



US006385619B1

(12) **United States Patent**  
**Eichstaedt et al.**

(10) **Patent No.:** **US 6,385,619 B1**  
(45) **Date of Patent:** **May 7, 2002**

(54) **AUTOMATIC USER INTEREST PROFILE GENERATION FROM STRUCTURED DOCUMENT ACCESS INFORMATION**

OTHER PUBLICATIONS

(75) Inventors: **Matthias Eichstaedt; Qi Lu**, both of San Jose, CA (US); **Shang-Hua Teng**, Champaign, IL (US)

Applying evolutionary algorithms to the problem of information filtering—Tjoa, A.M., Sep. 1997.\*

\* cited by examiner

(73) Assignee: **International Business Machines Corporation**, Armonk, NY (US)

*Primary Examiner*—Hosain T. Alam  
*Assistant Examiner*—Anh Ly

(74) *Attorney, Agent, or Firm*—Lumen Intellectual Property Services, Inc.

(\*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(57) **ABSTRACT**

(21) Appl. No.: **09/227,117**

(22) Filed: **Jan. 8, 1999**

(51) **Int. Cl.**<sup>7</sup> ..... **G06F 17/30**

(52) **U.S. Cl.** ..... **707/104; 707/1; 707/2; 707/5; 707/6; 707/10; 707/106**

(58) **Field of Search** ..... **707/1, 5, 6, 9–10, 707/104, 2, 200**

A system generates user interest profiles by monitoring and analyzing a user's access to a variety of hierarchical levels within a set of structured documents, e.g., documents available at a web site. Each information document has parts associated with it and the documents are classified into categories using a known taxonomy. The user interest profiles are automatically generated based on the type of content viewed by the user. The type of content is determined by the text within the parts of the documents viewed and the classifications of the documents viewed. In addition, the profiles also are generated based on other factors including the frequency and currency of visits to documents having a given classification, and/or the hierarchical depth of the levels or parts of the documents viewed. User profiles include an interest category code and an interest score to indicate a level of interest in a particular category. The profiles are updated automatically to accurately reflect the current interests of an individual, as well as past interests. A time-dependent decay factor is applied to the past interests. The system presents to the user documents or references to documents that match the current profile.

(56) **References Cited**

**U.S. PATENT DOCUMENTS**

5,933,827	A	*	8/1999	Cole et al.	707/10
5,943,670	A	*	8/1999	Prager	707/5
6,029,195	A	*	2/2000	Herz	709/219
6,038,560	A	*	3/2000	Wical	707/5
6,038,561	A	*	3/2000	Snyder et al.	707/6
6,066,542	A	*	5/2000	Nielsen et al.	707/104
6,078,866	A	*	6/2000	Buck et al.	707/2
6,108,645	A	*	8/2000	Eichstaedt et al.	707/1
6,154,783	A	*	11/2000	Gilmour et al.	707/6
6,233,575	B1	*	5/2001	Agrawal et al.	707/6

**8 Claims, 2 Drawing Sheets**

