(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(72) Inventors: DAI, James; One Microsoft Way, Redmond,
Washington 98052-6399 (US). FARAGO, Julia, H.; One
Microsoft Way, Redmond, Washington 98052-6399 (US).
MENEZES, Natala, J.; One Microsoft Way, Redmond,
Washington 98052-6399 (US). NAAM, Ramaz; One
Microsoft Way, Redmond, Washington 98052-6399 (US).
SATHE, Saleel; One Microsoft Way, Redmond, Wash-
ington 98052-6399 (US). WILLIAMS, Hugh, J.; One
Microsoft Way, Redmond, Washington 98052-6399 (US).

(54) Title: USING USER FEEDBACK TO IMPROVE SEARCH RESULTS
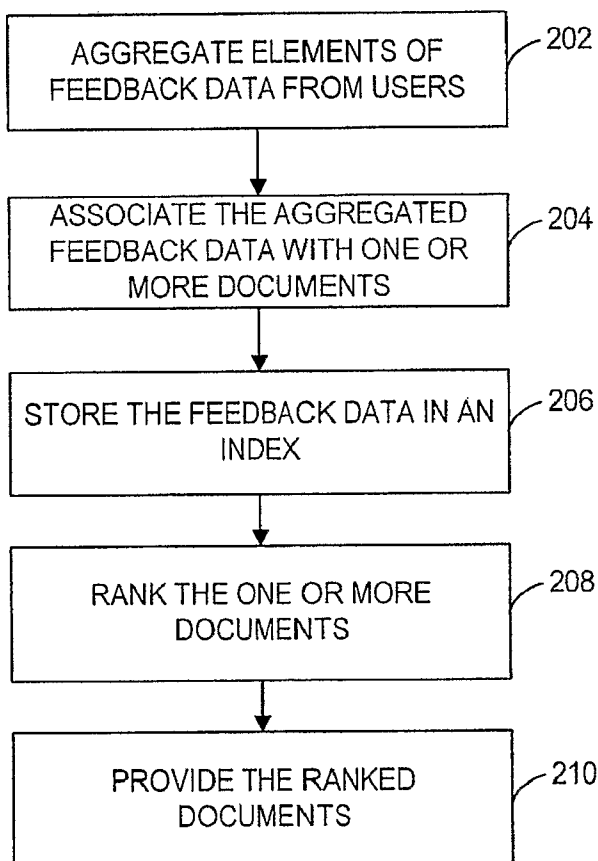
(57) Abstract: The invention discloses a system and method
for managing feedback data that will be used for ranking
search results. The invention can aggregate a plurality of
user feedback data from more than one user into a search
index. The user feedback data can be associated with one or
more documents within the index such that the one or more
documents can be ranked based on the type of feedback data
that is aggregated. Once the documents have been ranked, the
ranked documents can be provided to a requester.

Declarations under Rule 4.17:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*

— *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

— *with international search report*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

USING USER FEEDBACK TO IMPROVE SEARCH RESULTS

BACKGROUND

[0001]   Products and services evolve and improve over time through the comments,

suggestions, and feedback of the general public. The general public can be a good source

of feedback as it is the public that uses these products and services on a regular basis, and

the public may be able to observe most flaws within the product or service that may need to

be corrected.

[0002]   There are many searching services available to the public that can search for a wide-

range of information for a requesting user. Many of these searching services are based on

different algorithms that are used to find the most relevant search results for a user's search

request. However, these algorithms are not always perfect and may not always provide the

best search results for a user's search request. Additionally, a search result set for a

particular search query may be adequate for one person, however, the same search result

for the same search query may not be sufficient for another person because everyone's

needs and likes are not the same. Searching services, therefore, could be another service

that could benefit from being subjected to user comments and feedback. Incorporating user

feedback into the searching services' method of providing search results could lead to

better and more relevant search results as the public's own feedback can be used to offer a

better sense of what the public wants instead of just relying on algorithms to provide

results.

SUMMARY

[0003]   A system and method for managing feedback data to be used for ranking search

results is presented. The system and method discloses aggregating elements of feedback

data from a plurality of users, wherein the feedback data can be in an unstructured or

structured format. Moreover, the system and method discloses associating the feedback

1

data with one or more documents. Additionally, the system and method can rank the one or more documents based from the feedback data. Furthermore, the system and method can provide the ranked one or more documents as search results to a search request.

[0004]   This Summary is provided to introduce a selection of concepts in a simplified form that are further described below in the Detailed Description. This Summary is not intended to identify key features or essential features of the claimed subject matter, nor is it intended to be used as an aid in determining the scope of the claimed subject matter.

BRIEF DESCRIPTION OF THE DRAWINGS

[0005]   Fig. 1 illustrates an embodiment of a system for implementing the invention.

[0006]   Fig. 2 illustrates an embodiment of a method for managing feedback data that can be used for ranking search results.

[0007]   Fig. 3 illustrates an embodiment of a method for providing feedback for one or more documents.

DETAILED DESCRIPTION

[0008]   Fig. 1 illustrates an embodiment of system for implementing the invention. Client 102 can be a desktop or laptop computer, a network-enabled cellular telephone (with or without media capturing/playback capabilities), wireless email client, or other client, machine, device, or combination thereof, to perform various tasks including Web browsing, search, electronic mail (email) and other tasks, applications and functions. Client 102 may additionally be any portable media device such as digital still camera devices, digital video cameras (with or without still image capture functionality), media players such as personal music players and personal video players, and any other portable media device.

[0009]   Server 104 may be or can include a workstation running the Microsoft Windows®, MacOS™, Unix, Linux, Xenix, IBM AIX™, Hewlett-Packard UX™, Novell Netware™,

Sun Microsystems Solaris™, OS/2™, BeOS™, Mach, Apache, OpenStep™ or other

operating system or platform. In an embodiment, client 102 may also be a server.

[0010]   Client 102 can include a communication interface. The communication interface

can be an interface that allows the client to be directly connected to any other client or

device or that allows the client 102 to be connected to a client, server, or device over

network 106. Network 108 can include, for example, a local area network (LAN), a wide

area network (WAN), or the Internet. In an embodiment, the client 102 can be connected

to another client, server, or device via a wireless interface.

[0011]   Fig. 2 illustrates an embodiment of a method for managing feedback data that can

be used for ranking search results. In an embodiment, a search engine can be used as a

server that can manage such feedback data. The search engine can include an index that

can be, for example, an inverted index. In step 202, a search engine can aggregate one or

more elements of feedback data from one or more users. The feedback data can be any

type of information received from a user that is associated with a document. In an

embodiment, the feedback data can be used to rate a document according to the relevance

of a corresponding search query, wherein the search query can comprise at least one search

term. In another embodiment, the feedback data can generally describe the user's feelings

about a particular document. In such embodiments, the feedback data can describe how

pleased or how dissatisfied the user is with the document. A document can be, for

example, a web page, any type of multimedia content, or any other type of content that can

be indexed by a search engine. In an embodiment, a document can be presented to the user

as a search result.

[0012]   In step 204, the aggregated feedback data can be associated with one or more

documents. In an embodiment, the user can transmit feedback data to the search engine in

a structured format. In creating a structured format, the search engine can present an user

interface (UI) to the user in which the search engine can define what type of information it

would like to have from the user. For example, the search engine can create a UI in which it can request that the user provide feedback data regarding a particular document by rating the document on a scale of 1-5 wherein 5 could be the most favorable rating for the document and 1 could be the least favorable rating. Another example could be presenting a UI to the user in which the user can select from any number of buttons that correspond to predefined textual descriptions of a document. The textual description can describe how pleased or how dissatisfied the user is with the document. For example, the UI may contain such terms as "the best possible," "good," "fair," "poor," or "the worst possible" that the user can select as feedback data regarding a particular document. In yet another example, the search engine can present an UI to the user and ask the user provide feedback data by reordering documents within a search result to show which documents the user believes are more relevant to their search request than other documents. The invention can use a client-side scripting language that can offer a conventional means to allow the user to reorder the documents. For example, the user could click and drag icons, images, or links that represent documents above or below other icons, images, or links of other documents to create a reordered list of documents that correlate to an ascending or descending order of relevance.

[0013]   In any type of structured format for receiving feedback data in a pre-defined manner, the search engine can associate the feedback data with the particular document because the search already has knowledge of what type of information it will receive from the user. Additionally, if the feedback data was in response to a document included in a list of search results provided by the search engine, the search engine can associate the search query that corresponds to the search results, the feedback data, and document together with each other.

[0014]   In another embodiment, the user can transmit feedback data to the search engine in an unstructured format. In creating an unstructured format for the user to transmit feedback

4

data, the search engine can present an UI to the user in which the user can freely enter in any type of information that can reveal the user's feelings regarding a particular document. For example, the search engine can present a UI including a text box in which the user can type in whatever comments he/she has as feedback data concerning a particular document.

[0015]   In an embodiment, the feedback data can be directly correlated to a search query as the feedback data is provided in response to search results received from a search engine that the search engine provided after receiving the search query. For example, the feedback data can be linked to a searching experience wherein the user can provide feedback data that can be associated with a web page that the search engine provided as a search result to the user's search query. Since the feedback data is linked to a search query that the search engine provided search results for, the search engine can associate the feedback data with the particular document and the search query.

[0016]   In another embodiment, the feedback data may not be directly correlated with a search query and the search engine may not be able to directly link the feedback data with a document or a search query. In fact, a user may be able to provide feedback data that is not directly related to a search experience. More specifically, a user may be able to provide feedback data by contacting the search engine, a separate server that is connected to the search engine, or by contacting the entity that is managing the search engine. For example, the user can send an email as feedback data to the search engine or separate server that can describe their feelings about a particular search result that it received while conducting a search request, or the feedback data can generally describe the user's feelings about a particular document. In another example, the user can call the entity that is managing the search engine and can verbally provide feedback data that can describe their feelings about a particular search result that it received while conducting a search request, or again the feedback data can generally describe the user's feelings about a particular document. The entity can then convert the speech into text using conventional speech-to-text software. In

5

yet another example, the search engine or separate server can retrieve feedback data that may be found on a blog on the Internet. In each example, the search engine or separate server can parse the text using conventional parsing algorithms to associate the feedback data to a particular document and possibly a search query as well.

[0017]   In step 206, the search engine can store the aggregated feedback data in its index. Again, the index can be, for example, an inverted index. In an embodiment, the feedback data, one or more documents, and search query that were associated together in step 204 can be stored together in the index. In another embodiment, the feedback data can be stored in a separate index from other content that is indexed by the search engine from the network.

[0018]   In step 208, the search engine can rank one or more documents stored in its index based from the feedback data that it received. The feedback data can be one feature of a plurality of features that can be used to provide better search results for similar search queries that were associated with the feedback data. For example, the text indexed from documents found on the Internet can be one feature to determine the ranking of search results for particular search queries. Another feature can be the text received from the user feedback data. Yet another feature can be the text found specifically from the title of a document. Moreover, another feature can be the ratio of the size of a document to the number of times a particular word is found within the document. It should be noted that any type of feature can be used to determine how to rank one or more documents. The above-mentioned features are only a subset of the types of features that can be used to determine how to rank certain documents, therefore, the invention should not be limited to only those examples.

[0019]   When ranking a plurality of documents for a particular search query, the search engine can be configured to emphasize the importance of a word found within the feedback data more than the word being found within the context of any other feature. For example,

6

the search engine can emphasize the feedback data that contains a particular word as being three times more important than a document that contains that same word ten times within itself. In helping itself determine the importance of a word relative to what feature it is found in, the search engine can attach different identifiers with particular words of search queries that can inform the search engine as to where the word came from after it is indexed in the search engine. For example, the search engine may use an identifier to signal that the word came from a title of a document, an identifier to show that the word came from the body of the document, or another identifier to show that it came from feedback data.

[0020]　　　　　In other embodiments, search results, including documents, can be ranked based on a comparison of the context of a word within a particular feature. For example, the a word that is found toward the top of a particular item of feedback data can have more importance than the number of times the word in found within the same element of feedback data. The same example can be applied to a document, wherein the number of times a word is found within the document can have more importance than where the word is found within the document. The way in which the search engine is configured to emphasize certain features can determine how the search engine will subsequently rank and provide search results for similar search queries in view of the received feedback data. Once the one or more documents have been ranked, in step 210, the search engine can provide the ranked documents for similar search queries that correspond to the search queries that have already been associated with particular feedback data.

[0021]　　　　　The feedback data that is stored in the search engine's index may have a life span or a staleness rate. For example, in an embodiment, the invention can be configured to emphasize older feedback data less than feedback data that was received at a time closer to when a search request was received from a user. In another embodiment, the staleness rate of the feedback data can vary according to the crawl rate of the document that the feedback

7

data corresponds to. For example, if a document, a web page for example, has a high crawl

rate, meaning the data found on the web page is indexed by the search engine frequently

(every hour for example), then feedback data that is associated with that web page can

become stale and unimportant at the same rate as the web page is indexed. The staleness of

the feedback data can then become another feature for determining how to rank one or

more documents in view of the feedback data. However, in another embodiment, a

document that has a high crawl rate can be deemed to be more important than other

documents with lower crawl rates. In such an embodiment, any type of feedback data

received that is associated with a document with a high crawl rate can also be deemed

important and can become another feature for determining how to rank one or more

documents in view of the feedback data.

[0022]   Fig. 3 illustrates an embodiment of a method for providing feedback for one or

more documents. In step 302, the user can be presented with a first list of documents of

search results that correspond to search query that the user originally submitted. In step

304, the user can initiate a response to transmit feedback data that is associated with one or

more documents within the search results. As mentioned previously, the user may be

presented with a UI for inputting feedback data for one or more documents. The user can

input the data in a structured or unstructured format as detailed above. There may be many

different methods for initiating a response to transmit feedback data. In an embodiment,

the user can click on an image, icon, or link that is a representation of the document in

order to trigger the UI to be presented to the user to input feedback data. In another

embodiment, the user can hover their mouse cursor over the image, icon, or link that can

trigger a UI to be presented to the user to input feedback data. In yet another embodiment,

upon detection of an event that the will trigger displaying the UI, the UI can be displayed in

a side pane that can be to the positioned to the side of the window that contains the

documents of search results.

[0023]   In an embodiment, the search engine can receive the feedback data in step 306. In

another embodiment, a client-side application can be provided to receive the feedback data

in step 306. In both embodiments, the search engine and the client-side application can

instantly create an item of data based on the feedback data that was received. An item of

data can be a second list of search results that includes one or more different documents or

a reordered version of the first list of documents. Additionally, an item of data can be a

notification that can acknowledge that the search engine or client-side application received

the feedback data. When the item of data is a second list of search results, the client-side

application or search engine can instantly display the second list when it receives the

feedback data from the user at step 308. The client-side application can include a script

language that can provide a conventional means to allow the user to reorder or re-rank the

documents included in the first list of search results. For example, the user could click and

drag icons, images, or links that represent documents above or below other icons, images,

or links of other documents to create a reordered list of documents that correlate to an

ascending or descending order of relevance. Once the user reorders or re-ranks the

documents, the client-side application can instantly display the second list of search results.

Alternatively, the search engine can be used to instantly display a second list of search

results. The search engine can receive the feedback data, complete steps 204-208 as

discussed in Fig. 2, and can display a second list of search results in view of the feedback

data that was received.

[0024]   In an embodiment, a client-side application can allow a user to store preferences on

the client that can be used along with feedback data to allow the search engine to provide

search results that correspond to the user's preferences and feedback data. For example,

the user can customize his/her machine to inform a search engine that he/she prefers to

have news content, sports content, video content, or image content ranked higher than any

other type of content. Additionally, the client side application can inform the search engine

of particular physical properties of the user's machine that can be used along with feedback data to provide relevant search results. For example, the client-side application can inform the search engine that the user's machine is a mobile device. The search engine may have certain documents indexed that are better formatted for mobile devices that the search engine can send to the user as search results in view of the user's feedback data.

[0025]   While particular embodiments of the invention have been illustrated and described in detail herein, it should be understood that various changes and modifications might be made to the invention without departing from the scope and intent of the invention. The embodiments described herein are intended in all respects to be illustrative rather than restrictive. Alternate embodiments will become apparent to those skilled in the art to which the present invention pertains without departing from its scope.

[0026]   From the foregoing it will be seen that this invention is one well adapted to attain all the ends and objects set forth above, together with other advantages, which are obvious and inherent to the system and method. It will be understood that certain features and sub-combinations are of utility and may be employed without reference to other features and sub-combinations. This is contemplated and within the scope of the appended claims.

CLAIMS

We claim:

1.   A method for managing feedback data that will be used for ranking search results, comprising:

aggregating one or more elements of feedback data from a plurality of users;

associating the one or more elements of feedback data with one or more documents;

ranking the one or more documents based from the one or more elements of feedback data; and

providing the ranked one or more documents.

2.   The method according to claim1, further comprising associating a search query with the one or more elements of feedback data and the one or more documents.

3.   The method according to claim 1, wherein the one or more elements of feedback data is received in an unstructured format.

4.   The method according to claim 1, further comprising receiving a search request, wherein feedback data that is received at a time closer to when the search request is received is emphasized more than older feedback data when ranking the one or more documents.

5.   The method according to claim 4, further comprising emphasizing the feedback data based on how often the one or more document is indexed, wherein the feedback data is emphasized more when the feedback data corresponds to a document that is indexed more frequently than other documents.

6.   The method according to claim 1, further comprising ranking the document by identifying the position of one or more words within the feedback data.

7.   The method according to claim 1, wherein the feedback data is stored in an inverted   .

index.   .

8.   The method according to claim 1, wherein the feedback data is not associated with a search request.

9.   The method according to claim 1, wherein the documents are ranked based on the feedback data and a physical characteristic of a requesting device.

10.   A method managing feedback data that will be used for ranking search results, . comprising:

        receiving one or more elements of feedback data, the one or more elements of feedback data being in an unstructured format;

        associating the one or more elements of feedback data with one or more documents;

        ranking the one or more documents based from the one or more elements of feedback data; and

        providing the ranked one or more documents.

11.   The method of claim 10, further comprising associating a search query with the one or more elements of feedback data and the one or more documents.

12.   The method of claim 10, further comprising ranking the document by identifying the position of one or more words within the feedback data.

13.   The method of claim 10, wherein the feedback data is not associated with a search request.

14.   A method for managing feedback data that will be used for ranking search results, comprising:

        receiving a search query;

        ranking a plurality of documents based on the search query and at least one element of feedback data, the feedback data being associated with at least one document;

        providing the plurality of ranked documents.

15.   The method according to claim 14, wherein the one or more elements of feedback data is received in a structured format.

16.   The method according to claim 14, wherein the one or more elements of feedback data is received in an unstructured format.

17.   The method according to claim 14, wherein the feedback data is not associated with a search request.

18.   The method according to claim 14, further comprising ranking the document by identifying the position of one or more words within the feedback data.

19.   The method according to claim 14, further comprising providing a staleness rate for the at least one element of feedback data.

20. The method according to claim 19, wherein the staleness rate is varied according to a crawl rate of at least one of the plurality of documents.

1/3

FIG. 1

```
┌─────────────────────────────┐
│     AGGREGATE ELEMENTS OF    │─── 202
│   FEEDBACK DATA FROM USERS   │
└─────────────────────────────┘
               │
               ▼
┌─────────────────────────────┐
│   ASSOCIATE THE AGGREGATED   │─── 204
│  FEEDBACK DATA WITH ONE OR   │
│        MORE DOCUMENTS        │
└─────────────────────────────┘
               │
               ▼
┌─────────────────────────────┐
│ STORE THE FEEDBACK DATA IN AN│─── 206
│            INDEX             │
└─────────────────────────────┘
               │
               ▼
┌─────────────────────────────┐
│      RANK THE ONE OR MORE    │─── 208
│          DOCUMENTS           │
└─────────────────────────────┘
               │
               ▼
┌─────────────────────────────┐
│      PROVIDE THE RANKED      │─── 210
│          DOCUMENTS           │
└─────────────────────────────┘
```

FIG. 2

```
┌─────────────────────────────┐
│     PRESENT FIRST LIST OF    │  ⟋ 302
│        DOCUMENTS             │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│  INITIATE RESPONSE TO TRANSMIT │  ⟋ 304
│        FEEDBACK DATA         │
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│     RECEIVE FEEDBACK DATA    │  ⟋ 306
└─────────────────────────────┘
              │
              ▼
┌─────────────────────────────┐
│     DISPLAY SECOND LIST OF   │  ⟋ 308
│  DOCUMENTS BASED ON FEEDBACK │
│             DATA             │
└─────────────────────────────┘
```
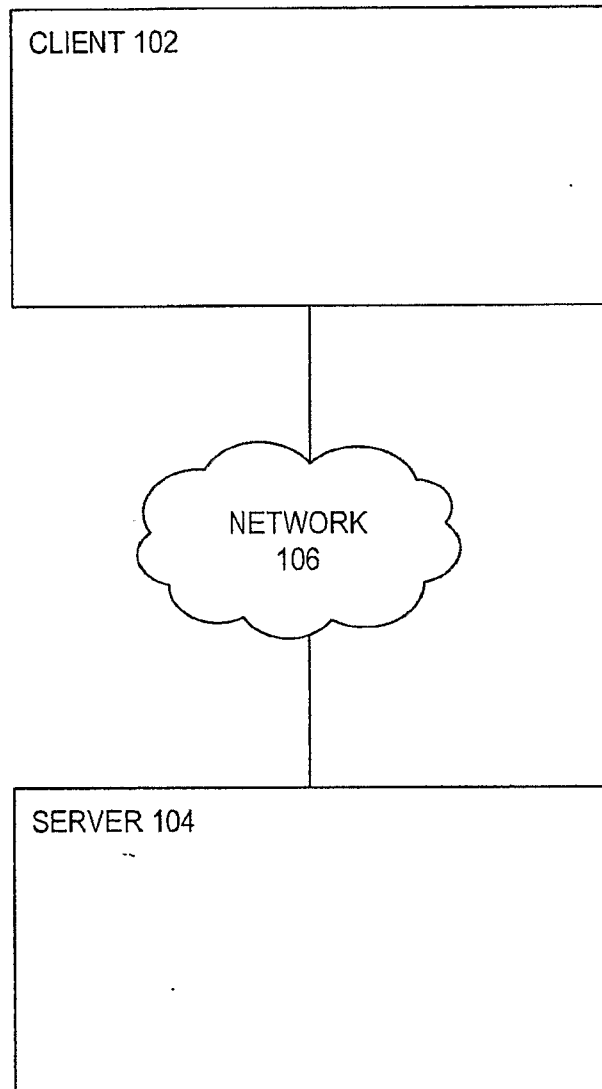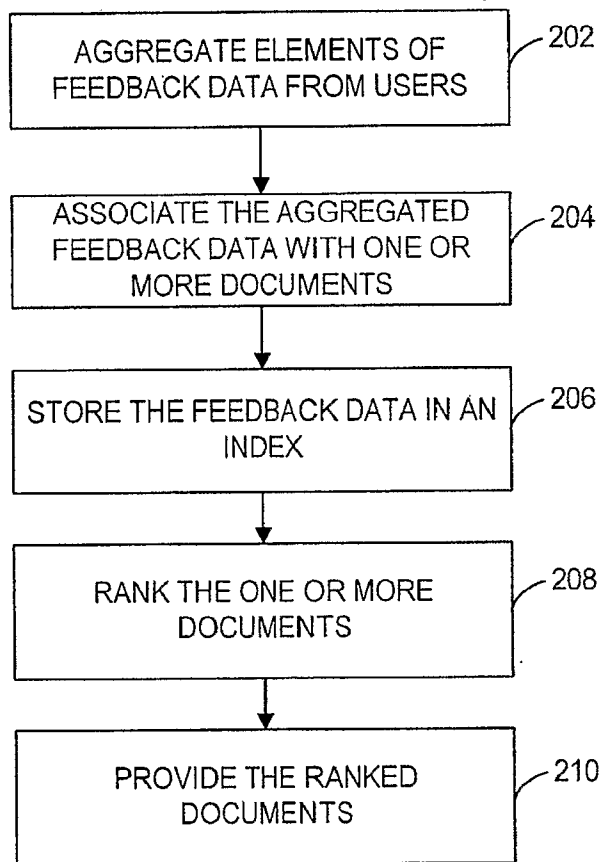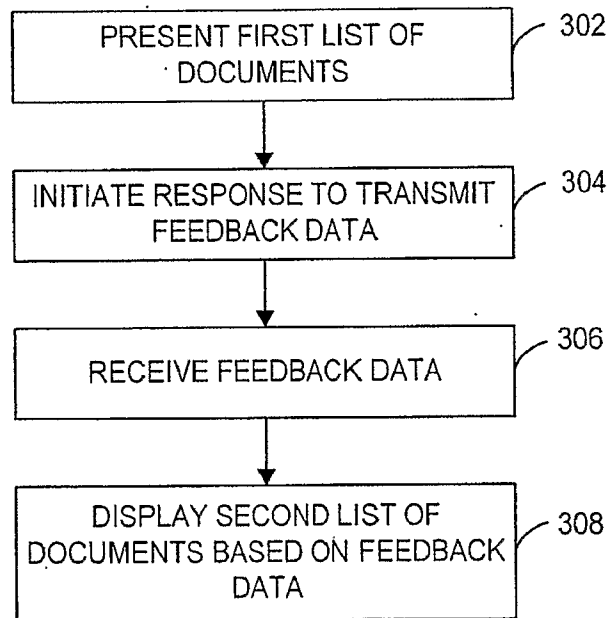
FIG. 3

A.   CLASSIFICATION OF SUBJECT MATTER

G06F 17/30(2006.01)i, G06F 17/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B.   FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 8: G06F G06Q G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
KOREAN PATENTS AND APPLICATIONS FOR INVENTIONS SINCE 1975

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
PAJ, FPD, USPAT, eKIPASS, IEEE, YAHOO, GOOGLE "SEARCH""FEEDBACK"RANK"

C.   DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2003/0033299 A1 (NEELAKANTAN SUNDARESAN) 13 FEBRUARY 2003 (2003-02-13) see the abstract, figures 2-5, paragraph [0032]-[0050]	1-3, 9-11, 13-16, 19, 20
Y	M.M. SUFYAN BEG, "User feedback based enhancement in web search quality", Information Sciences(an international journal), 2004 Elsevier Inc. published see the abstract, figure 1, "update relevance matrix from the user feedback(page 161- page 166)"	1-3, 9-11, 13-16, 19, 20
A	Steve Fox(Microsoft Corporation), "Evaluating Implicit Measures to Improve the Search Experience(presentation file)", Workshop SIGIR 2003(26th annual international ACM SIGIR conference), 1 AUGUST 2003 see the whole slides(especially, slides about "Approach")	1-20
A	US 2002/0120609 A1 (ANDREW K. LANG, DONALD M. KOSAK) 29 AUGUST 2002 (2002-08-29) see the abstract, figures 1-4, paragraph [0047]-[0062]	1-20

☐ Further documents are listed in the continuation of Box C.	☒ See patent family annex.

* Special categories of cited documents:	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A" document defining the general state of the art which is not considered to be of particular relevance	
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents,such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	
"P" document published prior to the international filing date but later than the priority date claimed	"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 JUNE 2007 (27.06.2007)	**27 JUNE 2007 (27.06.2007)**

Name and mailing address of the ISA/KR	Authorized officer
Korean Intellectual Property Office 920 Dunsan-dong, Seo-gu, Daejeon 302-701, Republic of Korea	CHEKAL Hyun
Facsimile No. 82-42-472-7140	Telephone No. 82-42-481-5667

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US20030033299A1	13.02.2003	US07099859 US2003033299A1 US2003033299AA	29.08.2006 13.02.2003 13.02.2003
US20020120609A1	29.08.2002	US06775664 US2002120609A1 US2002120609AA US6775664BB	10.08.2004 29.08.2002 29.08.2002 10.08.2004