



(12) 发明专利

(10) 授权公告号 CN 117252234 B

(45) 授权公告日 2024.03.01

(21) 申请号 202311531148.3

(22) 申请日 2023.11.16

(65) 同一申请的已公布的文献号
申请公布号 CN 117252234 A

(43) 申请公布日 2023.12.19

(73) 专利权人 之江实验室
地址 311121 浙江省杭州市余杭区中泰街
道科创大道之江实验室

(72) 发明人 金星 王永恒 王超 张文浩
王震 张杨 陈卫 曾洪海
杨亚飞 连建晓 鲁艺 王梦丝

(74) 专利代理机构 北京博思佳知识产权代理有
限公司 11415
专利代理师 孙毅俊

(51) Int. Cl.
G06N 3/006 (2023.01)
G06N 7/01 (2023.01)

(56) 对比文件

- CN 111294242 A, 2020.06.16
- CN 115333956 A, 2022.11.11
- CN 111082983 A, 2020.04.28
- CN 114266361 A, 2022.04.01
- KR 20010105163 A, 2001.11.28
- CN 116228449 A, 2023.06.06
- CN 112801299 A, 2021.05.14
- US 2019147438 A1, 2019.05.16
- US 2008091766 A1, 2008.04.17

李连阳. 基于演化博弈的多智能体系统社会
化行为研究. 中国优秀硕士学位论文全文数据库
(电子期刊). 2019, 全文.

王龙等. 合作演化中的反馈机制. 中国科学
. 2014, 全文. (续)

审查员 杜琳琳

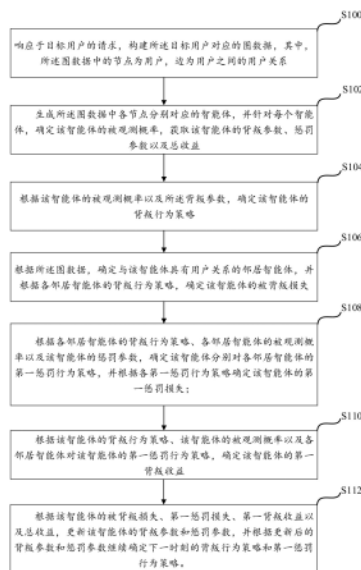
权利要求书3页 说明书16页 附图4页

(54) 发明名称

一种基于非合作博弈的策略生成方法及装
置

(57) 摘要

本说明书公开了一种基于非合作博弈的策略生成方法及装置, 通过构建目标用户对应的图数据, 针对每个智能体, 根据该智能体的被观测概率以及背叛参数, 确定该智能体的背叛行为策略。根据各邻居智能体的背叛行为策略和各邻居智能体的被观测概率, 通过该智能体的惩罚参数, 确定该智能体分别对各邻居智能体的第一惩罚行为策略。进而参考各智能体在背叛行为策略和第一惩罚行为策略的影响下, 更新该智能体的背叛参数和惩罚参数, 用于确定各智能体的行为策略。从上述方法可以看出, 通过构建目标用户对应的图数据, 参考了与目标用户相关用户的行为策略的影响, 以及实时更新目标用户的惩罚参数和背叛参数, 得到的行为策略, 更准确、实效性



CN 117252234 B

[接上页]

(56) 对比文件

Marco Casari等.Decentralized management of common property resources: experiments with a centuries-old institution.Journal of Economic Behavior & Organization.2003,全文.

裴华艳;闫光辉;王焕民.基于行为惩罚的合作演化研究.兰州交通大学学报.2020,(02),全文.

王浩云;张顺颐;赵振东;马燕玲.基于不完全信息博弈的P2P网络节点行为策略模型.应用科学学报.2008,(05),全文.

于建业;王元卓;靳小龙;程学旗.基于社交演化博弈的社交网络用户信息分享行为演化分析.电子学报.2018,(01),全文.

谢逢洁;武小平;崔文田;陈子凤.博弈参与水平对无标度网络上合作行为演化的影响.中国管理科学.2017,(05),全文.

1. 一种基于非合作博弈的策略生成方法,其特征在于,包括:

响应于社交网络中目标用户的请求,获取与所述目标用户相关的其他用户以及所述目标用户的用户数据,并根据获取的各用户数据,构建所述目标用户对应的图数据,其中,所述用户数据至少包括用户的活跃程度、被投诉频率、投诉频率以及点赞量,所述图数据中的节点为用户,边为用户之间的用户关系;

生成所述图数据中各节点分别对应的智能体,并针对每个智能体,根据该智能体对应的用户数据中的活跃程度确定该智能体的被观测概率,以及根据该智能体对应的用户数据中的被投诉频率、投诉频率以及点赞量,确定该智能体的背叛参数、惩罚参数以及总收益,其中,所述被观测概率指用户被与该用户具有用户关系的其他用户观测到的概率,所述背叛参数为不同背叛程度的概率分布,所述惩罚参数表征用户对与该用户具备用户关系的其他用户的惩罚倾向;

根据该智能体的被观测概率以及所述背叛参数,确定该智能体的背叛行为策略,其中,所述背叛行为策略包括用户传播不同质量内容的行为,以及所述背叛参数中的背叛程度和所述背叛行为策略为一一对应关系,所述背叛行为策略取决于背叛参数中背叛程度的概率分布;

根据所述图数据,确定与该智能体具有用户关系的邻居智能体,并根据各邻居智能体的背叛行为策略,确定该智能体的被背叛损失;

根据各邻居智能体的背叛行为策略、各邻居智能体的被观测概率以及该智能体的惩罚参数,确定该智能体分别对各邻居智能体的第一惩罚行为策略,当该智能体为所述目标用户对应的智能体时,将该智能体分别对各邻居智能体的第一惩罚行为策略返回给所述目标用户,用于给所述目标用户针对各邻居智能体对应的其他用户的投诉参考,以及根据各第一惩罚行为策略确定该智能体的第一惩罚损失,其中,所述第一惩罚行为策略包括用户投诉或不投诉其他用户的背叛行为策略的行为,以及用户的惩罚参数越高,用户越易做投诉其他用户的背叛行为策略的行为;

根据该智能体的背叛行为策略、该智能体的被观测概率以及各邻居智能体对该智能体的第一惩罚行为策略,确定该智能体的第一背叛收益;

根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益以及总收益,更新该智能体的背叛参数和惩罚参数,并根据更新后的背叛参数和惩罚参数继续确定下一时刻的背叛行为策略和第一惩罚行为策略。

2. 如权利要求1所述方法,其特征在于,根据该智能体的背叛行为策略、该智能体的被观测概率以及各邻居智能体对该智能体的第一惩罚行为策略,确定该智能体的第一背叛收益,具体包括:

根据该智能体的背叛行为策略,确定该智能体的背叛所得;

针对每个邻居智能体,根据该智能体的被观测概率以及该邻居智能体对该智能体的第一惩罚行为策略,确定该智能体在该邻居智能体的第一惩罚行为策略下的背叛损失;

根据所述背叛所得以及该智能体在各邻居智能体的第一惩罚行为策略下的背叛损失,确定该智能体的第一背叛收益。

3. 如权利要求1所述方法,其特征在于,根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益以及总收益,更新该智能体的背叛参数和惩罚参数,具体包括:

根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益以及总收益,确定该智能体在当前时刻的总收益;

根据各邻居智能体的当前时刻的总收益,确定平均总收益;

判断该智能体的当前时刻的总收益是否大于所述平均总收益;

若是,则根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数。

4.如权利要求3所述方法,其特征在于,根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数,具体包括:

在预设的取值范围内确定随机数,确定所述随机数与所述取值范围最大值的比值,并判断所述比值是否大于预设探索率;

若是,根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数;

若否,重新初始化该智能体的背叛参数和惩罚参数,作为当前时刻该智能体更新后的背叛参数和惩罚参数。

5.如权利要求4所述方法,其特征在于,根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数,具体包括:

根据该智能体的第一背叛收益与预设的背叛阈值的大小关系,确定该智能体对应的背叛参数的更新方向;

根据预设的背叛更新步长以及所述更新方向,更新该智能体的背叛参数;

判断该智能体的第一惩罚损失是否大于预设的惩罚阈值;

若是,则确定该智能体的惩罚参数与预设的惩罚更新步长的和值,作为该智能体更新后的惩罚参数;

若否,则确定该智能体的惩罚参数与预设的惩罚更新步长的差值,作为该智能体更新后的惩罚参数。

6.如权利要求1所述方法,其特征在于,根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益以及总收益,更新该智能体的背叛参数和惩罚参数之前,所述方法还包括:

针对该智能体的每个邻居智能体,根据所述图数据,确定该智能体与该邻居智能体均具有用户关系的共同邻居智能体;

针对每个共同邻居智能体,根据该智能体对该邻居智能体的第一惩罚行为策略以及该智能体的被观测概率,确定该共同邻居智能体对该智能体的第二惩罚行为策略;

根据该共同邻居智能体的第二惩罚行为策略以及该智能体对该邻居智能体的第一惩罚行为策略,确定该智能体的第二背叛收益;

根据该智能体对各邻居智能体的第二惩罚行为策略,确定该智能体的第二惩罚损失。

7.如权利要求6所述方法,其特征在于,更新该智能体的背叛参数和惩罚参数,具体包括:

根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益、第二背叛收益、第二惩罚损失以及总收益,确定该智能体在当前时刻的总收益;

根据各邻居智能体的当前时刻的总收益,确定平均总收益;

判断该智能体的当前时刻的总收益是否大于所述平均总收益;

若是,则根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失、第二惩罚损失和第二背叛收益,更新该智能体的惩罚参数。

8.如权利要求7所述方法,其特征在于,根据该智能体的第一惩罚损失、第二惩罚损失和第二背叛收益,更新该智能体的惩罚参数,具体包括:

将第一惩罚损失与第二惩罚损失之和,作为该智能体的总惩罚损失;

判断所述总惩罚损失是否大于该智能体的第二背叛收益;

若是,则确定该智能体的惩罚参数与预设的惩罚更新步长的和值,作为该智能体更新后的惩罚参数;

若否,则确定该智能体的惩罚参数与预设的惩罚更新步长的差值,作为该智能体更新后的惩罚参数。

9.一种计算机可读存储介质,其特征在于,所述存储介质存储有计算机程序,所述计算机程序被处理器执行时实现上述权利要求1~8任一项所述的方法。

10.一种电子设备,其特征在于,包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,其特征在于,所述处理器执行所述程序时实现上述权利要求1~8任一项所述的方法。

一种基于非合作博弈的策略生成方法及装置

技术领域

[0001] 本说明书涉及网络信息数据处理技术领域,尤其涉及一种基于非合作博弈的策略生成方法及装置。

背景技术

[0002] 随着互联网技术的发展,个体间的信息传播越发频繁导致社交网络呈现的问题大幅增长。一部分用户选择传播不真实或低质量内容的“搭便车”行为,以提高自身的曝光度和关注,同时,这种“搭便车”行为损害了其他用户的用户体验,所以,减少“搭便车”行为以提高社交网络中各用户的用户体验是至关重要的。

[0003] 在现有技术中,为用户在社交网络中提供行为策略的方案较为单一,一般是通过策略复制的方式确定用户的策略。即,先从各用户中确定高收益用户的行为策略,然后直接作为推荐给用户的策略。

[0004] 但是,这种方法存在较大局限性。首先,直接复用行为策略的方式,实际上并没有“生成”新的行为策略,盲目跟风高收益用户的策略。导致为用户推荐的行为策略单一,降低了探索其他行为策略的动机。其次,由于社交网络中用户个体的差异(投诉倾向、所处的社交网络范围等不同),导致高收益用户的行为策略可能不适合推荐给其他用户,导致推荐给用户的策略的效果不佳。

[0005] 基于此,本说明书提供一种基于非合作博弈的策略生成方法及装置。

发明内容

[0006] 本说明书提供一种基于非合作博弈的策略生成方法及装置,以部分的解决现有技术存在的上述问题。

[0007] 本说明书采用下述方案:

[0008] 本说明书提供一种基于非合作博弈的策略生成方法,所述方法包括:

[0009] 响应于目标用户的请求,构建所述目标用户对应的图数据,其中,所述图数据中的节点为用户,边为用户之间的用户关系;

[0010] 生成所述图数据中各节点分别对应的智能体,并针对每个智能体,确定该智能体的被观测概率,获取该智能体的背叛参数、惩罚参数以及总收益;

[0011] 根据该智能体的被观测概率以及所述背叛参数,确定该智能体的背叛行为策略;

[0012] 根据所述图数据,确定与该智能体具有用户关系的邻居智能体,并根据各邻居智能体的背叛行为策略,确定该智能体的被背叛损失;

[0013] 根据各邻居智能体的背叛行为策略、各邻居智能体的被观测概率以及该智能体的惩罚参数,确定该智能体分别对各邻居智能体的第一惩罚行为策略,并根据各第一惩罚行为策略确定该智能体的第一惩罚损失;

[0014] 根据该智能体的背叛行为策略、该智能体的被观测概率以及各邻居智能体对该智能体的第一惩罚行为策略,确定该智能体的第一背叛收益;

[0015] 根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益以及总收益,更新该智能体的背叛参数和惩罚参数,并根据更新后的背叛参数和惩罚参数继续确定下一时刻的背叛行为策略和第一惩罚行为策略。

[0016] 可选地,根据该智能体的背叛行为策略、该智能体的被观测概率以及各邻居智能体对该智能体的第一惩罚行为策略,确定该智能体的第一背叛收益,具体包括:

[0017] 根据该智能体的背叛行为策略,确定该智能体的背叛所得;

[0018] 针对每个邻居智能体,根据该智能体的被观测概率以及该邻居智能体对该智能体的第一惩罚行为策略,确定该智能体在该邻居智能体的第一惩罚行为策略下的背叛损失;

[0019] 根据所述背叛所得以及该智能体在各邻居智能体的第一惩罚行为策略下的背叛损失,确定该智能体的第一背叛收益。

[0020] 可选地,根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益以及总收益,更新该智能体的背叛参数和惩罚参数,具体包括:

[0021] 根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益以及总收益,确定该智能体在当前时刻的总收益;

[0022] 根据各邻居智能体的当前时刻的总收益,确定平均总收益;

[0023] 判断该智能体的当前时刻的总收益是否大于所述平均总收益;

[0024] 若是,则根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数。

[0025] 可选地,根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数,具体包括:

[0026] 在预设的取值范围内确定随机数,确定所述随机数与所述取值范围最大值的比值,并判断所述比值是否大于预设探索率;

[0027] 若是,根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数;

[0028] 若否,重新初始化该智能体的背叛参数和惩罚参数,作为当前时刻该智能体更新后的背叛参数和惩罚参数。

[0029] 可选地,根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数,具体包括:

[0030] 根据该智能体的第一背叛收益与预设的背叛阈值的大小关系,确定该智能体对应的背叛参数的更新方向;

[0031] 根据预设的背叛更新步长以及所述更新方向,更新该智能体的背叛参数;

[0032] 判断该智能体的第一惩罚损失是否大于预设的惩罚阈值;

[0033] 若是,则确定该智能体的惩罚参数与预设的惩罚更新步长的和值,作为该智能体更新后的惩罚参数;

[0034] 若否,则确定该智能体的惩罚参数与预设的惩罚更新步长的差值,作为该智能体更新后的惩罚参数。

[0035] 可选地,根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益以及总收益,更新该智能体的背叛参数和惩罚参数之前,所述方法还包括:

[0036] 针对该智能体的每个邻居智能体,根据所述图数据,确定该智能体与该邻居智能

体均具有用户关系的共同邻居智能体；

[0037] 针对每个共同邻居智能体,根据该智能体对该邻居智能体的第一惩罚行为策略以及该智能体的被观测概率,确定该共同邻居智能体对该智能体的第二惩罚行为策略；

[0038] 根据该共同邻居智能体的第二惩罚行为策略以及该智能体对该邻居智能体的第一惩罚行为策略,确定该智能体的第二背叛收益；

[0039] 根据该智能体对各邻居智能体的第二惩罚行为策略,确定该智能体的第二惩罚损失。

[0040] 可选地,更新该智能体的背叛参数和惩罚参数,具体包括:

[0041] 根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益、第二背叛收益、第二惩罚损失以及总收益,确定该智能体在当前时刻的总收益；

[0042] 根据各邻居智能体的当前时刻的总收益,确定平均总收益；

[0043] 判断该智能体的当前时刻的总收益是否大于所述平均总收益；

[0044] 若是,则根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失、第二惩罚损失和第二背叛收益,更新该智能体的惩罚参数。

[0045] 可选地,根据该智能体的第一惩罚损失、第二惩罚损失和第二背叛收益,更新该智能体的惩罚参数,具体包括:

[0046] 将第一惩罚损失与第二惩罚损失之和,作为该智能体的总惩罚损失；

[0047] 判断所述总惩罚损失是否大于该智能体的第二背叛收益；

[0048] 若是,则确定该智能体的惩罚参数与预设的惩罚更新步长的和值,作为该智能体更新后的惩罚参数；

[0049] 若否,则确定该智能体的惩罚参数与预设的惩罚更新步长的差值,作为该智能体更新后的惩罚参数。

[0050] 本说明书提供了一种计算机可读存储介质,所述存储介质存储有计算机程序,所述计算机程序被处理器执行时实现一种基于非合作博弈的策略生成方法。

[0051] 本说明书提供了一种电子设备,所述电子设备包括存储器、处理器及存储在存储器上并可在处理器上运行的计算机程序,所述处理器执行所述程序时实现一种基于非合作博弈的策略生成方法。

[0052] 本说明书采用的上述至少一个技术方案能够达到以下有益效果:

[0053] 在本说明书提供的一种基于非合作博弈的策略生成方法中,通过构建目标用户对应的图数据,针对每个智能体,根据该智能体的被观测概率以及背叛参数,确定该智能体的背叛行为策略。根据各邻居智能体的背叛行为策略和各邻居智能体的被观测概率,通过该智能体的惩罚参数,确定该智能体分别对各邻居智能体的第一惩罚行为策略。进而参考各智能体在背叛行为策略和第一惩罚行为策略的影响下,导致的总收益的变化,更新该智能体的背叛参数和惩罚参数用于确定下一时刻该智能体的背叛行为策略和第一惩罚行为策略以适应智能体之间行为策略的影响。

[0054] 从上述方法可以看出,通过构建目标用户对应的图数据,参考了与目标用户相关用户的行为策略的影响,以及实时更新目标用户的惩罚参数和背叛参数,得到的行为策略,更准确、实效性更强。

附图说明

[0055] 此处所说明的附图用来提供对本说明书的进一步理解,构成本说明书的一部分,本说明书的示意性实施例及其说明用于解释本说明书,并不构成对本说明书的不当限定。在附图中:

[0056] 图1为本说明书中提供的一种基于非合作博弈的策略生成方法的流程示意图;

[0057] 图2为本说明书中提供的目标用户对应的图数据的结构示意图;

[0058] 图3为本说明书中提供的智能体的背叛参数的示意图;

[0059] 图4为本说明书中提供的基于非合作博弈的策略生成的装置示意图;

[0060] 图5为本说明书中提供的用于实现一种基于非合作博弈的策略生成方法对应的电子设备结构示意图。

具体实施方式

[0061] 为使本说明书的目的、技术方案和优点更加清楚,下面将结合本说明书具体实施例及相应的附图对本说明书技术方案进行清楚、完整地描述。显然,所描述的实施例仅是本说明书一部分实施例,而不是全部的实施例。基于本说明书中的实施例,本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例,都属于本说明书保护的范围。

[0062] 目前,非合作博弈场景下的智能体一般是采用策略复制的方式生成其行为策略。以非合作博弈场景为社交网络(线上社交)的实际应用场景为例,策略复制指每个智能体均可复制高收益的智能体的行为策略,生成下一时刻的行为策略。而策略复制的方式存在局限性。首先,智能体通过复制高收益的智能体的策略,导致智能体群体的探索空间受限,陷入局部最优解,阻碍了智能体探索其他的行为策略的能力。其次,由于社交网络中个体的多样性,所以个体通过复制高收益智能体的行为策略可能不是自身的最优行为策略,准确性较差。

[0063] 基于此,本说明书提供了一种基于非合作博弈的策略生成方法,用以生成智能体的策略。本说明书提供的技术方案,首先,通过构建目标用户对应的图数据,考虑了智能体间行为策略上的影响,使得智能体通过更新后的背叛参数和惩罚参数更适应智能体之间在行为策略上的影响。其次,本技术方案中设置了探索率,使得智能体具有探索不同行为策略的能力,以及在生成智能体的策略时考虑了其他智能体的影响,使得生成智能体的策略更准确。

[0064] 以下结合附图,详细说明本说明书各实施例提供的技术方案。

[0065] 图1为本说明书中提供的一种基于非合作博弈的策略生成方法的流程示意图,包括以下步骤:

[0066] S100:响应于目标用户的请求,构建所述目标用户对应的图数据,其中,所述图数据中的节点为用户,边为用户之间的用户关系。

[0067] 在本说明书一个或多个实施例中,不限制具体由何种设备实行生成智能体在非合作博弈下的策略的过程,例如,个人电脑、移动终端以及服务器等。但由于后续步骤涉及到数据处理等操作,而这种对计算资源的要求较高的操作一般都由服务器来执行,因此本说明书后续也以服务器生成智能体在非合作博弈下的策略为例进行描述。其中,该服务器可以是单独的一台设备,或者由多台设备组成,例如,分布式服务器,本说明书对此不做限制。

[0068] 如前所述,本说明书提供了一种实现基于非合作博弈的策略生成方法,通过构建目标用户对应的图数据,生成图数据中各节点分别对应的智能体,以模拟目标用户对应的智能体在其他智能体行为策略的影响下,实时生成其行为策略的决策行为。因此,该服务器可响应于目标用户的请求,用以构建该目标用户的图数据,用以后续步骤生成目标用户对应的图数据中各节点分别对应的智能体。构建的图数据还可用于确定各智能体对应的邻居智能体,其中,目标用户对应的图数据是指以目标用户为中心的进行信息交互的社交网络。

[0069] 具体的,该服务器可响应于社交网络中目标用户的请求,通过识别该目标用户请求中携带的用户身份证明(User Identification,UID),从数据库中获取该目标用户对应的用户数据,以及与该目标用户相关的其他用户的用户数据,根据获取的用户数据,构建目标用户对应的图数据。

[0070] 其中,图数据中的节点为用户,边为用户之间的用户关系,具有用户关系的用户之间会在彼此行为策略上有较大的影响,并在后续步骤中将该影响量化为用户对应智能体收益上的变化,用于更新该智能体的背叛参数和惩罚参数。与目标用户相关的其他用户指的是在该目标用户对应的图数据中与该目标用户网络距离在预设步数内的用户,网络距离表示两个用户之间的连接距离,指的是在图数据中两个用户之间通过具有用户关系的用户所需的步数。

[0071] 需要说明的是,用户关系指两个用户在社交网络中在信息传播上相互关注的关系。同时,在目标用户的图数据中,边为用户之间的用户关系,即两个用户有边相连表示这两个用户具备用户关系,以及这两个用户的网络距离的步数应为1。由于网络距离的步数为1的两个用户在彼此行为策略上的影响较大,所以,对于目标用户而言,需要确保与目标用户网络距离的步数为1的用户的行为策略的准确性,以此类推,与目标用户网络距离的步数为1的各用户的行为策略的准确性取决于与目标用户网络距离的步数为2的各用户的行为策略的准确性。因此,至少应获取与目标用户的网络距离在预设步数内的用户对应的用户数据,以保障该目标用户的行为策略的准确性。

[0072] 其中,预设步数设置的越大表示在生成目标用户的策略时考虑的其他用户的范围越广,但预设步数设置的过大会导致该目标用户的图数据中包含了较多对生成该目标用户的行为策略影响较小的其他用户,反而对生成该目标用户的行为策略产生的干扰增多。因此,本说明书中并不限制设置预设步数的具体数值,可根据需求设置。比如,可设置该目标用户对应的图数据中的预设步数为3,也就是,获取与目标用户网络距离的步数在3以内其他用户的用户数据,构建该目标用户对应的图数据。

[0073] 如图2所示,图2为响应于目标用户A的请求,构建的目标用户A对应的图数据,以及各用户与目标用户的网络距离在预设步数3以内。图中颜色的深浅代表与目标用户A网络距离的远近,颜色越深,所需网络距离的步数越小,也就是,图中的各B用户与目标用户的网络距离的步数为1,即各B用户与目标用户具备用户关系,图中的各C用户为通过具有用户关系的各B用户与目标用户的间接连接的用户,依次类推,所以各C用户与目标用户的网络距离的步数为2,图中的各D用户与目标用户的网络距离的步数为3。

[0074] S102:生成所述图数据中各节点分别对应的智能体,并针对每个智能体,确定该智能体的被观测概率,获取该智能体的背叛参数、惩罚参数以及总收益。

[0075] 在本说明书一个或多个实施例中,如前所述,在步骤S100中构建了目标用户对应

的图数据,考虑了对生成目标用户的行为策略能产生影响的其他用户,但是在社交网络中,并不是具有用户关系的两个用户一定会观测到彼此所有的行为策略,即用户不一定能观测到其他用户的“搭便车”行为,以及在社交网络中各用户在做出“搭便车”行为并不完全相同,即不同的“搭便车”行对用户在用户体验上的影响程度不同(比如,社交网络中,对观测到用户传播虚假内容和传播低质量内容的其他用户来说,低质量内容对其他用户的用户体验造成的影响更小),因此,在本步骤需生成图数据中各用户对应的智能体,同时,设置各智能体对应的被观测概率以及获取各智能体的背叛参数、惩罚参数和总收益。

[0076] 具体的,该服务器可生成目标用户对应的图数据中各节点分别对应的智能体,并针对每个智能体,确定该智能体的被观测概率。同时,如果当前时刻为还未更新该智能体的背叛参数、惩罚参数以及总收益之前,那么获取该智能体初始化的背叛参数、惩罚参数和总收益;如果当前时刻之前已更新过该智能体的背叛参数、惩罚参数以及总收益,那么获取该智能体上一时刻更新后的背叛参数、惩罚参数和总收益。在本说明书中,智能体初始化的背叛参数、惩罚参数以及总收益,指的是该服务器可随机生成各智能体的该智能体的背叛参数、惩罚参数以及总收益。或根据各智能体对应的用户数据(“搭便车”行为次数/频率、被投诉频率、投诉频率、发布信息的点赞量等用户数据),量化的背叛参数、惩罚参数以及总收益。

[0077] 由于在社交网络中,用户的“搭便车”行为并不完全相同,那么用户的不同“搭便车”行为,也就是不同背叛程度的“搭便车”行为,将用户的“搭便车”行为划分为多种不同的背叛程度,以及确定各用户在多种不同的背叛程度下的概率值。比如,智能体*i*的背叛参数包含多种背叛程度对应的概率值,该智能体的背叛参数 B_i ,如下:

$$[0078] \quad B_i = [b_i^0, b_i^1, b_i^2, \dots, b_i^{n-1}]$$

[0079] 其中, B_i 表示智能体*i*的背叛参数,衡量了智能体*i*在不同背叛程度下的概率分布,背叛程度表征智能体做出“搭便车”行为对与其具有用户关系的其他智能体的影响程度,背叛程度的值越大,影响程度越高。 $[b_i^0, b_i^1, b_i^2, \dots, b_i^{n-1}]$ 为智能体*i*在*n*种不同背叛程度下对应的各概率值的集合, b_i^0 表示智能体*i*的背叛程度为0的概率值,并且智能体*i*在*n*种不同背叛程度下的各概率值的和值为1,在本说明书并不限制*n*的具体数值,可根据需求设置。如图3所示,图中的 $b_i^0, b_i^1, b_i^2, \dots, b_i^{n-1}$ 表示智能体*i*的背叛程度分别为0~*n*-1的概率值,该图表示 $b_i^0 \sim b_i^{n-1}$ 的概率值之和为1(圆的面积为1)。

[0080] 惩罚参数表征智能体对与该智能体具备用户关系的其他智能体的惩罚倾向,也就是说,在社交网络中,目标用户的惩罚参数表示目标用户对与该目标用户具备用户关系的其他用户做出“搭便车”行为的容忍度,惩罚参数越大该目标用户的容忍度越低,越倾向于投诉与其具备用户关系并做出“搭便车”行为的其他用户。比如,获取智能体*i*在当前时刻的惩罚参数 $V_i \in [0,1]$,针对每个与智能体*i*具备用户关系的其他智能体*j*,智能体*i*的惩罚参数的值越大,表示智能体*i*越倾向与对智能体*j*做出惩罚。

[0081] 需要说明的是,针对每个智能体,该智能体的被观测概率指的是该智能体的行为策略被与该智能体具有用户关系的智能体观测到的概率。即在社交网络中,用户的被观测概率表示用户在其社交网络中所发布的消息、帖子或其他更新内容被关注该用户的其他用

户所看到的概率。因此,可以基于该智能体对应的用户在社交网络中的活跃程度初始化设置该智能体的被观测概率。当然,还可根据该用户与其他用户的互动积极性或其他用户数据确定该智能体的被观测概率,被观测概率反映了社交网络中各用户参与信息传播和互动的复杂性,因此在确定各智能体的被观测概率时可综合考虑多种因素的影响,在本说明书并不限制。

[0082] S104:根据该智能体的被观测概率以及所述背叛参数,确定该智能体的背叛行为策略。

[0083] 在本说明书一个或多个实施例中,针对每个智能体,该服务器可根据在S102步骤中确定的该智能体的被观测概率以及背叛参数,确定该智能体的背叛行为策略,用以在后续步骤中,通过该智能体的背叛行为策略,得到与该智能体具有用户关系智能体的被背叛损失。

[0084] 具体的,该服务器可针对每个智能体,通过该智能体的背叛参数和被观测概率,确定该智能体的背叛行为策略。

[0085] 沿用上例,首先根据下式判断智能体*i*是否做出“搭便车”的背叛行为策略:

$$[0086] \quad 1 - b_i^0 > S_i$$

[0087] 其中, $1 - b_i^0$ 表示背叛程度大于0的概率, S_i 表示智能体*i*的被观测概率。在上述不等式成立的条件下,从 $b_i^0, b_i^1, b_i^2, \dots, b_i^{n-1}$ 随机选取 b_i^m ($0 < m < n-1$)作为该智能体的背叛行为策略。

[0088] 需要说明的是, b_i^0 为智能体*i*选择背叛程度为0的概率,那么也就是,当智能体*i*选择 b_i^0 为背叛行为策略,即智能体*i*选择不做“搭便车”行为的概率,相对应的 $1 - b_i^0$ 即为智能体*i*选择做“搭便车”行为的概率,当 $1 - b_i^0$ 大于被观测概率 S_i 时,代表该智能体选择做“搭便车”行为的倾向较高,所以,在该条件下,该智能体会随机选择背叛参数中的一种,作为背叛行为策略。

[0089] 需要说明的是,本说明书中的背叛行为策略是指通过传播低质量的热点信息的“搭便车”行为。这种“搭便车”行为使得大量的低质量内容充斥网络,损害了社交网络中各用户的用户体验,也就是说,用户本可以通过社交网络与其他用户交流有趣或有用的信息,但由于社交网络中盲目跟风“搭便车”行为,导致用户耗费了较多的时间精力在“搭便车”行为的低质量信息上,损害了用户的用户体验。

[0090] S106:根据所述图数据,确定与该智能体具有用户关系的邻居智能体,并根据各邻居智能体的背叛行为策略,确定该智能体的被背叛损失。

[0091] 在本说明书一个或多个实施例中,在社交网络中,用户的传播低质量内容或不实消息的“搭便车”行为,会对与该用户具备用户关系的其他用户的用户体验造成了损失。也就是说,针对每个智能体,由于该智能体的背叛行为策略会使该智能体的邻居智能体产生损失,所以,服务器可针对每个智能体,确定该智能体的邻居智能体,并通过各邻居智能体的背叛行为策略,确定该智能体的被背叛损失,并在后续步骤中用于更新背叛参数和惩罚参数。

[0092] 具体的,该服务器可针对每个智能体,先通过步骤S100中获取的图数据,确定与该智能体具有用户关系的邻居智能体,针对每个邻居智能体,通过该邻居智能体的背叛行为

策略,确定该智能体在该邻居智能体的背叛行为策略下的损失,将该智能体在各邻居智能体的背叛行为策略下的损失的和值,作为该智能体的被背叛损失。

[0093] 沿用上例,针对智能体*i*的每个邻居智能体*j*,该智能体*i*在该邻居智能体*j*的背叛行为策略 b_j^m 下的被背叛损失,如下:

$$[0094] \quad TS_{i,j} = -\frac{H * (n - 1) * m}{\sum_{k=0}^{n-1} k}, j \in Nb_i$$

[0095] 其中, $TS_{i,j}$ 表示智能体*i*在邻居智能体*j*的背叛行为策略下的损失, H 为预设的平均被背叛损失水平, $\sum_{k=0}^{n-1} k$ 表示0~ $n-1$ 的平均背叛程度, Nb_i 表示智能体*i*的邻居智能体的集合。邻居智能体*j*的背叛程度*m*越高,智能体*i*在邻居智能体*j*的背叛行为策略下的损失 $TS_{i,j}$ 的绝对值越大。

[0096] S108:根据各邻居智能体的背叛行为策略、各邻居智能体的被观测概率以及该智能体的惩罚参数,确定该智能体分别对各邻居智能体的第一惩罚行为策略,并根据各第一惩罚行为策略确定该智能体的第一惩罚损失。

[0097] 在本说明书一个或多个实施例中,如前所述,对于图数据中的每个用户而言,与该用户具备用户关系的其他用户的“搭便车”行为会使该用户产生损失,当其他用户中存在做出“搭便车”行为较多时,导致该用户的损失较大。因此,该用户为了避免其他用户的“搭便车行为”对自身造成的损失,该用户可通过消耗一定的时间精力去投诉“搭便车”行为的用户。所以,该服务器可针对每个智能体,通过确定该智能体对各邻居智能体的第一惩罚行为策略,确定该智能体的第一惩罚损失,用以后续步骤中更新背叛参数和惩罚参数。

[0098] 具体的,该服务器可针对每个邻居智能体,该智能体基于该邻居智能体的被观测概率,观测到该邻居智能体的背叛行为策略,当观测到该邻居智能体的背叛程度不为零时,该智能体会基于惩罚参数对该邻居智能体做出第一惩罚行为策略,并根据该第一惩罚行为策略确定对应的损失,通过该智能体对各邻居智能体的第一惩罚行为策略下的各损失,将各损失之和作为该智能体的第一惩罚损失。

[0099] 如步骤S102中所述,惩罚参数表征智能体对与该智能体具备用户关系的其他智能体的惩罚倾向。比如,针对智能体*i*的每个邻居智能体*j*,该智能体*i*基于该邻居智能体*j*的被观测概率,观测到该邻居智能体的背叛行为策略 b_j^m ,该智能体*i*基于对应的惩罚参数 V_i 对该邻居智能体做出第一惩罚行为策略,当该第一惩罚行为策略为智能体*i*对邻居智能体*j*做出惩罚时,确定对应的损失 $C_{i,j}$ 以及 $C_{i,j} = -C$,当该第一惩罚行为策略为智能体*i*对邻居智能体*j*不做出惩罚时,确定对应的损失 $C_{i,j}$ 以及 $C_{i,j} = 0$ 。其中, C 为执行成本以及该执行成本为预设的固定值且该值大于0,本说明书并不限制执行成本的具体数值,可根据需求设置。

[0100] S110:根据该智能体的背叛行为策略、该智能体的被观测概率以及各邻居智能体对该智能体的第一惩罚行为策略,确定该智能体的第一背叛收益。

[0101] 在本说明书的一个或多个实施例中,在社交网络中,用户在做出“搭便车”行为,该用户可通过“搭便车”行为获取一定的收益(关注度、点赞量等)。在上一步骤中提到对于每个用户来说,为了避免“搭便车行为”对自身造成的损失,该用户会投诉做出“搭便车”行为的用户。同时,当用户被其他用户投诉时,该用户会有被投诉的损失(封号时长等)。也就是

说,针对每个智能体,该智能体的背叛行为策略会使该智能体有一定的收益,但也会由于邻居智能体的第一惩罚行为策略产生一定的损失。所以,该服务器可根据该智能体的背叛行为策略、该智能体的被观测概率以及各邻居智能体对该智能体的第一惩罚行为策略,确定该智能体的第一背叛收益,用以后续步骤更新背叛参数和惩罚参数。

[0102] 具体的,该服务器可通过该智能体的背叛行为策略,确定该智能体的背叛所得,并针对每个邻居智能体,该邻居智能体基于该智能体的被观测概率,观测到该智能体的背叛行为策略,并基于该邻居智能体的第一惩罚行为策略,确定该智能体在该邻居智能体的第一惩罚行为策略下的背叛损失。确定该智能体在各邻居智能体的第一惩罚行为策略下的各背叛损失之和,并将各背叛损失之和与背叛所得的和值,作为该智能体的第一背叛收益。

[0103] 比如,根据该智能体的背叛行为策略 b_i^m ,确定该智能体的背叛所得 DS_i ,如下:

$$[0104] \quad DS_i = \frac{T * (n - 1) * m}{\sum_{k=0}^{n-1} k}$$

[0105] 其中, DS_i 表示智能体 i 通过背叛行为策略获得的背叛所得, T 为预设的平均背叛所得水平,每个智能体均有 n 种不同的背叛程度, $\sum_{k=0}^{n-1} k$ 表示 $0 \sim n-1$ 的平均背叛程度。邻居智能体 i 的背叛程度 m 越高,智能体 i 在背叛行为策略下的背叛所得 DS_i 越大。

[0106] 针对智能体 i 的每个邻居智能体 j ,确定该智能体 i 在该邻居智能体 j 的第一惩罚行为策略下的背叛损失,如下:

$$[0107] \quad PS_{i,j} = -\frac{P * (n - 1) * m}{\sum_{k=0}^{n-1} k}, j \in Nb_i$$

[0108] 其中, $PS_{i,j}$ 表示智能体 i 在邻居智能体 j 的第一惩罚行为策略下的背叛损失, P 为预设的平均背叛损失水平, Nb_i 表示智能体 i 的邻居智能体的集合,智能体 i 的背叛程度 m 越高,智能体 i 在该邻居智能体 j 的第一惩罚行为策略下的背叛损失 $PS_{i,j}$ 的绝对值越大。

[0109] S112:根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益以及总收益,更新该智能体的背叛参数和惩罚参数,并根据更新后的背叛参数和惩罚参数继续确定下一时刻的背叛行为策略和第一惩罚行为策略。

[0110] 在本说明书的一个或多个实施例中,该服务器可根据上述步骤中确定的被背叛损失、第一惩罚损失、第一背叛收益,以及获取的总收益,更新该智能体的背叛参数和惩罚参数,并根据更新后的背叛参数和惩罚参数继续确定下一时刻的背叛行为策略和第一惩罚行为策略。

[0111] 具体的,该服务器可将该智能体在当前时刻的被背叛损失、第一惩罚损失、第一背叛收益与获取的总收益之和,作为该智能体在当前时刻的总收益,以及根据各邻居智能体的当前时刻的总收益的平均值,为平均总收益。

[0112] 当该智能体的当前时刻的总收益大于该平均总收益时,该服务器可根据该智能体的第一背叛收益与预设的背叛阈值的大小关系,确定该智能体对应的背叛参数的更新方向,并根据预设的背叛更新步长以及该更新方向,更新该智能体的背叛参数。同时,该服务器可判断该智能体的第一惩罚损失是否大于预设的惩罚阈值,若是,则确定该智能体的惩罚参数与预设的惩罚更新步长的和值,作为该智能体更新后的惩罚参数,若否,则确定该智

能体的惩罚参数与预设的惩罚更新步长的差值,作为该智能体更新后的惩罚参数,以及更新后的惩罚参数小于0时取该惩罚参数为0,更新后的惩罚参数大于1时取该惩罚参数为1。

[0113] 当该智能体的当前时刻的总收益不大于该平均总收益时,则将上一时刻的背叛参数和惩罚参数,作为更新后的背叛参数和惩罚参数。

[0114] 沿用步骤S100中的例子,该服务器可该智能体的第一背叛收益与预设的背叛阈值的大小关系,确定该智能体对应的背叛参数的更新方向,根据预设的背叛更新步长以及该更新方向,更新该智能体的背叛参数,具体如下:

[0115] (1) 当智能体*i*的第一背叛收益大于预设的背叛阈值时,减小该智能体背叛程度为0的概率,均增大该智能体的背叛参数中其他背叛程度的概率。

$$[0116] \quad \begin{cases} b_i^0 = \max(b_i^{0'} - \delta_B, 0) \\ b_i^m = \min(b_i^{0'} + \delta_B/n, 1) \end{cases}$$

[0117] (2) 当智能体*i*的第一背叛收益不大于预设的背叛阈值时,增大该智能体背叛程度为0的概率,均减小该智能体的背叛参数中其他背叛程度的概率。

$$[0118] \quad \begin{cases} b_i^0 = \max(b_i^{0'} + \delta_B, 0) \\ b_i^m = \min(b_i^{0'} - \delta_B/n, 1) \end{cases}$$

[0119] 其中, $1 < m < n-1$, $b_i^{0'}$ 为步骤S102中获取的智能体*i*的背叛参数, δ_B 为预设的背叛更新步长,在得到该智能体各背叛程度的概率后,并对各概率值归一化处理,处理后的各概率值组成该智能体更新后的背叛参数。

[0120] 沿用上例,可根据该智能体的第一惩罚损失,更新该智能体的惩罚参数,如下:

[0121] (1) 当智能体*i*的第一惩罚损失大于预设的惩罚阈值时,则确定该智能体的惩罚参数与预设的惩罚更新步长的和值,作为该智能体更新后的惩罚参数 V_i , 如下:

$$[0122] \quad V_i = \max(V_i' + \delta_V, 0)$$

[0123] (2) 当智能体*i*的第一惩罚损失不大于预设的惩罚阈值时,则确定该智能体的惩罚参数与预设的惩罚更新步长的差值,作为该智能体更新后的惩罚参数 V_i , 如下:

$$[0124] \quad V_i = \min(V_i' - \delta_V, 1)$$

[0125] 其中, V_i 为步骤S102中获取的智能体*i*的惩罚参数, δ_V 为预设的惩罚更新步长。

[0126] 可选的,在步骤S102中,若各智能体的惩罚参数和背叛参数是服务器随机生成的,那么在根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数之前。该服务器可针对每个智能体,在预设的取值范围内确定随机数,确定该随机数与预设的取值范围最大值的比值。若该比值大于预设概率,根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数。若是该比值不大于预设概率,可重新随机生成该智能体的背叛参数和惩罚参数,作为当前时刻该智能体更新后的背叛参数和惩罚参数。

[0127] 在上述方法中,服务器通过构建目标用户对应的图数据,考虑了对目标用户产生较大影响的其他用户,以及参考各智能体在背叛行为策略和第一惩罚行为策略的相互影响,更新该智能体的背叛参数和惩罚参数。使通过各智能体更新后的背叛参数和惩罚参数确定下一时刻该智能体的背叛行为策略和第一惩罚行为策略,更准确、实效性高以及更适应各智能体的行为策略的动态变化对该智能体的影响。

[0128] 在上述步骤中,该服务器针对每个智能体,根据构建的图数据,确定与该智能体具有用户关系的邻居智能体,以及针对每个邻居智能体,当该智能体观测到该邻居智能体的“搭便车”行为,但基于惩罚参数较小以及为了避免执行成本,该智能体未对该邻居智能体的“搭便车”行为做出投诉时,会使该邻居智能体由于被投诉产生的损失较小继续“搭便车”行为,但邻居智能体持续性的“搭便车”行为会使该智能体继续产生被背叛损失,直到该邻居智能体不再做“搭便车”行为。

[0129] 因此,为了使该智能体避免持续受到该邻居智能体的“搭便车”行为的损害。为了使该智能体避免持续受到该邻居智能体的“搭便车”行为的损害。在S112步骤之前,在本说明书一个或多个实施例,该服务器可针对该智能体的每个邻居智能体,根据构建的图数据,确定该智能体与该邻居智能体均具有用户关系的共同邻居智能体。并针对每个共同邻居智能体,当该共同邻居智能体基于该智能体的被观测概率,观测到该智能体对该邻居智能体的未惩罚的第一惩罚行为策略,以及基于该邻居智能体的被观测概率,观测到该邻居智能体的“搭便车”行为的背叛行为策略时,该共同邻居智能体可基于自身的惩罚参数对该邻居智能体和该智能体均做出惩罚的第二惩罚行为策略。

[0130] 具体的,首先,该服务器可针对该智能体的每个邻居智能体,通过步骤S100中目标用户对应的图数据,确定该智能体与该邻居智能体均具有用户关系的共同邻居智能体。并针对每个共同邻居智能体,根据该智能体对该邻居智能体的第一惩罚行为策略以及该智能体的被观测概率,确定该共同邻居智能体对该智能体的第二惩罚行为策略。以及根据该共同邻居智能体的第二惩罚行为策略以及该智能体对该邻居智能体的第一惩罚行为策略,确定该智能体的第二背叛收益。根据该智能体对各邻居智能体的第二惩罚行为策略,确定该智能体的第二惩罚损失。

[0131] 其次,该服务器根据该智能体在当前时刻的被背叛损失、第一惩罚损失、第一背叛收益、第二背叛收益、第二惩罚损失与获取的总收益之和,作为该智能体在当前时刻的总收益。并将各邻居智能体的当前时刻的总收益的平均值,作为平均总收益。当该智能体的当前时刻的总收益大于该平均总收益,则与步骤S112中相同更新背叛参数相同,根据该智能体的第一背叛收益,更新该智能体的背叛参数。同时,将步骤S108中获取的第一惩罚损失与第二惩罚损失之和,作为该智能体的总惩罚损失,并判断总惩罚损失是否大于该智能体的第二背叛收益,若是,则确定该智能体的惩罚参数与预设的惩罚更新步长的和值,作为该智能体更新后的惩罚参数,若否,则确定该智能体的惩罚参数与预设的惩罚更新步长的差值,作为该智能体更新后的惩罚参数。

[0132] 沿用步骤S108中的例子,根据该智能体对各邻居智能体的第二惩罚行为策略,确定该智能体的第二惩罚损失。针对智能体*i*的每个邻居智能体,智能体*i*基于对应的惩罚参数 V_i 对邻居智能体*j*的第二惩罚行为策略,当该第二惩罚行为策略为智能体*i*对邻居智能体*j*做出惩罚时,确定对应的损失 $C_{i,j}$ 以及 $C_{i,j} = -C$,当该第一惩罚行为策略为智能体*i*对邻居智能体*j*不做出惩罚时,确定对应的损失 $C_{i,j}$ 以及 $C_{i,j} = 0$,其他各参数在步骤S108中均已说明,在此不再赘述。

[0133] 沿用步骤S110中的例子对确定智能体的第二背叛收益进行说明,该共同邻居智能体基于该智能体的被观测概率,观测到该智能体对该邻居智能体的第一惩罚行为策略,并基于该邻居智能体的第二惩罚行为策略,确定该智能体在该共同邻居智能体的第二惩罚行

为策略下的损失。确定该智能体在各共同邻居智能体的第二惩罚行为策略下的各损失之和,作为该智能体的第一背叛收益。针对智能体*i*和智能体*j*,确定智能体*i*和智能体*j*的共同邻居智能体*h*,智能体*h*对该智能体*i*针对智能体*j*的第一惩罚行为策略做出惩罚的第二惩罚行为策略,对应的损失 $POS_{i,j,h}$,如下:

$$[0134] \quad POS_{i,j,h} = -P, j \in Nb_i, h \in Nb_i \cap Nb_j$$

[0135] 其中, P 为预设的平均背叛损失水平, Nb_i 表示智能体*i*的邻居智能体的集合, Nb_j 表示智能体*j*的邻居智能体的集合。

[0136] 沿用步骤S112中的例子,可根据该智能体的第一惩罚损失、第二惩罚损失以及第二背叛收益,更新该智能体的惩罚参数。将步骤S108中获取的第一惩罚损失与第二惩罚损失之和,作为该智能体的总惩罚损失,并判断总惩罚损失是否大于该智能体的第二背叛收益,如下:

[0137] (1)当智能体*i*的总惩罚损失大于第二背叛收益时,则确定该智能体的惩罚参数与预设的惩罚更新步长的和值,作为该智能体更新后的惩罚参数 V_i ,如下:

$$[0138] \quad V_i = \max(V_i' + \delta_v, 0)$$

[0139] (2)当智能体*i*的总惩罚损失不大于第二背叛收益时,则确定该智能体的惩罚参数与预设的惩罚更新步长的差值,作为该智能体更新后的惩罚参数 V_i ,如下:

$$[0140] \quad V_i = \min(V_i' - \delta_v, 1)$$

[0141] 其中, V_i 为步骤S102中获取的智能体*i*的惩罚参数, δ_v 为预设的惩罚更新步长。

[0142] 通过该共同邻居智能体的第二惩罚行为策略,实现了对该智能体不惩罚的第一惩罚行为策略的惩罚,如前所述,该智能体仅是对该邻居智能体的“搭便车”行为未做出及时的惩罚,可能会使得邻居智能体继续做出“搭便车”行为对该智能体以及该共同邻居智能体造成了持续性的被背叛损失,所以对该智能体的第二惩罚行为策略可以在一定程度上避免邻居智能体的“搭便车”行为继续造成损失,间接考虑了下一时刻该智能体与邻居智能体的行为策略的影响,使得更新后的该智能体的背叛参数和惩罚参数,更准确,实效性强。

[0143] 以上为本说明书的一个或多个实施例提供的一种基于非合作博弈的策略生成方法,基于同样的思路,本说明书还提供了相应的基于非合作博弈的策略生成方法装置,如图4所示。

[0144] 图生成模块500,响应于目标用户的请求,构建所述目标用户对应的图数据,其中,所述图数据中的节点为用户,边为用户之间的用户关系;

[0145] 获取模块501,生成所述图数据中各节点分别对应的智能体,并针对每个智能体,确定该智能体的被观测概率,获取该智能体的背叛参数、惩罚参数以及总收益;

[0146] 背叛模块502,根据该智能体的被观测概率以及所述背叛参数,确定该智能体的背叛行为策略;

[0147] 被背叛损失模块503,根据所述图数据,确定与该智能体具有用户关系的邻居智能体,并根据各邻居智能体的背叛行为策略,确定该智能体的被背叛损失;

[0148] 第一惩罚模块504,根据各邻居智能体的背叛行为策略、各邻居智能体的被观测概率以及该智能体的惩罚参数,确定该智能体分别对各邻居智能体的第一惩罚行为策略,并根据各第一惩罚行为策略确定该智能体的第一惩罚损失;

[0149] 背叛收益模块505,根据该智能体的背叛行为策略、该智能体的被观测概率以及各邻居智能体对该智能体的第一惩罚行为策略,确定该智能体的第一背叛收益;

[0150] 第一更新模块506,根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益以及总收益,更新该智能体的背叛参数和惩罚参数,并根据更新后的背叛参数和惩罚参数继续确定下一时刻的背叛行为策略和第一惩罚行为策略。

[0151] 可选地,所述背叛收益模块505,具体用于根据该智能体的背叛行为策略,确定该智能体的背叛所得,针对每个邻居智能体,根据该智能体的被观测概率以及该邻居智能体对该智能体的第一惩罚行为策略,确定该智能体在该邻居智能体的第一惩罚行为策略下的背叛损失,根据所述背叛所得以及该智能体在各邻居智能体的第一惩罚行为策略下的背叛损失,确定该智能体的第一背叛收益。

[0152] 可选地,所述第一更新模块506,具体用于根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益以及总收益,确定该智能体在当前时刻的总收益,根据各邻居智能体的当前时刻的总收益,确定平均总收益,判断该智能体的当前时刻的总收益是否大于所述平均总收益,若是,则根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数。

[0153] 可选地,所述第一更新模块506,具体用于在预设的取值范围内确定随机数,确定所述随机数与所述取值范围最大值的比值,并判断所述比值是否大于预设探索率,若是,根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失,更新该智能体的惩罚参数,若否,重新初始化该智能体的背叛参数和惩罚参数,作为当前时刻该智能体更新后的背叛参数和惩罚参数。

[0154] 可选地,所述第一更新模块506,具体用于根据该智能体的第一背叛收益与预设的背叛阈值的大小关系,确定该智能体对应的背叛参数的更新方向,根据预设的背叛更新步长以及所述更新方向,更新该智能体的背叛参数;判断该智能体的第一惩罚损失是否大于预设的惩罚阈值,若是,则确定该智能体的惩罚参数与预设的惩罚更新步长的和值,作为该智能体更新后的惩罚参数;若否,则确定该智能体的惩罚参数与预设的惩罚更新步长的差值,作为该智能体更新后的惩罚参数。

[0155] 可选地,该装置还可设置第二惩罚模块507,所述第二惩罚模块507具体用于针对该智能体的每个邻居智能体,根据所述图数据,确定该智能体与该邻居智能体均具有用户关系的共同邻居智能体,针对每个共同邻居智能体,根据该智能体对该邻居智能体的第一惩罚行为策略以及该智能体的被观测概率,确定该共同邻居智能体对该智能体的第二惩罚行为策略,根据该共同邻居智能体的第二惩罚行为策略以及该智能体对该邻居智能体的第一惩罚行为策略,确定该智能体的第二背叛收益,根据该智能体对各邻居智能体的第二惩罚行为策略,确定该智能体的第二惩罚损失。

[0156] 可选地,该装置还可设置第二更新模块508,所述第二更新模块508具体用于根据该智能体的被背叛损失、第一惩罚损失、第一背叛收益、第二背叛收益、第二惩罚损失以及总收益,确定该智能体在当前时刻的总收益根据各邻居智能体的当前时刻的总收益,确定平均总收益,判断该智能体的当前时刻的总收益是否大于所述平均总收益,若是,则根据该智能体的第一背叛收益,更新该智能体的背叛参数,以及根据该智能体的第一惩罚损失、第二惩罚损失和第二背叛收益,更新该智能体的惩罚参数。

[0157] 可选地,该装置还可设置第二更新模块508,具体用于将第一惩罚损失与第二惩罚损失之和,作为该智能体的总惩罚损失;判断所述总惩罚损失是否大于该智能体的第二背叛收益,若是,则确定该智能体的惩罚参数与预设的惩罚更新步长的和值,作为该智能体更新后的惩罚参数,若否,则确定该智能体的惩罚参数与预设的惩罚更新步长的差值,作为该智能体更新后的惩罚参数。

[0158] 本说明书还提供了一种计算机可读存储介质,该存储介质存储有计算机程序,计算机程序可用于执行上述图1提供的一种基于非合作博弈的策略生成方法。

[0159] 本说明书还提供了图5所示的电子设备的示意结构图。如图5所述,在硬件层面,该设备包括处理器、内部总线、网络接口、内存以及非易失性存储器,当然还可能包括其他业务所需要的硬件。处理器从非易失性存储器中读取对应的计算机程序到内存中然后运行,以实现上述图1所述基于非合作博弈的策略生成方法。当然,除了软件实现方式之外,本说明书并不排除其他实现方式,比如逻辑器件抑或软硬件结合的方式等等,也就是说以下处理流程的执行主体并不限于各个逻辑单元,也可以是硬件或逻辑器件。

[0160] 在20世纪90年代,对于一个技术的改进可以很明显地区分是硬件上的改进(例如,对二极管、晶体管、开关等电路结构的改进)还是软件上的改进(对于方法流程的改进)。然而,随着技术的发展,当今的很多方法流程的改进已经可以视为硬件电路结构的直接改进。设计人员几乎都通过将改进的方法流程编程到硬件电路中来得到相应的硬件电路结构。因此,不能说一个方法流程的改进就不能用硬件实体模块来实现。例如,可编程逻辑器件(Programmable Logic Device, PLD)(例如现场可编程门阵列(Field Programmable Gate Array, FPGA))就是这样一种集成电路,其逻辑功能由用户对器件编程来确定。由设计人员自行编程来把一个数字系统“集成”在一片PLD上,而不需要请芯片制造厂商来设计和制作专用的集成电路芯片。而且,如今,取代手工地制作集成电路芯片,这种编程也多半改用“逻辑编译器(logic compiler)”软件来实现,它与程序开发撰写时所用的软件编译器相类似,而要编译之前的原始代码也得用特定的编程语言来撰写,此称之为硬件描述语言(Hardware Description Language, HDL),而HDL也并非仅有一种,而是有许多种,如ABEL(Advanced Boolean Expression Language)、AHDL(Altera Hardware Description Language)、Confluence、CUPL(Cornell University Programming Language)、HDCal、JHDL(Java Hardware Description Language)、Lava、Lola、MyHDL、PALASM、RHDH(Ruby Hardware Description Language)等,目前最普遍使用的是VHDL(Very-High-Speed Integrated Circuit Hardware Description Language)与Verilog。本领域技术人员也应该清楚,只需要将方法流程用上述几种硬件描述语言稍作逻辑编程并编程到集成电路中,就可以很容易得到实现该逻辑方法流程的硬件电路。

[0161] 控制器可以按任何适当的方式实现,例如,控制器可以采取例如微处理器或处理器以及存储可由该(微)处理器执行的计算机可读程序代码(例如软件或固件)的计算机可读介质、逻辑门、开关、专用集成电路(Application Specific Integrated Circuit, ASIC)、可编程逻辑控制器和嵌入微控制器的形式,控制器的例子包括但不限于以下微控制器:ARC 625D、Atmel AT91SAM、Microchip PIC18F26K20 以及Silicone Labs C8051F320,存储器控制器还可以被实现为存储器的控制逻辑的一部分。本领域技术人员也知道,除了以纯计算机可读程序代码方式实现控制器以外,完全可以通过将方法步骤进行逻辑编程来

使得控制器以逻辑门、开关、专用集成电路、可编程逻辑控制器和嵌入微控制器等的形式来实现相同功能。因此这种控制器可以被认为是一种硬件部件,而对其内包括的用于实现各种功能的装置也可以视为硬件部件内的结构。或者甚至,可以将用于实现各种功能的装置视为既可以是实现方法的软件模块又可以是硬件部件内的结构。

[0162] 上述实施例阐明的系统、装置、模块或单元,具体可以由计算机芯片或实体实现,或者由具有某种功能的产品来实现。一种典型的实现设备为计算机。具体的,计算机例如可以为个人计算机、膝上型计算机、蜂窝电话、相机电话、智能电话、个人数字助理、媒体播放器、导航设备、电子邮件设备、游戏控制台、平板计算机、可穿戴设备或者这些设备中的任何设备的组合。

[0163] 为了描述的方便,描述以上装置时以功能分为各种单元分别描述。当然,在实施本说明书时可以把各单元的功能在同一个或多个软件和/或硬件中实现。

[0164] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0165] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0166] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0167] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0168] 在一个典型的配置中,计算设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

[0169] 内存可能包括计算机可读介质中的非永久性存储器,随机存取存储器(RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

[0170] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括,但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除

可编程只读存储器 (EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器 (CD-ROM)、数字多功能光盘 (DVD) 或其他光学存储、磁盒式磁带, 磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质, 可用于存储可以被计算设备访问的信息。按照本文中的界定, 计算机可读介质不包括暂存电脑可读媒体 (transitory media), 如调制的数据信号和载波。

[0171] 还需要说明的是, 术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含, 从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素, 而且还包括没有明确列出的其他要素, 或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下, 由语句“包括一个……”限定的要素, 并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

[0172] 本说明书可以在由计算机执行的计算机可执行指令的一般上下文中描述, 例如程序模块。一般地, 程序模块包括执行特定任务或实现特定抽象数据类型的例程、程序、对象、组件、数据结构等等。也可以在分布式计算环境中实践本说明书, 在这些分布式计算环境中, 由通过通信网络而被连接的远程处理设备来执行任务。在分布式计算环境中, 程序模块可以位于包括存储设备在内的本地和远程计算机存储介质中。

[0173] 本说明书中的各个实施例均采用递进的方式描述, 各个实施例之间相同相似的部分互相参见即可, 每个实施例重点说明的都是与其他实施例的不同之处。尤其, 对于系统实施例而言, 由于其基本相似于方法实施例, 所以描述的比较简单, 相关之处参见方法实施例的部分说明即可。

[0174] 以上所述仅为本说明书的实施例而已, 并不用于限制本说明书。对于本领域技术人员来说, 本说明书可以有各种更改和变化。凡在本说明书的精神和原理之内所作的任何修改、等同替换、改进等, 均应包含在本说明书的权利要求范围之内。

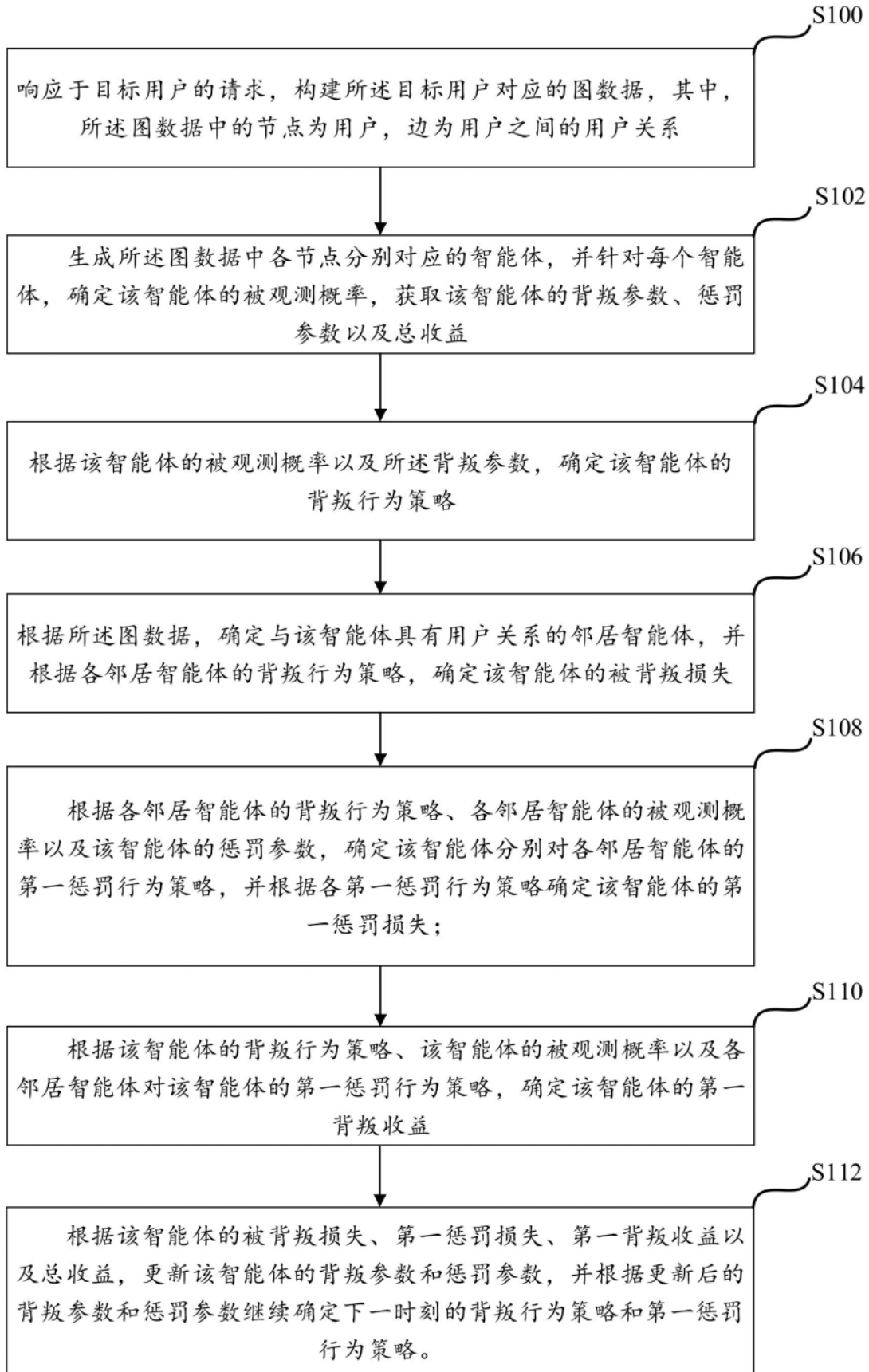


图1

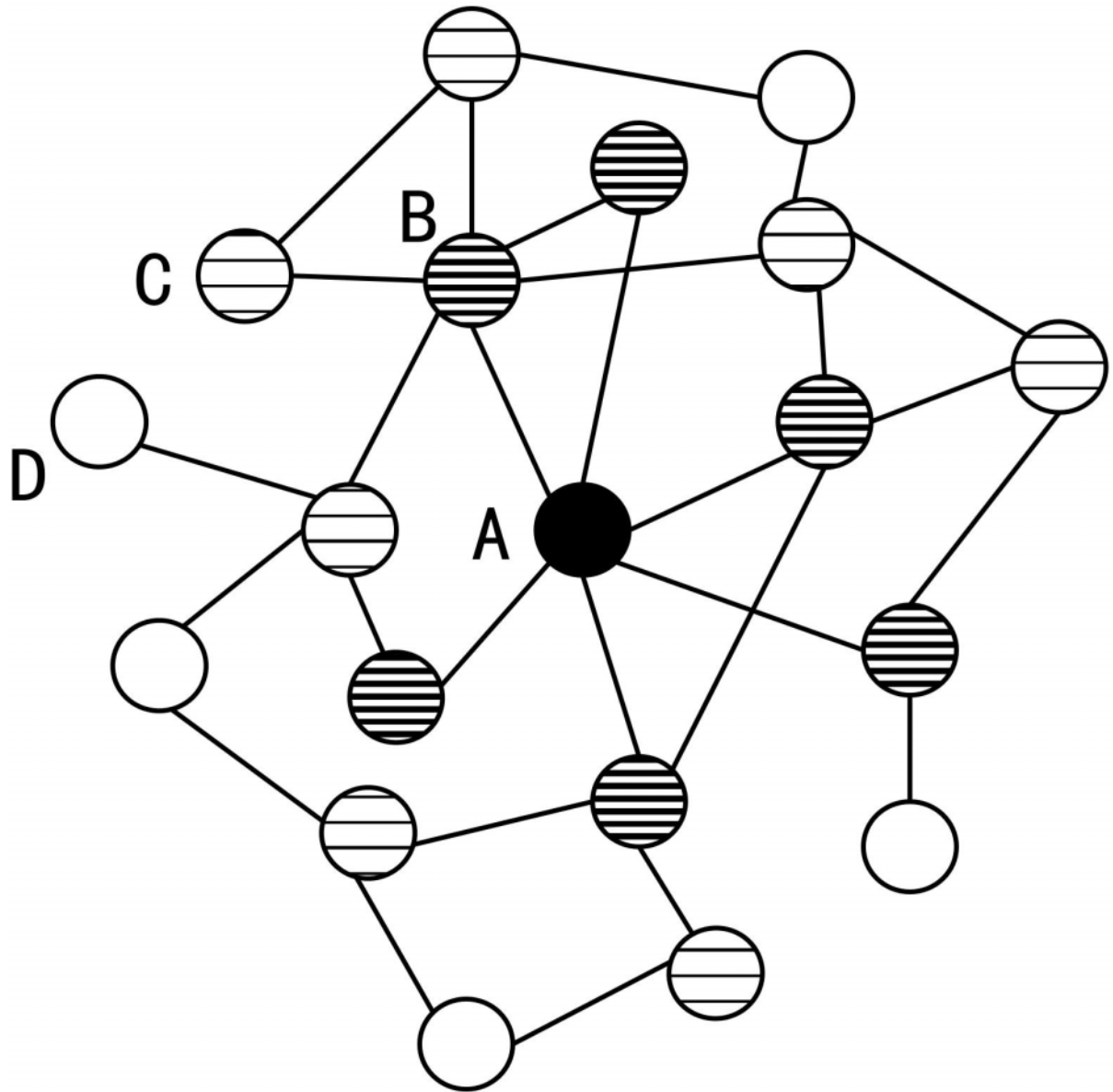


图2

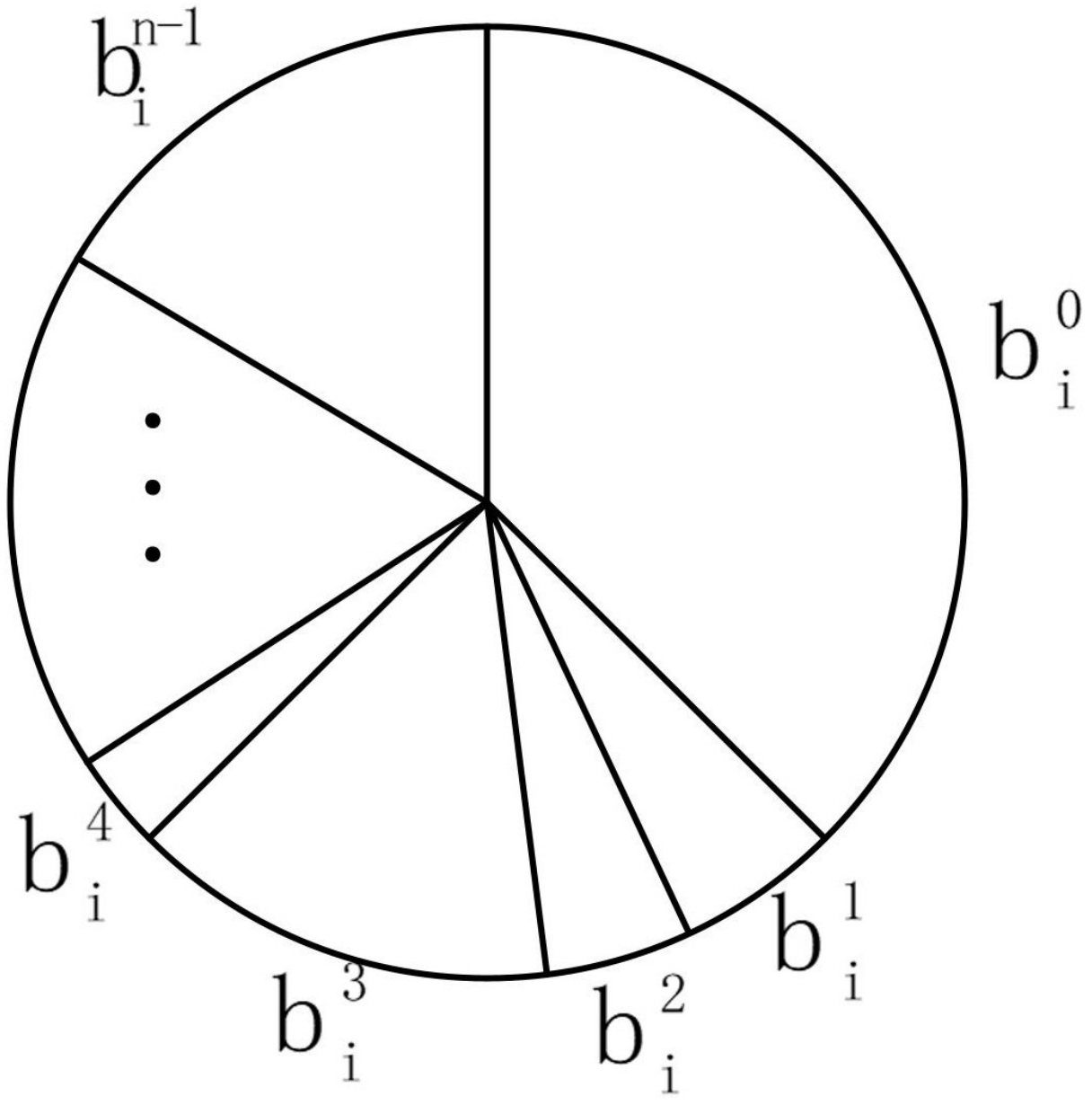


图3

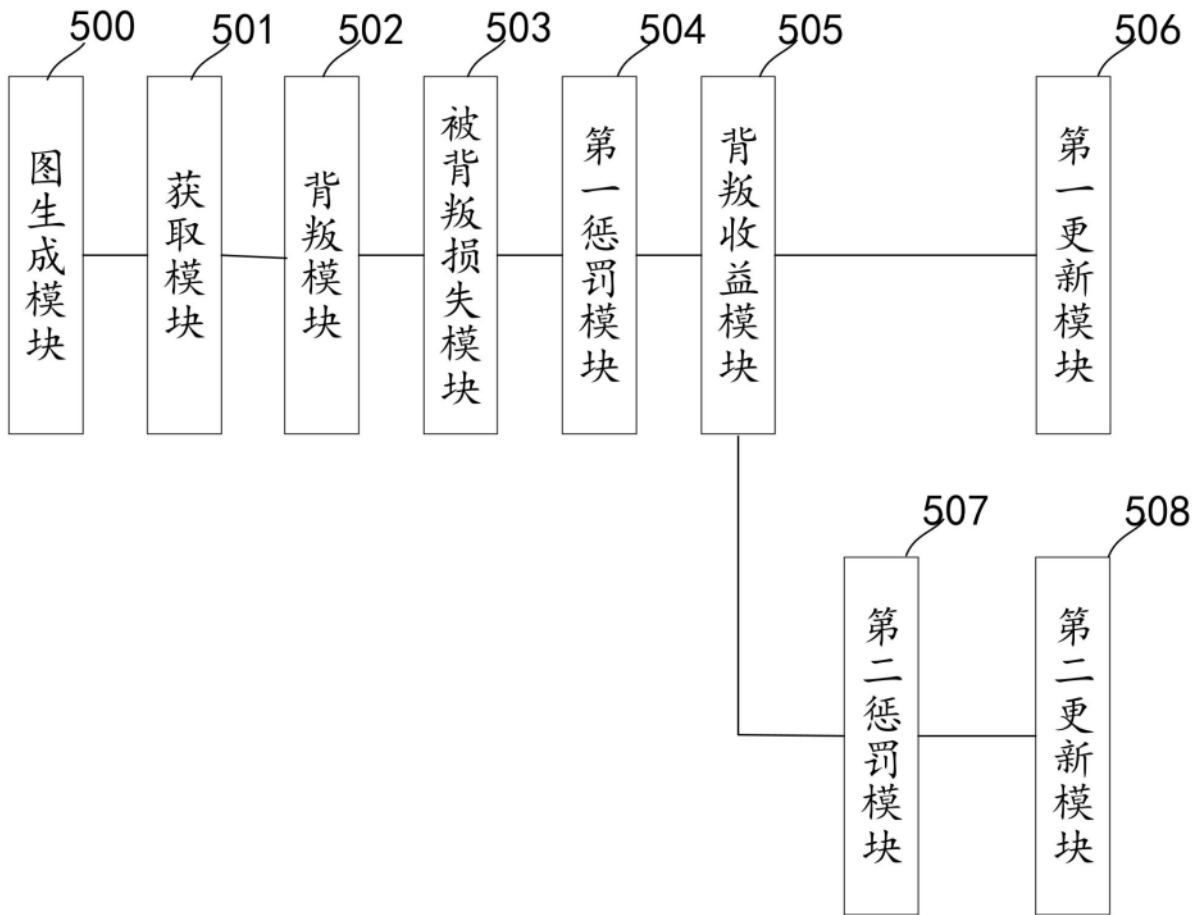


图4

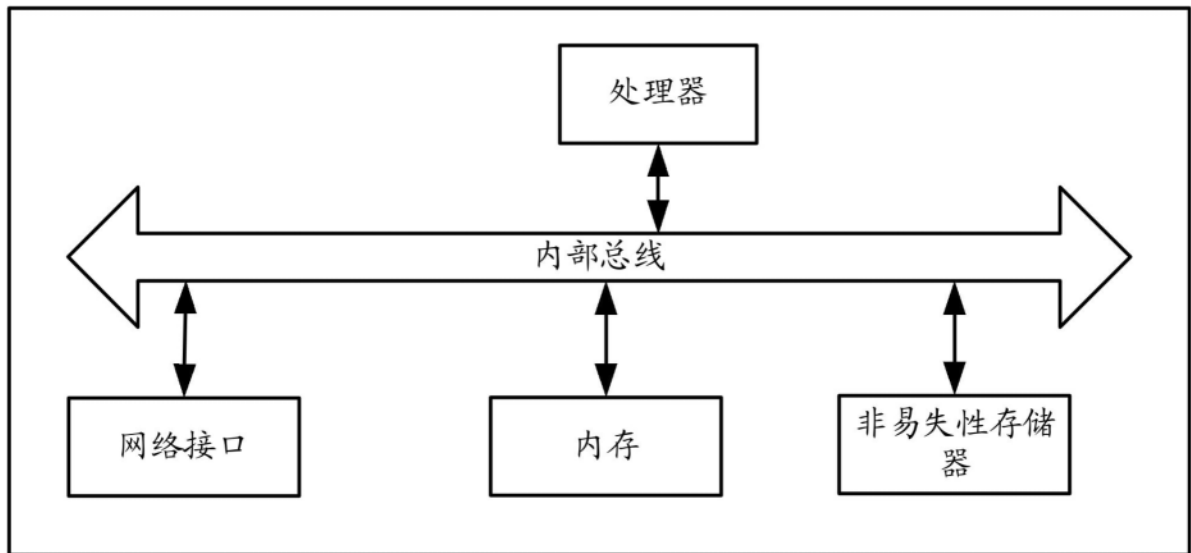


图5