



**HAL**  
open science

# Students' Tracking Data: an Approach for Efficiently Tracking Computer Mediated Communications in Distance Learning

Madeth May, Sébastien George, Patrick Prévôt

► **To cite this version:**

Madeth May, Sébastien George, Patrick Prévôt. Students' Tracking Data: an Approach for Efficiently Tracking Computer Mediated Communications in Distance Learning. International Conference on Advanced Learning Technologies, Jul 2008, Santander, Spain. p. 783-787. hal-00296796

**HAL Id: hal-00296796**

**<https://hal.science/hal-00296796>**

Submitted on 15 Jul 2008

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Students' Tracking Data: an Approach for Efficiently Tracking Computer Mediated Communications in Distance Learning

Madeth May, Sébastien George, Patrick Prévôt  
*LIESP Laboratory, INSA-Lyon, France*  
*La Doua, 21 Avenue Jean Capelle – Villeurbanne, F-69621*  
*{Madeth.May, Sebastien.George, Patrick.Prevot}@insa-lyon.fr*

### Abstract

*This paper presents an approach for closely observing the different levels of Human and Computer Interactions during student's communication activities on Computer Mediated Communication (CMC) tools. It focuses on how to efficiently track learners' activities, along with their outputs, and to exploit the collected tracking data in order to assist participants, both learners and teachers, in the learning process. Three experiments using the system and the result analysis are also presented.*

### 1. Introduction

According to Berge and Collins [1], Computer-Mediated Communication (CMC) promotes a type of interaction that is often lacking in the traditional teacher-based classroom. It allows learners the freedom to explore alternative pathways to find and to develop their own style of learning. CMC tools are also a part of distance learning environments. They are widely employed in distance learning platforms [2], to support communication activities between both learners and teachers. Discussion forums play an important role in knowledge sharing and in providing additional learning opportunities for students to exchange ideas for subject areas that are more discursive [3]. As Groeling [4] suggests, using discussion forums has the potential to improve teaching and learning experiences in traditional classrooms. As a means of providing more convenient support to learners and teachers, a research project involved in keeping track of learners' activities during the learning process is conducted. This research investigates the efficient ways to track students' activities and to exploit the collected tracking data to assist both students and teachers during and after their communications.

The rest of the paper is structured as follows. Related works are discussed in the second section. The third section is dedicated to a presentation of the research approach and its technical aspects. The fourth section presents some experiments, followed by the results analysis.

### 2. Related works and research issues

Tracking data of a user's activity can be referred to "traces of use", generally generated by a tracking system in accordance with its defined trace format or model [5]. The research work of Courtin and Talbo [6] focuses on the observation process of learners' activities. They introduced a model for trace analysis that allows teachers to understand different sequences of learners' activities. Després [7] presented a synchronous "Student Monitoring Tool" that observes the participation rate and the progress of each student during a learning session. In the research work of Komis *et al.* [8], the tracking process has been done through collaborative learning environments. They studied the traces of learners in order to provide the views on the communications exchanged between learners and instructors. Avouris *et al.*, proposed a set of Collaboration Analysis Tools [9] to analyze every step of the learner's activities within a collaborative learning session. Evidence of a similar approach can also be found in the research works of Riccardo and Dimitrova [10], where teachers could view three main aspects of learners which include social, behavioral and cognitive aspects. Donath *et al.*, [11] conducted one of the most interesting research works regarding the learners' tracking data visualization. This allows us to view student's communication activities and to evaluate the social aspect of each individual as well as groups of learners. Hardy *et al.*, [12] provided a set of "Learner Behavior Discovering" tools that allow learners to review their own behavioral aspects.

During the study of these related works, an important observation has been made. In the existing approaches, the tracking systems were designed to observe the user's activity on only the server side (e.g. where the communication platform is hosted); the user's interaction on the client side (e.g. user Web browser) is completely ignored. In this case, the granularity of traces is rather large and the information returned from the trace analysis might not be accurate enough to reflect the complete activities of users during their communications. Even with both interactions analysis methods [13] and content analysis methods [14], the quantitative and/or qualitative analysis can not be done efficiently if the recorded traces are not descriptive enough or when there is a lack of significant information inside the recorded traces (e.g. data indicators). In short, the root of the research issue has a higher degree of correlation to the collection of the tracking data.

An attempt is made to investigate "how to make tracking systems capable of producing traces that are not just a simple history of users' activities, but containing substantive information that can be useful to various users?". The research work described in this paper presents the observation of users' activities on both client and server sides, thus keeping track of (i) Human-Computer Interactions, (ii) Human-Human Interactions Mediated by Computer, and (iii) Computer action without user action. This allows us to have various compositions of traces with finer granularity in which user interactions and semantic aspects of the communications can both be found.

### 3. Proposed approach

#### 3.1. Conception

At the beginning of the study, three different levels of user's interactions in a CMC activity were distinguished which enables us to identify how to track different user's activities at each of the three levels. The observation process is performed at each level and the collected tracking data can be synchronized all together or merged into a single structure. The three levels of Human and Computer Interactions are:

**(1) The Human-Computer Interactions** which refers to the user's actions while using the Graphic User Interface of CMC tools. Using an example of an activity "Writing a new message", the interaction between a user and a CMC-tool interface can be "edit" message title or message content, "move" vertical scrollbars upward or downward, "drag & drop" smiles

into the message, etc. In Web-based applications, all of these actions occurred only on the user interface at the client side. In this study, we define the traces of user's activities by a composition of two parts of the traces: a part that represents activities on the server side, which is collected at the moment of exchanging user queries between client browser and server, and another part that represents the Human-Computer Interactions on the client side. The collected tracking data at this level are useful for rebuilding the successive processes of the past activity (e.g. what did a user do to write a new message) that allow us to identify the user's behavior while using a CMC-tool for the communications.

**(2) The Human-Human Interactions Mediated by Computer** refers to the content of the interaction exchanged between users. With the same example of "Writing a new message"; all the written text on the user interface will be submitted to the server via a server request query so that the message can be read by other users. Tracking the content of the communication makes the tracking data more descriptive, thus enabling us to know how a user interacts with other users and what the interaction content is about. The collected tracking data at this level will be exploited along with those at the first level. This makes it easier to reproduce not only the general context of the communication activity that describes the successive sequence of user's interactions, but also the content of the communication (i.e. the semantic aspect of the activity).

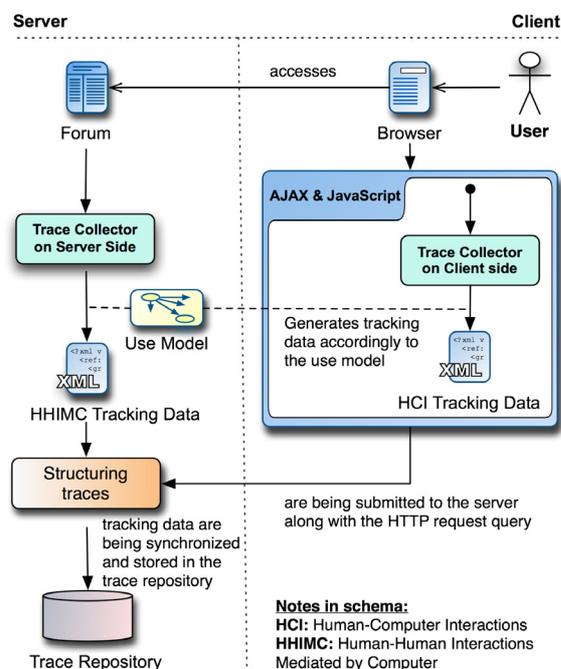
**(3) The computer action without user action:** during the user's communications there are many computer actions that occur automatically without the user action. Examples include a pop-up message indicating to the user that his/her session in the chat room will be expired in 5 minutes, or a jingle to alert that a new member has logged in to the forum, etc. Tracking such computer actions can be done on both the client and server sides. On the client side, we can capture events that showed up on the user interface, as on the server side, the events will be captured once the request query has been launched and executed. The collected tracking data at this level contain more information that describes what else happens during the users' interactions on the previous two levels.

#### 3.2. System design

We presented in detail the architecture of our tracking system for CMC tool and how each system component was built in [15]. The following is a brief description of the two main components of the system, the **use model** and the **observation component**.

**3.2.1. Use model.** Since there are a lot of CMC tools in Web-based learning environments, the wisest solution is not to build a tracking system for each single tool. The most appropriate solution is to study the common points (e.g. functional tools) and the particularities (e.g. different types of users' interactions) of each tool and to propose tracking system architecture, which is applicable to a variety of CMC tools. Therefore, we formalized the **use models** to describe the way users employ each tool to perform their communication activities. A use model enables us to (i) define the context of a user's activities and (ii) identify every **user action** on the **interaction objects** (i.e. user interface) and its associated **events** (i.e. what happens when there is an interaction). The main advantage of having use models for CMC activities is to make the tracking system able to take into account what and where to observe, when to capture, and what to generate as tracking data.

**3.2.2. Observation component.** The observation components were specifically designed with a number of "traces collectors", which ensure the observation of the user's interaction on the client side and the user's communications on the server side. Figure 1 shows the general architecture of the tracking system. To better understand how CMC activities are tracked, we give below example of a tracking process of an activity «Post a new message» in a discussion forum.



**Figure 1. Tracking system architecture on client and server sides.**

The user's interactions on their browser interface, such as "typing" a message, "drag & drop" smiles into the message, "moving" the scrollbar up or down will be captured by traces collectors on the client side. The tracking data is generated and temporarily stored on user's workstation. When the user clicks on the «Send» or «Submit» button, there is a HTTP server request query to submit the written message to the server (where the forum is hosted). The trace collector on the server side captures that request query and generates simultaneously the tracking data to represent the communication activity (i.e. "post a new message") as well as the content of the communication (i.e. written message). At each HTTP request, the temporary tracking data, previously stored on client workstations, is submitted to the server. These data are then synchronized with those on the server, and are structured and stored in the trace repository.

We developed the trace collector on the client side by using JavaScript language and AJAX (Asynchronous JavaScript And XML) technologies. It should be noted that we are not developing any spyware-type application and we do not need to install any tracking application on user machine. Using AJAX technologies enables the Web-based tracking system to be more flexible in terms of manipulating the tracking data directly on the client side. Thus, the generation and synchronization of the tracking data on both client and server sides are done in the background and without interrupting the users and CMC-tool interactions.

## 4. Experiments and results

In an attempt to observe how efficient the tracking system is and how the collected tracking data can help learners and teachers in distance learning situations, three semi-controlled experiment were conducted with the participation of students, teachers, and researchers from different disciplines.

### 4.1. Experiment set up

The Moodle Course Management System was used in the experiments. A random course was picked up by the research team and added to the Moodle platform so that the participants could carry out the discussions by concentrating on its contents. In order to make the discussions more relevant to the learning activities (e.g. the educational scenario that is implemented in the chosen course), a Contextual Forum, CONFOR [16] was integrated into the Moodle platform. Thanks to the tracking system architecture that is based on

“observation components” and “use models”, new “Traces Collectors” to CONFOR and Moodle were easily added. In each experiment, two different use scenarios were prepared to guide the participants in their communication activities. The first use scenario was distributed to the half of the participants, and the second one to the other half. The main idea of having two use scenarios was to make sure that a majority of participants performed various communication activities on different topics and in different orders. The experiments also aim to determine whether or not there will be two different communication behaviors corresponding to the two given scenarios.

#### 4.2. Results analysis

N.	Number of participants	Number of messages	Duration (minutes)	Number of records
1	8	62	30	1037
2	26	83	45	1399
3	26	147	45	2421

**Table 1. Information on the experiment set up and the records from the experiments.**

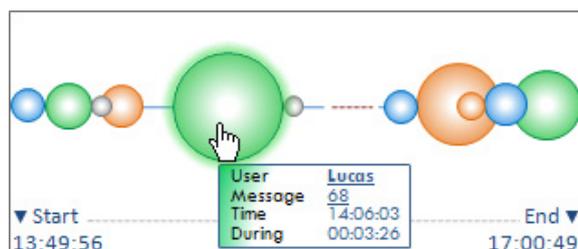
Table 1 gives us a summary on the experiment conditions and the number of records we have stored in the trace repository. We are particularly interested in analyzing the second and third case. As shown in table 1, there is a big difference for the number of records even the two experiments were conducted with the same number of participants and the same amount of time. By going through the tracking data, we found out that the participants in the third case have exchanged a lot of messages. Thus, there are a lot of traces for the communication activities such as “post new messages in the forum” or “reply messages in the forum”. A significant information that we found in the third experiment is that more than 80% of the participants did not really follow what described in the given use scenario. In addition, the extra communication activities that they have carried out were also tracked and that is why we came up at the end of the third experiment with 73% of extra records comparing to the second case.

In the second part of the result analysis, we are interested in exploiting the recorded traces. Apart of the traces exploitation is to extract from the traces the useful “data indicators”, with both quantitative and qualitative information that allows the participants in the learning process to analyze and/or to evaluate different aspects of their activities. We give below an example of exploiting students’ tracking data, inspired by the real needs of teachers in distance learning situations. The given example shows how CMC

tracking data are being analyzed in order to help the teachers evaluate student’s communication activities. Figure 2 shows a portion of the recorded traces for an activity “Display a message in the forum”. Figure 3 presents the data indicators for that activity.

HCI id	: 8096
User	: Lucas
Title	: Display a message in the forum
Attribute	: IDForum=4586, IDMsg=68
Date	: 19/06/2007
Time	: 14:06:03
Duration	: 00:03:26
HCI id	: 8097
User	: Lucas
Title	: Move scrollbar downward
Time	: 14:07:18
HCI id	: 8098
User	: Lucas
Title	: Move scrollbar downward (reach the bottom)
Time	: 14:08:49

**Figure 2. An example of records in the trace repository.**



**Figure 3. Data indicators for an activity “Reading messages in the forum”.**

“How do we know whether or not a displayed message is read?”. This question has been frequently asked, particularly by the teachers who regularly use discussion forums in their teaching activities. We are not pretending that we can prove if a message was really read by the user who displayed it, but there is a significant data indicator that can be used to justify that a message has not been entirely read. For example, if a user has only rapidly displayed the message (e.g. less than 3 seconds) without touching or moving the vertical scrollbar downward the bottom of the message, but performing another activity instead (e.g. clicked on another message), the displayed message must not have been entirely read by the user. Back to the recorded tracking data shown in figure 2, a user (Lucas) might have read till the end of the message since he has not only displayed the message, but also moved twice the vertical scrollbar downward and to the bottom of the message, and besides, he has spent 3 minutes and 26 seconds on it (i.e. the windows that displays the message has always been active right after

the message was displayed and the user has not performed another activity within 3 minutes and 26 seconds). Such information can be very substantial in automatic analysis and in helping the teachers look very closely at student's activity. As shown in figure 3, each sphere represents an activity of displaying a message and the diameter of the sphere is proportional to the time spent by each user reading the displayed message. The distance between two spheres represents the time gap between two different readings. The green sphere notifies us that the user read the message by having moved the vertical scrollbar downward to the end of the page (reading till the end of the message). The orange sphere indicates the fact that the user has simply displayed the message contents without moving the scrollbar. The blue sphere signifies that the user has displayed the message contents and has moved the vertical scrollbar downward, but not to the bottom of the page (i.e. partial reading). The grey sphere indicates that the message was only displayed; the window that displays the message has immediately been left inactive. Providing such useful data indicators is a lot better than giving only the statistical data like the number of hits on the message, or which user clicks on which message, etc.

## 5. Conclusions and future work

This paper presents an approach for efficiently tracking user's communication activities on CMC tools. The approach focuses on different levels of Human and Computer Interactions, where the tracking process must be carried out in order to collect as thoroughly as possible the necessary information about the CMC activities, including the content of the exchanged communications. The collected traces from the three levels of HCI are with finer granularity from which substantive data indicators can be easily extracted to reflect different aspects of the users' communication activities. We are still analyzing the recorded traces from the three experiments to look for other types of interaction indicators. Besides, we are planning to conduct more experiments with other CMC tools, in different learning situations, and to produce more significant data indicators for the participants in the learning process.

We are now developing a platform called TrAVis (Tracking Data Analysis and Visualization *tools*) for assisting both teachers and learners in their teaching and learning activities, particularly in the tasks of managing and exploiting the CMC tracking data.

## 6. References

- [1] Z. Berge, M. Collins, "Computer-Mediated Communication and the Online Classroom in Distance Learning", *Computer-Mediated Communication Magazine*, Vol 2, No. 4, 1995
- [2] S. Corich, Kinshuk, L. Hunt, "Assessing Discussion Forum Participation: In Search of Quality", *International Journal of Instructional Technology and Distance Learning*, 2004, TEIR Center, Duquesne University, Pittsburgh
- [3] M. Guzdial, J. Turns, "Effective Discussion Through a Computer-Mediated Anchored Forum", *Journal of the Learning Sciences*, Vol 9, No. 4, 2000, pp. 437-469
- [4] T. Groeling, *Virtual Discussion: Web-based Discussion Forums in Political Science*. Paper presented at the American Political Science Association, 1999, Atlanta, Georgia
- [5] C. Choquet, S. Iksal, "Modeling Tracks for the Model Driven Reengineering of a TEL System", *Journal of Interactive Learning Research (JILR)*, Vol. 18, No. 2, 2007, pp. 161-184
- [6] Courtin C., Talbot S. *Trace Analysis in Instrumented Collaborative Learning Environments*. IEEE ICALT, 2006, pp. 1036-1038
- [7] C. Després, *Synchronous Tutoring in Distance Learning*. AIED 2003, pp. 271-278
- [8] V. Komis, N. Avouris, C. Fidas, *Computer-supported Collaborative Concept Mapping: Study of Synchronous Peer Interaction*. *Education and Information Technologies*, 2002, pp. 69-188
- [9] N. Avouris, A. Dimitracopoulou, V. Komis, M. Margaritis, *Participatory Analysis of Synchronous Collaborative Problem Solving Using the OCAF Methodology and Tools*. CSCL 2003, pp. 232-234
- [10] M. Riccardo, V. Dimitrova, *CourseVis: Externalising Student Information to Facilitate Instructors in Distance Learning*. AIED, 2003, pp. 279-286
- [11] J. Donath, K. Karahalios, F.B. Viegas, *Visualizing Conversation*. *System Sciences*, Vol. 4, No. 4, 1999
- [12] J. Hardy, M. Antonioletti, S. Bates, *E-Learner Tracking: Tools for Discovering Learner Behavior*. *Proceedings of IASTED WBE*, 2004, pp. 458-463
- [13] F. Pozzi, S. Manca, D. Persico, L. Sarti, "A General Framework for Tracking and Analysing Learning Processes in Computer-Supported Collaborative Learning Environments". *Innovations in Education and Teaching International*, Vol. 44, No. 2, 2007, pp. 169-179
- [14] B. De Wever, T. Schellens, M. Valcke, H. Van Keer, "Content Analysis Schemes to Analyze Transcripts of Online Asynchronous Discussion Groups: A Review", *Computers & Education*, Vol. 46, No.1, 2006, pp. 6-28
- [15] M. May, S. George, P. Prévôt, *Tracking, Analyzing, and Visualizing Learners' Activities on Discussion Forums*. *IASTED WBE*, 2007, pp. 649-656
- [16] George, S., Labas, H. *E-learning Standards as a Basis for Contextual Forums Design*, *Computers in Human Behavior*, 2007. <http://www.sciencedirect.com>