Data Mining II, C.A. Brebbia & N.F.F. Ebecken (Editors)
© 2000 WIT Press, www.witpress.com, ISBN 1-85312-821-X

Attractability: an indicator for optimising the design of a web site

A. Platis[1], P. Tselios[2], G. Vouros[1]
[1]*Department of Mathematics*
[2]*Department of Information and Communication Systems*
University of the Aegean, Greece

Abstract

This paper, in a greater extend than previous work, proposes a method for identifying website pages that may be of high interest to website visitors but are not, somehow, directly accessible by them, and for providing navigational support. Towards this account we model website visitors activity in a site using a Markov chain and provide a method for measuring pages' *transitions attractabilities* on a time interval of site usage. In this way, we have an indicator for page usage depending on the information interest of the page, the probability of a user to follow a link to that page and the probability of a user to be in that page. The paper provides examples from a test case and discusses the utilization of attractabilities for re-organising a web site and providing navigational support.

1 Introduction

The designers of a website try to achieve their information-communication intentions by structuring the site in a way that is comprehensive and more or less intuitive (according to designers' judgement) to the average visitor. But it is not always possible to predict visitors' navigational behaviour within a site. Different visitors have different needs and goals. Even the same visitor has different needs at different times.

Towards making the websites more effective to their visitors, designers have to identify "problematic cases" within their site. Such cases comprise information that can be of website visitors' interest but, somehow, it is not accessible by them. Then, designers have to re-organise their site in order to make "hidden" information more accessible and provide navigational support to the website visitors.

Data Mining II, C.A. Brebbia & N.F.F. Ebecken (Editors)
© 2000 WIT Press, www.witpress.com, ISBN 1-85312-821-X

On the other hand, website visitors must be assisted with their navigation through the site so as to access information that is probably of high importance to them. However, it is very difficult to organise a site with a large number of pages in such a way that every visitor could easily access information.

As already stated in [11], towards helping designers of existing web sites to make their sites more accessible, we have to capture their intentions in a simple but effective way, provide them with tools that monitor visitors' behaviour, help them to identify problematic cases and monitor the progress achieved towards making a "better" site organisation. The organisation of a site is made "better" if information, which is of interest to the visitors and it is the intent of the site's designer to be made interesting (or the designer thinks that visitors should be interested to that) is more accessible.

On the other hand, intentions of website designers and visitors' behaviour within a site, across many visiting sessions and different types of users, can be exploited to provide dynamic navigational support to future website visitors. For instance, visitors, being in a page, need to know what are the links that are highly suggested to be followed from that page.

It must be pointed that this is an optimisation problem since it does not take into account the navigational behaviour of a particular site visitor, but the behaviour of all visitors across many visiting sessions.

Our aim is to create a mechanism for the effective optimisation of existing web sites. Towards this aim [11] proposes a method for re-organising websites by calculating the probability of visitors to be in a page (or in a set of pages) and page's attractability. *Attractability* of a page is a measure that is based on the probability of visitors to be in that page and the *information interest* that the site designer assigns to this page. Based on this method, we have developed *DesignersAdvisor* that utilises sites' log files in order to provide evidence for the behaviour of web sites visitors and compute the attractability of pages based on designers' intentions. These indicators and an extension of *DesignersAdvisor* are described in the sections that follow.

This paper extends our previous work towards measuring probabilities and attractabilities related to transitions made between pages. It proposes a new indicator for measuring the attractability of following a link to a page P, i.e. the *transition attractability* to P. This indicator is calculated by taking into account the *probability to request P* from any page of the site that provides a link to P, the *attractabilities of the pages* linked to P, and the *information interest* assigned to P by the designer of the site. Based on this indicator, the paper proposes a method for identifying website pages that may be of high interest to website visitors but are not, somehow, directly accessible by them. The goal is twofold: To help designers towards making their sites more effective to their visitors, and to provide dynamic navigational support to site visitors.

The paper is structured as follows: Section 2 sketches previous work towards re-organising web sites' structure and providing navigational support to website visitors. It identifies the key issues of our research. Section 3 presents the proposed framework, and describes the computation of pages' attractabilities and pages' transitions attractabilities. Section 4 briefly presents *DesignersAdvisor* II

that implements the framework described in section 3. The section proceeds to demonstrate the application of the proposed method to an existing web site. Finally, section 5 concludes the paper and presents plans for future work.

2 Related Work - Motivation

As far as we know, the only system that assists the re-organisation of web sites is the one presented in [2] and [3]. The system consults web site log files and uses the PageGather clustering algorithm in order to recognise *conceptually related* pages and create index pages for them. Conceptually related pages are those with no direct link from one to another, but visitors' navigation paths indicate that they may be related. The aim of that work is different from the one reported here. PageGather alters web site's structure by creating new index pages. Our aim is to develop a system that will take into account web site designers' intents and visitors' transitions to pages and provide evidence for altering web site structure.

Other work, which is closely related to the aims of the work reported in this paper include the following systems:

Webwatcher [6] is a system that uses visitors' remarks and comments in order to provide navigational support to new visitors. This requires web site visitors' feedback, which is an additional effort for visitors, and in the general case, is unreliable and subjective.

Footprints [4] computes visitors' access patterns. The basic idea is to visualise the paths that visitors have already followed and provide information concerning ways to proceed within the site. The aim is to help designers and end-users to operate with the site. The system computes navigational patterns taking into account requests to pages. It does not take into account web designers' intents and it does not provide any (quantitative or qualitative) information concerning pages' information interest.

Adaptive hypermedia systems take into account visitors' different needs and capabilities in order to provide information that is more appropriate to them. Many of these systems require visitors' feedback and require pages to be developed in a special way. In general, adaptive hypermedia systems [7] provide navigational support, as well as adaptation of information content. Most of these systems intent to create hypermedia pages more or less dynamically, by taking into account user's preferences and needs, or alter the navigational structure of a hypermedia system. However, the development of such systems needs additional development effort, which in the general case of world-wide-web sites does not outweigh their benefits.

We are focusing on the development of a tool that will assist the designers of any existing web site to re-organise her/his site in an effective way. The proposed tool does not pose any requirements concerning the form or the content of web pages.

Towards this aim we are focusing on methods that:

- Do not require additional effort from the visitors to provide feedback concerning their interest on pages' or sites' organisation.
- Neither add/remove pages in the site automatically, nor alter the structure of the site by adding new pages.

Data Mining II, C.A. Brebbia & N.F.F. Ebecken (Editors)
© 2000 WIT Press, www.witpress.com, ISBN 1-85312-821-X

- Suggest web pages that have a high probability to be visited but are not found easily by sites' visitors.
- Take into account site designers' intents.
- Can be integrated easily into current web sites with the minimal effort from the webmasters.

This paper extends the work reported in [11] in two ways: First, it does provide a method for navigational support. Second, it elaborates and extends the criteria used for identifying "problematic cases" in websites by measuring not only the probability of visitors to be in a website page, but also the probability to follow a link to that page from any place in the site.

3 Modelling Visitors and Attractability of pages

As it is also stated in [11], it is possible to remark that the evolution of a visitor in a web site seems markovian. In this case, the transition from one page to another depends only on the information received from the former page. In fact, the history of the visited pages does not influence the visitors' decision to choose the future pages. This assumption can lead us to model visitors with a Markov chain.

Although this is true in the general case, such a model is not appropriate for sites where the navigational history plays a major role (e.g. in educational hypermedia). However, this is true in information-providing sites, such as university sites, commercial sites etc. where "*looking for the set of pages with the needed information*" or "*attracting visitors attention to a set of pages*" drive visitors' and designers' behaviour, respectively.

3.1 Markov chain

Let E be a set of points that represent the space of pages within a site and $X=\{X_n,\ n\in\mathbb{N}\}$ an E-valued stochastic process on a probability space, whose measure is *Pr*. X represents visitors evolution inside the space of web pages.

If we assume that the transition from one page to another depends solely on the information contained in the current page, then X is considered to be a Markov chain. That is,

$$Pr\{X_{n+1}=j|X_0, X_1, X_2, \ldots, X_n\} = Pr\{X_{n+1}=j|\ X_n\}\ a.s.,$$

where $p(i,j)=Pr\{X_{n+1}=j|X_n=i\}$ is called the transition probability from i to j.

Let α be the initial distribution on E. X is completely defined by the initial distribution function α and the transition function p. Therefore, the probability of being in page j at time n is defined as: $P_j(n)=Pr(X_n=j)=[\alpha P^n](j)$.
The steady state probability distribution is given by

$$\lim_{n\to\infty} P_n(n) = \lim_{n\to\infty} (\alpha P^n) = \alpha \lim_{n\to\infty} P^n = \Pi, \qquad (1$$

where Π is the steady state probability vector (lim P^n tends to an ergodic matrix - all the rows are identical). In this way, the probability of being in a page i

Data Mining II, C.A. Brebbia & N.F.F. Ebecken (Editors)
© 2000 WIT Press, www.witpress.com, ISBN 1-85312-821-X

asymptotically, is given by $\Pi_i = [\lim_{n\to\infty}(\alpha P^n)](i)$.

Taking into account the probability of being in a page i, one may find the pages that are very probable to be visited by any visitor. However, it is very difficult to infer the pages that are of real interest to the visitors. Towards this in [11] we introduced web site designers' intents in order to have a near-to-the-complete view of the activity that occurs in the site. Designers' intents enable the calculation of pages' attractabilities, as well as pages' transition attractabilities.

3.2 Attractability

Pages' attractability, as defined in [11], combines information on the probabilistic activity of the visitors of the site, as well as information concerning web site designers' intents. The latter are defined by a relative *information-interest factor* assigned to each page, which will be explained in (4.2). Attractability of pages has been derived from the performability measure, which is used in reliability and performance evaluation ([8], [9], [10]).

In terms of requests made by web sites' visitors, a navigation scheme can be drawn with states corresponding to pages accessed by visitors. Here, the interest is the evaluation of the "performance" of each page and thereby, for the *combined* measurement of visitors' being in that page and the importance of the page's information. For this purpose, we define the *attractability of a page* i by:

$$attract_i = g(i) \sum_{0\le n\le N-1} \mathbf{Pr}(X_n = i), \qquad (2$$

where g(i) is the information interest of state i and $\sum_{0\le n\le N-1} \mathbf{Pr}(X_n = i)$ represents the total time spent in state i on the time interval [0, N].

Concerning information interest of pages, we have to bear in mind that it is essential to take the designer's point of view and not to let the system evolve chaotically leaded only by the visitors' requests. Therefore, the proposed method towards optimising the structure of a site and providing navigational support to website visitors uses a combination of both, designers' intentions in terms of page's information interest and data coming from visitors' navigation in the site.

Pages' attractability, as it is demonstrated in [11] provide an important indicator towards re-organising websites.

However, since the goal here is not only to make information accessible, but also to provide navigational support to website visitors, there is a great need to calculate the probability of users to request a page j, through a particular link, in other words, the frequency of accessing a page j from a page i. This probability, combined with a reward for making a transition to that page, is given by a new indicator called "*transition attractability*" from page i to page j, at time n.

$$T attract_{i,j}(n) = \mathbf{E}[g(i,j)\mathbf{1}_{\{X_{n-1}=i, X_n=j\}}] = g(i,j)\ \Pr\{X_{n-1} = i, X_n = j\},$$

where **E** is the expected value and $\mathbf{1}_{\{.\}}$ is the indicator function. It is possible to remark that this indicator is also a specific case of the general attractability

Data Mining II, C.A. Brebbia & N.F.F. Ebecken (Editors)
© 2000 WIT Press, www.witpress.com, ISBN 1-85312-821-X

indicator defined in [11].

We define the reward for the transition in this case as the information interest of the destination page, that is $g(i,j) = r_j$, and asymptotically, the resulting expression becomes: $Tattract_{i,j} = p(i,j) \cdot \Pi_i \cdot r_j$, where i is the origin page, j is the destination page, $p(i,j)$ is the transition probability from page i to page j, Π_i is the steady state probability of a user to be in page i and r_j is the information interest of page j. Therefore, the total attactability of transitions to page j from any page in the site is:

$$Tattract_j = \sum_{i \in E, i \neq j} p(i,j) \cdot \Pi_i \cdot r_j \tag{3}$$

4 *DESIGNERSADVISOR* II

DesignersAdvisor II, at this stage of implementation comprises four modules (figure 1): The *Log Analyzer*, the *PathCalculator*, the *SolutionFinder* and the *Navigator*, which is the new component added to *DesignersAdvisor* towards providing navigational support. However, as it will be demonstrated, in order DesignersAdvisor II to provide advice towards website re-organisation, it utilizes transitions attractabilities rather than pages' attractabilities.

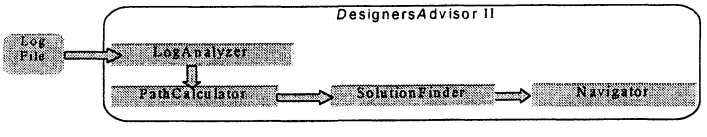


Figure 1: The framework of *DesignersAdvisor*

The input to the *DesignersAdvisor* II is the site's log file. *LogAnalyzer* is responsible for the representation of the Log File in a way that is exploitable by the system. We use this module to pre-process the log file and clean it from useless information, such as requests to image files or error entries. Analyzer's output includes the name of each requested page, the IP number of the visitor, the date and the time of the request.

PathCalculator, reads Analyzer's output and creates an (n+1)X(n+1) matrix, where n is the number of pages in the site. Each column and row from 1 to n in this matrix corresponds to a page, while row and column 0 correspond to a virtual page that represents visitor being out of the site. Each entry (i, j), $0 \leq i \leq n$, $0 \leq j \leq n$, takes an integer value that represents the number of transitions from page i to page j. For example, if there is a log entry indicating that a visitor has entered the site from the index page (page no 1), adds one to the element (0,1). If there is a log entry indicating that the visitor has left the site, the system increases the value of the entry (0, i) by one, where "i", was the last visited page. The previous procedure was repeated for all entries in the log file.

The *SolutionFinder* uses the resulting matrix to calculate the steady-state probabilities of visitors to be in each page of the site using eqn (1).

The *Navigator* uses the probabilities calculated by the SolutionFinder in

Data Mining II, C.A. Brebbia & N.F.F. Ebecken (Editors)
© 2000 WIT Press, www.witpress.com, ISBN 1-85312-821-X

order to calculate the transition attractabilities using eqn (3). Navigator exploits attractabilities to suggest pages that need to be accessible and to suggest pages that need not be so accessible as they currently are. These calculations are performed for the site as a whole, towards providing advice for site re-organisation, as well as for the pages linked to each website page separately, in order to assist website visitors. *DesignersAdvisor* II is a Win32 application and was implemented with Microsoft Visual C++.

4.1 Test Case

All tests were performed using the log files of the ACAI '99 web site. This log file contains entries covering the period time from 6th to 26th May of 1999. After the pre-process the log file contained 5595 valid records. These records represented requests for HTML pages made by 576 visitors. We assumed that every IP address corresponds to a single visitor, although, as we noted in [11], in general this is not true. We can make this assumption since we are interested in pages' requests made by all visitors.

Generally, a session is defined to be a series of requests made by a single visitor in a period of time. We conjecture that we cannot define definitely a session as a certain period of time. In [2] the time period of a session is defined to be one day, while in [5] the time period is considered to be one hour. We assume that during a session a visitor pursues a single goal. However, a visitor may pursue more than one goal in the period of a day, while she/he may not be able to achieve her/his goal in less than a single hour. Furthermore, when there is a large time gap between two subsequent requests made by a visitor, we can assume that either she/he starts pursuing a new goal or, in a way, restarts a previous one. Based on our experiments and experiments of others [5], we assume that this time gap should be approximately, 20 minutes. However this is an issue of debate. Therefore, we define a session to be a set of requests, where two subsequent requests within the session are made within a time interval of 20 minutes.

4.2 Calculated Probabilities and Attractabilities

The site has 41 pages and the system produced a 42x42 matrix, which represents the requests made by all visitors from 6th to 26th May of 1999.

The probabilities that the system calculates using data from the log file (and normalised by the sum of the probabilities) are indicated in Figure 2. Figure 3 presents the percentage of requests made per page for the whole period of time.

From the two histograms we can see that although some pages have a small percentage of visits, the probability to be requested is relatively high. For example, two such pages are "Workshops/hatziargyriou" and "organizations". Based totally on indications constructed by the percentage of requests we can easily assume that these pages are of no interest to the visitors. However, taking into account the probabilities indicated in figure 2, we can see that it is very probable for visitors to be in these pages. But it is very difficult to infer that these pages are of real interest to the visitors.

© 2000 WIT Press, www.witpress.com, ISBN 1-85312-821-X

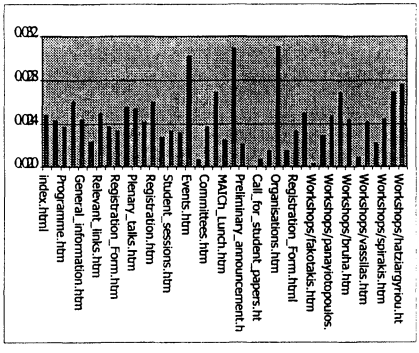


Figure 2: Page probabilities

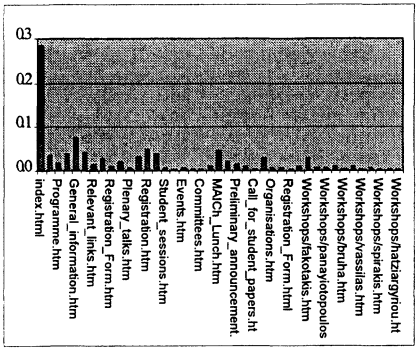


Figure 3: Page percentage requests

Towards this, as already explained in section 3.2, we use pages' "*information interest*", based on designer's intention. This factor expresses only interest on the information content for each page. The information interest factor is an integer from 1 to 100. Table 1 indicates the information interest assigned to each page. Based on the information interest factors, the system calculates the attractability of each transaction. Figure 4 indicates the results of these calculations.

Table 1. Information Interests assigned to pages

Page Name	Inf. Int	Page Name	Inf. Int	Page Name	Inf. Int
Index	1	Student_sessions	77	Registration_Form.html	50
Whats_New	61	Travel_information	45	ChaniaMap1	45
Programme	96	Events	41	Workshops/fakotakis	65
Participation	78	Local_information	52	ChaniaMap2	45
General_information	45	Committees	30	Workshops/panayiotopoulos	60
Announcements	100	Important_dates	100	Workshops/jessen	60
Relevant_links	65	MAICh_Lunch	10	Workshops/bruha	60
Workshops	100	Machine_learning_links	100	Workshops/moustakis	60
Registration_Form	70	Preliminary_announcement	57	Workshops/vassilas	60
School_schedule	100	Call_for_workshops	77	Workshops/magoulas	60
Plenary_talks	80	Call_for_student_papers	76	Workshops/spiraki	60
Accommodation	60	Call_for_participation	80	Workshops/hatziargyriou	60
Registration	100	Organisations	20	Workshops/Agents-Instructions	60
Grants.htm	77	Poster	15		

The results provided by the system may be used as a feedback to designers, in order to reconsider their judgement concerning pages' information interest, to help them indicate the "problematic cases" in their site and can be used for providing dynamic navigational support to the site's visitors. We will elaborate on these issues in the next paragraphs.

4.3 Site's Re-Organisation

A method for making information "more accessible" in web sites is to promote pages' links. While it is difficult in the general case to define such a mechanism [1], we will use the term as a synonym to making pages more accessible.

Data Mining II, C.A. Brebbia & N.F.F. Ebecken (Editors)
© 2000 WIT Press, www.witpress.com, ISBN 1-85312-821-X

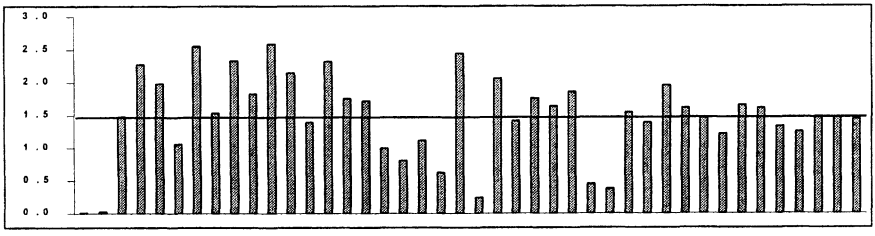


Figure 4: Transitions attractabilities

In order to create an optimisation mechanism, we have to utilise results concerning percentage of requests, probability of visitors to be in any website page, and pages' transitions attractabilities. We define pages to be *accessible wrt their information interest* when the

*(percentage of requests * transition attractability)/(probability to be in page)*

ratio of each page is equal to the corresponding ratio of all other pages in the site.

Pages that needed to be promoted should have a high probability, a low percentage of visits and high transition attractability. On the other hand, pages with a low probability, a high percentage of visits and low transition attractability, should be made less accessible. Otherwise, in the worst case, all links to pages will be included in a single web page.

The mean value of probabilities, the mean value of visits' percent and the mean value of pages' transitions attractabilities form the boundaries for judging whether the values of the corresponding indicators are "high" and "low".

In table 2 we can see the results of applying this method in our test case. From this table we can see that there are 6 pages that need to be promoted and two pages that should be denoted. The sets of pages indicated are subsets of the corresponding sets indicated in [11]. This is expectable, since the method proposed here utilise not only pages' attractabilities (which are used for the calculation of pages' transition attractabilites) but also transition attractabilities.

Table 2. Pages to be made more (less) accessible

Pages to be made more accessible		Pages to be made less accessible
Important_dates	Workshops/hatziargyriou	Workshops
Organisations	Workshops/Agents-Instructions	Registration_Form
Workshops/jessen	General_information	

The final step to be performed is to decide how and where to place links to the pages of the table 2. Since we aim to apply the method to any web site and we want the designer to have full control in his web site, we do not apply an automated procedure or site's updating. Hence, designers may choose any method they think suitable for promoting the indicated pages. However, the automated re-organisation of the navigational structure of a web site is an issue of further investigation.

Data Mining II, C.A. Brebbia & N.F.F. Ebecken (Editors)
© 2000 WIT Press, www.witpress.com, ISBN 1-85312-821-X

4.4 Navigational Support

In order to provide navigational support to website visitors, the above mentioned method for deciding on the pages that need to be more/less accessible is applied to the set of pages linked to each website page.

The application of this method to the set of pages linked from a page P results to three subsets of links: (a) links from P to pages, which the method suggests they should be made more accessible, constitute the *"Highly suggested"* group of links to be followed. (b) Links from P to pages that the method suggests they should be made less accessible, constitute the *"Weakly suggested"* group of links to be followed. (c) The rest of the links form the "*Moderately suggested*" group.

In table 3 we can see the results of this mechanism in two pages. We have to note that the number of "moderately suggested" pages is quite large, compared to the number of pages in the site. This is another indicator of a faulty design.

Table 3. The results of Promotion / Demotion

Highly suggested	Weekly Suggested	Moderated suggested
General_information, Events	Workshops, Registration_Form	Preliminary_announcement, Programme, Participation, Committees, Grants, Accommodation, Registration
Important_dates, Organisations, General_information, Workshops/hatziargyriou	Workshops, Registration_Form	Student_sessions, Committees, Whats_New, Programme, Participation, Plenary_talks, Registration, School_schedule, ChaniaMap2, Accommodation, Registration_Form.html, Announcements, Workshops/panayiotopoulos, Call_for_workshops, Call_for_student_papers

5 Conclusions – Future Work

In this paper we presented a method towards helping web designers to identify problematic cases in their sites. In this way, designers may re-organise their sites towards making information, which is of high interest and currently hidden, more accessible to future visitors. Furthermore, the paper proposed a method for providing navigational support to website visitors. Towards these goals the paper extends a previously proposed method based on the computation of pages' attractability [11]. The newly proposed method exploits pages' transitions attractabilities whose calculation is based on page's attractabilities, on the probability of visitors to request a page and on the information interest assigned to each page.

The proposed method is implemented in *DesignersAdvisor* II. So far, *DesignersAdvisor* is an interface between the log files and a spreadsheet application, which must be used to create the needed histograms. Our aim is to create an integrated tool that will be able to analyse web site's log files, let designers experiment with results provided, and give evidence regarding sites' effective re-organisation.

Furthermore, we presented an effective method for re-organising the structure of a web site towards making it optimal to its visitors with respect to designers'

intention. Although the effort needed by the designer to integrate the proposed methods in a web site seems to be an important factor, still a second thought can show us that integration could be made with less effort than the re-design of the site. We can use present navigational tools and standards in order to help users without altering the structure of the site, or of a single page. For instance, we can use layers, which will contain the groups of suggested links.

Last, but not least, another issue that is subject to investigation is the adequate representation of designers' intention.

6 References

[1] M. Perkowitz, O. Etzioni, 'Adaptive Websites: an AI Challenge' *In the proceedings of Fifteenth International Joint Conference on Artificial Intelligence*, 1997.

[2] M. Perkowitz, O. Etzioni 'Adaptive Websites: Automatically synthesizing web pages', *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, IAAA Press, (AAAI-98).

[3] M. Perkowitz, O. Etzioni 'Towards Adaptive Websites: Conceptual Framework and Case Study', *To Appear in AI Journal*

[4] *A. Wexelblat, P. Maes* , 'Footprints: Visualizing Histories for web browsing', *In proceedings of "5th RIAO Conference: Computer-Assisted Information Retrieval on the Internet"*, Montreal, 1997

[5] G. Paliouras, C. Papatheodorou, V. Karkaletsis, P. Tzitziras, C.D. Spyropoulos, 'Learning communities of the ACAI '99 website visitors', *In the proceedings of "Advanced Course on Artificial Intelligence 1999"*, 1999

[6] R. Armstrong, T. Freitag, T. Joachims T. Mitchell 'WebWatcher: A learning apprentice for the World Wide Web", In the proceedings of the 1995 AAAI Spring symposium on information gathering form heterogeneous, distributed environments, pp 6-12, 1995, AAAI Press

[7] *P. Brusilovsky* 'Methods and techniques of adaptive hypermedia', *User Modeling and user-adapted interaction*, vol 6: 87-129, 1996

[8] *J. F. Meyer* 'On evaluating performability of degradable computing systems', *IEEE, Transactions on computers*, vol: 29, pp 720-731, 1980.

[9] A. Platis, N. Limnios, M. Le Du, 'Performability of electrical power systems modeled by non-homogeneous Markov Chains', *IEEE Transactions on reliability*, vol 45(4), pp 605-610, 1996

[10] R. Pattipati Krishna, Y. Li, A. P. Blom, 'A unified framework for the performability evaluation of fault tolerance computer systems', *IEEE Transactions on computers*, vol 42(3), pp 312-326, 1993

[11] P.Tselios, A.Platis, G. Vouros, 'Providing advise to website designers towards effective websites structure re-organization', Submitted for publication.

Acknowledgements

We would like to express our sincere thanks to the Software and Knowledge Engineering Laboratory of the Institute of Informatics and Telecommunications of the NRCS Demokritos for providing to us the log files of the ACAI'99 site. It should be indicated that the log files provided had no visitors' personal information. IPs were replaced by small integers.