Evaluating Web Sites Based on GHAP

Bindu Madhuri .Ch
Department of CSE
Gitam University
Visakhapatnam, INDIA

Padmaja.M
Department of CSE
Gitam University
Visakhapatnam, INDIA

Dr.J.A.Chandulal
Department of CSE
Gitam University
Visakhapatnam, INDIA

ABSTRACT

We proposed a novel approach for "Evaluating Web Sites Based on GHAP", we evaluate websites based on web usage mining and try to make the evaluation practically, so that users can easily find a good website from WWW and access to high quality data and good services that attract users. Hence, evaluation methods for the effectiveness of web sites are a critical issue in both practice and research. Also we conducted experiments and confirmed the effectiveness of our approach and its potential in performing high quality web site evaluation.

KEYWORDS: *Grey clustering theory, Index system, Analytical Hierarchy Process (AHP), Web Site Evaluation, Grey Evaluation. Grey Clustering.*

1. INTRODUCTION

With the abundance of information available on the World Wide Web (WWW), it has become increasingly necessary for users to find the desired information resources, the need to automatically build a user profile to assist the user in everyday browsing and searching tasks is steadily increasing, mainly for web personalization purposes.

The web server maintains the web log to store the user's information in Common Log Format (CLF). By applying preprocessing on the web log we can detect the individual users easily. We can use not only individual browsing habits of users but also the patterns which are derived from web log are used to perform web site evaluation. Evaluating web sites is a significant task for web service because evaluated web sites provide useful information for users to estimate sites' validation and popularity.

We develop a method to obtain "Evaluating Web Sites Based on GHAP", evaluate websites based on web log data and try to make the evaluation practically, so that users can easily find a good website from WWW and access to high quality data and good services that attract users. In this method, we used AHP to construct the evaluation index and grey clustering is used to evaluate the web sites.

The rest of this paper is organized as follows. First, AHP method is reviewed, Grey Clustering theory explained in detail in section 3. An illustrating example is given in Section 4. Finally, the conclusion is drawn out in Section 5.

2. AHP METHODOLOGY

The analytical hierarchy process (AHP) developed by Saaty (1980)[1] is a powerful and flexible multi-criteria decision making tool for complex problems where both qualitative and quantitative aspects need to be considered. The AHP is aimed at integrating different measures into a single overall score for ranking decision alternatives. Its main characteristic is that it is based on pair wise comparison judgments. It is a decision-rule model that relaxes the measurement of related factors to subjective managerial inputs on multiple criteria. By reducing complex decisions to a series of simple comparisons and rankings, then synthesizing the results, the AHP not only helps the analysts to arrive at the best decision, but also provides a clear rationale for the choices made. The use of AHP does not involve cumbersome mathematics. AHP involves the principles of decomposition, pair wise comparisons, and priority vector generation and synthesis.

According to the principles of the analytic hierarchy process, this paper construct evaluation index system of website that include the target layer, attribute layer, indicators layer. Target layer is the highest level of evaluation index system; under target layer many evaluation indexes will be formed first grade indexes as to different attributes; indicators under each first grade indexes are divided into second indexes, second indexes are the lowest level of evaluation index system .A multi-level analytical structure as shown in Fig. 1.

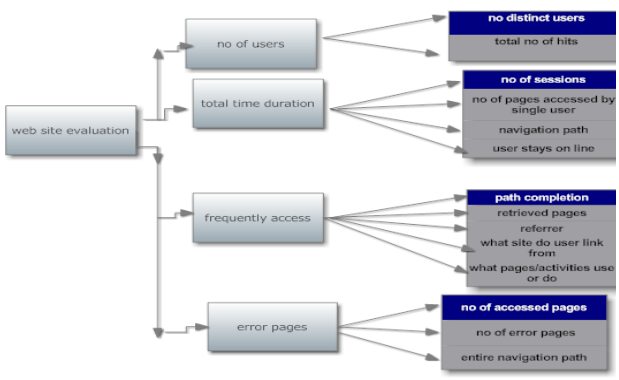

Fig.1. Multi-level analytical structure

Analytical Hierarchy Process(AHP) consists of following steps:

1) Define decision criteria in the form of a hierarchy of objectives.

AHP uses a standardized comparison scale for relative importance shown in Table 1 to make the pair-wise comparison, the comparisons done in any two indexes of commercial indexes by using "Proportion Criteria 1---9.

Table 1: Scale for Relative Importance

Intensity of importance	Definition	Intensity of importance	Definition
1	Equal importance	2	weak
3	Moderate importance	4	Moderate plus
5	Strong importance	6	Strong plus
7	Very strong importance	8	Very, very strong
9	Extreme importance		

If they are equally important then to take 1;If the former is slightly important than the latter then the former taking 3 and the latter taking 1 / 3; If the former is strongly important than the latter then the former taking 5 and the latter taking 1 / 5;If the former is very strongly important than the latter then the former taking 7 and the latter taking 1 / 7;If the former is extremely important than the latter then the former taking 9 and the latter taking 1 / 9; Between their values are for 2, 4, 6, 8.

2) Development of judgment matrices A by pair wise comparisons:

$$A=\{a_{ij}\}=\begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots\dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots. & a_{2n} \\ a_{31} & a_{32} & a_{33} & \cdots & a_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix} \quad (1)$$

3) After a judgment matrix, a priority vector to weight the elements of the matrix is calculated.

$$W_i = \sqrt[n]{\prod_{j=1}^{n} a_{ij}} \Big/ \sum_{i=1}^{n} \sqrt[n]{\prod_{j=1}^{n} a_{ij}} \quad (i,j=1,2,...,n) \quad (2)$$

4) After the generation of priority vector, inconsistency in pair-wise comparison may occur due to subjective human judgment error. Therefore, it is important to check the consistency in response through a consistency index (CI) by using the following equation.

$$CI = (\lambda max - n)/(n-1) \quad (3)$$

5) Finally, the consistency ratio (CR) is calculated as the ratio of the CI and the random consistency index (RCI), which is shown in Table 2.

$$CR = CI/RCI \quad (4)$$

Table 2: Random Consistency Index

Matrix Rank	1	2	3	4	5	6	7	8
RI	0.0	0.0	0.58	0.90	1.12	1.24	1.32	1.41

When the *CR<0.10,* we think the judging matrix has satisfying consistency. Otherwise, the comparison matrices are not consistent; we should adjust the elements in the matrixes and carry out a consistency test until they are consistent.

We can calculate the weights of the hierarchical structure for web site assessment by the AHP method. The exact weights of main criteria are obtained as:

w_1 =0.4816 (No. of users accessing);

w_2 =0.2591 (Total time spent);

w_3=0.1500 (Frequently accessed pages);

w_4 =0.0766 (Error state);

All sub-criteria are compared at the second level in terms of corresponding main criteria. By using the same prioritization method, the local weights of sub-criteria are calculated as the results of Table3.

Table 3: Results from Judgment Matrices of Sub-Criteria

Level -1	Weight	Level-2	Weight
No. of users	0.4816	Distinct users	0.426
		Total no. of hits	0.574
Time duration	0.2591	Sessions	0.468
		No of pages accessed by single user	0.211
		Navigation path	0.085
		How long users stay online	0.175
Frequently accessed pages	0.1500	Navigation path	0.256
		Retrieved pages	0.270
		What pages/activities use or do	0.210
		Referrer	0.185
		What site do user link from	0.075
Error pages	0.0766	Accessed pages	0.284
		Error pages	0.432
		Navigation path	0.284

3. GREY CLUSTERING EVALUATION

1) Constituting Comment Set of Evaluation Indicator.

We make out all the comment set of evaluation indicator. The assessment standard was divided into five grades: "very satisfactory", "satisfactory", "average", "unsatisfactory", "and very unsatisfactory". If the evaluation hierarchy of one index is between the two adjacent grades, its score is of 9,7,5,3 and 1 points.

2) Confirmation of Evaluation Sample Matrix.

Under the circumstance of determining the evaluation indicator system and the evaluation indicator weight, we can give l evaluation indicators' values according to evaluation indicator B_j. Then the evaluation sample matrix is as follows:

$$D = \begin{bmatrix} d_{111} & d_{112} & d_{113} & \dots\dots & d_{11l} \\ d_{121} & d_{122} & d_{123} & \dots. & d_{12l} \\ d_{131} & d_{132} & d_{133} & \cdots & d_{13l} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ d_{mn1} & d_{mn2} & d_{mn3} & \cdots & d_{mnl} \end{bmatrix} \begin{matrix} B_1 \\ B_2 \\ B_3 \\ \vdots \\ B_m \end{matrix} \quad (5)$$

3) Determining Evaluation Grey Cluster.

First, We divided the grey cluster into five grades: ""very satisfactory", "satisfactory", "average", "unsatisfactory", "very unsatisfactory", e =1, 2 ,3, 4, 5.

The corresponding grey cluster and the first grey cluster are as follows: For the first classification" very satisfactory ", the grey number is $\otimes_1 \in [0,9,\infty)$; for the second classification" satisfactory ",the grey number is $\otimes_2 \in [0, 7,10)$; for the third classification " average ", the grey number is $\otimes_3 \in [0,5,8)$; for the fourth classification " unsatisfactory ", the grey number is $\otimes_4 \in [0,3,6)$ and for the fifth classification "very unsatisfactory ", the grey number is $\otimes_5 \in [0,1,3)$. Their explicit weight functions are:

$$f_1(x) = \begin{cases} x/9, & 0 < x \le 9 \\ 1, & x > 9 \\ 0 & else \end{cases} \quad (6)$$

$$f_2(x) = \begin{cases} 1, & 0 < x \le 7 \\ \frac{10-x}{3}, & 7 < x \le 10 \\ 0 & else \end{cases} \quad (7)$$

$$f_3(x) = \begin{cases} 1, & 0 < x \le 5 \\ \frac{8-x}{3}, & 5 < x \le 8 \\ 0 & else \end{cases} \quad (8)$$

$$f_4(x) = \begin{cases} 1, & 0 < x \le 3 \\ \frac{6-x}{3}, & 3 < x \le 6 \\ 0 & else \end{cases} \quad (9)$$

$$f_5(x) = \begin{cases} 1, & 0 < x \le 1 \\ \frac{3-x}{2}, & 1 < x \le 3 \\ 0 & else \end{cases} \quad (10)$$

4) Calculating the Grey Evaluation Weight.

To one of the evaluation indicator *B*, candidate which belongs to the l[th] (*l* = 1, 2, 3, 4, 5) evaluation grey cluster has the grey assessment coefficient:

$$X_{ije} = \sum_{k=1}^{l} f_e(d_{ijk}) \qquad (11)$$

Then to the evaluation indicator *B*, candidate which belongs to the entire evaluation grey cluster has the total quantity of grey evaluation:

$$X_{ij} = \sum_{e=1}^{l}(X_{ije}) \qquad (12)$$

The grey evaluation weight of the eth evaluation grey cluster:

$$r_{ije=}X_{ije}/X_{ij} \qquad (13)$$

Therefore the indicator *B* which belongs to the grey evaluation weight vectors r_{ij}= (r_{ij1},r_{ij2} ,$r_{ij3,}$ r_{ij4}, r_{ij5}), A_i for all evaluation grey clusters has the grey evaluation weight matrix:

$$R_i=\begin{bmatrix} r_{i1} \\ r_{i2} \\ r_{i3} \\ \vdots \\ r_{ij} \end{bmatrix}\begin{bmatrix} r_{i11} & r_{i12} & r_{i13} & \cdots & r_{i15} \\ r_{i21} & r_{i22} & r_{i23} & \cdots & r_{i25} \\ r_{i31} & r_{i32} & r_{i33} & \cdots & r_{i35} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{in1} & r_{in2} & r_{in3} & \cdots & r_{in5} \end{bmatrix} \qquad (14)$$

5) Calculating the Total Assessment Value.

First, the grey assessment weight vector of every grey classification was is *Bi* :

$$B_i = W_i * R_i = (B_{i1}, B_{i2}, B_{i3}, B_{i4}, B_{i5}) \qquad (15)$$

Then we obtained the grey assessment weight vector *B* of this websites by integrating every assessment grey classification.

$$B = W * R = (b_1, b_2, b_3, b_4, b_5) \qquad (16)$$

According to the maximum principle, we can determine the grey grades of the websites. But sometimes judgments will be distorted because of losing too much information. At this time, we can deal with *B* further, make it single-value.

$$S = B * C^T \qquad (17)$$

4. APPLICATION EXAMPLE

A. Weight Calculation

According to the AHP method, the weights of the hierarchical structure for websites health assessment are calculated as follow:

$w = (w_1, w_2, w_3, w_4)$ = (0.4816, 0.2591, 0.1500, 0.0766);

w_{1i}= (w_{11}, w_{12}) = (0.574, 0.426);

w_{2i}= $(w_{21}, w_{22,}w_{23}, w_{24,})$ = (0.468, 0.211, 0.085, 0.175);

w_{3i}= $(w_{31}, w_{32,}w_{33}, w_{34,}w_{35})$ = (0.256, 0.270, 0.210, 0.185, 0.079);

w_{4i}=$(w_{41}, w_{42,}w_{43})$ = (0.284, 0.432, 0.304);

B. Calculating the Matrix of Sample Assessment

We performed some applications by taking the grades for the evaluator indexes for finding the performance of the web sites, and then completed the grade table (Table.4).We obtained the matrix of the sample assessment according to the assessment results.

C. *Calculating the Grey Assessment Coefficient*

To the evaluating indicatorv_{11}, in this system the grey assessment coefficient belong to the e[th] grey type is X_{l1e}:

e=1,

$$f_1(7) + f_1(6) + f_1(8) + f_1(6) + f_1(6)$$

=7/9+6/9+8/9+6/9+6/9=3.66;

e=2,

$$f_2(7) + f_2(6) + f_2(8) + f_2(6) + f_2(6)$$

=1+1+2/3+1+1=4.66;

e=3,

$$f_3(7) + f_3(6) + f_3(8) + f_3(6) + f_3(6)$$

=1/3+2/3+0+2/3+2/3=2.33;

e=4,

$$f_4(7) + f_4(6) + f_4(8) + f_4(6) + f_4(6)$$

=0+0+0+0+0=0;

e=5,

$$f_5(7) + f_5(6) + f_5(8) + f_5(6) + f_5(6)$$

=0+0+0+0+0=0;

The total grey assessment coefficient is: X_{11}

=3.66+4.66+2.33+0+0=10.6;

The other grey assessment coefficients and total assessment coefficient can be calculated by this way.

D. Calculating the Gray Assessment Weight Vector and the Matrix.

e=1, r_{111}=3.66/10.6=0.345;

e=2, r_{112}*=4.66/10.6=0.439;*

e=3, r_{113} *=2.33/10.6=0.219;& e=4,* r_{114}*=0; & e=5,* r_{115}*=0;*

So, by this way,

$$R_1=\begin{bmatrix} r_{11} \\ r_{12} \end{bmatrix}=\begin{bmatrix} 0.345 & 0.439 & 0.219 & 0 & 0 \\ 0.636 & 0.363 & 0 & 0 & 0 \end{bmatrix}$$

E. Calculating Total Assessment

We obtained the grey assessment weight vector was the secondary index *B;* for every assessment grey classification.

The grey assessment weight vector of grey classification:

$$B1=w_{1i} * R_1$$

$$=[0.574 \quad 0.426] * \begin{bmatrix} 0.345 & 0.439 & 0.219 & 0 & 0 \\ 0.636 & 0.363 & 0 & 0 & 0 \end{bmatrix}$$

$$=[0.468 \quad 0.405 \quad 0.125 \quad 0 \quad 0\ ]$$

The grey assessment weight vector of grey classification:

$$B2=w_{2i} * R_2$$

$$=[0.468 \quad 0.211 \quad 0.085 \quad 0.175] * \begin{bmatrix} 0.709 & 0.290 & 0 & 0 & 0 \\ 0.325 & 0.506 & 0.168 & 0 & 0 \\ 0.088 & 0.284 & 0.284 & 0.284 & 0.056 \\ 0.212 & 0.369 & 0.319 & 0.098 & 0 \end{bmatrix}$$

$$=[0.443 \quad 0.329 \quad 0.114 \quad 0.041 \quad 0.056\ ]$$

The grey assessment weight vector of grey classification:

$$B3=w_{3i} * R_3$$

$$=[0.256 \quad 0.270 \quad 0.210 \quad 0.185 \quad 0.079\ ] * \begin{bmatrix} 0.333 & 0.454 & 0.212 & 0 & 0 \\ 0.636 & 0.363 & 0 & 0 & 0 \\ 0.257 & 0.428 & 0.313 & 0 & 0 \\ 0.133 & 0.317 & 0.317 & 0.232 & 0 \\ 0.074 & 0.280 & 0.280 & 0.280 & 0.084 \end{bmatrix}$$

$$=[0.3382 \quad 0.383 \quad 0.257 \quad 0.062 \quad 0.084\ ]$$

The grey assessment weight vector of grey classification:

$$B4=w_{4i} * R_4$$

$$=[0.284 \quad 0.432 \quad 0.304] * \begin{bmatrix} 0.458 & 0.470 & 0.070 & 0 & 0 \\ 0.636 & 0.363 & 0 & 0 & 0 \\ 0.636 & 0.363 & 0 & 0 & 0 \end{bmatrix}$$

$$=[0.597 \quad 0.399 \quad 0.019 \quad 0 \quad 0\ ]$$

The grey assessment weight vectors of grey classification are:

B1 =[0.468 0.405 0.125 0 0];

B2 =[0.443 0.329 0.114 0.041 0.056];

B3 =[0.3382 0.383 0.257 0.062 0.084];

B4 = [0.597 0.399 0.019 0 0];

Grey evaluation weight matrix R is:

$$R=\begin{bmatrix} 0.468 & 0.405 & 0.125 & 0 & 0 \\ 0.443 & 0.329 & 0.114 & 0.041 & 0.056 \\ 0.338 & 0.383 & 0.257 & 0.062 & 0.084 \\ 0.597 & 0.399 & 0.019 & 0 & 0 \end{bmatrix}$$

Total grey assessment B=W*R

$$=[\ 0.4816 \quad 0.2591 \quad 0.1500 \quad 0.0766\ ] * \begin{bmatrix} 0.468 & 0.405 & 0.125 & 0 & 0 \\ 0.443 & 0.329 & 0.114 & 0.041 & 0.056 \\ 0.338 & 0.383 & 0.257 & 0.062 & 0.084 \\ 0.597 & 0.399 & 0.019 & 0 & 0 \end{bmatrix}$$

$$=[0.434 \quad 0.366 \quad 0.128 \quad 0.019 \quad 0.026\ ]$$

Finally, grey grade S=B*C^T

$$=(0.434 \quad 0.366 \quad 0.128 \quad 0.019 \quad 0.026) * \begin{bmatrix} 9 \\ 7 \\ 5 \\ 3 \\ 1 \end{bmatrix}$$

=7.191

Obviously, 7 < 7.1981< 9 and we can conclude that by considering the web log data we done the assessment and we conclude that the web site is "satisfactory" as per the user's perspective.

5. CONCLUSION

This paper proposes "Evaluating Web Sites Based on GHAP". It combines the measures of the grey evaluation and the AHP method to evaluate web site. We build a grey hierarchy evaluation model and this method confirms the elements of the evaluation matrix using grey numbers and the explicit weight functions. Finally we conclude that by using this method we can evaluate the web sites as well as personalize the web site.

REFERENCES

[1] T. L. Saaty, The Analytic Hierarchy Process. New York: McGraw-Hill , 1980.

[2] Youmin Xi, Youlong Zhang, W.Huang, Congguo She, Zhipeng Zhang,"The quality assessment and content analysis of corporate websites in China: an empirical study", International Journal of Information Technology & Decision Making, Vol. 6, No.2 . pp. 389-405,2007

[3] Xiaohong Liu, Da Ruan, Tianrui Li, "A credit evaluation in E-business through game theory ", New Mathematics and Natural Computation, Vol. 3, No. 1, pp. 91-109,2007

[4] H.Zhu and M.C.Zhou,"Roles—based collaboration and its kernel mechanisms". IEEE Trans, on Systems,Man and Cybernetics,Part C, Vl01..36,No.4,PP.578—589,July 2006.

[5] DU Yuyue, JIANG Changjun, "On the design and temporal Petri net verification of grid commerce architecture". Chinese Journal of Electronics,Vo1.17,No.2,PP.249-251, 2008

[6] Zu Lu,Bing Han, Yan Lili, "A two layer supply-demand analysis model for evaluation tourism online service" New Mathematics and Natural Computation, Vol. 3, No. 1, PP. 123-134,2007.

[7] Xui Li, Wayne Huang, "Design a knowledge-based system to automatically assess commercial websites", International Journal of Information Technology & Decision Making ,Vol. 6 ,No. 1, PP. 43-592,007

Table:4 Assessment And Weight Vector Of Grey System

Evaluator indicator	Grade					Assessment of Grey Weight Vector				
	1	2	3	4	5	r_{11}	r_{12}	r_{13}	r_{14}	r_{15}
v_{11}	7	6	8	6	6	0.345	0.439	0.219	0	0
v_{12}	9	8	9	8	8	0.636	0.363	0	0	0
v_{21}	9	9	8	9	9	0.709	0.290	0	0	0
v_{22}	7	6	5	6	5	0.325	0.506	0.168	0	0
v_{23}	3	2	3	2	4	0.088	0.284	0.284	0.284	0.056
v_{24}	6	5	6	5	4	0.212	0.369	0.319	0.098	0
v_{31}	7	6	7	6	7	0.333	0.454	0.212	0	0
v_{32}	9	8	9	8	8	0.636	0.363	0	0	0
v_{33}	7	5	7	4	4	0.257	0.428	0.313	0	0
v_{34}	3	4	5	3	4	0.133	0.317	0.317	0.232	0
v_{35}	3	2	3	2	2	0.074	0.280	0.280	0.280	0.084
v_{41}	7	8	7	8	8	0.458	0.470	0.071	0	0
v_{42}	9	8	8	8	9	0.636	0.363	0	0	0
v_{43}	9	8	9	8	8	0.636	0.363	0	0	0