



(12)发明专利申请

(10)申请公布号 CN 109728941 A
(43)申请公布日 2019.05.07

(21)申请号 201811560491.X

(22)申请日 2018.12.20

(71)申请人 北京比新科技有限公司
地址 100094 北京市海淀区永嘉北路6号5
幢2层东侧162

(72)发明人 王春峰 李军

(74)专利代理机构 北京邦创至诚知识产权代理
事务所(普通合伙) 11717
代理人 吴强

(51) Int. Cl.
H04L 12/24(2006.01)
H04L 29/08(2006.01)

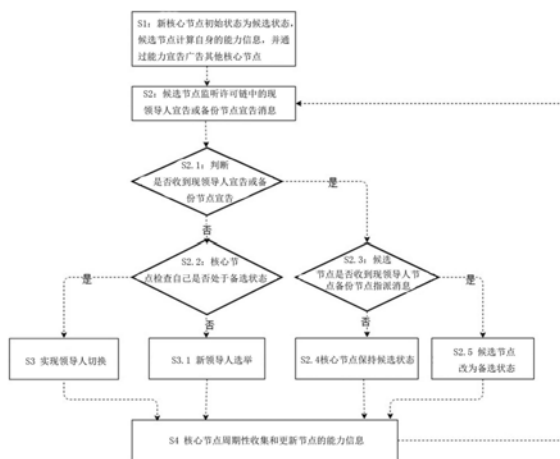
权利要求书3页 说明书9页 附图5页

(54)发明名称

一种区块链领导人选举方法及其装置

(57)摘要

本发明提供了一种区块链领导人选举方法及其装置,本发明方法根据节点能力和既定策略预先确定备份节点,从而在到期后或者突发情况下快速实现领导人节点更迭,核心节点可快速转为参选节点,并通过投票方式快速选出新的领导人节点,防止许可链中领导人缺位。由此,领导人更迭速度快,效率高,同时可以选择较为复杂的节点选择规则,兼顾效果、效率和灵活性。



1. 一种区块链领导人选举方法,其特征在于,包括:

步骤1:构成一个许可链的多个节点包括若干个核心节点,所述核心节点为经过认证授权可以参加共识算法的领导人选举,并有可能成为许可链中领导人的节点;所有的核心节点构成核心节点集合;

两两核心节点之间通过心跳消息方式维护核心节点间的连通,每个核心节点通过能力消息向其他核心节点广播自身的能力信息;

步骤2:核心节点集合包括一个唯一的现领导人节点;现领导人节点周期性地向所有其他核心节点广播关于核心节点集合的宣告信息,以使所有核心节点维护相同的核心节点集合信息;

步骤3:现领导人节点选出一个符合设定约束条件的核心节点作为备选节点;现领导人节点和备选节点之外的其他核心节点为候选节点;

现领导人节点向所有其他核心节点广播备选节点信息,被选作备选节点的核心节点将自身状态置为备选状态;

候选节点持续保持准备参加接替现领导人节点的共识算法领导人选举、进而成为许可链中新领导人的候选状态;

步骤4:现领导人检测到预设周期到期时,发出切换消息给备份节点;备份节点收到切换消息,发送切换确认信息给现领导人节点,现领导人节点转变为候选节点,备份节点转换成为新领导人节点,新领导人节点向所有候选节点发送领导人宣告信息;新领导人节点根据步骤3选择新备份节点;

步骤5:如果发生异常事件,备份节点在一定的时间内无法收到现领导人节点的宣告消息,则备份节点自动切换成为新领导人节点;新领导人节点向所有候选节点发送领导人宣告信息;新领导人节点根据步骤2选择新备份节点;

步骤6:候选节点持续监听许可链中的现领导人节点和备选节点的宣告信息,如在设定时间内未监听到现领导人节点和备选节点的宣告信息,一个或多个候选节点将自身状态变更为参选状态进而变更为参选节点;

步骤7:参选节点向许可链中其他核心节点广播自己的选举宣告,所述选举宣告为包括节点能力和节点状态的广告消息,许可链将在参选节点中选出新领导人节点;

任何收到选举宣告的核心节点根据预先设定策略或规则检查一个参选节点是否适合成为领导人节点,或者检查哪一个参选节点更合适成为领导人节点,进而向每个参选节点回复支持其当选新领导人的确认信息,或者向每个参选节点回复不支持其当选新领导人的拒绝信息,或者向所有参选节点回复投票支持哪个参选节点成为新领导人节点;

如果一个参选节点收到候选节点集合中2/3以上数量的核心节点的确认信息,该参选节点自动转换为领导人状态并成为新领导人节点,落选的参选节点自动将自身状态转变为候选状态,转变回为候选节点;

新领导人节点检查当前核心节点集合信息,依据步骤2宣告核心节点集合信息,以及根据步骤3选出备选节点。

2. 根据权利要求1所述的区块链领导人选举方法,其特征在于,步骤1中,如果能力消息或心跳消息无法送达一个核心节点,则从所述核心节点集合中删除该核心节点。

3. 根据权利要求1所述的区块链领导人选举方法,其特征在于,步骤3中所述设定约束

条件为:按预先设定的策略或规则,选取排在第一位的核心节点为备选节点。

4. 根据权利要求3所述的区块链领导人选举方法,其特征在于,所述预先设定的策略或规则为优先级顺序、加权排队或者根据节点能力信息的得出函数结果。

5. 根据权利要求4所述的区块链领导人选举方法,其特征在于,所述节点能力包括CPU速度、内存大小或存储空间大小。

6. 根据权利要求1所述的区块链领导人选举方法,其特征在于,所述许可链上还包括若干个非核心节点,非核心节点经通过认证授权后可成为核心节点,新加入的核心节点作为候选节点持续监听许可链中现领导人节点和备选节点的宣告信息。

7. 根据权利要求1所述的区块链领导人选举方法,其特征在于,一个许可链可分裂成若干个子链,所述子链分为:同时包括现领导人节点和备选节点的第一子链,包括现领导人节点而不包括备选节点的第二子链,包括备选节点而不包括现领导人节点的第三子链,以及既不包括现领导人节点又不包括备选节点的第四子链;其中,子链中的领导人选定步骤包括:

步骤8.1:第一子链依照步骤4-7进行领导人节点更迭;

步骤8.2:第二子链中现领导人依照步骤3选出第二子链的备选节点,然后参照步骤4-7进行领导人节点更迭;

步骤8.3:第三子链中,备份节点依照步骤5会切换成新领导人节点,新领导人节点根据步骤3选择新备份节点;

步骤8.4:第四子链中,候选节点因检测不到现领导人节点和备份节点而根据步骤6和7选举出新领导人节点,新领导人节点根据步骤3选出备选节点。

8. 根据权利要求1所述的区块链领导人选举方法,其特征在于,两个以上的许可链合并形成一个联合链,联合链中包括两个以上的现领导人节点,现领导人节点将自身状态变更为参选状态进而成为参选节点,联合链根据步骤7选举出新领导人节点,新领导人节点根据步骤3选出备选节点,后参照步骤4-7进行领导人节点更迭。

9. 一种用于权利要求1-8任一所述的区块链领导人选举方法的领导人选举装置,其特征在于,包括:节点状态维护模块、候选节点维护模块、节点能力计算模块、节点间通信模块、系统信息模块和配置策略管理模块;

节点状态维护模块管理和维护核心节点的状态转换,依据预先设定策略或规则,从节点能力计算模块获取许可链中节点的能力信息,并从候选节点维护模块和节点间通信模块获取核心节点的状态以及维护核心节点的状态转换;

候选节点维护模块主要负责维护链中所有候选节点的状态、能力及配置策略;候选节点维护模块会接收候选节点的能力宣告消息,添加并维护候选节点的能力信息;候选节点维护模块会接收来自配置管理模块的策略及规则配置信息,以用于候选节点能力的排序;候选节点维护模块根据各个候选节点的信息,及配置的策略和规则,对所有的候选节点进行排序,并把排序结果及各个候选节点的能力信息提供给节点状态维护模块;

节点能力计算模块计算核心节点的能力信息,从配置管理模块接收相应的能力参数配置,从系统信息模块获取系统信息,计算各个核心节点的能力值,并提供给向候选节点维护模块提供候选节点的能力信息,向节点状态维护模块提供所有核心节点的状态信息;

配置管理模块提供人机交互界面,允许用于输入并配置节点能力信息数据和领导人选

举的选择策略信息;配置管理模块将节点能力信息数据和领导人选举的选择策略信息输送给节点状态维护模块、候选节点维护模块和节点能力计算模块;

系统信息模块负责与操作系统交互获取需要的系统信息;

网络通信模块负责节点之间的信息传输和交换,其中包括领导人宣告、能力宣告、备份节点宣告、领导人选举消息;网络通信模块、节点状态维护模块和候选节点维护模块相互连接实现核心节点之间节点状态和能力信息的交互和统一。

一种区块链领导人选举方法及其装置

技术领域

[0001] 本发明涉及信息技术领域,尤其是涉及一种区块链领导人选举方法及其装置。

背景技术

[0002] 区块链是一个分布式系统,一致性是分布式系统的关键问题。所谓一致性是指:对于系统中的多个服务节点,在共识机制的保障下,多个服务节点对处理结果达成一致。因此,区块链借助共识机制解决了分布式场景中节点间互相信任和一致性的问题。

[0003] 现有区块链中采用的共识机制方法包括:工作量证明(POW)、权益证明(POS)、实用拜占庭容错算法(PBFT)以及Raft等。有的共识算法适用于公链,有的适用于许可链(联盟链和私有链)。在许可链中,虽然其共识机制运行于预先选定的服务节点至上,较公链的效率有非常大的提升,但如何进一步提升许可链共识机制性能依然是一个关键。其中,领导人选举算法是分布式的共识算法的关键。

[0004] 目前的领导人选举算法很难兼顾效率和灵活性;有的采用多简单的选举算法,如实用拜占庭容错算法(PBFT)中基于节点序号的方发;有的虽然灵活,依赖很多计算机硬件因素作为选举因子,比如IP地址、CPU核数、内存大小、自定义序列号等等,但需要经过竞争窗口的随机算选举机制,浪费系统时间,影响系统效率。

[0005] 因此,如何设计兼顾效率和灵活性的领导人选举算法对共识算法显得非常重要。

发明内容

[0006] 本发明的目的在于提出了一种基于热备方式的区块链领导人选举方法及其装置。此方法既可避免基于周期的随机选举方法的效率问题,又能有效地考虑节点的能力和系统策略,总是选择符合系统策略的最适合的备份领导人节点作为备选,从而通过快速领导人切换避免了随机选举的时间浪费,从而实现了一种高效、灵活的领导人选举机制。

[0007] 为实现上述目的,本发明提供的一种区块链领导人选举方法,包括

[0008] 步骤1:构成一个许可链的多个节点包括若干个核心节点,所述核心节点为经过认证授权可以参加共识算法的领导人选举,并有可能成为许可链中领导人的节点;所有的核心节点构成核心节点集合;

[0009] 两两核心节点之间通过心跳消息方式维护核心节点间的连通,每个核心节点通过能力消息向其他核心节点广播自身的能力信息;

[0010] 步骤2:核心节点集合包括一个唯一的现领导人节点(Leader-Status节点);现领导人节点周期性地向所有其他核心节点广播关于核心节点集合的宣告信息,以使所有核心节点维护相同的核心节点集合信息;

[0011] 步骤3:现领导人节点选出一个符合设定约束条件的核心节点作为备选节点(Standby-Status节点);现领导人节点和备选节点之外的其他核心节点为候选节点(Candidate-Status节点);

[0012] 现领导人节点向所有其他核心节点(包括所有候选节点和备选节点)广播备选节

点信息,被选作备选节点的核心节点将自身状态置为备选状态;

[0013] 候选节点持续保持准备参加接替现领导人节点的共识算法领导人选举、进而成为许可链中新领导人的候选状态;

[0014] 步骤4:现领导人检测到预设周期到期时,发出切换消息给备份节点;备份节点收到切换消息,发送切换确认信息给现领导人节点,现领导人节点转变为候选节点,备份节点转换成为新领导人节点,新领导人节点向所有候选节点发送领导人宣告信息;新领导人节点根据步骤3选择新备份节点;

[0015] 步骤5:如果发生异常事件(如现领导人节点网络故障、现领导人节点宕机等等),备份节点在一定的时间内无法收到现领导人节点的宣告消息,则备份节点自动切换成为新领导人节点;新领导人节点向所有候选节点发送领导人宣告信息;新领导人节点根据步骤2选择新备份节点;

[0016] 步骤6:候选节点持续监听许可链中的现领导人节点和备选节点的宣告信息,如在设定时间内未监听到现领导人节点和备选节点的宣告信息,一个或多个候选节点(认为许可链中缺少现领导人节点和备选节点)将自身状态变更为参选状态(Electing-Status)进而变更为参选节点;

[0017] 步骤7:参选节点向许可链中其他核心节点广播自己的选举宣告,所述选举宣告为包括节点能力和节点状态的广告消息,许可链将在参选节点中选出新领导人节点;

[0018] 任何收到选举宣告的核心节点根据所述预先设定策略或规则检查一个参选节点是否适合成为领导人节点,或者检查哪一个参选节点更合适成为领导人节点,进而向每个参选节点回复支持其当选新领导人的确认信息,或者向每个参选节点回复不支持其当选新领导人的拒绝信息,或者向所有参选节点回复投票支持哪个参选节点成为新领导人节点;

[0019] 如果一个参选节点收到候选节点集合(或核心节点集合)中2/3以上数量的核心节点的确认信息,该参选节点自动转换为领导人状态并成为新领导人节点,落选的参选节点自动将自身状态转变为候选状态(Candidate-Status),转变回为候选节点;

[0020] 新领导人节点检查当前核心节点集合信息,依据步骤2宣告核心节点集合信息,以及根据步骤3选出备选节点。

[0021] 其中,心跳消息(Heartbeat Message)是一种发送源发送到接收方的消息,这种消息可以让接收方确定发送源是否以及何时出现故障或终止。通常,心跳消息从发送源启动时开始发送,直到发送源关闭,期间发送源会不间断的发送周期性或重复消息。当接收方在某个消息接收周期内未收到消息,接收方可能会认为发送源已经关闭、出现故障、或者当前不可用。心跳消息可能常被用于高可用性或容错处理的目的。

[0022] 而能力消息是周期性广播式的消息,以反映节点能力的变化。能力消息和心跳消息可以在一起。

[0023] 进一步地,步骤1中,如果能力消息或心跳消息无法送达一个核心节点(即其他核心节点检测到该核心节点丢失),则从所述核心节点集合中删除该核心节点。

[0024] 进一步地,步骤3中所述设定约束条件为:按预先设定的策略或规则,选取排在第一位的核心节点为备选节点。

[0025] 进一步地,所述预先设定的策略或规则为优先级顺序、加权排队或者根据节点能力信息的得出函数结果。

[0026] 其中,所述节点能力包括CPU速度、内存大小或存储空间大小等等。

[0027] 进一步地,所述许可链上还包括若干个非核心节点,非核心节点经通过认证授权后可成为核心节点,新加入的核心节点作为候选节点持续监听许可链中现领导人节点和备选节点的宣告信息。

[0028] 其中,一个许可链由于网络或其他原因可能分裂为两个或多个子链,或者两个或多个许可链合并成一个联合链。

[0029] 进一步地,一个许可链可分裂成若干个子链,所述子链分为:同时包括现领导人节点和备选节点的第一子链,包括现领导人节点而不包括备选节点的第二子链,包括备选节点而不包括现领导人节点的第三子链,以及既不包括现领导人节点又不包括备选节点的第四子链;其中,子链中的领导人选定步骤包括:

[0030] 步骤8.1:第一子链依照步骤4-7进行领导人节点更迭;

[0031] 步骤8.2:第二子链中现领导人依照步骤3选出第二子链的备选节点,然后参照步骤4-7进行领导人节点更迭;

[0032] 步骤8.3:第三子链中,备份节点依照步骤5会切换成新领导人节点,新领导人节点根据步骤3选择新备份节点;

[0033] 步骤8.4:第四子链中,候选节点因检测不到现领导人节点和备份节点而根据步骤6和7选举出新领导人节点,新领导人节点根据步骤3选出备选节点。

[0034] 进一步地,两个以上的许可链合并形成一个联合链,联合链中包括两个以上的现领导人节点,现领导人节点将自身状态变更为参选状态进而成为参选节点,联合链根据步骤7选举出新领导人节点,新领导人节点根据步骤3选出备选节点,后参照步骤4-7进行领导人节点更迭。

[0035] 具体而言,现领导人节点作为参选节点向联合链中其他核心节点广播自己的选举宣告,联合链中任何收到能力宣告的核心节点根据所述预先设定策略或规则投票选出新领导人节点。

[0036] 采用上述技术方案,本发明具有如下有益效果:

[0037] 本发明提供的一种基于热备方式的区块链领导人选举方法,根据节点能力和既定策略预先确定备份节点,从而在到期后或者突发情况下快速实现领导人节点更迭,核心节点可快速转为参选节点,并通过投票方式快速选出新的领导人节点,防止许可链中领导人缺位。由此,领导人更迭速度快,效率高,同时可以选择较为复杂的节点选择规则,兼顾效果、效率和灵活性。

[0038] 另外,本发明还公开一种用于上述区块链领导人选举方法的领导人选举装置,该选举装置可以服务器中的硬件单元,也可以是软件模组。其中领导人选举装置包括:节点状态维护模块、候选节点维护模块、节点能力计算模块、节点间通信模块、系统信息模块和配置策略管理模块;

[0039] 节点状态维护模块管理和维护核心节点的状态转换,依据预先设定策略或规则,从节点能力计算模块获取许可链中节点的能力信息,并从候选节点维护模块和节点间通信模块获取核心节点的状态以及维护核心节点的状态转换;

[0040] 候选节点维护模块主要负责维护链中所有候选节点的状态、能力及配置策略;候选节点维护模块会接收候选节点的能力宣告消息,添加并维护候选节点的能力信息;候选

节点维护模块会接收来自配置管理模块的策略及规则配置信息,以用于候选节点能力的排序;候选节点维护模块根据各个候选节点的信息,及配置的策略和规则,对所有的候选节点进行排序,并把排序结果及各个候选节点的能力信息提供给节点状态维护模块;

[0041] 节点能力计算模块计算核心节点的能力信息,从配置管理模块接收相应的能力参数配置,从系统信息模块获取系统信息,计算各个核心节点的能力值,并提供给向候选节点维护模块提供候选节点的能力信息,向节点状态维护模块提供所有核心节点的状态信息;

[0042] 配置管理模块提供人机交互界面,允许用于输入并配置节点能力信息数据和领导人选举的选择策略信息;配置管理模块将节点能力信息数据和领导人选举的选择策略信息输送给节点状态维护模块、候选节点维护模块和节点能力计算模块;

[0043] 系统信息模块负责与操作系统交互获取需要的系统信息(系统可以是Linux、Windows、OSX等等以及其他任何嵌入式系统);

[0044] 网络通信模块负责节点之间的信息传输和交换,其中包括领导人宣告、能力宣告、备份节点宣告、领导人选举消息;网络通信模块、节点状态维护模块和候选节点维护模块相互连接实现核心节点之间节点状态和能力信息的交互和统一。

附图说明

[0045] 为了更清楚地说明本发明具体实施方式或现有技术中的技术方案下面将对具体实施方式或现有技术描述中所需要使用的附图作简单地介绍,显而易见地,下面描述中的附图是本发明的一些实施方式,对于本领域普通技术人员来讲,在不付出创造性劳动的前提下,还可以根据这些附图获得其他的附图。

[0046] 图1为本发明实施例1中许可链中节点部署结构图;

[0047] 图2为本发明实施例2中领导人选举装置的结构图;

[0048] 图3为本发明实施例3提供的一种领导人选举方法的主要流程图;

[0049] 图4为本发明实施例中核心节点的状态转换图;

[0050] 图5为本发明实施例中新领导人选举环节的流程图。

具体实施方式

[0051] 下面结合具体的实施方式对本发明做进一步的解释说明。

[0052] 实施例1

[0053] 本实施例提供的一种区块链领导人选举方法,包括

[0054] 步骤1:构成一个许可链的多个节点包括若干个核心节点,所述核心节点为经过认证授权可以参加共识算法的领导人选举,并有可能成为许可链中领导人的节点;所有的核心节点构成核心节点集合;

[0055] 图1为本实施例中许可链中节点部署结构图。如图1所示,在许可链中,有6个节点,其中节点101、102、103和104是核心节点,并构成核心节点集合 $100 = \{101, 102, 103, 104\}$,节点105、106是非核心节点。

[0056] 每个核心节点(即101-104所有核心节点)周期性的计算自己的节点能力,并向其他核心节点周期性广播其状态和能力,以使现领导人节点和整个网络了解这个节点的状态,并维护最新的核心节点集合信息。

[0057] 核心节点101、102、103、104中两两之间通过心跳消息方式维护核心节点间的连通,每个核心节点通过能力消息向其他核心节点广播自身的能力信息。具体而言,各个节点之间通过虚拟链路107-112维护任何两个节点间的消息通信。

[0058] 其中,如果能力消息或心跳消息无法送达一个核心节点(即其他核心节点检测到该核心节点丢失),则从所述核心节点集合中删除该核心节点。

[0059] 步骤2:核心节点集合包括一个唯一的现领导人节点(Leader-Status节点);现领导人节点周期性地向所有其他核心节点广播关于核心节点集合的宣告信息,以使所有核心节点维护相同的核心节点集合信息。

[0060] 其中,节点101是现领导人节点,处于领导人状态(Leader-Status);现领导人节点101周期性地向核心节点102、103和104广播关于核心节点集合的宣告信息,以使所有核心节点101、102、103和104之间维护相同的核心节点集合信息。

[0061] 步骤3:现领导人节点选出一个符合设定约束条件的核心节点作为备选节点(Standby-Status节点);现领导人节点和备选节点之外的其他核心节点为候选节点(Candidate-Status节点);

[0062] 现领导人节点向所有其他核心节点(包括所有候选节点和备选节点)广播备选节点信息,被选作备选节点的核心节点将自身状态置为备选状态;

[0063] 候选节点持续保持准备参加接替现领导人节点的共识算法领导人选举、进而成为许可链中新领导人的候选状态;

[0064] 当现领导人节点101产生的时候,现领导人节点101会立即选出一个备选节点,例如根据约定的规则和策略,核心节点102为被选为备份节点,现领导人节点101会发送备份消息给核心节点102,核心节点102转变为备份节点。

[0065] 核心节点103和104则成为候选节点,并处于候选状态(Candidate-status)。

[0066] 另外,许可链中的节点不断维护自身的状态转换,任何一个新加入的核心节点,初始状态自动处于候选状态,并成为候选节点。

[0067] 例如假设非核心104节点当复合设定条件后转变为核心节点,在此之前,核心节点集合100是{101,102,103},节点104加入后监听到现领导人节点101的领导人宣告,作为新加入的核心节点自动成为候选节点。

[0068] 104同时向其他核心节点发出自己的能力宣告,现领导人节点101收到104的能力宣告,现领导人节点101更新核心节点集合信息,将104加到候选节点中,候选节点集合={103,104},核心节点集合{101,102,103}变为核心节点集合{101,102,103,104},现领导人节点101将新的核心节点集合信息发布全许可链节点,统一全网信息。

[0069] 其中设定约束条件可以是按预先设定的策略或规则,选取排在第一位的核心节点为备选节点。例如预先设定的策略或规则为优先级顺序、加权排队或者根据节点能力信息的得出函数结果。其中节点能力包括CPU速度、内存大小或存储空间大小等等。

[0070] 步骤4:现领导人检测到预设周期到期时,发出切换消息给备份节点;备份节点收到切换消息,发送切换确认信息给现领导人节点,现领导人节点转变为候选节点,备份节点转换成为新领导人节点,新领导人节点向所有候选节点发送领导人宣告信息;新领导人节点根据步骤3选择新备份节点。

[0071] 即约定周期到期后,备选节点102自动成为新领导节点。

[0072] 步骤5:如果发生异常事件(如现领导人节点网络故障、现领导人节点宕机等等),备份节点在一定的时间内无法收到现领导人节点的宣告消息,备份节点自动切换成为新领导人节点;新领导人节点向所有候选节点发送领导人宣告信息;新领导人节点根据步骤2选择新备份节点;

[0073] 假如现领导人节点101宕机,备份节点102在一定的时间内无法收到来自101的领导人宣告消息,备份节点102自动切换成为新领导人节点(或者,备份节点102会给核心节点集合内的其他核心节点发备份节点的切换消息,待收到其他核心节点103和104的回复之后,可自动成为新的领导人节点)。

[0074] 步骤6:候选节点持续监听许可链中的现领导人节点和备选节点的宣告信息,如在设定时间内未监听到现领导人节点和备选节点的宣告信息,一个或多个候选节点(认为许可链中缺少现领导人节点和备选节点)将自身状态变更为参选状态(Electing-Status)进而变更为参选节点;

[0075] 步骤7:参选节点向许可链中其他核心节点广播自己的选举宣告,所述选举宣告为包括节点能力和节点状态的广告消息,许可链将在参选节点中选出新领导人节点;

[0076] 任何收到选举宣告的核心节点根据所述预先设定策略或规则检查一个参选节点是否适合成为领导人节点,或者检查哪一个参选节点更合适成为领导人节点,进而向每个参选节点回复支持其当选新领导人的确认信息,或者向每个参选节点回复不支持其当选新领导人的拒绝信息,或者向所有参选节点回复投票支持哪个参选节点成为新领导人节点;

[0077] 如果一个参选节点收到候选节点集合(或核心节点集合)中2/3以上数量的核心节点的确认信息,该参选节点自动转换为领导人状态并成为新领导人节点,落选的参选节点自动将自身状态转变为候选状态(Candidate-Status),转变回为候选节点;

[0078] 新领导人节点检查当前核心节点集合信息,依据步骤2宣告核心节点集合信息,以及根据步骤3选出备选节点。

[0079] 其中,心跳消息(Heartbeat Message)是一种发送源发送到接收方的消息,这种消息可以让接收方确定发送源是否以及何时出现故障或终止。通常,心跳消息从发送源启动时开始发送,直到发送源关闭,期间发送源会不间断的发送周期性或重复消息。当接收方在某个消息接收周期内未收到消息,接收方可能会认为发送源已经关闭、出现故障、或者当前不可用。心跳消息可能常被用于高可用性或容错处理的目的。

[0080] 而能力消息是周期性广播式的消息,以反映节点能力的变化。能力消息和心跳消息可以在一起。

[0081] 例如,现领导人节点101和备份节点102都出现问题,候选节点103在一定的时间内没有监控到101的领导人宣告和102备份节点宣告,则认为许可链内没有领导人和备份节点,需要选举新的领导人节点,则核心节点103将自己的状态改为参选状态(Electing-Status),向许可链中其他核心节点广播自己的选举宣告。如果收到核心节点104的支持其当选领导人的确认信息,候选节点103则转换为新领导人节点,否则仍转换为候选状态。

[0082] 一个许可链由于网络或其他原因可能分裂为两个或多个子链,或者两个或多个许可链合并成一个联合链。一个许可链可分裂成若干个子链,所述子链分为:同时包括现领导人节点和备选节点的第一子链,包括现领导人节点而不包括备选节点的第二子链,包括备选节点而不包括现领导人节点的第三子链,以及既不包括现领导人节点又不包括备选节点

的第四子链;其中,子链中的领导人选定步骤包括:

[0083] 步骤8.1:第一子链依照步骤4-7进行领导人节点更迭;

[0084] 步骤8.2:第二子链中现领导人依照步骤3选出第二子链的备选节点,然后参照步骤4-7进行领导人节点更迭;

[0085] 步骤8.3:第三子链中,备份节点依照步骤5会切换成新领导人节点,新领导人节点根据步骤3选择新备份节点;

[0086] 步骤8.4:第四子链中,候选节点因检测不到现领导人节点和备份节点而根据步骤6和7选举出新领导人节点,新领导人节点根据步骤3选出备选节点。

[0087] 而两个以上的许可链合并形成一个联合链,联合链中包括两个以上的现领导人节点,现领导人节点将自身状态变更为参选状态进而成为参选节点,联合链根据步骤7选举出新领导人节点,新领导人节点根据步骤3选出备选节点,后参照步骤4-7进行领导人节点更迭。

[0088] 具体而言,现领导人节点作为参选节点向联合链中其他核心节点广播自己的选举宣告,联合链中任何收到能力宣告的核心节点根据所述预先设定策略或规则投票选出新领导人节点。其中节点可以为服务器、台式机、笔记本电脑或者手持式智能机等。

[0089] 本发明提供一种基于热备方式的区块链领导人选举方法,根据节点能力和既定策略预先确定备份节点,从而在到期后或者突发情况下快速实现领导人节点更迭,核心节点可快速转为参选节点,并通过投票方式快速选出新的领导人节点,防止许可链中领导人缺位。由此,领导人更迭速度快,效率高,同时可以选择较为复杂的节点选择规则,兼顾效果、效率和灵活性。

[0090] 实施例2

[0091] 本实施例公开了一种用于领导人选举的领导人选举装置,该选举装置可以服务器中的硬件单元,也可以是软件模组。从图2中可以看出,用于领导人选举的领导人选举装置200包括节点状态维护模块201、候选节点维护模块202、节点能力计算模块203、节点间通信模块204、系统信息模块205和配置策略管理模块206组成。

[0092] 节点状态维护模块201管理和维护核心节点的状态转换,依据预先设定策略或规则,从节点能力计算模块203获取许可链中节点的能力信息,并从候选节点维护模块202和节点间通信模块204获取核心节点的状态以及维护核心节点的状态转换。

[0093] 候选节点维护模块202主要负责维护链中所有候选节点的状态、能力、及配置策略等。候选节点维护模块会接收候选节点的能力宣告消息,添加并维护候选节点的能力信息;候选节点维护模块会接收来自配置管理模块206的策略及规则配置信息,以用于候选节点能力的排序;候选节点维护模块202根据各个候选节点的信息,及配置的策略和规则,对所有的候选节点进行排序,并把排序结果及各个候选节点的能力信息提供给节点状态维护模块201。

[0094] 节点能力计算模块203计算核心节点的能力信息,从配置管理模块206接收相应的能力参数配置,从系统信息模块205获取系统信息,计算各个核心节点的能力值,并提供给向候选节点维护模块202提供候选节点的能力信息,向节点状态维护模块201提供所有核心节点的状态信息。

[0095] 配置管理模块206提供人机交互界面,允许用于输入并配置节点能力信息数据、领

领导人选举的选择策略等信息,领导人选举的选择策略包括但不限于优先级策略、加权节点能力策略等等。这些信息提供给节点状态维护模块201、候选节点维护模块202和节点能力计算模块203使用。

[0096] 系统信息模块205负责与系统交互获取需要的系统信息,系统可以是Linux、Windows、OSX等等以及其他任何嵌入式系统。

[0097] 网络通信模块204负责节点之间的信息传输和交换,其中包括领导人宣告、能力宣告、备份节点宣告、领导人选举消息等等。网络通信模块204、节点状态维护模块201和候选节点维护模块202相互连接用于节点之间节点状态和信息的交互。

[0098] 实施例3

[0099] 本实施例公开了一种的领导人选举方法的主要流程。如图3所示,基于此流程,领导人选举装置可以自动维护节点的状态,可以在各个核心节点间自动的运行共识算法,从而保证系统中信息的一致性。该算法包括:

[0100] 步骤S1:任何一个新节点加入核心节点集合,其初始状态都是候选状态,候选节点首先收集自身的节点能力信息,并通过能力宣告发送到其他所有核心节点。

[0101] 步骤S2:候选节点监听许可链中的现领导人宣告(现领导人节点发出的宣告或广告信息)或备份节点宣告(备选节点发出的宣告或广告信息);

[0102] 步骤S2.1:判断是否收到现领导人宣告或备份节点宣告,如果收到则转向步骤S2.3;如果没收到,则转向步骤S2.2。

[0103] 步骤S2.2:候选节点和备选节点没有收到现领导人宣告,每个核心节点检查自己是否处于备选状态,即自己是否为备选节点,如果是,则备选节点执行步骤S3实现领导人切换。如果不是,核心节点处于候选状态,候选节点也没有收到备选节点的宣告,则执行步骤S3.1进行新领导人选举。

[0104] 步骤S2.3:候选节点收到现领导人宣告,则会检查是否收到现领导人节点的关于备份节点的指派消息,如果一个候选节点收到指派信息,该候选节点则执行步骤S2.5,将自身状态改为备选状态;如果没有收到指派信息,则执行步骤S2.4核心节点保持候选状态。

[0105] 步骤S4:每个核心节点周期性收集和更新节点的能力信息,并发送能力信息宣告。

[0106] 图4为核心节点状态间转换图;如图4所示,任何核心节点有四种运行状态,且这四种状态是在动态转变中的,这四种状态是:候选状态(Candidate-Status) 401、参选状态(Electing-Status) 402、领导人状态(Leader-Status) 403和以及备选状态(Standby-Status) 404。其中,参选状态由候选状态转变而成,即候选节点存在候选状态和参选状态。

[0107] 其中,候选状态表示该核心节点是纯粹的候选节点,任何新加入的节点,其初始化状态为候选状态。处于候选状态的节点接收任何领导人宣告或备份领导人宣告,保持候选状态。

[0108] 如果候选状态的核心节点收到现领导人节点的备份领导人指派消息,则切换到备选状态。

[0109] 如果处于候选状态的核心节点在一定的时间内未检测到现领导人宣告或备份节点的宣告,则进入参选状态并发出领导人选举请求。

[0110] 如果处于参选状态的核心节点收到的核心节点中支持回复不足2/3,则表明该节点不能成为新的领导人,则切换回候选状态。

[0111] 如果处于参选状态的核心节点如果收到的2/3以上的核心节点的支持的确认信息,则该核心节点将成为新的领导人,则该核心节点切换到领导人状态。

[0112] 处于备选状态的节点,如果检测到现领导人节点无响应,则发生领导人切换,切换到领导人状态。如果处于领导人状态的节点,发生领导人切换之后,状态切换回候选状态。

[0113] 图5为新领导者选举环节的流程图。如图5所示,如果一个核心节点没有监控到许可链内的领导宣告和备份领导人宣告,则认为许可链内没有领导人和备份领导人节点,需要选举新的领导人节点,触发新领导人选举过程。该过程如下:

[0114] 步骤M1: 候选节点将自己的状态改为参选状态,向链中其他候选节点广播自己的选举宣告。选举宣告是包括节点能力信息和状态信息的宣告消息。并等待其他候选节点的回复。

[0115] 步骤M2: 在选举宣告超时之前检查收到选举确认消息(即支持其成为新领导人节点的确认消息)的数量是否大于候选节点数量的2/3;

[0116] 步骤M3: 如果收到的选举确认消息数量在候选节点数量的2/3以上,则该候选节点可以成为新的领导人节点,并发出领导人宣告。

[0117] 步骤M4: 如果选举宣告超时,或者收到的选举确认消息数量少于候选节点数量的2/3,则该候选节点则不能成为新的领导人节点;

[0118] 此时检查是否收到新领导人宣告或备份领导人宣告消息,如果都没收到,则开始新一轮新领导人选举,该候选节点仍保持参选状态,重复步骤1-4直到新的领导人节点选出。

[0119] 步骤M5: 如处于参选状态的候选节点收到一个新领导人宣告或备份领导人宣告,则切换回候选状态。

[0120] 最后应说明的是:以上各实施例仅用以说明本发明的技术方案,而非对其限制;尽管参照前述各实施例对本发明进行了详细的说明,本领域的普通技术人员应当理解:其依然可以对前述各实施例所记载的技术方案进行修改,或者对其中部分或者全部技术特征进行等同替换;而这些修改或者替换,并不使相应技术方案的本质脱离本发明各实施例技术方案的范围。

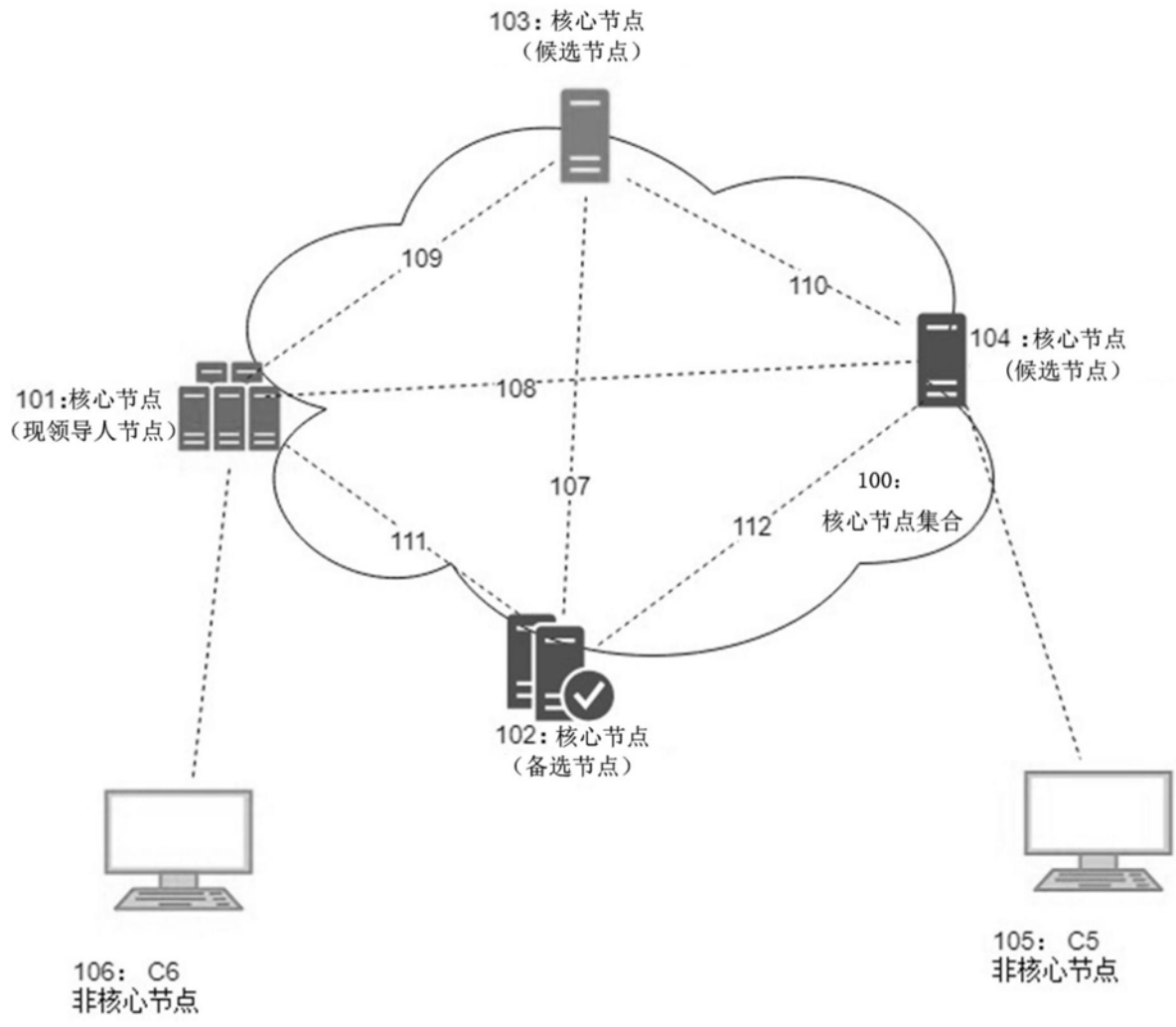


图1

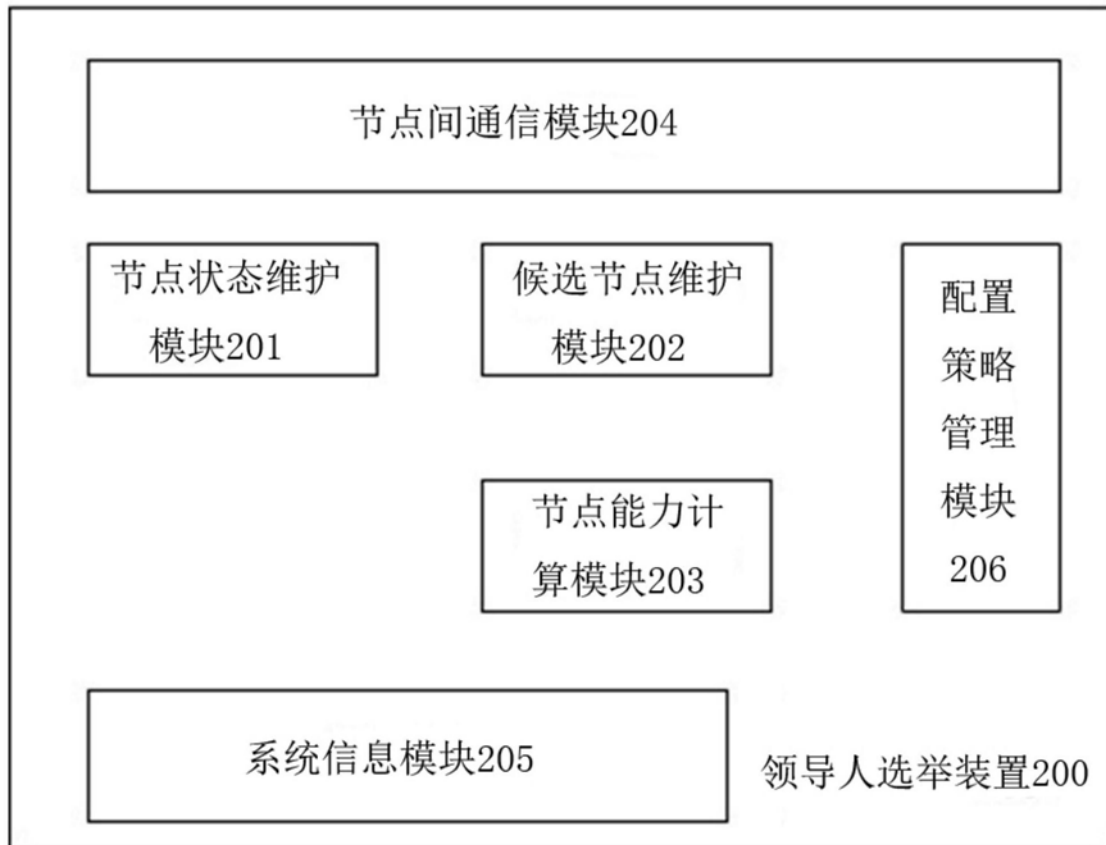


图2

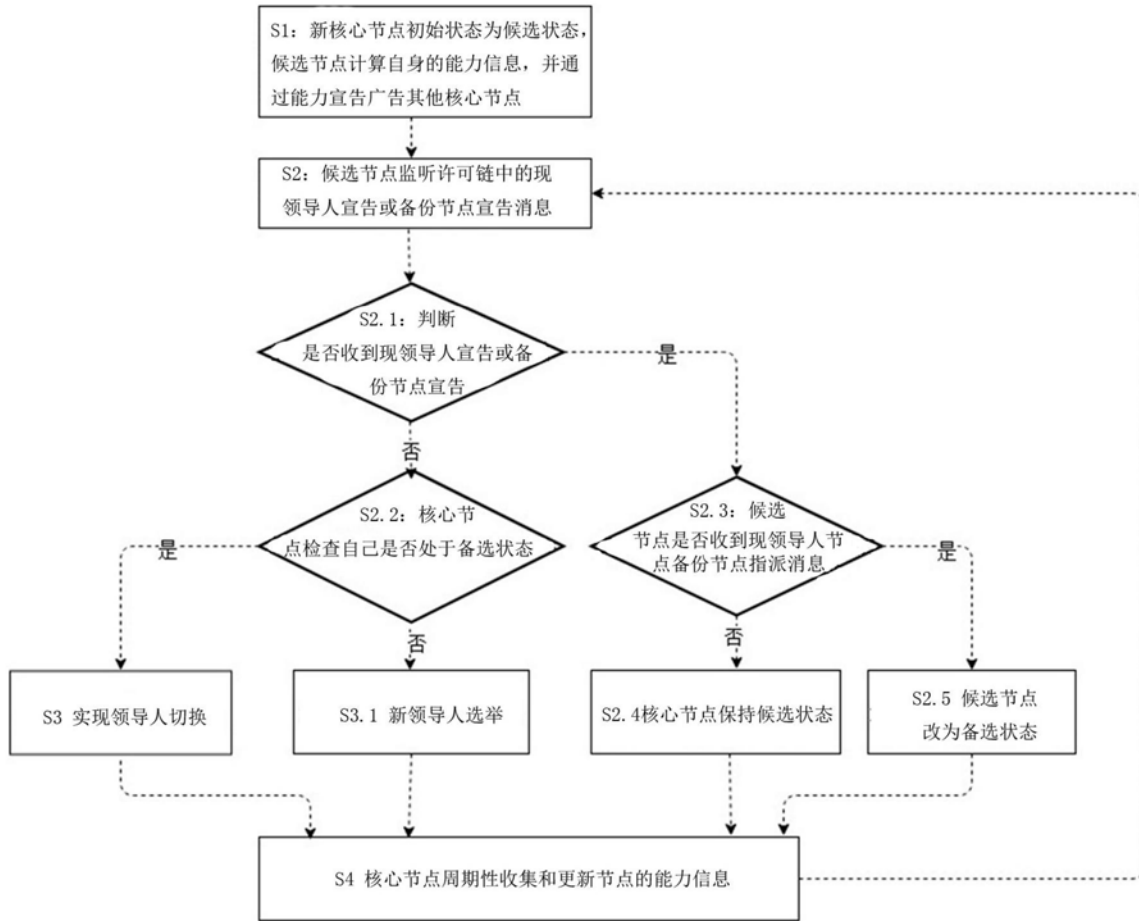


图3

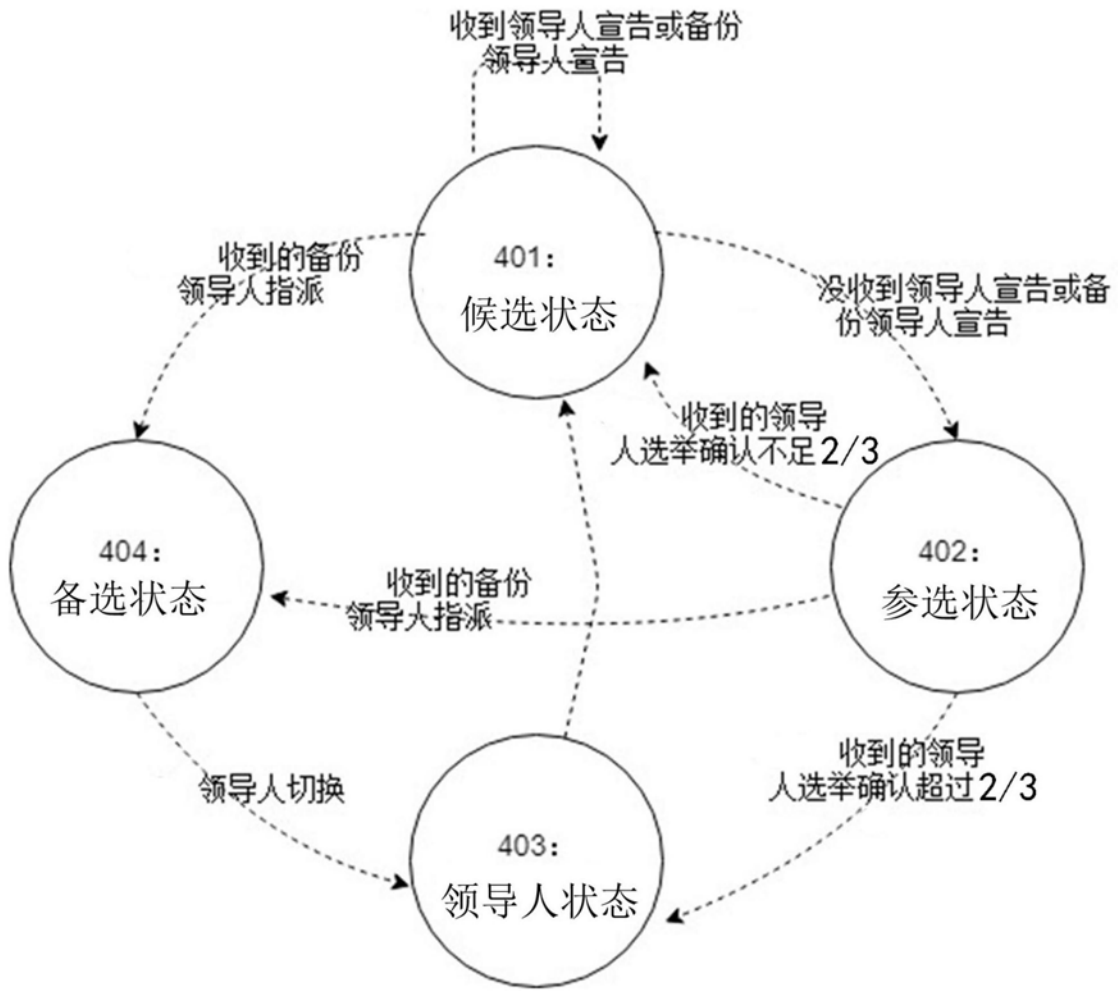


图4

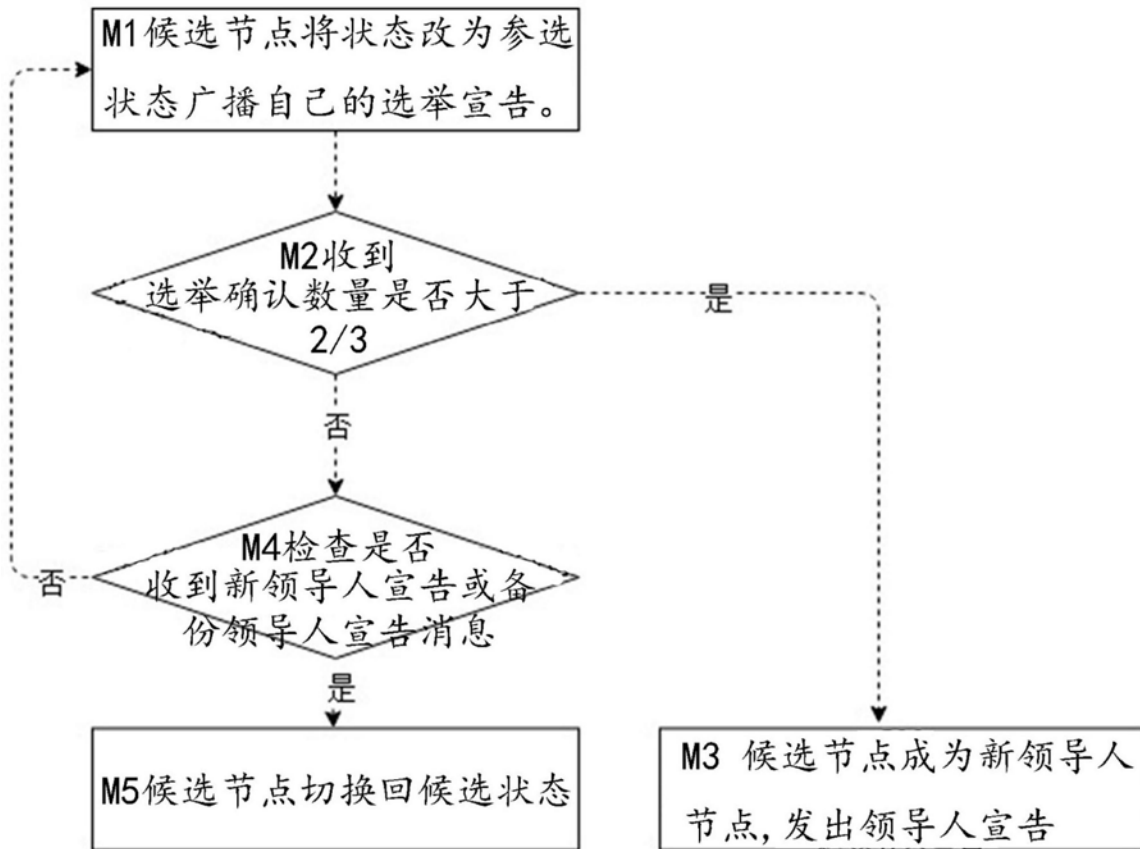


图5