

Improvement Personalization of Website Using Database and Cookies

Ekhlas Khalaf Gbashi 

Received on:30/4/2009

Accepted on:7/1/2010

Abstract

Personalization involves taking some set of inputs and returning recommendations to users through some types of computations. The method by which the personalization is performed depend on the type of data presented as input, the required representation of the output and (in some cases) the speed and accuracy by which the output must be computed.

This research tends to build the personalization in a new view, since it builds the personalization in the client in addition of the personalization in the server and using mobile agent. That by making the personalization for improvement the speed and efficiency of search in crawling, producing a technique, that technique involve recording the previous user's searches in cookies and databases then if the user try to make a search on other time the proposed system will produce to him all the previous search's results and the latest results from www. The tools used in the proposed system are SQL server 2000, HTML, ASP, Java script and two computers one presents server and the second presents the client.

Keywords: Search engine, mobile agents, personalization, web databases and cookies.

تحسين خصوصية الموقع باستخدام قواعد البيانات والمسجلات

الخلاصه

الخصوصية تتضمن اخذ بعض مجاميع المدخلات وتقوم بارجاع التوصيات الى المستخدمين عبر بعض الانواع من العمليات الاحتمالية. طريقة الخصوصية المعده تعتمد على نوع البيانات المقدمه كمدخل، والتمثيل المطلوب للمخرجات (وفي بعض الحالات) السرعه والدقه للحصول على المخرجات من خلال الحسابات المختلفه.

هذا البحث يهدف الى بناء الخصوصية برويه جديده، حيث ان الخصوصية تبنى عند الزبون بالاضافه الى الخصوصية عند الخادم واستخدام الوكيل الجوال، هذا يتم من خلال بناء الخصوصية لتحسين

السرعه و الكفاءه لعمليات البحث في برنامج النساج (crawling)، من خلال تقديم تقنية مقترحه هذه التقنيه تتضمن تسجيل نتائج بحث المستخدم (كل مستخدم) السابق في داخل قواعد البيانات والمسجلات (cookies) ثم في حالة ان المستخدم حاول ان يقوم بعملية بحث في وقت لاحق، النظام المقترح سوف يقدم له كل نتائج البحث السابق و اخر النتائج الموجوده في وقت البحث الحالي من الشبكة العنكبوتيه.

الادوات المستخدمه بالنظام المقترح هي SQL 2000 ولغة HTML ولغة JAVA SCRIPT ونظام ASP بالاضافه الى حاسبتين احدهما تمثل الزبون والاخرى الخادم.

1- Agent Philosophy

Intelligent agent is a program that maps percepts to actions. It acquires information from its environment ("perceives" the environment) and decides about its actions and performs them. While there is no real consensus about the definition of intelligent agents, the above one adapted from is intuitively clear and essentially describes the general concept of *generic agent*. Figure (1) All more specific intelligent agents can be derived from that concept.[2]

Mobile agents Are units of executing computation that can migrate between machines. The concept of mobile agents is based on remote programming (RP) for distributed systems, as opposed to remote procedure call (RPC). In the RPC case, Figure (2-a), a client program issues a procedure/function

call through a network to a remote server. In the RP case, the client program sends an agent to the remote

server, where the agent issues local calls in order to obtain the desired service. After it completes its task on the remote server, it can get back to the client machine and deliver the results to the client program. Most mobile agents today are based on Java applets' mobility and Java protocol for agent transport. There is a number shareware tools for mobile agents on the Internet, such as IBM Aglets. Such tools make possible to create, deploy, deactivate, and delete mobile agents easily, often no direct coding. However, for both technical and security reasons, deploying mobile agents in practice usually require an *agent's server* to be installed on every machine that can be visited by a mobile agent created by using a

specific tool. Agent server is the "gathering place" for mobile agents on a given machine. The agents work within that server, and migrate from one server to another carrying their states along. [1, 2]

2- Search Engines and Personalization

The *World Wide Web (WWW or Web)* has changed the way we work and the way we live. The Web is currently the most successful hypermedia system in existence, but "the Web is far from done", the inventor of the WWW stresses. We are moving to a more "intelligent", collaborative, and personalized Web. It is becoming more a personal environment for collaborative creation than just for browsing. Personalization is the combined use of technology and customer information to tailor electronic commerce interactions between a business and each individual customer. [3,4]

Personalization usually means gathering and storing information about website visitors, and analyzing this information in order to deliver the right content in a user preferred form and layout. Personalization helps to increase customer satisfaction, promote customer loyalty by establishing a one-to-one relationship between a website and its visitor, and

increases sales by providing products and services tailored to customers' individual needs. The goal of personalization is to better serve the customer by anticipating his needs; it customizes services and products, and establishes a long-term relationship encouraging the customer to return for subsequent visits. Although both customization and personalization refer to the delivery of information, products, or services tailored to users' needs, the two notions differ in several respects. [5] Customization is usually done manually by the user, according to his preferences. An example of customization is My Yahoo (my.yahoo.com) this is shown in Figure (3). In My Yahoo a user can customize content by selecting from various modules (portfolios, company news, weather, currency converter, market summary, etc.) and places them on a customized web page, which is updated periodically. In this case, the locus of control lies with the user who manually selects the modules on the page. In contrast, personalization is automatically performed by a website based on the history of previous interactions with the user, on the user's profile, or on the profiles of like-minded users. For example, Amazon.com recommends items to a user and creates personalized web pages

based on the user's navigation and Purchase history.[4] There are many strategies of web personalization:

Link Personalization This strategy involves selecting the links that are more relevant to the user, changing the original navigation space by reducing or improving the relationships between nodes. E-commerce applications use link personalization to recommend items based on the clients buying history or some categorization of clients based on ratings and opinions. Users who give similar ratings to similar objects are presumed to have similar tastes, so when a user seeks recommendations about products, the site suggests those that are most popular for his class, or those that best correlate with the given product for that class. Link personalization is widely used in www.amazon.com (see Figure (4)) to link the home page with recommendations, new releases, shopping groups, etc. that are personalized. Amazon.com has taken this approach to an extreme by building a "New for you" home page and presenting it to each user, with those new products in which he may be interested.[3.5]

Content Personalization In the web site the content is personalized when nodes (pages) present different information to different users. The difference with link customization is subtle since when links are personalized, part of the contents (the link anchors) presents different information. However, it will refer to content personalization when substantive information in a node is

personalized, other than link anchors. Content personalization can be further classified into two types: node structure customization and node content customization. Structure personalization usually appears in those sites that filter the information that is relevant for the user, showing only sections and details in which the user may be interested. The user may explicitly indicate his preferences, or it may be inferred (semi-) automatically from his profile or from his navigation activity. For example, in my.yahoo.com or in www.mycnn.com users choose a set of "modules" (from a large set including weather, news, music, etc...) and further personalize those modules choosing a set of attributes of the module to be perceived. In Figure (5) present an example of structure customization in my.yahoo.com. [1,3]

Although necessary, user profile management (creating, updating, and maintaining user profiles) is not sufficient for providing personalized services. Information in user profiles has to be analyzed in order to infer users' needs and preferences. In this section we will briefly explain the most popular personalization techniques: rule-based filtering, collaborative filtering, and content-based filtering. All these techniques are used for predicting customers' interests and making recommendations. Rule-based techniques provide a visual editing environment for the business administrator to specify business rules to drive personalization. This requires the administrator, most likely with the

help of a consultant, to figure out the appropriate rules. The rule-based approach provides a flexible mechanism to specify rules for business applications or marketing campaigns. IBM Web Sphere provides a set of tools and services that enable an e-business development team to easily create personalized Web sites. Cross-selling is an e-business example of the rule-based technique. For example, a rule could be specified to offer product X to a customer who has just bought product Y; for example, a customer of a book might be interested in current or previous books by the same author or in books on the same subject. **Collaborative filtering (CF)** was one of the earliest recommendation technologies. CF is used to make a recommendation to a user by finding a set of users, called a neighborhood that have tastes similar to those of the target user. Products that the neighbors like are then recommended to the target user. In other words, CF is based on the idea that people who agreed on their decisions in the past are likely to agree in the future. The process of CF consists of the following three steps: representing products and their rankings, forming a neighborhood, and generating recommendations. During the representation stage, a customer-product matrix is created, consisting of ratings given by all customers to all products. The customer-product matrix is usually extremely large and sparse. It is large because most online stores offer large product sets ranging into millions of products. The results are specified from the fact that each

customer has usually purchased or evaluated only a small subset of the products. To reduce the dimensionality of the customer-product matrix, different dimensionality reducing methods can be applied, such as latent semantic indexing and term clustering. [1, 5]

The neighborhood formation stage is based on computing the similarities between customers in order to group like-minded customers in one neighborhood. The similarity between customers is usually measured by either a correlation or a cosine measure.

Content-based filtering

another recommendation technique is the content-based recommendation. While collaborative filtering identifies customers whose tastes are similar to those of the target customer, content-based recommendation identifies items similar to those the target customer has liked or has purchased in the past. Content-based recommendation has its roots in information retrieval. For example, a text document is recommended based on a comparison between the content of the document and a user profile. The comparison is usually performed using vectors of words and their relative weights. In some cases, the user is asked for feedback after the document has been

shown to him. If the user likes the recommendation, the weights of the words extracted from the document are increased. This process is called relevance feedback.[4,5]

3- The proposed system

The architecture and components of proposed system can be explained in the following block diagram, see figure (6). As you see in above diagram the system contained three components (client side computer, server side computer and internet).

Client side computer

This side of system represent the user of system which use the browser to search for information in the internet . when the user search for keywords the engine system return lists of links to pages that contains this keyword, the browser save the keyword in the cookies ,in the second run of browser the home page display information and news about the keyword that saved in cookies.

Server side compute

This computer contains three parts (local database, search engine application and crawler).

- Local database: this database is collected by crawler program and it contains tables of links and related keywords

used for searching operation. This database (local database) is designed using *SQL* server 2000.

- search engine application :This part represent the main part which is searching program, this application designed and implement in ASP (active page for server)program .this application is search for keyword that requested from the user to local database and return the results as links in ASP page . And send the result page to user in browser and save the keyword in cookies(for second running). This system part can be represented by the following algorithm.

Algorithm search engine

Input: The user input the data that represent his request.

Output: The exact filtering request with previous result.

Step1: While there is a requested keyword from the

user do steps 1-8 do

Step2: Open database connection

Step3: Open record set of table in opened database

Step4: search for this keyword in record set

Step5: if current word is equal to keyword then

Add related link to list of link

Step6: move to next record in record set

Step7: close record set

Step8: close database connection

Step9: end of while

Step10: search for keywords in cookies

Step 10.a: if found

get all new links from internet and generate new home page

Step 10.b: add list of result to it,

Otherwise;

Step 11: return default page and result links to it.

End.

Running of our proposed system:

crawler:-this part of search engine is running automatically to download pages from internet and save it to local database and retrieve keywords from these pages and save it as table contains the keyword and link to that page.

4 Implementation of the proposed system

To explain the proposed system in practical, this research will present full view for implement the proposal.

There are two PC the first one called personalized client and the second called personalized server, see the proposed personalized web page of the client in figure (7).

In figure (7) if the user wants to search about specific subject he will introduce the keywords which will be written in the query. Then these keywords will be taken by the proposed mobile agent which will search inside the database, then the results will appear according the introduced keywords, see figure (8).

In the proposed system the suggestion was be each user has specific personalization, this mean each user has it is specific interface so user A interface will be differ from user B interface from the personalization view.

From that if the user who search in the previous figure (8) will search again for another subjects then the proposed system will introduce the proposed web page as the following, see figure (9):

- Display the default proposed personalized web page.
- Display all the subjects that the user searches in them previously.
- Display the results of the new subject lastly he want it.
- If the search of that user was in same subject the personalization will display the results of previous search and latest news and links about that subject.

Finally must explain who the personalization is done inside the personalized client, that by since there are keywords requested by the user A then the proposal web page will open the database connection and open the record set of table in the opened database. Then the program will search these keywords in record set and compare if a current word is equal to keyword then add related link to itself links. Then will move to next record in record set and continue until no one record is finding.

Searching for keywords in *cookies* , see figure (10) and figure (11) if found get all new links from internet and generate new home page and add list of result to it otherwise return default page and result links to it.

5- Discussion

The proposed system aims to simplify the process of search, so the search will reduce the results to display only the favorite's things. Will always get new information without users interactions (get the information automatically).

The proposed personalization will keep the main news in specific constant place on the user interface. This interface is not a part of the user's search but it will be always exist in any user's interface, that for making the web site view more general

6-Conclusions

This research concludes the following:

- In the proposed system there is a mobile agent would search inside the proposed database.
- In the proposed system the interface will appear as a newspaper, so any thing that the user opens it previously will be displayed as a interface.
- For the news, each user will has specific interface which will be differ from the other user's interfaces.
- The proposed personalization will give a specific view for the web site.

- There is a default shape for web site, this mean, if we search about cars for example, and shutdown the computer when the user turn on the computer in any other time the web site will display the all interfaces related to cars.

References

- [1] Lil Y., Well G. B. , Whigham P. and Mulgan N., “**An Agent-oriented software paradigm and the design of a new generation of spatial information system**”, University of Otago. Dunedin, New Zealand, 2000.
- [2] Nicholas R. Jennings and Michael Wooldridge, “**Agent-Oriented Software Engineering**”, Department of Electronic Engineering, Queen Mary & Westfield College University of London, 2002.
- [3] S. paulakis, c. lampos, m. eirinaki, m. Vazirgiannis, SEWep: “**a web mining system supporting semantic personalization**”, in proc. Of the ECML/PKDD 2004 CONFERENCE, pisa, Italy, September 2004.
- [4] G. Silber, k. McCoy, “**efficiently computed lexical chains an intermediate representation for automatic text summarization**”, computational linguistics, 29(1), 2003.
- [5] B. Sarwar, G. karypis, J.konstan, J. Riedl, “**item-based collaborative filtering recommendation algorithm**”, university of Hong Kong , CT 910, 2001,

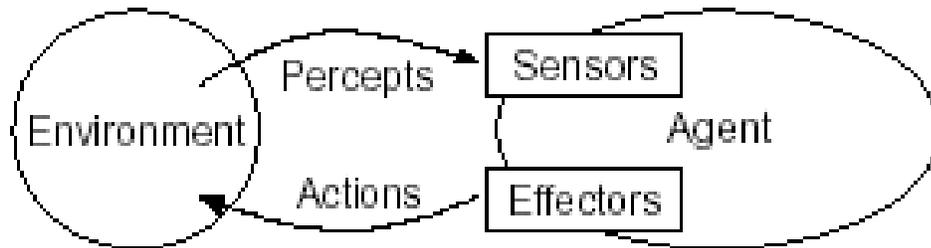


Figure (1) Generic agent

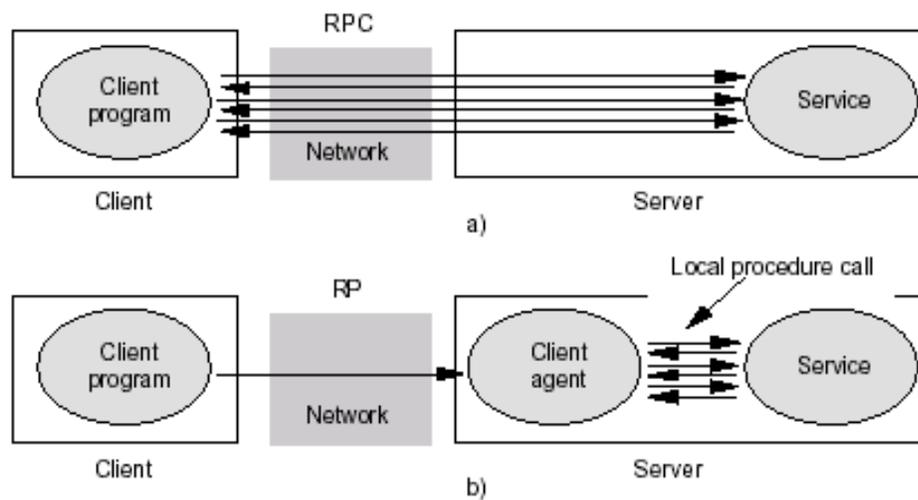


Figure (2) a Remote procedure call and b Remote programming



Figure (3) My Yahoo.com.

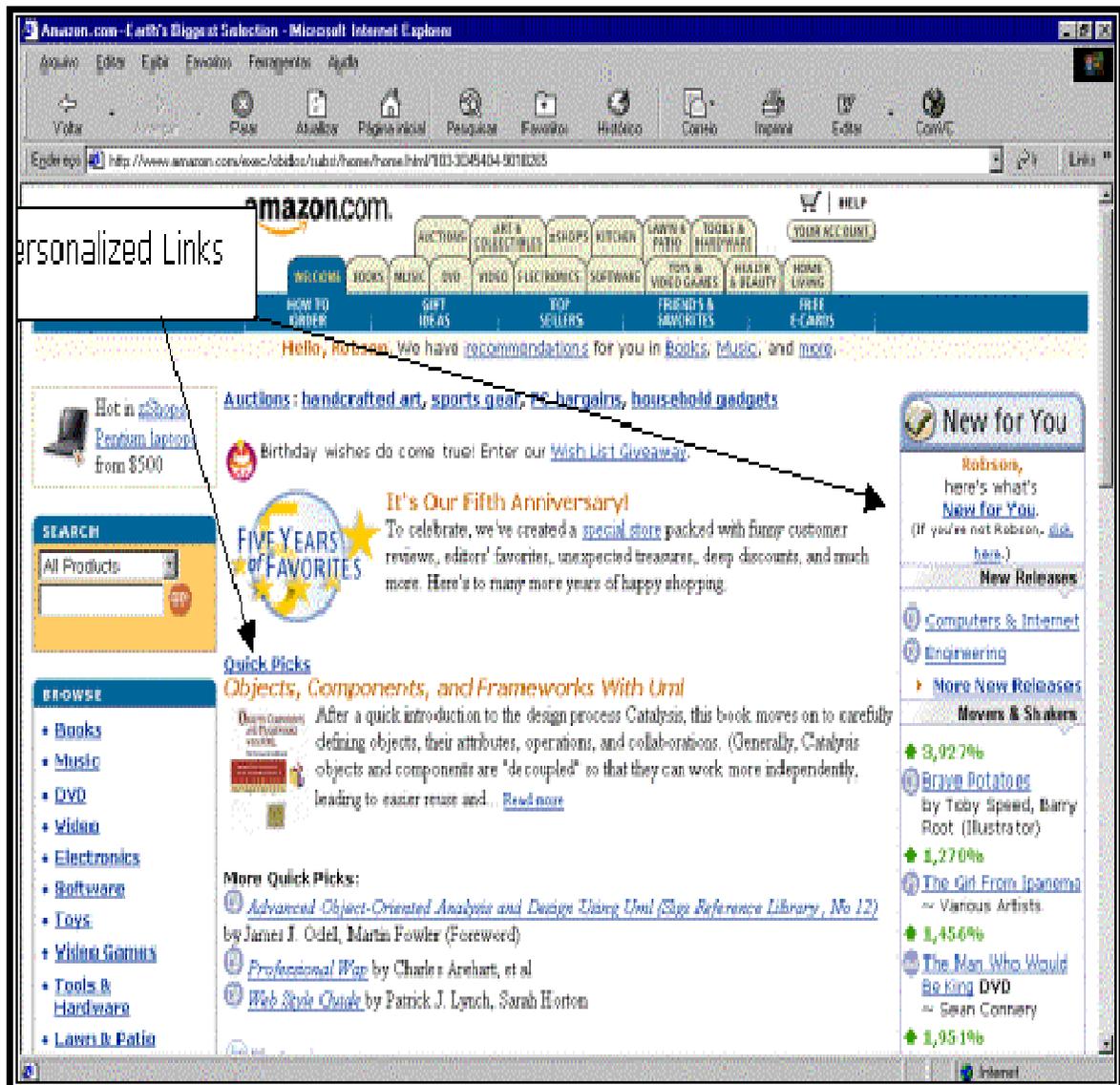


Figure (4) Using Link personalization in www.amazon.com



Figure (5) Structure customization in my.yahoo.com.

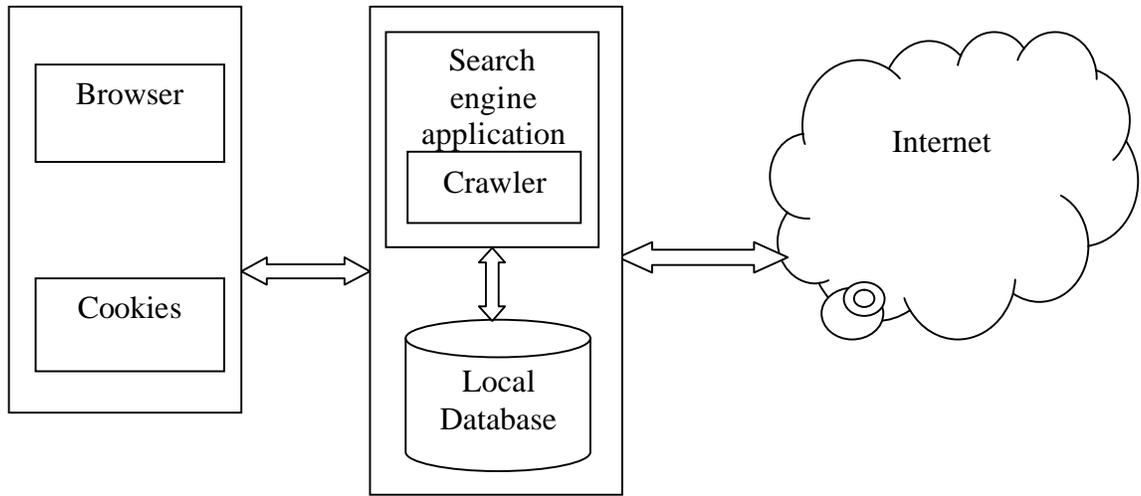


Figure (6) explain the architecture of the proposed system

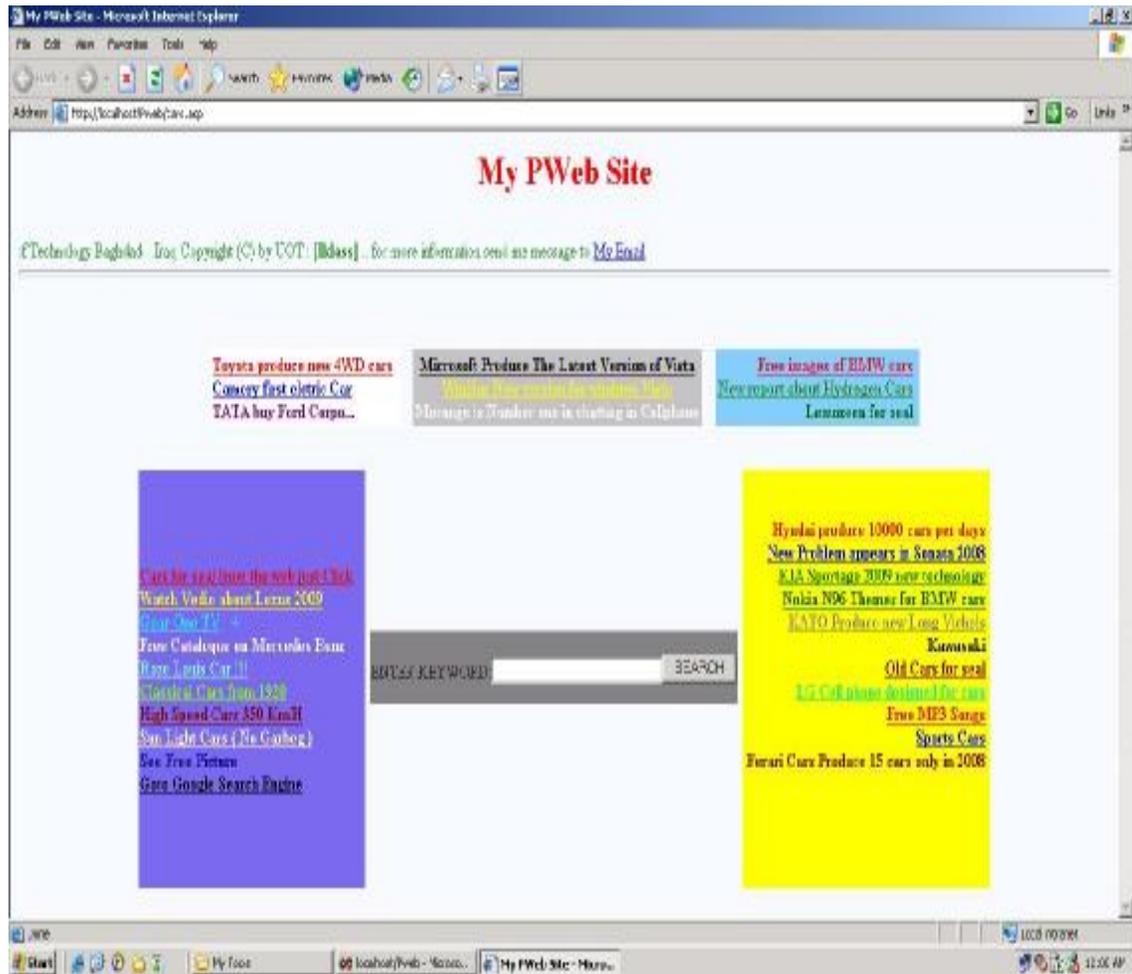


Figure (7) proposed personalized web site

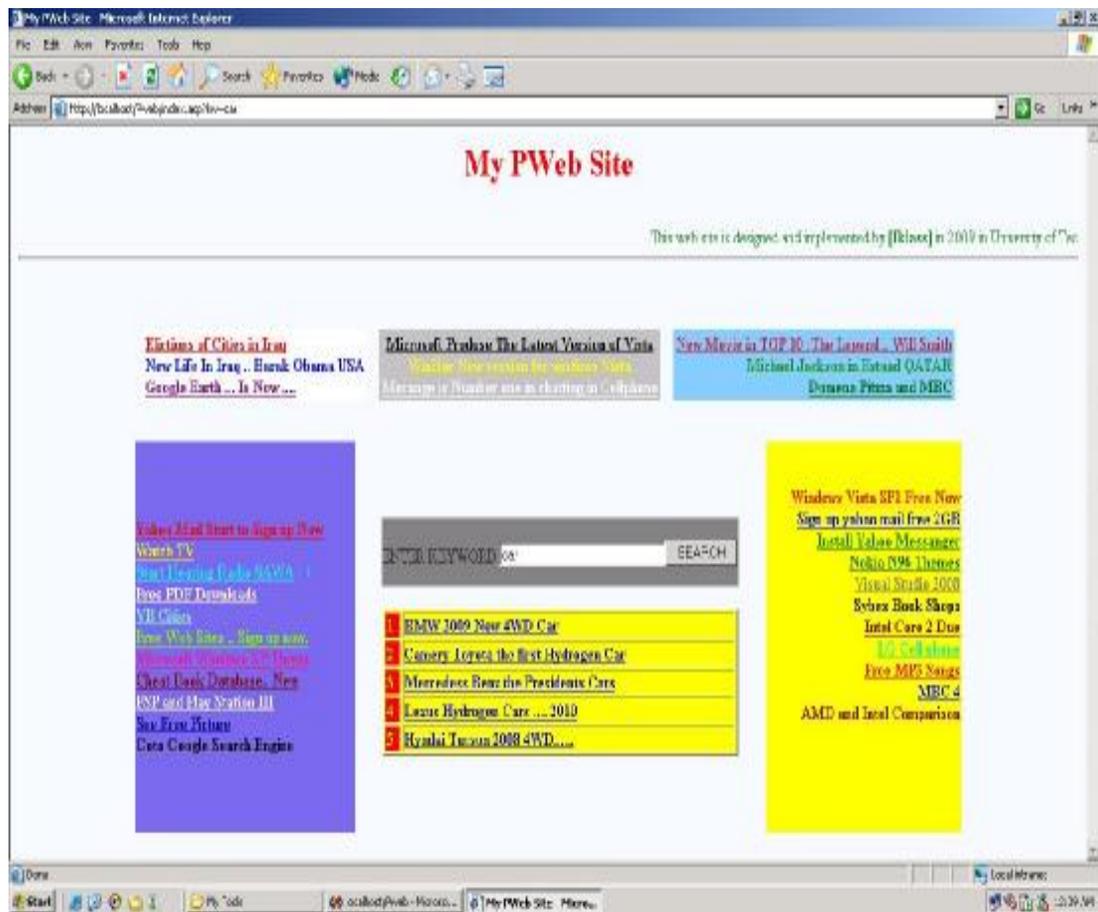


Figure (8) the results of introduced keywords.

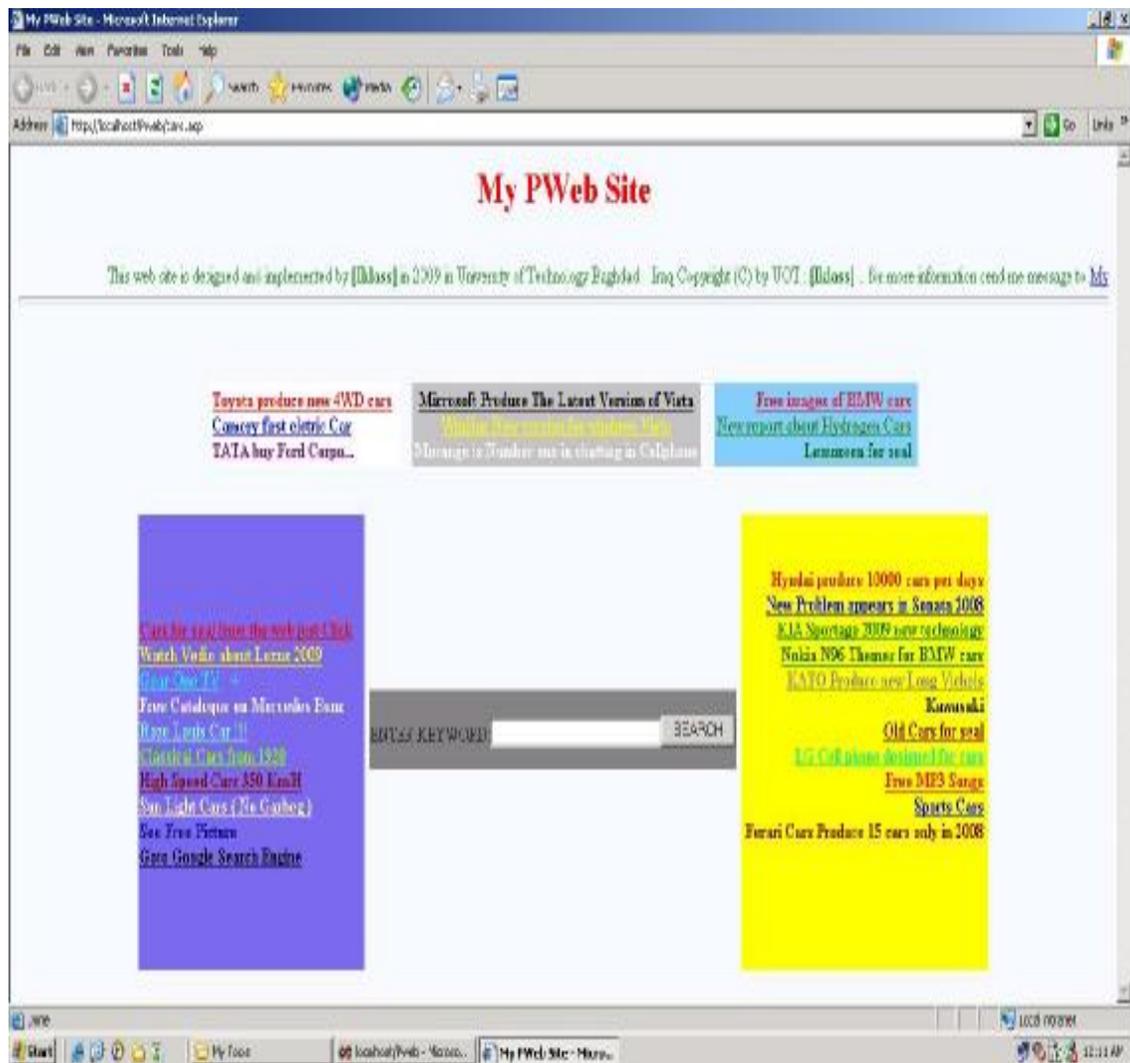


Figure (9) the personalization of user search in figure (8).

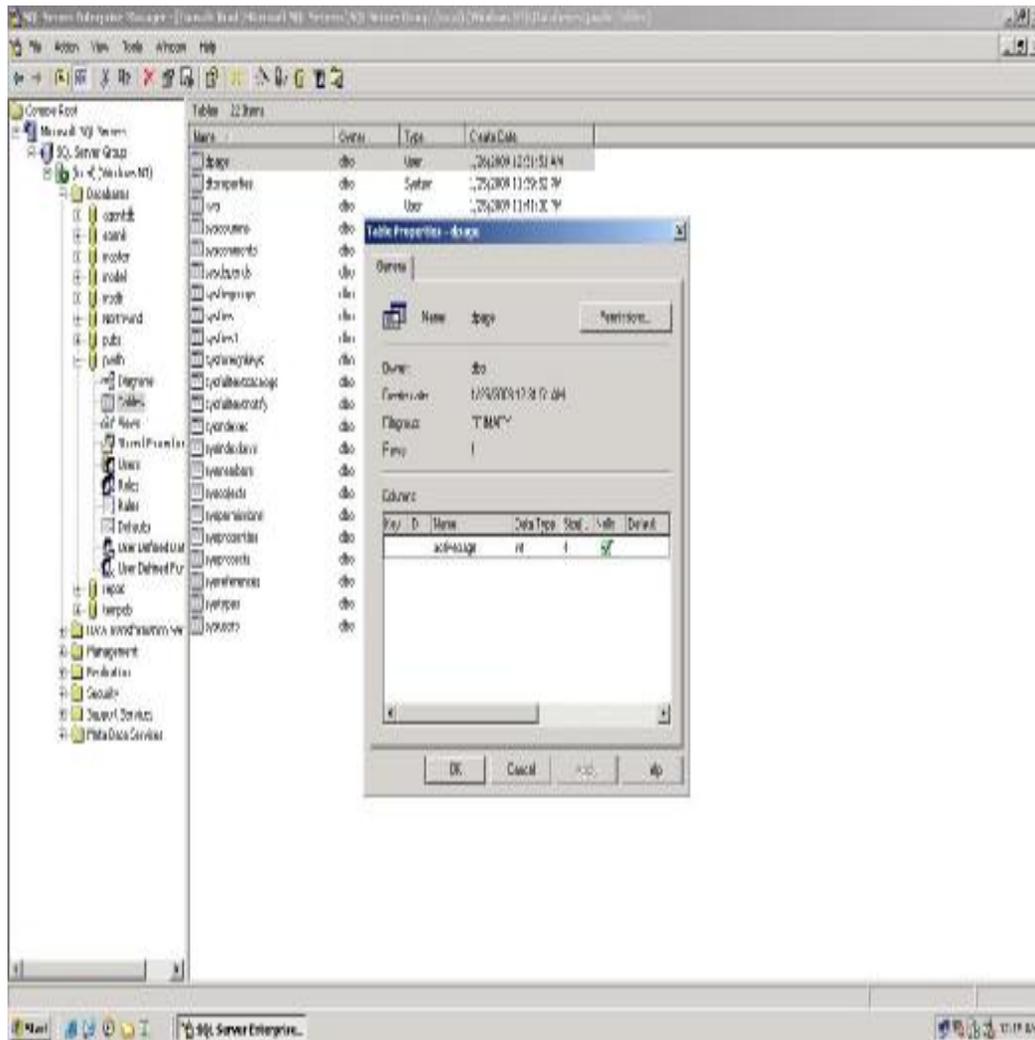


Figure (11) table properties of the personalized home page.