

Multi-Agent Web Recommendations

Joaquim Neto^{3,4}, A. Jorge Morais^{1,2,4}

¹Faculty of Engineering of the University of Porto

²Laboratory of Artificial Intelligence and Decision Support (LIAAD – INESC TEC L. A.)

³National Laboratory of Civil Engineering (LNEC)

⁴Universidade Aberta (Portuguese Open University)

Abstract Due to the large amount of pages in Websites it is important to collect knowledge about users' previous visits in order to provide patterns that allow the customization of the Website. In previous work we proposed a multi-agent approach using agents with two different algorithms (associative rules and collaborative filtering) and showed the results of the offline tests. Both algorithms are incremental and work with binary data. In this paper we present the results of experiments held online. Results show that this multi-agent approach combining different algorithms is capable of improving user's satisfaction.

1 Introduction

Recommender systems technology usually applies technologies and methods from other information system areas, such as Human Computer Interaction and Information Retrieval. On most of those kinds of systems was possible to find algorithms and techniques used in Data Mining and Knowledge Discovery areas. The motivation of these areas is related to the analysis and processing of large datasets [1]. Considering the large number of pages in the Web, it became natural to apply this concept to the Web scope, resulting in the new area of Web Mining [2][3]

Recommender systems [3] have had several improvements over the last decade and are increasingly present in the websites. The recommendation of a set of Web pages of potential interest for the visitor of a website, based on the users' access history, aims to improve the users' experience, providing them with effective search tools. One of the current solutions that are being proposed to address 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, given their characteristics, allow the combination of multiple recommendation algorithms, increasing the chances of the suggested recommendations are effectively of users interest. The multi-agent approach is being proposed 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]. In previous work [7] we presented the implementation of a multi-agent recommender system, and the results of offline experiments.

In this paper we present the results of online experiments of the multi-agent recommender system, using a Website with photographs that was built for this purpose. Different incremental algorithms based on binary data produce item-based recommendations and make bids to provide the next set of recommendations to the user. Agents are cooperative in the sense they base their bids on client's satisfaction instead of their own revenue and they share the same data. However, their results are not combined in order to provide recommendations. Our goal is to show that online experiments confirm the results of the ones taken offline.

The remaining of the paper starts by presenting previous approaches and applications in the area of recommender systems and multi-agent systems, followed by the description of our approach. The results of the experiments online, and some conclusions and future work complete the paper.

2 Previous approaches and applications

In [8] is described a global vision on adaptive Web sites, which is based on user interaction analysis. Some 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 several algorithms and techniques, such as 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 Website usability may be done by typical usage patterns visualization [11] or with objective metrics [19].

Some platforms already implemented, like WebWatcher, use previous users' knowledge to recommend links [20], while AVANTI implements an adaptive presentation based on a model constructed from user actions [21], and 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.

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. The use of autonomous agents in a similar context can be found in [5], where a multi-agent platform for personalization of

Web-based systems was proposed, due to its flexibility and its dynamic adaptation to Website needs. Another important usage of multi-agent systems in this context is the automatic collection and update of information in Websites [6].

Some related work including a web adaption platform [26] and an implementation of collaborative filtering using an incremental approach [27] were also important for the architecture of our system.

3 Multi-Agent Approach

The multi-agent system recommender [7] was implemented taking into account that agents should answer rapidly to any request from another agent and prepare in advance for the next request, and tasks that involve a large amount of time (like updating the model) should not interfere with the performance of the system.

Recapitulating what was described in that paper, two recommender agents were created. The first one generates single-condition association rules and the second one uses a collaborative filtering algorithm. Both algorithms share a matrix $A_{n \times n}$, where n is the number of items (Webpages) and each $a_{ij} \in A$ registers the total number of co-occurrences of items i and j in the same session. The matrix is updated each time a session ends.

The single-condition association rules agent checks all possible rules $i \rightarrow j$, where i and j are items, taking into account two values (k number of sessions):

$$Support_{i \rightarrow j} = \frac{a_{ij}}{k} \quad Confidence_{i \rightarrow j} = \frac{a_{ij}}{a_{ii}}$$

Therefore, if a set of n recommendation is requested, the n best recommendations according to the confidence that satisfy minimum confidence and support requirements are proposed.

The collaborative filtering agent uses the same matrix to compute similarity, returning the top n most similar items:

$$sim(i, j) = \frac{a_{ij}}{\sqrt{a_{ii}} \sqrt{a_{jj}}}$$

Agent biddings are based on an accumulated score for each given item obtained from previous ratings – the best N are sorted and if the next selected item was in that set it receives a score $N-p+1$, where p is the ordered position of the item. To this score we add the percentage of the overall score to untie equal biddings:

$$Bid_{agent,item} = Score_{item} + \frac{1}{\#requests * N} * \sum_{i \in Items} Score_i$$

The multi-agent approach was implemented in Java, using the JADE platform [28]. The communication with the browser is implemented using AJAX [29], us-

ing XMLHttpRequest interface, so that the user can consult the Web page without losing interest. The interaction between the user and the recommender system, as well as the recommender system architecture is presented in figure 1.

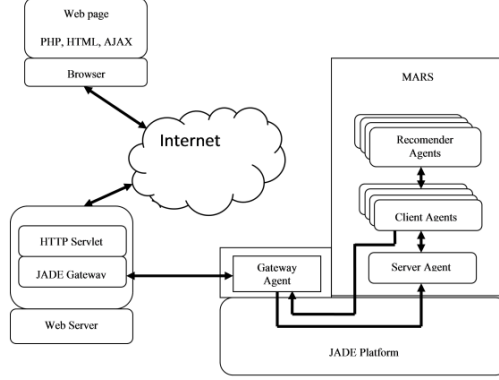


Fig. 1 Website and recommender system interaction

4 Experimental Results

In [7] we presented offline experiments focused on four datasets (obtained from real Web data records). In this paper we present and compare the results of online experiments with a photographs website [31] to the offline experiments.

Like in [7], we consider each time recommendations are made four possibilities:

- No item was followed (discarded, no implicit knowledge – end of session).
- The set of recommendations was empty.
- An item not in the recommendation set was followed.
- One of the recommendations was followed.

For evaluation of performance, since the algorithms are incremental, which means we do not have a fix split for train and test sets, the evaluation that fits better to our case is a per-user variant, where predictions are computed and the ranking metrics are calculated for each recommendation, and the average over all recommendations gives the final value [32].

There are two measures that we will use for evaluating recommendation: precision and recall [33]. Precision is the ratio of relevant items selected to number of items selected – it represents the probability that a selected item is relevant. Recall is the ratio of relevant items selected to total number of relevant items available. In our case, precision and recall are given by the following formulas (given N recommendations, and considering b , c and d of the list of possible situations above):

$$Recall = \frac{\#d}{\#b + \#c + \#d} \quad Precision = \frac{1}{N} \quad Recall = \frac{1}{N} \frac{\#d}{\#b + \#c + \#d}$$

This measure is also applied to the recommendation system, which combines agents' algorithms. When the recommendation set is incomplete or inexistent (because it is the first time the item appears, so there are no correlations yet), the system completes it with the most popular items.

In our online experiments we used a photographs website [31] that was built for this purpose. The website includes 344 photographs, distributed by 8 galleries, in the total of 353 Web pages. For recommendation we only consider the dataset of the 344 items with photographs. We decided to present 10 recommendations to the user, each time a Web page is being consulted.

The online experiments were undertaken in two different phases:

1. Recommendations were turned off, meaning that the system did everything as usual except presenting the recommendations to the user.
2. Recommendations were presented to the users and could be followed.

Table 1 shows the previous datasets used for offline experiments and new dataset used for online experiments – e1 and e2 refer to the respective phases.

Table 1 Datasets characteristics.

Dataset	#items	# records	#sessions	#records/#session	#records/#items
e-com	335	1409	413	3.411622	4.20597
pe200	200	2042	200	10.21	10.21
e1	344	2514	364	6.9066	7.3081
e2	344	2377	205	11.5951	6.9098

In order to analyse the results of our experiments for the two phases, we consider the tables 2 and 3, where the results are expressed in terms of the evaluation metrics (EM) recall and precision, considering association rules (AR), collaborative filtering (CF), the winner of auction (W) and the best possible result if the best recommendation was always chosen (B).

Table 2 Results of experiments online – considering all records, even those where the user follows an item that does not belong to the item set. The end of session records were discarded.

EM	AR		CF		W		B	
	e1	e2	e1	e2	e1	e2	e1	e2
Recall	10.60%	30.62%	13.30%	27.49%	12.05%	34.02%	15.30%	35.13%
Precision	1.06%	3.06%	1.33%	2.75%	1.20%	3.40%	1.53%	3.51%

Table 3 Results of experiments online – considering records where an eligible item was followed. All other records are treated as end of session.

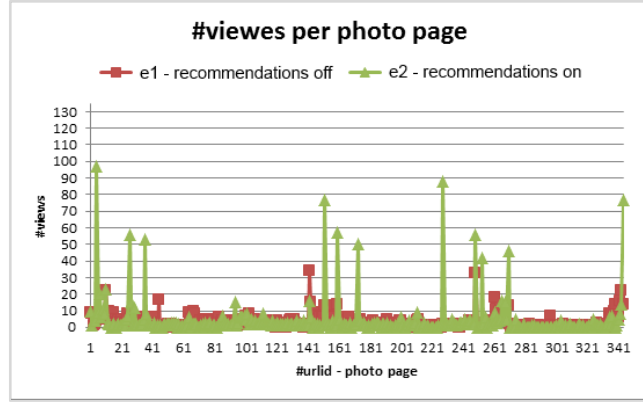
EM	AR		CF		W		B	
	e1	e2	e1	e2	e1	e2	e1	e2
Recall	27.91%	52.20%	35.01%	46.86%	31.70%	58.01%	40.27%	59.89%
Precision	2.79%	5.22%	3.50%	4.69%	3.17%	5.80%	4.03%	5.99%

Also of interest for analysis are the results presented in table 4, where we show the impact of the presence of the multi-agent recommender system (MARS) in terms of the percentage of views of items eligible for recommendation.

Table 4 Results of experiments online.

	MARS	
	e1	e2
%viewed items eligible for recommendation	47,33%	62,66%
%viewed items that follows a recommendation	23.76%	50.44%

The dispersion graph of Figure 2 shows the distributions of item views obtained on online experiments.

**Fig. 2** Item views distribution.

5 Discussion and future work

Comparing the two phases of the experiments, we observe that for phase one, with the recommendations turned off, the CF based agent gets a better result than MARS. However, for the phase two of experiments, with the recommendations turned on, the MARS is able to obtain better results than the individual algorithms. From table 3 we conclude that, with MARS, the times the users follow a recommended item grows from 31.70% to 58.01%. The score for MARS outperforms the AR based agent in 5.8% and the CF based agent in 11.15%.

Table 4 shows that with the recommendations turned on, the percentage of items (photographs) viewed, grows from 47.33% to 62.66%, the same happened with the percentage of those viewed items that followed a recommendation, 50.44% for e2 while for e1 is 23.76%.

Figure 2 shows that with recommendations turned on is possible to detect more peaks, as a consequence of the users followed the presented recommendations, however, besides this, we could see that the page view distribution is similar with that obtained with de recommendations turned off. So, the MARS does not significantly restrict the range of viewed items.

As we can observe, the MARS is able to clearly outperform the individual algorithms when the recommendations were presented to the user and could be fol-

lowed. We may also conclude that the increase in the items viewed, as well as in the followed items, when the recommendations were presented to the user, represents a benefit for the users as well for the website.

As future work, we intend to make more online experiments with this and other Websites, and to make some improvements to the MARS.

Acknowledgements

This work was funded by: Fundação para a Ciência e a Tecnologia (FCT), Ref. SFRH/BD/27571/2006; ERDF - European Regional Development Fund through the COMPETE Programme (operational programme for competitiveness), through FCT within project FCOMP -01-0124-FEDER-022701; PRODEI – FEUP.

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 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.
- 3 Jannach, D., Zanker, M., Felfernig, A., Friedrich, G., 2011. *Recommender systems: an introduction*. Cambridge University Press.
- 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 Morais, A. J., Oliveira, E., and Jorge, A. M., 2012. A Multi-Agent Recommender System. In S. Omatu et al. (Eds.): 9th International Conference on DCAI'12, AISC 151, pp. 281-288.
- 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: Integra-*

- tion 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 Massegli, 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 Massegli, 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 Miranda, C., and Jorge, A. M., 2008. Incremental Collaborative Filtering for Binary Ratings. In *Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. Volume 01 (WI-IAT '08), Vol. 1. IEEE Computer Society, Washington, DC, USA, 389-392.
 - 28 JADE (Java Agent DEvelopment Framework) Website: <http://jade.tilab.com>. Access date: 31/01/2014.
 - 29 Asynchronous Javascript And XML (AJAX), Mozilla Developer Center: <https://developer.mozilla.org/en/ajax>. Access date: 31/01/2014.
 - 30 Apache Derby Website: <http://db.apache.org/derby/>. Access date: 31/01/2014.
 - 31 1000 Palavras Fotografia: <http://www.1000palavras.pt/>. Access date: 31/01/2014.
 - 32 Herlocker, J. L., Konstan, J. A., Terveen, L. G., Riedl, J. T., 2004. Evaluating collaborative filtering recommender systems. In *ACM Transactions on Information Systems*. 22 (1): 5-53.
 - 33 Cleverdon, C., Kean, M., 1968. Factors Determining the Performance of Indexing Systems. *Aslib Cranfield Research Project*, Cranfield, England.