User Model User-Adap Inter (2010) 20:455–494
DOI 10.1007/s11257-010-9081-5

ORIGINAL PAPER

Learners' navigation behavior identification based on trace analysis

Nabalia Bousbia · Issam Rebaï · Jean-Marc Labat · Amar Balla

Received: 28 September 2009 / Accepted in revised form: 13 October 2010 /
Published online: 4 November 2010

Abstract Identifying learners' behaviors and learning preferences or styles in a Web-based learning environment is crucial for organizing the tracking and specifying how and when assistance is needed. Moreover, it helps online course designers to adapt the learning material in a way that guarantees individualized learning, and helps learners to acquire meta-cognitive knowledge. The goal of this research is to identify learners' behaviors and learning styles automatically during training sessions, based on trace analysis. In this paper, we focus on the identification of learners' behaviors through our system: Indicators for the Deduction of Learning Styles. We shall first present our trace analysis approach. Then, we shall propose a 'navigation type' indicator to analyze learners' behaviors and we shall define a method for calculating it. To this end, we shall build a decision tree based on semantic assumptions and tests. To validate our approach, and improve the proposed calculation method, we shall present and discuss the results of two experiments that we conducted.

Keywords Navigation type · Indicator · Trace · Web behavior analysis · Educational Hypermedia System

N. Bousbia (✉) · A. Balla
LMCS, Ecole nationale Supérieure d'Informatique (ESI), BP 68M, 16270 Oued-Smar, Algiers, Algeria
e-mail: n_bousbia@esi.dz

A. Balla
e-mail: a_balla@esi.dz

I. Rebaï
Telecom Bretagne, Computer Science Department, Technopôle Brest-Iroise,
CS 83818 29238, Brest cedex 3, France
e-mail: issam.rebai@telecom-bretagne.eu

J.-M. Labat
LIP6, Pierre & Marie Curie University, 4 place Jussieu, 75252, Paris cedex 05, France
e-mail: jean-marc.labat@lip6.fr

1 Introduction

In Web-based learning environments, identifying useful information about learners is an important factor for planning tracking activity. Several studies have been conducted on the collection and the interpretation of information concerning the learner's activities, also known as *traces* or *observables*. The analysis of these traces, defined as a sequence of observed actions, provides "knowledge" about the activity, which we call *learning indicators*. The aim of defining such indicators is to infer high-level information about the learner from low-level data (URLs, clicks, etc.).

Most e-learning platforms provide typical indicators for assessing learner activity. Examples of indicators include the number of connections to the course, the time spent on it, the last part visited. Assessments are generally based on achieved scores. A survey with teachers involved in e-learning environments (Bousbia and Labat 2007), reveals that - although useful - this information needs to be summarized to obtain an overview of the learner's behavior and facilitate the data analysis process. Specifically, two questions were raised by teachers: "*What does the learner really do during an e-learning session?*" and "*What are his/her learning preferences?*". The first question relates to the learner's behavior. Answering this question helps the teacher perceive the learner's activity, as in traditional learning. The second question is about the student's learning style, defined as a set of behaviors and strategies regarding how to manage and organize information (Page-Lamarche 2004). Answering this question allows regulating and personalizing training. Consequently, answering these questions both helps the teachers and enables the Educational Hypermedia System (EHS) to adapt the learning content and to ensure individualized learning.

Several studies (Papanikolaou and Grigoriadou 2004) show the link between the learner's navigation behavior and his or her learning style in a hypermedia environment. Therefore, in this paper, we shall initially focus on answering the first question, i.e. the identification of the learner's behavior. The results will then be used to address the second question.

To identify the learner's behavior, we intend to produce informative behavioral indicators from traces in a Web-based learning context. Web-based e-learning courses are generally provided as online tutorials using Web pages. They do not necessarily contain evaluations, which are typically provided separately. Since the learner's behavior is similar to the navigational behavior of Internet users, we are interested in studies about Web browsing semantics. These studies generally use the navigation typology proposed by Canter et al. (1985), which is used to classify users' behavior. We propose to adapt this typology to our learning context using the "navigation type" indicator that describes how learners navigate through educational hypermedia resources.

The present paper aims at describing our system architecture and at validating the calculation process of the navigation type indicator, to analyze the learner's browsing path on Web-based learning. Learners work on activity-structured hypermedia courses available online in an e-learning platform.

The rest of the paper is organized as follows. Section 2 presents traces, indicators in EHS and Web browsing semantics. Section 3 describes our approach and our system architecture called IDLS (Indicators for the Deduction of Learning Styles). Section 4 presents the navigation typology that we propose and the calculation process of the navigation type indicator. Section 5 presents the results of the two experiments conducted to validate our approach. The first experiment involved 33 participants and produced 97 observations. The second one analyzes 387 observations produced by 116 students. Finally, Section 6 brings our paper to a conclusion.

2 Trace analysis and EHS

2.1 Trace

Over the last few years, several approaches dedicated to the observation of user activities within EHS have emerged. Each approach has adopted a different definition for what a "trace" is depending on its role and its utilization. In this paper, we consider the definition given by Settouti et al. (2006): "*The trace is defined as a temporal sequence of observed elements recorded from a user's interaction and navigation*".

Depending on the interaction environment, the data observed vary from simple mouse clicks, visited URLs and browsing time, to the recorded voice, chronological browsing within an activity, the use of communication services and answers to exercises and questionnaires. The selection of data for observation depends on observation objectives, such as characterizing the user's activity and analyzing behavior (Georgeon et al. 2006; Sanderson and Fisher 1994); interpreting the user's interaction with the system and with other e-learning actors (Siebra et al. 2005; Avouris et al. 2007; Heraud et al. 2004); identifying common behaviors among learners (Cheype 2006; Hofmann 2006), etc.

This variety of objectives and associated data types increases the need to help and assist trace users (teachers, researchers, etc.) in their production, exploitation, and analysis. Thus, trace processing is formalized, using Trace-Based Systems (TBS) (Settouti et al. 2006; Cram et al. 2007). A TBS is composed of several interconnected components. The most important ones are: the *collection system* and the *transformation system*.

The collection system captures the interactions from the observation sources and creates the trace. These sources are often based on the Web server logs, or on the use of specific software. The Web server log files have no information about the learner's activities on his or her machine outside this server, whereas the specific software saves only the trace of the interactions, conducted in a dedicated environment, in a proprietary format. Approaches based on the analysis of log files on the client side can overcome these disadvantages. They allow the collection of traces even in heterogeneous systems. The resulting traces provide information, not only on the learner's activity within the course or the e-learning platform, but also on the inactivity time, requested by the teachers to fulfill the tracking task. A trace model is required for defining what is traced (defining *a priori* which interactions will be traced) or filter the entities observed (defining *a posteriori* what information should be kept). Among the trace models studied, the "Musette approach" (Champin and Prié 2002) proposes a trace representation in an alternate sequence of states and transitions, allowing the reconstruction of the browsing path. The meta-language UTL (Choquet and Iksal 2006)

Table 1 Examples of indicators

Aspect	Type	Examples
Environment	Individual	Categories of posts per learner (Chen 2006)
	Collaborative	Interaction level of a group (Schummer et al. 2005)
Considered characteristics	Cognitive	Average depth level of a discussion tree (Gerosa et al. 2005)
	Social	Group cohesion (Reffay and Lancieri 2006)
	Affective	Individual motivation (Reimann 2003)
Value types	Quantitative	Collaborative factor (Fessakis et al. 2004)
	Qualitative	Actor degree centrality (Martinez et al. 2003)
Interpretation level	High	Argumentation quality (Barros and Verdejo 2000)
	Intermediate	Average visit duration of a course page
	Low	Visit duration of a page

allows the definition of traces to be considered at the scenario level by specifying their source and utilization.

The transformation system is the TBS kernel. It allows getting information from the collected traces, according to the observation objective (displaying indicators to users, identify or regulate their behaviors, etc.).

The observation objective establishes whether and when the results are visible to the user (Ollagnier-Beldame and Mille 2007): (*i*) without visualization, e.g., TPAnglais (Talbot and Courtin 2007); (*ii*) visualization during the activity, e.g., Teamframes[1]; (*iii*) visualization after the activity, e.g., COLAT (Avouris et al. 2007). We are interested in the third case where we aim at displaying the indicators after the activity. First, we shall define and study the learning indicator concept.

2.2 Learning indicators

According to Dimitracopoulou et al. (2004): "*learning indicators are variables that indicate …the mode, the process or the 'quality' of the considered 'cognitive system' activity, the features or the quality of the interaction product, the mode or the quality of the collaboration*".

The values of the indicators are determined by the processing of one or more traces. Each indicator has a name, a definition, possibly a natural language description, a goal, a domain of values, a duration validity, and an interpretation, naturally depending on the context in which it is calculated.

The ICALT (Dimitracopoulou et al. 2004) and DPULS (Lejeune et al. 2005) projects proposed a large number of indicators (e.g., 32 indicators in the ICALT project). We have listed some indicators in Table 1.

Throughout our study, we have classified the indicators according to different aspects: *(i)* the context or environment of validity (*individual*, *collaborative*); *(ii)* the

[1] http://teamframes.epfl.ch/.

personal dimension described (*cognitive*, *social*, and *affective*); *(iii)* the type of values (*quantitative*, *qualitative*); *(iv)* and their interpretation level (*low*, having no meaning alone and generally deducted directly from the raw data; *intermediate* or composite; and *high-level*, with an interpretative value, often derived by a complex treatment of the traces). We mainly use this last classification (*low*, *intermediate*, *high*) to describe the indicators defined in this paper.

Most indicators described in the literature are low-level, quantitative, and social (e.g., participation rates (Dimitracopoulou and Bruillard 2006)). Few indicators analyze the cognitive aspects. This analysis requires expertise, and is time-demanding. For example, the cognitive indicator: 'identifying misconceptions of students in solving electricity exercises' has been the subject of a whole Ph.D. thesis (Michelet et al. 2006). It is hence useful to reuse indicators. Diagne (2009) proposes a Pattern of Reusable Indicators and an open multi-agent architecture named EMAGIIR that allows the reuse of indicators.

Concerning the automatic determination of learning styles, few studies exist in the literature (Popescu et al. 2007). According to Papanikolaou and Grigoriadou (2004), indicators studied to distinguish between learning styles are: *(i)* time indicators, *(ii)* performance indicators, and *(iii)* navigation indicators (number of visits, navigation structure, etc.).

We will focus on navigation indicators. We aim to produce a high-level navigation indicator describing the browsing behavior. In the following section we present a literature review concerning Web browsing semantics.

2.3 Web browsing semantics

The aim of Web browsing semantics is to provide a more detailed description of possible user paths on the Web. A relatively abundant literature deals with this research domain and analyzes the user's browsing path in several ways. Damez-Fontaine (2008); Beauvisage (2004), and Hernandez (2004) present an overview of this research. These studies can be classified in two categories.

The first category focuses on how a user builds a perception of the Web, based on the use of a particular navigation strategy. Several authors proposed typologies of navigation strategies. The work by Canter et al. (1985), which remains the most studied one, is based on the discovery of navigation patterns to infer user intentions in terms of information search. The authors proposed a typology that distinguishes five navigation strategies: *scanning*, *exploring*, *searching*, *browsing* and *wandering*. In the same vein, we can quote Gall and Hannafin (1994); Catledge and Pitkow (1985) and Baeza-Yates and Ribeiro-Neto (1999).

The second set of studies analyzes the web browsing path considering navigation aspects. Thuring et al. (1995) consider that the navigation has two aspects: direction (forward/backward) and distance (step/jump). Another approach proposed by Lawless and Kulikowich (1996) focuses on two criteria to differentiate the way users navigate: their interests and knowledge about the area covered by the document.

To analyze the web-browsing path, the authors use statistical and machine learning studies: Neural networks, Markov Chain (MC), Bayesian networks, based on navigation traces. Beauvisage (2004) classifies the Web browsing semantic studies that deal with the path analysis into two categories: *site-centric* and *user-centric*. Several approaches fall in the first category including the *WebSIFT* system (Cooley et al. 1999), *WUM* (Spiliopoulou et al. 1999), the studies of Acharyya and Ghosh (2003); Heer and Chi (2002), and Mroué and Caussanel (2006). These *site-centric* tools and methods often consider that the content is known. These approaches are interesting because they identify relevant navigation patterns. However, as mentioned by Beauvisage (2004), the *site-centric* view is partial and biased. A semantic path requires a user-focused methodology that takes into account the dynamics of use in the analysis. The second category includes a few studies Catledge and Pitkow (1985); Tauscher and Greenberg (1997); Cockburn and Mckenzie (2000); Beauvisage (2004), and Damez-Fontaine (2008).

Our review of the literature shows that many studies exist on Web browsing path analysis. In our research, we focus on navigation types, and in particular, we rely on the typology proposed by Canter et al. (1985), which we consider to be particularly representative of navigation strategies. With regards to navigation analysis, *server-centric* research suggests interesting methods, but needs to place the browsing path in the general context of a user's navigation. Therefore, we are more interested in a *user-centric* approach.

3 Approach

The literature review presented in Sect. 2, reveals that a great amount of research is currently being done on traces and their modeling, as well as on the extraction of learning indicators, based on these traces.

In our research, we aim to have indicators, or a TBS system, which automatically deduces the learners' behaviors and learning styles based on trace analysis (Bousbia et al. 2008). We have in mind a Web-based learning environment, such as an e-learning platform, which provides online educational content. The learning material does not necessarily contain assessment.

We shall focus on the digital traces, registered in log files, on the learner side. This approach allows us to trace all the learner's activities, even those made outside the learning system, i.e., learner interactions, during his/her browsing path, which are not necessarily prescribed by the content author. These interactions include navigation within the educational objects and between them, Web browsing, personal production, and any activity done in parallel. Note that it is for the learner to decide when the recording of his activities on her machine begins and ends. This recording defines a session.

To define the indicators describing learners' behaviors, we based our work on the results of two main studies. First, the analysis of feedback that teachers want to have about their students. We have conducted this analysis through a survey, which we carried out with teachers (Bousbia and Labat 2007). Second, the state of the art on digital traces, resulting from learning situations, as well as the work on the influence of learning styles on the learner's behavior in a digital environment (Chen and Liu 2008).

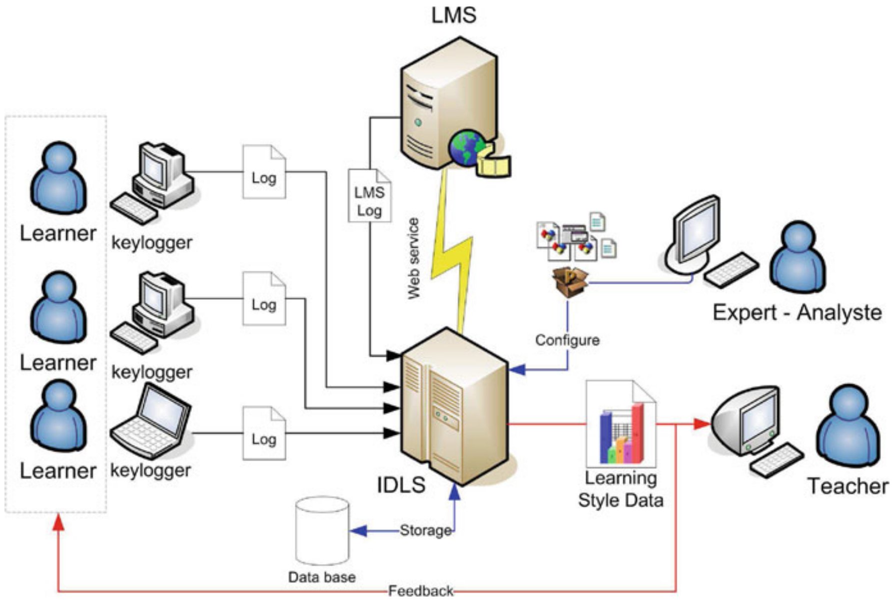

Fig. 1 IDLS architecture

To compute these indicators, we implemented a trace-based system called IDLS (Indicators for the Detection of Learning Styles). (Fig. 1)

On the learner's side, a *keylogger* (an application to collect traces) has been installed. It is activated by the learner at the beginning of the session. It records all the learner's interactions in a log file in XML format[2].

At the end of the session, the file is sent to the server to be stored and processed. The system can also analyze the log files, generated by the platform server, used by learners (LMS for Learning Management System).

To analyze and interpret the log files, IDLS has been installed on a server that can be independent from the LMS server. The system needs metadata about the educational content, obtained from the LMS via a Web service. It also requires the configuration data carried out during the system initialization. In fact, for the calculation of indicators, we need an "Indicator Dependence Graph (IDG)", supplied by an expert or a researcher in learners' trace analysis. This graph defines the dependencies between the indicators (*low*, *intermediate* and *high*) and designates their calculation methods. It also allows specifying the input or additional data[3] (e.g., thresholds) needed to calculate each indicator. The system execution allows deducing first the indicators and then the learning style values. These results are stored in a database. According to the receiver, they can be encapsulated in summary sheets for both learners and teachers.

[2] The traces are recorded for each learner between the times of the beginning and the end of the learning session. For each visited page, we store its URI (Uniform Resource Identifier), its order (first access) and all learner interactions (mouse action, keyboard, etc.).

[3] Data not recorded by the system during sessions, but needed to compute the indicators.

They can also be used to enrich the learner model in an LMS for adaptation, or used to implement an ITS (Intelligent Tutoring System).

In the following, we shall define the "navigation type" indicator that can describe learner navigational behavior.

4 Navigation type indicator

To describe the learner's behavior, it is necessary to interpret his or her navigation types. Research on the interpretation of Web navigation behavior often refers to the popular taxonomy defined by Canter et al. (1985):

- *Scanning*: seeking an overview of a theme (i.e. subpart of the hypermedia) by requesting an important proportion of its pages but without spending much time on them.
- *Browsing*: going wherever the data leads the navigator until catching an interest.
- *Exploring*: reading the viewed pages thoroughly.
- *Searching*: seeking for a particular document or information.
- *Wandering*: navigating in an unstructured fashion without any particular goal or strategy.

To identify a way to detect these types and check if all these values are necessary to describe the learner's behavior on Web-based learning, we carried out *preliminary tests*[4]. Through the analysis of these preliminary tests, we observed that learners generally go back and forth in the educational resources. We reached the same conclusion stated by Bidel et al. (2003), who proposed to merge *exploring* and *browsing* classes, which are fairly similar. Thus, as there is no consensus on a navigation strategy typology (Sect. 2.3), we propose a navigation typology from which the "navigation type" indicator is defined to describe a learner's navigation behavior.

4.1 Navigation typology

Based on Bidel et al. (2003), and preliminary tests, we adapted the Canter et al. (1985) typology in our learning context to propose a navigation typology with four navigation types (Bousbia et al. 2009):

- *Overviewing*: this value is close to the Canter "*scanning*" value. It implies that the learner is covering a large proportion of course pages. Through this phase of *fast-reading*, the user seeks to acquire an overall view of the course.
- *Flitting*: close to "*wandering*". It reflects a browsing activity without a strategy or a particular goal. The main difference with the *overviewing* type is the lack of focus on the course.

[4] Tests done on laboratory, on different e-learning courses, involving between 8 and 12 volunteer students and also by ourselves, in supervised and non supervised situations. Among these tests we can quote the one done to test the *Prox* indicator. It involved ten students, who worked on an English online course (Khatraoui 2008) in the same framework as the experiments presented in this paper. Other preliminary tests were performed, also in the same framework, to evaluate the other indicators, as the one done by Ait-Adda et al. (2008) with eight students.

- *Studying*: corresponds to a partial or complete reading of the course pages where the learner spends time on each page.
- *Deepening*: This describes a learner who spends relatively long time on a course, checking details, and seeking Web documents related to the course topics. The main difference with *studying* is the Web search part that the learner uses to obtain a deeper understanding of the course.

We have discussed this typology with students and they found it obvious. They recognized the various paths they take when navigating the course content and the different navigation behaviors. We have also discussed this typology with teachers who had a similar opinion. They recognized the way their students navigate the course. Moreover, they were able to classify the student navigation according to this typology, either when they observed the students in a lab or after examining the log files. However, getting such information is not easy because the user's navigation is obviously versatile. For example, a learner browsing a course may suddenly decide to look for a specific item of information.

Therefore, to identify the navigation behavior, Canter et al. (1985) proposed to use measures that characterize the graph associated with the navigation path. These characteristics, however, are not sufficient in our learning context. The behavior of any random user is different from that of a learner. Accordingly, the next section presents the intermediate indicators selected to calculate the navigation type high-level indicator.

4.2 Calculation of the intermediate indicators

To define the different types of the navigation typology described above, it is necessary to identify: *(i)* the visit duration of the course pages, used to distinguish a quick reading of the course pages from a complete one; *(ii)* the importance of the percentage of visited pages; *(iii)* the course browsing pattern, to distinguish an *overviewing* or *flitting* from searching activities; and *(iv)* the semantic link between the visited pages to know whether the learner focuses on the same topic or not.

Note that we consider a course to have a tree structure. The tree is composed of nodes, consisting of activities. These activities can include sub-activities. Leaf activities point to the resources (hypermedia content). We are particularly interested in HTML pages, thus, we will only consider the reading activities of a chapter, or a sub-chapter, as shown in Fig. 2.

The choice of this tree structure stems from the fact that it has been adopted by most online courses, and by the widely used SCORM[5] standard. However, having such a structure can influence learners' behaviors. To avoid this problem, we do not impose any scenario or visit order of the pages through the use of a course index. We mainly focus on navigation from page to page.

To calculate the navigation type indicator at the course interaction level, we propose the use of the following intermediate indicators: the average visit duration for a course page Dmc; the course consultation type Tc; the browsing pattern Fpc; and the

[5] http://www.adlnet.org.

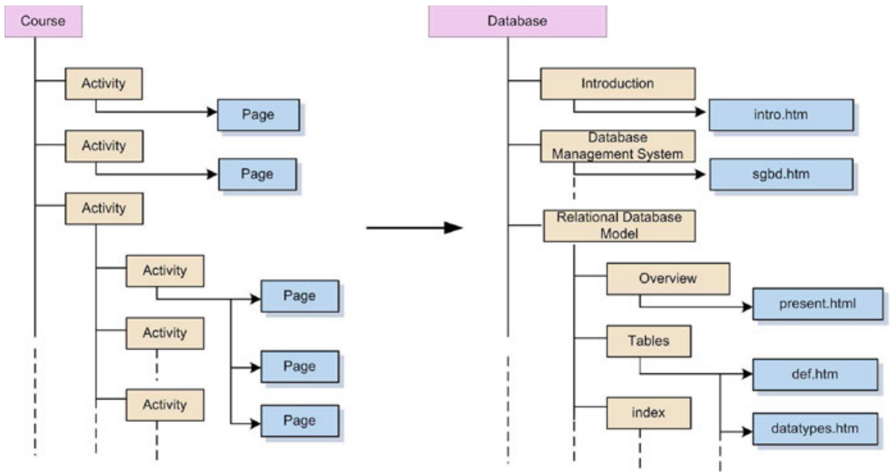

Fig. 2 The course tree structure

semantic proximity between the course and the content of the pages visited outside the course, $Prox$.

The following subsections present the proposed calculation method of each intermediate indicator.

4.2.1 Average visit duration for a course page

It is equal to the total visit duration of the course Dc, divided by the number of different viewed course pages N_{pdc}:

$$Dmc = Dc/N_{pdc} \tag{1}$$

To compute the total course visit duration Dc, we can proceed in two ways: summing the visit durations of the visited course pages, or summing the visit durations of the sub-activities of the course D_{Act_i}. We chose the second method since the calculation of intermediate indicators is based on durations of sub-activities. The course visit duration Dc is given by the formula (2).

$$Dc = \sum_{i=1}^{N} D_{Act_i} \tag{2}$$

$$D_{Act} = \sum_{i=1}^{K} D_{Node_i} \tag{3}$$

Dc: the course visit duration; D_{Act_i}: the visit duration of an activity Act_i; N: the number of sub-activities of the course; K: the number of the visited nodes of the activity Act_i; D_{Node_i}: the visit duration of a node $Node_i$ (a page or a sub-activity).

Note that the visit duration of a page is the effective duration. In other words, during calculation, we sum the durations between two actions and we subtract the non-activity time. Additional data indicates a period of time where the learner performs no action. This parameter can be set by the teacher or the course designer. The purpose of detecting inactivity is to distinguish, as much as possible, a learner who attempts to study and thus makes at least one mouse move, from the learner who merely opens the course window but then leaves. Nevertheless, in certain cases students do not make any action when reading a Web page. To deal with such cases, we use a similar principle to the one of chat tools regarding preferences about idle status. Although the default value for the additional data, indicating the maximum time of inactivity, is given, we allow the teacher or the course designer to change this setting depending on the specific course.

As teachers ask for informative indicators about their learners to help them in their tracking task, we qualified the average visit duration *Dmc* as *high*, if the value exceeds a threshold (also defined as additional data), or *low* otherwise.

4.2.2 Consultation type

As revealed by the survey analysis about feedback and indicators, teachers need to know whether learners deepen their knowledge about the visited parts of the course rather than going through them superficially. To address this, we proposed the "consultation type" indicator (Ait-Adda et al. 2008). Three values are assigned to this indicator: *superficial*, *intermediate* or *thorough*. However, it is necessary to take into account the course structure. In our case, we considered a tree structure. To calculate the consultation type at the course level, we first calculated its value at the leaf activity level, and then at the parent activity level, as explained below.

4.2.2.1 Calculation of the consultation type indicator at the leaf activity level It is the ratio of the number of consulted pages of the activity to the total number of pages associated to this activity:

$$Tc_a = N_{pca}/N_p \tag{4}$$

Tc_a is the consultation type indicator at the leaf activity level; N_{pca} is the number of visited pages of the activity; N_p is the number of pages of the activity.

4.2.2.2 Calculation of the consultation type indicator at the internal activity level To compute the consultation type indicator at a parent node, we computed the average of the values of its child nodes:

$$Tc_a = \sum_i Tc_{a_i}/N_a \tag{5}$$

Tc_{a_i} is the consultation type indicator value at an activity i; N_a is the number of child activities of the parent activity.

We proceeded in the same way to find the Tc value at the course level. At this level, we considered only the activities with at least one branch visited. The interpretation of this value is as follows:

- If ($Tc \geq 0.8$), the consultation is *thorough*;
- If ($0.5 \leq Tc < 0.8$) the consultation is *intermediate*;
- Otherwise ($Tc < 0.5$), the consultation is *superficial*.

These thresholds were selected after several preliminary tests in which we compared the consultation type realized with the computed quantitative value of Tc.

4.2.3 Browsing pattern

As mentioned earlier, to detect the navigation types, Canter et al. (1985) proposed using measures to characterize the graph associated with the navigation browsing path. Among these characteristics, it was deemed relevant to consider the sequential dimension of the browsing path. Formalizing the browsing sessions as sequences of nodes representing visited pages, Canter et al. (1985) distinguish four browsing patterns (Fig. 3):

- *Pathiness*: a path is a route that does not visit the same node twice;
- *Ringiness*: a ring is a route that returns to the starting node;
- *Loopiness*: a loop is a ring that does not contain other rings;
- *Spikiness*: a spike is a route that retraces the original path on the return journey.

The recognition of these four forms is far from trivial, especially the extraction of loops and rings including the nested ones can be complex. In addition, preliminary tests have shown that these forms are often mixed. Indeed, learners usually go to-and-from pages according to the learning task. Thus, we propose three forms related to our learning context (Fig. 4):

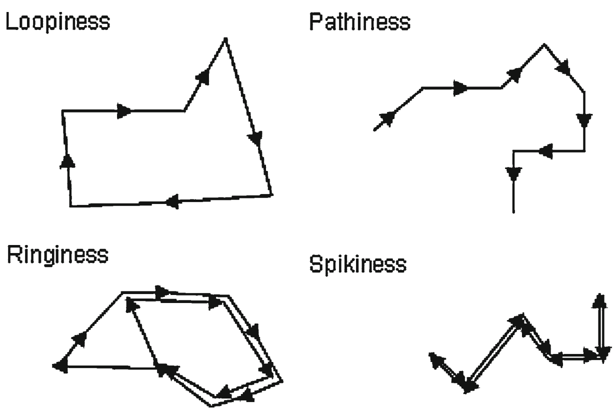

Fig. 3 Canter et al. (1985) browsing pattern

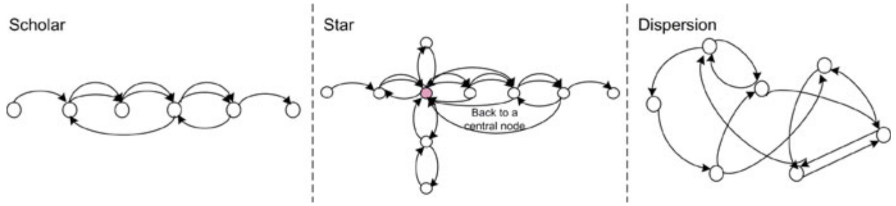

Fig. 4 IDLS browsing pattern

- *Scholar*: This value is retained when the dominant Canter pattern is a path with a few returns back (a few loops), or when there is a spike;
- *Star*: This form occurs when the learner tends to often return to the same node(s) (many loops to one or some nodes). It often appears in a search query;
- *Dispersion*: This is a ring or a mixture of Canter forms; in other words, learners tend to move in all directions.

To build this "Browsing Pattern" indicator Fp, it is necessary to filter the graph, depending on the interaction level considered so as to determine the indicators on which it depends. The course level is considered here.

In general, this indicator mainly depends on the consultation order, following the browsing graph analysis and the intermediate indicator TL or "Path Linearity". To compute TL, we calculated the ratio of the number of *different* consulted pages (number of nodes, N_{pdc}), to the number of consulted pages (number of steps, N_{pc}):

$$TL = N_{pdc}/N_{pc} \tag{6}$$

The number of steps is higher than or equal to the number of nodes. Thus, the value of TL belongs to the interval [0,1]. If the value is close to 1 (above 0.8), the browsing path is linear. Thus, the browsing pattern indicator takes the value "*Scholar*". For non linear paths, we developed an algorithm to find the dominant pattern by analyzing the existing forms in the sub-graphs. We first extract the graph sub-sequences corresponding to the four patterns of Canter et al. (1985). We also identified the number of nodes to which the learner often goes back to (Fig. 4). These nodes are called "*central nodes*". If the number of central nodes is lower than or equal to half of the sub-sequences, the browsing pattern indicator takes on the value "*Star*". Otherwise, if the percentage of steps in linear sub-sequences is the highest one, the browsing pattern is "*Scholar*". In all other cases, the browsing pattern indicator takes the value "*Dispersion*".

At the course level, we noted Fpc the browsing pattern indicator and Tlc the path linearity.

4.2.4 Semantic proximity

To understand what a learner has done outside the course, it is useful to detect whether the pages viewed outside the course are related to the subject of the course or not. To capture this information, we propose the $Prox$ or "proximity" indicator. For a given course, this indicator allows estimating the degree of proximity or similarity with the Web pages visited in the same time interval (i.e., visited between the time of connection and disconnection of the course, in other frames or windows). The calculation of $Prox$ is carried out using a dedicated sub-system. For this, we represent each visited Web page by the keywords assigned to it. The course is indexed by the teacher or the content designer and is represented by its keyword list. Note that the pages viewed by the learner outside the course are not necessarily indexed. Thus, we proposed to index the Web pages using a two stage process (Khatraoui et al. 2008):

- *Page rebuilding.* This phase aims at finding the complete text of each page visited by the learner. From the log file, the rebuilding is carried out by downloading the pages and by handling HTML tags to extract meaningful content.
- *Indexing.* The indexing system extracts the terms from each page and computes their weights by taking into account the term importance with regards to the tag in which it appears (see Appendix A). We then proceeded to the translation of the terms to English to use the original version of the WordNet[6] ontology. The terms were then converted into concepts by a projection onto WordNet. We opted for a general ontology because, pages on the Web, belong to different domains. As it is possible to find several meanings for a term, a disambiguation process examines the various meanings associated with the keyword and locates the closest sense to the context of the Web document, following the technique of Baziz et al. (2003). We chose this method because of its use of the WordNet ontology, which we also use.

The proximity indicator is calculated as the average of similarity values between the course and each Web page. It is given by the formula (7) (Khatraoui et al. 2008):

$$Prox(\text{Course}) = \frac{\sum_i D_i \times \text{Sim}(\text{Course}, P_i)}{\sum_i D_i} \tag{7}$$

$Course$: the course studied; P_i: a Web page visited at the same time as the course; D_i: the visit duration of the page P_i; Sim(Course, *Pi)*: the semantic similarity between the course and the page P_i. To compute this similarity we propose the formula (8) Khatraoui et al. (2008) inspired from the work of Varelas et al. (2005).

$$\text{Sim}(q, d) = \frac{\sum_i q_i \times d_j \times \text{sim}(i, j)}{\sum_i \sum_k q_i \times d_k} \tag{8}$$

$$j = \underset{k \in \{\text{conceptsof} d\}}{\text{argmax}} \ \text{sim}(i, k) \tag{9}$$

[6] http://www.cogsci.princeton.edu/~wn/.

i represents a concept of the document q (the course $Course$); j is the concept of the document d (the page P) having the maximum similarity with i (formula 9); q_i is the weight[7] of the concept i in the document q; d_j is the weight of the term j in the document d; $sim(i, j)$ is the semantic similarity between the two concepts i and j, calculated using the WordNet ontology. We chose Lin's measure (Lin 1998) due to the good performances that it provides.

As similarity measures between concepts/documents usually take values in the interval [0,1], the proximity indicator also takes values in this interval. It reaches the maximum value 1 when the concepts visited by the learner on the Web are identical to those of the studied course. The value 0 is reached otherwise.

To give an interpretative value to the teachers, we consider that the proximity is *high* when the $Prox$ value is greater than a given threshold, *low* otherwise. After several tests during which we checked the $Prox$ values of related and unrelated Web pages to the course, we have chosen 0.5 as default value for this threshold. However, as this indicator depends on the course indexing, we allow the possibility to change the threshold value. Associating navigation duration to the pages provides them with a weight, i.e., a page that has been visited for a short time will have less influence than another, which has been visited for a longer while. This is to highlight, for example, the pages visited as responses of search engines, but are irrelevant. Thus, we have added the "*null*" value to the $Prox$ indicator that has to be distinguished from the *low* value identified when visited pages are not semantically close to the course content. This "*null*" value is assigned, not only when no document is visited in parallel to the course, but when visited documents are pages viewed during a very short time (eliminated during the pretreatment before calculating $Prox$), or when these documents are LMS pages that should not be taken into account (home or index page of the e-learning platform).

4.2.5 Course duration rate Dc/Ds

Following preliminary tests, we noticed that some learners do not consult any Web page outside the course during their navigation. Therefore, we calculated the ratio of the course visit duration Dc compared to the session total duration Ds (time elapsed since the connection of the learner to the e-learning platform until she disconnects). When this ratio is very close to 0, it is needless to calculate the other indicators as the learner has not visited the course. Similarly, when the ratio Dc/Ds is close to 100%, it is needless to calculate the $Prox$ indicator because its value is *null*; the learner has not visited pages outside the course.

Dc/Ds can be "*low*" or "*high*" depending on whether its value is lower or higher than a given threshold. By default, the values below 50% are considered *low* and those above are considered *high*. We also consider the "*null*" value if this ratio tends to 0 by comparing its value with another threshold also specified as additional data. By default, it is equal to 5%. If Dc/Ds is less than this threshold it is considered "*null*".

[7] Weights are calculated using the TF.IDF method (see Appendix A).

The values of the Dc/Ds indicator are often proportional to the Dmc values (average visit duration of a course page), for example, when Dc/Ds is *low*, it is more likely to find Dmc *low*. These values can also explain the behavior when they are disproportional, for example, when Dc/Ds is *high* and Dmc is *low*, this indicates more likely an *overviewing* behavior.

4.3 Calculation of the navigation type indicator

In Sect. 4.2, we presented the calculation method of each intermediate indicator. These indicators are used to find the navigation type indicator value that classifies the learner's navigation behavior. Similar approaches (Bidel et al. 2003; Beauvisage 2004; Canter et al. 1985) are based on experimental studies on which are applied, depending on the context and the objective, statistical analysis or data mining techniques for the classification. In our context, we have a total of 5 indicators and 14 modalities, as shown in Table 2.

The interpretation of each behavior definition depends on the qualification we associate to intermediate indicators, not their numeric values. For example, the diagnosis that a learner overviews the course through a *fast-reading* depends on the interpretation that is assigned to the concept of "*fast-reading*". This interpretation can vary from one teacher to another and is related to the course content. The teachers' need of interpretative indicators motivates our choice of associating qualitative values to each indicator. Moreover, after setting default values to thresholds from which we get the indicator qualitative values, we allow the possibility to change these values according to the teacher's opinion. The most likely to be modified are the maximum duration of inactivity, thresholds for $Prox$ and Dmc. How can we predict the navigation type from these values?

Initially, no learning information was available to help identify navigation type indicator value for each possible combination of the intermediate indicator values. Thus, we have developed a set of rules summarized by the decision tree shown in Fig. 5. To improve these rules, we examined experimental data from preliminary tests. The *null* value of the Dc/Ds indicator is not included in the decision tree; because in this case the navigation type indicator cannot be deduced (the learner has not visited the course).

To understand the interpretation of this tree, consider an example. If the intermediate indicator calculation results for a learner's log file are: (Dmc: High, Dc/Ds:

Table 2 Indicator modalities

Indicator	Quantitavie values	Qualitative values (Modalities)
Dmc	R+	Low, high
Tc	[0,1]	Superficial, intermediate, thorough
Fpc	–	Scholar, star, dispersion
$Prox$	[0,1]	Null, low, high
Dc/Ds	[0,1]	Null, low, high

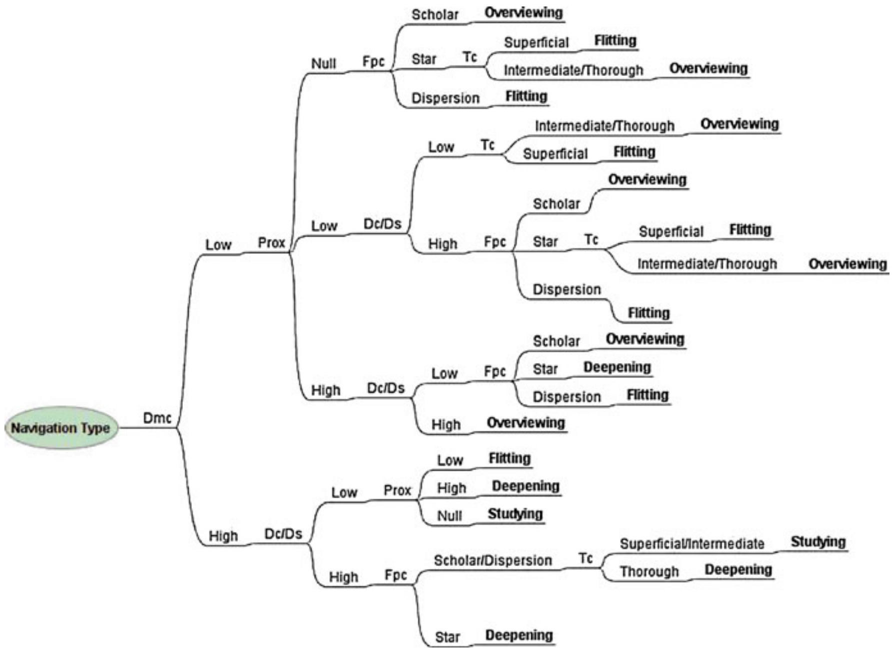

Fig. 5 Inferring the navigation type based on intermediate indicators

High, Fpc: Star, $Prox$: High, Tc: Thorough), the associated navigation type will be *Deepening*.

From this tree, we notice that the first indicator having a strong impact on the navigation type is the average visit duration Dmc. In general, when Dmc value is *high*, we infer a *Studying* or *Deepening* behavior. When Dmc is *high* with *low* values of Dc/Ds and $Prox$, we can infer that it is *Flitting*. We will present, in the next section, the results obtained during the tests.

5 Tests and results

In the previous sections, we presented our approach to infer the learner's navigation type from his/her browsing traces. In fact, we first tried to use the navigation typology proposed by Canter et al. (1985), but the observations of students' behaviors during preliminary tests led us to the navigation typology proposed in Sect. 4.

To validate our research hypothesis, that it is possible to find an algorithm that can classify the students' navigation according to this typology, we shall present the results obtained by conducting two main experiments. The aim of the first experiment was to validate the approach proposed. The last experiment aimed at evaluating and improving our model by comparing its results to data mining methods.

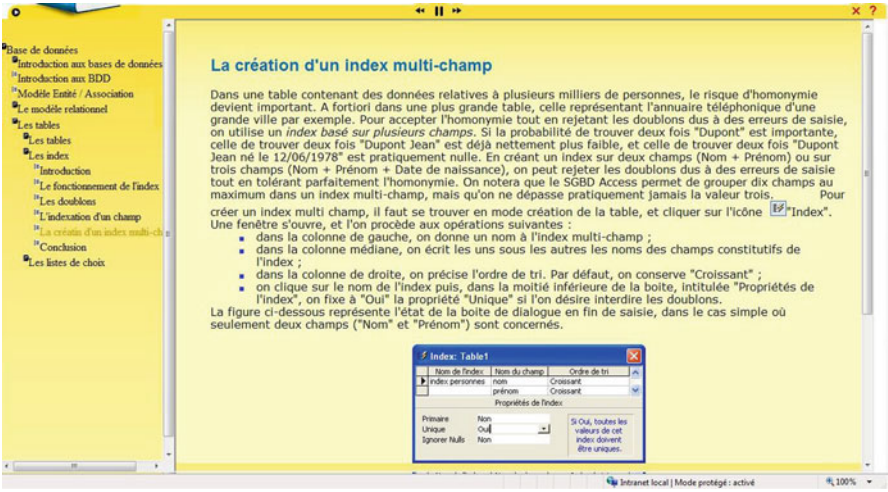

Fig. 6 A snapshot of the database course on the eFAD platform

5.1 Methodology

5.1.1 Context

To test our system and validate our approach, we carried out two experiments with two online courses that we designed in accordance with the SCORM standard (Sharable Content Object Reference Model). A SCORM course is a content aggregation of resources, from primary (assets or sco) to more important ones. The "organization" element is the root of the tree and each of its "items" is a pedagogical activity. This course organization is represented by its "Manifest", which also provides the course metadata. We choose SCORM for two reasons. First, the tree structure it employs is the most common in online courses and is compatible with our indicators (see Tc, Sect. 4.2.2). Second, SCORM requires a metadata file to include the course keywords that we need to calculate $Prox$.

The course content is in HTML format and mostly contains text. The consultation was carried out on the eFAD platform (Bousbia 2005), (Fig. 6). The runtime environment includes: (*i*) the course index, located in the left frame of the browser, offering free access to the resources (the learner can hide this index using the button located above the text); (*ii*) the platform navigation buttons in the top to visit the pages in a specific order; and (*iii*) the course content in the central frame. The navigation environment also includes the browser interaction objects (the navigation buttons, print, search, next, back, copy/past, etc.) using the keyboard and/or the mouse.

5.1.2 Procedure

The students in the experiment worked on computers equipped with a keylogger, with a personal account on the eFAD platform. The observation started at their connection to the e-learning platform after activation of the keylogger.

Table 3 Additional data

Max inactivity time threshold	Dmc threshold	Prox threshold	Dc/Ds threshold
3 mn	1 mn	0.5	50%

The experiments were conducted in four sessions, with the presence of a human observer (the experimenter) to ensure everything was being done correctly. In our case, the experimenter was the teacher. In each session, we asked students to adopt one of the four defined behaviors. This allows evaluating the four behaviors in the same course and with the same group of learners. Details about each experiment scenario will be provided later before giving its results.

As already mentioned, the indicator calculation requires additional data, including the thresholds for the discretization of the indicator values. But, how are thresholds chosen?

As stated earlier, the threshold values were set using preliminary tests. As the proposed qualitative indicator values can all be ordered (e.g., low, high), we first stored the indicator quantitative values, then we chose the value that split the sorted values. We also asked teachers to evaluate the interval chosen for each categorical value, which helped us to adjust the thresholds according to their interpretation. The chosen additional data are listed in Table 3.

5.1.3 Data analysis

We performed the trace analysis in 3 steps: first we checked whether the student behaved according to the scheduled sessions (manual analysis); then we made a statistical analysis (study validation); and finally we compared the prediction against the navigation type actually performed by the students.

– *Manual analysis*: The manual analysis consists of annotating each observation by the navigation type the student did. In fact, as we previously noted, the navigation behavior frequently changes and depends on several factors. Moreover, the human observer noticed that the behavior requested was not adopted by all the students. Thus, three teachers interpreted the navigation type from the log files according to the definitions given in Sect. 4.1. The teachers replayed the student actions according to his/her log file, in order to understand which behavior s/he adopted without knowing to which session the observation was scheduled. They observed the sequence of the student actions, the time spent for each action, the transitions from one page to another, and when necessary, the teachers visited the URLs of these pages to identify the behavior and label each observation (log file).
– *Validation of the study*: We checked: the discretization of the indicator values (*are the chosen thresholds optimal?*); and the distinction between the navigation types (*do these types correspond to the learners' navigation behavior?*). To answer these questions, we first calculated the intermediate indicator values, on which we

applied a statistical analysis using XLSTAT[8]. The study is presented in several stages.

To assess the chosen thresholds, for each indicator, we calculated the mean, the standard deviation, the maximum, the minimum and the median values to check the consistency of the chosen thresholds and the dispersion of the samples. These calculations were done on the *quantitative* (*numerical*) values of the indicators. For the browsing pattern indicator Fpc, as it does not take numerical values (see Table 2), we considered the quantitative values of the path linearity indicator Tlc, used in its calculation. The box plot diagram was used to give an overview of these characteristics and the sample dispersion. In this diagram, each box corresponds to the central part of the distribution of the indicator values (half of the values between the first and third quartile Q1 and Q3). The box plot extends to minimum and maximum values.

To verify the existence of differences between the groups, identified through the manual analysis, we carried out a factorial discriminant analysis (DA) on our indicator *qualitative* (*nominal*) values.

The purpose of this analysis is to study the relationship between a qualitative variable and a set of explanatory variables. In our case, the explanatory variables are the five intermediate indicators, while the variable to explain is the navigation type that is manually labeled on each observation. However, to implement such analyses, it is necessary that the variables have quantitative values. This condition is not satisfied because our indicators have qualitative values. To resolve this problem, we chose to use the DISQUAL method (Saporta 2006) because of its trustworthy results. DISQUAL is applied in two phases (Saporta 2006). First, we applied Multiple Correspondence Analysis (MCA) to replace the variable values by their coordinates on the factorial axes of the MCA. These axes are considered quantitative variables for the second stage, which consists of applying discriminant analysis to these new variables. The results of applying the DISQUAL method were presented using the group centroids to ensure the difference between groups. This statistical analysis was used further to compare and discuss its results to the results obtained with our model.

– *Automatic analysis*: After validating the study, the next step is to infer, through the IDLS system, the navigation type of each student. The navigation types are inferred from the values of intermediate indicators using the proposed decision tree (automatic analysis). To assess the effectiveness of the system, we compared the results of the manual analysis to those of our system.

In the next two subsections we will present two experiments that followed the above-described steps.

[8] http://www.xlstat.com.

5.2 First experiment

This first experiment involved 33 volunteers who were undergraduate students of the 3rd year of Information Systems at the National School of Computer Science (ESI-Algiers). These students belonged to the same age group (19–21 years). We analyzed their navigation traces on the course "Database", as part of their curriculum. However, they did not previously work with the platform used or visit the proposed hypermedia course. This course consisted of several activities composed of sub-activities (mainly chapters containing sections to be consulted). It contained 21 pages.

The experiment was held in 4 sessions of 15 min, in which we asked each student to adopt a different behavior respecting the following order: *Overviewing*, *Studying*, *Deepening*, and then *Flitting*. This order was chosen according to the findings made during the preliminary tests. We noticed that students began with *Overviewing* the course to discover its contents, and then explored and studied it, to go deeper into the topic. Finally they *flitted* when they got tired or started to get bored.

After processing the log files, we have noticed that the number of valid log files can be different from one session to another. In fact, in some cases, the logs were not generated at all or empty. This was due to the non-activation of the collection tool by participants, or the *null* value of Dc/Ds when the learners did not visit the course at all, particularly in the *Flitting* session. As a result, we obtained 29 observations for the *Overviewing* session, 28 for *Studying*, 27 for *Deepening* and 13 for *Flitting*: a total of 97 valid observations.

Following the data analysis methodology, we will first present the results of the manual analysis.

5.2.1 Manual analysis

After the manual analysis performed by three independent teachers, we got 77.32% joint-probability of agreement between the raters. We kept the 22.68% hesitation cases without discussing their disagreement in order to better analyze the results presented in Table 4. The rows represent the identified behaviors, while the columns represent the expected ones (scheduled sessions).

The table shows that, for the first session, *Overviewing* (1st column), only 17 students adopted this navigation type (1st row). Thus, only 58.62% of the students complied with the navigation constraint that we had imposed. Reading the three other columns in the same way, we obtain the proportion of students that complied with the behavior requested. We observe that most students did not comply with the behavior requested, in particular for *Deepening*, where only 4 students among the 27 adopted this behavior.

In addition, users often change their navigation strategy. For 22 observations, it was difficult to distinguish between *Overviewing* and the three other navigation types *Flitting*, *Studying* and *Deepening*. This is because the learners spent a period of time adopting one of the three behaviors, before or after *Overviewing* other course pages, for nearly the same laps of time in the two behaviors.

Table 4 Results of the manual analysis - 1st Experiment

Navigation type	Session				
	O	S	D	F	Total (Found behavior)
O: overviewing	**17**	8	3	0	28
S: studying	2	**15**	9	0	26
D: deepening	2	0	**4**	0	6
F: flitting	2	2	1	**10**	15
Overviewing/ flitting	4	1	4	3	12
Overviewing/studying	1	1	1	0	3
Overviewing/deepening	1	1	5	0	7
Total (expected behavior)	29	28	27	13	97

This perfectly reflects the difficulty of behavior classification when users frequently change navigation strategy. In order to explain why we did not find the expected behavior, we asked the participants and they confirmed that some students did not comply with the requested behavior.

This added noise to our data did not facilitate the behavior classification task, but improved ecological validity, which could lead to a more robust system in real use. However, before testing our model, the next section presents a statistical analysis that we made to check the proposed thresholds and classes.

5.2.2 Validation of the study

Following the procedure explained in Sect. 5.1.3, we first checked the threshold estimation and then we applied discriminant factorial analysis.

5.2.2.1 Threshold estimation After the manual analysis, we calculated the values of the five intermediate indicators. Descriptive statistics on the quantitative values of these indicators are presented using the box plot diagram in Fig. 7. We considered the values found for each class of behavior labeled manually, without considering the 22 disagreement cases.

Through this diagram, we observed that concerning the *Dmc* indicator, the mean value for the *Overviewing* and the *Flitting* behaviors (0.58 and 0.27, respectively) are lower than the *Dmc* threshold (1 mn), however those for the *Studying* and *Deepening* behaviors (1.72 and 1.09, respectively) are higher than the *Dmc* threshold. These results show that most observations, labeled as *Flitting* or *Overviewing*, have a *Dmc* value lower than the *Dmc* threshold, whereas those labeled as *Studying* or *Deepening* have a *Dmc* value higher than the *Dmc* threshold, which corresponds to the definitions we gave to these behaviors. By doing so for each indicator, in each class of behavior, we showed that the thresholds were adequately defined.

5.2.2.2 Validation of the classification To verify the existence of the groups found manually, we applied the DISQUAL method explained in Sect. 5.1.3. Figure 8 presents

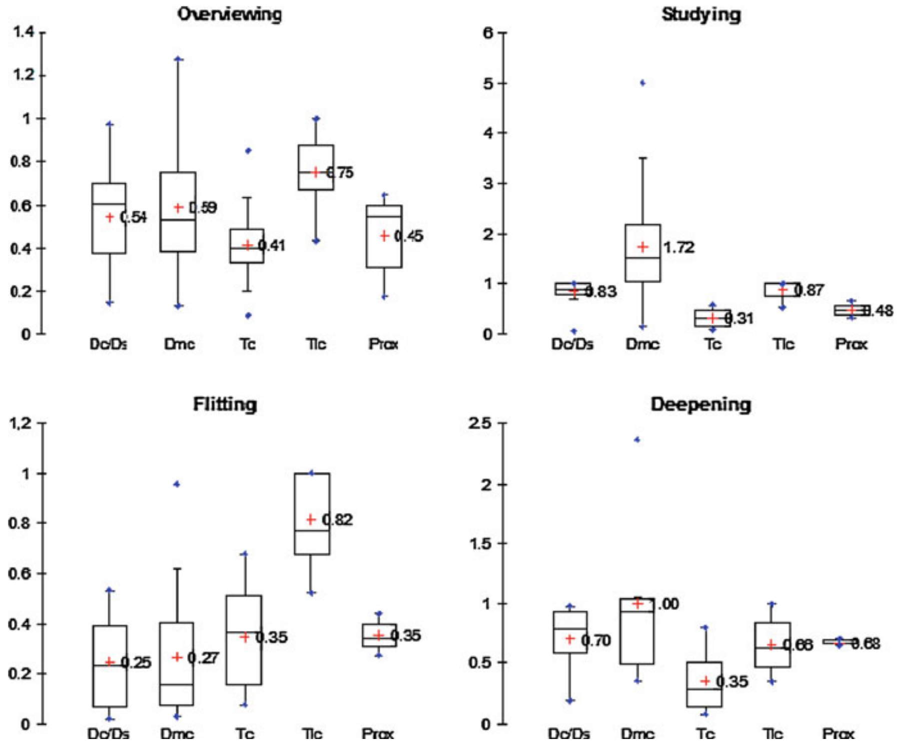

Fig. 7 Box plots—1st experiment

the centroids of the classes found using manual analysis, on the first two factorial axes of the DA, which allows us to account for 92.98% of the data.

We noted here that the four classes, identified in the manual analysis, do exist as their centroids are distinct and they do not intersect (bold circles). Concerning the hesitant cases, we observe that the ones found between *Overviewing* and *Flitting*, express rather *Overviewing* than *Flitting*. The observations classified between *Overviewing* and *Studying* and those between *Overviewing* and *Deepening*, intersect with the other classes. These findings confirm that the students adopted a different behavior in the session, leading to a different classification by the teachers, involved in the manual analysis. However, we observed that these hesitant cases are closer to *Overviewing* than the other classes. For this reason, we consider these hesitant cases as *Overviewing*. Finally, we noted that the DISQUAL method used here were accurate 70.10% of the time, without considering the 22 discordant cases.

5.2.3 Automatic analysis

To assess the effectiveness of our decision tree, described in Sect. 4.3, we first predicted the navigation class based on the values of the intermediate indicators computed using the IDLS system (automatic analysis). Then, we compared the results of the automatic and the manual analysis.

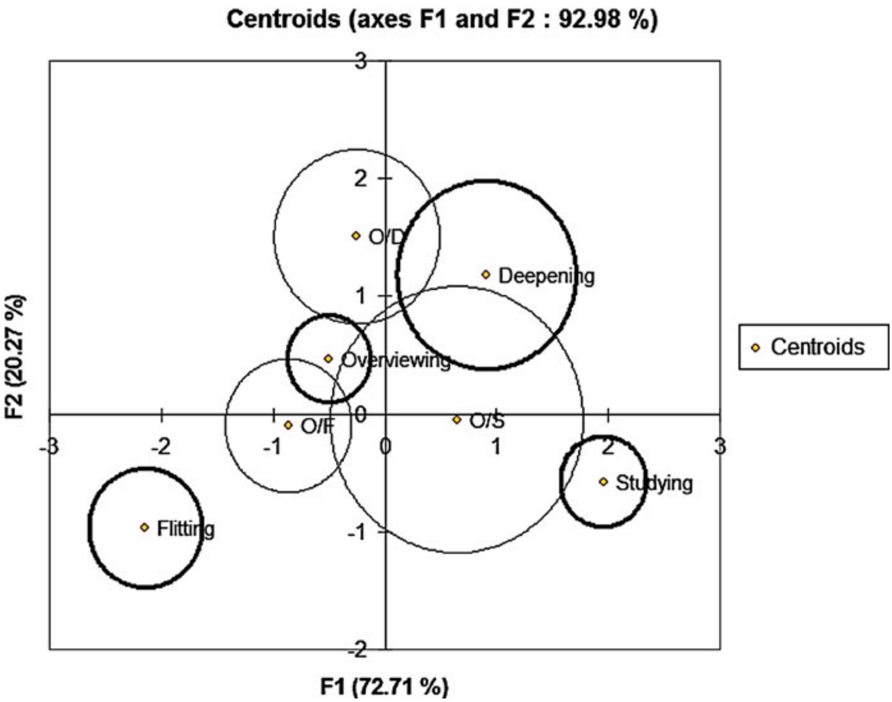

Fig. 8 Class centroids according to the DISQUAL—1st experiment

Table 5 provides a summary of the results. The main rows correspond to the four sessions. Each main row is sub-divided into secondary rows to take into account the results of the manual analysis and check if the automatic analysis correctly classifies the observations in the behavior found. Thus, the columns of the table are divided into two groups: the first group of columns corresponds to the navigation types, found in the manual analysis (see Table 6). The columns of the second group show, row by row, the classification result of the automatic analysis.

We noted here that most of the observations were identically classified by the two types of analysis. However, concerning the cases of uncertainty between *Overviewing* and the other behaviors, the system classified these observations as *Overviewing*, except one observation, which was classified as *Flitting*. This result corresponds to the findings of Fig. 8.

Comparing the number of observations, where we found the same behavior by manual and automatic analysis, we obtain the confusion matrix given in Table 6. The columns correspond to the results of the automatic analysis. The rows represent the results of the manual analysis.

The last column informs us about the proportion of observations, correctly classified by the automatic analysis over the four sessions, according to the manual analysis. For example, for the *Overviewing* navigation type (1st row), we have 28 observations interpreted as *Overviewing* (5th column), of which 23 were well-classified by the automatic analysis (1st column). Consequently, we obtained 82.14% of correct results

Table 5 Manual versus automatic analysis results—1st experiment

Session	Manual analysis							Automatic analysis			
	O	*S*	*D*	*F*	*O/F*	*O/S*	*O/D*	*O*	*S*	*D*	*F*
O: overviewing	**17**							**12**	1	1	3
		2							2		
			2							2	
				2							2
					4			4			
						1		1			
							1	1			
S: studying	8							8			
		15						1	**12**	2	
				2							2
					1			1			
						1		1			
							1	1			
D: deepening	3							**3**			
		9							**9**		
			4					2		**2**	
				1				1			
					4			3			1
						1		1			
							5	5			
F: flitting				**10**				4			**6**
					3			3			

Table 6 Recognition rate of analyzed behavior—1st experiment

Manual	Automatic					
	O	S	D	F	Total	% Correctly found
O: overviewing	**23**	1	1	3	28	82.14
S: studying	1	**23**	2	0	26	88.46
D: deepening	2	0	**4**	0	6	66.67
F: flitting	5	0	0	**10**	15	66.67
Total	31	24	7	13	75	80

(last column). Proceeding in the same way, we computed the system recall for each navigation type.

The sum of observations, correctly classified for each behavior, indicates that 60 log files have been correctly classified over the 75 log files clearly identified by the manual analysis (last column). Thus, we concluded that, during this experiment, the system identified the navigation type with 80% accuracy.

These results are promising in comparison to those of the discriminant analysis (70.10% in this experiment), and the state-of-the-art in the field of behavior interpretation. Indeed, as we mentioned above and found in the manual analysis, the behavior is expectedly variable. Our approach seeks to identify the dominant behavior (majority attitude), during the whole learning session. The goal is to provide a summary of the learners' behavior since the teacher is not interested in knowing all the details of the behavioral changes over a session.

Finally, this experiment validated our approach by confirming that we can, from low-level data, automatically detect high-level behavior. In order to improve our model, we carried out another experiment with a larger number of students.

5.3 Second experiment

For this second experiment, conducted in the same framework as the previous one, we designed a SCORM course about "Computer Security". The course contained 53 Web pages, and had two main chapters: "Threats" and "Vulnerability". To analyze the navigation behaviors, we used the first chapter, with 116 undergraduate student volunteers at the ESI School.

This experiment was also conducted in four sessions, each one dedicated to one of the four values of the navigation type indicator: *Overviewing*, *Deepening*, *Studying*, and *Flitting*. The *Deepening* session was done before the *Studying* session, so that the students could focus on the concept on which they were doing the search activity, and read it in the course before doing research on the Internet. Moreover, as we found that students had difficulty understanding the behaviors during the first experiment, we asked them questions for which the answers were in the course. The aim was to induce for each question one of the four behaviors in each session:

- *Overviewing session*: This session lasted 10 min. To simulate this behavior, we asked the students to extract "*the attack and threat nouns*", or "*extract the words in bold*". In other words, to browse quickly multiple pages to extract the required information without having to read them through completely. We began the experiment with the first question, but noticed that students took time reading the course to classify names. Therefore, we discarded these sessions and we opted for the second question.
- *Deepening session*: We planned a longer time for this session, 10–20 min, allowing learners to do research on the Internet after going deeper into the content of the course. We prepared three questions: "*Search on the course and the Web for the difference between a virus and a worm*", "*between SCAM, SPAM and mail bombing" or "the definition of phishing*". The second question was the most investigated. Some of the answers to these questions were provided in the course so that the students could delve deeper into the course concepts before searching the Web.
- *Studying session*: 20 min were spent on this session, during which we asked the students to read the first course chapter "Threats", and answer a multiple-choice questionnaire at the end of the session. The objective of this test was to help us in

the manual annotation. It allowed us to know if the learners have actually studied the course or not.

– *Flitting session*: We planned 10 min for this session. We did not ask questions, but explained the behavior expected, personalities or situations in which this behavior usually appears: "*you look at the course contents without focusing on its parts and at the same time you browse in the Internet consulting messages, other Web sites, not necessarily related to the course …you go back and forth between the course and Web pages*".

Theoretically, the user adopted the expected behavior for each session. However, in reality, things were not so simple. Even in this controlled experimental setting, some students claimed they had been distracted from the session instructions by their immediate interest for a course part, introducing periods such as in-depth reading (*Studying*), searching (*Deepening*) or *Overviewing*, among other behaviors. This perfectly reveals the difficulty of behavior classification, as users frequently change navigation strategy, even in the case of a unique and well-defined objective.

That is why we performed manual analysis as in the first experiment, and followed the same data analysis methodology. As we had a larger sample in this experiment than the previous one, we also used data mining algorithms to improve our approach and find a better way to identify the learner's navigation behavior.

5.3.1 Manual analysis

At the end of the experiment, we collected 387 log files: 98 for the *Overviewing* session, 113 for *Deepening*, 97 for *Studying*, and 79 for *Flitting*. We noted here that the session distribution is not uniform. This is due to the non-participation of all the students in all the sessions, which explains why the last session had the lowest number of observations. On the other hand, some participants' sessions were not used because the participants made a search on the Internet or visited Web sites without visiting the course, which led us to cancel these sessions. We also discarded the log files corresponding to the first question dedicated to the *Overviewing* session.

Moreover, as noted by the human observer and confirmed by the participants, the behavior expected for each session was not the one really adopted. That is why we first carried out a manual analysis to label each log file correctly. Our analysis was reiterated to eliminate all uncertainty cases, considering the results of the first experiment, the observer and participants' notes, and the test scores. Table 7 (as Table 4) summarizes the results of the manual analysis.

We can see that 68% of the participants performed the required behavior: the expected behavior was different from the one actually performed especially in the *Overviewing* session. In fact, since the course proposed was on "Computer Security", a topic that interests most of the students, the observer noted that some students read the pages before extracting the words in bold, following the scenario question. This result influenced the other sessions, where the participants who had already read parts of the course, performed *Overviewing* during these sessions, as shown in row 1.

Table 7 Results of the manual analysis—2nd experiment

Navigation type	Session				
	O	S	D	F	Total (found behavior)
O : overviewing	**62**	35	5	19	121
S : studying	34	**60**	19	4	117
D : deepening	2	2	**87**	2	93
F : flitting	0	0	2	**54**	56
Total (expected behavior)	98	97	113	79	387

5.3.2 Validation of the study

For this second experiment, we followed the same strategy as in the previous one. First, we tested the validity of the thresholds chosen. To this end, we conducted a statistical study based on the *numerical* values of the intermediate indicators presented using the box plots. As we had a large sample in this experiment, we also tested the validity of the threshold choice with a data mining method (C4.5 decision tree), also based on indicator *numerical* values. Second, we pursued the validation of the study, by applying the DISQUAL method on the indicator *nominal* values, on which our study is mainly based.

5.3.2.1 Threshold estimation Following the procedure explained in Sect. 5.1.3, we checked the threshold estimation. We calculated for each indicator the mean, the standard deviation, the maximum, the minimum, and the median values from the five intermediate indicators using their quantitative values. The box plot diagram, in Fig. 9, gives an overview of these characteristics, for each class of behavior labeled manually.

Through this diagram, we observed that the mean and median values (see Appendix B) are generally close to the chosen threshold (lower or higher as defined by the navigation type), except for the semantic proximity indicator $Prox$. The mean and the median calculated for this indicator $Prox$, for each navigation type or overall, were below the threshold 0.5 used in the first experiment. These small values of the proximity indicator stem from students spending more time in the course than outside. On the other hand, the calculation of this indicator depends on several factors, such as the indexing of the course and the pages visited, in addition to the disambiguation of indexed words on these pages, especially when they do not exist in the general Wordnet ontology. These noisy values of the $Prox$ indicator led us to redefine and assign the threshold 0.3 to distinguish the navigation types better since it is close to the medians and means, and is almost half of the maximum value recorded.

The choice of thresholds is also justified using data mining methods. Consequently, we first checked if the intermediate indicators selected are relevant features to identify the navigation type values using the Weka[9] data mining tool. We used the 10-fold cross validation method. We applied three of the search methods provided by

[9] http://www.cs.waikato.ac.nz/ml/weka.

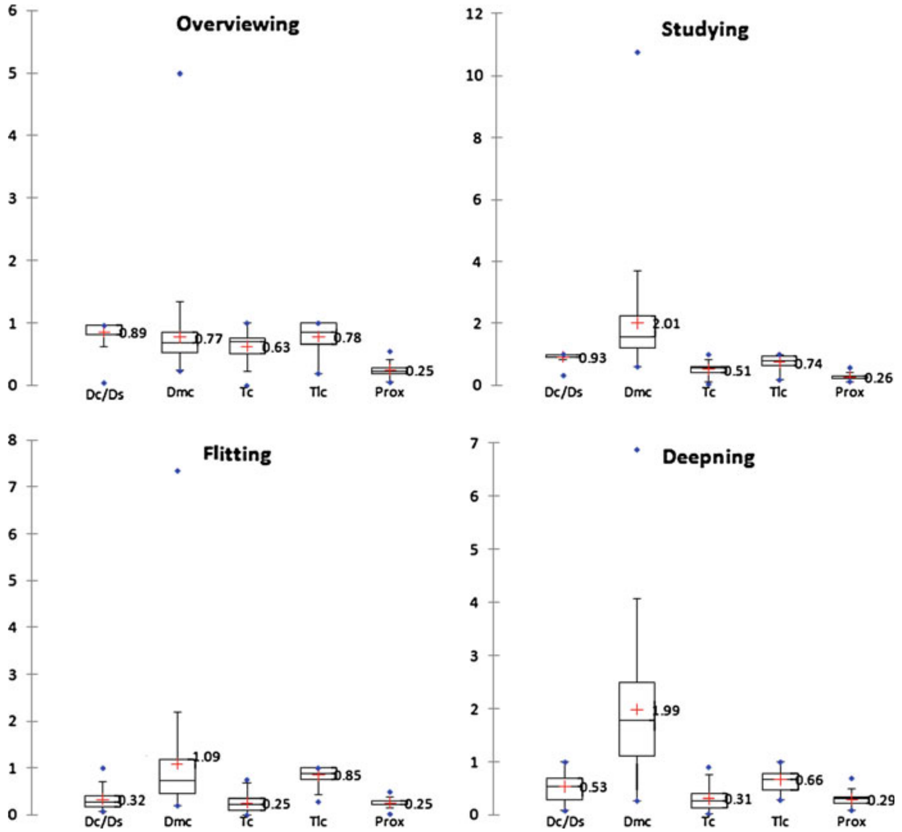

Fig. 9 Box plots—2nd experiment

Weka: *(i) Ranker* with *InfoGainAttributeEval*, *ReliefFAttributeEval*, *GainRatioAttributeEval*, *SymmetricalUncertAttributeEval*, *OneRAttributeEval*, *ChiSquaredAttributeEval* algorithms; *(ii) Best First* with *CfsSubsetEval* and *ConsistencySubsetEval* algorithms; *(iii) Exhaustive Search* with *ConsistencySubsetEval* algorithm. Using these algorithms, on both the *quantitative* and *qualitative* indicator values, the five intermediate indicators were selected, where the Dmc and the Dc/Ds indicators were always at the top position, followed by the other three indicators that had variable orders. This means that all of the five indicators are important features in identifying the navigation type. Moreover, comparing this order to the one proposed in Fig. 4, we found that Dmc is the root, and Dc/Ds, is one of the first indicators taken into account, which is coherent.

Next, to check the validity of the threshold estimation, we generated a decision tree using our sample with *numerical* values of the indicators. We applied the C4.5 algorithm known for its good result using the Weka data mining tool, and the 10-fold cross validation method. When generating the decision tree with the default settings of Weka for the C4.5 algorithm (confidence threshold to 25%), we had a complicated tree in which all the indicators appeared. To obtain a more pruned and generalized tree

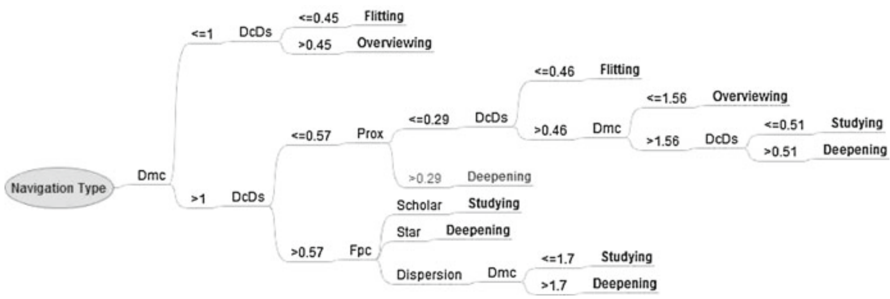

Fig. 10 Decision tree generated by the C4.5 algorithm based on the indicator numeric values

(not specific to the training data), we reduced the confidence threshold to 20%. The decision tree produced after 10 runs is presented in Fig. 10.

We noted that the first indicator considered in this tree is the average duration for a course page consultation, Dmc, which is compared first at 1 min, the same threshold that we chose in our experiments. This indicator is also displayed in internal branches of the tree, adding more precision, where it is compared to 1.56 and 1.7 thresholds. These values are higher but not far from the chosen Dmc threshold. The second indicator is the ratio Dc/Ds which is compared to 45%, 46% and 57%, which are also close to the Dc/Ds threshold. Regarding the proximity indicator $Prox$, we note that it is compared to 0.29, explaining why we redefined this threshold to 0.3 during this experiment. This leads us to ask questions about this indicator, which is noisy, and thus needs to be improved. Moreover, one can note the absence of the consultation type indicator Tc, in this pruned tree. This reveals that Tc is less important than Dmc or Dc/Ds. We will see that we obtain the same result using C4.5 with *nominal* values.

Note that this tree with *numerical* indicator values, provided correct classification 76.2% of the time: 85.1% for *Overviewing*, 66.7% for *Deepening*, 82.9% for *Studying*, and 71.7% for *Flitting*.

Once our thresholds had been justified by this *numerical* analysis, we pursued our analysis based on the *nominal* values of the indicators.

5.3.2.2. Validation of the classification As in the previous experiment, we applied the DISQUAL method (see Sect. 5.2.3) based on *nominal* values. The results are shown in Fig. 11. This figure illustrates the class centroids, on the first two axes of the DA, which allows us to account for 98.23% of the data.

The existence of the four classes identified in the manual analysis is confirmed, as their centroids are clearly separated. We concluded that the manual analysis is valid. The classification obtained with this statistical method will be compared in the following sections to the IDLS classification.

5.3.3 Automatic analysis

The next step is to infer, through the system, the "*navigation type*" of each student in each session from the intermediate *nominal* indicator values. Taking into account

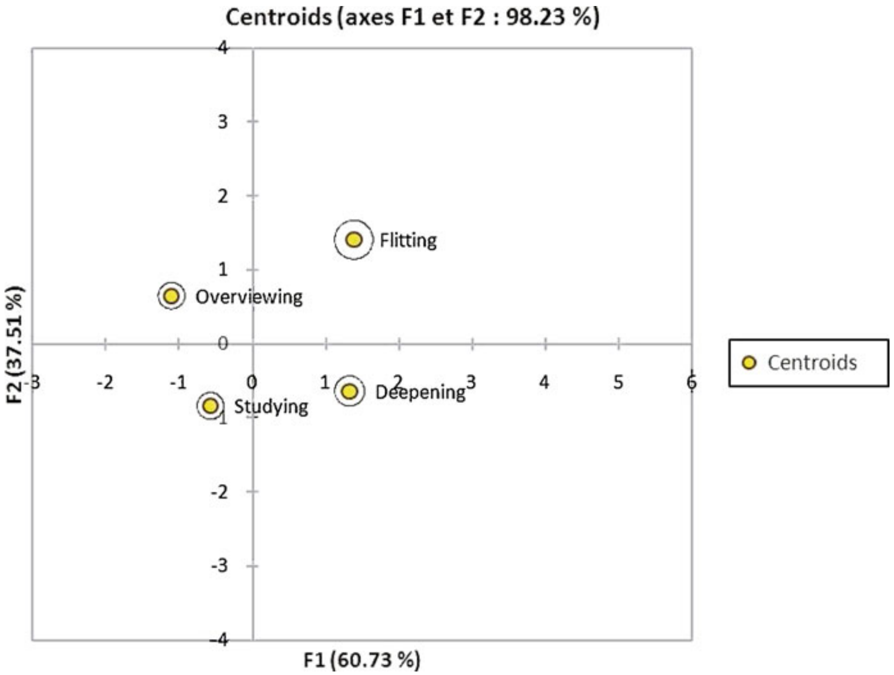

Fig. 11 Class centroids according to DISQUAL—2nd experiment

Table 8 Recognition rate of analyzed behavior—2nd Experiment

Manual	Automatic					
	O	S	D	F	Total	% Correctly found
O: overviewing	**98**	13	2	8	121	81
S: studying	3	**106**	6	2	117	90.6
D: deepening	14	26	**49**	4	93	52.7
F: flitting	13	3	2	**38**	56	67.9
Total	128	148	59	52	387	75.2

all the observations, without separating each session, we obtain the confusion matrix, presented in Table 8 (read as Table 6). This matrix compares the IDLS classification to that of the manual analysis.

We obtained 75.2% of correct classification results, using our decision tree presented in Sect. 4.3. This result is lower than the one obtained in the first experiment, especially for *Deepening*. This is mainly due to the noisy values of the $Prox$ indicator. In some cases, the $Prox$ yields the *null* value because the learner visited local files that cannot be collected during the processing.

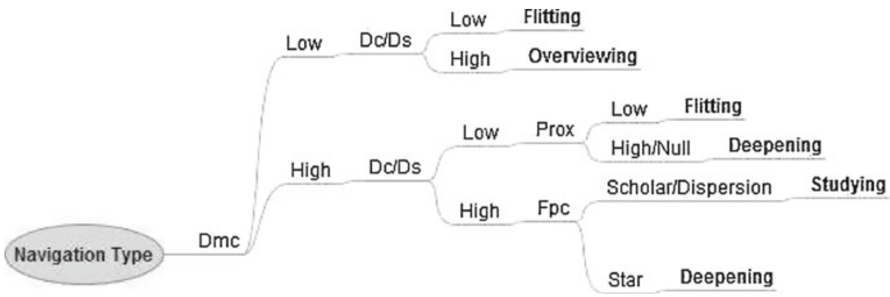

Fig. 12 Decision tree generated by the C4.5 algorithm on indicator nominal values—2nd experiment

5.3.4 Comparison and discussion

As we have a larger sample size than the previous experiment, it is useful to compare the tree we have developed semantically according to our assumptions, to the tree that can be generated on this sample with data mining methods.

To generate a new tree, using the *nominal* indicator values on which our study is mainly based, we also used the C4.5 algorithm provided by the Weka data mining tool. We split the log files generated in this experiment into 20/80 test-train split with 310 training case (97 *Overviewing*, 94 *Studying*, 75 *Deepening*, 44 *Flitting*) and 77 test cases (24 *Overviewing*, 23 *Studying*, 18 *Deepening*, 12 *Flitting*). The choice of the test-train data was random, but the distribution of the classes in the training and test samples reflects the starting sample based on the manual analysis.

We did not perform an iterative cross validation, as our goal was to compare the results of the tree we already built (known algorithm), to the results produced using data mining algorithms from the existing data. C4.5 produced the pruned decision tree presented in Fig. 12.

One can notice the absence of the *consultation type* indicator Tc. The C4.5 algorithm considers that the detail it adds can be ignored, which is a coherent result with the decision tree obtained, based on *numerical* indicator values (see Fig. 10). We also note that when Dmc is *high* and Dc/Ds is *low*, with *null* values of $Prox$, the behavior is considered to be *Deepening*. These cases occur when the learner spends time in local files that we cannot collect.

We compared the classification results of the manually annotated observations, obtained by this tree (Fig. 12) and the Disqual method, to those of IDLS as shown in Table 9. We also applied the classical data mining methods: K-Nearest Neighbor (KNN), and Naïve Bayes. We presented the results on the test sample (20%) using three metrics:

– The recall *Rec*: is the fraction of relevant observations that are retrieved.

$$Rec = \frac{Number\ of\ observations\ correctly\ classified}{Total\ number\ of\ observations\ that\ should\ be\ found} \tag{10}$$

Table 9 Comparison between classification methods

Method	Overviewing			Studying			Deepening			Flitting			Average		
	Rec (%)	Prec (%)	Fm (%)	Rec (%)	Prec (%)	Fm (%)	Rec (%)	Prec (%)	Fm (%)	Rec (%)	Prec (%)	Fm (%)	Rec (%)	Prec (%)	Fm (%)
IDLS	**79.2**	79.2	79.2	78.3	69.2	73.5	**61.1**	**78.6**	**68.8**	75	69.3	72	74.03	74.05	74.04
DISQUAL	66.7	88.9	76.2	73.9	68	70.8	**66.7**	66.7	66.7	**91.7**	**68.8**	**78.6**	72.7	73.1	72.9
C4.5	**79.2**	**95**	**86.4**	78.3	69.2	73.5	55.6	71.4	62.5	**91.7**	64.7	**75.9**	75.3	77.1	75.3
K-NN (K=6)	**79.2**	**95**	**86.4**	**82.6**	70.4	76	**61.1**	**78.6**	**68.8**	83.3	62.5	71.4	76.6	78.7	76.8
Naïve bayes	**79.2**	**95**	**86.4**	**82.6**	73.1	**77.6**	**61.1**	**78.6**	**68.8**	**91.7**	64.7	**75.9**	**77.9**	**79.9**	**78**

– The precision $Prec$: is the fraction of observations retrieved that are relevant.

$$Prec = \frac{Number\ of\ observations\ correctly\ classified}{Total\ number\ of\ observations\ found\ by\ the\ classifier} \tag{11}$$

– The F-measure Fm: is a trade-off between precision and recall.

$$Fm = \frac{2 \times (Prec \times Rec)}{(Prec + Rec)} \tag{12}$$

The results show that the *Overviewing* behavior was well appropriated by all classifiers, where the C4.5 algorithm, the K-NN and the Naïve Bayes obtained the highest results ($Fm = 86.4\%$). The IDLS obtained the same recall ($Rec = 79.2\%$) but a lower value for the precision ($Prec = 79.2\%$) and the f-measure ($Fm = 79.2\%$). Concerning the *Studying* navigation type, the Naïve Bayes received the best results ($Rec = 82.6\%$, $Fm = 77.6\%$). Regarding the *Deepening* navigation type, the best accuracy was reached by IDLS, the K-NN and the Naïve Bayes ($Fm = 68.8\%$). We noted that for this behavior all the classifiers got recognition rates lower than 70%. This is mainly due to the noisy $Prox$ values, on which we based our semantic assumption to identify the *Deepening* behavior. For the *Flitting* behavior, the Disqual statistical method obtained the best result ($Fm = 78.6\%$). The best overall accuracy was achieved by the Naïve Bayes with 78%. The overall accuracy obtained by the IDLS system is not so far from the Naïve Bayes results (74.03% compared to 78%), and the C4.5 tree (75.3%). Therefore, as we are in a learning context and our aim is to follow-up learners better, we preferred to keep the pruned tree generated by the C4.5 algorithm that is easier for teachers to understand than a Naïve Bayes model.

6 Conclusion

In this paper, we presented a study that aims at contributing to the field of computer-assisted learning. We proposed an approach for automatic identification of a learner's behavior based on trace analysis. While many other studies looked at behaviors in terms of how often and how long students visited specific kinds of learning objects/activities, this study has focused on the navigational behavior in a web-based learning context. This context, commonly used in computer-assisted learning, makes the findings widely applicable.

We presented our indicator production system, called IDLS, and our calculation process through the calculation of a high-level indicator: "*navigation type*". This indicator assigns the behavior of the learner to one of the following four classes: *Overviewing*, *Studying*, *Deepening* and *Flitting*. These classes have been induced from the literature, and successively manually identified during the tests by three independent teachers who came to an agreement at 77.32%, which is decent for three independent raters.

To calculate the values of this indicator, we proposed five intermediate indicators, which are also informative and can be visualized by the teacher to get more details about the learner's behavior. In fact, we chose to discretize the indicator values in order to provide teachers with comprehensive information. This choice aims also at generalizing our calculation methods. However, given the difficulty of discretizing indicator values, a task which depends on several factors, such as the resource types and the activities included in the course, we left scope for the teacher or the researcher to set the thresholds. As it is difficult to set these thresholds, we proposed default values for these parameters based on extensive tests. Finally, to infer the learner's behavior (the *navigation type* indicator value) from the intermediate indicators, we defined rules that we summarized in a decision tree.

To confirm and improve this approach, we conducted two experiments. On the first one, with 97 log files, our method achieved 80% accuracy. In the second experiment, using 387 log files, the accuracy dropped to 74%. Comparing our results with those of discriminant analysis, and classical data mining algorithms, we found that we obtained competitive results.

We may be able to improve our results in two ways: First, as the learner's behavior can change from one topic to another, we need to look at the behavior of the learner using a more fine-grained model (e.g., course chapter). Second, the values of the $Prox$ indicator are too noisy. We need to improve its calculation by finding a way to index external Web pages more accurately, probably using small domain ontologies. Generally, a large use of our system could allow estimating automatically precise threshold values using data mining methods. Furthermore, although we have already reduced the number of navigation behavior classes from five to four, maybe we should check whether we can present *Studying* and *Deepening* in one class. To do so, other experiments should be performed in supervised and unsupervised ways.

When we started this study, our goal was twofold: The first one was to provide tutors with information about the behavior of learners using online courses. We believe that this information could be also used by other stakeholders (learners, parents). It can give the learners a reflexive glance on their learning methods, acquiring meta-cognitive knowledge. We believe that this goal is easily reachable using our system. The second goal was to incorporate the findings of this study in the automatic student modeling process to provide adaptive and personalized navigation support, as many works made reference to "learner's learning style" when they presented adaptive educational hypermedia systems. This work suggests that it is not as easy as it might appear. Much future research is necessary to identify the learning styles automatically from the learner's navigational behavior in order to allow the IDLS system to provide such a service to Adaptive Educational Hypermedia Systems (AEHS).

Acknowledgments The authors wish to especially thank Mrs. Naima Bessah, Vu Huyen-Trang, Agathe Merceron, Mouna Kacimi, Mr. Mohammed Bellal and Dr. John Wisdom for their guidance, as well as all ESI's students that have agreed to take part in the experiments, particularly the 3SI, 2009. The authors also acknowledge the anonymous reviewers for all their effort in improving the paper, especially the "Reviewer 3".

Appendices

Appendix A: Weighting of terms

The weighting method follows TF.IDF. The method is to calculate the occurrence number of each stem, in other words, all words with the same stem are merged into a single row in the output table. The stem weight calculation takes into account the preliminary weight calculated earlier in the extraction phase of single words. We keep the list of lemmas corresponding to each word after gathering in the same stem row.

Also, since we are in a web page context (usually HTML), and since "tags do not have the same importance", the idea proposed by Desmontils and Jacquin (2002) aims to assign higher weight to terms that are found in more important tags such as tag <TITLE> which contains the title of the document, <u> which contains underlined words, etc. The table below contains the weights assigned to each tag Desmontils and Jacquin (2002).

HTML tag description	HTML tag	Weight
Document title	*<TITLE></TITLE>*	*10*
Keywords	*<META name = keywords></META>*	*10*
Link	*<A HREF = ...></A>*	*8*
Font size 7	*<FONT SIZE = 7></FONT>*	*5*
Font size 4	*<FONT SIZE = "+4"></FONT>*	*5*
Font size 6	*<FONT SIZE = 6></FONT>*	*4*
Font size 3	*<FONT SIZE = "+3"></FONT>*	*4*
Font size 2	*<FONT SIZE = "+2"></FONT>*	*3*
Header level 1	*<H1></H1>*	*3*
Header level 2	*<H2></H2>*	*3*
Title image	*<IMG...ALT = "...">*	*3*
BIG marker	*<BIG></BIG>*	*2*
Underscore	*<U></U>*	*2*
Italics	*<I></I>*	*2*
Bold	*<B></B>*	*2*

Thus, when a term appears in a tag, its occurrence number is incremented by the value of this weight (for example, if a term appears in the title, it adds to the occurrences number a value of ten (10), what is equivalent to ten (10) appearances in the text).

The term weights are defined by applying statistical methods, and according to the weight, the more descriptive terms are retained.

$$\begin{aligned}
&(1+\log(tf_{ij}))\times tf\\
&\quad = Weight(t_{ij})\times\log\frac{N}{n}\times\frac{frequency}{frequency+0.5+1.5\times\dfrac{document_length}{average_document_length}}\\
&\quad = (1+\log(tf_{ij}))\times\log\frac{N}{n_j}
\end{aligned}$$

Appendix B

See Table 10.

Table 10 Statistics on the indicator numerical values—2nd experiment

Behavior	Function	Dc/Ds	Dmc	Tc	Tlc	Prox
Overviewing	Minimum	0.08	0.23	0.00	0.19	0.05
	Median	1.00	0.69	0.69	0.85	0.22
	Maximum	1.00	5.00	1.00	1.00	0.55
	Mean	0.89	0.77	0.63	0.78	0.25
	Standard deviation (SD)	0.20	0.56	0.22	0.23	0.11
Studying	Minimum	0.32	0.60	0.03	0.17	0.12
	Median	1.00	1.54	0.58	0.80	0.22
	Maximum	1.00	10.77	1.00	1.00	0.57
	Mean	0.93	2.02	0.51	0.74	0.26
	SD	0.14	1.43	0.20	0.22	0.10
Deepening	Minimum	0.08	0.27	0.03	0.29	0.09
	Median	0.52	1.77	0.27	0.67	0.31
	Maximum	1.00	6.88	0.90	1.00	0.69
	Mean	0.53	1.99	0.31	0.66	0.29
	SD	0.28	1.41	0.19	0.22	0.09
Flitting	Minimum	0.08	0.20	0.00	0.29	0.02
	Median	0.28	0.73	0.23	0.88	0.23
	Maximum	1.00	7.35	0.75	1.00	0.49
	Mean	0.33	1.09	0.26	0.85	0.26
	SD	0.24	1.14	0.18	0.17	0.09
Over the sample	Minimum	0.08	0.20	0.00	0.17	0.02
	Median	0.92	1.13	0.48	0.79	0.26
	Maximum	1.00	10.77	1.00	1.00	0.69
	Mean	0.73	1.49	0.46	0.75	0.27
	SD	0.31	1.31	0.25	0.22	0.10

References

Acharyya, S., Ghosh, J.: Context-sensitive modeling of web-surfing behavior using concept trees. In: Proceedings of the Fifth WEBKDD Workshop: Webmining as a Premise to Effective and Intelligent Web Applications, WEBKDD'2003, Washington, USA (2003)

Ait-Adda, S., Bousbia, N., Balla, A.: Détection des comportements des apprenants à base d'indicateurs de parcours'. In: Proceedings of the 1st International Conference on Web and Information Technologies, ICWIT08, Sidi Bel Abes, Algeria, 29–30 June 2008

Avouris, N., Fiotakis, G., Kahrimanis, G., Margaritis, M., Komis, V.: Beyond logging of fingertip actions: analysis of collaborative learning using multiple sources of data. J. Interact. Learn. Res. JILR **18**(2), 231–250 (2007)

Baeza-Yates, R., Ribeiro-Neto, B.: Modern Information Retrieval. ACM Press, USA (1999)

Barros, B., Verdejo, F.: Analysing student interaction processes in order to improve collaboration. The DEGREE approach. Int. J. Artif. Intell. Educ. **11**, 221–241 (2000)
Baziz, M., Boughanem, M., Nassr, N.: La recherche d'information multilingue : désambiguïsation et expansion de requêtes basés sur WordNet. In: Proceedings of the Sixth International Symposium on Programming and Systems, ISPS'2003, pp. 175–186, Algiers, Algeria, 5–7 May 2003
Beauvisage, T.: Sémantique des parcours des utilisateurs sur le Web. Ph.D. thesis, Université Paris X, France, (2004)
Bidel, S., Lemoine, L., Piat, F., Artires, T., Gallinari, P.: Statistical machine learning for tracking hypermedia user behavior. In: Proceedings of the 9th International Conference on User Modeling, UM'03, pp. 48–57, Johnstown, PN, 22–26 June 2003
Bousbia, N.: Contribution théorique et méthodologique à l'élaboration d'un environnement de FOAD (Formation Ouverte et à Distance). Magister en Informatique, Ecole nationale Supérieure d'Informatique (ESI), Algiers, Algeria (2005)
Bousbia, N., Labat, J.-M.: Perception de l'activité de l'apprenant dans un environnement de formation. In: INRP (Eds.) Environnements Informatiques pour l'Apprentissage Humain, EIAH 2007, pp. 233–238 (2007)
Bousbia, N., Labat, J.-M., Rebai, I., Balla, A.: How to determine the Learners' Learning Styles in e-Learning Situation? In: Proceedings of the 16th International Conference on Computers in Education, ICCE 2008, Taipei, Taiwan 27–31 Oct 2008
Bousbia, N., Labat, J.-M., Rebai, I., Balla, A.: Indicators for deducting the learners' learning styles: case of the navigation typology indicator. In: Proceedings of the 9th IEEE International Conference on Advanced Learning Technologies, ICALT 2009, Riga, Latvia, 14–18 July 2009
Canter, D. , Rivers, R., Storrs, G.: Characterizing user navigation through complex data structures. Behav. Inform. Technol. **4**, 93–102 (1985)
Catledge, L.D., Pitkow, J.E.: Characterizing browsing strategies in the World-Wide Web. In: 3rd International World Wide Web Conference, Darmstadt, Germany (1995)
Champin, P.-A., Prié, Y.: MUSETTE: un modèle pour réutiliser l'expérience sur le Web sémantique. Journées scientifiques Web sémantique, Paris, France, Oct 2002
Chen, W.: Suporting teachers' interventions in collaborative knowledge building. J. Netw. Comput. Appl. **29**(2–3), 200–215 (2006)
Chen, S.Y., Liu, X.: An integrated approach for modeling learning patterns of students in Web-based instruction: A cognitive style perspective. ACM Trans. Comput.-Hum. Interact. (TOCHI) **15**(1), 1–28 (2008)
Cheype, A.: Recherche de motifs séquentiels pour guider l'interprétation des traces d'apprentissage dans un EIAH. In: Proceedings of the Conference Rencontre de Jeunes Chercheurs en EIAH, RJC-EIAH'2006, pp. 123–130, Evry, France, 11–12 May 2006
Choquet, C., Iksal, S.: Usage tracking language: a meta-language for modeling tracks in TEL systems. In: Proceedings of the International Conference on Software and Data Technologies (ICSOFT), pp. 133–138, Setubal (Portugal), 11–14 Sept 2006
Cockburn, A., Mckenzie, B.: What do Web users do? An empirical analysis of Web use. Int. J. Hum.-Comput. Stud. **54**, 903–922 (2000)
Cooley, R., Tan, P.-N., Srivastava, J.: Websift: The Web site information filter system. In: Proceedings of the 1999 KDD Workshop on Web Mining, San Diego, CA (1999)
Cram, D., Jouvin, D., Mille, A.: Visualisation interactive de traces et réflexivité : application à l'EIAH collaboratif synchrone eMédiathèque. Sciences et Technologies de l'Information et de la Communication pour l'Education et la Formation (STICEF) **14** (2007)
Damez-Fontaine, M.: De l'apprentissage artificiel pour l'apprentissage humain : de la récolte de traces à la modélisation utilisateur. Ph.D. in Computer Science, Université Pierre et Marie Curie, Paris 6, France (2008)
Desmontils, E., Jacquin, C.: Indexing a web site with a terminology oriented ontology. In: Cruz, I.F., Decker, S., Euzenat, J., McGuinness, D.L. (eds.) The Emerging Semantic Web, pp. 181–197. IOS Press, Amsterdam (2002)
Diagne, F.: Instrumentation de la supervision par la réutilisation d'indicateurs : Modèles et Architecture. Ph.D. in Computer Science, Université Joseph Fourier—Grenoble, France, (2009)
Dimitracopoulou, A. et al.: State of the art on interaction analysis: interaction analysis indicators. JEIRP-ICALTS Deliverable 26.1, Kaleidoscope, (2004)

Dimitracopoulou, A., Bruillard, E.: Enrichir les interfaces de forums par la visualisation d'analyses automatiques des interactions et du contenu. Sciences Et Technologies de l'Information Et de la Communication Pour l'Education Et la Formation (STICEF) **13**, 345–397 (2006)

Fessakis, G., Petrou, A., Dimitracopoulou, A.: Collaboration activity function: an interaction analysis'tool for computer supported collaborative learning activities. In: 4th IEEE International Conference on Advanced Learning Technologies (ICALT 2004), Joensuu, Finland (2004)

Gall, J.E., Hannafin, M.J.: A framework for the study of hypertext. Instr. Sci. **22**, 207–232 (1994)

Georgeon, O., Mille, A., Bellet, T.: Abstract: un outil et une méthodologie pour analyser une activité humaine médiée par un artefact technique complexe. In: Proceedings of the Conference Ingénierie des Connaissances, IC (2006)

Gerosa, M.A., Pimentel, M. Gomes, Fuks, H., Lucena, C.: No need to read messages right now: helping mediators to steer educational forums using statistical and visual information. In: Proceedings of the Conference on Computer Support for Collaborative Learning, CSCL 2005, pp. 160–169 Taipei, Taiwan (2005)

Heer, J., Chi, E.: Separating the swarm: categorization methods for user sessions on the Web. In: Proceedings of ACM CHI 2002 Conference on Human Factors in Computing Systems, Minneapolis, USA (2002)

Heraud, J.-M., France, L., Mille, A.: Pixed: An ITS that guides students with the help of learners' interaction logs. In: Proceedings of the 7th International Conference on Intelligent Tutoring Systems. Maceio, Alagoas, Brazil (2004)

Hernandez, N.: Description et détection automatique de structures de texte. Ph.D. thesis, Université d'Orsay, Paris XI, France, (2004)

Hofmann, K.: Subsymbolic user modeling in adaptative hypermedia, a case study in educational context. Master of Science in Computer Science, California State University (2006)

Khatraoui, M., Bousbia, N., Balla, A.: Détection de similarité sémantique entre pages visitées durant une session d'apprentissage. In: Proceedings of Atelier Mesures de Similarité Sémantique (EGC'08). 8èmes Journées Francophones Extraction et Gestion des Connaissances, Sophia-Antipolis, France, 29th jan 2008

Khatraoui, M.: Analyse et assistance á l'apprentissage: Proposition d'une méthode pour la détection des relations sémantiques entre les pages visitées en dehors du contenu lors d'une session d'apprentissage. Magister in Computer Science, Ecole nationale Supérieure d'Informatique (ESI), Algiers (2008)

Lawless, K.A., Kulikowich, J.M.: Understanding hypertext navigation through cluster analysis. J. Educ. Comput. Res. **14**(4), 385–399 (1996)

Lejeune, A., David, J.-P., Luengo, V., Pernin, J.-P., Diagne, F., Adam, J.-M., Choquet, C.: State of art of tracking and analysing usage. DPULS Deliverable 32.3. (2005)

Lin, D.: An information-theoretic definition of similarity. In: Proceedings of the 15th International Conference on Machine Learning, pp. 296–304, San Francisco, CA, 1998

Martinez, A., Dimitriadis, Y., Rubia, B., Gḿez, E., de la Fuente, P.: Combining qualitative and social network analysis for the study of social aspects of collaborative learning. Comput. Educ. **41**(4), 353–368 (2003)

Michelet, S., Adam, J.M., Luengo, V.: Adaptive learning scenarios for detection of misconceptions about electricity and remediation. In: Proceedings of the International Conference ICL'2006 Interactive Computer Aided Learning, Villach, Autriche, 27–29 Sept 2006

Mroué, A., Caussanel, J.: Recherche de patrons de navigation pour les systèmes de recommandations. In: Proceedings of the Workshop Modélisation utilisateur et personnalisation d'interface Web, 8èmes Journées Francophones Extraction et Gestion des Connaissances, Sophia Antipolis, France (2008)

Ollagnier-Beldame, M., Mille, A.: Faciliter l'appropriation des EIAH par les apprenants via les traces informatiques d'interactions. Rapport de recherche RR-LIRIS-2007-023, (2007)

Page-Lamarche, V.: Styles d'apprentissage et rendement académique dans les formations en ligne. Ph.D. en sciences de l'éducation, Faculté des études supérieures, Université de Montréal, (2004)

Papanikolaou, K.A., Grigoriadou, M.: Accommodating learning style characteristics in adaptive educational hypermedia systems. In: Proceedings of Individual Differences in Adaptive Hypermedia Workshop at the Third International Conference on Adaptive Hypermedia and Adaptive Web-based systems, AH'04, Eindhoven, Netherlands (2004)

Popescu, E., Badica, C., Trigano, P.: Rules for learner modeling and adaptation provisioning in an educational hypermedia system. In: Proceedings of the Ninth International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, SYNASC, pp. 492–499 (2007)

Reffay, C., Lancieri, L.: Quand l'analyse quantitative fait parler les forums de discussion. Sciences et Technologies de l'Information et de la Communication pour l'Education et la Formation (STICEF), Numéro spécial forum de discussion en éducation, pp. 255–288 (2006)
Reimann, P.: How to support groups in learning: more than problem solving. In: Artificial Intelligence in Education (AIED 2003). Supplementary Proceedings, Invited talk, pp. 3–16, Sydney, Australia, 2003
Sanderson, P.-M., Fisher, C.A.: Exploratory sequential data analysis: foundations. Hum.-Comp. Interact. **9**(3), 251–317 (1994)
Saporta, G.: Probabilitées, analyse des données et statistiques. Editeur Technip (2006)
Schummer, T., Strijbos, J.-W., Berkel, T.: A new direction for log-files analysis in CSCL: Examples with a spatio-temporal metric. In: Proceedings of the 2005 Conference on Computer Support for Collaborative Iearning, CSCL2005, pp. 567–576, Taipei, Taiwan, 2005
Settouti, L., Prié, Y., Mille, A., Marty, J.-C.: Système à base de traces pour l'apprentissage humain. In: Proceedings of the Colloque international des Technologies de l'Information et de la Communication dans l'Enseignement Supérieur et l'Entreprise, TICE (2006)
Siebra, S., Salgado, A.C., Brézillon, P., Tedesco, P.: A learning interaction memory using contextual information. In: Proceedings of the CONTEXT-05 Workshop on Context and Groupware, Paris, France, 2005
Spiliopoulou, M., Faulstich, L.C., Winkler, K.: A data miner analyzing the navigational behaviour of web users. In: Proceedings of the Workshop on Machine Learning in User Modeling of the ACAI'99, Creta, Greece, 1999
Talbot, S., Courtin, C.: Trace analysis in instrumented learning groupware: an experiment in a practical class at the university. In: Proceedings of IASTED, the International Conference on Computers and Advanced Technology in Education (CATE'07), Beijing, China, 8–10 Oct 2007
Tauscher, L., Greenberg, S.: How people revisit Web pages: empirical findings and implications for the design of history systems. Int. J. Hum.-Comput. Stud. **47**, 97–137 (1997)
Thuring, M., Hannemann, J., Haake, J.M.: Hypermedia and cognition, designing for comprehension. Commun. ACM **38**(8), 57–66 (1995)
Varelas, G., Voutsakis, E., Raftopoulou, P., Petrakis, E.G.M., Milios, E.E.: Semantic similarity methods in WordNet and their application to information retrieval on the Web. In: Proceedings of the 7th ACM International Workshop on Web Information and Data Management, WIDM 2005, pp. 10–16, Bremen, Germany, 5 Nov 2005

Author Biographies

N. Bousbia is a teaching assistant of Computer Science at the School of Computer Science (ESI), Algeria. She received her M.S. degree from the ESI, in 2005. Then she joined the UFC University in Algeria, as a project coordinator for the e-learning platforms. Currently she is a Ph.D. candidate in Computer Science under a joint supervision of the ESI in Algeria, and UPMC University in France. Her primary interests lie in the areas of trace analysis, user modeling and learning styles. Her paper summarizes the current state of her thesis work on learners' behavior analysis.

I. Rebaï is an associate professor in the Department of Computer Science at Telecom Bretagne Engineering School in France. He received his Ph.D. from the University of Paris Descartes, in 2006. His dissertation focused on the design and the development of a platform allowing the capitalization of software components for interactive learning environments. He then joined the CNRS in 2008 as a post-doctoral fellow at the LIMSI Computer Science Laboratory. He conducted research on ambient environments, emotions, and robotics. His research interests include learning metadata, affective and ubiquitous computing.

J.-M. Labat is a full Professor of Computer Science at UPMC University in Paris. He is the head of a research team within the LIP6 laboratory and the leader of the UTES, a university service. He is also the current president of the ATIEF association. He received his Ph.D. from UPMC. His research focuses generally on cognitive modeling of users experiencing problem solving. His current research goals are to investigate serious games.

A. Balla is an associate professor of Computer Science at the School of Computer Science (ESI) in Algeria. He is working in the area of adaptive educational hypermedia systems. He completed his Ph.D. in 2005 at the ESI. He is the head of a research team within the LMCS laboratory and the department leader at the ESI. His research interests mainly focus on adaptive hypermedia and student modeling.