Web Sites as Agents' Environments: General Framework and Applications

Stefania Bandini, Sara Manzoni, Giuseppe Vizzari

Dipartimento di Informatica, Sistemistica e Comunicazione
Università degli Studi di Milano–Bicocca
Via Bicocca degli Arcimboldi 8 20126 Milan - Italy
tel +39 02 64487835 fax + 39 02 64487839
{bandini,manzoni,vizzari}@disco.unimib.it

Abstract. A web site presents an intrinsic graph–like spatial structure composed of pages connected by hyperlinks. This structure may represent an environment in which agents related to visitors of the web site are positioned and moved in order to track their navigation. To consider this structure and to keep track of these movements allows the monitoring of the site and its visitors, in order to support the enhancement of the site itself through forms of adaptivity, but also to introduce new forms of interaction among registered visitors. This paper presents a model supporting the collection of information related to user's behaviour in a web site, and an application supporting the proposal of hyperlinks based on the history of user's movement in the web site environment. Moreover the paper briefly describes a system implementing a context-aware form of interaction, supporting the communication among visitors of a web site through the exploitation of its structure.

1 Introduction

A web site presents an intrinsic graph–like spatial structure composed of pages connected by hyperlinks. However, this structure is generally not considered by web servers, which essentially act as a sort of extended and specific File Transfer Protocol servers [1], receiving requests for specific contents and supplying the related data. Several web–based applications instead exploit the structure of the sites itself to support users in their navigation, generating awareness of their position. For instance, many e–commerce sites emphasize the hierarchical structure linking pages related to categories (and possibly subcategories), included products and their specific views, and remind users' relative position (i.e. links to higher level nodes in the tree structure). Some specific web–based applications, mainly bulletin boards and forums (see, e.g., phpBB[1]), are also able to inform users about the presence of other visitors of the web site or even, more precisely, of the specific area of the site that they are currently viewing. Web site structure and users' context represent thus pieces of information that can be exploited to supply visitors a more effective presentation of site contents.

[1] http://www.phpbb.com/

Different visitors, however, may have very different goals and needs, especially with reference to large web sites made up of several categories and subcategories. This consideration is the main motivation for the research in the area of adaptive web sites [2]. The various forms of adaptation may provide a customization of site's presentation for an individual user or even an optimization of the site for all users. There are various approaches supporting these adaptation activities, but they are generally based on the analysis of log files which store low–level requests to the web server: this kind of file is generally made up of entries including the address of the machine that originated the request, the indication of the time and the resource associated to the request. In order to obtain meaningful information on users' activities these raw data must be processed (see, e.g., [3]), for instance in order to collapse requests related to various elements of a single web page (e.g. composing frames and images) into a single entry. Moreover this kind of information must be further processed to detect groups of requests that indicate the path (web pages connected by hyperlinks) that a user followed in the navigation. Recent results [4] show that this kind of analysis, also referred to as web usage mining, could benefit from the consideration of site contents and structure.

This paper proposes to exploit the graph-like structure of a web site as a Multi–Agent System (MAS) environment [5] on which agents representing visitors of the web site are positioned and moved according to their navigation. In particular, in this case, the environment is a virtual structure which allows the gathering of information on user's activities in a more structured way, simplifying subsequent phases of analysis and adaptation of site contents. Furthermore, part of the adaptivity could be carried out without the need of an off-line analysis, but could be the result of a more dynamic monitoring of users' activities. In particular, the paths that are followed by users are often related to recurrent patterns of navigation which may indicate that the user could benefit from the proposal of additional links providing shortcuts to the terminal web pages. Index pages may thus be enhanced by the inclusion of links representing shortcuts to the typical destinations of the user in the navigation of the web site. Users without a relevant history (and also anonymous or unrecognized ones) may instead exploit the paths that are most commonly followed by site visitors. Moreover such an information could also be communicated to the webmaster suggesting possible modifications to the static predefined structure of the site. This approach provides thus both a support for site optimization, but also for the customization to specific visitor's needs and preferences.

The metaphor of a web site as an environment on which users move in search for information is not new (see, e.g., [6] but also more recent approaches such as [7]), and its application to web site adaptation resembles the emergent, collective phenomenon of trail formation [8] which is can be identified in several biological systems. However this proposal provides more than just gather information on users' behaviours for sake of web pages adaptation or navigation support, but exploits the MAS environment to provide users a means for mutual perception and interaction. In fact information related to users' positions on the

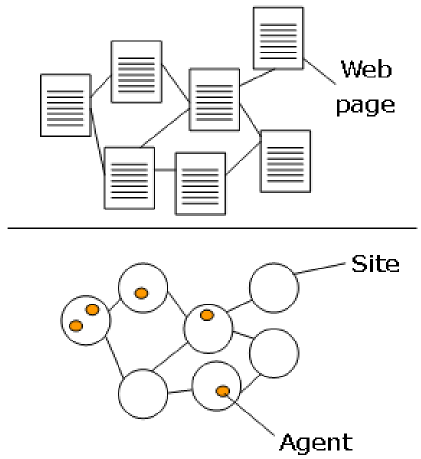

Fig. 1. The diagram shows a mapping between a web site structure and an agent environment.

environment representing the web site can also be used to supply them awareness information on other visitors which are currently browsing the same page or area of the site. Moreover to keep track of this information allows the conception of a form of interaction among users that is based on their positions on the site. Essentially, more than just showing a user the other registered visitors that are "nearby" (i.e. viewing the same page or adjacent ones), the system should also allow to communicate with them. This form of interaction clearly requires the adoption of a supporting technology that goes beyond the request/response form (e.g. a Java applet). This kind of interaction represents a hybrid between a common web site and an instant messenger (see, e.g., ICQ2Go![2]).

The following section describes the general framework of this approach, the mapping between the web site structure and agents' environment, while Section 3 describes the kind of gathered information on agents' movement in their environment. Section 4 describes an application providing the exploitation of this kind of information for the adaptation of web pages, both for customization and optimization, while Section 5 briefly introduces a system supporting a context-aware form of interaction among visitors of the same web site. Concluding remarks and future developments will end the paper.

2 Site Structure and Agents' Positions

A web site is made up of a set of HTML pages (generally including multimedia contents) connected by means of hyperlinks. It is possible to obtain a graph-like structure mapping pages to nodes and hyperlinks to edges interconnecting these

[2] http://go.icq.com/

nodes. This kind of spatial structure could be exploited as an *environment* on which agents related to site visitors are placed and move according to the related users' activities. A diagram showing a sample mapping among a web site and this kind of structure is shown in Figure 1.

This structure can be either static or dynamic: for instance it could vary according to specific rules and information stored in a database (i.e. database driven web sites). However, this kind of structure (both for static and dynamic web sites) can generally be obtained by means of a crawler (see, e.g., Sphinx [9] and the related WebSphinx project[3]); then it could be maintained by having periodic updates.

Given this spatial structure, a multi-agent model allowing an explicit representation of this aspect of agents' environment is needed to represent and exploit this kind of information. Environments for Multi Agent Systems [10] and situated agents represent promising topics in the context of MAS research, aimed at providing first class abstractions for agents environment (which can be more than just a message transport system), towards a clearer definition of concepts such as *locality* and *perception*. There are not many models for situated agents, which provide an explicit representation of agent's environment. Some of them are mainly focused on providing mechanisms for coordinating situated agent's actions [11], other provide the interaction among agents through a modification of the shared environment (see, e.g., [12,13]). An interesting approach that we adopted for this work is represented by the Multilayered Multi Agent Situated Systems (MMASS) [14] model. MMASS allows the explicit representation of agents' environment through a set of interconnected layers whose structure is an undirected graph of nodes (also referred to as sites in the model terminology; from now on we will use the term node to avoid confusion with web sites). The model was adopted given the similarity among the defined spatial structure of the environment and the structure underlying a web site. Moreover the model defines a set of allowed actions for agents' behavioural specification (including a primitive for agents' movement); for this specific application, however, the constraint which limits the number of agents positioned in a node was relaxed. In fact there is no limit to the number of users that are viewing the same web page.

Moreover a platform for the specification and execution of simulations based on the MMASS model [15] was exploited to implement the part of the system devoted to the management of agents in their environments. The definition of spatial structure of the environment was supplied by the previously introduced crawler, while agents' movement is guided by external inputs generated by the requests issued by the related web site visitor. The general architecture of the system is shown in Figure 2: the *Agent server* module is implemented through the MMASS platform, while the *Web server* is represented by SnipSnap[4], a Java-based weblog and wiki software. The highlighted *Tracker* module is a implemented through a Java Servlet, which is invoked by every page of the site but does not produce a visible effect on the related web page. Contents created

[3] http://www-2.cs.cmu.edu/ rcm/websphinx/
[4] http://snipsnap.org

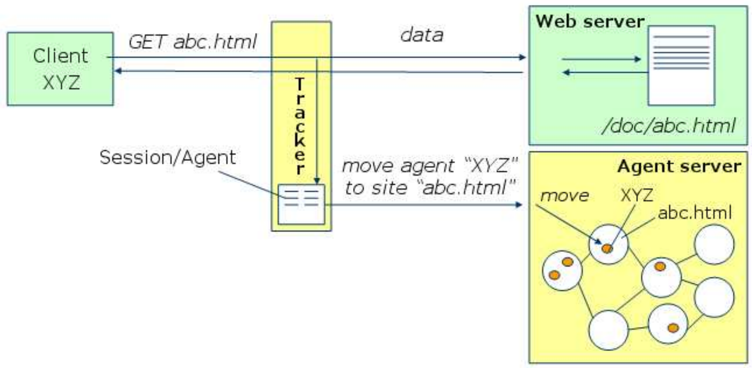

Fig. 2. A diagram showing how user actions influence the related agent through the capture of requests by the Tracker module.

through the SnipSnap page creation facility automatically include this invocation, which is specified in the template for new pages, while other contents (e.g. specific HTML pages, Java Server Pages) should explicitly state it in order to be included in the site spatial structure. The Tracker is responsible for the management of user authentication and requests, but it is also responsible for the creation of agents related to visitors and for the triggering of their movement in the environment related to the web site.

In particular, when a user makes his/her first page request, the Tracker tries to set a cookie on the client including the session information. If the cookie is accepted, then it is possible to use the session information and the *Referer* parameter of the HTTP request header to track user's movement in the graph related to the web site. Requests from clients not accepting cookies will thus not be monitored.

Moreover, the management of agents creation and movement is not as simple as its intuitive description might indicate. In fact, the same user could be using different browser pages or tabs to simultaneously view distinct pages of the site. In other words, a user might be simultaneously following different trajectories in his/her web site navigation. In order to manage these situations, a user can be related to different agents, and his/her requests must be associated to the correct agent (possibly a new one). Finally, agents related to finished (or interrupted) user navigation should be eliminated by the system, storing the relevant part of their state in a persistent way, until the related user requires again a page of the site. In particular, remote users' requests may be divided into two main classes, according to their effects on the Tracker and Agent server:

– *creating a new agent*: whenever a new user requires a web page, the Tracker will invoke the Agent Server requiring the creation of an agent whose starting position is the node related to the required page; the same effect is generated by a request coming from an already registered user which was not present in the system, but in this case information related to previous user agents

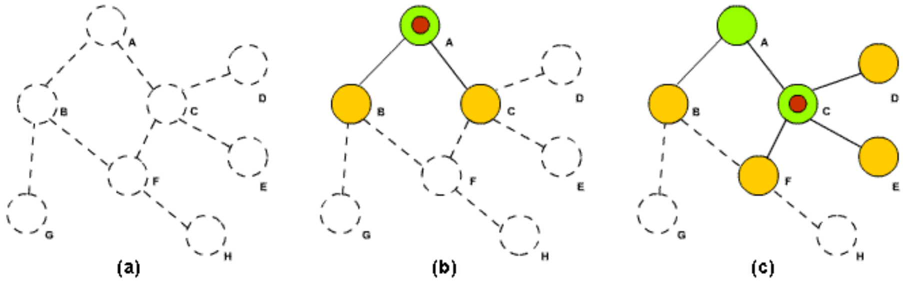

Fig. 3. The diagram shows the dynamic management of the web space according to actual requests captured by the Tracker.

must be retrieved in order to determine the new agent's state; finally, when an already registered and active user requires a page that is not adjacent to its current one, a new agent related to the new browsing activity must also be created;

- *generating the movement of an agent*: when the viewer of a page follows one of the provided links, the related web browser will generate a request for a page that is adjacent to one of the related agents which must be moved to the node related to the required page; whenever there are two or more agents in positions that are adjacent to the required page, in order to solve the ambiguity and choose the agent to be moved, the Tracker will invoke the Session object in which it stores the current URL related to the viewed page.

Finally, the Tracker can also represent a resource in the management of the spatial structure of the environment. Whenever the high dynamism of the structure (and in particular the addition of new pages) or particular architectural choices or web server configurations prevent the adoption of the web crawler solution, the Tracker can actually create sites related to pages that were not previously present in the environment, interacting with the Agent Server. Figure 3 shows a sample situation in which the spatial structure of the environment is not present until actual requests are captured by the Tracker. In (a) no page of the web site is still present in the environment, while in (b) a user just requested the page A. The latter was analyzed and outgoing links were identified, so three sites were created, respectively related to pages A, B and C; user's request of page C generates in turn the situation shown in (c), that is, the creation of sites related to pages D, E and F.

A prototype implementation of this approach to the creation and management of the spatial structure of the web site (as well as the web crawling solution) was realized and tested in a small scale web site, but it must be noted that this work is not meant as a contribution to solving the "invisible web" [16] issues (i.e. pages on the World Wide Web that are not simply indexed by common search engines and related technologies). It represents instead a possible approach to the design and development of web sites exploiting a MAS architecture supporting a novel form of users' monitoring, a simple adaptation approach and a new form interaction among users.

The following section will describe how the raw information that can be gathered thanks to the above described framework can be processed in order to obtain higher level indications on users' behaviours.

3 Gathered Information: Users' Traces

This system allows to gather and exploit two kinds of information: first of all situated agents related to web site visitors have a perception of their local context, both in terms of relative position, adjacent nodes and presence of other visitors; second, agents may gather information related to the paths defined by the browsing activities or the related user in the site itself.

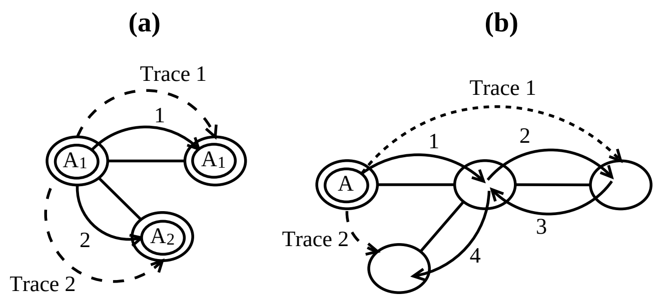

Fig. 4. A diagram describing two traces that are derived by a sequence of user requests.

There are inherent issues in determining in a precise way the actual users' activities on the web site, due to the underlying request/response model: the only available indications on these activities can be obtained by requests captured by the Tracker. In particular, we have an indication of the page that was required by a user and the time-stamp of the request. Starting from this raw information we can try to detect *emerging links*, which are hyperlinks that are not provided by the structure of the site but can be derived by the behaviour of specific visitors. To this purpose, the concept of *trace* was introduced as a higher level information describing the behaviour of a user. A trace synthesizes a path followed by a user, from the web page representing his/her entry point, to a different point of the environment (i.e. another web page) which may represent an interesting destination. Every agent related to a visiting user is associated to a *temporary* trace, and it may generate several actual traces (also called *closed* traces) in the course of its movement in the environment.

Formally a trace is a three-tuple $\langle A_{Id}, Start, Dest \rangle$, where A_{Id} represents the identifier of the agent to which the trace is related, while $Start$ and $Dest$

indicate the starting and destination node related to the browsing sequence which generated the trace. A new trace is generated when a user enters the site, triggering the creation of a related agent. The starting trace has a null value for the destination node. Subsequent requests by the user generated following hyperlinks will bring the related agent to an adjacent node, and the the *Dest* field of the corresponding trace will be modified in order to reflect user's current position. Non trivial traces provide *Start* and *Dest* nodes that are not directly connected by means of a hyperlink.

There are two relevant exceptions to the basic rule for trace update, that are related respectively to the *duplication* of a trace and to its *closing*. According to the previously introduced informal definition, a trace should be coherent in time and space. In fact, whenever the same user requires simultaneously two or more different pages he/she is probably following distinct search trajectories, possibly even related to different goals. In this case, as previously introduced, the Tracker will detect this situation and create additional agents that refer to the same user. Figure 4 shows two sample situations providing respectively trace duplication and closing: in (a) the user has chosen to open a hyperlink in a new browser page (request 1) and then has followed another link in the first browser page (request 2). According to the previously described Tracker behaviour, two agents are now associated to the user, and they are associated to different traces sharing the *Start* field.

In (b), instead, the user has followed links 1 and 2 from the starting page, then he/she made a step back (request 3) and eventually moved to the last known position (request 4). The step back causes the closure of the temporary trace associated to the agent (Trace 1 in the Figure), and the creation of a new temporary one with the same *Start* field (Trace 2). In this case the step back may have different interpretations: it could refer to a negative evaluation of the page contents but it could also indicate the fact that the user has found what he/she was searching for. An information that could be exploited to determine if the *Dest* field of the trace was interesting for the user is the time interval between request 2 and 3: for instance, given Δt_d a threshold indicating the minimum time required to reasonably inspect the content of a specific web page, if $timestamp(3) - timestamp(2) < \Delta t_d$ then Trace 1 could be ignored. However, the mere interval between the two requests is not a safe indicator of the fact that the page was actually viewed and considered interesting.

In fact, the time spent on a web page is also important in order to determine when a temporary trace must be closed. In fact, whenever a user does not issue requests for a certain time we could consider that his/her browsing activity has stopped, possibly because he/she is reading the page related to the *Dest* field of the trace associated to the related agent. In other words, every agent has a timer, set to the previously introduced threshold Δt_d, which is set when the agent is created and it is reset whenever it moves. The action associated to this timer specifies that its temporary trace becomes closed, and a new timer is set: the action associated to this second timer caused the disappearance of the agent from the system, and the storage of the related state.

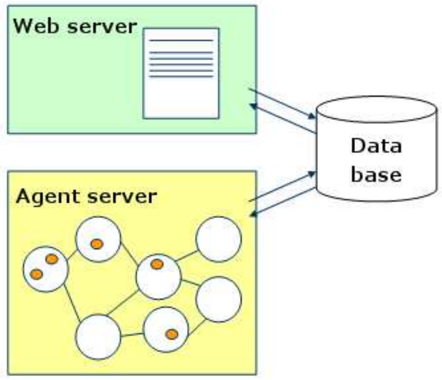

Fig. 5. Users' actions, captured by the related agents, influence the construction of subsequent web pages.

It is important to note that even anonymous visitors (i.e. non authenticated ones) whose clients are accepting cookies, can be tracked and can thus generate traces, although anonymous ones. The latter can be exploited for sake of web optimization but are not relevant for sake of user specific site customization.

Figure 5 shows how the information generated by user agents, and in particular traces, can be used to influence the new pages that will be generated by the Web server, and more precisely by the SnipSnap based Content Management system. In fact the latter uses information stored in a database to compose the required web pages; agents store information related to closed traces into this database, and a specific dynamic user interface element exploits this information to propose links that are not included in the basic structure of the site that are considered interesting, according to the previous user's behaviour. The following section will more thoroughly discuss the application of this framework for web site adaptivity.

4 Web Site Adaptation

4.1 Proposed Approach

The adopted instrument for the dynamic generation of web pages based on the content of a database organizes the structure of pages in blocks. The implemented system provides a static header block, including relevant areas of the web site, a left column providing dynamic additional information, such as the current user position in the structure of the web site and relevant links, and a main central area in which the specific current content is shown in details. The area which is interested in the first experimentation of this approach to content adaptivity is included in the left column. It is aimed at showing a visitor emerging links, that are hyperlinks not included in the predefined structure of the site but are considered interesting according to the history of the related user. These emerging links have some kind of relationship with the previously introduced traces, which represent behaviours and movements of a user in a web site. The strategy which is adopted to select the most relevant traces to be presented to a given user in a given situation represents the behavior of an interface agent whose responsibility is the management of this adaptive sub-block of the user interface related to the web site. Figure 6 presents a screenshot of a sample adapted web page: the visitor is recognized and his/her movement are monitored by the Tracker. The lower part of the left column presents three links related to stored traces related to the same user.

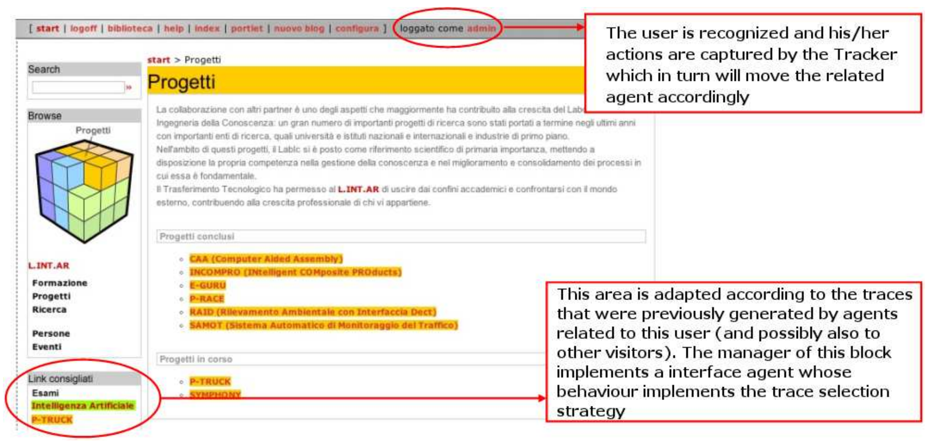

Fig. 6. A screenshot of a web page adapted according to gathered traces and interface agent selection strategy.

A first element of this strategy is adopted when new users (or non authenticated ones) enter the site. In this case the user has no previous history (or it is not possible to correlate the user with his/her history), and the adopted strategy considers all stored traces, not considering the user which generated them. An additional information that is stored with traces is the number of times that the related trace was effectively selected and shown to a user and the number of times that the related link was effectively exploited by a user. This kind of information allows to obtain an indication of the success rate of the hyperlinks that were chosen by the interface agent, and can be exploited by this agent to select the traces to be shown in the adaptive block. Furthermore this success rate can be used by the web master to consider which traces should be considered as emerging links to be included in the predefined site structure. Summarizing, the interface agent, in order to select which traces must be proposed as emerging links, considers two kinds of information: the occurrence of trace generation and the success rate of the traces that were proposed.

When the interface agent has an indication of the user which issued the request, it may focus the selection activity to those traces that compose the history of user's activities in the web site, in a web customization framework. In fact traces include an indication of the agent which generated them, and in turn agents are related to registered users. As for the anonymous or new user case, also this strategy must consider both the occurrence of traces and their success rate. Moreover, in order to focus on a specific user's history but do not waste the chance to exploit other users' experiences, just two of the three available slots for emergent links are devoted to traces that were generated by that user and one is selected according to the strategy adopted for anonymous or new users.

These strategies for the exploitation of the gathered and stored traces, based on users' behaviours and movement in the web site environment, represent a very simple way of exploiting this kind of information without requiring an off-line analysis of the logs generated by the web server. The design, implementation and test of more complex strategies, for instanced based on details of the outcomes of emerging link proposals (e.g. which user effectively followed the suggested adaptive hyperlink) are object of future works.

4.2 Related Works

There are several approaches and relevant experiences in the area web site adaptation. The Avanti project [17] provides an automated customization of web site contents, basing on user modelling techniques and analysis of their behaviours. It also provided a specific attention to specific needs of elderly and partially disabled users.

The Footprints system [6] instead provides a site optimization through the metaphor of site visitors leaving traces in their navigation. These signals accumulate in the environment, generating awareness information on the most frequently visited areas of the web site. No user profile is needed, as visitors are essentially provided this information which could represent an indicator of the most interesting pages to visit. The metaphor of the structure of the web site as an environment on which visitors move in their search for information is very similar to the one on which the proposed framework is based, but we also propose the exploitation of the gathered information on users' paths for user specific customization.

Other approaches provide instead the generation of index pages [3], that are pages containing links to other pages covering a specific topic. These pages, resulting from an analysis of access logs aimed at finding clusters grouping together pages related to a topic, are proposed to web masters in a computer-assisted site optimization scheme. A different approach provides the real-time generation of shortcut links [18], through a predictive model of web usage based on statistical techniques and the concept of expected saving of a shortcut, which considers both the probability that the generated link will be effectively used and the amount of effort saved (i.e. intermediate links to follow). In particular this framework is very similar to the one proposed here with reference to the aims of the overall system, but it incorporates a complex algorithm for off-line analysis of logs, while the proposed approach provides a light and dynamic generation of most probable useful links and the storage of these proposals and high level information on site usage for a possible further off-line analysis.

In the agent area, a relevant approach provides the adoption of information agents supporting users in their navigation [19], considering both his/her specific behaviour and the actions of other visitors and adopting multiple strategies for making recommendations (e.g. similarity, proximity, access frequency to specific documents).

A different approach to web site adaptation provides the adoption of a learning network to model the evolution of a distributed hypertext network, such as a web site [20]. Also in this case the adaptation provides a modification in the structure of a web site, and the concept of emergent link and the underlying mechanisms present a similarity with the learning rules adopted for that kind of learning network. However that approach also provides a modification in the architecture of the site and modifications in the web protocols, while this work aims at providing a solution that can be easily integrated with a traditional web architecture.

From this point of view, the introduced system supporting web site adaptation seems more similar to a recommendation system. A relevant type of recommender exploiting users' behaviours to decide which contents could be interesting for a certain visitor is represented by the collaborative filter approach [21]. The latter has been adopted in different recommendation systems, filtering mail messages, newsgroup articles and web contents in general, but typically requires users to rate these items. Moreover it generally provides a concept of explicit users descriptions through profiles which can be compared to determine similarity among them. The idea is that contents that received a high rating by a certain user could be considered interesting by a similar user. The introduced system instead does not require an explicit rating of contents, but it rather observes the frequency of specific navigation paths, and exploits emergent links for customization or optimization of site structure. However, the adaptive block of the page can include emerging links that are not related to the specific visitor who is currently browsing that page, but were generated by other users which frequently followed paths that the current one still did not follow. From this point of view, the system provides a very basic collaborative browsing scheme, but a more through analysis of a possible integration with this approach is object of current and future works.

5 Towards New Forms of Interaction

The metaphor of a web site as an environment on which agents related to visitors move according to their browsing activity allows to gather and exploit information on users' behaviours for sake of web pages adaptation in a collaborative agents framework. Another interesting possibility offered by this framework provides the exploitation of this structure and information to provide users a of context aware interaction form. While several web based applications are able to provide users an awareness information related to their position on the web site (e.g. the category of products their are currently viewing in an e-commerce site), and also an indication of other users that are currently viewing the same page, we propose to exploit this information on user context to support user interaction. Such a system represents a hybrid among a web application and an instant messenger, and could be fruitfully exploited in sites related to relatively small communities. In fact, the number of visitors that are viewing a single web page may be relatively high and difficult to present[5] in an effective and usable way in a block of a web page or in a separate page that however should not occupy a large portion of user screen. In order to support this new form of interaction among web site visitors, the general architecture of the system must be modified. In particular, a client-side component able to establish and maintain a connection with the Agent server must be included. A possible approach to tackle this problem provides the inclusion of an applet as a block of the web page structure which is constantly presented to the users (see Figure 7 for a diagram of the modification of the general architecture).

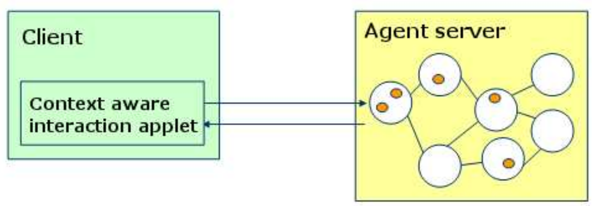

Fig. 7. In order to support interaction among web page visitors a client-side component (Context aware interaction applet) must be included in the system architecture.

The proposed form of interaction is context aware as a visitor is able to perceive other users currently viewing the same web page or adjacent ones and can try to interact with them. The perception of other users and the interaction is thus mediated by the environment, and given the fact that generally the structure of a web site reflects an underlying semantics (e.g. index pages which connect elements of a category) the concept of adjacency can be a relevant contextual element.

The MMASS multi-agent model provides both a concept of agent perception and two mechanisms for interaction: a direct form of interaction among adjacent agents is provided through the *reaction* operation, but agents are also able to *emit fields* which propagate in the spatial structure of the environment and may be perceived by other agents. The first mechanism may be invoked by a user which tries to establish an interaction with another one, provided that a preliminary agreement phase is successfully carried out. This phase represents

[5] See, e.g., the "What's Going On?" section of RPGnet forums (http://forum.rpg.net).

a possibility for a user to ignore incoming interaction requests. The other form of interaction instead provides the diffusion of an information conveyed by a field which may represent a message of general interest for visitors of a specific area of the site. Such message could represent a help request to other visitors of the area which could be interested in the same subject. In this framework, the design of diffusion strategies for this kind of field should take into account the underlying conceptual structure of the web site. For instance, in an e-commerce site, fields generated in a page related to a product could be related to a request of information on that subject. These fields should thus be diffused in pages related to the product category and other instances of that category, but should probably not be diffused into areas related to other categories. Figure 8 shows this sample diffusion strategy.

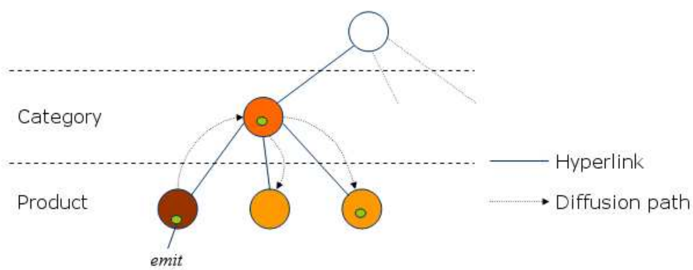

Fig. 8. A sample diffusion strategy in an e-commerce web site: fields are propagated inside a category but do not reach other ones.

In addition to the agreement phase which precedes the effective reaction, the model provides the definition of the perceptive capabilities of every agent, and thus of every user. The actual sensitivity to fields also depends on agent state, so there is the possibility to model and implement different levels of availability of a user to incoming interaction requests. For instance, in the e-commerce example, some agents could be related to operators of the product support service, and could thus be more sensitive to help requests, while the casual visitor of the site could be less sensitive to these messages.

6 Conclusions and Future Developments

This paper introduced a general framework providing the adoption of a web site as an environment on which agents related to visitors move and possibly interact. This approach allows the gathering of a more structured form of information on users' behaviours and activities in the web site. The concept of emerging links and traces have been introduced in order to support an application exploiting information on users' browsing history for sake of web pages adaptation. The introduced framework and the application to web site adaptation have been designed and implemented, exploiting a platform supporting systems based on the MMASS model.

A campaign of tests aimed at evaluating the effectiveness of the adaptation approach, and also for sake of tuning the involved parameters (e.g. timings, number of presented possible emerging links) is under way, in the context of a collaboration with the Italian company Cosmovision Srl. This evaluation will provide both forms for user interviews and the exploitation of the gathered information of the success rate of proposed adaptive hyperlinks. The results of this evaluation might also lead to consider the modelling, design and implementation of more complex trace selection strategies, and thus a more complex behaviour for the interface agent.

An application exploiting the data gathered by the system in order to support the monitoring and visualization of the web site structure and utilization is currently being developed. Future works will be focused on the introduction and exploitation of higher level semantic information related to the site structure and contents, and thus agents' environment, aimed at providing additional forms of adaptation, including images and multimedia contents. A further development provides the design and implementation of a prototype supporting the context-aware interaction among web site visitors.

References

1. Tanenbaum, A.S.: Computer Networks - third edition. Prentice Hall (1996)
2. Perkowitz, M., Etzioni, O.: Adaptive Web Sites: an AI Challenge. In: Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence. (1997) 16–23
3. Perkowitz, M., Etzioni, O.: Adaptive Web Sites. Communications of the ACM **43** (2000) 152–158
4. Cooley, R.: The Use of Web Structure and Content to Identify Subjectively Interesting Web Usage Patterns. ACM Transactions on Internet Technology **3** (2003)
5. Weyns, D., Parunak, H.V.D., Michel, F., Holvoet, T., Ferber, J.: Environments for Multiagent Systems State-of-the-Art and Research Challenges. In: Environments for Multi-Agent Systems, First International Workshop (E4MAS 2004). Volume 3374 of Lecture Notes in Computer Science., Springer–Verlag (2005) 1–47
6. Wexelblat, A., Maes, P.: Footprints: History-Rich Tools for Information Foraging. In: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM Press (1999) 270–277
7. Liu, J., Zhang, S., Yang, J.: Characterizing Web Usage Regularities with Information Foraging Agents. IEEE Transactions Knowledge and Data Engineering **16** (2004) 566–584
8. Helbing, D., Schweitzer, F., Keltsch, J., Molnár, P.: Active Walker Model for the Formation of Human and Animal Trail Systems. Physical Review E **56** (1997) 2527–2539
9. Miller, R.C., Bharat, K.: Sphinx: a Framework for Creating Personal, Site-Specific Web Crawlers. Computer Networks and ISDN Systems **30** (1998) 119–130

10. Weyns, D., Michel, F., Van Dyke Parunak, H., eds.: Environments for Multi-Agent Systems, First International Workshop (E4MAS 2004). Volume 3374 of Lecture Notes in Artificial Intelligence., Springer–Verlag (2004)
11. Weyns, D., Holvoet, T.: Model for Simultaneous Actions in Situated Multi-Agent Systems. In: First International German Conference on Multi-Agent System Technologies, MATES. Volume 2831 of Lecture Notes in Computer Science., Springer–Verlag (2003) 105–119
12. Mamei, M., Zambonelli, F., Leonardi, L.: Co-fields: Towards a Unifying Approach to the Engineering of Swarm Intelligent Systems. In: Engineering Societies in the Agents World III: Third International Workshop (ESAW2002). Volume 2577 of Lecture Notes in Artificial Intelligence., Springer–Verlag (2002) 68–81
13. Hadeli, K., Valckenaers, P., Zamfirescu, C., Van Brussel, H., Saint Germain, B., Hoelvoet, T., Steegmans, E.: Self-Organising in Multi-Agent Coordination and Control Using Stigmergy. In: Engineering Self-Organising Systems: Nature-Inspired Approaches to Software Engineering. Volume 2977 of Lecture Notes in Computer Science., Springer–Verlag (2004) 105–123
14. Bandini, S., Manzoni, S., Simone, C.: Dealing with Space in Multi–Agent Systems: a Model for Situated MAS. In: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems, ACM Press (2002) 1183–1190
15. Bandini, S., Manzoni, S., Vizzari, G.: Towards a Specification and Execution Environment for Simulations Based on MMASS: Managing At–a–distance Interaction. In Trappl, R., ed.: Proceedings of the 17th European Meeting on Cybernetics and Systems Research, Austrian Society for Cybernetic Studies (2004) 636–641
16. Sherman, C., Price, G.: The Invisible Web: Uncovering Information Sources Search Engines Can't See. CyberAge Books (2001)
17. Fink, J., Kobsa, A., Nill, A.: User-Oriented Adaptivity and Adaptability in the Avanti Project. Technical report, Microsoft Usability Group (1996)
18. Anderson, C.R., Domingos, P., Weld, D.S.: Adaptive Web Navigation for Wireless Devices. In: Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence. (2001) 879–884
19. Pazzani, M.J., Billsus, D.: Adaptive Web Site Agents. Autonomous Agents and Multi-Agent Systems **5** (2002) 205–218
20. Bollen, J., Heylighen, F.: Algorithms for the Self-Organisation of Distributed, Multi-User Networks. Possible Application to the Future World Wide Web. In Trappl, R., ed.: Proceedings of the 13th European Meeting on Cybernetics and Systems Research, Austrian Society for Cybernetic Studies (1996) 911–916
21. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J.: Grouplens: an Open Architecture for Collaborative Filtering of Netnews. In: CSCW '94: Proceedings of the 1994 ACM conference on Computer Supported Cooperative Work, ACM Press (1994) 175–186