

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
4 March 2004 (04.03.2004)

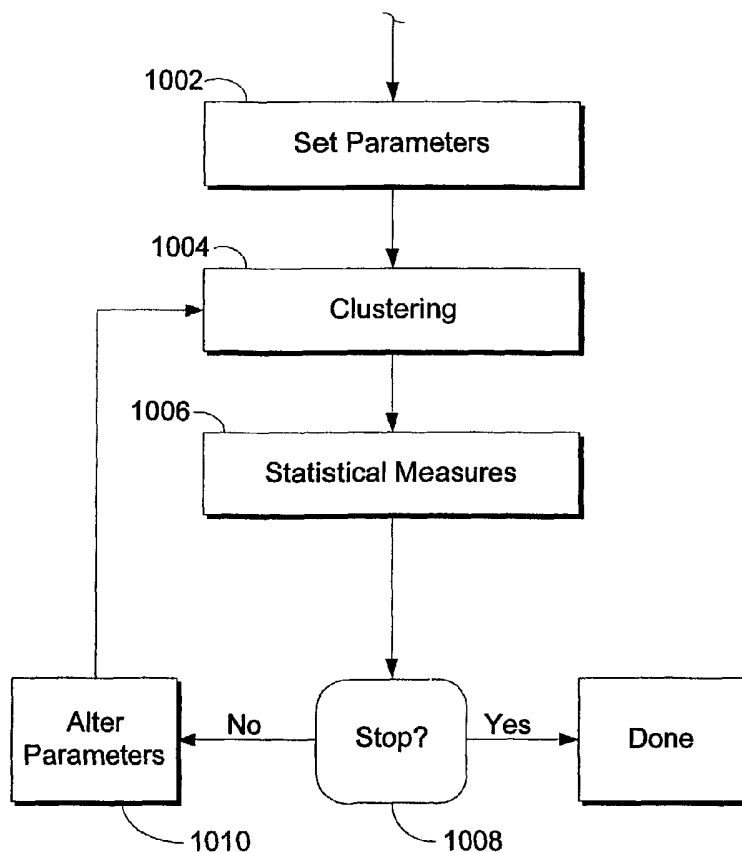
PCT

(10) International Publication Number
WO 2004/019169 A3

- (51) International Patent Classification⁷: G06F 17/30
- (21) International Application Number: PCT/US2003/026239
- (22) International Filing Date: 22 August 2003 (22.08.2003)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data: 10/226,542 22 August 2002 (22.08.2002) US
- (71) Applicant: RIGHTNOW TECHNOLOGIES, INC. [US/US]; 77 Discovery Drive, Bozeman, MT 59718 (US).
- (72) Inventors: WARNER, Douglas, K.; 943 Cougar Drive, Bozeman, MT 59718 (US). RICHTER, James, Neal; 2411 B Wheeler, Bozeman, MT 59718 (US). BANERJEE, Bikramjit; 1320 Lowerline St., Apartment D, New Orleans, LA 70118 (US).
- (74) Agents: KIRCHER, William, B. et al.; Shook, Hardy & Bacon L.L.P., 2555 Grand Blvd., Kansas City, MO 64108-2613 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: METHOD FOR CLUSTERING AUTOMATION AND CLASSIFICATION TECHNIQUES



(57) Abstract: A method and system for use in an information retrieval system that provides related informational items that can be automatically classified and summarized in a hierarchical manner. The system comprises the steps and means for the presentation and analysis of mined data through the application of four distinct processes, namely, feature selection, clustering, classification and summarization (Figure 1). The implementation of each of these techniques is a single process without regard to order. Individual informational items are analyzed, and relevant features extracted (Figure 3). Individual groups of informational items, otherwise known as clusters, are identified, and rules are determined to allow the assignment of informational items to the previously mentioned groups. Descriptive and meaningful titles are then identified for the groups. Various methods of presenting searched information to an end user are also provided. A method of adaptive clustering is also provided.

WO 2004/019169 A3



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.

(88) Date of publication of the international search report:

29 April 2004

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US03/26239

A. CLASSIFICATION OF SUBJECT MATTER		
IPC(7) : G06F 17/30		
US CL : 707/3, 4, 5, 6, 7, 10, 101, 102, 103, 104.1; 706/47, 48, 56, 60, 61		
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)		
U.S. : 707/3, 4, 5, 6, 7, 10, 101, 102, 103, 104.1; 706/47, 48, 56, 60, 61		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
ACM		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	TIAN ZHANG et al. An Efficient Data Clustering Method for Very Large Databases ACM SIGMOD, June 1996 pages 103-112	1-19
X	US 6,263,337 B1 (FAYYAD et al.) 17 June (17.06.01), column 2, lines 45-67; column 3, lines 1-67; column 4, lines 1-67; column 5, lines 1-33.	1-11
Y	US 6,289,354 B1 (AGGRAWAL et al) 11 September 2001 (11.09.2001), column 3, lines 30-67; column 4, lines 1-67; column 6, lines 32-76; column 8, lines 32-67.	1-19
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input type="checkbox"/> 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	"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
"E"	earlier application or patent 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 another 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	"&" document member of the same patent family
"P"	document published prior to the international filing date but later than the priority date claimed	
Date of the actual completion of the international search	Date of mailing of the international search report	
20 February 2004 (20.02.2004)	05 MAR 2004	
Name and mailing address of the ISA/US	Authorized officer	
Mail Stop PCT, Attn: ISA/US	for <i>Michelle R. Gasser</i>	
Commissioner for Patents	Jean Bolte Fleurantin	
P.O. Box 1450	Telephone No. 703-305-3900	
Alexandria, Virginia 22313-1450		
Facsimile No.		