

(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
2020 年 2 月 6 日 (06.02.2020)

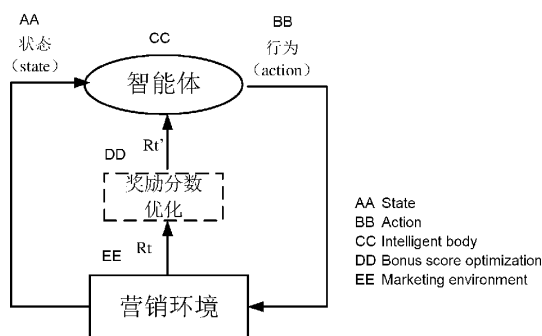


(10) 国际公布号
WO 2020/024717 A1

- (51) 国际专利分类号:
G06Q 30/02 (2012.01) *G06Q 30/06* (2012.01)
- (21) 国际申请号: PCT/CN2019/091713
- (22) 国际申请日: 2019 年 6 月 18 日 (18.06.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201810879908.2 2018 年 8 月 3 日 (03.08.2018) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 何建杉 (HE, Jianshan); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。
- (74) 代理人: 北京博思佳知识产权代理有限公司 (BEIJING BESTIPR INTELLECTUAL PROPERTY LAW CORPORATION); 中国北京市海淀区上地三街 9 号嘉华大厦 B 座 409, Beijing 100085 (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, 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, 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,

(54) **Title:** METHOD AND APPARATUS FOR RECOMMENDATION MARKETING VIA DEEP REINFORCEMENT LEARNING

(54) 发明名称: 通过深度强化学习进行推荐营销的方法及装置



(57) **Abstract:** Provided is a method for recommendation marketing via deep reinforcement learning. A deep reinforcement learning system comprises an intelligent body and an execution environment, wherein the intelligent body is used for determining, according to a marketing policy, a marketing action for state information of the execution environment. The method comprises: acquiring, from an execution environment and with respect to multiple effect targets on a marketing effect link, multiple execution results respectively corresponding to users; according to the acquired multiple execution results, determining a bonus score of reinforcement learning; and then returning the bonus score to the intelligent body for the intelligent body to update the marketing policy thereof.

(57) **摘要:** 本说明书实施例提供一种通过深度强化学习系统进行推荐营销的方法, 其中深度强化学习系统包括智能体和执行环境, 智能体用于根据营销策略, 针对执行环境的状态信息确定营销行为。上述方法包括: 从执行环境中获取, 针对营销效果链路上的多个效果目标, 用户分别对应的多项执行结果, 根据所获取的多项执行结果, 确定强化学习的奖励分数; 然后, 将奖励分数返回给智能体, 用于智能体更新其营销策略。

[见续页]

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))。