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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 January 2007 (25.01.2007)

PCT

(10) International Publication Number
WO 2007/011140 A1

(51) International Patent Classification:
G06F 17/30 (2006.01)

(21) International Application Number:
PCT/KR2006/002787

(22) International Filing Date: 14 July 2006 (14.07.2006)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
10-2005-0064515 15 July 2005 (15.07.2005) KR

(71) Applicant *(for all designated States except US)*: **CHUTNOON INC.** [KR/KR]; 6th Fl. Poonglim Bldg., 823 Yeoksam-dong, Gangnam-gu, Seoul 135-784 (KR).

(72) Inventors; and

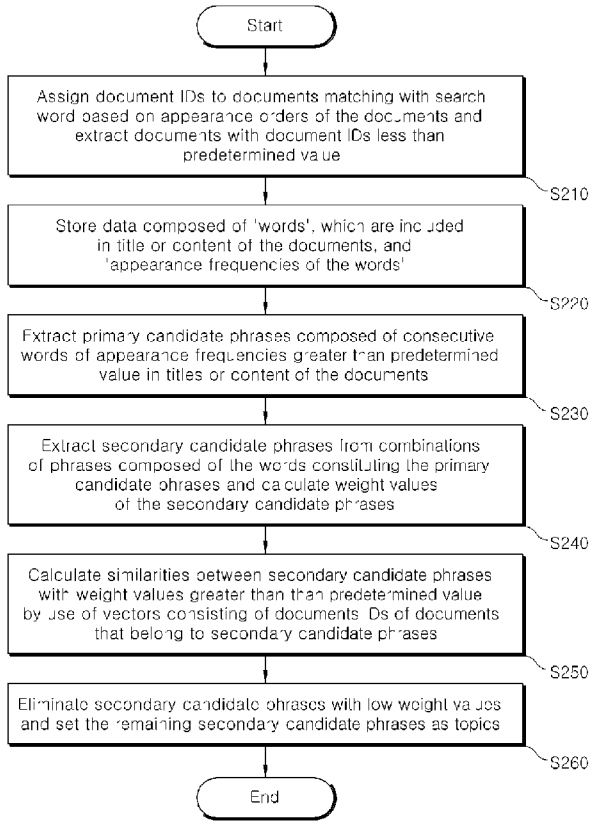
(75) Inventors/Applicants *(for US only)*: **LEE, Eun-young** [KR/KR]; 1411-201 Mokdong Apt., Mok-dong, Yangcheon-gu, Seoul 158-050 (KR). **HAN, Mi-na** [KR/KR]; 1404 Bundang Obelisk, 245-1 Seohyeon-dong, Bundang-gu, Seongnam-si, Gyeonggi-do 463-050 (KR). **PARK, Eui-vin** [KR/KR]; 1411 Hyoseong Olympic County, Bangi-dong, Songpa-gu, Seoul 138-050 (KR). **LEE, Sung-jin** [KR/KR]; 707-606 Jugong Apt., Gaepo-dong, Gangnam-gu, Seoul 135-240 (KR). **SON, Hoon-seok** [KR/KR]; 102-402 Youngpung Apt., 911 Dongcheon-dong, Yongin-si, Gyeonggi-do 449-513 (KR). **SHIN, Joong-ho** [KR/KR]; 110-103 Samsung Apt., Suseo-dong, Gangnam-gu, Seoul 135-521 (KR).

(74) Agents: **LEE, Houn-Su** et al.; 701 BYC Bldg., 648-1 Yeoksam-dong, Gangnam-gu, Seoul 135-080 (KR).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KZ, LA, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG,

[Continued on next page]

(54) Title: METHOD OF EXTRACTING TOPICS AND ISSUES AND METHOD AND APPARATUS FOR PROVIDING SEARCH RESULTS BASED ON TOPICS AND ISSUES



(57) Abstract: Disclosed is a method of displaying search results with respect to a search word, including: (a) referring to words contained in titles or content of search results matching with the search word to calculate similarities between the search results according to a predetermined similarity calculation method, and extracting representative phrases among combinations of words repeatedly contained in similar search results; and (b) displaying the representative phrases and the search results that belong to each of the representative phrases.

WO 2007/011140 A1

NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*

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.

Description

METHOD OF EXTRACTING TOPICS AND ISSUES AND METHOD AND APPARATUS FOR PROVIDING SEARCH RESULTS BASED ON TOPICS AND ISSUES

Technical Field

[1] The present invention relates to an information search technology and, more particularly, to a method and apparatus for extracting topics from search results and providing the search results based on the topics, and a method and apparatus for selecting and providing frequently appearing search results as issues.

Background Art

[2] A conventional search system groups search results into groups based on their types, sequentially provides the search results based on similarities with search words, or places search results that are most similar to the search words at the top of search pages.

[3] However, there is a problem in the conventional search system in that too many redundant search results appear and most of the search results are useless since users tend to view only a few search results appearing at the top of the search pages.

Disclosure of Invention

Technical Solution

[4] The present invention provides a method and apparatus for searching for information based on topics by extracting phrases constituting search results to select topics and outputting the search results topic-by-topic so that users can obtain desired information more easily.

[5] The present invention further provides a method of searching for information based on issues by outputting Internet search results as issues in order of appearance frequencies of the search results.

Advantageous Effects

[6] According to the present invention, users can use search results more efficiently since the users can easily grasp the search results and are not provided with repeated search results.

Brief Description of the Drawings

[7] The above and other features and advantages of the present invention will become more apparent by describing in detail exemplary embodiments thereof with reference to the attached drawings in which:

[8] Fig. 1 is a view for explaining a method of providing search results based on topics

according to an embodiment of the present invention;

[9] Fig. 2 is a flow chart of a method of extracting topics according to an embodiment of the present invention;

[10] Figs. 3 to 9 are views for explaining a method of extracting topics according to an embodiment of the present invention;

[11] Fig. 10 is a flow chart of a method of extracting issues according to an embodiment of the present invention;

[12] Figs. 11 to 13 are views for explaining a method of extracting issues according to an embodiment of the present invention;

[13] Fig. 14 is an issue output result;

[14] Fig. 15 is another issue output result; and

[15] Fig. 16 is a block diagram of an information search apparatus according to an embodiment of the present invention.

Best Mode for Carrying Out the Invention

[16] According to an aspect of the present invention, there is provided a method of displaying search results with respect to a search word, including: (a) referring to words contained in titles or content of search results matching with the search word to calculate similarities between the search results according to a predetermined similarity calculation method, and extracting representative phrases among combinations of words repeatedly contained in similar search results; and (b) displaying the representative phrases and the search results that belong to each of the representative phrases.

[17] According to another aspect of the present invention, there is provided a method of extracting topics, including: (a) assigning document IDs to documents with respect to a search word based on appearance orders of the documents, and extracting documents with document IDs less than a predetermined value; (b) extracting words contained in titles or content of the extracted documents and appearance frequencies of the words; (c) extracting primary candidate phrases composed of words of appearance frequencies greater than a predetermined value appearing consecutively in the titles or content of the documents; (d) generating secondary candidate phrases from combinations of phrases composed of the words constituting the primary candidate phrases, and calculating weight values of the secondary candidate phrases; (e) calculating similarities between secondary candidate phrases with weight values greater than a predetermined value by use of vectors consisting of document IDs of documents belonging to the secondary candidate phrases; and (f) eliminating secondary candidate phrases with low weight values among the secondary candidate phrases with similarities greater than a predetermined value, and setting the remaining secondary

candidate phrases as topics.

[18] According to another aspect of the present invention, there is provided a method of extracting issues, including: (a) extracting the same or similar data the number of which is greater than a predetermined threshold value among stored data; and (b) extracting as issue data a plurality of high-ranking data among the extracted data and displaying the issue data in order of writing time of the issue data or in order of a number of similar documents.

[19] According to another aspect of the present invention, there is provided an apparatus for providing search services based on extracted topics, including: a searching unit searching for stored documents; a primary candidate phrase extracting unit sequentially assigning document IDs to searched documents based on appearance orders of the searched documents, and extracting documents with document IDs less than a pre-determined value; a secondary candidate phrase extracting unit extracting words contained in titles or content of the extracted documents and appearance frequencies of the words, extracting primary candidate phrases composed of words of appearance frequencies greater than a predetermined value appearing consecutively in the titles or content of the documents, generating secondary candidate phrases from combinations of phrases composed of the words constituting the primary candidate phrases, and calculating weight values of the secondary candidate phrases; and a similar candidate phrase eliminating unit calculating similarities between secondary candidate phrases with weight values greater than a predetermined value by use of vectors consisting of document IDs of documents belonging to the secondary candidate phrases, eliminating secondary candidate phrases with low weight values among the secondary candidate phrases with similarities greater than a predetermined value, and setting the remaining secondary candidate phrases as topics.

Mode for the Invention

[20] Exemplary embodiments in accordance with the present invention will now be described in detail with reference to the accompanying drawings.

[21] Fig. 1 is a view for explaining a method of providing search results based on topics according to an embodiment of the present invention.

[22] Referring to Fig. 1, search results are grouped into groups having similar phrases, and topics are extracted from the groups. For instance, it is assumed that 'A=Neowiz', 'B= Separated_search_serive', 'C=Pmang', 'D=Special_force', 'E=Founding_anniversary', and 'F=Party'. When various search results including 'A, B, C, D, E, and F' are output, 'ABE', 'ABF', and 'ABD' may be grouped into a group, 'CDE', 'CDF' and 'CDG' may be grouped into a group, and 'AEFG', 'AEFH', and 'AEFI' may be grouped into a group. In this case, 'AB' becomes a topic 100, 'CD' becomes a

topic and 'AEF' becomes a topic. The term 'topic' implies an expression indicating a subject of search results.

[23] Fig. 2 is a flow chart of a method of extracting topics according to an embodiment of the present invention.

[24] Referring to Fig. 2, when a search word is input, similarities between search results are calculated according to a similarity calculation method by referring to words that are included in titles or content in the search results and match with the search word. Further, representative phrases are extracted among a combination of duplicate words in similar search results, and search results are displayed according to the extracted representative phrases.

[25] In more detail, document IDs are sequentially assigned to documents matching with a search word based on appearance orders of the documents, and documents with documents IDs less than a predetermined value are extracted (operation S210). The predetermined value may vary based on the number of search results, i.e., documents, or the like. Data composed of 'words' which are included in titles or content of the documents, and 'Appearance frequencies of the words' are stored (operation S220). Next, primary candidate phrases composed of words of appearance frequencies greater than a predetermined value in the titles or content of the documents are extracted (operation S230). The predetermined value may vary according to the number of primary candidate phrases to be extracted.

[26] Next, secondary candidate phrases are generated from combinations of phrases composed of the words constituting the primary candidate phrases, and weight values of the secondary candidate phrases are calculated (operation S240). The weight values of secondary candidate phrases are calculated by referring to document IDs included in the secondary candidate phrases, appearance frequencies of words constituting the secondary candidate phrases, and the number of primary candidate phrases used in the secondary candidate phrases. For instance, since a document with a low document ID is important, its weight value becomes high. In addition, if appearance frequency of words constituting the secondary candidate phrase is high, it is regarded as an important document. Further, if document ID included in the secondary candidate phrase is low, it is regarded as an important document.

[27] Next, similarities between secondary candidate phrases with weight values greater than a predetermined value are calculated by use of vectors consisting of document IDs of documents that belong to the secondary candidate phrases (operation S250). That is, when there are several document IDs, the similarities are calculated by referring to the number of the same document IDs. Among secondary candidate phrases having similarities greater than a predetermined value, secondary candidate phrases with low weight values are eliminated and the remaining secondary candidate phrases are

determined as topics (operation S260).

[28] The topics and the documents belonging to individual topics are displayed.

[29] Figs. 3 to 9 are views for explaining a method of extracting topics according to an embodiment of the present invention.

[30] As shown in Fig. 3, when a search word 'Neowiz' is entered, titles 320 appear as search results and document IDs 310 are assigned to the titles based on appearance orders of the titles.

[31] As shown in Fig. 4, a database 330 is obtained from words constituting the titles 320 and appearance frequencies of the words. It can be seen from Fig. 4 that a word 'Neowiz' appears thirteen times and a word 'Yogurting' appears four times in the titles 320. The appearance frequencies of the other words are obtained in this manner. Words of appearance frequencies less than a predetermined value are eliminated. In Fig. 4, a word 'Showdown' appears once and is eliminated.

[32] Next, phrases composed of words of appearance frequencies greater than a predetermined value are extracted from the titles 320 to make primary candidate phrases 340. It can be seen from Fig. 5 that there are six titles each composed of a string of consecutive words 'Neowiz', 'Yogurting', 'RPG', 'Search_corporation', 'Jukeon', Popularized', 'Announces', 'Music', 'Service', and 'Mobile_carrier' among the fourteen titles 320 in Fig. 3.

[33] Next, as shown in Fig. 6, secondary candidate phrases 350 are created with a combination of phrases composed of the words. Appearance frequencies 351 of phrases including the secondary candidate phrases 350 in the primary candidate phrases 340 are extracted. As described in Fig. 2, weight values 352 of the secondary candidate phrases 350 are calculated by referring to document IDs included in the secondary candidate phrases 350, appearance frequencies of words constituting the secondary candidate phrases 350, and the number of primary candidate phrases 340 used in the secondary candidate phrases 350. It can be seen form Fig. 7 that the phrase 'Announces RPG yogurting popularized' has a weight value of 1732, the phrase 'Neowiz Jukeon' has a weight value of 1720, the phrase 'Neowiz search_corporation' has a weight value of 1710, and the phrase 'Neowiz Jukeon mobile_carrier' has a weight value of 1320. The phrase 'Jukeon mobile_carrier music' having a weight value of 1200 is discarded. Thus, a reference weight value to eliminate phrases is 1200.

[34] Referring to Fig. 8, strings 353 of document IDs of documents including the secondary candidate phrases 350 are extracted to calculate similarities between the secondary candidate phrases 350. For instance, it is assumed that documents containing the phrase 'Announces RPG yogurting popularized' are (7, 10), documents containing the phrase 'Neowiz yogurting' are (1, 5, 7, 10), documents containing the phrase 'Neowiz search_corporation' are (2, 4, 12), and documents containing the phrase

'Neowiz search' are (2, 4, 8, 12). In this case, since the similarity between the phrases 'Announces RPG yogurting popularized' and 'Neowiz yogurting' is 66%, the similarity is regarded to be low. Since the similarity between the phrases 'Neowiz search_corporation' and 'Neowiz search' is 82%, the phrase 'Neowiz search' having a lower weight value is eliminated from the secondary candidate phrases. In this manner, topics 361 and search results topic-by-topic are obtained as shown in Fig. 9.

[35] Fig. 10 is a flow chart of a method of extracting issues according to an embodiment of the present invention.

[36] First, data having the same or similar data greater than a predetermined threshold value is extracted from stored data. A plurality of high-ranking data is extracted as issue data from the extracted data. The issue data is displayed in order of writing time of the issue data or in order of a number of similar documents. The stored data may be all of the Internet documents, specific blogs, data on news sites, or data obtained from predetermined search methods.

[37] In more detail, target documents on the Internet or target documents matching with a search word are extracted (operation S410). The extracted documents may be the same or similar to one another. After the number of same or similar documents is calculated, documents having appearance frequencies greater than a predetermined value are extracted (operation S420).

[38] High-ranking documents having a number of the same or similar documents are extracted as issues (operation S430). The extracted issues are output in order of writing time of the documents or the number of same or similar documents (operation S440).

[39] Figs. 11 to 13 are views for explaining a method of extracting issues according to an embodiment of the present invention.

[40] When there are Internet data 510 as shown in FIG. 11, the data 510 are arranged in order of document title 520 and its appearance frequency 521 as shown in FIG. 12. Documents of appearance frequencies less than a predetermined value are eliminated. In this case, documents of appearance frequencies less than two hundreds are eliminated. The remaining documents are selected as issues and output in order of recent writing date as shown in Fig. 13.

[41] Fig. 14 is an issue output result.

[42] Issues may be extracted from the whole target documents on the Internet and displayed as described above. As described in Figs. 2 to 9, topics may be extracted from the target documents and issues may be extracted from the topics and displayed.

[43] Fig. 15 is another issue output result.

[44] Issues and topics may be displayed as shown in Fig. 15. For instance, issues 720 and topics 730 corresponding to a search word 'Neowiz' 710 may be displayed at different positions.

[45] Fig. 16 is a block diagram of an information search apparatus according to an embodiment of the present invention.

[46] The information search apparatus includes a web data storage unit 810, a searching unit 820, a primary candidate phrase extracting unit 830, a secondary candidate phrase extracting unit 840, a similar candidate phrase eliminating unit 850, and a topic output unit 860.

[47] The web data storage unit 810 collects and stores documents on the Internet. The searching unit 820 uses typical search methods to search for the documents. The primary candidate phrase extracting unit 830 sequentially assigns document IDs to the documents in appearance order of the documents, and extracts documents having document IDs less than a predetermined value. A method of extracting the primary candidate phrases is described above in detail with reference to Fig. 2. The secondary candidate phrase extracting unit 840 extracts words contained in titles or content of the documents and appearance frequencies of the words, extracts documents containing words of appearance frequencies greater than a predetermined value in the titles or content as primary candidate phrases, generates secondary candidate phrases composed of combinations of phrases obtained from the words constituting the primary candidate phrases, and calculates weight values of the secondary candidate phrases.

[48] The similar candidate phrase eliminating unit 850 uses vectors consisting of document IDs of documents belonging to secondary candidate phrases with weight values greater than a predetermined value to calculate similarities between the secondary candidate phrases. The similar candidate phrase eliminating unit 850 eliminates secondary candidate phrases with lower weight values among secondary candidate phrases with similarities greater than a predetermined value, and sets the remaining secondary candidate phrases as topics. The topic output unit 860 sets the topics as titles and outputs the topics and documents corresponding to the topics.

[49] The above-mentioned methods of extracting topics and issues may be written with computer programs. Codes and code segments constituting the programs can be easily deduced by computer programmers skilled in the art. In addition, the programs are stored in computer readable media, read and executed by computers, thereby implementing the methods of extracting topics and issues. Examples of the computer readable media include magnetic recording media, optical recording media, and carrier wave media.

[50] While the present invention has been described with reference to exemplary embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the scope of the present invention as defined by the following claims.

Industrial Applicability

[51] The present invention can be efficiently applied to industrial fields related to a method and apparatus for extracting topics from search results and providing the search results based on the topics, and a method and apparatus for selecting and providing frequently appearing search results as issues.

Claims

[1] A method of displaying search results with respect to a search word, comprising:
(a) referring to words contained in titles or content of search results matching with the search word to calculate similarities between the search results according to a predetermined similarity calculation method, and extracting representative phrases among combinations of words repeatedly contained in similar search results; and
(b) displaying the representative phrases and the search results that belong to each of the representative phrases.

[2] The method of claim 1, wherein the operation (a) comprises:
(a1) extracting words contained in titles or content of the search results matching with the search word, and extracting primary candidate phrases in which at least one of the words consecutively appears; and
(a2) generating secondary candidate phrases from words constituting the primary candidate phrases, calculating significance of the secondary candidate phrases based on appearance orders of the search results, appearance frequencies of the words, and the number of primary candidate phrases used in the secondary candidate phrases, and extracting representative phrases by eliminating similar candidate phrases from the secondary candidate phrases of higher significance.

[3] A method of extracting topics, comprising:
(a) assigning document IDs to documents with respect to a search word based on appearance orders of the documents, and extracting documents with document IDs less than a predetermined value;
(b) extracting words contained in titles or content of the extracted documents and appearance frequencies of the words;
(c) extracting primary candidate phrases composed of words of appearance frequencies greater than a predetermined value appearing consecutively in the titles or content of the documents;
(d) generating secondary candidate phrases from combinations of phrases composed of the words constituting the primary candidate phrases, and calculating weight values of the secondary candidate phrases;
(e) calculating similarities between secondary candidate phrases with weight values greater than a predetermined value by use of vectors consisting of document IDs of documents belonging to the secondary candidate phrases; and
(f) eliminating secondary candidate phrases with low weight values among the secondary candidate phrases with similarities greater than a predetermined value, and setting the remaining secondary candidate phrases as topics.

[4] The method of claim of 3, further including (g) displaying the topics as titles and documents that belong to each of the topics.

[5] The method of claim of 3, wherein the operation (d) comprises:
generating secondary candidate phrases from combinations of phrases composed of the words constituting the primary candidate phrases; and
calculating weight values of the secondary candidate phrases based on document IDs contained in the secondary candidate phrases, appearance frequencies of the words constituting the secondary candidate phrases, and the number of the primary candidate phrases used in the secondary candidate phrases.

[6] A method of extracting issues, comprising:
(a) extracting the same or similar data the number of which is greater than a pre-determined threshold value among stored data; and
(b) extracting as issue data a plurality of high-ranking data among the extracted data and displaying the issue data in order of writing time of the issue data or in order of a number of similar documents.

[7] The method of claim 6, wherein the stored data is data obtained by a pre-determined search method.

[8] The method of claim 6, wherein the operation (a) includes determining the same or similar data based on words contained in titles or content of stored data, and extracting the same or similar data the number of which is greater than a pre-determined threshold value.

[9] An apparatus for providing search services based on extracted topics, comprising:
a searching unit searching for stored documents;
a primary candidate phrase extracting unit sequentially assigning document IDs to searched documents based on appearance orders of the searched documents, and extracting documents with document IDs less than a predetermined value;
a secondary candidate phrase extracting unit extracting words contained in titles or content of the extracted documents and appearance frequencies of the words, extracting primary candidate phrases composed of words of appearance frequencies greater than a predetermined value appearing consecutively in the titles or content of the documents, generating secondary candidate phrases from combinations of phrases composed of the words constituting the primary candidate phrases, and calculating weight values of the secondary candidate phrases; and
a similar candidate phrase eliminating unit calculating similarities between secondary candidate phrases with weight values greater than a predetermined value by use of vectors consisting of document IDs of documents belonging to

the secondary candidate phrases, eliminating secondary candidate phrases with low weight values among the secondary candidate phrases with similarities greater than a predetermined value, and setting the remaining secondary candidate phrases as topics.

[10] The apparatus of claim 9, further including a topic output unit displaying the topics as titles and documents that belong to each of the topics.

[11] The apparatus of claim 9, wherein the secondary candidate phrase extracting unit generates secondary candidate phrases from combinations of phrases composed of the words constituting the primary candidate phrases, and calculates weight values of the secondary candidate phrases based on document IDs contained in the secondary candidate phrases, appearance frequencies of the words constituting the secondary candidate phrases, and the number of the primary candidate phrases used in the secondary candidate phrases.

[12] Computer readable media storing programs for executing on a computer the method of claim 1 or 2.

[Fig. 1]

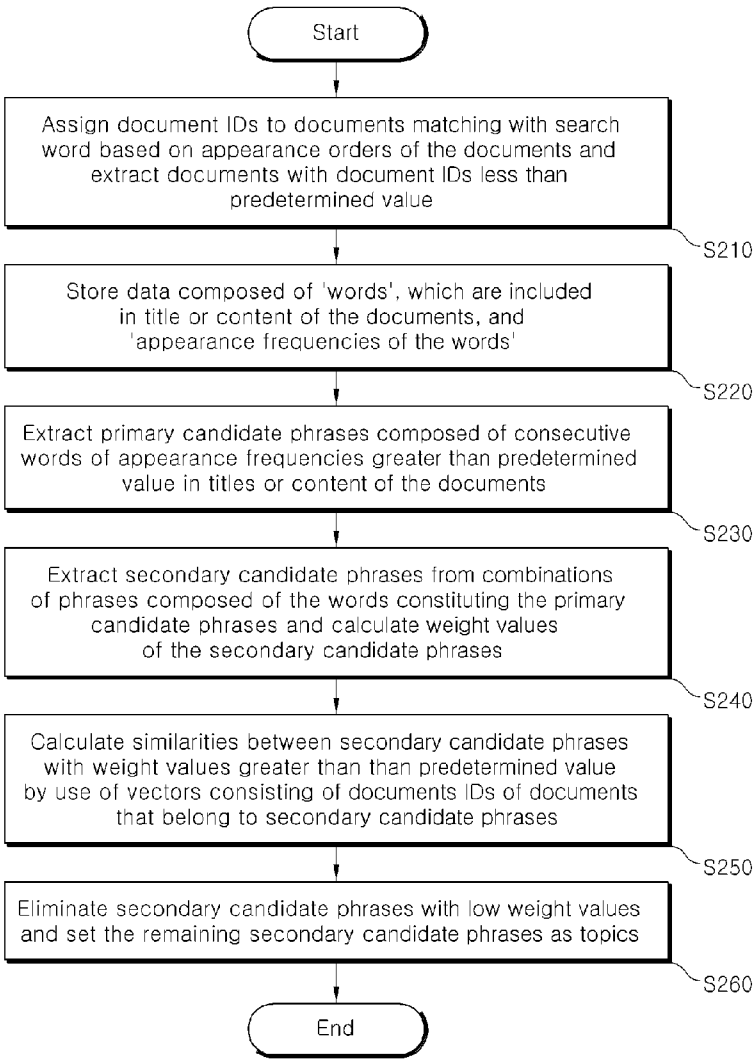
Topic(100)

AB(Neowiz Separate_Search_corporation)
-ABE
-ABF
-ABD

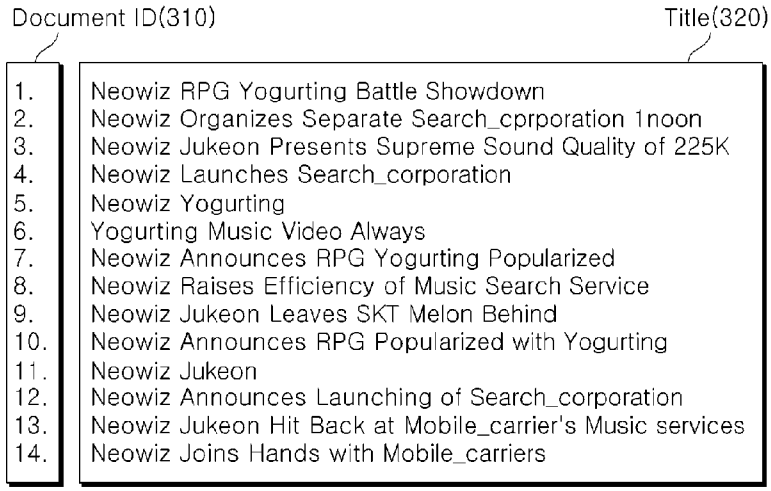
CD(Pmang Special_force)
-CDE
-CDF
-CDG

AEF(Neowiz Founding_anniversary Party)
-AEFG
-AEFH
-AEFI

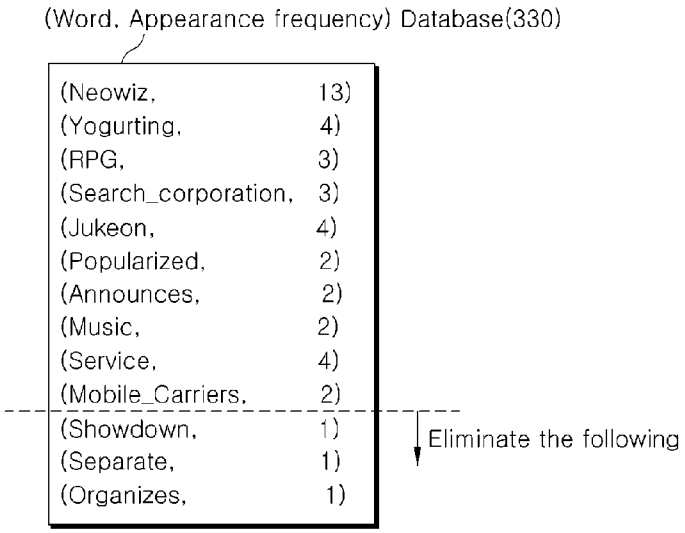
[Fig. 2]



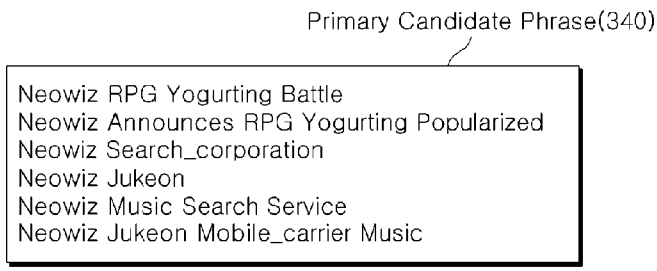
[Fig. 3]



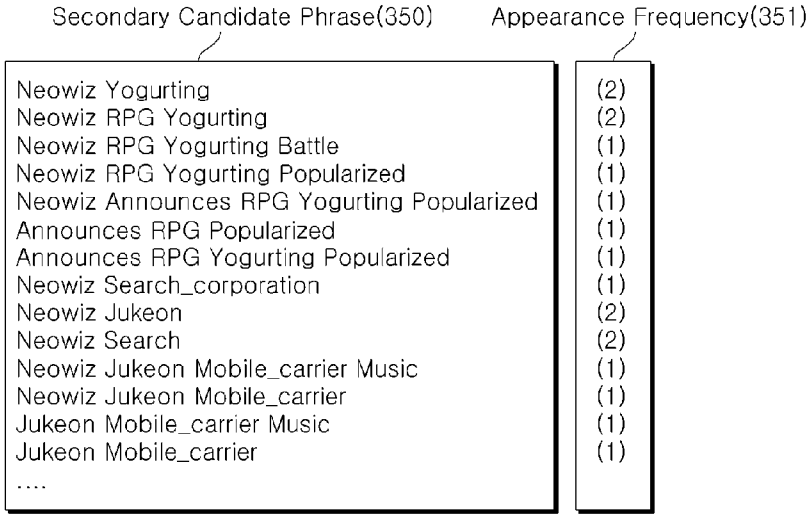
[Fig. 4]



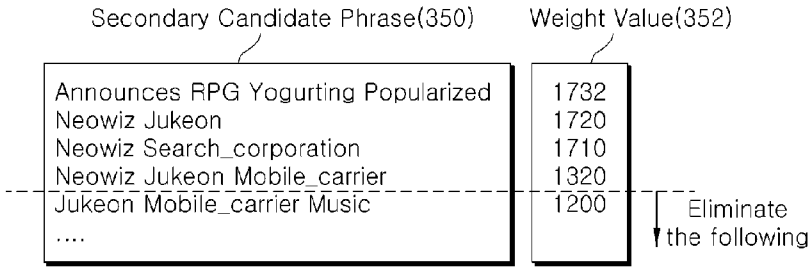
[Fig. 5]



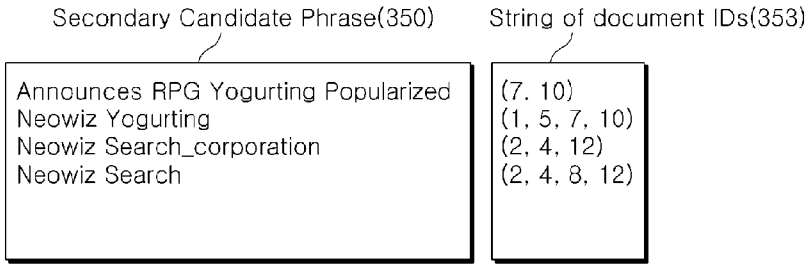
[Fig. 6]



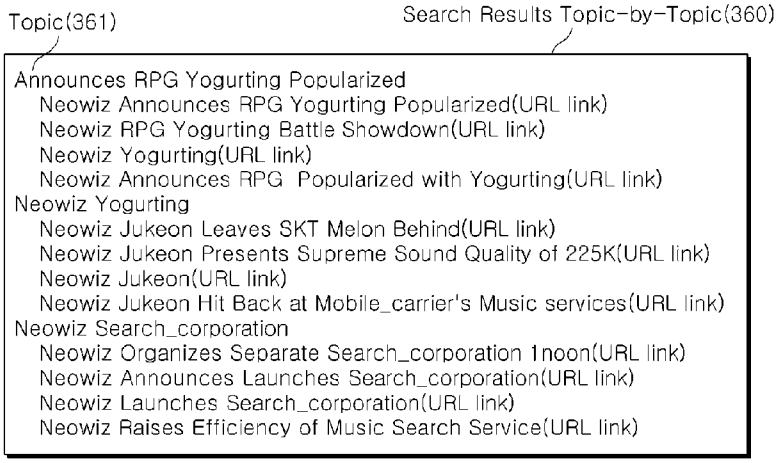
[Fig. 7]



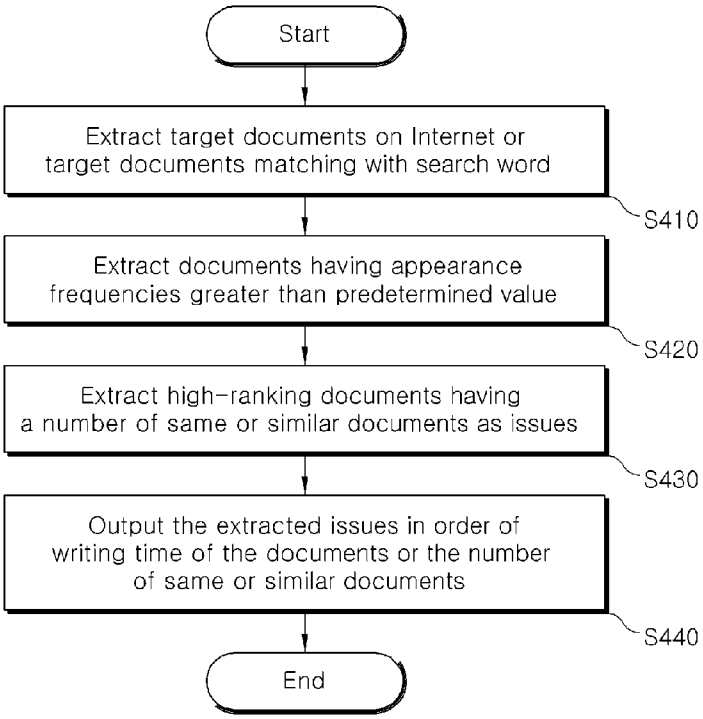
[Fig. 8]



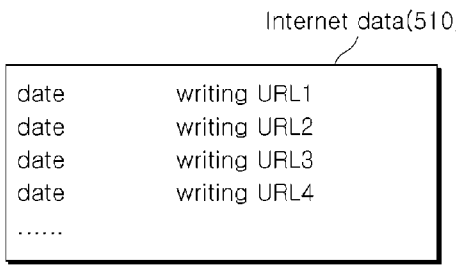
[Fig. 9]



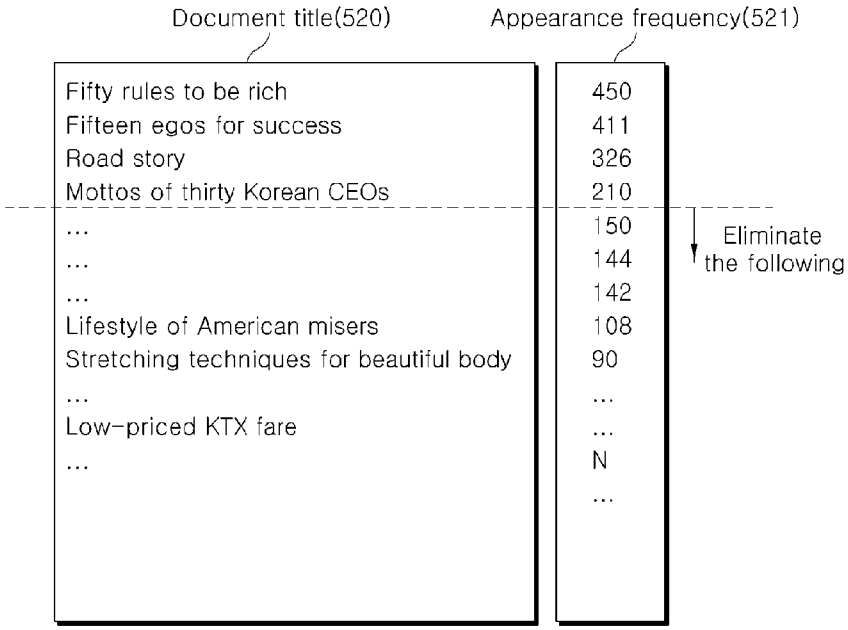
[Fig. 10]



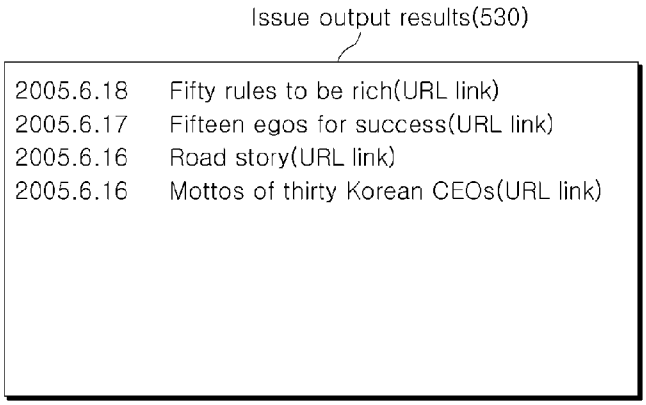
[Fig. 11]



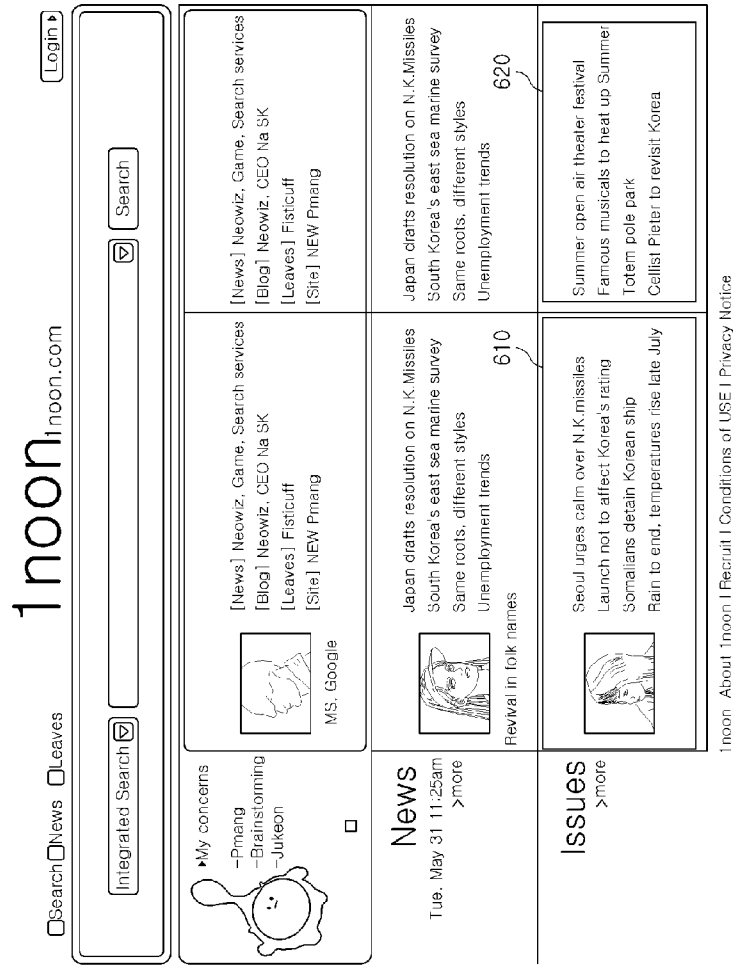
[Fig. 12]


Document title(520)
Appearance frequency(521)
Fifty rules to be rich
Fifteen egos for success
Road story
Mottos of thirty Korean CEOs
...
...
...
Lifestyle of American misers
Stretching techniques for beautiful body
...
Low-priced KTX fare
...
450
411
326
210
150
144
142
108
90
...
...
N
...
Eliminate the following

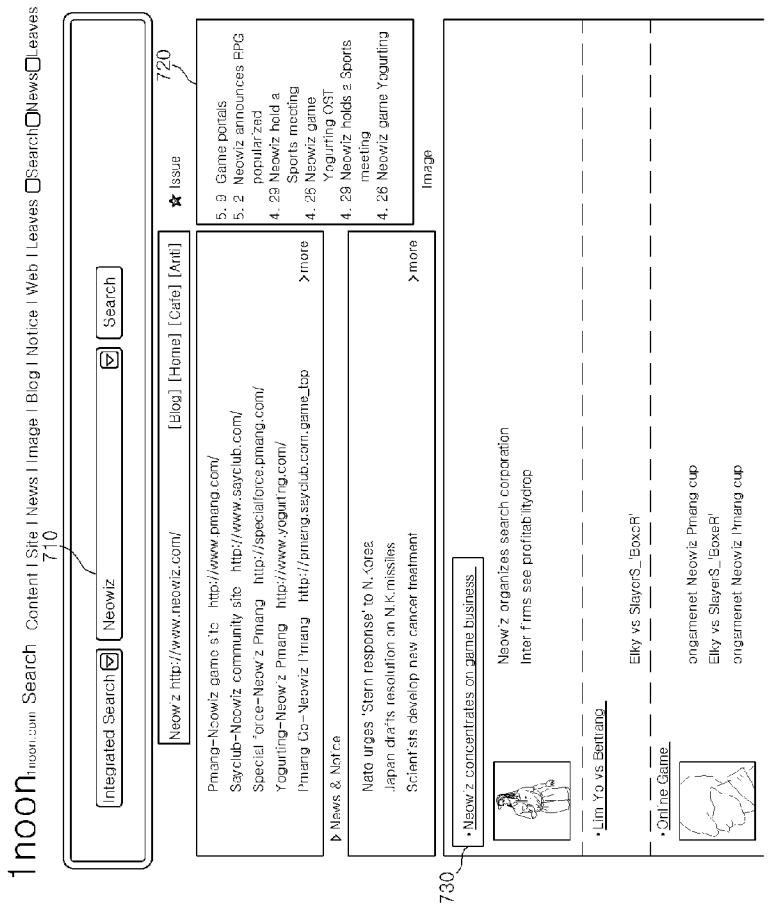
[Fig. 13]


Issue output results(530)
2005.6.18 Fifty rules to be rich(URL link)
2005.6.17 Fifteen egos for success(URL link)
2005.6.16 Road story(URL link)
2005.6.16 Mottos of thirty Korean CEOs(URL link)

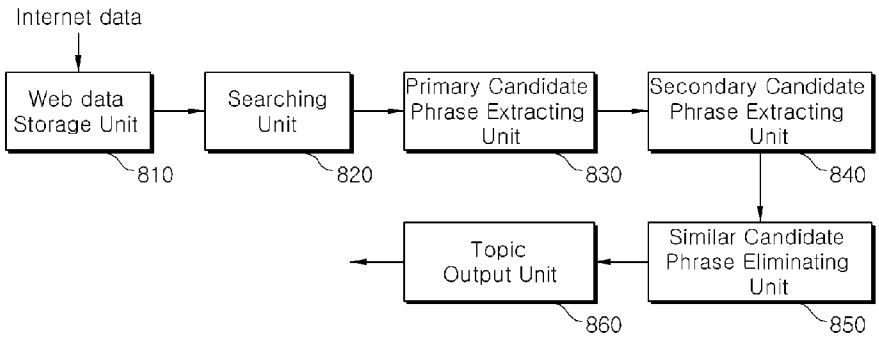
[Fig. 14]


Search News Leaves
1noon 1noon.com
Login
Integrated Search
Search
My concerns
-Pmang
-Brainstorming
-Jukeon
[News] Neowiz, Game, Search services
[Blog] Neowiz, CEO Na SK
[Leaves] Fisticuff
[Site] NEW Pmang
MS, Google
[News] Neowiz, Game, Search services
[Blog] Neowiz, CEO Na SK
[Leaves] Fisticuff
[Site] NEW Pmang
News
Tue. May 31 11:25am
>more
Japan drafts resolution on N.K.Missiles
South Korea's east sea marine survey
Same roots, different styles
Unemployment trends
Revival in folk names
610
Japan drafts resolution on N.K.Missiles
South Korea's east sea marine survey
Same roots, different styles
Unemployment trends
620
Issues
>more
Seoul urges calm over N.K.missiles
Launch not to affect Korea's rating
Somalians detain Korean ship
Rain to end, temperatures rise late July
Summer open air theater festival
Famous musicals to heat up Summer
Totem pole park
Cellist Pieter to revisit Korea
1noon About 1noon I Recruit I Conditions of USE I Privacy Notice

[Fig. 15]


1noon
Search Content I Site I News I Image I Blog I Notice I Web I Leaves Search News Leaves
710
Integrated Search
Neowiz
Search
Neowz http://www.neowiz.com/ [Blog] [Home] [Cafe] [Anti]
Pmang-Neowiz game site http://www.pmang.com/
Sayclub-Neowiz community site http://www.sayclub.com/
Special force-Neowz Pmang http://specialforce.pmang.com/
Yogurting-Neowz Pmang http://www.yogurting.com/
Pmang Co-Neowiz Pmang http://pmang.sayclub.com.game_top
>more
News & Notice
Nato urges 'Stern response' to N.Korea
Japan drafts resolution on N.K.missiles
Scientists develop new cancer treatment
>more
Issue
720
5. 3 Game portals
5. 2 Neowiz announces RPG popularized
4. 29 Neowiz hold a Sports meeting
4. 26 Neowiz game Yogurting OST
4. 29 Neowiz holds a Sports meeting
4. 26 Neowiz game Yogurting
Image
730
• Neowz concentrates on game business
Neowz organizes search corporation
Inter firms see profitability drop
• Lim Yo vs Bertrang
Elky vs SlayerS_'EoxeR'
• Online Game
ongamenet Neowiz Pmang cup
Elky vs SlayerS_'EoxeR'
ongamenet Neowiz Pmang cup

[Fig. 16]


Internet data
Web data Storage Unit
810
Searching Unit
820
Primary Candidate Phrase Extracting Unit
830
Secondary Candidate Phrase Extracting Unit
840
Similar Candidate Phrase Eliminating Unit
850
Topic Output Unit
860

INTERNATIONAL SEARCH REPORT

International application No.
PCT/KR2006/002787

A. CLASSIFICATION OF SUBJECT MATTER

G06F 17/30(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)
IPC8 G06F 17/30

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
KR, JP : IPC as above

Electronic data base consulted during the internationl search (name of data base and, where practicable, search terms used)
eKIPASS(KIPO Internal) "Keyword : extract*, topic, keyword, similarity, appearance, count, order and similar terms"

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5,924,090 A (Marc F. Krellenstein) 13 July 1999 See figure 2; column 2, line 51 - column 3, line 16.	1-12
A	US 6,012,053 A (Sangam Pant et al.) 4 January 2000 See figures 3-5; column 1, line 51 - column 3, line 49.	1-12
A	US 6,212,517 B1 (Mitsuhiro Sato et al.) 3 April 2001 See figures 5, 7; column 1, lines 5-63.	1-12
A	KR 10-2004-0029895 A (CCR INC.) 8 April 2004 See page3; claims 1-6.	1-12
A	KR 10-2000-0050225 A (NIBSOFT CO., LTD.) 5 August 2000 See page3 and abstract.	1-12

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:
"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
"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)
"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
"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
"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
"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
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
30 OCTOBER 2006 (30.10.2006)	**30 OCTOBER 2006 (30.10.2006)**

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	SON, Young Tae
Facsimile No. 82-42-472-7140	Telephone No. 82-42-481-5748

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/KR2006/002787

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US05924090	13.07.1999	AU199872717B2	24.11.1998
		AU736428B2	26.07.2001
		CA2288745AA	05.11.1998
		EP00979470A1	16.02.2000
		EP979470A1	16.02.2000
		JP13522496	13.11.2001
		US5924090A	13.07.1999
		W09849637A1	05.11.1998
US06012053	04.01.2000	None	
US06212517	03.04.2001	CN1198225C	20.04.2005
		DE6983323BC0	06.04.2006
		DE69833238T2	28.09.2006
		EP00889419A2	07.01.1999
		EP00889419A3	22.03.2000
		EP00889419B1	18.01.2006
		EP0889419A2	07.01.1999
		EP889419B1	18.01.2006
		EP889419A3	22.03.2000
		JP11025108A2	29.01.1999
		JP3607462B2	05.01.2005
		KR1019990013736	25.02.1999
KR 10-2004-0029895	08.04.2004	None	
KR 10-2000-0050225	05.08.2000	None.	