A Multi-Agent Approach for Web Adaptation

A. Jorge Morais

Universidade Aberta (Portuguese Open University), Lecturer
Faculty of Engineering of the Univeristy of Porto, PhD Student
Laboratory of Artificial Intelligence and Data Analysis (LIAAD – INESC Porto L. A.), PhD Student Researcher
jmorais@univ-ab.pt

Abstract Web growth has brought several problems to users. The large amount of information that exists nowadays in some particular Websites turns the task of finding useful information very difficult. Knowing users' visiting pattern is crucial to owners, so that they may transform or customize the Website. This problem originated the concept known as Adaptive Website: a Website that adapts itself for the purpose of improving the user's experience. This paper describes a proposal for a doctoral thesis. The main goal of this work is to follow a multi-agent approach for Web adaptation. The idea is that all knowledge administration about the Website and its users, and the use of that knowledge to adapt the site to fulfil user's needs, are made by an autonomous intelligent agent society in a negotiation environment. The complexity of the problem and the inherently distributed nature of the Web, which is an open, heterogeneous and decentralized network, are reasons that justify the multi-agent approach. It is expected that this approach enables real-time Web adaptation with a good level of benefit to the users.

1 Introduction

The World Wide Web has experienced a quick growth during the past decade. Nowadays, having a Website is almost mandatory for organizations, in order to easily deliver information to the general audience. With the advent of e-commerce (electronic commerce), organizations became more concerned with the problem of organizing all the information efficiently, so that it may be easy to find every product a user is searching or the organization wants clients to purchase, or to present useful information to the public.

Let us consider, for instance, that a client is searching a laptop with some particular characteristics (processor, memory, hard disk, screen size, etc.). The Website structure may be organized by different suppliers, which means clients will have to search inside each one in order to compare models. Suppose that this is a standard procedure of most of the users. Discovering such a pattern might lead to considering a solution like suggesting similar laptops from other suppliers in the same page. Another solution might be reorganizing the Website structure to organize laptops by characteristics instead of suppliers.

The problem of dealing with large sets of data was already the motivation for the area of Data Mining and Knowledge Discovery [1]. The idea was to take advantage of the

large quantity of data from previous transactions that were kept in organizations, finding useful information that is not easily reachable. This led to the use and development of a set of algorithms that automatically extracted and discovered important patterns, and represented them in an understandable way. Considering the large number of pages in the Web, it became natural the application of Data Mining and Knowledge Discovery to the Web scope, which resulted in the new area of Web Mining [2][3]. Web Mining can be defined as the application of data mining algorithms to extract and discover useful information (documents and services) from the Web. The results may enable the Website owner to improve its structure by reorganizing the Website or providing users navigation assistance.

There are three main research sub areas of Web Mining. Web Content Mining is focused on the content of Web pages, mainly the search of documents satisfying a given criteria, and summarization. Web Structure Mining takes into account the structure of the Website, its pages, and hyperlinks between them, enabling categorization of Web pages and finding similarities and relationships between Websites. Web Usage Mining tries to extract and discover useful information from previous visits to the Website, and is useful for Web adaptation and market-based analysis.

The problem of Web adaptation is not new. Some work has already been made in this particular area in the last years. One of the current solutions that are being proposed for this problem is using autonomous agents. Multi-Agent Systems [4] is a research area that has been in great development over the last decade, and has some particular characteristics that fit in this problem. In fact, it was already proposed to use a multi-agent approach, because of its flexibility and its capability of dynamic adaptation to the Web applications needs [5]. Moreover, Multi-Agent Systems are already used for automatic retrieval and update of information in Websites [6]. An implementation of a recommender system using this approach was already proposed in [7].

This paper is organized as follow. We start by presenting previous approaches and applications in the area of recommender systems and multi-agent systems. Then we present an overview of the project, showing its scope and previous work. Finally, we describe some conclusions and future work.

2 Previous approaches and applications

A global vision on adaptive Web sites based on user interaction analysis is given in [8]. In fact, only less ambitious approaches were proposed, such as reorganization of the Website [9], use of recommendations in the pages [10], automatic categorization of user actions [11], or seek of relevant Web sequence paths using Markov models [12].

Recommendation systems include the combination of clustering with nearest neighbour algorithms [13], Markov chains and clustering [14], association rules [15], and collaborative and content-based filtering [16].

Web dynamics has been controlled, for instance, by efficient incremental discovery of sequential Web usage patterns [17], and on-line discovery of association rules [18].

Data-driven categorization of a Website usability may be done by typical usage patterns visualization [11] or with objective metrics [19].

Some platforms, like WebWatcher, use previous users' knowledge to recommend links [20]. AVANTI implements an adaptive presentation based on a model constructed from user actions [21]. WUM infers a tree structure from log records enabling experts to find patterns with predefined characteristics [22]. In [23]

it was proposed an integrated tool (HDM) to discover access patterns and association rules from log records in order to automatically modify hypertext organization.
In [5] it was proposed the use of a multi-agent platform for personalization of Web-based systems, given the flexibility of this approach and its dynamic adaptation to Website needs. Multi-agent approaches for developing complex systems, like Web adaptation, were defended in [24]. Intelligent agents may also be an important contribution for autonomic computing [25]. Such systems main characteristics are being complex systems with self-administration, self-validation, self-adjustment and self-correction. Web adaptation systems should also have these characteristics, because Website environment dynamics requires either a high degree of system automation or high allocation of human resources. Another important usage of multi-agent systems in this issue is the automatic collection and actualization of information in Websites [6].
There is already an implementation of a recommendation system [7]. In this work the author implements a market-based recommender system that is stated to be Pareto efficient, to maximize social welfare and all agents are individually rational. The author distinguishes between Internal Quality (INQ) and User Perceived Quality (UPQ). INQ of a specific recommendation is the sum of the weighted evaluation scores made of different techniques on different properties of a recommendation while UPQ is introduced by users. Agents are unaware of UPQ, and they must infer from the reward mechanism. A generalized first-price sealed-bid auction protocol is used to minimize the time for running the auction and the amount of communication generated.

3 Project overview

This project is a part of an undergoing work in the area of Adaptive Websites. We will now describe its scope, the preliminary work and compare it with related work in the same area.

3.1 Scope of the project

The work to develop during this thesis is to be implemented in a web adaption platform [26] that was developed during the site-o-matic project[1], and is already implemented and running (except the part described in this paper). The architecture of the platform is presented in figure 1.

Fig. 1 – Web adaption platform.

As the client opens the Website page in a browser, the information about the current page content and the actions performed are recorded in the data warehouse. The client browser also performs an automatic adaption request to broker. The broker acts as an intermediary between the client and the adapters, requesting hyperlinks to be shown in the browser. This activity is recorded in a logger, which also feeds the data warehouse. Finally the broker sends suggested hyperlinks to the client browser. The process will be repeated whenever the user chooses another hyperlink or performs some particular action in the browser (e.g., selecting a portion of the text with the mouse).
Our propose is that the construction and management of the Knowledge Base with information about the Website and its users, and the use of that knowledge to perform the adaptation according to users behaviour, should be done by a multi-agent system with negotiation. Each agent goal is to optimize the quality of the Website user experience by choosing adequate contents and to represent them in an appropriate manner (for instance, with an appropriate format that calls the attention of the user). They must collect and produce knowledge about users, eventually specializing themselves in a given segment of users or contents. Depending on the recommender system algorithm, agents will try to adapt a particular client using its or all users (or part of them) past behaviour.
Therefore, agents will have to compete for the limited space in the browser and the user's attention. They also have to compete for modelling each user. This attribution may be made randomly, centralized, or as result of a negotiation or an auction between agents. Each agent will have its virtual financial resources

[1] Supported by the POSC/EIA/58367/2004/Site-o-Matic Project (Fundação Ciência eTecnologia), FEDER e Programa de Financiamento Plurianual de Unidades de I & D.

(utilities) that rise when good decisions are made and decrease with the acquisition of user modelling and knowledge acquisition.
In the scope of the architecture of the platform in figure 1, the broker will be modified in order to perform an auction and manage the utilities received from and paid to the agents. The adapters will have to be given the additional capability of performing intelligent bids in order to maximize their revenues. When a request is made to the broker, all adapter agents will make a single bid, and the broker choose the best offer, informing the winner and the losers (with the value of the winning proposal, so that they can use this information in future bids). The agent that wins the auction will be given the task of adapting the Website for the respective client until the end of the session and will be rewarded according to the success of its recommendation.

3.2 Preliminary work

A small project was made in the scope of the course Multi-Agent Systems, in the Doctoral Programme in Informatics Engineering at Faculty of Engineering of the University of Porto, and consisted in a simple recommendation system, developed using the JADE platform [27].
The system used an applet in the client-side and a multi-agent system in the server-side. After a request from the user browser, three agents with different recommendation algorithms competed for the recommendation of hyperlinks that were sent to the applet in the client browser, offering an amount of its virtual financial resources. If users followed the hyperlinks, the agent was rewarded according to the number of pages and the order in which the respective hyperlinks were presented.
The project was important for studying and using the platform, and also showed that the approach was pertinent. However, some significant amount of time was usually spent while loading the applet. For this reason another concept will be used: AJAX [28]. It is not yet possible to estimate the time that will take to make an adaption, but since AJAX acts asynchronously and the adaption is processed in the server side, it is expected that it takes a very small amount of time.

3.3 Comparison with related work

This proposal approach is different in many aspects from the recommender system proposed in [7], although some knowledge about that work may be useful (for instance, some tests that were already made about the number of hyperlinks that should be used in recommendation, and some parameters that were already tuned). The cited work does not take into account the evolution of the market, which is not suitable for a daily updated Website. For instance, in an online news journal a user might classify a Web page as very interesting today but tomorrow it could loose all of its interest. But next week, if the issue of that Web page was again on the top of the news, it would be interesting again.
There are also other aspects that our work can improve such as taking into account client behaviour during each session and obtaining user's quality score implicitly rather than expecting their explicit scores.

4 Conclusions and future work

The current work is in an early stage. Rather than final conclusions we will present the expected contribution of the thesis and the respective work plan.

4.1 Expected contribution of the thesis

The main contribution of this thesis is the use of a market-based approach, using autonomous agents. Multi-Agent Systems [4] are a research area that has been in great development over the last decade, and have some particular characteristics that fit in this problem. In fact, in [5] it was proposed the use of a multi-agent approach, because of its flexibility and its capability of dynamic adaptation to the Web applications needs. Moreover, Multi-Agent Systems are already used for automatic retrieval and update of information in Websites [6].
It is our belief that design, construction, validation and management of a dynamic Website, capable of adapting itself automatically to the user needs, following high-level specifications, are tasks sufficiently complex to justify this approach. It is even more pertinent if we consider large Websites, where manual organization or definition of explicit rules is not easy, making the task of improving user experience harder.

4.2 Work plan

After defining the Web adaptation strategy, using Web mining algorithms and Multi-Agent Systems, and also the evaluation strategy for measuring its success, the primary goals are the design of the knowledge base and the definition of agents, possibly with some pilot implementations.

The platform has already some recommender systems implemented, using algorithms such as collaborative filtering, association rules, etc. The agents to be implemented will use these algorithms to perform adaption. The bidding strategy will be important for the success of the approach: if the market fails, an agent may become in a situation of monopoly, which means that the approach becomes useless. This means that all agents must have intelligent bid strategies, in order to maintain the equilibrium of the market.

Afterwards, a preliminary evaluation that may lead to important changes in the previous design phase will be held, resulting in the final specification for the implementation phase, with the integration of the Knowledge Base with log analysis problems and the implementation of the agents. Finally, the system will be evaluated, with eventual refinement of methods.

References

1 Fayyad, U. M., Piatetsky-Shapiro, G., Smyth, P., and Uthurusamy, R., editors, 1996. *Advances in Knowledge Discovery and Data Mining*. AAAI/MIT Press, Menlo Park, California.

2 Etzioni, O. 1996. The World Wide Web: Quagmire or gold mine? In *Communications of the ACM*, vol. 39, nr. 11, pages 65-68.

3 Cooley, R., Mobasher, B., and Srivastava, J. 1997. Web mining: Information and patterns discovery on the world wide Web. In *Proceedings of the ninth IEEE International Conference on Tools with Artificial Intelligence*, pages 558–567, Newport Beach, California.

4 Wooldridge, M., 2002. *An Introduction to MultiAgent Systems*. John Wiley & Sons.

5 Ardissono, L., Goy, A., Petrone, G., and Segnan, M., 2005. A multi-agent infrastructure for developing personalized web-based systems. ACM Trans. Inter. Tech., 5(1):47–69.

6 Albayrak, S., Wollny, S., Varone, N., Lommatzsch, A., and Milosevic, D., 2005. Agent technology for personalized information filtering: the pia-system. In *SAC'05: Proceedings of the 2005 ACM symposium on Applied computing*, pages 54–59, New York, NY, USA. ACM Press.

7 Wei, Y. Z., 2005. A Market-Based Approach to Recommendation Systems, PhD thesis, University of Southampton.

8 Perkowitz, M. and Etzioni, O., 2000. Towards adaptive web sites: Conceptual framework and case study. *Artificial Intelligence*, 118(2000):245–275.

9 Ishikawa, H., Ohta, M., Yokoyama, S., Nakayama, J., Katayama, K., 2002. Web usage mining approaches to page recommendation and restructuring. In *International Journal of Intelligent Systems in Accounting, Finance & Management*, 11(3), Pages: 137-148.

10 El-Ramly, M., Stroulia., E., 2004. Analysis of Web-usage behavior for focused Web sites: a case study. In *Journal of Software Maintenance and Evolution: Research and Practice*, 16(1-2), Pages: 129-150.

11 Berendt, B., 2002. Using Site Semantics to Analyze, Visualize, and Support Navigation. In *Data Mining and Knowledge Discovery*, Volume 6, Issue 1, Pages 37 – 59.

12 Borges, J. L., 2000. A Data Mining Model to Capture User Web Navigation Patterns, PhD thesis, University College London, University of London.

13 Mobasher, B., Dai, H., Luo, T., Nakagawa, M., 2002. Discovery and Evaluation of Aggregate Usage Profiles for Web Personalization. In *Data Mining and Knowledge Discovery*, Kluwer Publishing, Vol. 6, No. 1, pp. 61-82.

14 Cadez, I., Heckerman, D., Meek, C., Smyth, P., White, S., 2003. Model-Based Clustering and Visualization of Navigation Patterns on a Web Site. In *Data Mining and Knowledge Discovery*, Volume 7, Issue 4, Pages 399 – 424.

15 Jorge, A., Alves, M. A., Grobelnik, M., Mladenic, D., and Petrak, J., 2003. Web Site Access Analysis for A National Statistical Agency. In *Data Mining And Decision Support: Integration And Collaboration*, Mladenic, D., Lavrac, N., Bohanec, M., Moyle, S., Kluwer Academic Publishers.

16 Basilico, J., Hofmann, T., 2004. Unifying collaborative and content-based filtering. In *Proceedings of ICML'04, Twenty-first International Conference on Machine Learning*. ACM Press, New York.

17 Masseglia, F., Teisseire,, M., Poncelet, P., 2003. HDM: A client/server/engine architecture for real time web usage mining. In Knowledge and Information Systems (KAIS), Volume 5, Number 4, pp 439 - 465.

18 Lin, W., Alvarez, S. A., Ruiz, C., 2002. Efficient Adaptive-Support Association Rule Mining for Recommender Systems. In *Data Mining and Knowledge Discovery*, 6:83-105.

19 Spiliopoulou, M., and Pohle, C., 2001. Data mining for measuring and improving the success of web sites. In Ronny Kohavi and Forest Provost, editors, *Journal of Data Mining and Knowledge Discovery, Special Issue on E-commerce*, volume 5(1-2), pages 85-114. Kluwer Academic Publishers.

20 Armstrong, R., Freitag, D., Joachims, T., and Mitchell, T., 1995. WebWatcher: A learning apprentice for the world wide web. In *Proceedings of the AAAI Spring Symposium on Information Gathering from Heterogeneous, Distributed Environments*, pages 6–12, California.
21 Fink, J., Kobsa, A., and Nill, A., 1996. User-oriented adaptivity and adaptability in the AVANTI project. In *Designing for the Web: Empirical Studies*, Microsoft Usability Group, Redmond, Washington.
22 Spiliopoulou, M. and Faulstich, L. C., 1998. WUM: a tool for web utilization analysis. In *Proceedings of the International Workshop on the Web and Databases*, pages 184–203, Valencia, Spain.
23 Masseglia, F., Teisseire, M., and Poncelet, P., 2001. Real Time Web Usage Mining: a Heuristic Based Distributed Miner., In *Second International Conference on Web Information Systems Engineering (WISE'01)*, Volume 1, p. 0288.
24 Jennings, N. R., 2001. An agent-based approach for building complex software systems. In *Communications of the ACM*, vol. 44, no. 4, pp. 35-41.
25 Kephart, J. O., 2005. Research challenges of autonomic computing. In *ICSE '05: Proceedings of the 27th International Conference on Software Engineering*, pages 15–22, New York, NY, USA. ACM Press.
26 Domingues, M. A., Jorge, A. M., Soares, C., Leal, J. P., and Machado, P., 2007. A data warehouse for web intelligence. In *Proceedings of the 13th Portuguese Conference on Artificial Intelligence (EPIA 2007)*, 487–499.
27 JADE (Java Agent DEvelopment Framework) Website: http://jade.tilab.com. Access date: 01/11/2008.
28 Asynchronous Javascript And XML (AJAX), Mozilla Developer Center: http://developer.mozilla.org/en/docs/ajax. Access date: 01/11/2008.