enough to reliably measure quality.
 - Both of the above reasons apply for me. If I really like an article, I like to rate it up so that other people will check it out.
 - When I read an article, it made sense to rate it, since it takes just 2 seconds to rate.
 - earn cpoints and trying to find an interesting article to write a

summary for, and deciding if you're going to read it you might as well rate it.

2. Do you think the final earned rating of your posts (the sum of others' ratings for your posts) can fairly reflect the quality of your posts?
 - a. Yes 31.3%
 - b. No 68.8%

3. Do you believe that you could manage to make your posts earn more ratings from others?
 - a. No 37.5%
 - b. Yes 62.5%

4. If yes, how do you think one can succeed best in doing it?
 - a. By sharing really good articles 6.3%
 - b. By sharing papers early in the week. 56.3%
 - c. By viewing my own paper many times to increase its views counter 0%

5. Did you try to use any of these strategies?
 - a. No, I did not try to do that. 62.5%
 - b. Yes, I used: (please, list all that you used in order of frequency): 37.5
 - share early and good
 - viewing my own paper many times
 - share good
 - I figured other users would rate the articles based on if they liked them or not. how many times an article was viewed shouldn't have mattered towards how it was rated. (answer: a)
 - share early
 - Only submitted really good articles.
 - share early

6. When you tried to find an article to read or to write a summary, which kind of sorting functions in the "Search" window did you find was most useful?
 - a. The default sequence by share time is good enough. 18.8%
 - b. Sorting by paper title. 6.3%
 - c. Sorting by earned rating. 31.3%
 - d. Sorting by view times. 6.3%
 - e. I did not use any of these sorting facilities. 37.5%

7. Which one in the following do you think is most important to the thriving of the Comtella system:
 - a. Quality of the articles shared. 62.5%
 - b. Quantity of the articles shared. 6.3%
 - c. Quality and quantity are equally important to the system. 31.3%

8. How many articles did you read per week in average in Comtella?
Max – 20; Min – 2; Average – 8.38.
9. Do you think it is hard to find the interesting articles in the system?
- | | |
|---|-------|
| a. There were many links in the system and I had to spend much time filtering out the ones I was not interested in. | 37.5% |
| b. It is okay. | 50% |
| c. Easy to find good articles. | 12.5% |
10. Do you think Comtella's mechanism to determine the users' membership grades is fair or not? (please rank it in a 1 to 5 scale, 5 is excellent and 1 is very poor)
- | | | | | |
|----|-------|-------|-------|-------|
| 1, | 2, | 3, | 4, | 5 |
| 0% | 18.8% | 31.3% | 37.5% | 12.5% |
11. How would you classify the typical weekly shared papers in quality categories?
- | | |
|----------------|--------|
| High-quality | _____% |
| Medium-quality | _____% |
| Low-quality | _____% |

The result of this question is listed in Table 8.

12. Did you sense the following connections in the Comtella system? (you can select more than one answer)
- | | |
|--|-------|
| a. When your posts were highly rated by others, the system would expect more articles from you in the next week. | 25% |
| b. If your posts did not get many ratings from others, the system would expect fewer articles from you in the next week. | 12.5% |
| c. If your ratings were usually different from the ratings given by others, the points you can get for subsequent rating would decrease. | 6.3% |
| d. If your ratings were usually similar to other ratings, the points you can earn for subsequent rating would increase. | 12.5 |

Did you sense any other connections: (please, explain)

- If you contributed lots of papers or gave lots of ratings (regardless of quality), your score would go up. Getting good ratings also helped. If you used up your rating points one week, you got more of them the next week. I never noticed anything about system expectations.
- The ratings could largely effect what membership you were.
- I only noticed that the more articles you and if you use all your ratings, then your membership increases.

13. Would you like to give any comments, criticism, or suggestions
- | |
|---------------------------------------|
| a. About the community visualization? |
|---------------------------------------|

- There are SO many ways to present visualization of contribution. I have no idea why the system is so convoluted and bizarre. Take a look at forums and see how they rate members. Its easy and intuitive to read.
- it stimulates members to participate or contribute more, but doesn't represent quality of contribution
- The brightness of a star is not that apparent unless one is looking for it. Consider changing the way the brightness is expressed.
- Well, I think it is unfair to list all the links provided by one person because what if this student wants to cheat? I mean one student could view what links that his friends contributed and give good ratings to his friends. I don't know how many people would really go to do this but this could happen.
- I didn't like how it turned into a competition to see who could post the most links. I think it promoted posting garbage.
- It was interesting to look at but had the potential for malicious users to rate the submissions of someone they didn't like poorly. As far as I heard, someone in the group I wasn't in was having their submissions rated poorly by others in the class that didn't like them personally. The user(s) could see which papers that person submitted from the visualization.
- I don't think community visualization would be much of a use in a system when marks are not involved.
- It is a good idea and innovative.
- I think the community visualization was okay. It seemed that the stars would not change if someone received a lot of negative ratings. A lot of people posted a lot of poor articles and yet their star was large.
- It is ok for me.

b. About the membership mechanism?

- no comment
- -The ratio between the number of articles and the number of members is not balance. Therefore when articles submitted are much more than members, some of them are not read. -There is no requirement to read some number of article, and subjective opinions/judgments
- I cannot for the life of me understand the mechanism that boosted someone up to gold status in the first 1 and 1/2 weeks of the class using the system. I would like to know what that person did to get such an early boost.
- There should be more things to consider in terms of determining one's membership:
 - when those papers were contributed.
 - how many people other than the link provider finally summarize those papers?

- Most of people will summarize a paper that he/she is really interested in.
 - the ratings the user get (if the rating itself could be more fair...)
 - I think this needs some reworking. Some users were putting up massive amounts of articles to see their star get bigger. This had the potential to make other users who didn't submit 20+ articles "look bad". I think that the membership mechanism degraded the quality of articles submitted. Too many articles were submitted just to increase the quantity. This is easily seen with how many of the numerously submitted articles received poor ratings. There was no regard for quality.
 - I think point wise there should be some benefit for having highly rated papers.
 - I first the membership mechanism was interesting. As I had to keep working harder to keep my status, I become uninterested.
 - not quite care about it
- c. About Comtella in general?
- Comtella is so skewed because no one would use it in the real world. The fact that we get marks for this and is not voluntary just messes with it so much. How else would you get testing? That is a difficult question, and I unfortunately have no answers.
 - The system fails to identify the same article from the same url/address.- The system destructs text layout in summary section using .txt files
 - Comtella is a nice community system. Whether it is unique is debatable, as there are other more intuitive community systems on the web. I am not sure if computer science students are an adequate enough test bed, since most of us are busy, and we are more likely to find the system easier to use than computer illiterate people. I also find Comtella easily manipulated, like what some users do to earn a higher status.
 - Comtella is great. I think real time chat should be added. And ranking procedures should change. It should be based on quality.
 - Comtella is a good system. Sometimes it crashed when I tried to select summaries to review and I ended up getting three summaries to review in some week....there must be some poor other students didn't get two summary because i took the extra one. So I think the system should put a limit on how many summaries that one student could review in one week.
 - Too many persistent popup windows. Comtella usually had about a dozen windows open by the time I finished using it (especially for reviewing summaries), which made it difficult to find the page I wanted and it took me awhile to close all of them when I finished.
 - I really hated how many articles were posted. Right before break, there were way too many articles to sift through.

- I liked the system, however I think that using "get" instead of "post" was a grave security oversight. I was not please to see that my password was visible in the location bar as a parameter when I logged into the system.
- It's difficult to figure out. At first, you don't know what you're committing yourself when you click a link, like for reviewing summaries. Perhaps having the most recent posts from the discussion forum visible on the main page would increase awareness of the activity there, and increase participation in it.
- It is a very good system to read about and discuss various topics. We can get updated news and articles on what is happening in the world related to that topic.
- It works very well for Cmpt 408, however I think it might be useful sometimes to limit the number of links one person can post because I think you should give everyone the chance to share equally.
- There were numerous problems at the beginning of the term, that should have been ironed out ahead of time. I found it questionable that this program was more of a research project than a collaboration tool.
- Displaying the username and password in the address bar is a big security issue. Anyone could get someone's username and password, and login as them and mess around with their Comtella account.
- if it is used by academic, not matter membership whatever, I will use it

Appendix G: The Process of the Mann-Whitney Test

The Mann-Whitney test was used in the second study to test whether the users in the test group contributed more ratings than the users in the control group did during the nine-week study. First, I made sure that the case can meet all the preconditions of the test. The three preconditions that the Mann-Whitney test can be applied are listed below:

1. The two samples are randomly and independently drawn;
2. The dependent variable (in this case is the number of ratings) is intrinsically continuous, capable in principle, if not in practice, of producing measures carried out to the n^{th} decimal place;
3. The measures within the two samples have the properties of an ordinal scale of measurement, so that it is meaningful to speak of "greater than", "less than", and "equal to".

Obviously, all the three preconditions were met in this case. The users were randomly sorted into the two groups and the number of ratings a user may contribute during the study was intrinsically continuous and comparable. After making sure that the conditions were met, I followed the Mann-Whitney procedure to perform the test.

Let us assume that Sample A is the numbers of ratings contributed by the users in the test group and Sample B is the numbers of ratings contributed by the users in the control group. So the sizes of the two samples (n_A and n_B) are 15 and 16 respectively. Firstly, the two samples are assembled into one single set, the size of which (N) is 31. Then the numbers of the set are rank-ordered from highest (rank#1) to lowest (rank#31). Tied ranks have to be included where appropriate.

After that, the numbers, together with rankings, are returned to the original sample to which they belong. Then the sum of the 15 ranks in Sample A (T_A), the sum of the 16 ranks in Sample B (T_B) and the sum of the 31 ranks in both samples (T_{AB}) are calculated. T_A , T_B , and T_{AB} are equal to 196.5, 299.5 and 496 respectively. Table 12 lists the

numbers, the ranks, the sum of the ranks and the average of the ranks in the two groups.

Table 12. The numbers, the ranks, the sum and the average of the ranks in both groups

Test Group		Control Group		Both Group
Number	Rank	Number	Rank	
140	1	132	2	
128	3	67	11	
118	4	65	12	
116	5	52	13	
103	6	48	14	
95	7	44	15	
88	8.5	43	16	
88	8.5	41	17	
80	10	27	19	
40	18	24	20	
23	21	21	23	
22	22	12	25	
18	24	6	26.5	
5	28.5	6	26.5	
1	30	5	28.5	
		0	31	
Sum of ranks	196.5		299.5	496
Average of Ranks	13.1		18.7	16

The average of the ranks in both samples (A_{AB}) is equal to $\frac{N+1}{2}$, which is 16. If we assume that the users in both groups do not have difference in the activeness in rating articles, then the observed value of $T_A=196.5$ should belong to a sampling distribution whose mean (μ_A) is equal to $A_{AB} * n_A$, which is 240 and the observed value of $T_B=299.5$ should belong to a sampling distribution whose mean (μ_B) is equal to $A_{AB} * n_B$, which is 256. The sampling distributions for T_A and T_B both have the same standard deviation (σ), which is calculated through Formula 9. In our case, σ is equal to 24.90. z-ratio is calculated through Formula 10, which in our case was equal to 1.727.

$$\sigma = \sqrt{\frac{n_A n_B (N+1)}{12}} \quad (9)$$

$$z - ratio = -\frac{(T_A - \mu_A) + 0.5}{\sigma} = \frac{(T_B - \mu_B) - 0.5}{\sigma} \quad (10)$$

Finally, z-ratio is referred to the normal distribution table to get the corresponding level of statistical significance (Table 11). It can be seen that the result falls into the level of 0.05, which shows that the hypothesis that the users in the test group contributed more rating during the study is true with statistical significance beyond 0.05 level.

For more information about the Mann-Whitney test, please refer to Subchapter 11a of Concepts and Applications of Inferential Statistics, authored by Richard Lowry (1999). The book is available online at <http://faculty.vassar.edu/lowry/webtext.html>.