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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 September 2010 (02.09.2010)

PCT

(10) International Publication Number
WO 2010/098722 A1

(51) International Patent Classification:
H04L 12/56 (2006.01) *H04L 12/26* (2006.01)

(21) International Application Number:
PCT/SG2009/000068

(22) International Filing Date:
25 February 2009 (25.02.2009)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **NANYANG POLYTECHNIC** [SG/SG]; 180 Ang Mo Kio Avenue 8, Singapore 569830 (SG).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **TEO, Yong King** [SG/SG]; 233 Paya Lebar Road, #16-06, Singapore 409044 (SG). **LO, Siaw Ling** [MY/SG]; Blk 866 Yishun Street 81, #08-57, Singapore 760866 (SG).

(74) Agent: **ALLEN & GLEDHILL LLP**; #28-00, One Marina Boulevard, Singapore 018989 (SG).

(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, BR, BW, BY, BZ, CA, CH, 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, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, 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, HR, HU, IE, IS, IT, LT, LU, LV,

[Continued on next page]

(54) Title: DATA LOSS PREVENTION SYSTEM

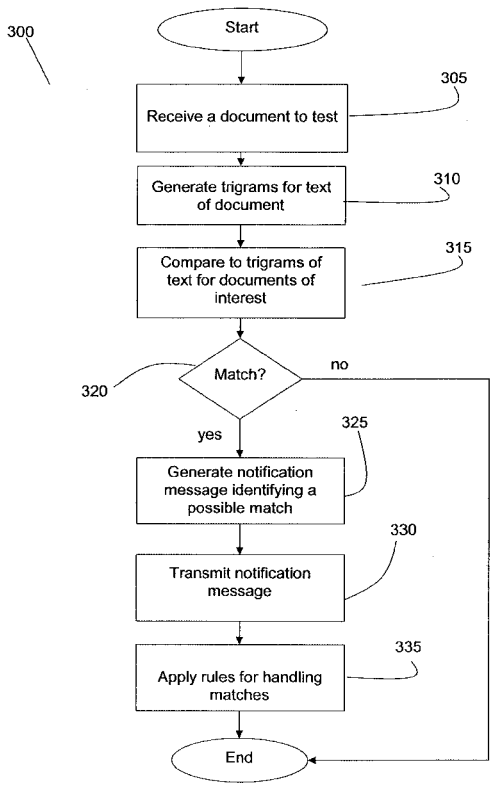


Figure 3

(57) Abstract: A system that detects the possible transmission of information of interest. The system receives a document being transmitted and generates a set of trigrams for the text of the document. The set of trigrams for the document is then compared to sets of trigrams for documents of interest. If the number of matching trigrams between the text of the received document and one of the documents of interest is above a certain threshold value, a possible match is found and indicated by the system.

WO 2010/098722 A1

MC, MK, MT, NL, NO, 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 (Art. 21(3))*

— *with amended claims (Art. 19(1))*