US006526440B1

(12) **United States Patent**
Bharat

(10) **Patent No.: US 6,526,440 B1**
(45) **Date of Patent: Feb. 25, 2003**

(54) **RANKING SEARCH RESULTS BY RERANKING THE RESULTS BASED ON LOCAL INTER-CONNECTIVITY**

(75) Inventor: **Krishna Bharat**, Santa Clara, CA (US)

(73) Assignee: **Google, Inc.**, Mountain View, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 185 days.

(21) Appl. No.: **09/771,677**

(22) Filed: **Jan. 30, 2001**

(51) **Int. Cl.**[7] **G06F 13/00**
(52) **U.S. Cl.** **709/219**; 707/3; 707/7
(58) **Field of Search** 709/217, 219, 709/328; 707/3, 10, 7

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,913,208	A *	6/1999	Brown et al.	707/3
6,070,158	A *	5/2000	Kirsch et al.	707/3
6,263,329	B1 *	7/2001	Evans	707/3
6,286,000	B1 *	9/2001	Apte et al.	707/5

OTHER PUBLICATIONS

"Automatic Resource Compilation by Analyzing Hyperlink Structure and Associated Text"; Soumen Chakrabarti et al.; 1998; pp. 1–14.

"The Anatomy of a large–Scale Hypertextual Web Search Engine"; Sergey Brin et al.; print date Aug. 7, 2000; pp. 1–20.

"Hilltop: A Search Engine Based on Expert Documents"; Krishna Bharat; Feb. 2000; pp. 1–12.

"Does "Authority" Mean Quality" Predicting Expert Quality Ratings of Web Documents; Brian Amento et al.; Jul. 2000; pp. 296–303.

* cited by examiner

Primary Examiner—Viet D. Vu
(74) *Attorney, Agent, or Firm*—Harrity & Snyder, L.L.P.

(57) **ABSTRACT**

A search engine for searching a corpus improves the relevancy of the results by refining a standard relevancy score based on the interconnectivity of the initially returned set of documents. The search engine obtains an initial set of relevant documents by matching a user's search terms to an index of a corpus. A re-ranking component in the search engine then refines the initially returned document rankings so that documents that are frequently cited in the initial set of relevant documents are preferred over documents that are less frequently cited within the initial set.

14 Claims, 3 Drawing Sheets

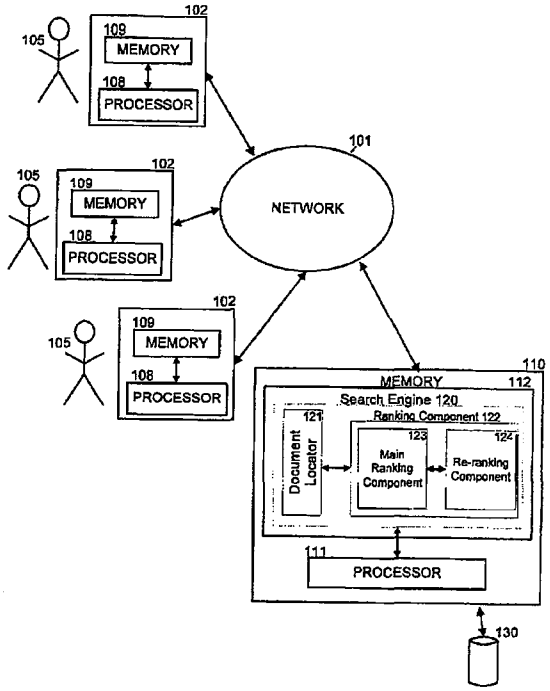

EXHIBIT 2110
Facebook, Inc. et al.
v.

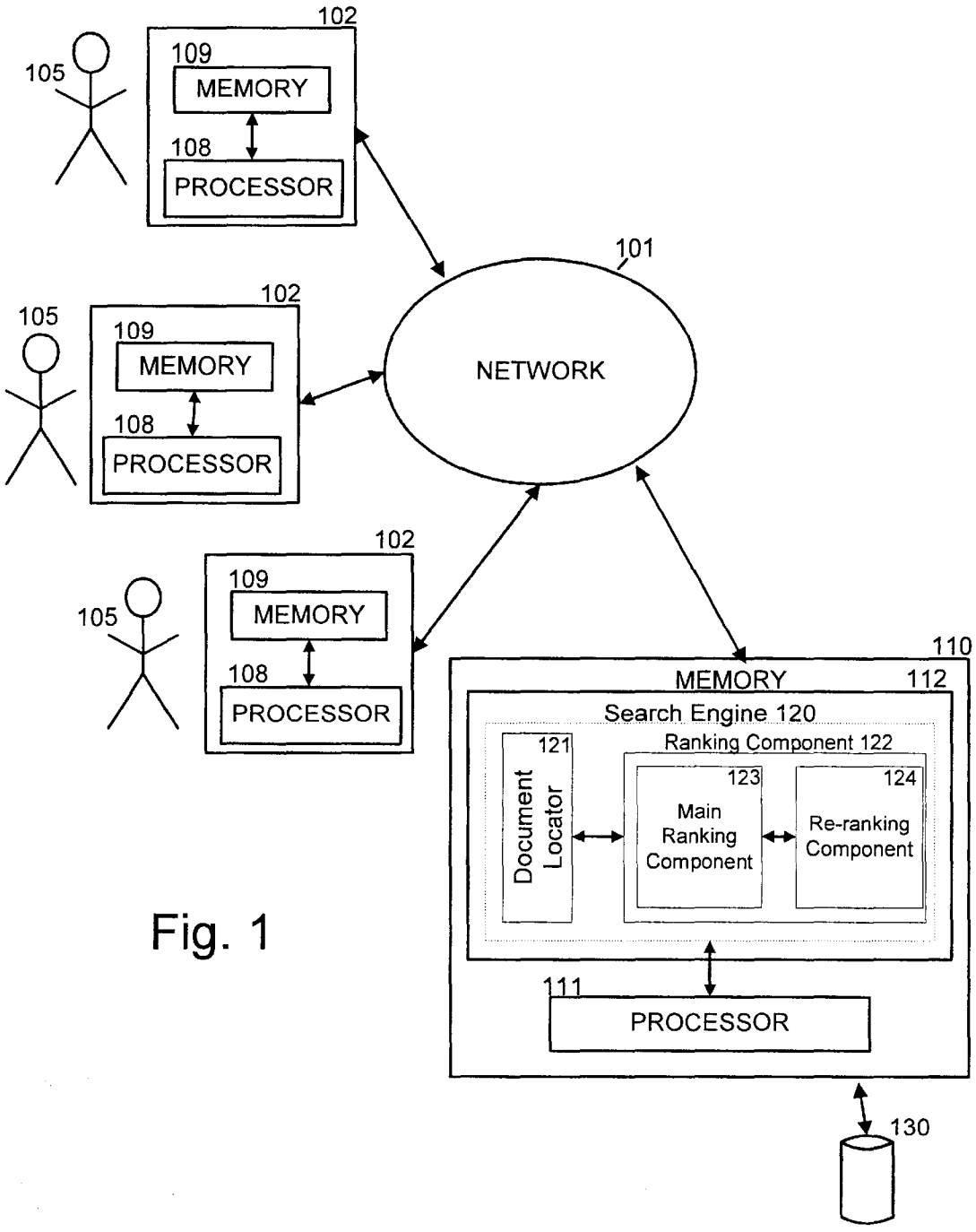

Fig. 1

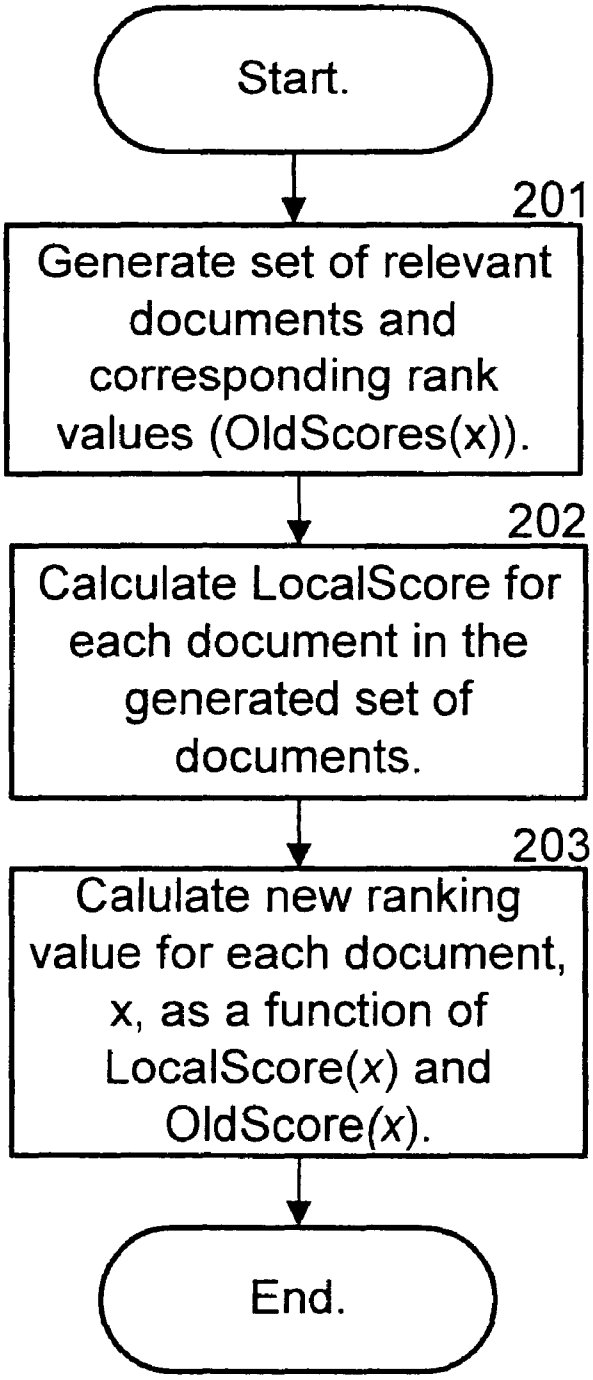

Fig. 2

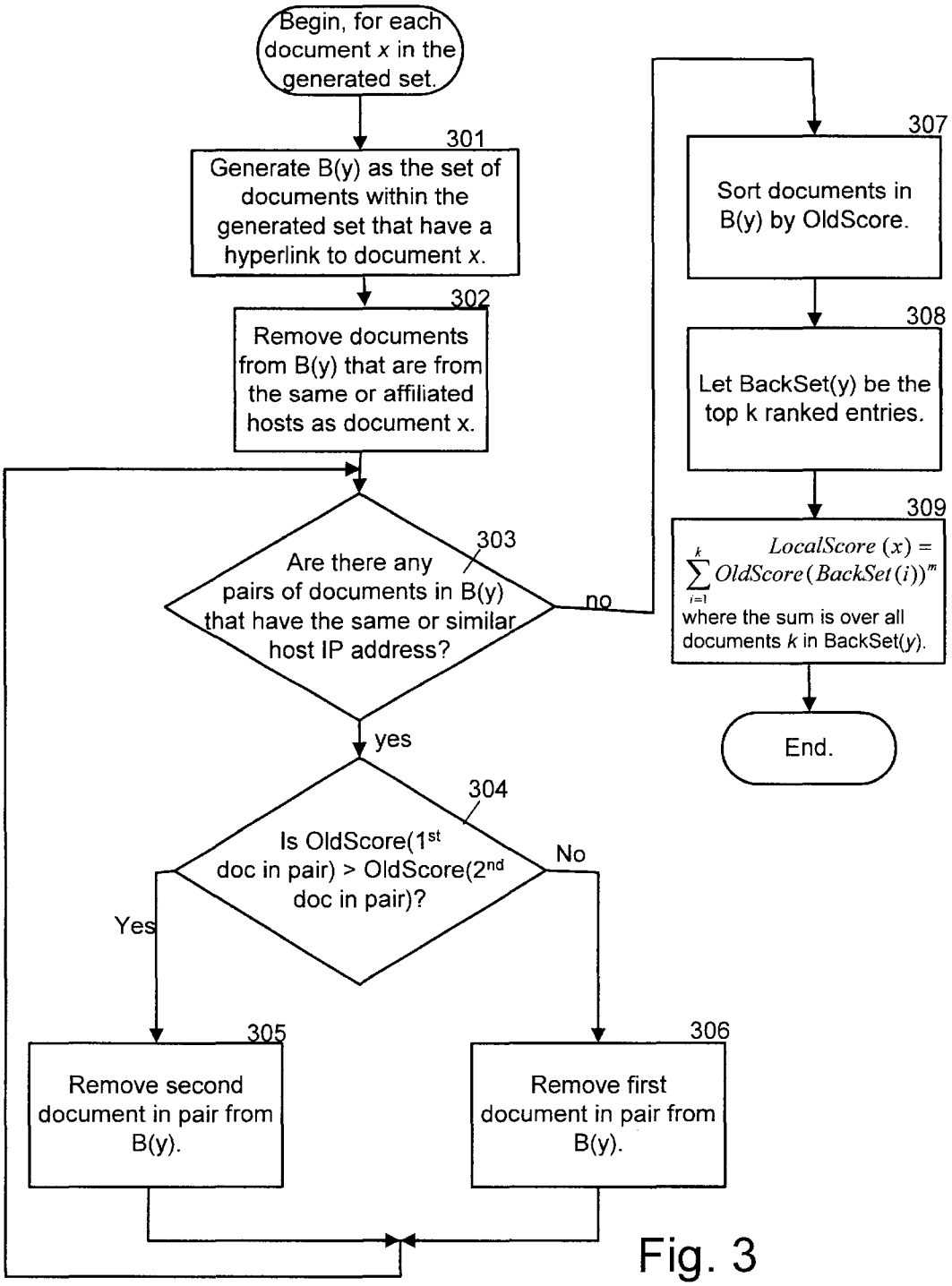

Fig. 3

RANKING SEARCH RESULTS BY RERANKING THE RESULTS BASED ON LOCAL INTER-CONNECTIVITY

BACKGROUND OF THE INVENTION

A. Field of the Invention

The present invention relates generally to the ranking of search results and, more particularly, to search engines that intelligently rank web pages based on a search query.

B. Description of Related Art

The World Wide Web ("web") contains a vast amount of information. Locating a desired portion of the information, however, can be challenging. This problem is compounded because the amount of information on the web and the number of new users inexperienced at web searching are growing rapidly.

Search engines attempt to return hyperlinks to web pages in which a user is interested. Generally, search engines base their determination of the user's interest on search terms (called a search query) entered by the user. The goal of the search engine is to provide links to high quality, relevant results to the user based on the search query. Typically, the search engine accomplishes this by matching the terms in the search query to a corpus of pre-stored web pages. Web pages that contain the user's search terms are "hits" and are returned to the user.

In an attempt to increase the relevancy and quality of the web pages returned to the user, a search engine may attempt to sort the list of hits so that the most relevant and/or highest quality pages are at the top of the list of hits returned to the user. For example, the search engine may assign a rank or score to each hit, where the score is designed to correspond to the relevance or importance of the web page. Determining appropriate scores can be a difficult task. For one thing, the importance of a web page to the user is inherently subjective and depends on the user's interests, knowledge, and attitudes. There is, however, much that can be determined objectively about the relative importance of a web page. Conventional methods of determining relevance are based on the contents of the web page. More advanced techniques determine the importance of a web page based on more than the content of the web page. For example, one known method, described in the article entitled "The Anatomy of a Large-Scale Hypertextual Search Engine," by Sergey Brin and Lawrence Page, assigns a degree of importance to a web page based on the link structure of the web page. In other words, the Brin and Page algorithm attempts to quantify the importance of a web page based on more than just the content of the web page.

The overriding goal of a search engine is to return the most desirable set of links for any particular search query. Thus, it is desirable to improve the ranking algorithm used by search engines and to therefore provide users with better search results.

SUMMARY OF THE INVENTION

Systems and methods consistent with the present invention address this and other needs by providing an improved search engine that refines a document's relevance score based on inter-connectivity of the document within a set of relevant documents.

In one aspect, the present invention is directed to a method of identifying documents relevant to a search query. The method includes generating an initial set of relevant documents from a corpus based on a matching of terms in a search query to the corpus. Further, the method ranks the generated set of documents to obtain a relevance score for each document and calculates a local score value for the documents in the generated set, the local score value quantifying an amount that the documents are referenced by other documents in the generated set of documents. Finally, the method refines the relevance scores for the documents in the generated set based on the local score values.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate an embodiment of the invention and, together with the description, explain the invention. In the drawings,

FIG. **1** is a diagram illustrating an exemplary system in which concepts consistent with the present invention may be implemented;

FIG. **2** is a flow chart illustrating methods consistent with the present invention for ranking documents within a search engine; and

FIG. **3** is a flow chart illustrating, in additional detail, methods consistent with the present invention for ranking documents within a search engine.

DETAILED DESCRIPTION

The following detailed description of the invention refers to the accompanying drawings. The detailed description does not limit the invention. Instead, the scope of the invention is defined by the appended claims and equivalents.

As described herein, a search engine modifies the relevance rankings for a set of documents based on the inter-connectivity of the documents in the set. A document with a high inter-connectivity with other documents in the initial set of relevant documents indicates that the document has "support" in the set, and the document's new ranking will increase. In this manner, the search engine re-ranks the initial set of ranked documents to thereby refine the initial rankings.

FIG. **1** is a diagram illustrating an exemplary system in which concepts consistent with the present invention may be implemented. The system includes multiple client devices **102**, a server device **110**, and a network **101**, which may be, for example, the Internet. Client devices **102** each include a computer-readable medium **109**, such as random access memory, coupled to a processor **108**. Processor **108** executes program instructions stored in memory **109**. Client devices **102** may also include a number of additional external or internal devices, such as, without limitation, a mouse, a CD-ROM, a keyboard, and a display.

Through client devices **102**, users **105** can communicate over network **101** with each other and with other systems and devices coupled to network **101**, such as server device **110**.

Similar to client devices **102**, server device **110** may include a processor **111** coupled to a computer readable memory **112**. Server device **110** may additionally include a secondary storage element, such as database **130**.

Client processors **108** and server processor **111** can be any of a number of well known computer processors, such as processors from Intel Corporation, of Santa Clara, Calif. In general, client device **102** may be any type of computing platform connected to a network and that interacts with application programs, such as a digital assistant or a "smart" cellular telephone or pager. Server **110**, although depicted as

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS
Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS
Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS
Sync your system to PACER to automate legal marketing.