

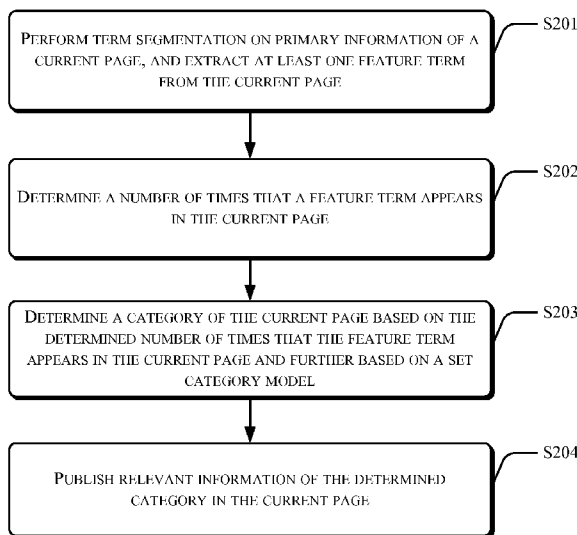


- (51) **International Patent Classification:**
G06F 17/30 (2006.01)
- (21) **International Application Number:**
PCT/US2013/033376
- (22) **International Filing Date:**
21 March 2013 (21.03.2013)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
2012 10078439.7 22 March 2012 (22.03.2012) CN
- (71) **Applicant: ALIBABA GROUP HOLDING LIMITED**
[—/US]; Fourth Floor, One Capital Place, P.O. Box 847, Grand Cayman (KY).
- (72) **Inventors: LIU, Yizhe;** c/o Alibaba Group Legal Department, 10/F, Building A, the West Lake International, Plaza of S&T, No.391 Wen'er Road, Hangzhou, 310013 (CN).
QIU, Guang; c/o Alibaba Group Legal Department, 10/F, Building A, the West Lake International, Plaza of S&T, No.391 Wen'er Road, Hangzhou, 310013 (CN).

- (74) **Agents: NELSON, Brett, L.** et al; Lee & Hayes, PLLC, 601 W. Riverside Ave, Suite 1400, Spokane, WA 99201 (US).
- (81) **Designated States** (*unless otherwise indicated, for every kind of national protection available*): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, 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, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM,

[Continued on nextpage]

(54) **Title:** METHOD AND APPARATUS OF PUBLISHING INFORMATION



(57) **Abstract:** Method of publishing information to solve the problems of low efficiency and accuracy of published information in existing technology. The method segments primary information of a current page, extracts at least one feature term from the current page, determines a number of times that the extracted feature term appears in the current page, determines a category of the current page based on the determined number of times that the feature term appears in the current page and a set category model, and publishes relevant information that belongs to the determined category in the current page. By directly extracting a feature term from a current page and determining a category of the current page based on a number of times that the feature term appears in the current page and a set category model, the exemplary embodiments do not need to perform manual labeling for the current page.

FIG. 2

WO 2013/142732 A3

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US20 13/033376

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - G06F 17/30 (2013.01)
USPC - 707/737

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 17/18, 17/20, 17/21, 17/30 (2013.01)

USPC - 707/736, 737, 738, 739, 739, 748, 750

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

CPC - G06F 17/18, 17/20, 17/21, 17/30, 17/30705, 17/30707, 17/3071, 17/30616 (2013.01)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

PatBase, Orbit, Google Patent, Google Scholar

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2012/001 1475 A1 (HONTZ, JR) 12 January 2012 (12.01.2012) entire document	1-20
Y	CHEN et al. Web Page Classification Based on a Support Vector Machine Using a Weighted Vote Schema. Expert Systems with Applications, Vol. 31, 2006, pages 427-435. [retrieved on 2013-10-22]. Retrieved from the Internet: <URL:http://ccc.inaop.mx/~villasen/bib/web%20page%20categorization.pdf>. entire document	1-20
Y	US 7,725,424 B1 (PONTE) 25 May 2010 (25.05.2010) entire document	2-4, 7, 9-14, 16-18
Y	US 2005/0021324 A1 (BRANTS et al.) 27 January 2005 (27.01.2005) entire document	5-6, 19-20
Y	US 2010/0306229 A1 (TIMM et al.) 02 December 2010 (02.12.2010) entire document	4, 11-13, 18

Further documents are listed in the continuation of Box C.

* 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	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search	Date of mailing of the international search report
22 October 2013	08 NOV 2013

Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-3201	Authorized officer: Blaine R. Copenheaver PCT Helpdesk: 571-272-4300 PCT OSP: 571-272-7774
---	---