Abstract—Any activity in a computer-supported cooperative working environment produces a set of traces. In a collaborative working context, such traces may be very voluminous and heterogeneous. They reflect all the interactive actions among the actors themselves and between the actors and the system. This paper examines what is required to study traces in the context of a collaborative working environment (CWE). The objective is to propose a definition of the different kinds of traces and to build a model for classifying and analyzing the interactions with respect to both individual needs and group needs. In a CWE, the different types of traces can be divided into two categories: Collaborative Traces and Private Traces. As a basis of our model, the Collaborative Trace represents the shared personal experiences in a group and can be filtered to improve the collaborative and personal work. For practical applications, we validate our model using the E-MEMORa2.0 platform.

Index Terms—collaborative working environment, trace-based system, collaborative trace, collaborative engineering.

I. INTRODUCTION

With the rapid development of information technologies and the popularity of smart devices, collaborative work is much simpler and more common than ever. People can work together irrespective of their geographical location and time limitation. Furthermore, the computer-supported tasks are widely available and many tools customizable exist for performing the tasks. However, the closer the relationship between individual and group, the more complicated the group model. Consequently, the interaction modes in groups or the sharing of experiences among users brings up critical research issues in different fields of CSCW, like cooperative work in design (Forlizzi & Battarbee [1], Hao et al. [2]), e-learning (Naidu & Ip [3], Bird et al. [4]), social networking (Vaarmanen-Vainio-Mattila et al. [5]), or e-business (Kim & Lee [6], Tomaz et al. [7]).

In a web-based collaborative working environment (CWE) interactions facilitate sharing information. Almost all the past interactions represent a kind of trace that can be regarded as the user’s working experience (Laflaquière [8]). According to Clauzel and his colleagues, an interaction trace is defined as: “histories of users’ actions collected in real time from their interaction with the software” [9]. In their project: “Trace-Based Management Systems (TBMS) (systems devoted to the management of modeled traces)”, the researchers focused on the personal interaction trace. They mentioned the concept: "Synchronous Collaborative Traces", but do not offer further discussion of its definition. In 2003, Mille and his colleagues proposed an approach, MUSSETTE (Modelling USEs and Tasks for Tracing Experience), to "capture a user trace conforming to a general use model, describing the objects and relations handled by the user of the computer system" [10]. MUSSETTE considers the trace as "a task-neutral knowledge base" that can be reused by the system assistants. The researchers of the TRAILS project (Personalized and Collaborative Trails of Digital and Non-Digital Learning Objects) consider the trace in hypermedia as a sequence of actions and use them to identify the overall objective of the user. In a different way, Settouti and his colleagues defined a numerical trace as a "trace of the activity of a user who uses a tool to carry out this activity saved on a numerical medium" [11]. Zarca and his colleagues define an interaction trace as "a record of the actions performed by a user on a system, in other words, a trace is a story of the user’s actions, step by step" [12]. These considerable research works emphasize the personal aspect, however, they provide little insight for answering the question on "how to share and reuse the users’ experiences in a group" and do not provide an effective method for directing the practical design in a CWE.

In collaborative working environments, the users’ traces are numerous and varied, for example resulting from sharing calendars and documents, assigning tasks, charting history, sending email, writing wikis and so on. Such traces typically deal with two kinds of activities: personal activity and group activity. However, telling the differences between each member’s activities and the group is inefficient. Moreover, compared with e-learning and e-business environments, the users’ interactive activities in CWE are more complex, due for example to privacy rights or group leadership. Although every trace is produced by a single user, with respect to his objectives the traces can be separated into two categories: Private Trace and Collaborative Trace. Even though we are not interested in

---

1Unless said otherwise, in this article, we do not make a difference between trace and interaction trace.

2http://www.noe-kaleidoscope.org/telearc/
private traces directly, they nevertheless have to be taken into account for two reasons: (i) to contribute to the theoretical framework, which otherwise would be incomplete; and (ii) to be mined in case some personal work is redistributed among different participants, e.g., in order to smooth the workload.

The objective of our research is to define and model the users’ collaborative traces for representing, sharing, classifying and visualizing their working experience in the collaborative working environment. In this article, we define the concept of "collaborative trace" to specify and classify the interactions with the environment and the relations between group and individuals. It is based on the intention of "interactive trace". Hence, collaborative traces enable us to capture the users’ working experiences. We propose a formal model to analyze and organize the collaborative traces. With this model, a general architecture of the system based on collaborative traces is introduced. Our approach stands within the theory of trace based systems and focuses on the collaborative working environment.

This paper is organized as follows: Section 2 presents and discusses the definition of collaborative trace, which is the theoretical foundation of our model. Section 3 presents our model with a practical example. Necessary comparisons between models of traces are discussed and analyzed. Section 4 presents the application of our model in a collaborative working environment: E-MEMORAE 2.0. Finally, Section 5 concludes and discusses future perspectives.

II. DEFINITION OF COLLABORATIVE TRACE

Our approach concerns collaborative working environments (CWE). The concept comes from the use of collaborative software in a workspace (for example: e-work and virtual workspaces, see Shaffers et al. [13] or Prinz et al. [14]). A CWE can support both the individual work and the cooperative work (Wangsa et al.[15]). It is a subclass of collaborative software. As the trace is produced by the interactive activities and naturally, it is strongly affected by the environment properties. Before giving the definition of a collaborative trace, it is necessary to explain the difference between "Collaborative Software" and "Computer-Supported Collaboration Work".

Collaborative software (also referred to as Groupware) is a controversial concept which has a great overlapping with Computer-supported Cooperative Work (CSCW). Paul Cashman and Irene Grief coined the term "computer-supported cooperative work" (or "CSCW") at a workshop in 1984, in order to find out how the technology could support people in their work [16]. Many researchers have their own definition of CSCW. Carstensen and Schmidt [17] consider that CSCW addresses "how collaborative activities and their coordination can be supported by means of computer systems". Baeccker [18] defined CSCW as "computer-assisted coordinated activity carried out by groups of collaborating individuals". Probably, one of the most general definition is: "in its most general form, CSCW examines the possibilities and effects of technological support for humans involved in collaborative group communication and work processes" by Bowers and Benford [19].

The concept of "groupware" was coined by Peter and Trudy Johnson-Lenz in 1978 [20]. Their definition is: "intentional group processes plus software to support them". The objective of groupware is to assist a group of users in communicating, in collaborating and in coordinating their activities [21]. Ellis, Jacobson and Horvitz claim that "Groupware is a computer-based system that supports groups of people engaged in a common task (or goal) and that provides an interface to a shared environment" [21]. Differences between CSCW and collaborative software (groupware) come from the different definitions and understandings of the two concepts. Wilson [22] gives a clear explanation of the discrepancy: "CSCW is a generic term, which combines the understanding of the way people work in groups with the enabling technologies of computer networking, and associated hardware, software, services and techniques." CWE is nearer to groupware and pays more attention to "web-based" conditions and practical features of collaboration (see Shaffers et al. [13] or Prinz et al. [14]).

CWE derives from core concepts of groupware, for example, supports both individual and group work, facilitates information sharing and improves communications. All types of users’ actions reveal the interactions among the actors themselves and between actor and environment. In consideration of the objectives of such actions (see Speech Act theory, Austin 1962 [24], Searle 1965 [25], Winograd and Flores 1986 [23]), generally, the “actions” can be divided into three types: (i) individual actions; (ii) transfer actions; and (iii) group actions as shown Figure 1. The framework that we are proposing does not require further detailing of the types of actions. In practice, users’ actions depend on the particular environment they are working with. Their interactions are captured by various devices, stored, then abstracted through filters. The results of actions can be saved either in the private space, e.g. the personal calendars, or in the group space, e.g. the shared documents and so on. Here, the word space refers to a production space (private or shared workspace). Since every finished action and corresponding result denotes an interactive trace. From Figure 2, Ontology of Trace in CWE, we could distinguish the existing bulk traces into two categories and four basic traces. In our article, a single user can be regarded as an elementary group or a set containing a single user.

Consider the means of collaboration and the correlation of
A Collaborative Trace is a set of traces produced by a user belonging to a group and aimed at that group.

Two points about this definition need to be clarified: (i) "a user belonging to a group"; (ii) "a set of traces";

A user belonging to a group means the traces in a group strongly rely on the group structure. Once the collaboration relation changes, the group member’s collaborative traces are rebuilt. It concerns the theory of groupware model (or team modeling) which is a complex issue in CWE theory, and may involve the group size, the framework of the group, and many other features. More details can be found in the work of Sartori [26], Levi [27] or Pankiewicz [28]. Our particular interest here is to answer the question: "how to define and model these collaborative traces?" A mapping from the group structure to the trace space is proposed in Section IV to find the corresponding collaborative traces.

Trace order. The analyzed traces have an order, for instance: a temporal series or an importance series. In general, users' activities are saved and organized according to a time line. Although the "time sequence" is the most common choice, we could use other standards like geographical position, importance level, or urgency level to define this sequence. Thus, some kind of filtering is needed to classify the traces for a specific usage. For example, in a group, we may want to see which documents have been most used by whom in a given time interval.

III. Collaborative Trace Model

From the definition of collaborative trace in Section II we can set up a formal model to further study this subject. In order to better understand the model of collaborative trace, a simple example is introduced first.

Suppose there are several engineers working together for a project in a certain web-based CWE. One of them, Peter, sends a message to a colleague, Tom, for a technical problem and then writes an entry into his personal wiki about the issue. Tom thinks that the question is crucial and useful for their next step. So he shares his answer and also posts the question and answer to the group.

Taking a comprehensive view of interactive traces and of the example, we can see that one trace involves four basic factors: (i) "Identity", the person who is the agent (does this "action"); (ii) "Actions", the type of action, a transfer action, personal action or group action; for example, "send a message" is a transfer action and "post a message to share it" is a group action; (iii) "Content", is a description of the action and of its result. It depends on the capturing ability and could be a vector with several values, for example, image, video, text, or geographical position. However, collecting all the possible contents characterizing actions is not practical and is usually restricted by technical conditions. In a web-based environment, it is necessary to refer to the technique of web data mining for collecting and analyzing the data concerning traces, for instance machine learning, statistical method, or ILP (Inductive Logic Programming); (iv) "Index", an identifier depending on the trace sequence. A common index is "time", in practical situations the geographical location could also be chosen. Based on this four basic factors, in our example, Peter’s collaborative trace is: <"Peter"(user ID), "Sends a message"(commands), "Content of message"(string), "2011-09-02, 10:23:45"(time)>. Tom’s private trace is: <"Tom"(user ID), "Write an entry for Wiki"(commands), "Content of entry"(string and images), "2011-09-02, 14:03:35"(time)>. Besides, his collaborative trace is: <"Tom"(user ID), "Share an article"(commands), "Content of article"(string and images), "2011-09-02, 15:30:52"(time)>

A. Definition of the collaborative trace

According to the above example, a trace of the i'th user can be defined as a vector with four attributes:

\[ t_i = <\text{Identity}, \text{Action}, \text{Content}, \text{Index} >, k \in N^+ \]

In practice, a trace can be considered as a vector with several values: \[ t_i = <v_1, v_2, ..., v_{k-1}, v_k > \]. Every value \( v_i \) is defined by some environment conditions and by the capacity of capturing it. Logically, depending on the goal of the corresponding action, a trace can be a private trace or a collaborative trace. For example: "edit the personal wiki" is a "private trace", since personal wikis cannot be read or edited by other users, and "post an article to the group project" is a "collaborative trace", because the shared article can be used by all the group members. Therefore, for the i'th user, we have formally: \[ t_i = [I_i; T_i] \], where \( I_i \) is the private trace and \( T_i \) is the collaborative trace (Figure 3).

Suppose now that the group is composed by subgroups and other members (Figure 4). Every group is a set of users: \( G_j = \{ u_i, i \in N^+ \}, j = 1, ..., m \). Where \( m \) is the total number of users in the group and \( u_i \) is the i'th user in the group. Mapping the group structure onto traces, it is not difficult to find the collaborative traces (Figure 5). For the j'th group, the mapping \( F_j \) can be defined as: \[ GT_j = F_j(G_j) \], where \( GT_j \) is the set of members’ collaborative traces and...
formally $GT_j = \{ \bigcup_{u_i \in G_j} \bigcup_{k \in N^+} T_i^k \}$. The mapping $F_j$ reflects the group composition in the set of traces. Indeed, it is a function from the group structure space to the trace spaces. In some practical cases, for example: as for databases, these mappings can be regarded as a sequence of queries to select data and create new tables. We can regard a single user as a particular case of group, a group containing a single element. With this extension the mapping also covers the single user’s private traces.

For the group $G_j$, the i’th member’s collaborative trace is defined as: $T_i = \bigcup_{k \in Q_i} T_i^k$ where $Q_i$ is a set of total number of collaborative traces. $T_i$ totally relies on the group structure. For instance: if a user belongs to two independent groups at the same time, for each group, his collaborative trace is different and cannot be shared between the groups. Regarding privacy, the private traces cannot be shared nor managed except by the user himself. On the contrary, the collaborative trace can be considered as a shared working experience of a particular group.

### B. Definition of collaborative trace filter

A filter of collaborative traces can help analyze the synchronous and asynchronous interactions in this collaborative working environment. In the most general situation, a filter of collaborative traces can be defined as:

$$\psi = \zeta_{p_m \in P_m} (^< \xi_m^q, v_{m,q} >)$$

Where $P_m \subset \wp(P)$ and $P = \{p_m\}$.

In this definition:
- $\zeta$: a combination of elementary operators $\xi^q_m$
- $\xi^q_m$: an elementary operator
- $v_{m,q}$: a value of reference
- $p_m$: a property of the trace (component of the trace)
- $\wp(P)$: the power set of $P$
- $P_m$: a member of $\wp(P)$
- $q$: the index of a filter associated with $p_m$
- $m$: the index of the property $p_m$

The complete combination of all the elementary operators is defined as:

$$\Xi = \bigcup_{p_m \in P_m} \bigcup_{q \in N^+} \xi^q_m$$

From the definition of collaborative traces, the output of the filter can be written as:

$$T_j = \psi(GT_j) = \bigcup_{u_i \in G_j} \bigcup_{k \in N^+} \psi(T_i^k)$$

For the reference values $\{v_{m,q}\}$, here is an example, $\{v_{m,q}\}$ can be time points or geographical positions. The elementary operator $\xi^q_m$ can be regarded as the basic operation for the property $p_m$, for example: a sentence of programs. Since not all of the properties are indispensable for a certain demand, a filter is acceptably defined as a combination of the elementary operators.

The filter keeps the information that we gather from the observing process of the collaborative trace. For the group, this kind of filter has three fundamental functions: (i) transformation of data, which means that the distinct contents of collaborative trace (txt, photo, video etc.) can be analyzed in different manners, and then output the transformed data; for example: a filter that transforms the resources (the shared files "doc" in group) into an html document; (ii) analysis of certain kinds of data, for example: the contribution of every member in a group to the wiki; (iii) classification of data, this function can be applied in the transformed data and also the untransformed data. The rules of classification can be defined by explicit usage of group members, for example: in a time interval, or between two members. In the actual situations, a filter is usually designed as an algorithm with several components of the trace object.

To conclude this section, the triple structure $(T, F, \psi)$ is defined as the basic collaborative trace model. Compared with the trace model proposed by Clauzel and his colleagues [9], our model focuses on the classification and on the different levels of traces, moreover, in the context of collaborative working environment. The model introduced by Sehaba [32]
on the other hand, is used to deal with the topic of how to share the traces from different operating systems.

IV. APPLICATION

Before explaining our application, it is necessary to give a sketch of a general framework of trace-based system (Figure 6). Normally, a trace-based system is composed of three processes: (i) collection; (ii) transformation; and (iii) presentation.

During collection a trace results from the observation of the interactions and is composed by a set of values depending on the level of capturing. In a web-based environment, generally, the techniques and models of collection are an important issue of Web Mining. The main data is text documents, hypertext documents, links structure, server logs, browser logs and so on. In our application, we mainly collect the "concepts of ontology" and "shared documents" that are stored in a MYSQL server.

Transformations are done with the traces collected in the first process. We can classify, analyze, merge and edit these traces automatically or manually. Indeed, the transformation of traces is directly affected by the system environment and the programming language. For different sizes of data, the operating time costs vary with the input, which is why we must clarify the group size. Besides, our proposed filters (one is personal filter, and the other is for the group) would be activated in this process.

Presentation is the last process and concerns the interface with the environment. Although visualization of the results form transformation is the mainstream in CWE applications, audio presentations can also be helpful and powerful. We prefer tables and figures that can be clearer and more accurate for explaining the results.

We apply our model to a web-based collaborative platform E-MEMORAe2.0[33]. E-MEMORAe2.0 is used to support organization learning within the MEMORAe approach [34]. The model of this platform associates (i) Knowledge engineering and educational engineering; and (ii) Semantic Web and Web 2.0 technologies. It can manage the fields of expertise of the organization and favor collaboration. In order to define, structure and capitalize explicit knowledge, the learning organizational memory is structured by means of ontologies that define knowledge within the organization in this platform [35] (Figure 7). The user can navigate through the application ontology. Besides, he can also organize and capitalize the resources. In a group, members can share documents, write wikis, edit the group calendar and so on. E-MEMORAe2.0 meets the demands of the collaborative work and also assists the personal work.

We give the definition of the four trace attributes for the collaborative trace in our application. The "Identity" is the user id in the list "per_id" from the table "mem_personne". The "Actions" that we capture are "examine the ontology concept" and "edit the resources". The "Content" is the name of a "Concept" or "Resource" that can be captured from the tables of "the concept" and "the resources" in data base. The "Index" is defined as time and date. All these variables are recorded in a MYSQL data base. The test involves an ontology of a lecture on probability. Group members are Qiang, Adeline, Marie-Hélène and Jean-Paul (Figure 8). This test group is formed by two subgroups: one that has three members and the other owns a single user. It is taken for granted that all members are equal in the group A and their collaborative traces are within the group’s activities.

For private traces, the filter classifies the concerned "Concepts" and "Resources". In the case of "Concepts", as shown Figure 9, the outputs include two parts: the upper chart shows the five most consulted concepts during one month (from 09/12/2011 to 10/12/2011); and the lower figure presents the corresponding service conditions. From Figure 9, one can see easily that the most consulted concept is "Ensemble Fini" (Finite Set), and in Oct.07, the user examined this concept three times. From the private trace, we could obtain user’s preferences and the relevant details based on the time-line. For the "Resources", it is almost the same, but focus on the private document service conditions.
The filter of collaborative traces analyzes the shared documents in the group. In our application, we captured two categories of shared files: one including pdf and doc documents, the other involving videos and images. The upper outputs of filter are presented in Figure 10 showing the quantity of each category that were shared during one month (from 09/12/2011 to 10/12/2011) for the three most concerned concepts. For the concept "Probabilité" (Probability), this member shared three pdf, three doc, one video and three images. The lower table shows the feedbacks of the service conditions of the three shared pdf documents about the concept "Probabilité" (Probability) in group A. The "frequency" means the number of times: "reading the document". As we see, the PDF2: "Note I de Probabilité" (Note I of Probability) was of no interest and had never been opened by Adeline, however, it was read several times by Qiang and Jean-Paul. With the feedbacks of the shared documents, every member can distinctly know his contributions to the group and also the preferences and needs of other members during a certain period.

As a favorable method to enhance the sharing working experiences in groups, the filtered collaborative traces could help understand the group preferences (the most relevant problems or the concepts of highest interest) and some potential collaborative relations between group members (for instance, we could propose a communication between Qiang and Jean-Paul about the PDF2 in a next step). With the recommendation techniques that are based on the group preferences (as in our application, the frequency and the number of shared documents both can be regarded as the "ranking" for group preference), documents, links and other kind of knowledge recommendations become possible. Last but not least, apparently, the outputs of filters are directly affected by the "Content" of the trace but different forms of presentation are very crucial to make the traces that could be intelligible and convenient to use. For the CWE, transforming different data types in the processes of presentation, for example, the content of images and videos is a difficult problem. In E-MEMORae2.0, through the application of the collaborative trace, the personal collaborative working experiences are recorded and classified to enrich the group experiences and could be applied to support the Tendering process (in railway transport)[36][37]. Indeed, in this case, several teams need to collaborate in order to propose the best solution to the customer. An example of collaborative situation during Tendering is the customer RFP documents analysis. The collaborative trace could be a considerable solution for these challenges, for example: short time, distributed teams, making the right decisions, deciding on the strategy etc. Besides, as shown in Section III, our model can be expanded and easily ported, for instance: in an agent-based CWE, filters can be implemented through different agents.

V. CONCLUSIONS AND PERSPECTIVES

In this paper, a definition of collaborative traces and a corresponding model have been proposed and discussed in the context of collaborative working environment in Section II and III. Furthermore, to validate the definition and test the
model, a typical application based on the E-MEMORAe2.0 platform was introduced and explained in Section IV. The concept of collaborative trace was introduced to meet several issues in CWE, which can be summarized in three key points: (i) Classify and organize users’ interactions a posteriori to understand the use of the CWE; (ii) share working experiences: the collaborative trace, which can assist both personal and group work; (iii) support the design of CWE - the different aspects of group modeling and user experience.

As a matter of fact, the issues of exploiting the collaborative traces is currently in progress. The recommendation based on the traces is one possible and practical way to facilitate this process and needs more consideration. Furthermore, the group collaboration interactive pattern can be studied by the model of collaborative trace.

REFERENCES