Using ants' behavior based simulation model AntWeb to improve website organization

Li Weigang[*a], Marcos Vinícius Pinheiro Dib[b], Wesley Martins Teles[a], Vlaudemir Morais de Andrade[b], Alba Cristina Magalhães Alves de Melo[a], Judas Tadeu Cariolano[a]

Department of Computer Science, The University of Brasilia [a]
Politec Informatic Ltd. Brazil[b]

ABSTRACT

Some web usage mining algorithms showed the potential application to find the difference among the organizations expected by visitors to the website. However, there are still no efficient method and criterion for a web administrator to measure the performance of the modification. In this paper, we developed an AntWeb, a model inspired by ants' behavior to simulate the sequence of visiting the website, in order to measure the efficient of the web structure. We implemented a web usage mining algorithm using backtrack to the intranet website of the Politec Informatic Ltd., Brazil. We defined throughput (the number of visitors to reach their target pages per time unit relates to the total number of visitors) as an index to measure the website's performance. We also used the link in a web page to represent the effect of visitors' pheromone trails. For every modification in the website organization, for example, putting a link from the expected location to the target object, the simulation reported the value of throughput as a quick answer about this modification. The experiment showed the stability of our simulation model, and a positive modification to the intranet website of the Politec.

Keywords: Ant system, AntWeb, intranet, pheromone, web mining.

1. INTRODUCTION

To attract a client to open an on-line account in a bank may be a daunting task for a bank's web developer. We realized a simple experiment to measure the website performance in this sense. 12 system analysts from the Software Factory of the Politec Informatic Ltd. were asked to visit the website of two largest private banks in Brazil (A and B). The aim of this investigation is to open an on-line account (to reach the web page that contains an application form). On average, a user took 22 seconds to finish the task on the site of the bank A and went through 3 pages to reach the application form. For the bank B, the average time to finish the task was 106 seconds. Some of them even took 291 seconds and passed through 6 pages to reach the target paper. The successful secret of the back A is that there is an attracted banner in its main homepage as shown in Figure 1.

The successful banner in a web page may orient the users to access the object page directly. This may be something as the foraging behavior of ant colonies: ants can find the shortest path between the food sources and their nest [3]. While walking, ants deposit on the ground a substance called pheromone. The pheromone trail allows the ants to find their way back to the nest (and vice versa). The following trail behavior can give rise to the emergence of the shortest paths [4]. In website organization, a suitable design of the link or banner, may equal the effect of the pheromone trail. In this paper, we try to study about that relationship.

On the other hand, the web usage mining is encouraging to discover user access patterns and improve the website organization. Some developed algorithms showed the potential applications in this area. Srikant and Yang [13] developed an algorithm to find the difference from the organizations expected by visitors to the website. They proposed the idea of backtrack, expected location and etc. Perkowitz and Etzioni [9,10] proposed the adaptive website. Some studies have been conducted on finding interesting navigation patterns using the web utilization miner [11], converting the original sequence

---

[*]weigang@cic.unb.br; phone 55 61 3072702-215; fax 55 61 2735839: http://www.cic.unb.br/~weigang/; C.P. 4466, UnB, 70919-970.

of log data into a set of maximal forward references [2], the web access pattern tree for efficient mining of access pattern [8], and others [7,12]. But there were some limitations in practical implementation. Firstly, there are still no efficient method and criterion for the web administrator to measure the performance of the website. Secondly, visitor cannot be oriented to reach the target object directly with intelligent manner on the majority of websites. The error of visitor will always take place if a web administrator does not modify the organization using logs mining results.

Fig. 1: The banner in homepage of the bank A: open your account

In this paper, we implement a web usage mining algorithm using backtrack [13] to the intranet website of the Politec. We find some expected locations and illness of the website organizations. After the modification of the links between the expected locations and the target objects, we propose a model inspired by ants' behavior to simulate the visit sequence for the website. We define throughput (the number of visitors to reach their target page per time unit relates to the total number of visitors) [15] as an index to measure the website's performance. We also use the link in a web page to represent the effect of visitors' pheromone trails [3]. The link in a web page as in Figure 1 may be used to represent the effect of the pheromone trails. For every modification in the website organization, for example putting a link, the simulation reports the value of throughput to a quick answer about this modification. The experiment shows the stability of the simulation model, and a positive modification to the intranet website of the Politec. At the moment, our model just serves as the situation of visitors with one single target page.

This paper is organized as follows. Section 2 shows the concepts of the ant system and proposed simulation model inspired by ants' behavior to the web usage mining. Section 3 describes the implementation of the web usage mining algorithm. In section 4, we report the application of the implemented system to the website of the intranet of the Politec and resume some commendations to the web administrator. Conclusions are in section 5.

2. SIMULATION MODEL INSPIRED BY ANTS' BEHAVIOR

We describe the basic concept of the ant system and introduce the ants' behavior to develop a new model to measure the website performance. Using the developed algorithm to a simple web structure shows the process of the application.

2.1 Ant system

Real ants are capable of finding the shortest path from a food source to their nest [1, 5] without using visual cues [6]. They are also capable of adapting to the changes of the environment, for example finding a new shortest path once the old one is no longer feasible due to a new obstacle. While walking from the food sources to nest and vice versa, ants deposit on the ground a substance called pheromone, forming in this way a pheromone trail. Ants can smell pheromone and, when choosing their way, they tend to choose, in probability, paths marked by the strong pheromone concentrations. The pheromone trail allows the ants to find their way back to the food source (or to the nest) [4].

Schoonderwoerd, Holland, Bruten, and Rothkrantz [14] have recently proposed an interesting application of the ant system to dynamic problems. They developed an algorithm called ABC for routing and load balancing in circuit switched telecommunications networks. Di Caro and Dorigo [15] developed AntNet; the other ant algorithm applied to the route in packet switched telecommunications networks like the Internet.

2.2 Proposal of ants' behavior-based model to measure the website performance

In case of website visiting, the situation is worse than ant system at the moment. There is no way of direct communication among visitors. Some visitors reach their object pages directly and some others do not. The web usage mining is developed to try to modify the website organization to reduce the visitor's error. But this sequence is not adaptable as ant system. In this section, we show the simulation to measure the performance of the website based on a simple website in Figure 2 [13]. Firstly, we use a simple simulation model. Secondly, we use a simulation model inspired by ants’ behavior for this website.

2.2.1 The simulation with a simple model

In this experiment, let initial time of the simulation be 0 and time interval generating a visitor be 5 seconds. Each time when generating the visitor, a random variety is also generated. The probability $P_{(1A,2A)}$ to go along the route (1A 2A) is fixed as 0.3, $P_{(1A,2B)}$ as 0 and $P_{(1A,2C)}$ as 0.7. The values of 0.3 and 0.7 are obtained from the results of the web usage mining. If the random variety of the visitor (A(i,t)) is lesser than or equal to the actual probability P(1A,2A), the visitor goes into network following the route (1A 2A), otherwise along the route (1A 2C). The time token from one page to another is supposed as 5 seconds.

From the experiment (Table 1), we are able to observe the following results. During 60 seconds, 12 visitors are generated, 4 visitors go through the network from 2A, 8 visitors go through the network from 2C, 8 visitors reach the target page (67%), and 4 visitors are still moving within the network (33%). If we define the number of visitors to reach their target page per time unit associated with the total number of visitors as throughput. The throughput in this case is 67%.

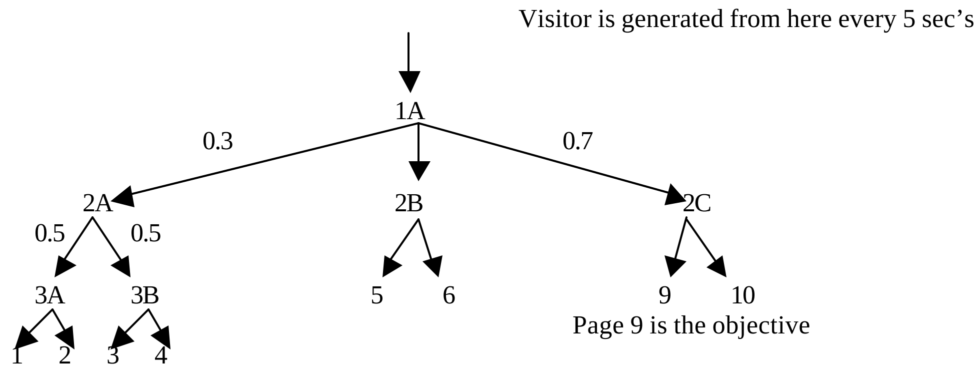


Fig. 2: The structure of a simple website

Table 1 The simulation with a simple model

Δ t = 5 sec.

Time	Path of visitor											
	5	10	15	20	25	30	35	40	45	50	55	60
t	1	2	3	4	5	6	7	8	9	10	11	12
A(1,t)	0.974	2C	-									
A(2,t)		0.182	2A	3A	2A	3B	2A	1A	2C	-		
A(3,t)			0.917	2C	-							
A(4,t)				0.175	2A	3B	2A	3A	2A	1A	2C	-
A(5,t)					0.435	2C	-					
A(6,t)						0.778	2C	-				
A(7,t)							0.571	2C	-			
A(8,t)								0.210	2A	3A	2A	3B
A(9,t)									0.791	2C	-	
A(10,t)										0.298	2A	3B
A(11,t)											0.322	2C
A(12,t)												0.655

2.2.2 The simulation with AntWeb, a model inspired by ants’ behavior

Similar to the first experiment (Figure 2), the initial time of the simulation equals to 0, time interval generating a visitor is 5 seconds. At every time to generate the visitor (in this section we call the visitor as ant) a random variety is also generated. In this paper, we just consider that all of the visitors just want to visit a page as destination (Fig. 2), where 9 is the destination page and 2A is an expected point.

In an ant system, the amount of pheromone trail $\tau_{ij}(p)$ associated with link (i,j) is intended to represent the learned desirability of the choosing page *j* when in page *i* [4]. Ants deposit a amount of pheromone proportional to the quality of the solutions they produced: the less number of links in the tour visited by an ant, the greater the amount of pheromone it deposits on the links. The initial amount of pheromone $\tau_{ij}(0)$ at iteration *p*= 0 is set to a small positive constant value or zero on all links. In our case the initial probability with which an ant(k) chooses to go from the page *i* to page $j \in N_i^d$, *d* is the destination page that gets from the web log mining. The iteration of simulation, *p*, is defined during certain simulation time. Time of simulation, *t*, is the time within an iteration at regular interval Δt. *m* ants (*k* = 1, …, *m*) are generated during an iteration. As the experiment of section 2.2.1, during iteration (60 seconds), the system generates an ant in every 5 seconds, and 12 ants in total.

The memory of each ant, the routing table, contains the already visited pages and is called tabu list. The memory is used to define, for each ant *k*, the set of pages that an ant located on the page *i* still has to visit. In the routing table, a probability value p_{ij}^d which expresses the goodness of choosing *j* as next page when the destination page is *d*, is stored for each pair (d,j) with the constraint:

$$\sum p_{ij}^d = 1,\ d \in [1, N],\ j \in N_i^d,\ N_i^d = \{N \text{ neighbors}(i)\} \quad (2.1)$$

The modification of the routing table $A_i^d = [a_{ij}(p)]_{|Ni|}$ of the page *i* is obtained by the composition of the local pheromone trail values with the local heuristic values as follows:

$$a_{ij}(p) = [\tau_{ij}(p)]^\alpha [\eta_{ij}(p)]^\beta / \sum\{[\tau_{il}(p)]^\alpha [\eta_{il}(p)]^\beta\}\ \forall\ j, l \in N_i \quad (2.2)$$

where $\tau_{ij}(p)$ is the amount of pheromone trail on the link (i, j) at iteration *p*. N_i is the set of neighbors of the page *i*, and α and β are both parameters that control the relative weight of pheromone trail and heuristic value. $\eta_{ij} = 1/wt_j$ is the heuristic value of moving from the page *i* to page *j*:

$$wt_j = lt_j + vt_j \quad (2.3)$$

where, wt_j is the estimated time at the page *j*; it includes lt_j the estimated time to get all of information of the page *j* to the browser at some velocity of the process; and vt_j is the estimated time to visit the page *j*.

The probability with which any ant chooses to go from the page *i* to page $j \in N_i^d$ while building its route at the *p*-th algorithm iteration is:

$$p_{ij}^d = a_{ij}(p) / \sum a_{il}(p),\ \forall\ l \in N_i^d \quad (2.4)$$

where $N_i^d \subseteq N_i$ is the set of pages in the neighborhood of the page *i* that the ant has not visited yet.

After an ant has completed his tour, pheromone evaporation on all links is triggered, and then the ant(k) deposits a quantity of pheromone $\Delta\tau_{ij}^k(t)$ on each link that it has visited:

$$\Delta\tau_{ij}^k(p) = \begin{cases} 1/nl^k(p) & \text{if } (i, j) \in T^k(p) \\ 0 & \text{if } (i, j) \notin T^k(p) \end{cases} \quad (2.5)$$

where $T^k(p)$ is the tour done by ant(k) at iteration *p*, and $nl^k(p)$ is the number of links in the $T^k(p)$.

$$\Delta\tau_{ij}(p) = \sum \Delta\tau_{ij}^k(p) \text{ for } k = 1, \ldots, m,\ (i, j) \in T^k(p) \quad (2.6)$$

The addition of the new pheromone by ants and pheromone evaporation is implemented by the following rule applied to all links.

$$\tau_{ij}(p) = (1-\rho)\, \tau_{ij}(p-1) + \Delta\tau_{ij}(p) \quad (2.7)$$

and ñ ∈ (0 ,1] is the pheromone trail decay coefficient. While á, â and ñ are respectively set to 1, 5 and 0.5 by Dorigo[4].

Using equation (2.1)-(2.7), we develop the AntWeb algorithm as follows.

(1) Applying the web usage mining algorithm to get the initial probability of the desirability of choosing the page *j* when in the page *i*; Iteration $p = 0$;

(2) Within an iteration, at the homepage of the site (root) *s* the system launches the ant(k) at every time *t* with a randomly selected destination page *d* (in this paper, we still limit in one destination page).

(3) Each ant(k) selects the next hop node using the information stored in the routing table. The route is selected, following a random scheme, proportionally to the goodness (probability) of each neighbor node and to the local queue status.

(4) The identifier of every visited page *i* and the time elapsed since its launching time to arrive at this *j*-th page are pushed onto a memory stack $S_{s \to d}$ (i) carried by the ant (k).

(5) When the ant (k) reaches the destination page *d*, it dies (in this paper, we assume, the visitor stops at the destination page). It deposits a quantity of pheromone $\Delta\tau_{ij}^{k}(t)$ on each link that it has visited in the tour $T^{k}(p)$ (2.5).

(6) The system updates the pheromones on the link along the tour of the ant (k). The routing table is changed by incrementing the probability p_{if}^{d} associated with the page *i* and page *f*, and decreasing (by normalization) the probabilities p_{ij}^{d} associated with the other neighbor pages *j*.

(7) $p = p+1$, return to step (2) if necessary.

Table 2 shows the form of the result of the application of the AntWeb algorithm. We put a link from expected point 2A to page 9 (Figure 2). The initial probability $P_{(1A,2A)}$ to go along the route (1A 2A) is 0.3, $P_{(1A,2B)}$ is 0 and $P_{(1A,2C)}$ is 0.7. When $t = 5$, the ant (1) is generated. Because A(1) = 0.974, it goes along the path of (1A, 2C). When arriving the page 9 (the target object), it deposits a amount of pheromone $\Delta\tau_{1A,2C}(1) = 0.5$, and then dies. During 60 seconds, 12 ants are generated. After an iteration, $\Delta\tau_{1A,2C}(1) = \sum \Delta\tau_{ij}^{k}(p) = 3$, and $\Delta\tau_{1A,2A}(1) = \sum \Delta\tau_{ij}^{k}(p) = 2$, k = 1, …, 12. From (2.3) and (2.5), we have the probability $P_{(1A,2A)} = 0.4$, $P_{(1A,2B)} = 0$ and $P_{(1A,2C)} = 0.6$. 4 ants go through the network from 2A and 8 ants go through the network from 2C. 10 ants reach the target page (83%) and 2 ants are still moving within the network (17%). The throughput in this case is 83%. So, with the new link (2A,9), it makes the web structure more efficient. Using the developed algorithm, we can measure the efficiency of the modification.

Table 2 The simulation with a model inspired by ants' behavior

$\Delta t = 5$ sec.

Time	Path of visitor											
	5	10	15	20	25	30	35	40	45	50	55	60
T	1	2	3	4	5	6	7	8	9	10	11	12
A(1,t)	0.974	2C	-									
A(2,t)		0.182	2A	-								
A(3,t)			0.917	2C	-							
A(4,t)				0.175	2A	-						
A(5,t)					0.435	2C	-					
A(6,t)						0.778	2C	-				
A(7,t)							0.571	2C	-			
A(8,t)								0.210	2A	-		
A(9,t)									0.791	2C	-	
A(10,t)										0.298	2A	-
A(11,t)											0.322	2C
A(12,t)												0.655

3. IMPLEMENTATION OF ANTWEB

In this section we describe the implementation of the proposed model – AntWeb in four parts: the database of website, determination of backtracks, main process, and interface of system.

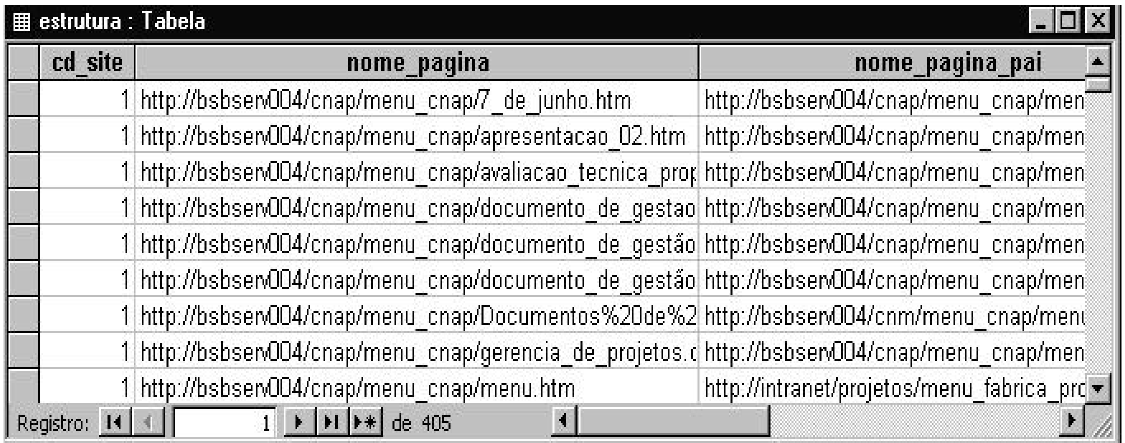
3.1 Database of website

The implementation of the database is based on the real information of the website, which consists of three main tables: the website, pages and links.

The table of the website consists of the information about the website and web mining result such as the objective information for that site (Fig. 3). For example, during the last 15 days, the visitors select the statistic results (percentage) of that page as the objective. This information is used in the simulation of AntWeb.

The table of pages contains the information of each page in the website: estimated time for download, estimated time for a visit and the relationship in the website.

The table of links is designed to represent the relationship between the pages, for example, the father and son of the actual pages. Three are some other kinds of information about parameters of pheromone and the route probability etc.


estrutura : Tabela

cd_site	nome_pagina	nome_pagina_pai
1	http://bsbserv004/cnap/menu_cnap/7_de_junho.htm	http://bsbserv004/cnap/menu_cnap/men
1	http://bsbserv004/cnap/menu_cnap/apresentacao_02.htm	http://bsbserv004/cnap/menu_cnap/men
1	http://bsbserv004/cnap/menu_cnap/avaliacao_tecnica_prop	http://bsbserv004/cnap/menu_cnap/men
1	http://bsbserv004/cnap/menu_cnap/documento_de_gestao	http://bsbserv004/cnap/menu_cnap/men
1	http://bsbserv004/cnap/menu_cnap/documento_de_gestão	http://bsbserv004/cnap/menu_cnap/men
1	http://bsbserv004/cnap/menu_cnap/documento_de_gestão	http://bsbserv004/cnap/menu_cnap/men
1	http://bsbserv004/cnap/menu_cnap/Documentos%20de%2	http://bsbserv004/cnm/menu_cnap/menu
1	http://bsbserv004/cnap/menu_cnap/gerencia_de_projetos.c	http://bsbserv004/cnap/menu_cnap/men
1	http://bsbserv004/cnap/menu_cnap/menu.htm	http://intranet/projetos/menu_fabrica_pro

Registro: 1 de 405

Fig. 3: The intranet website of the Politec

3.2 Determination of backtracks

To get the initial parameter of the AntWeb simulation, we also implemented Srikant and Yang´s algorithm which is used to automatically find pages in a website whose location is different from where visitors expect to find them [13]. The implementation of Srikant and Yang´s algorithm is divided into three main processes: to filtrate the information of log file, to construct the table of *backtracks* and to extract the information of the table of *backtracks*.

3.3 The main process

The main process is developed to implement the AntWeb algorithms. There are six principal classes and three interface classes. Figure 4 shows the diagrams of the relationship among these classes.

The class of the colonial is used to coordinate the ant activities.

The class of the control colonial is used to organize the interactions among the Colonial, Principal Window, Website and Parameters.

The class of Ants is used to generate ants. All ants will go through a website to try to find their objective page. After leaving from their homepage (start point), the ants will choice their next page based on pheromone on every link.

The class of the link is used to manage the connection between the pages and save the defined quantities of pheromone.

The class of the page is used to manage the pages in the website. There are some links to connect every page.

The class of the website is designed to save the references of all pages and the objective pages to the website.

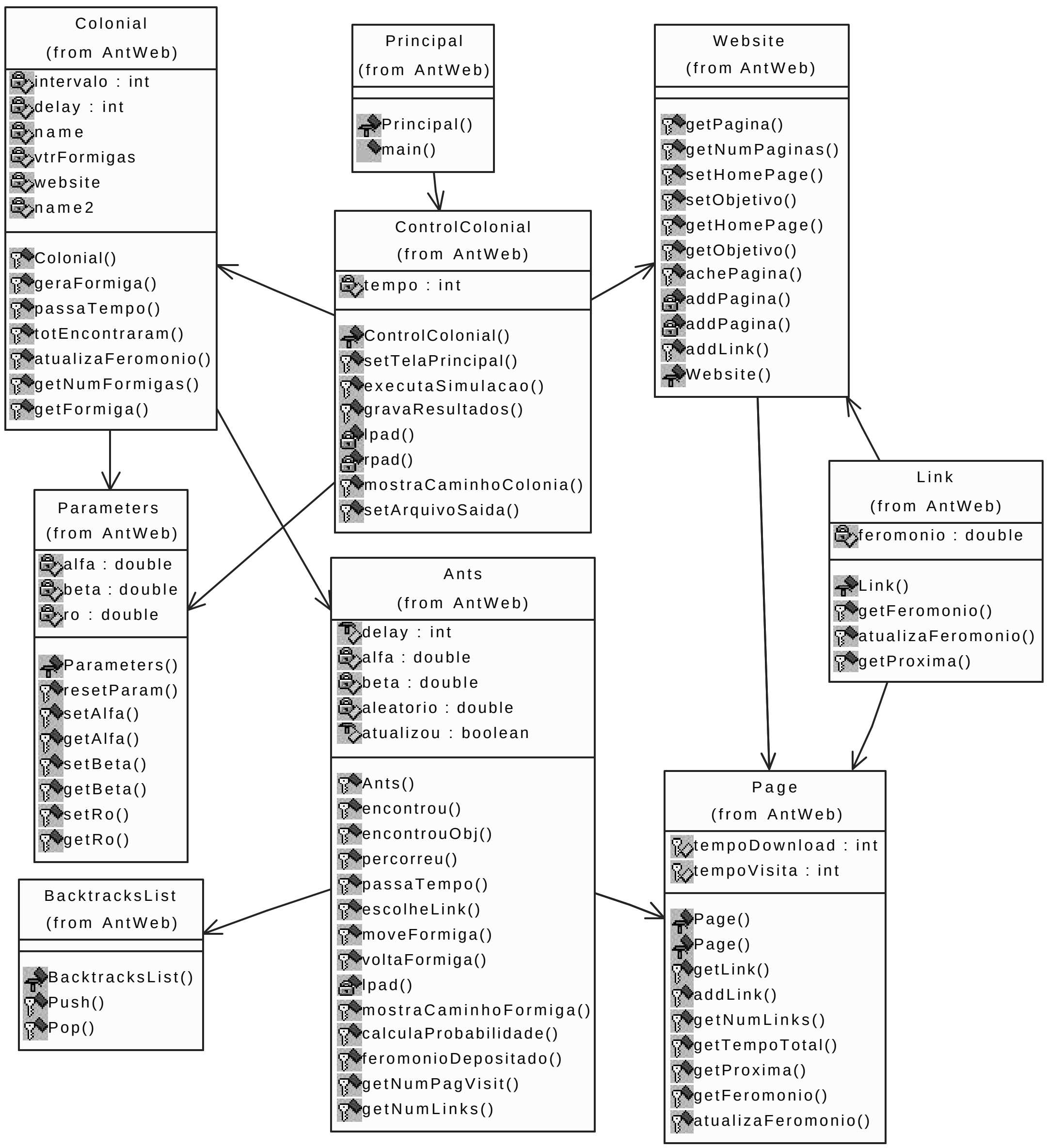


Fig. 4: The diagrams of the main process of AntWeb

3.4 Interface of system

For an easy application, an interface for user is developed in Java. This interface consists of three windows: the principal, registration and parameters windows. Figure 5 shows the AntWeb simulation in testing phase.

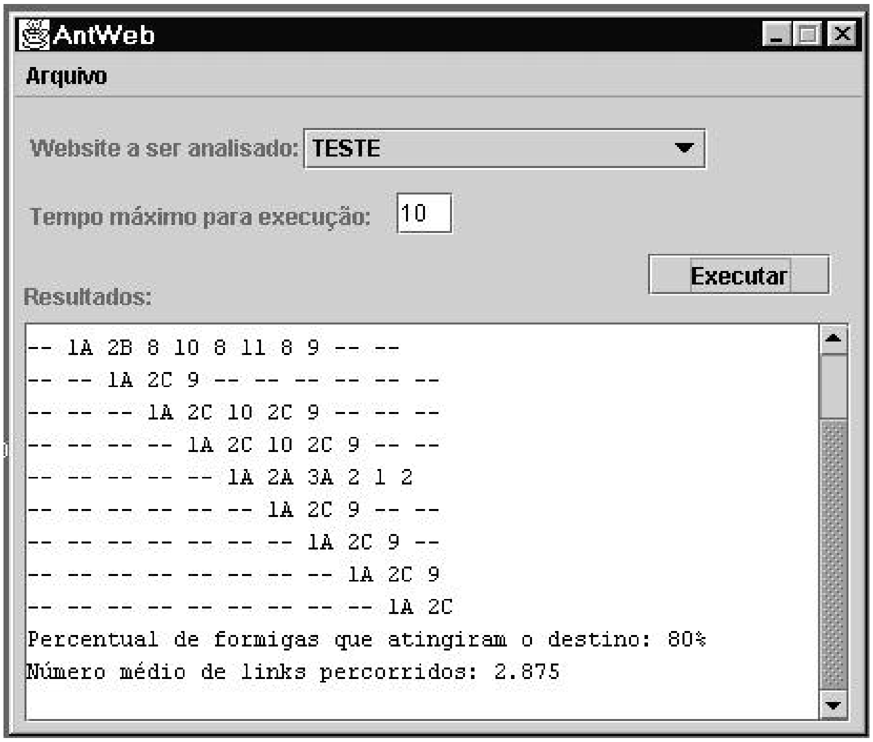


Fig. 5: the AntWeb simulation window

4. RESULTS OF THE ANTWEB SIMULATION

The Politec is a middle-scale information service company with 4200 employees. The Intranet website of the Politec is with a hierarchical structure. If considering the homepage (http://www.politec.com.br/) as the first level, the second level consists of 14 directories and 8 links to some special groups (4 shortcut links and 2 links for other mirror websites of the Politec). Under the level 3, there are 93 directories in the level 4 that expands to a great number of page (392) in the tree form. The log file is collected with 40500 visits during 15 days.

4.1 Web usage mining

Using web usage mining to the intranet website of the Politec, we get the following results:

- 17% user's accesses to the internet website are from the employees of the the Politec;
- Most of the visits is basically for updating of professional profiles and to search the information about training programs and courses;
- The website is well organized and main links of this site are also put in a suitable manner.

At the same time, we still find some issues in this site. For example, the page about workflow is organized within the subdirectory of product_service. But some visitors go first to the subdirectory of resources_ technology and then backtrack to the target page. The following is the report from our system.

Target page: http:/ /intranet/a_politec/produtos_servicos/servicos/workflow1.htm
Current location: http:/ /intranet/a_politec/produtos_servicos/servicos/
Backtrack point: http:/ /intranet/a_politec/resources_ technology/
Total visits to backtrack point: 23
Total visits: 76

There are 30% of visits for information about workflow in the directory of technological resources and the information actually is located in the directory of products and services. Consequently, the technology directory should have a link to the page about workflow.

4.2 The AntWeb simulation

The objective of the AntWeb simulation is to evaluate the modification of the structure of the website from the above implementation. At this moment we just found only one objective page: http://intranet/recursos_tecnologicos/relacao_equipamentos/gerencia_filial/quadro_geral_máquinas.htm. An ant is generated in every 5 seconds; every link reset with an initial pheromone rate of 1. The pheromone evaporation rates are configured as 1, and the parameters of both α and β are configured as 0.5 separately. Three cases are studied in the simulation experiments, which are described as following sections.

4.2.1 Case 1: One route with 5 links to the objective page

In this case, the website with the original structure, the position of objective page is located in the 5th level from homepage (Fig. 6), there is only one possible route to arrive at the objective page through 5 links. Table 3 shows the simulation results of the throughput of ants for this structure.

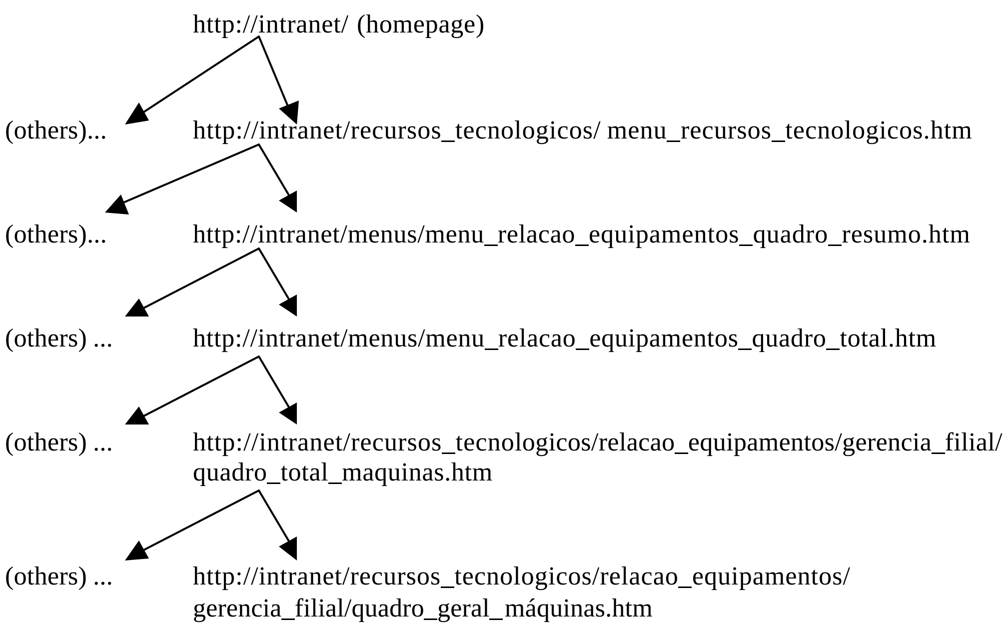


Fig. 6: One route with 5 links to the objective page (case 1)

Table 3. The simulation results from case 1: throughput of Ants

Time (sec)	**60**	**360**	**660**	**960**	**1260**	**1560**	**1860**	**2160**	**2460**
Simulation 1	0,00	34,72	68,94	76,56	83,73	85,58	87,90	89,81	91,26
Simulation 2	0,00	34,72	68,18	77,08	84,52	88,46	88,44	91,20	90,04
Simulation 3	0,00	34,72	66,67	73,96	82,94	85,58	88,44	90,28	91,06
…	…	…	…	…	…	…	…	…	…
Simulation 20	0,00	38,89	67,42	79,69	80,95	87,50	88,98	89,58	91,06
Mean	**0,00**	**39,17**	**67,95**	**76,85**	**82,60**	**86,33**	**88,24**	**89,90**	**90,90**
Standard Deviation	**0,00**	**3,79**	**2,47**	**1,55**	**1,39**	**0,94**	**0,84**	**0,71**	**0,49**

4.2.2 Case 2: one route with 5 links and another route with 4 links to the objective page

In this case, the website with the modified structure, the position of objective page is located in the 5th level from homepage, one route with 4 links is added to the page http://intranet/menus/menu_relacao_equipamentos _quadro_total.htm. With this modification, there are 2 routes for ants to arrive at the objective page (Fig. 7). Table 4 shows the simulation results of the throughput of ants for this structure.

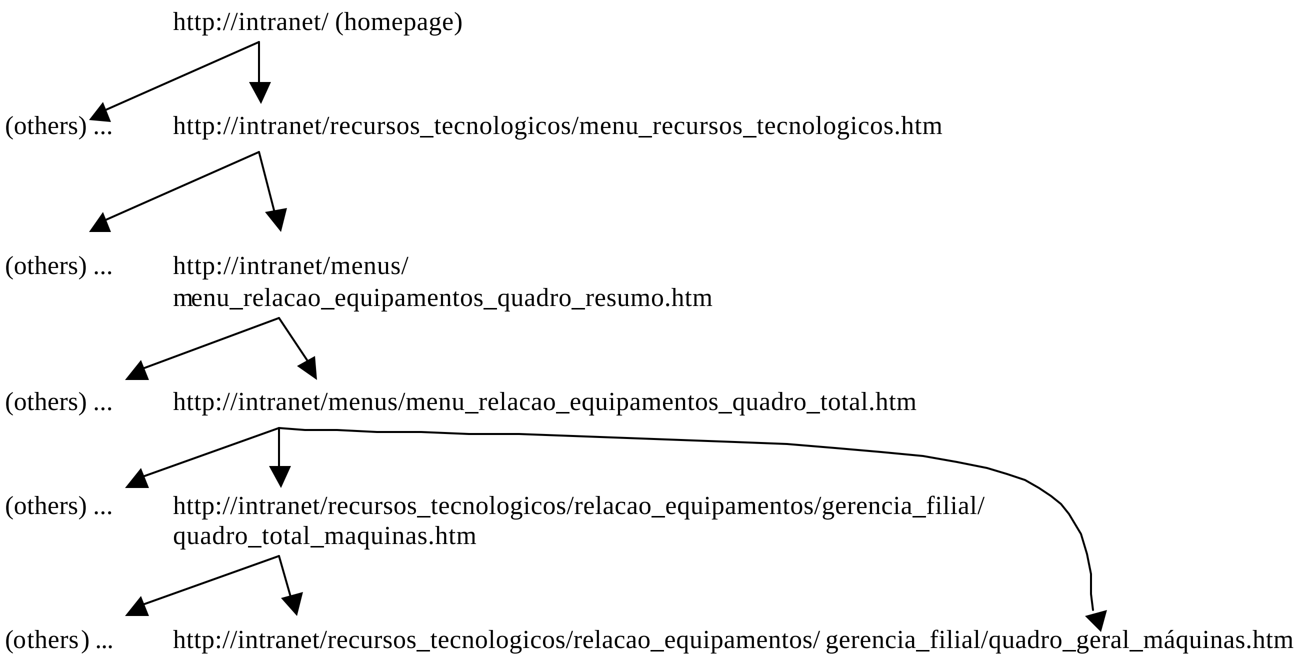


Fig. 7: One route with 5 links and another route with 4 links to the objective page (case 2)

Table 4 The simulation results from case 2: throughput of Ants

Time (sec)	**60**	**360**	**660**	**960**	**1260**	**1560**	**1860**	**2160**	**2460**
Simulation 1	0,00	45,83	71,21	76,56	82,14	86,54	88,98	89,81	92,48
Simulation 2	0,00	44,44	68,18	80,73	85,71	85,90	88,71	90,28	92,07
Simulation 3	0,00	38,89	71,21	78,65	81,75	87,18	89,52	91,20	91,26
…	...	...	...	...	...	...	...	...	...
Simulation 20	0,00	45,83	70,45	80,21	84,52	88,14	88,98	90,74	91,67
Mean	**0,00**	**44,31**	**69,47**	**79,19**	**83,51**	**87,12**	**89,03**	**90,74**	**91,69**
Standard Deviation	**0,00**	**3,55**	**2,23**	**1,50**	**1,38**	**1,10**	**0,78**	**0,89**	**0,67**

4.2.3 Case 3: one route with 5 links and another route with 2 links to the objective page

In this case, the website with the modified structure, the position of objective page is located in the 5^{th} level from homepage, one route with 2 links is added to the page *http://intranet/menus/menu_relacao_ equipamentos_quadro_resumo.htm*. With this modification, there are 2 routes for ants to arrive at the objective page (Fig. 8). Table 5 shows the simulation results of the throughput of ants for this structure.

4.2.4 The results analyses

From the analyses of the mean throughput of the three cases as Table 6, we have the following conclusions. In case 2, with the new link from the *page http://intranet/menus/ menu_relacao_equipamentos_quadro_total.htm* to the objective page, increase the efficient of ants to visit the website. Within 360 seconds simulation, the throughput is better than the original case 1, from 39.17 to 44.31 with the increasing of 13.12% (see table 6).

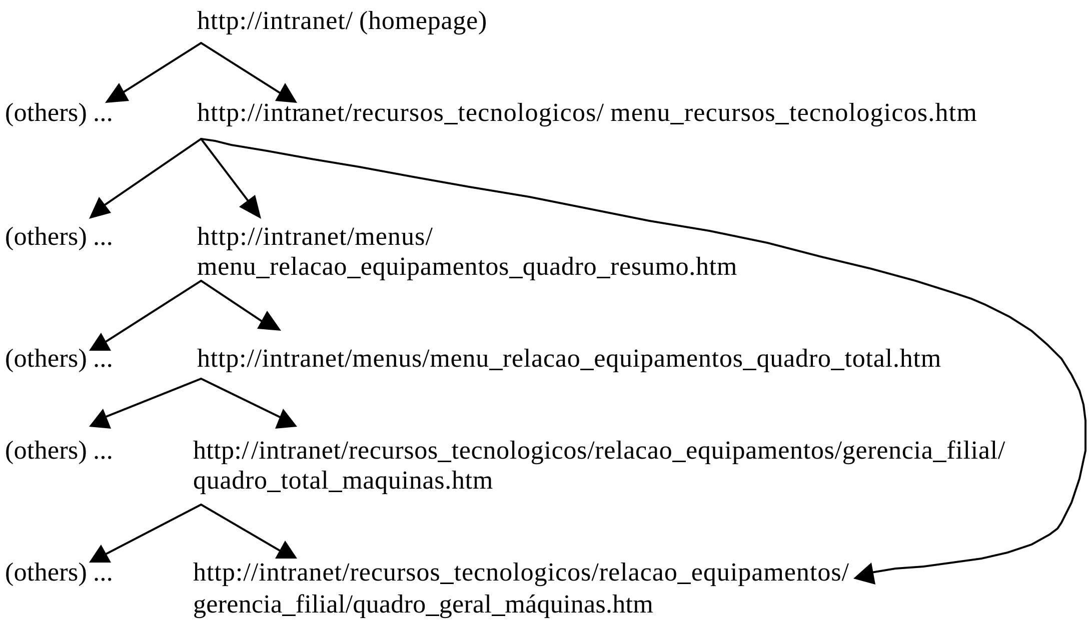


Fig. 8: One route with 5 links and another route with 2 links to the objective page (case 3)

Table 5 The simulation results from case 3: throughput of Ants

Time (sec)	**60**	**360**	**660**	**960**	**1260**	**1560**	**1860**	**2160**	**2460**
Simulation 1	0,00	44,44	69,70	77,08	86,90	88,14	88,71	91,67	92,07
Simulation 2	8,33	47,22	70,45	80,73	86,51	89,10	89,52	91,20	92,89
Simulation 3	16,67	48,61	72,73	81,77	83,73	88,46	91,67	91,90	91,06
…	...	...	...	...	...	...	...	...	...
Simulation 20	8,33	41,67	75,76	79,69	86,51	89,10	89,52	91,90	92,07
Mean	**10,83**	**49,58**	**72,54**	**80,08**	**86,31**	**88,97**	**90,39**	**91,59**	**92,48**
Standard Deviation	**10,15**	**5,16**	**2,46**	**2,01**	**1,16**	**0,82**	**0,85**	**0,63**	**0,81**

In case 3, with the new link from the page *http://intranet/menus/ menu_relacao_equipamentos_quadro_resumo.htm* to the objective page, increase the efficient of ants to visit the website. Within 360 seconds simulation, the throughput is also better than the original case 1, from 39.17 to 49.58 with the increasing of 26.58% (see table 6).

Table 6 Three cases of the simulation results analyses (mean)

Time(sec)	**60**	**360**	**660**	**960**	**1260**	**1560**	**1860**	**2160**	**2460**
Simulation 1	0	39,17	67,95	76,85	82,6	86,33	88,24	89,9	90,9
Simulation 2	0	44,31	69,47	79,19	83,51	87,12	89,03	90,74	91,69
Simulation 3	10,83	49,58	72,54	80,08	86,31	88,97	90,39	91,59	92,48

5. CONCLUSIONS

We propose the AntWeb model, with the support of the ants' system theory, to describe the web organization performance. The update function of the pheromone trails to describe the sequence of the web visit process is formulated based on the ant rule: the more the links visited, the greater the amount of pheromone deposits on the link which is used to generate the tour in the website. We use throughput as an index to measure the performance of the website. The paper shows the implementation of the AntWeb to an actual intranet website of the Politec.

The AntWeb concerns with the simulation to measure the web organization and also tries to study the visitor's behavior using ant system with the dynamic and adaptive manner. In our application, the simulation from three cases shows the interesting results. In the cases of modified model, the throughput is 13.12% more comparing case 2 with case 1; and 26.58% more comparing case 3 with case 1. It is also observed that the developed model is more sensitive to measure the efficient of the web organization. In the further research, we will consider the following directions:

- the multi-destination page case;
- the quantitative description of the relationship between the banner and *pheromone* effect;
- the adaptive modification for the web organization;

ACKNOWLEDGEMENTS

This research is partially supported by CNPq, a science and technology foundation of Brazilian government under contract no. 300585/94-2 and CAPES, an advanced education foundation of Brazilian government under contract with graduate program of Department of Computer Science of University of Brasília. The authors thank Dr. Zhizhai Hu for valuable suggestion to the paper.

REFERENCES

1. R. Beckers, J. L. Deneubourg and S. Goss, "Trails and U-turns in the selection of the shortest path by the ant Lasius niger", Journal of theoretical biology, **159**, 397-415, 1992.
2. M.-S. Chen, J. S. Park, and P. S. Yu, "Data mining for path traversal patterns in a web environment", In Proc. of the 16th International Conference on Distributed Computing Systems, pages 385–392, May 1996.
3. M. Dorigo, V. Maniezzo and A. Colorni, "The Ant System: Optimization by a Colony of Cooperating Agents", IEEE Transactions on Systems, Man, and Cybernetics-Part B, **26(1),** 29-41, 1996.
4. M. Dorigo, G. Di Caro and L. M. Gambardella, "Ant Algorithms for Discrete Optimization. Artificial Life", **5(2)**, 137-172, 1999.
5. S. Goss, S. Aron, J. L. Deneubourg and J. M. Pasteels, "Self-organized shortcuts in the Argentine ant", Naturwissenschaften, **76**, 579-581, 1989.
6. B. Hölldobler and E. O. Wilson, "The ants", Springer-Verlag, Berlin, 1990.
7. T. Nakayama, H. Kato and Y. Yamane, "Discovering the gap between web site designers' expectations and users' behavior", In Proc. of the Ninth Int'l World Wide Web Conference, Amsterdam, May 2000.
8. J. Pei, J. Han, B. Mortazavi-asl and H. Zhu, "Mining access patterns efficiently from web logs", In Proc. of the 4th Pacific-Asia Conf. on Knowledge Discovery and Data Mining, pages 396–407, April 2000.
9. M. Perkowitz and O. Etzioni, "Adaptive web sites: Automatically synthesizing web pages", In Proc. of the Fifteenth National Conf. on Artificial Intelligence (AAAI), pages 727–732, 1998.
10. M. Perkowitz and O. Etzioni, "Towards adaptive sites: Conceptual framework and case study", In Proc. of the Eighth Int'l World Wide Web Conf, Toronto, Canada, May 1999.
11. M. Spiliopoulou and L. C. Faulstich. Wum, "A web utilization miner", In Proc. of EDBT Workshop WebDB98, Valencia, Spain, March 1998.
12. C. Shahabi, A. M. Zarkesh, J. Abidi, and V. Shah, "Knowledge discovery from users web-page naviagtion", In Proc. of the 7th IEEE Intl. Workshop on Research Issues in Data Engineering (RIDE), pages 20–29, 1997.
13. R. Srikant and Y. Yang, "Mining Web Logs to Improve Website Organization", In Proc. of the Tenth International World Wide Web Conference, Hong Kong, May 2001.
14. R. Schoonderwoerd, O. Holland, J. Bruten and L. Rothkrantz, "Ant-based Load Balancing in Telecommunications Networks, Adaptive Behavior", **5(2)**, 169-207.
15. G. Di Caro and M. Dorigo. "AntNet: Distributed Stigmergetic Control for Communications Networks", Journal of Artificial Intelligence Research (JAIR), 9:317-365, 1998.