



(12) 发明专利申请

(10) 申请公布号 CN 105589795 A

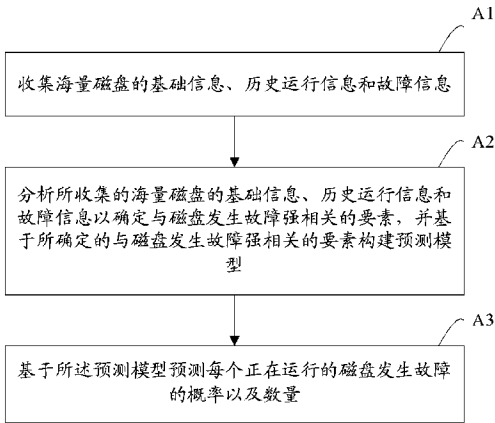
(43) 申请公布日 2016. 05. 18

(21) 申请号 201410845353. 1
(22) 申请日 2014. 12. 31
(71) 申请人 中国银联股份有限公司
地址 200135 上海市浦东新区含笑路 36 号
银联大厦
(72) 发明人 何东杰 张凌毅
(74) 专利代理机构 中国专利代理(香港)有限公司
司 72001
代理人 方世栋 汤春龙
(51) Int. Cl.
G06F 11/34(2006. 01)

权利要求书2页 说明书4页 附图2页

(54) 发明名称
基于预测模型的磁盘故障预测方法及装置

(57) 摘要
本发明提出了基于预测模型的磁盘故障预测方法及装置,其中,所述方法包括:(A1) 收集海量磁盘的基础信息、历史运行信息和故障信息;(A2) 分析所收集的海量磁盘的基础信息、历史运行信息和故障信息以确定与磁盘发生故障强相关的要素,并基于所确定的与磁盘发生故障强相关的要素构建预测模型;(A3) 基于所述预测模型预测每个正在运行的磁盘发生故障的概率以及数量。本发明所公开的基于预测模型的磁盘故障预测方法及装置能够预测磁盘故障发生概率以及数量。



1. 一种基于预测模型的磁盘故障预测方法,所述基于预测模型的磁盘故障预测方法包括下列步骤:

(A1) 收集海量磁盘的基础信息、历史运行信息和故障信息;

(A2) 分析所收集的海量磁盘的基础信息、历史运行信息和故障信息以确定与磁盘发生故障强相关的要素,并基于所确定的与磁盘发生故障强相关的要素构建预测模型;

(A3) 基于所述预测模型预测每个正在运行的磁盘发生故障的概率以及数量。

2. 根据权利要求 1 所述的基于预测模型的磁盘故障预测方法,其特征在于,所述预测模型是线性预测模型。

3. 根据权利要求 2 所述的基于预测模型的磁盘故障预测方法,其特征在于,所述方法还包括:周期性地检测并记录每个正在运行的磁盘的运行状态信息,并且在磁盘发生故障时记录故障发生时间、故障原因以及发生故障时该磁盘的运行状态。

4. 根据权利要求 3 所述的基于预测模型的磁盘故障预测方法,其特征在于,所述基础信息包括磁盘类型、磁盘生产商、磁盘出厂信息,所述历史运行信息包括运行时间、运行环境以及基于时间序列的运行状态信息,所述故障信息包括故障发生时间、故障原因。

5. 根据权利要求 4 所述的基于预测模型的磁盘故障预测方法,其特征在于,所述步骤(A2)进一步包括:分析磁盘的基础信息、历史运行信息和故障信息意之间的相关性以确定与磁盘发生故障强相关的要素,并将所确定的强相关要素作为采样点而构建所述线性预测模型。

6. 根据权利要求 5 所述的基于预测模型的磁盘故障预测方法,其特征在于,所述步骤(A3)进一步包括:获取各个磁盘的基础信息和当前运行信息并基于所述线性预测模型计算并预测每个磁盘未来一天或一周内发生故障的概率,并且基于所预测的概率确定未来一天或一周内可能会发生故障的磁盘的数量。

7. 一种基于预测模型的磁盘故障预测装置,所述基于预测模型的磁盘故障预测装置包括:

信息收集单元,所述信息收集单元收集海量磁盘的基础信息、历史运行信息和故障信息,并将所收集的海量磁盘的基础信息、历史运行信息和故障信息传送至模型构建单元;

模型构建单元,所述模型构建单元分析所述海量磁盘的基础信息、历史运行信息和故障信息以确定与磁盘发生故障强相关的要素,并基于所确定的与磁盘发生故障强相关的要素构建预测模型,以及将所述预测模型提供给故障预测单元;

故障预测单元,所述故障预测单元基于所述预测模型预测每个正在运行的磁盘发生故障的概率以及数量。

8. 根据权利要求 7 所述的基于预测模型的磁盘故障预测装置,其特征在于,所述预测模型是线性预测模型。

9. 根据权利要求 8 所述的基于预测模型的磁盘故障预测装置,其特征在于,所述装置还包括信息记录单元,所述信息记录单元周期性地检测并记录每个正在运行的磁盘的运行状态信息,并且在磁盘发生故障时记录故障发生时间、故障原因以及发生故障时该磁盘的运行状态。

10. 根据权利要求 9 所述的基于预测模型的磁盘故障预测装置,其特征在于,所述基础信息包括磁盘类型、磁盘生产商、磁盘出厂信息,所述历史运行信息包括运行时间、运行环

境以及基于时间序列的运行状态信息,所述故障信息包括故障发生时间、故障原因。

11. 根据权利要求 10 所述的基于预测模型的磁盘故障预测装置,其特征在于,所述模型构建单元以如下方式构建所述线性预测模型:分析磁盘的基础信息、历史运行信息和故障信息意之间的相关性以确定与磁盘发生故障强相关的要素,并将所确定的强相关要素作为采样点而构建所述线性预测模型。

12. 根据权利要求 11 所述的基于预测模型的磁盘故障预测装置,其特征在于,所述故障预测单元以如下方式预测每个正在运行的磁盘发生故障的概率以及数量:获取各个磁盘的基础信息和当前运行信息并基于所述线性预测模型计算并预测每个磁盘未来一天或一周内发生故障的概率,并且基于所预测的概率确定未来一天或一周内可能会发生故障的磁盘的数量。

基于预测模型的磁盘故障预测方法及装置

技术领域

[0001] 本发明涉及磁盘故障预测方法及装置,更具体地,涉及基于预测模型的磁盘故障预测方法及装置。

背景技术

[0002] 目前,随着计算机和网络应用的日益广泛以及不同领域的业务种类的日益丰富,基于海量数据存储的云计算技术的应用变得越来越广泛,由此,随着用于存储海量数据的机械式硬盘(例如 SAS 硬盘以及 SATA 硬盘)的数量的显著增长,其发生故障的概率以及不良影响也越来越大,因此,对磁盘故障进行处理和维护变得越来越重要。

[0003] 在现有的技术方案中,通常通过有限的冗余备份的方式来在磁盘发生故障的情况下确保数据处理和系统运行的连续性。

[0004] 然而,现有的技术方案存在如下问题:有限的冗余备份的方式仅能够在少量磁盘损坏的情况下保证数据处理和系统运行的连续性,而在较多磁盘发生故障的情况下,该方式难于保证系统的稳定运行和数据处理效率,此外,由于不能预估发生故障的磁盘的数量,因此难于准确地准备用于应急和备份的磁盘的数量,由此使得维护成本上升。

[0005] 因此,存在如下需求:提供能够预测磁盘故障发生概率以及数量的基于预测模型的磁盘故障预测方法及装置。

发明内容

[0006] 为了解决上述现有技术方案所存在的问题,本发明提出了能够预测磁盘故障发生概率以及数量的基于预测模型的磁盘故障预测方法及装置。

[0007] 本发明的目的是通过以下技术方案实现的:

一种基于预测模型的磁盘故障预测方法,所述基于预测模型的磁盘故障预测方法包括下列步骤:

(A1) 收集海量磁盘的基础信息、历史运行信息和故障信息;

(A2) 分析所收集的海量磁盘的基础信息、历史运行信息和故障信息以确定与磁盘发生故障强相关的要素,并基于所确定的与磁盘发生故障强相关的要素构建预测模型;

(A3) 基于所述预测模型预测每个正在运行的磁盘发生故障的概率以及数量。

[0008] 在上面所公开的方案中,优选地,所述预测模型是线性预测模型。

[0009] 在上面所公开的方案中,优选地,所述方法还包括:周期性地检测并记录每个正在运行的磁盘的运行状态信息,并且在磁盘发生故障时记录故障发生时间、故障原因以及发生故障时该磁盘的运行状态。

[0010] 在上面所公开的方案中,优选地,所述基础信息包括磁盘类型、磁盘生产商、磁盘出厂信息,所述历史运行信息包括运行时间、运行环境以及基于时间序列的运行状态信息,所述故障信息包括故障发生时间、故障原因。

[0011] 在上面所公开的方案中,优选地,所述步骤(A2)进一步包括:分析磁盘的基础信

息、历史运行信息和故障信息意之间的相关性以确定与磁盘发生故障强相关的要素,并将所确定的强相关要素作为采样点而构建所述线性预测模型。

[0012] 在上面所公开的方案中,优选地,所述步骤(A3)进一步包括:获取各个磁盘的基础信息和当前运行信息并基于所述线性预测模型计算并预测每个磁盘未来一天或一周内发生故障的概率,并且基于所预测的概率确定未来一天或一周内可能会发生故障的磁盘的数量。

[0013] 本发明的目的也可以通过以下技术方案实现:

一种基于预测模型的磁盘故障预测装置,所述基于预测模型的磁盘故障预测装置包括:

信息收集单元,所述信息收集单元收集海量磁盘的基础信息、历史运行信息和故障信息,并将所收集的海量磁盘的基础信息、历史运行信息和故障信息传送至模型构建单元;

模型构建单元,所述模型构建单元分析所述海量磁盘的基础信息、历史运行信息和故障信息以确定与磁盘发生故障强相关的要素,并基于所确定的与磁盘发生故障强相关的要素构建预测模型,以及将所述预测模型提供给故障预测单元;

故障预测单元,所述故障预测单元基于所述预测模型预测每个正在运行的磁盘发生故障的概率以及数量。

[0014] 在上面所公开的方案中,优选地,所述预测模型是线性预测模型。

[0015] 在上面所公开的方案中,优选地,所述装置还包括信息记录单元,所述信息记录单元周期性地检测并记录每个正在运行的磁盘的运行状态信息,并且在磁盘发生故障时记录故障发生时间、故障原因以及发生故障时该磁盘的运行状态。

[0016] 在上面所公开的方案中,优选地,所述基础信息包括磁盘类型、磁盘生产商、磁盘出厂信息,所述历史运行信息包括运行时间、运行环境以及基于时间序列的运行状态信息,所述故障信息包括故障发生时间、故障原因。

[0017] 在上面所公开的方案中,优选地,所述模型构建单元以如下方式构建所述线性预测模型:分析磁盘的基础信息、历史运行信息和故障信息意之间的相关性以确定与磁盘发生故障强相关的要素,并将所确定的强相关要素作为采样点而构建所述线性预测模型。

[0018] 在上面所公开的方案中,优选地,所述故障预测单元以如下方式预测每个正在运行的磁盘发生故障的概率以及数量:获取各个磁盘的基础信息和当前运行信息并基于所述线性预测模型计算并预测每个磁盘未来一天或一周内发生故障的概率,并且基于所预测的概率确定未来一天或一周内可能会发生故障的磁盘的数量。

[0019] 本发明所公开的基于预测模型的磁盘故障预测方法及装置具有以下优点:能够预测磁盘故障发生概率以及数量,由此能够确保系统的稳定运行和数据处理的效率,并且能够预估发生故障的磁盘的数量,从而能够更准确地准备用于应急和备份的磁盘,使得维护便捷且成本较低。

附图说明

[0020] 结合附图,本发明的技术特征以及优点将会被本领域技术人员更好地理解,其中:

图1是根据本发明的实施例的基于预测模型的磁盘故障预测方法的流程图;

图 2 是根据本发明的实施例的基于预测模型的磁盘故障预测装置的示意性结构图。

具体实施方式

[0021] 图 1 是根据本发明的实施例的基于预测模型的磁盘故障预测方法的流程图。如图 1 所示,本发明所公开的基于预测模型的磁盘故障预测方法包括下列步骤:(A1) 收集海量磁盘的基础信息、历史运行信息和故障信息;(A2) 分析所收集的海量磁盘的基础信息、历史运行信息和故障信息以确定与磁盘发生故障强相关的要素,并基于所确定的与磁盘发生故障强相关的要素构建预测模型;(A3) 基于所述预测模型预测每个正在运行的磁盘发生故障的概率以及数量。

[0022] 优选地,在本发明所公开的基于预测模型的磁盘故障预测方法中,所述预测模型是线性预测模型。

[0023] 优选地,本发明所公开的基于预测模型的磁盘故障预测方法还包括:周期性地(例如 10 秒或者 1 分钟)检测并记录每个正在运行的磁盘的运行状态信息,并且在磁盘发生故障时记录故障发生时间、故障原因以及发生故障时该磁盘的运行状态。

[0024] 优选地,在本发明所公开的基于预测模型的磁盘故障预测方法中,所述基础信息包括但不限于磁盘类型、磁盘生产商、磁盘出厂信息,所述历史运行信息包括但不限于运行时间、运行环境(即使用场景,例如,如果是海量数据处理环境则涉及大量的读写操作,如果是虚拟化环境,则仅涉及少量的读写操作)以及基于时间序列的运行状态信息,所述故障信息包括但不限于故障发生时间、故障原因。

[0025] 优选地,在本发明所公开的基于预测模型的磁盘故障预测方法中,所述步骤(A2)进一步包括:分析磁盘的基础信息、历史运行信息和故障信息意之间的相关性以确定与磁盘发生故障强相关的要素(例如运行时间、访问频度等等),并将所确定的强相关要素作为采样点而构建所述线性预测模型(例如,所述线性预测模型的时间单位为“天”)。

[0026] 优选地,在本发明所公开的基于预测模型的磁盘故障预测方法中,所述步骤(A3)进一步包括:获取各个磁盘的基础信息和当前运行信息并基于所述线性预测模型计算并预测每个磁盘未来一天或一周内发生故障的概率,并且基于所预测的概率确定未来一天或一周内可能会发生故障的磁盘的数量(例如发生故障的概率大于预定阈值的磁盘的数量)。

[0027] 由上可见,本发明所公开的基于预测模型的磁盘故障预测方法具有下列优点:能够预测磁盘故障发生概率以及数量,由此能够确保系统的稳定运行和数据处理的效率,并且能够预估发生故障的磁盘的数量,从而能够更准确地准备用于应急和备份的磁盘,使得维护便捷且成本较低。

[0028] 图 2 是根据本发明的实施例的基于预测模型的磁盘故障预测装置的示意性结构图。如图 2 所示,本发明所公开的基于预测模型的磁盘故障预测装置包括信息收集单元 1、模型构建单元 2 和故障预测单元 3。所述信息收集单元 1 收集海量磁盘的基础信息、历史运行信息和故障信息,并将所收集的海量磁盘的基础信息、历史运行信息和故障信息传送至模型构建单元 2。所述模型构建单元 2 分析所述海量磁盘的基础信息、历史运行信息和故障信息以确定与磁盘发生故障强相关的要素,并基于所确定的与磁盘发生故障强相关的要素构建预测模型,以及将所述预测模型提供给故障预测单元 3。所述故障预测单元 3 基于所述预测模型预测每个正在运行的磁盘发生故障的概率以及数量。

[0029] 优选地,在本发明所公开的基于预测模型的磁盘故障预测装置中,所述预测模型是线性预测模型。

[0030] 优选地,本发明所公开的基于预测模型的磁盘故障预测装置还包括信息记录单元 4,所述信息记录单元 4 周期性地(例如 10 秒或者 1 分钟)检测并记录每个正在运行的磁盘的运行状态信息,并且在磁盘发生故障时记录故障发生时间、故障原因以及发生故障时该磁盘的运行状态。

[0031] 优选地,在本发明所公开的基于预测模型的磁盘故障预测装置中,所述基础信息包括但不限于磁盘类型、磁盘生产商、磁盘出厂信息,所述历史运行信息包括但不限于运行时间、运行环境(即使用场景,例如,如果是海量数据处理环境则涉及大量的读写操作,如果是虚拟化环境,则仅涉及少量的读写操作)以及基于时间序列的运行状态信息,所述故障信息包括但不限于故障发生时间、故障原因。

[0032] 优选地,在本发明所公开的基于预测模型的磁盘故障预测装置中,所述模型构建单元 2 以如下方式构建所述线性预测模型:分析磁盘的基础信息、历史运行信息和故障信息意之间的相关性以确定与磁盘发生故障强相关的要素(例如运行时间、访问频度等等),并将所确定的强相关要素作为采样点而构建所述线性预测模型(例如,所述线性预测模型的时间单位为“天”)。

[0033] 优选地,在本发明所公开的基于预测模型的磁盘故障预测装置中,所述故障预测单元 3 以如下方式预测每个正在运行的磁盘发生故障的概率以及数量:获取各个磁盘的基础信息和当前运行信息并基于所述线性预测模型计算并预测每个磁盘未来一天或一周内发生故障的概率,并且基于所预测的概率确定未来一天或一周内可能会发生故障的磁盘的数量(例如发生故障的概率大于预定阈值的磁盘的数量)。

[0034] 由上可见,本发明所公开的基于预测模型的磁盘故障预测装置具有下列优点:能够预测磁盘故障发生概率以及数量,由此能够确保系统的稳定运行和数据处理的效率,并且能够预估发生故障的磁盘的数量,从而能够更准确地准备用于应急和备份的磁盘,使得维护便捷且成本较低。

[0035] 尽管本发明是通过上述的优选实施方式进行描述的,但是其实现形式并不局限于上述的实施方式。应该认识到:在不脱离本发明主旨和范围的情况下,本领域技术人员可以对本发明做出不同的变化和修改。

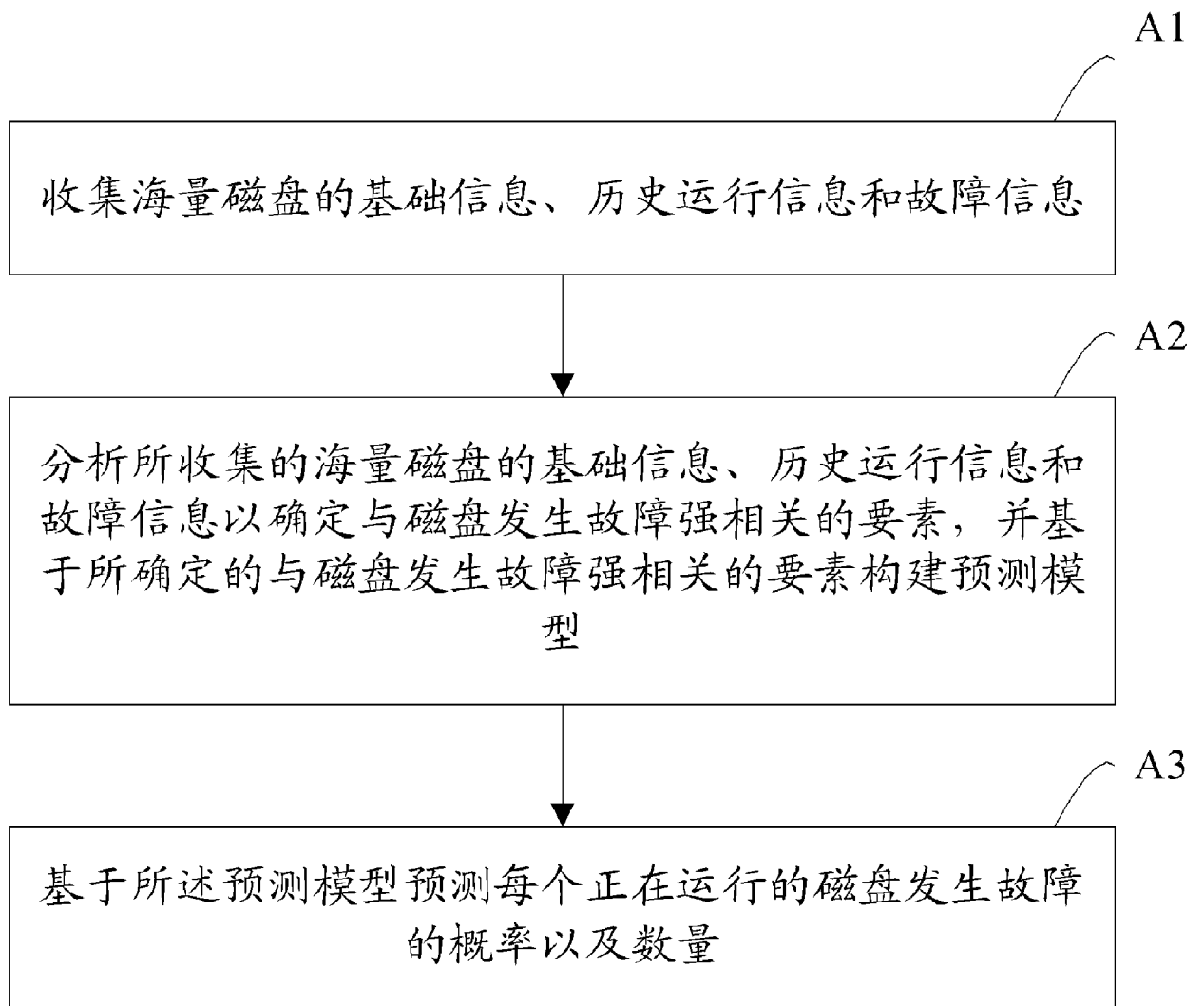


图 1

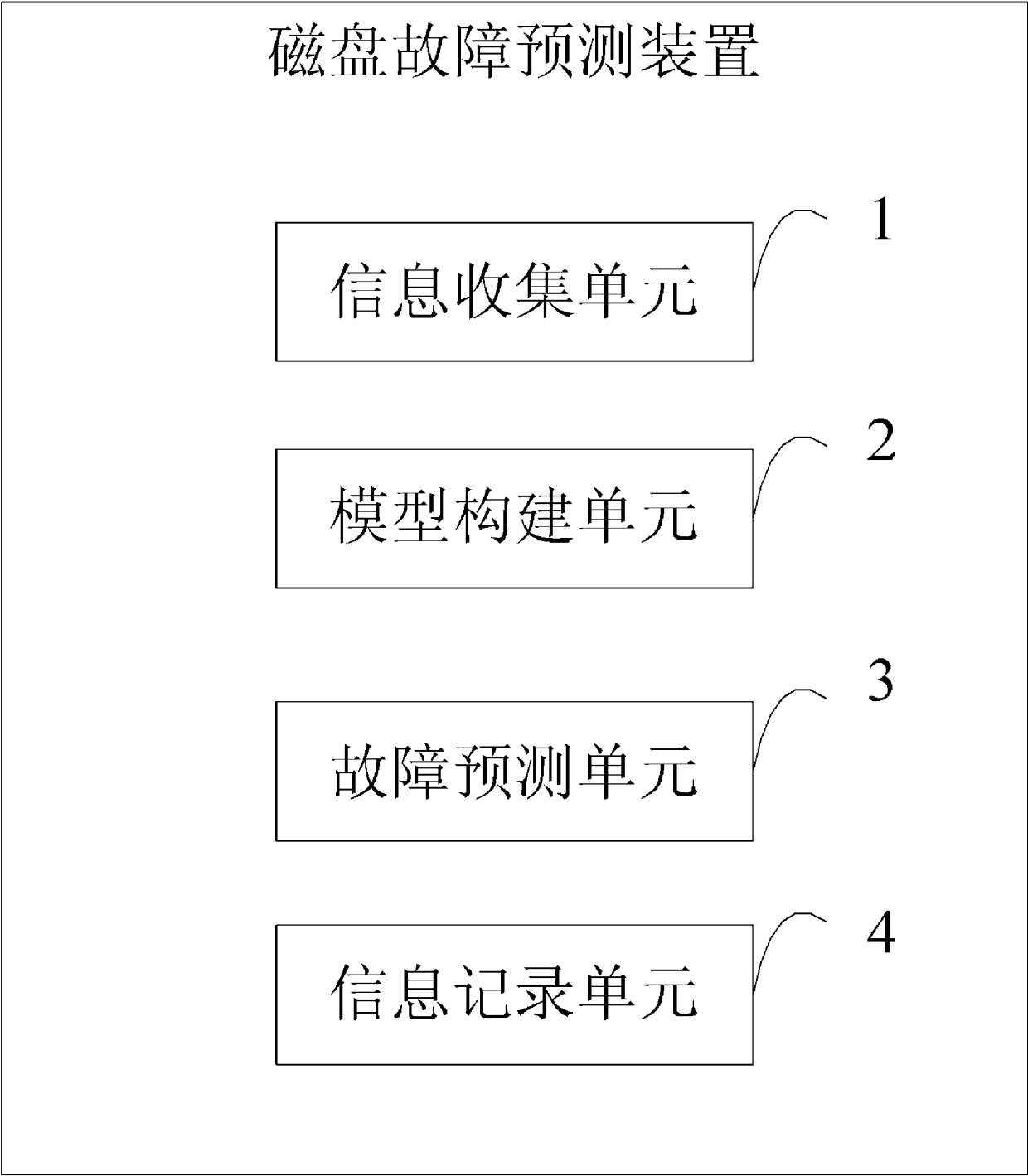


图 2