

(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



(10) 国际公布号
WO 2023/185186 A1

(43) 国际公布日
2023 年 10 月 5 日 (05.10.2023)

(51) 国际专利分类号:
G06F 16/31 (2019.01) *G06F 16/36* (2019.01)

(21) 国际申请号: PCT/CN2023/070483

(22) 国际申请日: 2023 年 1 月 4 日 (04.01.2023)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202210312004.8 2022 年 3 月 28 日 (28.03.2022) CN

(71) 申请人: 支付宝 (杭州) 信息技术有限公司 (ALIPAY (HANGZHOU) INFORMATION TECHNOLOGY CO., LTD.) [CN/CN]; 中国浙江省杭州市西湖区西溪路 556 号 8 层 B 段 801-11, Zhejiang 310000 (CN)。

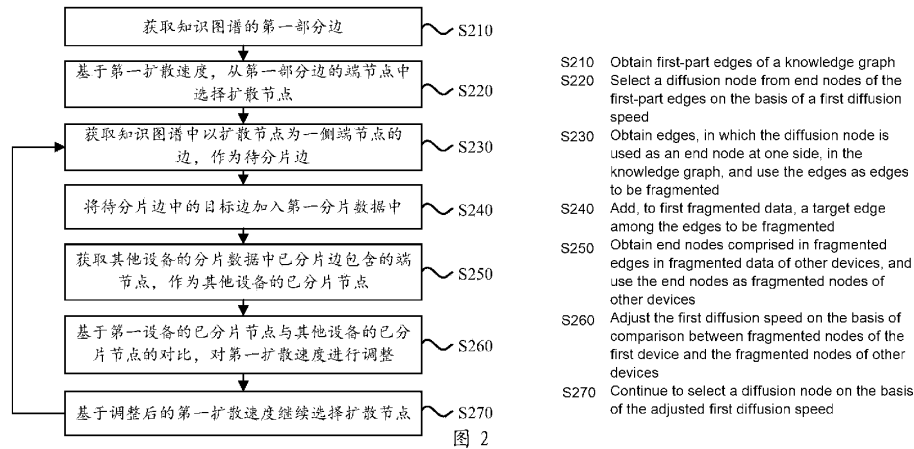
(72) 发明人: 万小培 (WAN, Xiao Pei); 中国浙江省杭州市西湖区西溪路 556 号 8 层 B 段 801-11, Zhejiang 310000 (CN)。

(74) 代理人: 北京亿腾知识产权代理事务所 (普通合伙) (E-TONE INTELLECTUAL PROPERTY FIRM (GENERAL PARTNERSHIP)); 中国北京市海淀区中关村紫金数码园 3 号楼 707, Beijing 100190 (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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE,

(54) Title: METHOD AND APPARATUS FOR PERFORMING DATA FRAGMENTATION ON KNOWLEDGE GRAPH

(54) 发明名称: 一种对知识图谱进行数据分片的方法及装置



(57) Abstract: A method and apparatus for performing data fragmentation on a knowledge graph. The method is used for splitting a knowledge graph into a plurality of pieces of fragmented data respectively belonging to a plurality of devices, and comprises: first, initially splitting a plurality of edges in the knowledge graph, so that the devices respectively obtain some of the edges; any first device selecting, on the basis of a first diffusion speed, a diffusion node from end nodes of first-part edges possessed by the first device, obtaining edges, in which the diffusion node is used as an end node at one side, in the knowledge graph, and using the edges as edges to be fragmented; adding, to first fragmented data of the first device, a target edge among the edges to be fragmented; then, the first device obtaining fragmented nodes in fragmented data of other devices, adjusting the first diffusion speed on the basis of comparison between fragmented nodes of the first device and the fragmented nodes of other devices, continuing to select a diffusion node on the basis of the adjusted first diffusion speed, and executing in a loop the step of obtaining edges, in which the diffusion node is used as an end node at one side, in the knowledge graph.

[见续页]

PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE,
SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ,
UA, UG, US, UZ, VC, VN, WS, 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, CV, GH, GM, KE, LR, LS, MW,
MZ, NA, RW, SC, 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, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL,
NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ,
CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE,
SN, TD, TG)。

本国际公布:

— 包括国际检索报告 (条约第21条(3))。

(57) 摘要: 一种对知识图谱进行数据分片的方法及装置。该方法用于将知识图谱拆分成分别归属于多个设备的多个分片数据。首先, 对知识图谱中的多个边进行初始拆分, 使得各个设备分别得到部分边。任一个第一设备基于第一扩散速度从其拥有的第一部分边的端节点中选择扩散节点, 获取知识图谱中以扩散节点为一侧端节点的边, 作为待分片边; 将待分片边中的目标边加入第一设备的第一分片数据中。接着, 第一设备获取其他设备的分片数据中的已分片节点, 基于第一设备与其他设备的已分片节点的对比, 对第一扩散速度进行调整, 基于调整后的第一扩散速度继续选择扩散节点, 循环执行获取知识图谱中以扩散节点为一侧端节点的边的步骤。