Tracking User Participation in a Large Scale Team Collaboration Environment

Dionysis Karaiskakis, Dimitris Kalles, Thanassis Hadzilacos

Hellenic Open University
Laboratory of Educational Material and Educational Methodology
Patras, Greece
{karaisk,kalles,thh }@eap.gr

Abstract. All students of the Hellenic Open University (HOU) attend undergraduate and postgraduate courses at a distance. The lack of a live academic community is reported by many as a drawback in their studies. Systematic exploitation of new communication and collaboration technologies is desirable in the HOU but cannot be imposed universally as the average student's IT competence level is relatively low. In this work we present the methodology for the development of an integrated communication environment in which collaboration spaces serving as open communities play a key role in user engagement in the whole communication environment. To track and evaluate user participation we propose analytic metrics which, when combined with our detailed knowledge of the internal workings of user groups, provide concrete evaluation of the community online activity.

Keywords: team collaboration, user participation, distance learning technologies

1 Introduction

The Hellenic Open University (HOU) provides education at a distance taking into consideration a tenet for the universal access of students to educational res ources. HOU is thus formally based on traditional practices (by mailing books and educational material, by encouraging students to personally communicate with their tutor, and by organizing a small number of student-tutor consulting sessions attendance in a small number of common advisory meetings per year). Thus, the use of new communication and collaboration technologies is not mandatory for students to complete their studies. Still, such technologies are being systematically used for publishing announcements and information of a general nature, and for providing basic supplementary electronic material and sources for further study.

Moving from a model where web technologies are used for publishing information to a model where such technologies constitute a basic working tool in the everyday life of at-a-distance-learning students is a huge undertaking, which addresses both

E. Tomadaki and P. Scott (Eds.): Innovative Approaches for Learning and Knowledge Sharing,
EC-TEL 2006 Workshops Proceedings, ISSN 1613-0073, p. 323-332, 2006.

technical and cultural issues. Both types of issues are closely linked to the diversity of the background of the students and of the tutors as well as the availability and ease of use of the underlying infrastructure.

As the only entry requirement of HOU students is the successful completion of high school studies, these students reflect the mean level of experience and competence in the use of electronic services in Greece which, to date, is not particularly high (2005: 59% of the population aged 25-54 has no basic computer skills [1]). This problem is aggravated in the uptake of collaboration or e-learning services, which demand the existence of a certain attitude by the users (beyond usage skills). Thus, planning for the development of electronic services should address the following problems:

- The need for universal access in services of *stratified* complexity (suitable for each team level in order for all to accept their use).
- The organizational aspects of scaling up in numbers and in complexity.

In this work we present aspects of our emerging methodology for designing the entire communication environment provided to the students and tutors as a supplementary service to help them in their everyday work.
The basic unit in HOU studies is the Thematic Unit (TU). One TU consists of one or more teaching groups (a tutor is assigned to each group, which must have at least 10 students, up to just over 30). Small TUs do exist with one tutor and just over 10 students. There are also some very large ones with about 1,250 students in over 40 groups. Currently ~200 TUs are offered and about 1,070 tutors are assigned to various groups in these TUs, encompassing in total about 28,000 students.
Collaboration spaces constitute a focal point in our environment. In those, users can engage in asynchronous communication, publishing content and opinions related to their work (content management and forum services). Given that access to these spaces is allowed for every student (and centrally managed) but that attendance and participation are by and large optional, these spaces function as emerging communities of practice.

Our aim is to define metrics to evaluate user participation in the communities. A comparative evaluation of the community online activity at the TU level will help us propose actions to promote user engagement and participation.
In particular, we explore aspects of a methodology for the quantitative and qualitative follow-up and evaluation of users' participation in combination with the participation of tutors who act as expert users providing advanced knowledge and guidance.

This rest of this paper is structured in five sections. Next, we offer a coarse description of the infrastructure. Following that, we elaborate on metrics for the role of the expert in communities of practice. We then analyze specific groups with respect to their comparative evaluation in terms of online collaboration and proceed to qualitative remarks on the impact of personal attitudes of tutors towards communication on the uptake of the collaboration infrastructure. We conclude by highlighting our research directions.

2 A High-Level Description of the Communication System

In HOU, a substantial part of the mandatory administrative procedures followed by students is done through a portal platform; a key example is the selection of TUs in which a student will be enrolled in the coming academic year.

Typically, such portal platforms do not support specialized services for educational purposes, thus paving the way for specialized LMS (Learning Management System) applications to be deployed. However, the latter tend to serve well advanced users only and are seldom harnessed to their potential.

Because of the (just) average level of IT literacy of students, the acceptance and exploitation of LMSs presents significant difficulties, when attempted at an almost universal scale. On the other hand, the exploitation of electronic services in organization and administration is more acceptable (experience in EU countries shows that the use of new technologies in the educational domain is first noticed for organizational purposes and later for educational ones [2]).

HOU tutors who manage to promote the emergence of student communities often rely on problem based learning as a constructivist learning instructional model [3] (even, subconsciously so). On the other hand the lack of a vibrant academic community in HOU constitutes an important problem for the students; in that respect HOU cannot match traditional campus-based universities. A high percentage of student drop out in HOU (at least, as far as the Informatics undergraduate program is concerned) is related to academic factors, especially a lack of confidence to pursue university-level studies and the perceived lack of adequate assistance (compared to what was initially expected) [4].

To address these needs, an integrated common communication environment was developed, based on a portal infrastructure. To-date it supports (see Figure 1) information services, content management services, and asynchronous team collaboration services, real time services and further education specific services.

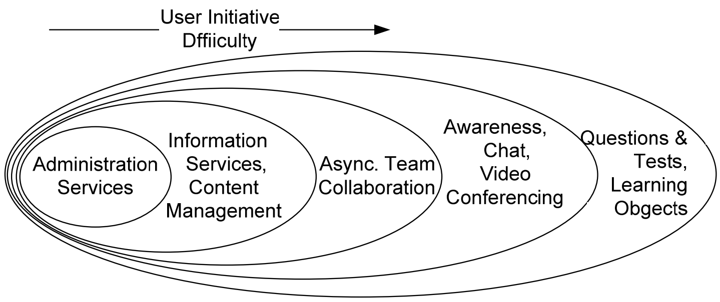

Figure 1: A hierarchy of services

All users and groups are updated in an LDAP server on an annual basis, with data drawn from the Student Registry MIS. Based on those user and group structures, working places were deployed for every TU, to support the communication and collaboration among students, with their group tutor, but also among tutors in the same TU. For each TU a content management space was created, along with a forum accessed by all TU members and a special forum accessed only by the TU tutors. In the collaboration spaces of large TUs additional spaces (*inner rooms*) were created to facilitate the private collaboration within one teaching sub-group (a tutor and all assigned students).

Videoconferencing services were initially provided by an independent application (with its own user and group management infrastructure). A new service has been installed and is now pilot tested to help users access and use the service in a seamless fashion, through the existing (unified) LDAP-based authentication scheme. The service provides video conferencing, chat and awareness services. Additionally, the (open source) Moodle LMS was installed and integrated; subsequently it has been extensively used by one TU to manage the submission and (automatic) grading of a large part of its homework assignments.

Note that all administrative services, content management, team collaboration spaces, teleconferencing and chatting services are hosted on different platforms but are all integrated through a common multi server Web Single Sign On domain to provide authentication. Figure 2 shows a high-level diagram of the overall infrastructure.

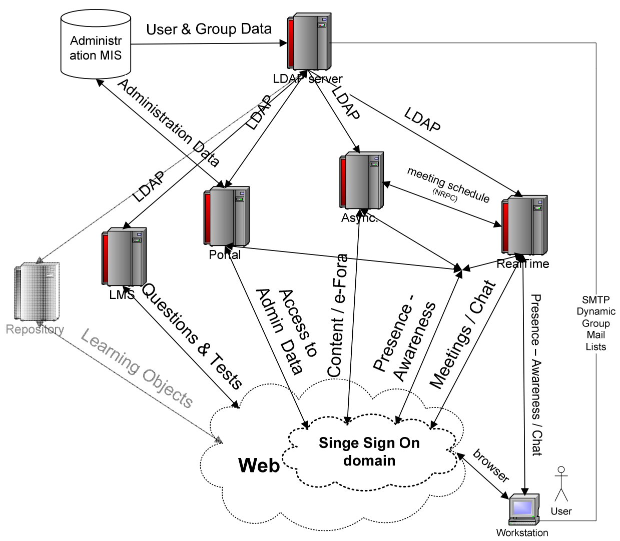

Figure 2: The server-services architecture

3. Measuring the Role of the Expert

We will start discussing some aspects of measuring the role of the expert by drawing on statistics generated by our platform. We will first introduce the concepts using a couple of examples before presenting the detailed results for all TUs.

Participation of group members is defined as the average number of visits per month per community member ($P_m = \Sigma\ V_n/n$), where a visit is defined as a sequence of successive page visits, with each page visit at most thirty minutes apart from the previous one.
While there is a substantial qualitative difference between passive and active user contribution in the community, we believe that such differentiation is only significant in the scope of individual user assessment [5]. When the focus is on the overall comperative evaluation of the community activity (as in our case), the total number of reads and posts is a sufficient metric.
Participation was examined in correlation with the activity of the expert (which is expressed as a percentage figure: *Exp_Activity* = *Exp_Visits* / 100* *All_Visits*).

For example, with reference to Figure 3, we note that the members of group G_{37} visit the workplace on average 20 times per month (roughly once per working day), whereas that rate is about 5 visits per month for the members of G_{188} (*y*-axis). A group index denotes the size of the group (as does the corresponding circle area). Furthermore, we also note that, within G_{188}, about 6% of its overall traffic was generated by the tutors whereas in G_{37}, this climbs up to about 9% (*x*-axis). Last, the dark filling of the G_{37} circle denotes a postgraduate group. At this point we urge the cautious reader to treat the above as a gentle introduction to the nomenclature and defer a comparative discussion (of groups G_{37} and G_{188}, among others) to Section 4.

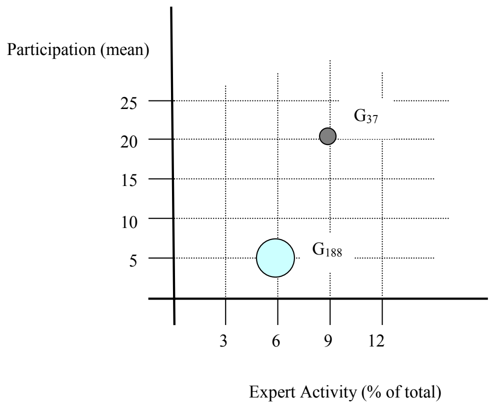

Figure 3: A measurement example

Figure 4 now shows the aggregate results. Data regarding an undergraduate program (consisting of 13 TUs) and an affiliated postgraduate program (5 TUs) were analyzed.

In 7 of those TUs the use of collaboration services was almost null and thus we analyzed the activity in the remaining 11 (6 undergraduate and 5 post graduate), accounting for a total of 2,086 engaged users.

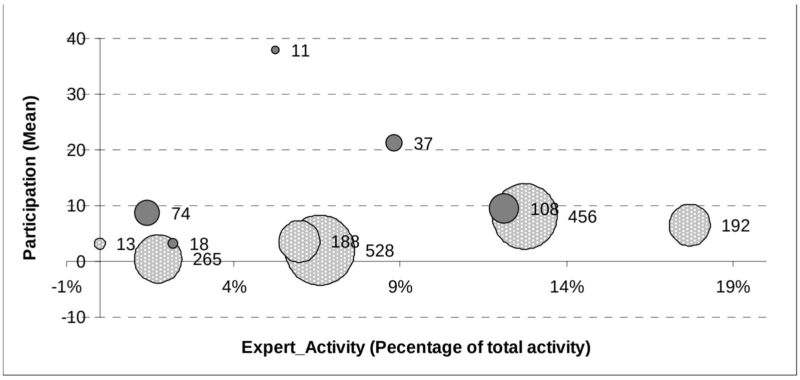

Figure 4: The measurement results

The distributions of visits within each group are not identical (not surprisingly). As a side-product we computed two standard statistical measures of these datasets, namely *kurtosis* and *skewness*. *Kurtosis* as a metric for tail size in a distribution provides a way to estimate the homogeneousness in the distribution of participation in each group. We report the *kurtosis*, in **Figure 5.**

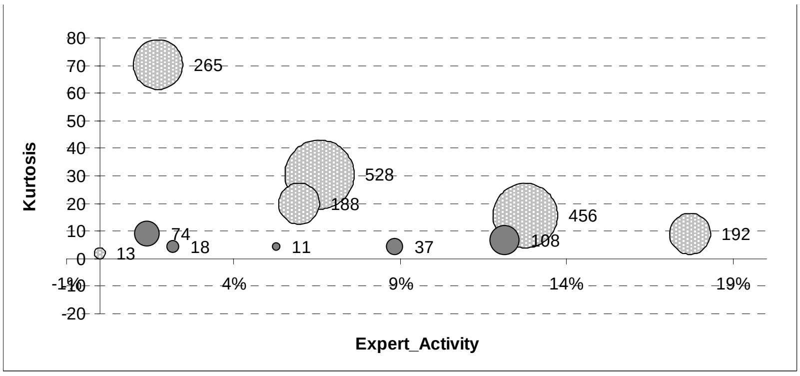

Figure 5: Data set *kurtosis* – small numbers indicate more even distributions

Skewness provides a direct way to estimate the relation between the number of users who are strong participators and those who are not. In all cases Skewness is positive, (ranging from 2 to 7) meaning that very active members are significantly outnumbered by the less active ones (especially in undergraduate groups). The differentiation here between groups is less pronounced than in kurtosis case, suggesting that this pattern is traced in all groups.

4. Discussion vis-à-vis a Detailed Analysis per Group

Before we discuss the results, it is useful to remind the reader that the systematic recording and analysis of activity in these spaces directly aims at tracking characteristic access patterns and at depicting problematic situations or highlighting efficient models of operation. In a working place, interaction between all the members of teams is desirable, particularly so for students. The role, however, of the tutor may be decisive since he, as an expert among other members, may be able to also open up new subjects and not simply respond to questions. Encouragement and participation by an instructor helps a community form more readily [6].

The interpretation of the particular results is facilitated by the fact that we have a detailed knowledge of the internal workings of the reported groups. Such knowledge is easily diffused among people who regularly share their tutoring experiences.

There are several axes of interpretation, which we will attempt to follow. Some finding will be recurring and we urge the reader to interpret these as non-orthogonal indications of the dynamics that exist in group collaborations. At this stage of our research, we seek to strengthen these indications by pointing out the common issues wherever they may be detected.

We start by discussing groups G_{108}, G_{74}, G_{11}, G_{37} and G_{18} (with reference to Figure 4). These groups all refer to postgraduate modules; we enumerate them in the respective expected order that a student would enroll in them. The figure reflects a strong indication that increased tutor activity raises student participation but group size adversely affects such participation (which is not unexpected since it is difficult to mobilize all group individuals when working at a distance).

It is intriguing that G_{74} and G_{108} are relatively close in the respective student participation axis yet so far apart in the tutor activity axis. We believe this is because tutors in the G_{108} are consistently active in their workplace involvement, both in terms of communicating between them and with their groups. Frequent communication raises issues which, from time to time, transcend the boundaries of a discussion forum and may re-appear in a neighboring forum, generating new rounds of collaboration.

A further, subtler, reason is that the study module related with G_{108} is the first module that these postgraduates take. This instills a community culture and when these students move on to the study module related with G_{74}, they are highly (and recently) aware of the benefits of community collaboration and presence is reinforced even without tutor involvement. This also refers to committed students who enroll in those study modules at the same year; they seem to be able to easily spot a good practice and stick with it. We thus note the flow of benefits from a module to another.

Such flow is also apparent, yet more subtly so, when analyzing the apparent strong student involvement of (senior postgraduate) groups G_{11}, G_{18} and G_{37}. It might be tempting to compare G_{11} with G_{18} based on tutor involvement (undoubtedly, measurably apart) but subtler issues arise. It is interesting to note that G_{11} is a module with a heavy software project management component, where the successful carrying out of assignments sometimes dictates the collaboration between students. That those students were already aware of the benefits of workplace collaboration facilitated their electing of the workplace to communicate during assignments. Note that both G_{11} and G_{18} refers to one student group per module (and, hence, one tutor) and therefore there is no room for intra-tutor collaboration. This is in contrast to G_{37} where two tutors were involved in student tutoring and two further tutors are involved in developing educational material for the module, as well as communicating with the students as regards educational matters. So, a substantial part of the traffic generated by the tutor component of G_{37} does in fact refer to communication between tutors. In the G_{18} group, the tutor has not embraced workplace collaboration and, hence, the students have been consulting the workplace for relatively static information (for example, meeting dates and venues) and no academic discussions were made.

Summarizing the postgraduate case, a unifying theme seems to emerge. This theme is that having instilled a collaboration culture in earlier modules has been fundamental in sustaining student workplace involvement. It is reasonable to assert that we must invest as early as possible to educate the student population in workplace collaboration. Such indirect knowledge is only gained by example but is exploited in subsequent study years where tutors may ease their activity without a negative impact on student participation (allowing for obvious deviations in tutoring style); the system seems to have gained momentum. We note that the emergence of this common qualitative characteristic is best demonstrated by the kurtosis figure, which demonstrates that irrespective of tutor activity (after an initial investment), students' access of the workplace more closely resembles that of a normal distribution. Interestingly enough, the *kurtosis* figure also suggests that the postgraduate groups demonstrate a more balanced way of how they access the workplace.

We now turn to discuss groups G_{528}, G_{265}, G_{456}, G_{188}, G_{192} and G_{13}, which all refer to undergraduate modules (the first three ones being junior modules and the latter three being advanced modules). As observed in the postgraduate modules, the larger the module the smaller the student participation. However, in the undergraduate modules, which are on average substantially larger than the postgraduate ones, we also observe that the collaboration workplace is mostly frequented by tutors in advanced modules. The first year modules display erratic performance which can be also traced to their nature and educational content. For example, G_{265} is a mathematics foundation module where the near-zero student participation can be attributed to a number of factors. Most important and influential among these are, the lack of maturity in students' perception of the subject and of academic study requirements in general, as well as the limited know-how of students and tutors in collaboration technologies. That only 2 tutors (out of 25) engage in some collaboration activity is best captured, again, by the kurtosis figure, where that group is a clear outlier.

A similar behavior is also demonstrated by the G_{528} group which, again, contains students at the start of their academic path and contains informatics foundations subjects. From then on, two clearly different paths are obvious. The first refers to the G_{456} group. Students in that group have been typically exposed to the learning curve (in terms of academic and attitude requirements) demanded by the mathematics and informatics foundations and coupled with a strong tutor investment in collaborative technologies display the relative emergence of a collaboration culture (with a healthy kurtosis figure) even at such a relatively large group size.

It is most instructing to see that such a culture is readily harnessed by the G_{192} group which has a reasonable participation index that is based on the majority of the student members. However, this is not the case with the G_{188} group and we are considering the possibility that this may be linked to the educational content of that module. The module covers theoretical computer science and it may be argued that modules with a relatively strong mathematics component are less suitable for collaborative work.

5. Qualitative Issues in the Tutoring Communities of Practice

Since HOU communication is traditionally based on Email and telephone, attendance in the working places is not obligatory. In HOU, the tutor has a mainly supporting and advisory role. However, HOU students are in general professionals that do not easily engage in activities which do not carry a direct practical profit. The emergence and the evolution of the collaboration spaces of TUs as communities of practice is closely linked to how much these can satisfactorily address the real needs of their users.
We have noted several problems that may limit user engagement and participation:

- Access problems (lack of basic skills and/or adequate infrastructure).
- Lack of time (full-time or part-time employment and family matters may limit the availability of time to study to just some time-chunks during weekends).
- Lack of apparent activity in the collaboration space by others is aggravated by physical isolation [7].

In the previous section we offered some insight as to why some student groups seem to be more active than others. We will now slightly deviate from analyzing the above data based on numbers and will try to shed some more light into the qualitative aspects of why some groups seem to shun online collaboration. In doing so we again exploit our intimate knowledge of the internal workings of those groups, however, we urge the cautious reader to note that no part of our analyses does in any way publicize individual data about any participant.
The starting point for our qualitative discussion is group G_{74}. It is very interesting to note that this group has a very low tutor activity because one of its most active tutors is strongly opposed to the use of collaboration technologies due to his strong preference of Email in the organization and carrying out of tutoring activities. This was, thus, a negative result.
How does one counter such a negative stance? The answer might lie within deploying a symmetrically strong opposition. Such behaviour was first spotted in group G_{108} (but not in this particular academic year that these results are based on). Specifically, one of the most active tutors was strongly opposing the deployment of the portal-based collaboration spaces due to his strong preference to a then-existing open-source

system for forum discussions. That opposition was unfortunately aggravated by several "teething" problems in the operation of the portal, at that time. It took a very focused and sustained contribution by at least one other tutor, in terms of generating fruitful discussions in the collaboration place forum, to establish a culture of actually using the collaboration place for further work (coupled, of course, with increased system availability). As the portal gained credibility and opposition grew smaller, it turned up that group participation was sustained even if fruitful discussions were now forthcoming at a more relaxed pace compared to the initial phase.

6. Further Work Directions

There are a number of limitations in our approach. For example, we know that a small number of sub-groups frequently engage in collaboration based on technologies that have not been integrated into our infrastructure, apart from email (text or voice) chat mechanisms or virtual classrooms. Such collaboration statistics are much more difficult to collect reliably and we believe that this (pessimistically) skews our results. Our recent infrastructure upgrade that allows chat and meeting sessions to be organized tightly integrated with the collaboration software will increase the seamless availability of such services to our academic community and will also boost our ability to collect essential usage statistics. After all, we hope to use our detailed knowledge of some modules to progressively refine our indices to also reflect as accurately as possible the situation in all other modules (currently at about 200), without requiring us to invest in understanding all of them. Not surprisingly, we are approaching the problem of the technology uptake in a rather conventional fashion, first trying several approaches on rather receptive users before applying the new concepts to more reluctant (subconsciously so) ones.

References

1. Eurostat, the Statistical Office of the European Communities. News Release 20/6/2006: The e-society in 2005, http://ec.europa.eu/eurostat/
2. Eurydice:The information network on education in Europe: 'Key Data on Information and Communication Technology in Schools in Europe 2004 Edition' http://www.eurydice.org
3. J.R. Savery, T.M. Duffy(1995). Problem based learning: An instructional model and its constructivist framework. *Educational Technology*, 35(5): 31-38.
4. M. Xenos, Ch. Pierrakeas and P. Pintelas (2002). A survey on student dropout rates and dropout causes concerning the students in the Course of Informatics of the Hellenic Open University. *Computers & Education* 39: 361–377.
5. Hazari, Sunil (2004) "Strategy for Assessment of Online Course Discussions", *Journal of Information Systems Education*, Vol. 15, No. 4. Winter 2004, pp. 349-356
6. R.E. Brown (2001). The process of community-building in distance learning classes. *Journal of Asynchronous Learning Networks*, 5(2).
7. M. Taplin (2000). Problem-based learning in distance education: Practitioners' beliefs about an action learning project. *Distance Education* 21(2): 284-307