Fuzzy Based User Dependency Analysis of a Search Session with Search Engine

Shruti Kohli

Birla Institute of Technology, Mesra, Noida Centre,

Department of Computer Science, A-47 Sector -1

Noida, India

Phone No. : +91-0120-2553661

Fax : +91-0120-2554146

e-mail: kohli.shruti@gmail.com

Ela Kumar

YMCA Institute of Engineering

Sector-6, Mathura Road,Faridabad

Phone: +91-129-2242141

Fax +91-129-2242143

e-mail: ela_kumar@rediffmail.com

Fuzzy Based User Dependency Analysis of a Search Session with Search Engine

Shruti Kohli: Shruti Kohli has done her MCA, Mphil Operational Research. She is actively involved in academic and research activities. Currently she is conducting her doctoral study in the area of Knowledge Engineering. She is actively involved in academic and research activities. She has published almost 15 research papers in international / national journal and national / international conferences. Her current research interests are Information Retrieval, Search Engine, Online User Behaviour, Web spam, Web 2.0. She has presented her research work in many conferences and has written many articles online. She is a regular blogger and maintains her own blogs and website. She has almost 8 years of teaching experience of teaching MCA,BCA,MBA students. At present she is working as Lecturer in Computer Science Department, Birla Institute of Technology, Mesra (Noida Centre),India.

Ela Kumar: Ela Kumar has done B.E. (Electronics and communication) and M. Tech (Computer science) from I I T Roorkee in 1988 and 1990 respectively. She did Ph.D. from Delhi University in Natural language processing (Artificial Intelligence) in 2003. She has published almost 25 research papers in international / national journal and national / international conferences. She has authored three books on (i) Artificial Intelligence, (ii) Information systems and (iii) Multiple choice questions in computer science. She has written two advanced study course material on (i) advanced computer architecture and (ii) Data communication and networking under AICTE NANCE scheme. She has written six course material on software engineering and parallel computing , web designing for MCA fifth and sixth semester students of Indira Gandhi National Open University , New Delhi, India. She has almost 18 years of teaching experience of teaching to B. Tech and M. Tech students. At present she is working as Assistant Professor in Computer Engineering Department, YMCA Institute of Engineering, Faridabad, India.

Fuzzy Based User Dependency Analysis of a Search Session with Search Engine

Abstract

This research analyzes User behavior in performing the search through the search engine. It develops an algorithm to find out a quantitative value of "*user dependency*" on Search Engine. The term '*user dependency'* indeed means the psychological satisfaction of user with the search result during a search session. It is an indicative measure of user's trust in search engine and will prompt the user to use the same search engine in future. Hence the user dependency model will effectively track the user behaviour in the search engine use. The paper investigated factors influencing a search session and determined user dependency on Search Engine in quantitative terms. It uses a fuzzy based approach to determine the dependency and overall faith of user in Search Engine. The proposed algorithm accepts 'user rating for the search session' as input and based on the 'user satisfaction with search' provides a user dependency value. The validity of algorithm and correctness of its result is judged according to survey conducted with a sample of users. Results have been observed to be accurate and matching according to sampled user's satisfaction.

Keywords: Search Engine, User Satisfaction, User Dependency on Search Engine, User rating of Search Session., Search Algorithm,User Dependency Algorithm.

Fuzzy Based User Dependency Analysis of a Search Session with Search Engine

INTRODUCTION

Web is an excellent medium for sharing information and for delivering products and services. This platform is, to some extent, mediated by Search Engines in order to meet the needs of users seeking information. Each of the Search Engine namely Google, Yahoo, MSN use their own algorithms for accurate and fast search (Xing Bo, Lin Zhangxi,2006). Improving the accuracy of search result had always been the most important goal for Search Engine. Much research has been conducted by researchers in this area. Border (Broder A, 2002) identified that the user queries can be broadly classified as (i) Informational (ii) Navigational (iii) Transactional query. It was determined that type of query has important impact on the information required by user(Broder A, 2002). This may influence the psychological satisfaction of the user. To judge this psychological behaviour the intension of the user need to be identified. As a normal practice the click through behavior of user is being taken by Search Engine as the most prominent approach to determine user intention (Joachims T.,2005). However there can be some kind of vagueness in the behavior of user as he is clicking on search results. Research has shown that humans are more consistent at giving relative relevance statements (Rees and Schultz, 1967). By taking relative rating for satisfaction it is possible to get a true opinion of users about various aspects of the search session (Kohli Shruti, Kumar Ela,2007). In this work, a User Dependency Model has been developed and a dependency algorithm has been formulated to determine the dependency of user on a Search Engine. User dependency is a newly coined term by authors to quantize the psychological satisfaction of user with the search session. If user is more satisfied with search results in the current session, it will be more likely that he will use the same search engine in future search sessions. Search Engines are always making efforts for understanding their user's needs. Many of them have deployed feedback system to capture user opinion. User dependency algorithm developed in this paper can be used for taking factor ratings as feedback from user. This can

facilitate Search Engine in determining searches which exhibited improved user dependency and concentrating on searches which showed low user dependency.

The model developed uses the concept of Fuzzy logic to determine quantitative value of user dependency. Fuzzy approach has been used by many researchers namely Saremi (Saremi H.Q, Montazer Gholam Ali 2006) who proposed fuzzy approach to optimize the website architecture. The model extends the search capabilities of existing methods and can answer more complex search requests. Zhang used fuzzy description logic (DL) IR model to develop an enhanced model that extends the search capabilities of existing methods and can answer more complex search requests. The use of fuzzy based approach is most appealing in this application because the user behaviour is most uncertain and quantitatively giving exact numerical values to user *satisfaction* will not develop accurate dependency model. This model will give the search engine developers an insight view to modify the search algorithms so as to provide best satisfaction to the users and thus increase the popularity of their search engine. The dragging of more users to the search engine will increase revenue also because the paid advertisements will also be enhanced accordingly.

The research presented in this paper addresses the following issues arising in the preceding discussion:

1. Investigating factors that impact user satisfaction with a Search Session.

2. Development of User Dependency Model to determine user dependency on Search Engine for his information needs

3. Applying Fuzzy Theory For Determining User Dependency

4. Determining a quantitative value which represents "*user dependency*" on Search Engine for his future needs.

DEVELOPMENT OF USER DEPENDENCY MODEL

To analyze the psychological behaviour of user, it is must to understand the intension of user during search. For finding this intension of the user a survey s conducted to develop a "User Satisfaction Metrics" that included factors influencing "user satisfaction" . Observations shows that in case the user is not satisfied with current search results, he may try (i) to query using different inputs, (ii) abandon the first Search Engine and try another Search engine or (iii) leave the option of searching using Search Engines, (Kohli Shruti,Kumar Ela,2007). Table [1] lists the factors included in User Satisfaction metrics. It was observed that "Degree of Influence of factors" varies for different queries i.e Informational/ Navigational/ Transactional. E.g. If a user needs to pay bill online, then website stability is an important factor. For a transaction query this factor is important. Similarly for a navigational query, factors like technology innovations and business conduct are important so that user is taken to latest web portal (Kohli Shruti,2007).

Notation	**Factor Name**	**Description**
F1	Response Time	Reflects the temporal capability of Search Engine to retrieve fast. It may fluctuate for different type of queries /complexity of task/ geographical locality of user/ data being requested.
F2	URL correctness	Ensures that URL in search result is correct and takes user to the desired website. User tends to get frustrated when non-working URL are displayed in search results.
F3	Result Display	Reflects the "perceived ease of use" characteristics of search result. Efficient ranking of features like Google's vertical search lures user for using it.
F4	Over All Impression	This factor captures the over all opinion of user regarding the Search Engine. This opinion may be framed according to user's experience with Search Engine, Search Engine's features like ease of use, response time etc..
F5	Technologic	Evaluating the technology expertise of Search Engine in showing results

	-al Expertise	relevant to user e.g. Geographical relevance. A person querying about newspaper from India should be shown Indian news portals before listing the links news portals from all over the world.
F6	Search Result relevancy	Reflects the relevancy of the search results with respect to user's need. Relevancy of result cannot be determined by correctness of result. It is dependent on individual user and its need. Thus it may differ for different search sessions even when keyword is same.
F7	Search result justification	Reflects correctness of search results for given user needs. This factor captures user's opinion regarding justification of results displayed for a particular query.
F8	Effectiveness of Site description	Determines effectiveness of site description in helping user to decide the relevancy of the result with respect to his needs.
F9	Business culture and conduct	Reflects updated ness of Search Engine with the current business updates. E.g. if a user want to install windows operating system then sites containing latest versions should be listed first.
F10	Website stability/reliability	Reflects user's opinion about the reliability of results being shown by Search Engine. This factor plays key role for Transactional query where user intends to make a deal on the website.

Table [1]: Factors impacting User Satisfaction Metrics

The prime factors impacting various kinds of queries are:

(i) *'Informational query':* F1….F7.

(ii) *'Transactional query* factors [F1….F7], F10

(iii) '*Navigational query*': [F1…F9].

To quantize User Dependency according to these findings mathematical model for User Dependency is given by following equation:

$$U = D(S) \quad (1)$$

where U: User , S: Search Engine.

D(S): User Dependency user on Search Engine.

Since this dependency varies for different type of queries {Informational, Transactional, Navigational}. Breaking equation (1) to incorporate various components, we get following equation:

$$U = D(w_I)/D(w_N)/D(w_T) \quad (2)$$

Here w_I/ w_T/ w_N: Weight associated with Informational /Transactional/Navigational type of query. As we have already mentioned the factors affecting these individual weights, hence

$$w_I \quad \alpha \quad Fi \ (1 \leq i \leq 7) \quad (3)$$

$$w_T \quad \alpha \quad Fi \ (1 \leq i \leq 9) \quad (4)$$

$$w_N \quad \alpha \quad Fi \ (1 \leq i \leq 7, i=10) \quad (5)$$

RANKING FACTORS ACCORDING TO PRIORITY RATINGS

This section discusses the quantitative ranking of different factors. To develop these rankings a survey, (Kohli Shruti,Kumar Ela,2007) was conducted and results acquired through the survey were evaluated using SPSS software. The importance of dependency factors were found to vary as follows:

High Priority Factors: The most important factors determined after factor analysis were [F2], [F3], [F5], [F6] which corresponds to physical display of search results.

Other Factors: [F1], [F4], [F7], [F8], [F9], [F10] were the second priority factors. The value of factor [F7] was determined by taking mean of average rating of all other factors since it is used for determining over all impression of the user. These factors were further categorized on the basis of their importance as:

Middle priority factors: [F1], [F4], [F7]

Low priority factors: [F8], [F9], [F10]

On the basis of this priority ranking each of the factors were assigned weight. This weight depicts the importance of the factor in decision making. Table [2] depicts the weights assigned to high, middle and low priority factors. Weights have been normalized such that w(F1)+ w(F2)+…… + w(F10)=1.

High Priority		**Middle Priority**		**Low Priority**	
Factor	**Factor Weight**	**Factor**	**Factor Weight**	**Factor**	**Factor Weight**
F2	0.2	F4	0.10	F8	0.020
F6	0.2	F1	0.075	F9	0.015
F3	0.15	F7	0.075	F10	0.015
F5	0.15				

Table [2]: Weight assignment according to priority

APPLYING FUZZY THEORY FOR DETERMINING USER DEPENDENCY

To determine user dependency it is essential to capture user opinion by taking his ratings for various factors in "User Satisfaction Metrics". In earlier quantitative model developed by same authors, Shruti Kohli (2007), *user rating* had been taken as crisp set of three values namely 'not satisfied', 'partially satisfied', 'completely satisfied, and their associated factor weights were 0, .5, 1. However, it was observed further that 'user satisfaction' is 'continuously varying phenomenon' which is 'uncertain' at times. It is possible that user himself is not sure about his needs. Thus, satisfaction may vary from keyword to keyword and degree of satisfaction is impacted by depth of information required by the user. This cannot be completely determined and thus it is unfair to measure satisfaction in crisp set. Moreover, we can not say that at factor .9 'User' is only partially satisfied. Hence, to develop a more deep and accurate quantitative mathematical model the application of fuzzy logic would be appropriate. In this model, while calculating "User

Dependency", the user ratings are taken on a scale of [1..10] to quantize the opinion of user. Ratings [7,8,9,10] identified high,[4,5,6] identified middle and [1,2,3] identified low satisfaction of user.

FUZZY RELATIONS BETWEEN USER AND SEARCH ENGINE

Presenting mathematically, the strength of discussed relations between the user feedback through factor ratings and User Dependency for future need is not instinctively crisp. For any search session relation between them can have different strength which can be expressed linguistically as "weak", "medium" or "strong". These relations can be interpreted as different levels of dependency a user can have on Search Engine. This brings up the notion of membership value for each relation, which determines the existence intensity of each relation. The fuzzy relations can be illustrated as below:

$U \rightarrow S$ where U: User querying Search Engine

S: Search Engine

To determine the User Dependency depicted in the above relation a fuzzy inference system has been proposed in the figure below [fig1]. The user dependency can take values (High, Inter-mediate, low dependency). This has to be determined on the basis of user's rating to satisfaction factors discussed above. This proposed fuzzy inference system has been used in the User Dependency algorithm to compute dependency of User on Search Engine for his future use.

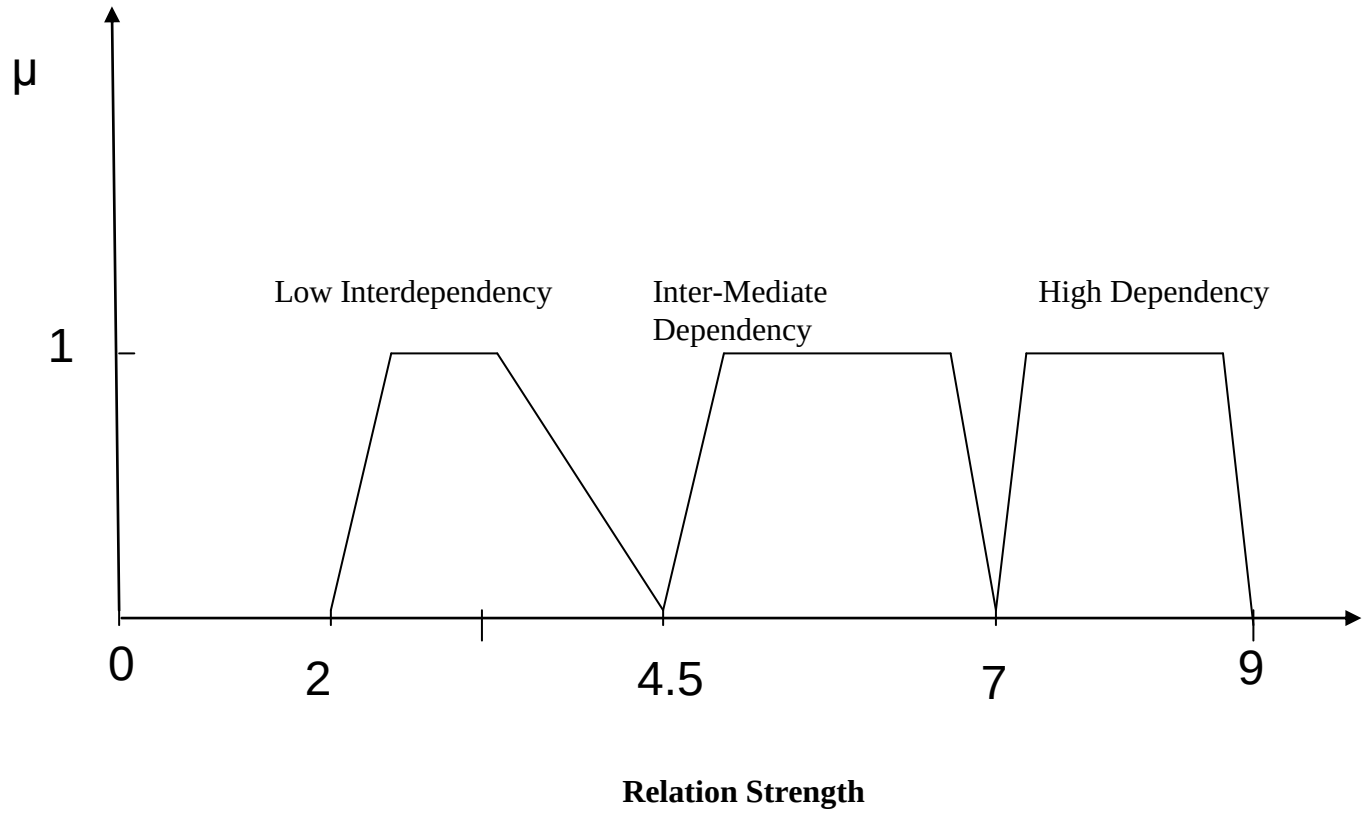


Fig 1. Fuzzy Linguistic rules for the given relation

DEVELOPING USER DEPENDENCY ALGORITHM

This section discusses the development of a "User Dependency Algorithm"(UDA). The algorithm has been developed to determine dependency of a user by calculating the influential weight of a query. User dependency has been classified into three level dependencies.(i) **Level I (L_I)**: implies 'High level' user dependency. Such dependency exhibit strong (ii) **Level II (L_{II})**: 'Inter-mediate level' user dependency (iii) **Level III (L_{III})**: 'Low level' user dependency. These are judged on the basis of ratings given by user to the high, middle and low priority factors. A user is required to rate the factors of a "User Satisfaction Metrics", Table [1], on the basis of his experience in the search session. Since F1…F7 factors are generic factors that impact all the three types of queries discussed in section 2 and these factors are listed as high or middle priority UDA applies to all type of user queries. Once factor ratings are taken they need to be segregated on basis of user rating and kept in 3 arrays A_H, A_M, A_L. Factors given high rating are put in A_H, factors provided middle ratings A_M and factors given low range ratings are kept in A_L. Depending on how these

factors are scattered in the three arrays minimum threshold weights are assigned for calculating dependency of a particular level. User Dependency value determines membership of user in a fuzzy set representing dependent users. Minimum and minimum user dependency value lie in range of {0.2,…, 0.9} where 0.2 is the minimum dependency accredited to user for any search session.

Rules Formulation For Determining "Level of User Dependency"

Terminology Used:

H_p: Factors rated as high priority ie rating [7,8,9,10]

M_p: Factors rated as high priority ie rating [4,5,6]

L_p: Factors rated as high priority ie rating [1,2,3]

A_H,A_M,A_L:Represents array for keeping H_p,M_p,L_p

w(A_H), w(A_M) :Total weight of factors in A_H and A_M

$w_{TH}(A_H)$,$w_{TH}(A_M)$:Threshold dependency weight of A_H ,A_M

Level I Dependency (L_I): *This kind of dependency reflects great satisfaction of user with Search Engine and his intention of using Search Engine for future need.*

This level of dependency is exhibited by the user when all high priority factors are in A_H or at least three lie in A_H and one in A_M. These conditions have been quantized below:

Condition I: All four high priority factors are in A_H and middle priority factors lie in A_M

Determining $w_{TH}(A_H)$: We need to find highest weight of A_H attained only when all four H_p lie in A_H.

w(A_H)=w(F2)+w(F3)+w(F5)+w(F6) =0.2+0.2+0.15+0.15 =0.70 [Weights have been assigned in Table [2]]

Hence, **$w_{TH}(A_H)=0.70$**

Determining $w_{TH}(A_M)$:Highest weight of A_M can be attained when all middle priority factors are in A_M i.e.

w(A_M) = w(F4) + w(F1)+ w(F7)= 0.1 + 0.5+ 0.5 =0.2

Hence, **$w_{TH}(A_M)=0.20$**

Condition II: It includes two conditions

(i) Three H_p factors are in A_H and one H_p is in A_M.

(ii) M_p factors scattered randomly between A_H, A_M, A_L such that only one of M_p factor lie in A_L.

Determining $w_{TH}(A_H)$: We need to find lowest possible weight when three minimum weight H_p factors and one minimum weight M_p factor lie in A_H. There are many possible combinations depicting L_I E.g. {F2, F3, F5, F1},{F2, F3, F6,F4}, {F3, F5,F6, F7}

E.g. $w_{TH}(A_H)$ =w(F3)+w(F5)+ w(F6) +w(F4) =0.2 +0.15+0.15 +0.5 =0.55

Hence, **$w_{TH}(A_H)=0.55$**

*Determining $w_{TH}(A_M)$***;** We need minimum weight required for incorporating one H_p factor and two M_p factor selected randomly in A_M. Different combinations possible are:

(i) All M_p factors lie in A_M , $w(A_M)$ =w(F4)+w(F1)+w(F7)= 0.1+0.75+0.75 =0.25

(ii) One H_p factor and two M_p factor lie in A_M e.g. $w(A_M)$ =w(F2)+w(F4)+w(F7) =0.37>0.25

(iii)One H_p factor, one M_p factor and any L_p factor lie in A_M , $w(A_M)$=w(F2)+w(F1)+w(F8)=0.33>0.25

Hence, **$w_{TH}(A_M)=0.25$**

Determining threshold weights from condition I and II

$w_{TH}(A_H)$=min(0.55,0.55)=0.55

$w_{TH}(A_M)$=min(0.25,0.20)=0.20

Rule for determining level I dependency is:

Rule: If $w(A_H)$ >= 0.55 then

If $w(A_M)$ >=0.20 then UD=0.9 else UD=0.7

User Dependency Calculation: *User dependency of 0.9 is the highest user dependency value attained for a user. It shows that user has given high ratings to at least three H_p factors. This exhibits his satisfaction. He may intent to use Search Engine frequently for his informational need. 0.7 reveals user has given high rating to some H_p factors but must have given low rating to some M_p and H_p factors. This exhibits his satisfaction with some features like result display, site description but may not be over all impressed. Still*

he may consider using Search Engine for future needs. Flow chart depicting level I dependency calculation is depicted in fig1:

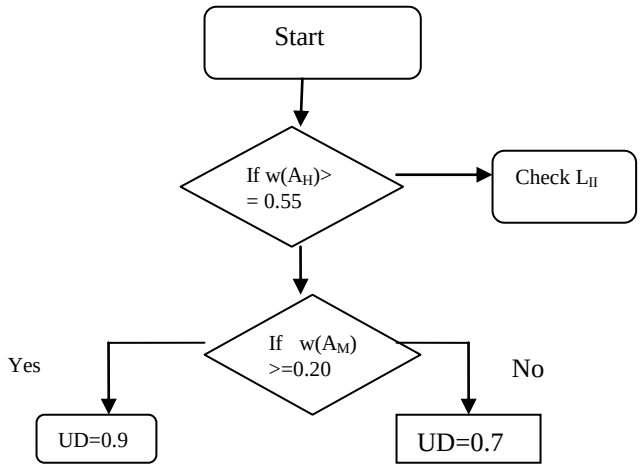


Fig 1: Flow Chart for level I dependency

Level II Dependency (L_{II}): User which does not exhibit L_I is checked for L_{II}. This is exhibited by the user when H_p factors and M_p factors are randomly distributed in A_H and A_M. This implies some H_p factors lie in A_M and some M_p lie in A_H. This reflects user's intermediate level satisfaction with Search Engine. Since H_p factors may lie in A_M or A_L such that $w(A_M)$ or $w(A_L)$ higher then $w(A_H)$. Therefore, $w_{TH}(A_H)$, is decreased and $w_{TH}(A_M)$ A_M is increased to incorporate these changes. We can say that:

$w_{TH}(A_H)$ for L_I < $w_{TH}(A_M)$ for L_I

$w_{TH}(A_M)$ for L_{II} > $w_{TH}(A_M)$ for L_{II}

L_{II} dependency is reflected in following conditions:

Condition I: At least two H_p and two M_p factors lie in A_H, two high priority factor lie in A_M

Determining $w_{TH}(A_H)$: User must have rated two H_p factors in middle or low range. It is possible that some M_p factors are in A_H Different possible combinations are {F2,F3,F4,F1},{F2,F3,F4,F7}, F6,F3,F4,F1} etc.. such that $w(A_H)$>=.50.

E.g. $w(A_H) = w(F2) + w(F3) + w(F4) + w(F1) = 0.2+0.15+0.075+0.1=0.525 > 0.50$ or

$w(A_H)=w(F2)+w(F3)+w(F4)+w(F7) =0.2+0.15+0.075+0.75=0.50.$

Hence, $\mathbf{w_{TH}(A_H)=0.50}$

Determining $w_{TH}(A_M)$: Condition I is fulfilled when at least two H_p factors lie in A_M. Other factors are immaterial they can be H_p or M_p factors. In such case different combinations of high priority are possible like {F2,F3},{F2,F6},{F2,F5} etc…such that $w(A_M)>=0.30$ e.g. $w(A_M)= w(F3)+w(F5) =0.30$

Hence, $\mathbf{w_{TH}(A_M)=0.30}$

Condition II: Two high weighted high priority factors lie A_H and other two high priority factors and one highly weighted middle priority factor lie in A_M.

Determining $w_{TH}(A_H)$: This a specific case where user has rated two highly weighted Hp factors {F2,F6} in high range. So, minimum acceptable weight is:

$w(A_H)=w(F2)+w(F6) =0.2+0.2=0.4$

Hence, $\mathbf{w_{TH}(A_H)=0.30}$

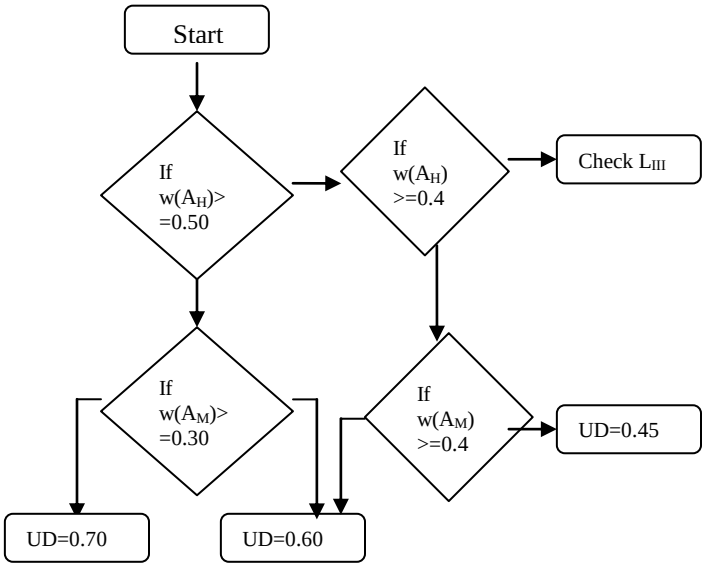
Determining $w_{TH}(A_M)$: Since two H_p factors and one highly weighted M_p factor lie in A_M. Minimum acceptable weight for A_M is $w(A_M)=w(F3)+w(F5)+w(F4) =0.15+0.15+0.1=0.4$

Hence, $\mathbf{w_{TH}(A_M)=0.40}$

Threshold weights of conditions I and II cannot be merged as will lead to acceptance of some unsatisfied cases so separate rules developed for them. Rules for weight calculation:

Rule I: If $w(A_H) >= 0.50$ then If $w(A_M) >= 0.30$ then UD=0.70 else UD =0.65

Rule II: If $w(A_H) >= 0.40$ then If $w(A_M) >= 0.40$ then UD=0.60 else UD =0.45



User Dependency Calculation: *User dependency of 0.65 reflects high rating to some of H_p factors and hence is acceptable value of User Dependency. 0.6 reflects intermediate satisfaction with H_p factors. User must have given middle level rating to H_p factors. A value of 0.45 reflects user's low ratings to some middle or high priority factor. Although some intermediate level satisfaction may be achieved low rating to others depict some kind of dissatisfaction with some features of the Search Engine.*

Level III Dependency(L_{III}): This dependency is the minimum level dependency that a user will have with Search Engine. User which do not exhibit L_{II} dependency is checked for L_{III}. User dependency decreases as user moves from one level to other. Two type of conditions may lead to L_{III} dependency.

Condition I: At least 1 H_p factor is in A_H /At least three H_p lie in A_M /At least two M_p factors lie in A_H and some H_p factors are in A_M

Determining $w_{TH}(A_H)$: A_H may contain at least one H_p factor followed by any other M_p or L_p factors or it may contain any two M_p factors. So minimum $w(A_H)$ is minimum of weights of all four H_p factors and minimum sum of any two M_p factors.

Min rating for H_p = min(F2,F3,F5,F6) = min(0.2,0.2,0.15,0.15)=0.15.

Similarly, minimum sum of any two M_p factors

=Sum(F4,F7) =Sum(0.075,0.075)=0.15

Hence, $\mathbf{w_{TH}(A_H)=0.15}$

Determining $w_{TH}(A_M)$: Since A_M should contain either at least three combinations of H_p or M_p factors the ma

ximum weight of any combination of any three H_p factors like {F2,F6,F3},{F2,F6,F5} is 0.55. if $w(A_M)$ is 0.55 it depicts some kind of satisfaction with H_p as they are rated in M_p. In case weight less then 0.55 it means high priority factors are rated in low range showing dissatisfaction with these factors hence user dependency decreased to value of 0.3.

Condition II: Minimum level of dependency

This condition occurs when high priority factors may lie in A_L. In such case a minimum user dependency is accredited to the user with an optimistic view that in spite of dissatisfaction user may try using Search Engine for his future needs. This user dependency has been determined to be 0.2.

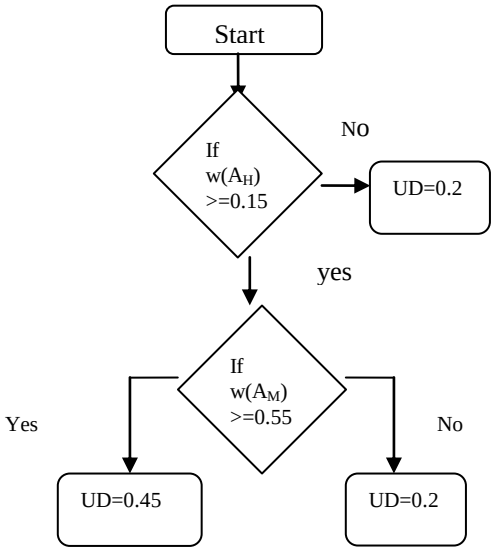
Combining condition I and II rules are:

Rule I: If $w(A_H) >= 0.15$ then

If $w(A_M) >= 0.55$ then UD=0.45 else UD=0.2

Rule II: If $w(A_H) < = 0.15$ then UD=0.2

User Dependency Calculation: *User dependency value 0.45 previously shows that some high priority factors are in middle range showing some kind of satisfaction. In case some high priority factors given low rating user dependency drops to a meager value 0.45. In case no condition is fulfilled user dependency given minimum value 0.2. This is to include the fact that the user when in major need may try using Search Engine again even though he is not satisfied in the previous session.*



USER DEPENDENCY ALGORITHM

Rules developed in calculating the three levels of dependency are used for developing 'User Dependency Algorithm'. This algorithm takes user factor ratings and type of query as input and outputs User dependency, UD value. This value determines user membership among Users dependent on Search Engine for their future needs. Algorithm has been depicted as flowchart in fig. 3.

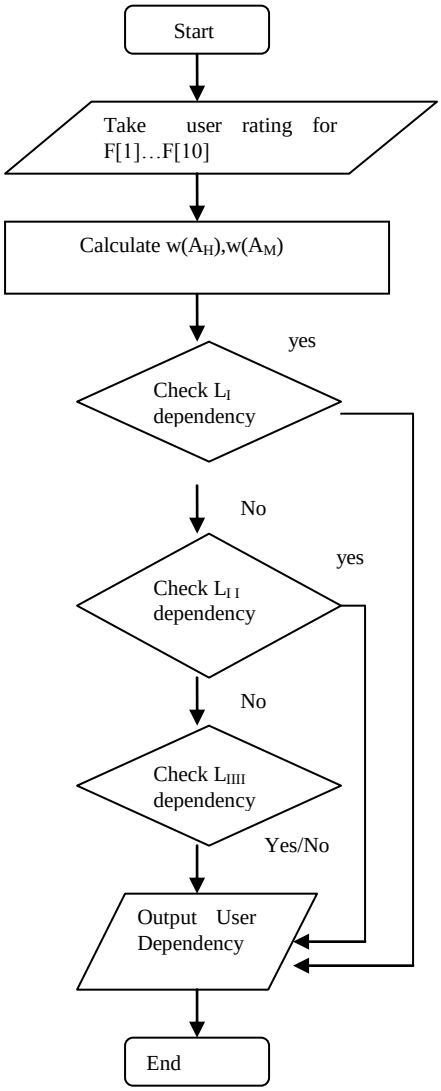


Fig 2: User Dependency Algorithm

Algorithm can be summarized as follows:

1. Take an input of factor ratings in range of [1..10] for factors F1…F10.
2. Put factors with high rating [7,8,9,10] in A_H , middle rating [4,5,6] in A_M,low rating [1,2,3] in A_L
4. Calculate $w(A_H),w(A_M),w(A_L)$
5. Check for L_I
6. If Not(L_I) Check for L_{II} and determine UD.
7. If Not(L_{II}) Check for L_{III} and determine UD.
8. Output UD value.

Illustrative Example: Suppose a user wants some information about Bhutan. He enters "bhutan" in the search box.

Applying UDA:

User Ratings:F1=10, F2=7, F3=9, F4=6,F5=4, F6=3,

F7=8, F8=2, F9=1 ,F10=5

R={F1,10},{F2,7},{F3,9},{F4,6},{F5,4},F6,3},{F7,8}, {F8,2},{F9,1},{F10,5}

$\mathbf{A_H}$={F1,F2,F3,F7},A_M={F4,F5,F10}, A_L={F6,F8,F9}

Output: UD=0.60.

Inference: User has a membership of 0.60 among users dependent on Search Engine for Informational needs. Since membership is above 0.5 it can be assumed that user may use Search Engine for his future needs. This is evident from the factor ratings provided by the users. Search Engine can use this output for determining user opinion. Search Engine can take user as some what satisfied user who may use Search Engine for future needs.

EXPERIMENT

To study the impact of the algorithm it was simulated for 100 users. Various factors were randomly given rating on scale of [1..10] and its impact on "User dependency" was determined. *'0.5' was assumed to be the threshold value for accepting user as member of "set of dependent users".* Fig. 3 depicts the result for simulated users.

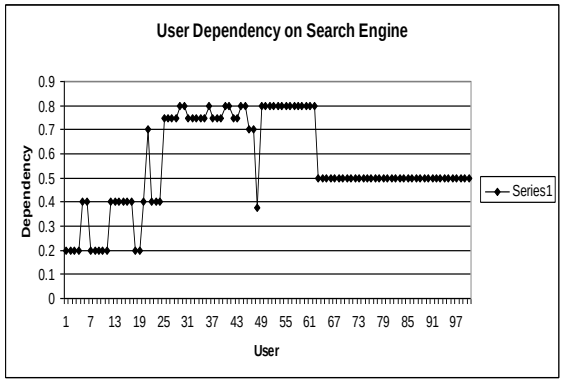


Fig 3. User dependency for 100 users

Simulation experiment was conducted on a excel sheet. This was done to explore the various ways through which high dependency of user can be obtained. It was determined that dependent users gave high or middle priority rating to high priority factors [F2,F3,F5,F6] where as users found to be non-dependent gave low priority rating to these factors. This is in accordance with our survey calculation which suggested that these factors were rated as most important factors for achieving customer satisfaction.

Applications of the User Dependency Model

Search Engines are always making efforts for understanding their user's needs. Many have deployed feedback system to capture user opinion. 'User dependency algorithm can be used for taking factor ratings as feedback from user. This algorithm can be used for following tasks:

1. Determining User Dependency on Search Engine

This mathematical model facilitates the analysis of *User Satisfaction* with Search engine and helps to determine "User Dependency" on Search Engine. This can be inferred from the fact that satisfaction will increase "User faith" in Search Engine and user is more likely to be dependent on it for future needs. Fig. 4 illustrates the usage of this user dependency model for as a feedback tool for determining User Dependency.

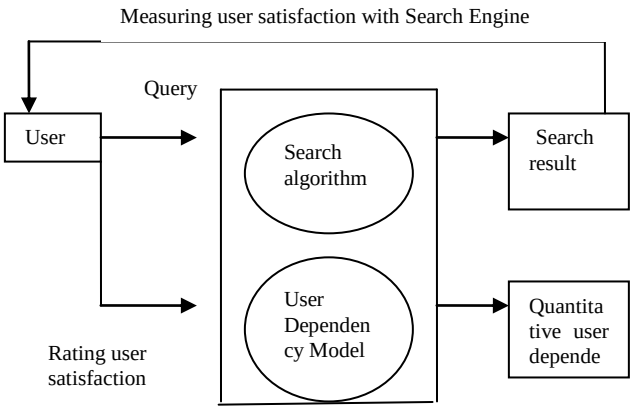


Fig 4: User Dependency Model as a Feedback Tool

This model can be implemented by Search Engine while taking user feedback. Rating various "User Satisfaction Factors" on scale of 1-10 and determining the membership of user among satisfied users will provide Search Engine a broad framework of user satisfaction. The sessions of users with low membership can be further studied through Search Engine logs to identify reason of dissatisfaction. Such exercise done periodically for few users will definitely help to improve user dependency on Search Engine.

2. Rating Search Engine Capabilities on basis of User factor Rating

Search Engine capabilities can be broadly classified as *'Temporal capability', Functional Capability, Geographical Capability, Technical Capability.* Temporal capability is reflected by Response time[F1],Functional capabilities are reflected by factors like Search result justification[F4], Result display[F5],URL correctness,[F2]. Geographical capability can be determined by factors like Technology Innovations [F8], Business culture and conduct [F9], Technical capability is reflected by factors like Search Result Relevancy[F3], Website stability/reliability [F10],Site Description,[F6]. User factor rating can be used to determine user's opinion regarding various capabilities of Search Engine. E.g. A factor rating of 3 for factor [F1] reflects user's dissatisfaction with the time taken by Search Engine in responding the query. Similarly a user rating of 9 for [F3] exhibits user's acceptance of search results as relevant documents to the input query. Search Engine can collect these ratings for different factors and can determine user's opinion regarding various capabilities of Search Engine.

3. Determining Keyword specific User Satisfaction

The User dependency model can be used for determining "keyword specific" user satisfaction with search results". E.g. User entering query "purchase laptop" may be looking buying for tips to buy laptop or needs list of shopping stores to purchase laptop. Once he rates various factors user satisfaction and his future dependency can be determined. Low rating to functional capability (F2/F4/F5) implies user did not find

desired information in search results. Similarly, low rating to Technical capabilities imply user did not find relevant information or was dissatisfied with result display.

Conclusion

This work present a new methodology to determine user satisfaction by evaluating user dependency on search Engine. *User Dependency Model* has been developed to determine the future dependency of a user on Search Engine. This dependency has been measured by applying fuzzy theory and hence provides approach to measure membership of user in set of dependent users. It is different from the other approaches being used currently by Search Engine like studying query logs or observing user clicks. This approach can be used as a feedback process where quick ratings by user can help to determine his future intentions of using Search Engine. This paper is the part of work being conducted to develop a User Satisfaction Tool.

References

- Xing Bo, Lin Zhangxi "The Impact of Search Engine Optimization on Online Advertising Market", *ACM International Conference Proceeding Series*,2006,Vol. 156.
- Broder A, "A Taxonomy Of Web Search", *SIGIR Forum*, 2002, 36(2) Pg: 3–10.
- Saremi H.Q, Montazer Gholam Ali "Web Usability: A Fuzzy Approach to the Navigation Structure Enhancement in a Website System, Case of Iranian Civil Aviation Organization Website.", *Proceedings of world academy of science, engineering and technology,* Waset, November 2006, volume (16), issn 1307-6884 123.
- Lei Zhang et al "An Enhanced Model for Searching in Semantic Portals.",*Proceedings of the 14th international conference on World Wide Web Conference*, ACM. Pages: 453 - 462
- Rose D.E. and Levinson D. "Understanding User Goals in Web Search.", *Thirteenth International World Wide Web Conference Proceedings*., WWW2004. pg: 13–19.

- Joachims T. et al ”Accurately Interpreting Clickthrough Data as Implicit Feedback.”, *Proceedings of the ACM Conference on Research and Development on Information Retrieval (SIGIR),*2005.
- Rees and D. Schultz ”A field experimental approach to the study of relevance assessments in relation to document searching.”, *NSF Report,*1967.

Kohli Shruti, Kumar Ela(2007) “Evolution of Interdependency Model Based on behavioral analysis of User, Search Engine and advertiser in Online Advertising.”, *Proceedings of the 4th International Conference on Innovations in Information Technology*, Dubai,2007, IEEE Xplore,6-18 Dec. 2008, page(s): 327-331.

- Au Irene et al. ”User experience at Google: Focus on the user and all else will follow.”, *Proceedings of Conference on Human Factors in Computing Systems*,2008, ACM,Pg:3681-3686.
- Lee U. et al,“Automatic identification of user goals in web search.”, *Proceedings of WWW 2005*,ACM Press.
- Kohli Shruti et al (2007),”Quantitative Dependency Analysis of User, Search Engine and Advertiser Behavior.”, *Proceedings of WWW/ Internet 2007*,IADIS International Conference,Portugal