



US 20060136378A1

(19) **United States**

(12) **Patent Application Publication**
Martin

(10) **Pub. No.: US 2006/0136378 A1**
(43) **Pub. Date: Jun. 22, 2006**

(54) **SEARCH ENGINE FOR A COMPUTER NETWORK**

Publication Classification

(51) **Int. Cl.**
G06F 17/30 (2006.01)
(52) **U.S. Cl.** **707/3**

(75) **Inventor: Anthony G. Martin, Los Altos, CA (US)**

Correspondence Address:
OKAMOTO & BENEDICTO, LLP
P.O. BOX 641330
SAN JOSE, CA 95164 (US)

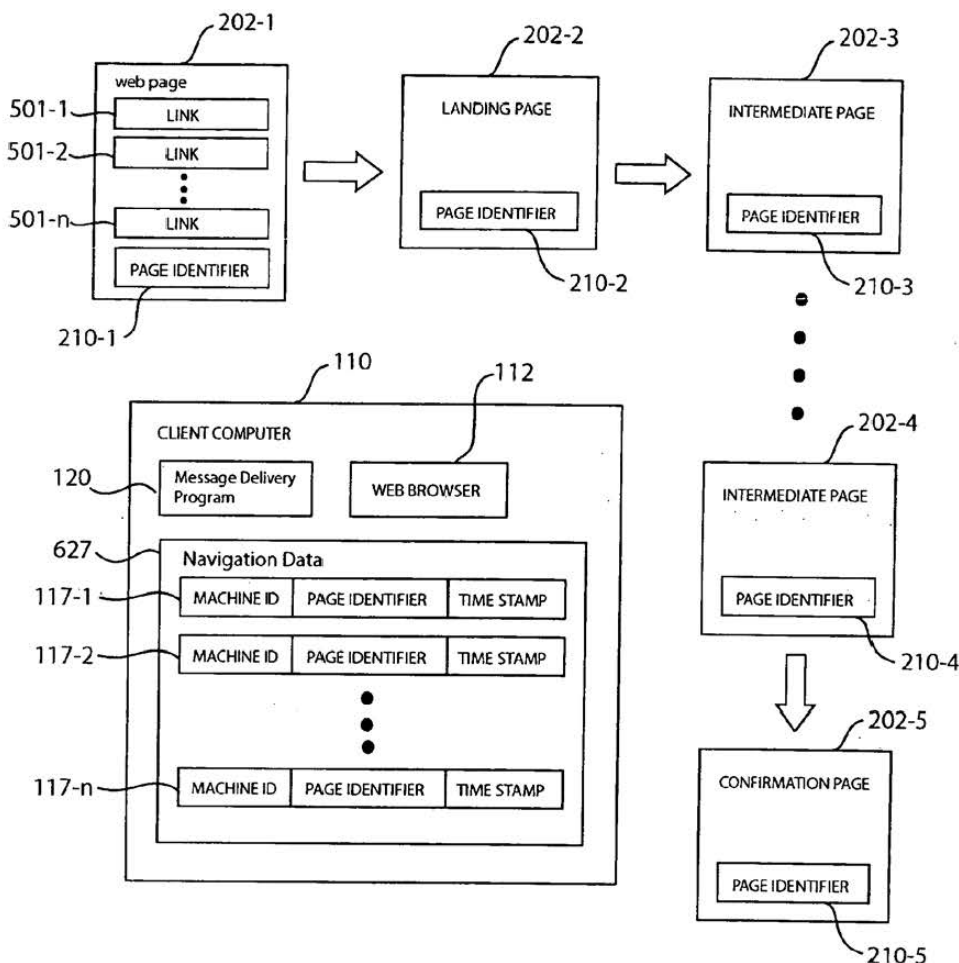
(73) **Assignee: CLARIA CORPORATION**

(21) **Appl. No.: 11/015,583**

(22) **Filed: Dec. 17, 2004**

(57) **ABSTRACT**

In one embodiment, client-side programs are employed to observe the navigation of consumers to various websites. Addresses of web pages viewed by consumers may be used to fetch the web pages. A fetched web page may be parsed for one or more keywords. The relevance of the web page to a keyword may be ranked according to consumer preferences, which may relate to consumer interaction with the web page. Web pages and their ranking information may be stored in an index. The index may be consulted to find links to web pages relevant to a keyword employed in a search request.



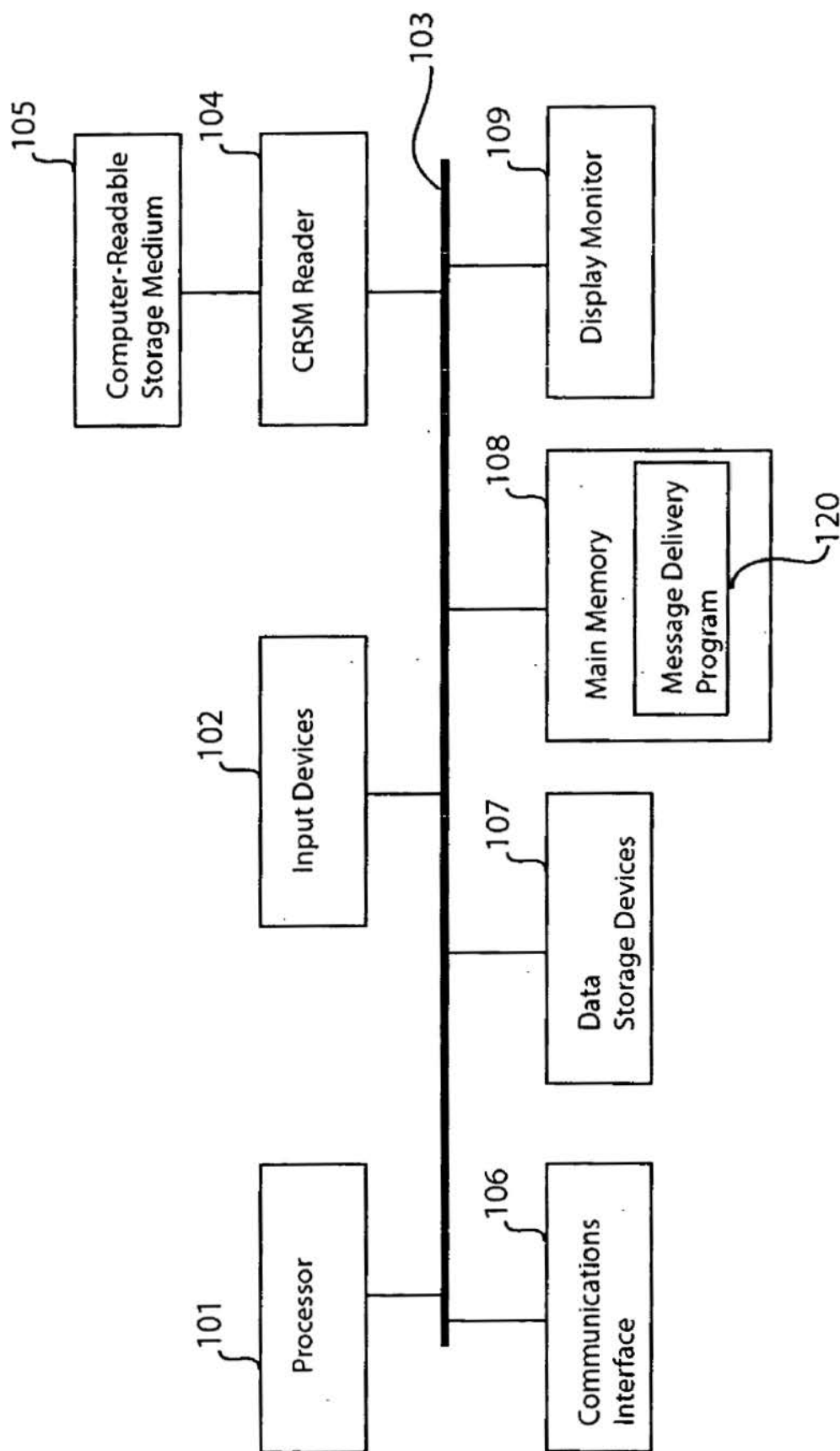
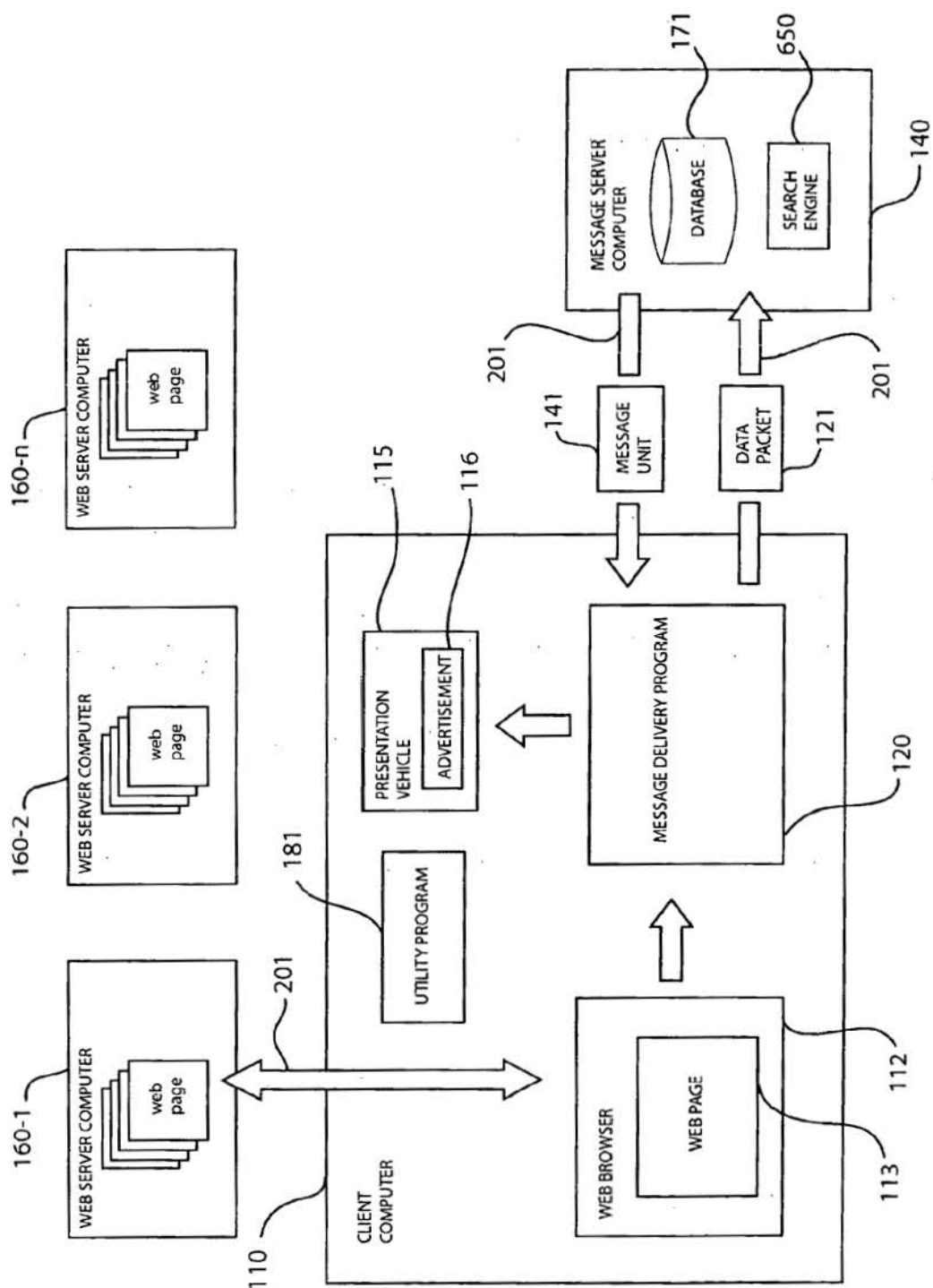


FIG. 1



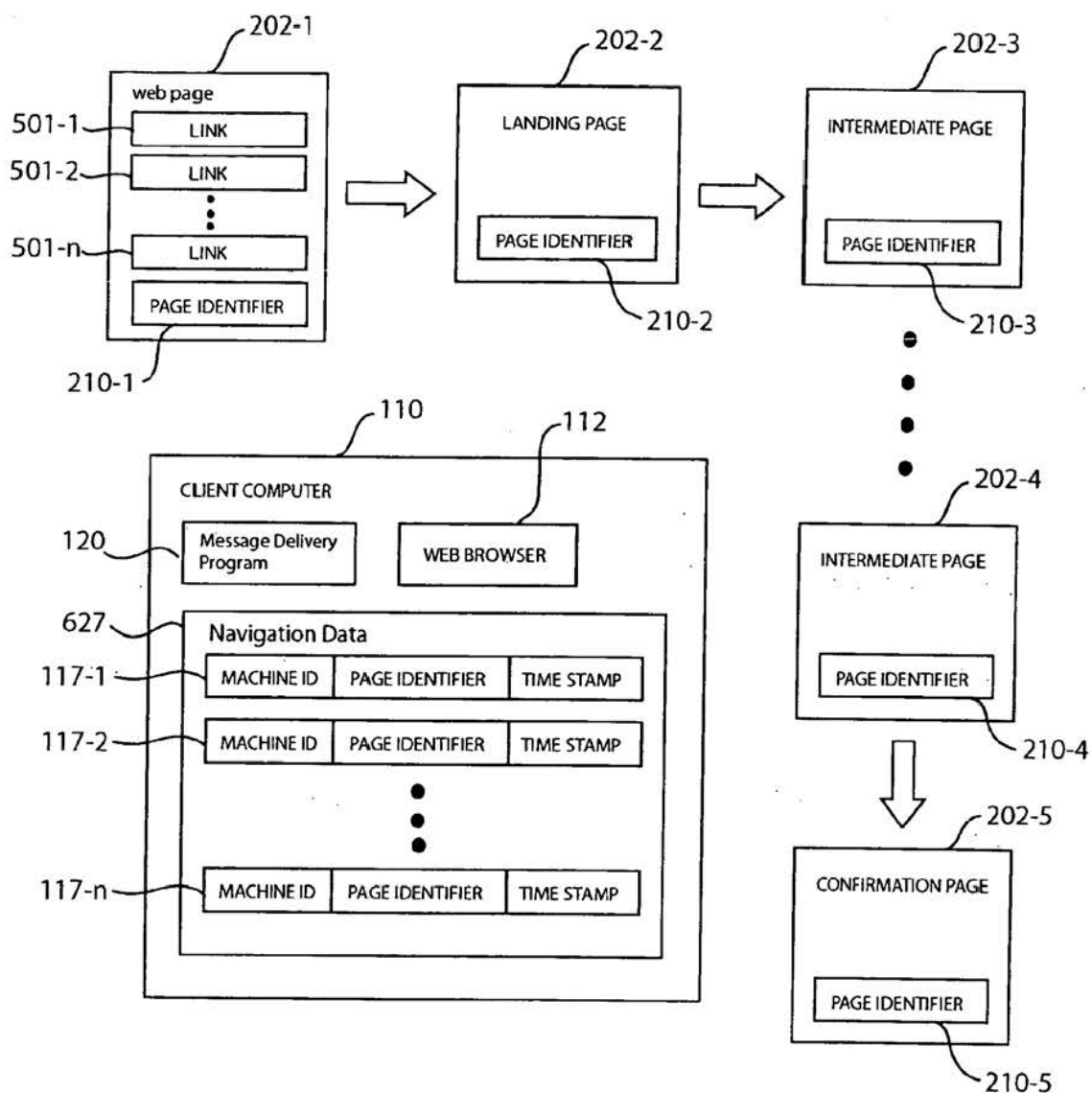


FIG. 3

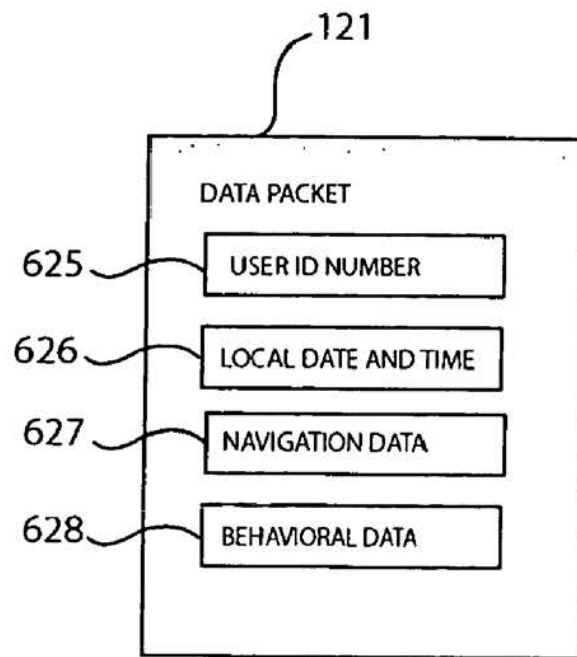


FIG. 4

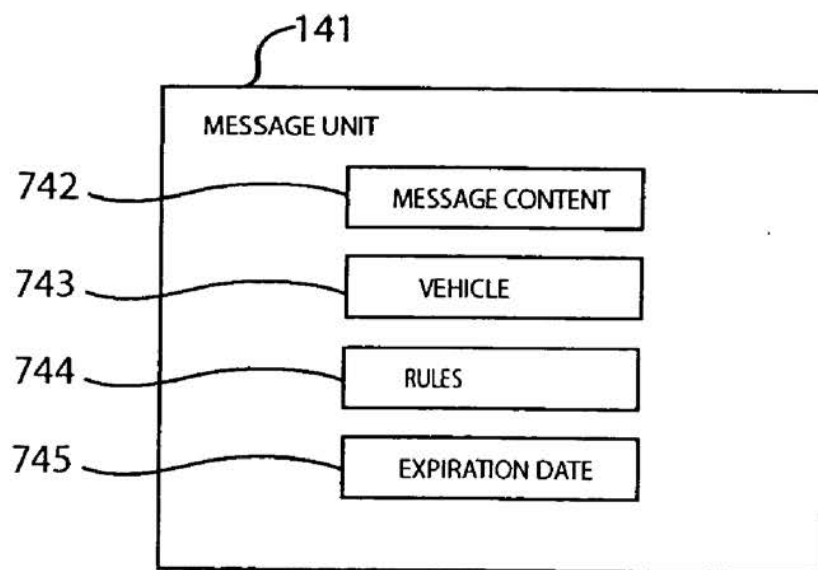


FIG. 5

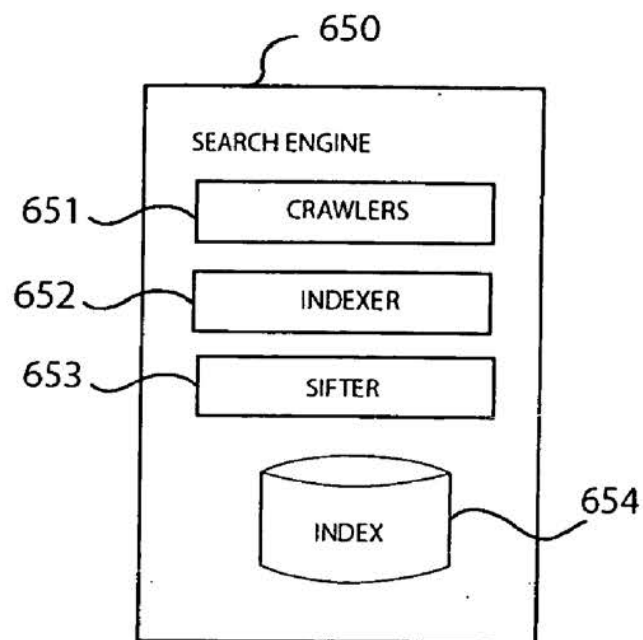


FIG. 6

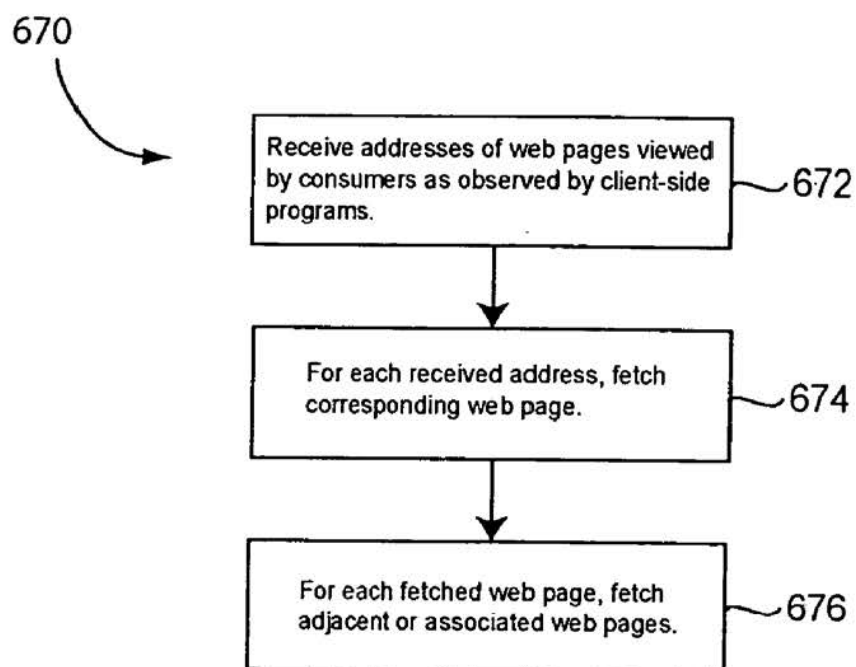


FIG. 7

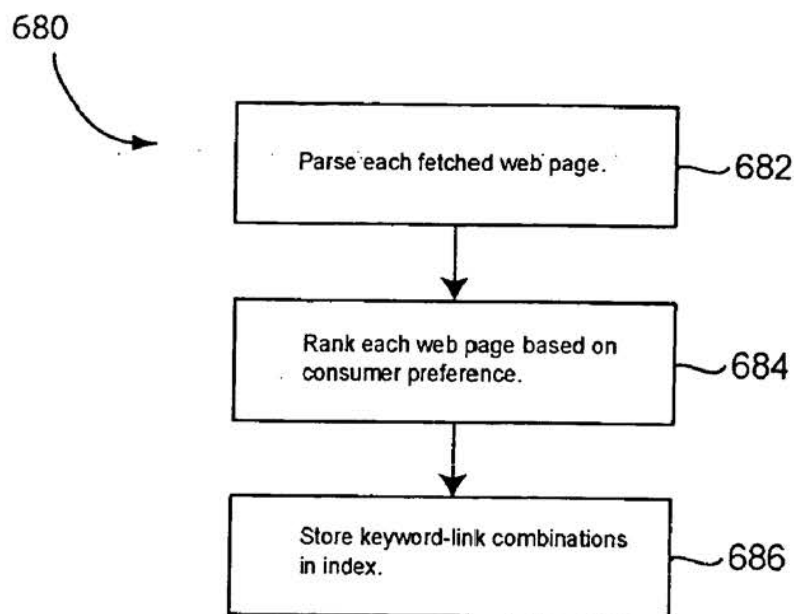


FIG. 8

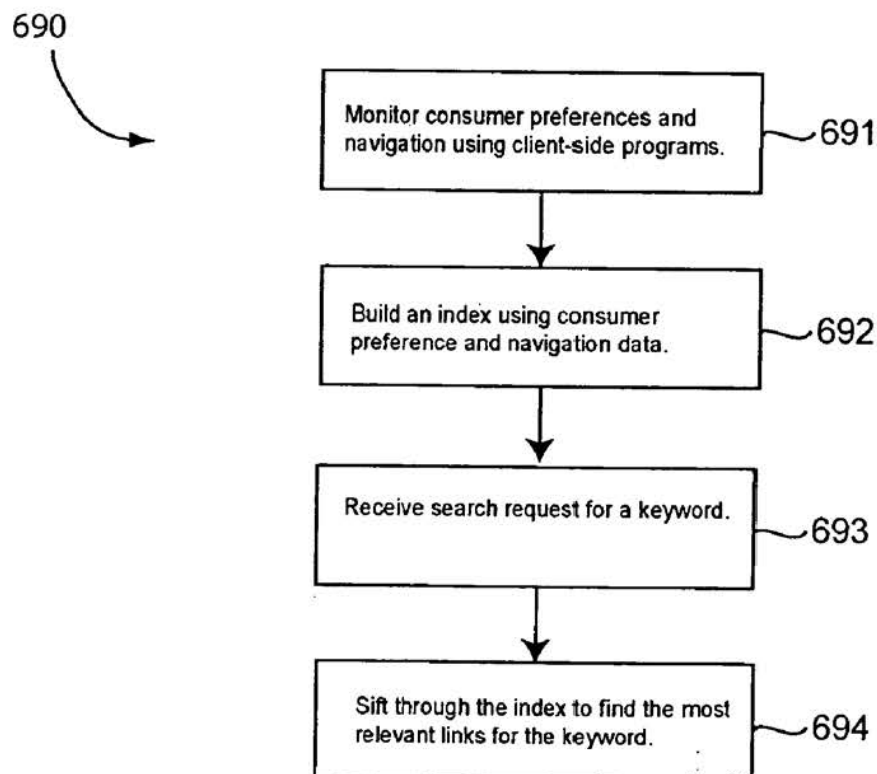


FIG. 9

SEARCH ENGINE FOR A COMPUTER NETWORK

BACKGROUND OF THE INVENTION

[0001] 1. Field of the Invention

[0002] The present invention relates generally to computer systems, and more particularly but not exclusively to search engines.

[0003] 2. Description of the Background Art

[0004] The Internet is an example of a computer network. On the Internet, consumers on client computers may access various types of information resident in server computers. A server computer that provides information over the Internet is also referred to as a "web server" or a "website". A website may provide information about various topics, or offer goods and services. A website may comprise a plurality of downloadable documents, such as web pages and files. A consumer may use a web browser to receive and view a web page.

[0005] Because of the large number of websites on the Internet, searching the Internet for specific information usually requires the services of a search engine. Generally speaking, a search engine helps consumers look for relevant web pages. A typical search engine accepts a word or a phrase, referred to herein as a "keyword." The search engine employs the keyword in conjunction with its search algorithm to find matching web pages. The links to matching web pages are presented to the consumer in the form of a listing referred to as "search results." Typical search results list the matching web pages as clickable links, with each link pointing to a corresponding web page. Examples of websites with search engines include Yahoo, Google, and Alta Vista.

[0006] A search engine includes an index (also referred to as a "catalog") containing links to web pages responsive to particular keywords. Entries to the index may be obtained by accepting submissions from website operators, as in the case of so-called "human-powered search engines," by using a crawler, or both. Crawler-based search engines employ a crawler (also referred to as a "spider" or "robot") to automatically fetch web pages over the Internet. A crawler may be given an address of a web page, fetch that web page, then fetch web pages pointed to by that web page, and so on. Web pages fetched by the crawler are stored in the index.

[0007] A keyword and a link responsive to the keyword are also collectively referred to as a "keyword-link combination." A typical search engine ranks keyword-link combinations stored in an index based on the relevance of the link to the keyword. A keyword-link combination may be ranked based on the number of times the keyword appears on the web page pointed to by the link. The web page pointed to by the link may be further ranked based on the number and credibility of other web pages pointing to it. For example, when a keyword appears the same number of times on two different web pages, the web page referenced by more, credible web pages gets the higher ranking. When a consumer employs a search engine to search the Internet using a keyword, the search engine sifts through the index to find the highest ranked keyword-link combinations having that keyword. Links of the highest ranked keyword-link combinations are presented to the consumer as search results.

SUMMARY

[0008] In one embodiment, client-side programs are employed to observe the navigation of consumers to various websites. Addresses of web pages viewed by consumers may be used to fetch the web pages. A fetched web page may be parsed for one or more keywords. The relevance of the web page to a keyword may be determined according to consumer preferences, which may relate to consumer interaction with the web page. Web pages and their ranking information may be stored in an index. The index may be consulted to find links to web pages relevant to a keyword employed in a search request.

[0009] These and other features of the present invention will be readily apparent to persons of ordinary skill in the art upon reading the entirety of this disclosure, which includes the accompanying drawings and claims.

DESCRIPTION OF THE DRAWINGS

[0010] FIG. 1 shows a schematic diagram of an example computer that may be used in embodiments of the present invention.

[0011] FIG. 2 shows a schematic diagram of a computing environment in accordance with an embodiment of the present invention.

[0012] FIG. 3 schematically illustrates the tracking of a conversion process, in accordance with an embodiment of the present invention.

[0013] FIG. 4 schematically shows a data packet in accordance with an embodiment of the present invention.

[0014] FIG. 5 schematically shows a message unit in accordance with an embodiment of the present invention.

[0015] FIG. 6 shows a schematic diagram of a search engine in accordance with an embodiment of the present invention.

[0016] FIG. 7 shows a flow diagram of a method that may be performed by a crawler, in accordance with an embodiment of the present invention.

[0017] FIG. 8 shows a flow diagram of a method that may be performed by an indexer, in accordance with an embodiment of the present invention.

[0018] FIG. 9 shows a flow diagram of a method for responding to a search request, in accordance with an embodiment of the present invention.

[0019] The use of the same reference label in different drawings indicates the same or like components.

DETAILED DESCRIPTION

[0020] In the present disclosure, numerous specific details are provided, such as examples of apparatus, components, and methods, to provide a thorough understanding of embodiments of the invention. Persons of ordinary skill in the art will recognize, however, that the invention can be practiced without one or more of the specific details. In other instances, well-known details are not shown or described to avoid obscuring aspects of the invention.

[0021] Being computer-related, it can be appreciated that the components disclosed herein may be implemented in

hardware, software, or a combination of hardware and software (e.g., firmware). Software components may be in the form of computer-readable program code stored in a computer-readable storage medium, such as memory, mass storage device, or removable storage device. For example, a computer-readable medium may comprise computer-readable program code for performing the function of a particular component. Likewise, computer memory may be configured to include one or more components, which may then be executed by a processor. Components may be implemented separately in multiple modules or together in a single module.

[0022] Referring now to FIG. 1, there is shown a schematic diagram of an example computer that may be used in embodiments of the present invention. Depending on its configuration, the computer shown in the example of FIG. 1 may be employed as a client computer, a server computer, or other data processing device. The computer of FIG. 1 may have less or more components to meet the needs of a particular application. As shown in FIG. 1, the computer may include a processor 101, such as those from the Intel Corporation or Advanced Micro Devices, for example. The computer may have one or more buses 103 coupling its various components. The computer may include one or more input devices 102 (e.g., keyboard, mouse), a computer-readable storage medium (CRSM) 105 (e.g., floppy disk, CD-ROM), a CRSM reader 104 (e.g., floppy drive, CD-ROM drive), a display monitor 109 (e.g., cathode ray tube, flat panel display), a communications interface 106 (e.g., network adapter, modem) for coupling to a network, one or more data storage devices 107 (e.g., hard disk drive, optical drive, FLASH memory), and a main memory 108 (e.g., RAM). Software embodiments may be stored in a computer-readable storage medium 105 for reading into a data storage device 107 or main memory 108. Software embodiments in main memory 108 may be executed by processor 101. In the example of FIG. 1, main memory 108 is shown as having a message delivery program 120, which is further described below. Message delivery program 120 and other programs in main memory 108 may be loaded from a computer-readable storage medium 105, a data storage device 107, or over the Internet by way of communications interface 106, for example. Message delivery program 120 and other programs in main memory 108 may be executed by processor 101.

[0023] FIG. 2 shows a schematic diagram of a computing environment in accordance with an embodiment of the present invention. In the example of FIG. 2, the computing environment includes one or more web server computers 160 (i.e., 160-1, 160-2, . . .), one or more client computers 110, one or more message server computers 140, and other computers not specifically shown. In the example of FIG. 2, a client computer 110 communicates with server computers (e.g., a web server computer or a message server computer) over the Internet. As such, arrows 201 denote Internet connections in this example. Intermediate nodes such as gateways, routers, bridges, Internet service provider networks, public-switched telephone networks, proxy servers, firewalls, and other network components are not shown for clarity.

[0024] A client computer 110 is typically, but not necessarily, a personal computer such as those running the Microsoft Windows™ operating system, for example. A consumer may employ a suitably equipped client computer

110 to get on the Internet and access computers coupled thereto. For example, a client computer 110 may be used to access web pages from a web server computer 160.

[0025] A web server computer 160 may be a server computer hosting a website, which comprises web pages designed to attract consumers surfing on the Internet. A web server computer 160 may include web pages supporting advertisements, downloadable computer programs, products available for online purchase, and so on. As can be appreciated, a website may be on one or more server computers.

[0026] A message server computer 140 may include the functionalities of a web server computer 160. In one embodiment, a message server computer 140 further includes a database 171. Database 171 may be a commercially available database, such as those available from the Oracle Corporation. Database 171 may store client data received from message delivery programs 120 running in client computers 110. The client data may be transmitted from a client computer 110 to message server computer 140 in a data packet 121. The client data may include navigation and behavioral data obtained by a message delivery program 120 by monitoring a consumer's online activities. In the example of FIG. 2, message server computer 140 is shown as communicating with one client computer 110 for clarity of illustration. In practice, message server computer 140 receives data packets 121 containing client data from a plurality of client computers 110, each of which has a message delivery program 120. A message server computer 140 may also include downloadable computer programs and files for supporting, updating, and maintaining software components on a client computer 110.

[0027] Web server computers 160 and message server computers 140 are typically, but not necessarily, server computers such as those available from Sun Microsystems, Hewlett-Packard, or International Business Machines. A client computer 110 may communicate with a web server computer 160 or a message server computer 140 using client-server protocol. It is to be noted that client-server computing is well known in the art and will not be further described here.

[0028] As shown in FIG. 2, a client computer 110 may include a web browser 112 and a message delivery program 120. Web browser 112 may be a commercially available web browser or web client. In one embodiment, web browser 112 comprises the Microsoft Internet Explorer™ web browser. A web browser allows a consumer on a client computer to access a web page. In the example of FIG. 2, web browser 112 is depicted as displaying a web page 113 from a web server computer 160. A web page, such as web page 113, has a corresponding address referred to as a "URL" (Uniform Resource Locator). Web browser 112 is pointed to the URL of a web page to receive that web page in client computer 110. Web browser 112 may be pointed to a URL by entering the URL at an address window of web browser 112, or by clicking a link pointed to that URL, for example.

[0029] In one embodiment, message delivery program 120 is downloadable from a message server computer 140 or a web server computer 160. Message delivery program 120 may be downloaded to a client computer 110 in conjunction with the downloading of another computer program. For example, message delivery program 120 may be downloaded to client computer 110 along with a utility program

181 that is provided free of charge or at a reduced cost. Utility program **181** may be an e-wallet or calendar program, for example. Utility program **181** may be provided to a consumer in exchange for the right to deliver advertisements to that consumer's client computer **110** via message delivery program **120**. In essence, revenue from advertisements delivered to the consumer helps defray the cost of creating and maintaining the utility program. Message delivery program **120** may also be provided to the consumer along with free or reduced cost access to an online service, for example.

[0030] Message delivery program **120** is a client-side program in that it is stored and run in a client computer **110**. Message delivery program **120** may comprise computer-readable program code for displaying advertisements in a client computer **110** and for monitoring the online activity of a consumer on the client computer **110**. It is to be noted that the mechanics of monitoring a consumer's online activity, such as determining where a consumer is navigating to, the URL of web pages received in client computer **110**, the domain names of websites visited by the consumer, what the consumer is typing on a web page, what keyword the consumer is providing to a search engine, whether the consumer clicked on a link or an advertisement, when the consumer activates a mouse or keyboard, and the like, is, in general, known in the art and not further described here. For example, message delivery program **120** may learn of consumer online activities by receiving event notifications from web browser **112**.

[0031] Message delivery program **120** may record the consumer's online activity for reporting to message server computer **140**. The recorded consumer online activity is also referred to as "client data," and provided to message server computer **140** using data packets **121**. Message server computer **140** may use the client data to provide targeted advertisements to the consumer. Message server computer **140** may include the advertisement or data for displaying the advertisement in a message unit **141**. In the example of FIG. 2, the targeted advertisement is labeled as advertisement **116** and displayed in a presentation vehicle **115**. Presentation vehicle **115** may be a pop-under, pop-up, separate browser window, custom browser window, or other means for displaying an advertisement on a computer screen. Techniques for delivering advertisements to client computers using a client-side program are also disclosed in commonly-owned U.S. application Ser. No. 10/152,204, entitled "Method and Apparatus for Displaying Messages in Computer Systems," filed on May 21, 2002 by Scott G. Eagle, David L. Goulden, Anthony G. Martin, and Eugene A. Veteska, which is incorporated herein by reference in its entirety.

[0032] As will be more apparent below, message delivery programs are primarily used in embodiments of the present invention to obtain client data for building a search engine index, not necessarily to display presentation vehicles in a client computer **110**. That is, a message delivery program does not necessarily have to display advertisements in a client computer **110**. This is advantageous in that consumers may be allowed to obtain a free or reduced cost utility program **181** (or other benefits) without having to see advertisements from the provider or sponsor of the utility program.

[0033] FIG. 3 schematically illustrates the tracking of a conversion process, in accordance with an embodiment of

the present invention. A "conversion" occurs when a consumer makes an online purchase in response to an advertisement presented to her. As will be more apparent below, embodiments of the invention may use conversion as a factor in ranking a keyword-link combination. In the example of FIG. 3, web pages **202** (i.e., **202-1**, **202-2**, . . .) may be sequentially displayed in the same or separate windows of web browser **112**. Each web page **202** includes a page identifier **210** (i.e., **210-1**, **210-2**, . . .), which may be a URL. Message delivery program **120** records the URLs of web pages **202** viewed by the consumer as well the amount of time the consumer spent with each web page as navigation data. For purposes of the present disclosure, the web pages "viewed" by consumers are those web pages received in client computers. In the example of FIG. 3, navigation data **627** comprises log entries **117** (i.e., **117-1**, **117-2**, . . .). Each log entry **117** includes a machine ID anonymously identifying the client computer **110** (or the consumer), a page identifier, and a time stamp indicating when the log entry **117** was made. The time stamps between log entries **117** provide an estimate of the amount of time the consumer spent viewing the indicated web page. A log entry **117** may be created for each web page **202** viewed by the consumer. For example, a log entry **117-1** may be created when the consumer clicks on a link **501** to receive web page **202-2** in client computer **110**, a log entry **117-2** may be created when the consumer receives web page **202-3** in client computer **110**, and so on.

[0034] In the example of FIG. 3, web page **202-2** is also referred to as a "landing page" because it is the web page directly pointed to by an advertisement having a link in web page **202-1**. A website that sells products online may also have a "confirmation page" **202-5**. A confirmation page is a web page provided to the consumer to confirm a just completed online purchase. A website may have "intermediate pages" **202-3**, **202-4**, and so on between a landing page **202-2** and a confirmation page **202-5**. An intermediate page may be an online product catalog, shopping cart, and other types of web pages. The page identifiers of landing and confirmation pages of popular or partner websites may be stored in database **171** of message server computer **140**. The page identifiers of the landing and confirmation pages may be compared to those in navigation data **627** of client computers **110** to determine the number of consumers making online purchases from a particular web page. As can be appreciated, web pages from which consumers make purchases may be highly relevant, and may thus be deemed a consumer preferred page. Techniques for monitoring consumer purchase behavior are also disclosed in commonly-owned U.S. application Ser. No. 10/464,419, entitled "Generation of Statistical Information In a Computer Network," filed by David L. Goulden and Dominic Bennett on Jun. 17, 2003, which is incorporated herein by reference in its entirety.

[0035] FIG. 4 schematically shows a data packet **121** in accordance with an embodiment of the present invention. A data packet **121** may include a user ID number **625** anonymously identifying the consumer or her client computer, local date and time **626** indicating when the data packet **121** was sent from the client computer **110** to message server computer **140**, navigation data **627**, and behavioral data **628**. Navigation data **627** include navigation related information, such as addresses of web pages viewed by the consumer. Example navigation data **627** has been discussed in connec-

tion with FIG. 3. Behavioral data 628 may contain information indicative of consumer online behavior, such as advertisements and links the consumer clicked on, advertisements converted into a purchase, and the like.

[0036] Referring to FIG. 5, there is schematically shown a message unit 141 in accordance with an embodiment of the present invention. A message unit 141 may include a message content 742, a vehicle 743, rules 744, and an expiration date 745. Message content 742 may include computer-readable program code, text, images, audio, video, hyperlink, and other information. A message content 742 may be an advertisement or computer-readable program code for receiving an advertisement from an ad server, for example.

[0037] Vehicle 743 indicates the presentation vehicle to be used in presenting the message content indicated by message content 742. For example, vehicle 743 may call for the use of a pop-up, banner, message box, text box, slider, separate window, window embedded in a web page, or other presentation vehicle to display a message content. In the example of FIG. 2, advertisement 116 and presentation vehicle 115 may be specified in a message content 742 and vehicle 743, respectively, of a message unit 141.

[0038] Rules 744 indicate one or more triggering conditions for processing a message unit 141. Rules 744 may specify to display a message content 742 when a consumer navigates to a specific web page or as soon as the message unit 141 is received in a client computer 110. For example, a car company may contract with the operator of a message server computer 140 to deliver a message unit 141 containing an advertisement for a minivan (hereinafter, "minivan message unit"). The rules 744 of the minivan message unit may specify that the minivan advertisement is to be displayed to consumers viewing the minivan web page of "cars.com". In this example, the minivan web page of cars.com has the URL "www.cars.com/minivans". When a consumer visits the main page (or any web page) of "cars.com", message delivery program 120 (see FIG. 2) will send a data packet 121 to message server computer 140 indicating that the consumer is on "cars.com". In response, message server computer 140 will send the minivan message unit to client computer 110. When the consumer navigates to the URL "www.cars.com/minivans", message delivery program 120 will detect that the minivan message unit has been triggered for processing (i.e., rules 744 of the minivan message unit have been satisfied). Accordingly, message delivery program 120 will process the minivan message unit by displaying it.

[0039] Rules 744 may also include: (a) a list of domain names at which the content of a message unit 141 is to be displayed, (b) URL sub-strings that will trigger displaying of the content of the message unit 141, and (b) time and date information. As shown in FIG. 5, a message unit 141 may also include an expiration date 745. Expiration date 745 indicates the latest date and time the message unit 141 can still be processed. In one embodiment, expired message units 141 are not processed even if their rules 744 have been satisfied. Expired message units 141 may be removed from client computer 110.

[0040] FIG. 6 shows a schematic diagram of a search engine 650 in accordance with an embodiment of the present invention. A search engine 650 may be implemented in a message server computer 140 as shown in FIG. 2, or in one

or more dedicated server computers, for example. A search engine 650 may be employed to provide search results to a consumer performing a keyword search. A keyword may be a word or a phrase, for example. Search results may be presented to a consumer as a web page containing one or more links responsive to a search request for one or more keywords. In the example of FIG. 6, a search engine 650 may include one or more crawlers 651, an indexer 652, a sifter 653, and an index 654.

[0041] Crawlers 651 may comprise computer-readable program code running on one or more computers. A crawler 651, also referred to as a "spider", "robot", or "bot", performs a discovery function for a search engine 650. A search engine 650 employs a crawler 651 to fetch web pages over the Internet. Note that a search engine 650 may divide the web page fetching function to one or more crawlers 651. The fetched web pages may be indexed, ranked and stored in an index 654.

[0042] FIG. 7 shows a flow diagram of a method 670 that may be performed by a crawler 651, in accordance with an embodiment of the present invention. In step 672, the crawler 651 receives addresses of web pages viewed by consumers as observed by client-side programs, such as a message delivery program 120. The addresses of web pages viewed by consumers may be extracted from navigation data contained in data packets 121, which may be stored in a database 171. In step 674, using the addresses received in step 672, the crawler 651 fetches the web pages viewed by the consumers. In step 676, for each fetched web page, a crawler 651 may fetch adjacent or associated web pages. An adjacent web page may be a web page one level below or above the fetched web page. A web page one level above a fetched web page may be a clickable link on the fetched web page, for example. A web page one level below the fetched web page may have a URL cut from the URL of the fetched web page. For example, if the fetched web page has the URL "http://main_page/level1.htm/level2.htm", a web page one level below the fetched web page would have the URL "http://main_page/level1.htm/". Associated web pages are those that are linked to the fetched web page through one or more web pages.

[0043] In light of the present disclosure, it can be appreciated that crawling to web pages actually viewed by consumers and adjacent web pages advantageously limit the contents of an index 654 to useable web pages. In other words, determining web pages viewed by consumers and using those web pages for search purposes, allow the search engine 650 to provide search results containing links to web pages actually employed by consumers. In marked contrast, conventional search engines have crawlers that follow a long chain of links from a particular web page, regardless of whether or not consumers actually use those web pages. While the conventional approach may yield an index containing lots of fetched web pages, those web pages are not necessarily relevant or useful to consumers. In other words, these conventional search engines may generate search results containing many links to not quite useful web pages. Furthermore, the approach taken by conventional search engines requires a relatively large amount of computer resources to perform crawling and indexing functions.

[0044] An indexer 652 may comprise computer-readable program code for parsing, ranking, and indexing web pages

fetched by a crawler 651. An indexer 652 may parse a fetched web page for the occurrence of one or more keywords for indexing purposes. A fetched web page may be identified by its address, which may be used as a link in a search result responsive to a keyword. Each keyword may be associated with a link, referred to as a “keyword-link” combination. A keyword-link combination may be ranked according to the relevance of the link to the keyword. Typical algorithms for ranking keyword-link combinations include taking into account the number of times the keyword appears on the link (i.e., the web page pointed to by the link), proximity of the keyword to other keywords, and use of regular expressions or other text-matching techniques.

[0045] To increase traffic on their websites, webmasters have come up with ways to “defeat” common ranking algorithms and artificially increase the ranking of their web pages. For example, web pages may be filled with non-visible popular texts, trademarked names, and so on to make them appear relevant to a search engine. In response, some conventional search engines also take into account the number of web pages pointing to a particular web page. That is, a keyword-link combination may also be ranked based on the number of other web pages pointing to the link and the credibility (e.g., university web page versus porn web page) of the pointing web pages. While this approach improves the quality of the ranking, it still has at least two major flaws. Firstly, this approach may inaccurately rank highly relevant web pages that are not pointed to by most web pages. Secondly, deeming certain sites as more credible than others is elitist and may or may not be correct. A site that is credible for one subject may not be for others.

[0046] In one embodiment, an indexer 652 takes into account consumer preferences in ranking a keyword-link combination. Consumer preferences include factors relating to consumer interaction with a web page. Consumer preferences may be measured in terms of how long consumers spend on a web page, the number of consumers who made an online purchase from a web page, the number of times a particular consumer viewed a web page within a given period of time, and the number of times consumers clicked on a link on a web page, for example. Consumer preferences may be taken into account along with conventional ranking factors in ranking keyword-link combinations. Weights may be assigned to particular consumer preferences and other ranking factors (which may be conventional) to come up with a score for a keyword-link combination. The scores of keyword-link combinations may be employed to determine their ranking. In responding to a search request for a particular keyword, the search engine 650 may list the responsive keyword-link combinations by rank on a search results web page.

[0047] The ranking of keyword-link combinations based on consumer preferences for search purposes is also disclosed in commonly-assigned U.S. application Ser. No. 10/815,112, filed on Mar. 31, 2004, which is incorporated herein by reference in its entirety.

[0048] FIG. 8 shows a flow diagram of a method 680 that may be performed by an indexer 652, in accordance with an embodiment of the present invention. In step 682, the indexer 652 parses web pages fetched by a crawler 651. A fetched web page may be parsed for the presence of one or more keywords. In step 684, the relevance of a parsed web

page to one or more keywords is ranked according to a ranking process that takes into account consumer preferences as discussed above. In step 686, the indexer 652 stores keyword-link combinations and their ranking information in an index 654.

[0049] In contrast to ranking algorithms employed by conventional search engines, taking into account consumer preferences advantageously allow for ranking that is based on actual consumer interaction with the web page. Taking into account consumer preferences is a democratic, as opposed to elitist, approach that increases ranking accuracy by effectively giving consumers a “vote” on which web pages are relevant for a particular keyword.

[0050] As shown in FIG. 6, a search engine 650 may also include a sifter 653. A sifter 653 may comprise computer-readable program code for sifting through keyword-link combinations and corresponding ranking information stored in an index 654. An index 654 may comprise a searchable, commercially-available database program, for example. Upon receiving a keyword as part of a search request from a consumer, the sifter 653 may consult the index 654 for links having a keyword-link combination with that keyword. The sifter 653 may order the responsive keyword-link combinations according to their ranking, and provide a corresponding search results web page.

[0051] Referring now to FIG. 9, there is shown a flow diagram of a method 690 for responding to a search request, in accordance with an embodiment of the present invention. In step 691, consumer preferences and navigations to various websites are monitored using programs running on client computers. Client-side programs may monitor consumer preferences, such as the number of consumers who received a web page, the number of clicks made on a web page, and so on. Client-side programs may also note the addresses of web pages viewed by the consumer. Data indicative of consumer preferences and navigation may be sent out of a client computer and sent to a message server computer associated with a search engine, for example.

[0052] In step 692, consumer preference and navigation data reported by client-side programs may be employed to build an index for a search engine. In one embodiment, addresses of web pages viewed by consumers are extracted from navigation data reported by client-side programs. One or more crawlers may be employed to fetch those web pages, which may be parsed to build keyword-link combinations. Keyword-link combinations may be ranked using a ranking process that takes into account consumer preferences. Keyword-link combinations and their ranking information may be stored in the index.

[0053] In step 693, a search engine receives a search request for a keyword. The keyword may be received from a consumer performing a search over the Internet. The search engine may be implemented in an Internet portal or another web server.

[0054] In step 694, the search engine may consult the index to find top ranked keyword-link combinations that include that keyword. In response to the search request, links from top ranked keyword-link combinations responsive to the keyword may be included in search results provided to the consumer.

[0055] As can be appreciated, embodiments of the present invention may be employed to provide a search engine

service over the Internet and other computer networks. Embodiments of the present invention may be incorporated in existing search engines to provide more responsive search results.

[0056] While specific embodiments of the present invention have been provided, it is to be understood that these embodiments are for illustration purposes and not limiting. Many additional embodiments will be apparent to persons of ordinary skill in the art reading this disclosure.

What is claimed is:

1. A method of responding to a search request, the method comprising:

using client-side programs to monitor navigation of consumers on the Internet;

building a search engine index using navigation data from the client-side programs;

receiving a search request for a keyword; and

sifting through the index to find links to web pages responsive to the keyword.

2. The method of claim 1 wherein the navigation data include addresses of web pages viewed by consumers, and wherein the index contains links to the web pages viewed by consumers.

3. The method of claim 2 wherein the links to the web pages viewed by consumers in the index are ranked based on consumer preferences.

4. The method of claim 1 further comprising:

using the client-side programs to monitor consumer preferences.

5. The method of claim 4 wherein the consumer preferences include information on the number of consumers who viewed a particular web page.

6. The method of claim 4 wherein the consumer preferences include information on the amount of time consumers spend viewing a particular web page.

7. The method of claim 1 wherein building the search engine index comprises:

determining a web page viewed by a consumer;

fetching the web page viewed by the consumer;

parsing the fetched web page; and

storing the parsed web page in the index.

8. A search engine comprising:

a crawler configured to receive addresses viewed by consumers as observed by client-side programs and to fetch the viewed web pages;

an indexer configured to parse web pages fetched by the crawler and to store links to the fetched web pages in an index; and

a sifter configured to sift through the index to find a stored link relevant to a keyword employed in a search request.

9. The search engine of claim 8 wherein the indexer is further configured to rank parsed web pages based on consumer preferences.

10. The search engine of claim 9 wherein the consumer preferences include information on the number of consumers who viewed a particular web page.

11. The search engine of claim 9 wherein the consumer preferences include information on the amount of time consumers spend viewing a particular web page.

12. The search engine of claim 8 wherein the crawler is further configured to fetch web pages associated with a web page having an address received from one of the client-side programs.

13. The search engine of claim 12 wherein the adjacent web pages include a web page one level above or one level below the web page having an address received from one of the client-side programs.

14. A method of responding to a search request, the method comprising:

observing navigation of consumers to websites to generate navigation data;

using the navigation data to build an index for a search engine;

receiving a keyword as part of a search request; and

responding to the search request by searching the index for links to web pages relevant to the keyword.

15. The method of claim 14 wherein the navigation data includes addresses of web pages viewed by consumers.

16. The method of claim 14 wherein client-side programs are employed to observe the navigation of the consumers.

17. The method of claim 14 wherein the keyword is received over the Internet.

18. The method of claim 14 wherein links relevant to the keyword are ranked in the index according to consumer preferences.

19. The method of claim 14 wherein using the navigation data to build the index comprises:

fetching web pages having addresses included in the navigation data;

parsing the fetched web pages for a keyword;

ranking the relevance of the parsed web page to the keyword; and

storing ranking information and a link to the parsed web page in the index.

20. The method of claim 19 further comprising:

fetching a web page adjacent to another web page having an address included in the navigation data.

* * * * *