

(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 2020/024715 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/CN2019/091649
- (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:
201810879896.3 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,

(54) Title: METHOD AND APPARATUS FOR CARRYING OUT RECOMMENDATION MARKETING BY MEANS OF DEEP REINFORCEMENT LEARNING

(54) 发明名称: 利用深度强化学习进行推荐营销的方法及装置

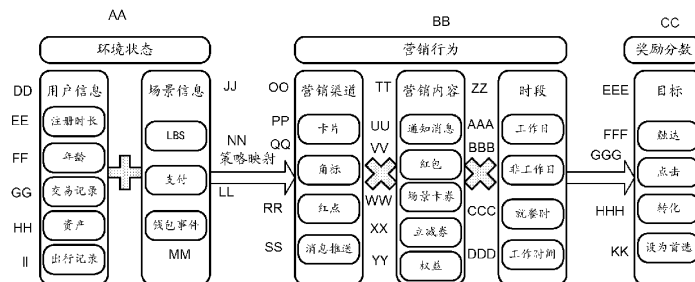


图 5

AA	Environment state	RR	Red dot
BB	Marketing action	SS	Message pushing
CC	Bonus score	TT	Marketing content
DD	User information	UU	Notification message
EE	Registration duration	VV	Red packet
FF	Age	WW	Scene coupon
GG	Transaction record	XX	Direct-discount coupon
HH	Asset	YY	Benefit
II	Travelling record	ZZ	Period
JJ	Scene information	AAA	Workday
KK	Set as preference	BBB	Non-workday
LL	Payment	CCC	Mealtime
MM	Wallet event	DDD	Work time
NN	Policy mapping	EEE	Target
OO	Marketing channel	FFF	Reach
PP	Card	GGG	Click
QQ	Badge	HHH	Transform

(57) Abstract: Provided are a method and apparatus for carrying out recommendation marketing by means of deep reinforcement learning. The method comprises: comprehensively acquiring state information of an execution environment, wherein the state information at least comprises user information of a current user, and may also comprise scene information of a scene that the user is in; then, according to a marketing policy, determining a corresponding marketing action according to the state information, wherein the marketing action comprises a combination of a marketing channel, marketing content and a marketing period; and next, acquiring a

[见续页]

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, 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))。

current bonus score of the execution environment for the marketing action, and updating the marketing policy according to the current bonus score, so as to comprehensively learn a marketing service process.

(57) 摘要: 本说明书实施例提供一种利用深度强化学习进行推荐营销的方法和装置, 方法包括, 全面获取执行环境的状态信息, 其至少包括当前用户的用户信息, 还可以包括用户所处场景的场景信息。然后根据营销策略, 针对上述状态信息确定相应的营销行为, 营销行为包括营销渠道、营销内容和营销时段的组合。接着, 获取执行环境针对营销行为的当前奖励分数, 并根据当前奖励分数更新营销策略, 从而全面学习营销业务过程。