

(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 2021/164317 A1

(43) 国际公布日
2021 年 8 月 26 日 (26.08.2021)

(51) 国际专利分类号:
G06F 16/2458 (2019.01) *G06Q 40/02* (2012.01)

(21) 国际申请号: PCT/CN2020/125898

(22) 国际申请日: 2020 年 11 月 2 日 (02.11.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
202010099547.7 2020 年 2 月 18 日 (18.02.2020) CN

(71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(72) 发明人: 陶冶(TAO, Ye); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong

518057 (CN)。金欢(JIN, Huan); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。金洪波(JIN, Hongbo); 中国广东省深圳市南山区高新区科技中一路腾讯大厦 35 层, Guangdong 518057 (CN)。

(74) 代理人: 北京三高永信知识产权代理有限公司(BEIJING SAN GAO YONG XIN INTELLECTUAL PROPERTY AGENCY CO., LTD.); 中国北京市海淀区学院路蓟门里和景园 A 座 1 单元 102 室, Beijing 100088 (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, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,

(54) Title: SEQUENCE MINING MODEL TRAINING METHOD, SEQUENCE DATA PROCESSING METHOD AND DEVICE

(54) 发明名称: 序列挖掘模型的训练方法、序列数据的处理方法及设备

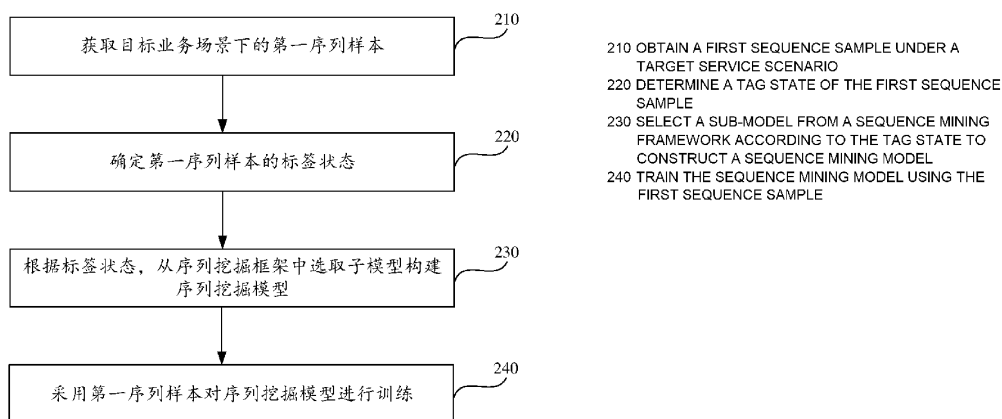


图 2

(57) Abstract: A sequence mining model training method, a sequence data processing method and a device. The method comprises: obtaining a first sequence sample under a target service scenario (210); determining a tag state of the first sequence sample (220); selecting a sub-model from a sequence mining framework according to the tag state to construct a sequence mining model (230); and training the sequence mining model using the first sequence sample (240). According to the method, corresponding sub-models can be selected from the sequence mining framework to construct a sequence mining model according to different tag states of historical sequence data in different service scenarios. Compared with the prior art that only a single model with a fixed structure can be adopted for training, the method makes full use of all historical sequence data in a service scenario, and improves the resource utilization rate.

[见续页]

WO 2021/164317 A1

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,
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, 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, IS, IT, LT, LU, LV, MC, 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: 一种序列挖掘模型的训练方法、序列数据的处理方法及设备。所述方法包括: 获取目标业务场景下的第一序列样本(210); 确定第一序列样本的标签状态(220); 根据标签状态, 从序列挖掘框架中选取子模型构建序列挖掘模型(230); 采用第一序列样本对序列挖掘模型进行训练(240)。该方法针对不同的业务场景下历史序列数据不同的标签状态, 可以从序列挖掘框架中选取对应的子模型构建序列挖掘模型。相比于只能采用结构固定的单一模型进行训练, 该方法充分利用了业务场景中的所有历史序列数据, 提升了资源利用率。