

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2017年9月21日 (21.09.2017)



(10) 国际公布号
WO 2017/157038 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2016/105948
- (22) 国际申请日: 2016年11月15日 (15.11.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201610152630.X 2016年3月17日 (17.03.2016) CN
- (71) 申请人: 华为技术有限公司 (HUAWEI TECHNOLOGIES CO., LTD.) [CN/CN]; 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。
- (72) 发明人: 冷继南 (LENG, Jinan); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 丹尼斯·帕尔霍姆科 (DENIS, Parkhomenko); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 牛进保 (NIU, Jinbao); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 沈建强 (SHEN, Jianqi-

ang); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 王工艺 (WANG, Gongyi); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。 伊万·马祖连科 (IVAN, Mazurenko); 中国广东省深圳市龙岗区坂田华为总部办公楼, Guangdong 518129 (CN)。

- (81) 指定国 (除另有指明, 要求每一种可提供的国家保护): 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, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, 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, SA, 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) 指定国 (除另有指明, 要求每一种可提供的地区保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE,

[见续页]

(54) Title: DATA PROCESSING METHOD, APPARATUS AND EQUIPMENT

(54) 发明名称: 数据处理的方法、装置和设备

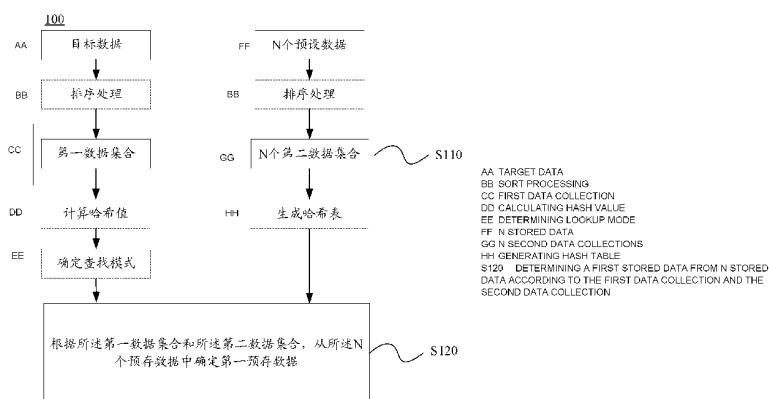


图 1

(57) Abstract: A data processing method, apparatus and equipment, the method comprising: building a first data collection according to a similarity threshold j and target data, the target data including T_1 first bit sets, the first data collection including M_1 first data, the M_1 first data corresponding one-to-one to M_1 combinations of selecting j first bit sets from among the T_1 first bit sets; generating N second data collections according to j and N stored data; the N stored data corresponding one-to-one to the N second data collections, each stored data including T_2 second bit sets, each second data collection including M_2 second data, each second data in the i th second data collection including the T_2 second bit sets in the i th stored data, the M_2 second data in the i th second data collection corresponding one-to-one to the M_2 combinations of selecting j second bit sets from among the T_2 second bit sets. Determining a first stored data from among N stored data according to the first data collection and second data collection may reduce the complexity of the lookup process of similar data.

(57) 摘要:

[见续页]



WO 2017/157038 A1