



(12) 发明专利申请

(10) 申请公布号 CN 117413238 A

(43) 申请公布日 2024. 01. 16

(21) 申请号 202280038926.2

(22) 申请日 2022.01.21

(30) 优先权数据

63/223,491 2021.07.19 US

(85) PCT国际申请进入国家阶段日

2023.11.29

(86) PCT国际申请的申请数据

PCT/US2022/013319 2022.01.21

(87) PCT国际申请的公布数据

W02023/003595 EN 2023.01.26

(71) 申请人 萨思学会有限公司

地址 美国北卡罗来纳州

(72) 发明人 D·V·卡克德 王浩宇

A·M·麦吉尔克

(74) 专利代理机构 北京律盟知识产权代理有限公司

责任公司 11287

专利代理师 章蕾

(51) Int.Cl.

G06E 1/00 (2006.01)

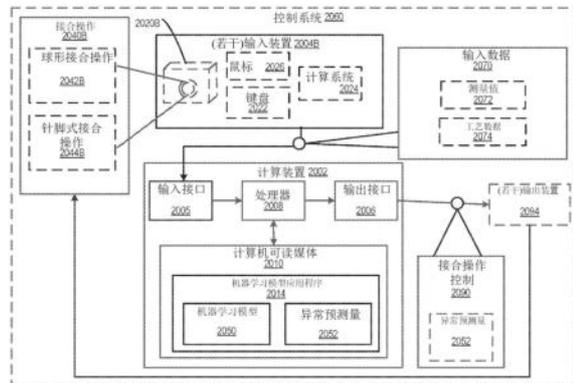
权利要求书5页 说明书41页 附图33页

(54) 发明名称

使用工艺数据的质量预测

(57) 摘要

计算装置 (2002) 存取在第一接合操作 (1308, 2040A) (例如, 球形及/或针脚式接合) 的训练数据 (2032) 上训练的机器学习模型 (2050)。所述第一接合操作包括用以将第一组导线 (1504) 接合到第一组表面 (1506, 1508) 的操作。通过监督式学习来训练所述机器学习模型。所述装置接收指示从第二接合操作 (2040B) 的测量值产生的工艺数据 (2074) 的输入数据 (2070)。所述第二接合操作包括用以将第二组导线接合到第二组表面的操作。所述装置根据所述机器学习模型对所述输入数据进行加权。所述装置基于根据所述机器学习模型对所述输入数据进行加权来产生指示所述第二接合操作中的异常发生的风险的异常预测量 (2052)。所述装置输出所述异常预测量以控制所述第二接合操作。



1. 一种有形地体现于非暂时性机器可读存储媒体中的计算机程序产品,所述计算机程序产品包含指令,所述指令能够操作以使计算系统(2002、2060):

存取在第一接合操作(1308、2040A)的训练数据(2032)上训练的机器学习模型(2050),其中所述第一接合操作包括用以将第一组的多条导线(1504)接合到第一组表面(1506、1508)的操作,且其中通过监督式学习来训练所述机器学习模型,包括:

接收所述训练数据(2032),其中所述训练数据包括:

从所述第一接合操作(2040A)的测量值产生的工艺数据(2034);及

所述多条导线在接合到所述第一组表面后的状态(2036),所述状态的每一状态包括与检测所述第一接合操作中之一或多个异常有关的目标之一或多个候选结果;及

产生所述工艺数据之一或多个权重,使得输入到所述机器学习模型的所述工艺数据预测所述目标的所述之一或多个候选结果;

接收指示从第二接合操作(2040B)的测量值产生的工艺数据(2074)的输入数据(2070),其中所述第二接合操作包括用以将第二组的多条导线接合到第二组表面的操作,其中所述第二组的多条导线不同于所述第一组的多条导线,且其中所述第二组表面不同于所述第一组表面;

根据所述机器学习模型对所述输入数据进行加权;

基于根据所述机器学习模型对所述输入数据进行加权来产生指示所述第二接合操作中的异常发生的风险的异常预测量(2052);及

输出所述异常预测量以控制所述第二接合操作。

2. 根据权利要求1所述的计算机程序产品,

其中所述之一或多个候选结果包括对所述第一接合操作中的个别导线(1504)或接合件(1510、1512)之一或多个破坏性质量保证测试(1800、1900);且

其中所述异常预测量(2052)是所述第二接合操作中的个别导线或接合件中的异常的预测量,而未对所述个别导线或接合件执行破坏性质量保证测试。

3. 根据权利要求2所述的计算机程序产品,其中所述第一接合操作包括以下中的一者:

球形接合操作,且其中所述破坏性质量保证测试包括用于测试球形接合件(1510)的球剪切测试(1800);及

针脚式接合操作,且其中所述破坏性质量保证测试包括用于测试针脚式接合件(1512)的针脚拉拔测试(1900)。

4. 根据权利要求1所述的计算机程序产品,

其中所述第一接合操作将所述第一组的多条导线的导线(1504)接合到所述第一组表面的对应表面以形成集成电路芯片(1700);

其中所述之一或多个候选结果包括所述第一接合操作中的所述集成电路芯片之一或多个有缺陷芯片结果;且

其中所述指令能够操作以使所述计算系统产生所述第二接合操作中的集成电路芯片制造过程中的所述异常的风险的所述异常预测量。

5. 根据权利要求1所述的计算机程序产品,其中所述之一或多个异常与以下中的一或多个者相关联:

浮动状况;

引线框架污染;及
裸片倾斜。

6. 根据权利要求1所述的计算机程序产品,

其中所述第二接合操作接合所述第二组的多条导线的导线以形成集成电路芯片(1700);且

其中从所述第二接合操作的所述测量值产生的所述工艺数据整体上与所述集成电路芯片的特定芯片有关,且从与所述特定芯片相关联且在所述第二接合操作中接合的导线的测量数据导出。

7. 根据权利要求1所述的计算机程序产品,

其中所述输入数据包括在所述第二接合操作期间接收的实时传感器测量值(2200);且

其中针对所述第二组的多条导线的给定导线,所述实时传感器测量值包括以下中的一或多者:

热量测量值;

功率测量值;

力测量值;

放电结球(EFO)测量值;及

超声波测量值。

8. 根据权利要求1所述的计算机程序产品,

其中所述输入数据包括在所述第二接合操作期间接收的实时传感器测量值(2200);

其中所述传感器测量值包括在所述第二接合操作中涉及的接合系统的测量值;及

其中所述异常预测量控制所述第二接合操作以校正所述第二接合操作中涉及的所述接合系统中的所述一或多个异常或减少所述一或多个异常的所述发生。

9. 根据权利要求1所述的计算机程序产品,

其中所述输入数据包括用指示在所述第二接合操作中涉及的特定导线、裸片或芯片的标识符或位置中的一或多者的原始信息标记的经接收传感器测量值;且

其中所述异常预测量识别在所述第二接合操作中发生的所述异常,且与所述原始信息相关以指示所述异常的位置。

10. 根据权利要求1所述的计算机程序产品,其中所述指令能够操作以使所述计算系统:

接收指示所述异常预测量正确地或不正确地预测在所述第二接合操作中制造的特定芯片中的所述异常的反馈;及

基于所述反馈更新所述机器学习模型。

11. 根据权利要求1所述的计算机程序产品,

其中所述第二接合操作由芯片制造系统执行;且

其中所述指令能够操作以使所述计算系统基于以下中的一或多者来调整继由所述芯片制造系统进行的所述第二接合操作后的接合操作:

指示所述第二接合操作中的所述异常发生的所述风险的所述异常预测量;及

指示所述异常预测量正确地或不正确地预测在所述第二接合操作中制造的特定芯片中的所述异常发生的反馈。

12. 根据权利要求1所述的计算机程序产品,其中所述输入数据包含包括以下中的一或多者的经导出数据:

指示与在所述第二接合操作中接合于特定芯片(1700)中的多条导线有关的测量值的中值(2353)或平均值的经产生值;

包括与所述多条导线中的每一者的所述值的偏差的一组经产生偏差(2354);及考虑所述一组偏差的所述芯片的经产生度量。

13. 根据权利要求1所述的计算机程序产品,

其中所述训练数据包括从经导出数据产生的所述工艺数据,所述经导出数据包括与多种不同类型的测量值相关且与所述第一接合操作中的相同导线有关的经产生奇异数据值;且

其中指示从所述第二接合操作的测量值产生的所述工艺数据的所述输入数据包括经产生奇异数据值,其中所述经产生奇异数据值与所述多种不同类型的测量值相关且与所述第二接合操作中的相同导线有关。

14. 根据权利要求1所述的计算机程序产品,

其中所述训练数据包括通过从多种不同测量类型导出考虑所述第一接合操作中的测量类型之间的关系的信息而产生的所述工艺数据;及

其中所述指令能够操作以接收指示通过从多种测量类型导出考虑所述第二接合操作中的测量类型之间的关系的信息而从所述第二接合操作的测量值产生的所述工艺数据的所述输入数据。

15. 根据权利要求1所述的计算机程序产品,

其中所述指令能够操作以使所述计算系统通过选择性地选取在所述第一接合操作中观察的参数类型的子集来接收所述训练数据;且

其中所述第一接合操作的所述测量值是针对参数类型的所述子集的测量值。

16. 根据权利要求1所述的计算机程序产品,其中所述工艺数据的所述一或多个权重经产生用于所述训练数据的梯度提升模型。

17. 根据权利要求1所述的计算机程序产品,其中所述机器学习模型进一步由多个经产生的机器学习模型训练且基于k折交叉验证来选择。

18. 根据权利要求1所述的计算机程序产品,其中所述第一及第二接合操作的所述测量值包括与形成球形接合件(1510)的工艺相关联的测量值。

19. 根据权利要求1所述的计算机程序产品,其中所述第一及第二接合操作的所述测量值包括与形成针脚式接合件(1512)的工艺相关联的测量值。

20. 根据权利要求1所述的计算机程序产品,其中所述异常预测量是所述第二组的多条导线的导线与所述第二组表面的引线框架或裸片之间的有缺陷接合件的预测量。

21. 一种计算机实施方法(2130),其包括:

存取(2131)在第一接合操作(1308、2040A)的训练数据上训练的机器学习模型(2050),其中所述第一接合操作包括用以将第一组的多条导线(1504)接合到第一组表面(1506、1508)的操作,且其中通过监督式学习来训练所述机器学习模型,包括:

接收(2101)所述训练数据(2032),其中所述训练数据包括:

从所述第一接合操作(2040A)的测量值产生的工艺数据(2034);及

所述多条导线在接合到所述第一组表面后的状态(2036),所述状态的每一状态包括与检测所述第一接合操作中之一或多个异常有关的目标之一或多个候选结果;及

产生(2102)所述工艺数据之一或多个权重,使得输入到所述机器学习模型的所述工艺数据预测所述目标的所述之一或多个候选结果;

接收(2132)指示从第二接合操作(2040B)的测量值产生的工艺数据(2074)的输入数据(2070),其中所述第二接合操作包括用以将第二组的多条导线接合到第二组表面的操作,其中所述第二组的多条导线不同于所述第一组的多条导线,且其中所述第二组表面不同于所述第一组表面;

根据所述机器学习模型对所述输入数据进行加权(2133);

基于根据所述机器学习模型对所述输入数据进行加权来产生(2134)指示所述第二接合操作中的异常发生的风险的异常预测量(2052);及

输出(2135)所述异常预测量以控制所述第二接合操作。

22. 根据权利要求21所述的计算机实施方法,

其中所述之一或多个候选结果包括对所述第一接合操作中的个别导线(1504)或接合件(1510、1512)之一或多个破坏性质量保证测试(1800、1900);且

其中所述异常预测量(2052)是所述第二接合操作中的个别导线或接合件中的异常的预测量,而未对所述个别导线或接合件执行破坏性质量保证测试。

23. 根据权利要求21所述的计算机实施方法,

其中所述第一接合操作接合所述第一组的多条导线的导线(1504)以形成集成电路芯片(1700);

其中所述之一或多个候选结果包括所述第一接合操作中的所述集成电路芯片之一或多个有缺陷芯片结果;且

其中产生所述异常预测量包括产生在所述第二接合操作中制造的集成电路芯片中的所述异常的风险的所述异常预测量。

24. 根据权利要求21所述的计算机实施方法,

其中所述第二接合操作接合所述第二组的多条导线的导线以形成集成电路芯片(1700);且

其中从所述第二接合操作的所述测量值产生的所述工艺数据整体上与所述集成电路芯片的特定芯片有关,且从与所述特定芯片相关联且在所述第二接合操作中接合的导线的测量数据导出。

25. 根据权利要求21所述的计算机实施方法,

其进一步包括在所述第二接合操作期间接收实时传感器测量值(2200);及

其中针对所述第二组的多条导线的给定导线(1504),所述实时传感器测量值包括以下中之一或多个者:

热量测量值;

功率测量值;

力测量值;

放电结球(EFO)测量值;及

超声波测量值。

26. 根据权利要求21所述的计算机实施方法，
其中所述接收输入数据包括在所述第二接合操作期间接收实时传感器测量值(2200)，且其中所述传感器测量值包括在所述第二接合操作中涉及的接合系统的测量值；且
其中所述异常预测量控制所述第二接合操作以校正所述第二接合操作中涉及的所述接合系统中的所述一或多个异常或减少所述一或多个异常的所述发生。
27. 根据权利要求21所述的计算机实施方法，
其中所述输入数据包括用指示在所述第二接合操作中涉及的特定导线、裸片或芯片的标识符或位置中的一或多者的原始信息标记的经接收传感器测量值；且
其中所述异常预测量识别在所述第二接合操作中发生的所述异常，且与所述原始信息相关以指示所述异常的位置。
28. 根据权利要求21所述的计算机实施方法，其进一步包括：
接收(2161)指示所述异常预测量正确地或不正确地预测在所述第二接合操作中制造的特定芯片中的所述异常的反馈；及
基于所述反馈更新(2162)所述机器学习模型。
29. 根据权利要求21所述的计算机实施方法，其中所述输入数据包含包括以下中的一或多者的经导出数据：
指示与在所述第二接合操作中接合于特定芯片中的多条导线有关的测量值的中值(2353)或平均值的经产生值；
包括与所述多条导线中的每一者的所述值的偏差的一组经产生偏差(2354)；及
考虑所述一组偏差的所述特定芯片(1700)的经产生度量。
30. 一种包括处理器(2008)及存储器(2010)的计算装置(2002)，所述存储器含有可由所述处理器执行的指令，其中所述计算装置经配置以：
存取在第一接合操作(1308、2040A)的训练数据上训练的机器学习模型(2050)，其中所述第一接合操作包括用以将第一组的多条导线(1504)接合到第一组表面(1506、1508)的操作，且其中通过监督式学习来训练所述机器学习模型，包括：
接收所述训练数据(2032)，其中所述训练数据包括：
从所述第一接合操作(2040A)的测量值产生的工艺数据(2034)；及
所述多条导线在接合到所述第一组表面后的状态(2036)，所述状态的每一状态包括与检测所述第一接合操作中的一或多个异常有关的目标的一或多个候选结果；及
产生所述工艺数据的一或多个权重，使得输入到所述机器学习模型的所述工艺数据预测所述目标的所述一或多个候选结果；
接收指示从第二接合操作(2040B)的测量值产生的工艺数据(2074)的输入数据(2070)，其中所述第二接合操作包括用以将第二组的多条导线接合到第二组表面的操作，其中所述第二组的多条导线不同于所述第一组的多条导线，且其中所述第二组表面不同于所述第一组表面；
根据所述机器学习模型对所述输入数据进行加权；
基于根据所述机器学习模型对所述输入数据进行加权来产生指示所述第二接合操作中的异常发生的风险的异常预测量(2052)；及
输出所述异常预测量以控制所述第二接合操作。

使用工艺数据的质量预测

[0001] 相关申请的交叉参考

[0002] 本申请基于35 U.S.C. §119规定主张2021年7月19日申请的第63/223491号美国临时申请案及2022年1月21日申请的第17/581,113号美国非临时申请案的权利及优先权,所述申请案中的每一者的公开内容的全文以引用的方式并入本文中。

技术领域

[0003] 本公开大体上涉及用于接合操作的计算机产生的预测。

背景技术

[0004] 例如制造及建设的工业使用接合技术来将材料结合在一起(例如,熔接技术)。质量保证测试可用于确定接合件的质量。例如,在破坏性测试中,破坏接合件的子集以进行关于未破坏的接合件的质量的预测。在常规非破坏性测试中,测试员手动检验接合件以进行关于接合件的质量的预测。

发明内容

[0005] 在实例实施例中,提供一种有形地体现于非暂时性机器可读存储媒体中的计算机程序产品。所述计算机程序产品包含可操作以使计算系统存取在第一接合操作的训练数据上训练的机器学习模型的指令。所述第一接合操作包括用以将第一组的多条导线接合到第一组表面的操作。通过监督式学习来训练所述机器学习模型,包括接收所述训练数据。所述训练数据包括从所述第一接合操作的测量值产生的工艺数据;及所述多条导线在接合到所述第一组表面后的状态。所述状态的每一状态包括与检测所述第一接合操作中之一或多个异常有关的目标之一或多个候选结果。通过监督式学习来训练所述机器学习模型,包括产生所述工艺数据之一或多个权重,使得输入到所述机器学习模型的所述工艺数据预测所述目标的所述之一或多个候选结果。所述计算机程序产品包含可操作以使计算系统接收指示从第二接合操作的测量值产生的工艺数据的输入数据的指令。所述第二接合操作包括用以将第二组的多条导线接合到第二组表面的操作。所述第二组的多条导线不同于所述第一组的多条导线。所述第二组表面不同于所述第一组表面。所述计算机程序产品包含可操作以使计算系统根据所述机器学习模型对所述输入数据进行加权的指令。所述计算机程序产品包含可操作以使计算系统基于根据所述机器学习模型对所述输入数据进行加权来产生指示所述第二接合操作中的异常发生的风险的异常预测量(predictor)的指令。所述计算机程序产品包含可操作以使计算系统输出所述异常预测量以控制所述第二接合操作的指令。

[0006] 在一或多个实施例中,第一接合操作将第一组的多条导线的导线接合到第一组表面的对应表面以形成集成电路芯片。一或多个候选结果包括第一接合操作中的集成电路芯片之一或多个有缺陷芯片结果。指令能够操作以使计算系统产生第二接合操作中的集成电路芯片制造过程中的异常的风险的异常预测量。

[0007] 在一或多个实施例中,指令能够操作以使计算系统接收指示异常预测量正确地或

不正确地预测在第二接合操作中制造的特定芯片中的异常的反馈。指令能够操作以使计算系统基于反馈更新机器学习模型。

[0008] 在一或多个实施例中,第二接合操作由芯片制造系统执行。指令能够操作以使计算系统基于以下中的一或多者来调整继由芯片制造系统进行的第二接合操作后的接合操作:指示第二接合操作中的异常发生的风险的异常预测量;及指示异常预测量正确地或不正确地预测在第二接合操作中制造的特定芯片中的异常发生的反馈。

[0009] 在一或多个实施例中,指令能够操作以使计算系统通过选择性地选取在第一接合操作中观察的参数类型的子集来接收训练数据。第一接合操作的测量值是针对参数类型的所述子集的测量值。

[0010] 在一或多个实施例中,训练数据包括通过从多种不同测量类型导出考虑第一接合操作中的测量类型之间的关系的而产生的工艺数据。指令能够操作以接收指示通过从多种测量类型导出考虑第二接合操作中的测量类型之间的关系的的信息而从第二接合操作的测量值产生的工艺数据的输入数据。

[0011] 本文中的实施例也包含对应计算机程序产品、设备及方法。

[0012] 例如,在一个实施例中,提供一种有形地体现于非暂时性机器可读存储媒体中的计算机程序产品。所述计算机程序产品包含可操作以使计算系统通过接收训练数据来训练机器学习模型的指令。所述训练数据包括从第一接合操作的测量值产生的工艺数据;及多条导线在接合到第一组表面后的状态。所述状态的每一状态包括与检测第一接合操作中的一或多个异常有关的目标的一或多个候选结果。计算机程序产品包含可操作以使计算系统通过产生工艺数据的一或多个权重使得输入到机器学习模型的工艺数据预测所述目标的所述一或多个候选结果来训练所述机器学习模型的指令。

[0013] 在另一实例实施例中,提供一种计算装置。所述计算装置包含但不限于处理器及存储器。所述存储器含有在由所述处理器执行时控制所述计算装置以存取在第一接合操作的训练数据上训练的机器学习模型的指令。所述第一接合操作包括用以将第一组的多条导线接合到第一组表面的操作。通过监督式学习来训练所述机器学习模型,包括接收所述训练数据。所述训练数据包括从所述第一接合操作的测量值产生的工艺数据;及所述多条导线在接合到所述第一组表面后的状态。所述状态的每一状态包括与检测所述第一接合操作中的一或多个异常有关的目标的一或多个候选结果。通过监督式学习来训练所述机器学习模型,包括产生所述工艺数据的一或多个权重,使得输入到所述机器学习模型的所述工艺数据预测所述目标的所述一或多个候选结果。所述存储器含有在由所述处理器执行时控制所述计算装置以接收指示从第二接合操作的测量值产生的工艺数据的输入数据的指令。所述第二接合操作包括用以将第二组的多条导线接合到第二组表面的操作。所述第二组的多条导线不同于所述第一组的多条导线。所述第二组表面不同于所述第一组表面。所述计算机程序产品包含可操作以使计算装置根据所述机器学习模型对所述输入数据进行加权的指令。所述存储器含有在由所述处理器执行时控制所述计算装置以基于根据所述机器学习模型对所述输入数据进行加权来产生指示所述第二接合操作中的异常发生的风险的异常预测量的指令。所述存储器含有在由所述处理器执行时控制所述计算装置以输出所述异常预测量以控制所述第二接合操作的指令。

[0014] 在另一实例实施例中,提供一种计算装置。所述计算装置包含但不限于处理器及

存储器。所述存储器含有在由所述处理器执行时控制所述计算装置以通过接收训练数据来训练机器学习模型的指令。所述训练数据包括从第一接合操作的测量值产生的工艺数据；及多条导线在接合到第一组表面后的状态。所述状态的每一状态包括与检测第一接合操作中之一或多个异常有关的目标之一或多个候选结果。计算机程序产品包含可操作以使计算装置通过产生工艺数据之一或多个权重使得输入到机器学习模型的工艺数据预测所述目标的所述之一或多个候选结果来训练所述机器学习模型的指令。

[0015] 在一或多个实施例中，计算装置是计算系统或计算系统的部分。

[0016] 在另一实例实施例中，提供一种计算机实施方法。所述方法包括存取在第一接合操作的训练数据上训练的机器学习模型。所述第一接合操作包括用以将第一组的多条导线接合到第一组表面的操作。通过监督式学习来训练所述机器学习模型，包括接收所述训练数据。所述训练数据包括从所述第一接合操作的测量值产生的工艺数据；及所述多条导线在接合到所述第一组表面后的状态。所述状态的每一状态包括与检测所述第一接合操作中之一或多个异常有关的目标之一或多个候选结果。通过监督式学习来训练所述机器学习模型，包括产生所述工艺数据之一或多个权重，使得输入到所述机器学习模型的所述工艺数据预测所述目标的所述之一或多个候选结果。所述方法包括接收指示从第二接合操作的测量值产生的工艺数据的输入数据。所述第二接合操作包括用以将第二组的多条导线接合到第二组表面的操作。所述第二组的多条导线不同于所述第一组的多条导线。所述第二组表面不同于所述第一组表面。所述方法包括根据所述机器学习模型对所述输入数据进行加权。所述方法包括基于根据所述机器学习模型对所述输入数据进行加权来产生指示所述第二接合操作中的异常发生的风险的异常预测量。所述方法包括输出所述异常预测量以控制所述第二接合操作。

[0017] 在另一实例实施例中，提供一种计算机实施方法。所述方法包括通过接收训练数据来训练机器学习模型。所述训练数据包括从第一接合操作的测量值产生的工艺数据；及多条导线在接合到第一组表面后的状态。所述状态的每一状态包括与检测所述第一接合操作中之一或多个异常有关的目标之一或多个候选结果。所述方法包括通过产生工艺数据之一或多个权重使得输入到机器学习模型的工艺数据预测所述目标的所述之一或多个候选结果来训练所述机器学习模型。

[0018] 在这些计算机程序产品、设备及方法中的任何者中，一或多个候选结果可包括对第一接合操作中的个别导线或接合件之一或多个破坏性质量保证测试。异常预测量可为第二接合操作中的个别导线或接合件中的异常的预测量，而未对所述个别导线或接合件执行破坏性质量保证测试。

[0019] 在这些计算机程序产品、设备及方法中的任何者中，第一接合操作可包括以下中的一者：球形接合操作，其中破坏性质量保证测试包括用于测试球形接合件的球剪切测试(ball shear test)；及针脚式接合操作，其中破坏性质量保证测试包括用于测试针脚式接合件的针脚拉拔测试(stitch pull test)。

[0020] 在这些计算机程序产品、设备及方法中的任何者中，一或多个异常可与以下中的一或多者相关联：浮动状况；引线框架污染；及裸片倾斜。

[0021] 在这些计算机程序产品、设备及方法中的任何者中，第二接合操作可接合第二组的多条导线的导线以形成集成电路芯片。从第二接合操作的测量值产生的工艺数据可整体

上与集成电路芯片的特定芯片有关,且从与所述特定芯片相关联且在第二接合操作中接合的导线的测量数据导出。

[0022] 在这些计算机程序产品、设备及方法中的任何者中,输入数据可包括在第二接合操作期间接收的实时传感器测量值。对于第二组的多条导线的给定导线,传感器测量值可包括以下中的一或多个:热量测量值;功率测量值;力测量值;放电结球 (electric flame-off) (EFO) 测量值;及超声波测量值。

[0023] 在这些计算机程序产品、设备及方法中的任何者中,输入数据可包括在第二接合操作期间接收的实时传感器测量值。传感器测量值可包括在第二接合操作中涉及的接合系统的测量值。异常预测量可控制第二接合操作以校正在第二接合操作中涉及的接合系统中的一或多个异常及/或减少一或多个异常的发生。

[0024] 在这些计算机程序产品、设备及方法中的任何者中,输入数据可包括用指示在第二接合操作中涉及的特定导线、裸片或芯片的标识符或位置中的一或多者的原始信息标记的经接收传感器测量值。异常预测量可识别在第二接合操作中发生的异常,且与原始信息相关以指示异常的位置。

[0025] 在这些计算机程序产品、设备及方法中的任何者中,输入数据可包含包括以下中的一或多者的经导出数据:指示与在第二接合操作中接合于特定芯片中的多条导线有关的测量值的中值或平均值的经产生值;包括多条导线中的每一者与所述值的偏差的一组经产生偏差;及考虑所述一组偏差的所述芯片的经产生度量。

[0026] 在这些计算机程序产品、设备及方法中的任何者中,训练数据可包括从经导出数据产生的工艺数据,所述经导出数据包括与多种不同类型的测量值相关且与第一接合操作中的相同导线有关的经产生奇异数据值。指示从第二接合操作的测量值产生的工艺数据的输入数据可包括经产生奇异数据值。所述经产生奇异数据值可与多种不同类型的测量值相关且与第二接合操作中的相同导线有关。

[0027] 在这些计算机程序产品、设备及方法中的任何者中,工艺数据的一或多个权重可经产生用于训练数据的梯度提升模型。

[0028] 在这些计算机程序产品、设备及方法中的任何者中,机器学习模型可进一步由多个经产生的机器学习模型训练且基于k折交叉验证 (k-fold cross-validation) 来选择。

[0029] 在这些计算机程序产品、设备及方法中的任何者中,第一及第二接合操作的测量值可包括与形成球形接合件的工艺相关联的测量值。

[0030] 在这些计算机程序产品、设备及方法中的任何者中,第一及第二接合操作的测量值可包括与形成针脚式接合件的工艺相关联的测量值。

[0031] 在这些计算机程序产品、设备及方法中的任何者中,异常预测量可为第二组的多条导线的导线与第二组表面的引线框架或裸片之间的有缺陷接合件的预测量。

[0032] 下文在结合本申请案所呈现的附图进行阅读时,在具体实施方式中呈现实例实施例的其它特征及方面。

附图说明

[0033] 图1说明提供根据本技术的至少一个实施例的计算系统的硬件组件的说明的框图。

[0034] 图2说明根据本技术的至少一个实施例的包含经由交换系统且经由网络彼此通信的一组实例装置的实例网络。

[0035] 图3说明根据本技术的至少一个实施例的通信协议系统的概念模型的表示。

[0036] 图4说明根据本技术的至少一个实施例的包含多种控制及工作节点(worker node)的通信网络计算系统。

[0037] 图5说明展示根据本技术的至少一个实施例的用于在节点的失败后调整通信网络或通信网络中的工作项目的实例处理过程(process)的流程图。

[0038] 图6说明根据本技术的至少一个实施例的包含控制节点及工作节点的通信网络计算系统的一部分。

[0039] 图7说明展示根据本技术的至少一个实施例的用于执行数据分析或处理项目的实例处理过程的流程图。

[0040] 图8说明根据本技术的至少一个实施例的包含事件流处理引擎(ESPE)的组件的框图。

[0041] 图9说明展示根据本技术的至少一个实施例的包含由事件流处理引擎执行的操作的实例处理过程的流程图。

[0042] 图10说明根据本技术的至少一个实施例的对接于发布装置与多个事件订阅装置之间的ESP系统。

[0043] 图11说明根据本技术的至少一个实施例的用于产生及使用机器学习模型的处理过程的实例的流程图。

[0044] 图12说明根据本技术的至少一个实施例的作为神经网络的机器学习模型的实例。

[0045] 图13到14说明根据本技术的至少一个实施例的用于制造集成电路芯片的实例流程图。

[0046] 图15A及15B说明根据本技术的至少一个实施例的集成电路芯片的一些实例组件及相应接合。

[0047] 图16A到16D说明根据本技术的至少一个实施例的在制造集成电路芯片时涉及的实例接合操作。

[0048] 图17说明根据本技术的至少一个实施例的实例集成电路芯片。

[0049] 图18A到18C说明用于球形接合件的破坏性球剪切测试过程。

[0050] 图19A到19C说明用于针脚式接合件的破坏性针脚拉拔测试过程。

[0051] 图20A说明在本技术的至少一个实施例中的训练系统的实例框图。

[0052] 图20B说明在本技术的至少一个实施例中的控制系统的实例框图。

[0053] 图21A是说明根据本技术的至少一个实施例的用于训练机器学习模型的实例方法的流程图。

[0054] 图21B是说明根据本技术的至少一个实施例的用于控制接合操作的实例方法的流程图。

[0055] 图21C是说明根据本技术的至少一个实施例的用于更新用于控制接合操作的机器学习模型的实例方法的流程图。

[0056] 图22A是说明根据本技术的至少一个实施例的在运动特征模式与对应芯片在引线框架上的位置之间的关系的关系的图表。

[0057] 图22B说明根据本技术的至少一个实施例的对应于图22A的图表的芯片在引线框架上的实例位置。

[0058] 图23A说明比较用于在接合操作期间形成的接合件的破坏性测试的预测模型的训练数据及测试数据的实例质量保证(QA)数据。

[0059] 图23B说明根据本技术的至少一个实施例的用于产生经导出处理数据的实例方法。

[0060] 图24是说明根据本技术的至少一个实施例的用于导出与接合操作系统相关联的工艺数据的实例图表的图表。

[0061] 图25是说明根据本公开实施例建模的经预测球剪切值与由破坏性测试获得的实际球剪切值之间的对应性的图表。

[0062] 图26是说明根据本公开实施例建模的经预测针脚拉拔值与由破坏性测试获得的实际针脚拉拔值之间的对应性的图表。

[0063] 图27A是说明根据本技术的至少一个实施例的事件流处理(ESP)系统的堆叠的功能框图。

[0064] 图27B是说明根据本技术的至少一个实施例的用于产生机器学习模型的实例方法的流程图。

[0065] 图28是根据本技术的至少一个实施例的计算机程序产品的功能框图。

具体实施方式

[0066] 在以下描述中,出于说明目的,阐述具体细节以提供对本技术的实施例的透彻理解。但是,将明白,可在没有这些具体细节的情况下实践各种实施例。图及描述并不希望为限制性的。

[0067] 随后描述仅提供实例实施例,且并不希望限制本公开的范围、适用性或配置。实际上,实例实施例的随后描述将向所属领域的技术人员提供用于实施实例实施例的详尽描述。应理解,在不脱离如权利要求书中阐述的技术的精神及范围的情况下,可在元件的功能及布置方面进行各种改变。

[0068] 在以下描述中给出具体细节以提供对实施例的透彻理解。但是,所属领域的技术人员将理解,可在没有这些具体细节的情况下实践实施例。例如,可将电路、系统、网络、处理过程及其它组件展示为框图形式的组件,以免在不必要的细节方面使实施例不清楚。在其它例子中,可在没有不必要的细节的情况下展示熟知电路、处理过程、算法、结构及技术以免使实施例不清楚。

[0069] 而且,应注意,个别实施例可被描述为处理过程,所述处理过程被描绘为流程图(flowchart/flow diagram)、数据流程图、结构图或框图。尽管流程图可将操作描述为循序处理过程,但许多操作可并行或同时执行。另外,可重新布置操作的顺序。处理过程在其操作完成时终止,但可具有未包含于图中的额外操作。处理过程可对应于方法、函数、过程、子例程、子程序等。当处理过程对应于函数时,其终止可对应于所述函数返回到呼叫函数或主函数。

[0070] 在一些图中描绘的系统可以各种配置提供。在一些实施例中,系统可配置为分布式系统,其中系统的一或多个组件跨云计算系统中的一或多个网络分布。

[0071] 图1是提供根据本技术的实施例的数据传输网络100的硬件组件的说明的框图。数据传输网络100是可用于处理大量数据(其中需要大量计算机处理循环)的专用计算机系统。

[0072] 数据传输网络100也可包含计算环境114。计算环境114可为处理在数据传输网络100内接收的数据的专用计算机或其它机器。数据传输网络100也包含一或多个网络装置102。网络装置102可包含尝试与计算环境114通信的客户端装置。例如,网络装置102可将数据发送到计算环境114以进行处理,可将信号发送到计算环境114以控制计算环境或其正在处理的数据的不同方面,以及其它原因。网络装置102可通过若干方式(举例来说,例如经由一或多个网络108)与计算环境114交互。如图1中所展示,计算环境114可包含一或多个其它系统。例如,计算环境114可包含数据库系统118及/或通信网络120。

[0073] 在其它实施例中,网络装置可经由网络108将大量数据提供到计算环境114,一次全部或在时段内流式传输(例如,使用事件流处理(ESP),关于图8到10进一步描述)。例如,网络装置102可包含可将数据传输或以其它方式提供到计算环境114的网络计算机、传感器、数据库或其它装置。例如,网络装置可包含局域网装置,例如路由器、集线器、交换机或其它计算机网络链接装置。这些装置可提供多种经存储或经产生数据,例如网络数据或特定于网络装置自身的数据。网络装置也可包含监测其环境或其它装置以收集关于所述环境或所述装置的数据的传感器,且此类网络装置可提供其随时间收集的数据。网络装置也可包含在物联网内的装置,例如在家庭自动化网络内的装置。一些这些装置可被称为边缘装置,且可涉及边缘计算电路系统。数据可由网络装置直接传输到计算环境114或到网络附接式数据存储库(例如网络附接式数据存储库110)以进行存储,使得数据可稍后由计算环境114或数据传输网络100的其它部分检索。

[0074] 数据传输网络100也可包含一或多个网络附接式数据存储库110。网络附接式数据存储库110用于将待由计算环境114处理的数据以及由计算系统产生的任何中间或最终数据存储于非易失性存储器中。但是,在特定实施例中,计算环境114的配置允许其操作经执行使得中间及最终数据结果可仅存储在易失性存储器(例如,RAM)中,而不需要将中间或最终数据结果存储到非易失性类型的存储器(例如,磁盘)。此在特定情形中可为有用的,例如当计算环境114从用户接收特定(ad hoc)查询时及当需要实时地产生通过处理大量数据而产生的响应时。在此非限制性情形中,计算环境114可经配置以将经处理信息留存于存储器内,使得可针对用户以不同细节层次产生响应以及允许用户交互地查询此信息。

[0075] 网络附接式数据存储库可存储以多种不同方式组织及来自多种不同来源的多种不同类型的数据。例如,网络附接式数据存储库可包含除定位于计算环境114内的可由定位于其中的处理器直接存取的主要存储器外的存储器。网络附接式数据存储库可包含次要、三次或辅助存储器,例如大型驱动器、服务器、虚拟存储器以及其它类型。存储装置可包含可携式或非可携式存储装置、光学存储装置,及能够存储、含有数据的各种其它媒体。机器可读存储媒体或计算机可读存储媒体可包含其中可存储数据且不包含载波及/或暂时性电子信号的非暂时性媒体。非暂时性媒体的实例可包含例如磁盘或磁带、光学存储媒体(例如光盘或数字多功能光盘)、快闪存储器、存储器或存储器装置。计算机程序产品可包含可表示过程、函数、子程序、程序、例程、子例程、模块、软件封装、类别,或指令、数据结构或程序叙述的任何组合的代码及/或机器可执行指令。代码段可通过传递及/或接收信息、数据、自

变量、参数或存储器内容而耦合到另一代码段或硬件电路。信息、自变量、参数、数据等可经由任何合适手段(包含存储器共享、消息传递、符记传递、网络传输等)传递、转送或传输。此外,数据存储库可保持多种不同类型的数据。例如,网络附接式数据存储库110可保持未结构化(例如,原始)数据,例如制造数据(例如,含有识别用例如色彩及模型的每一产品的参数数据制造的产品的记录的数据库)或产品销售数据库(例如,含有识别个别产品销售的细节的个别数据记录的数据库)。

[0076] 未结构化数据可以不同形式(例如一般文件(flat file)或数据记录集聚)呈现给计算环境114,且可具有数据值及伴随时间戳。计算环境114可用于以多种方式分析未结构化数据,以确定用于(例如,阶层式地)结构化所述数据使得经结构化数据是针对用户希望对数据执行的类型的进一步分析定制的最佳方式。例如,在被处理后,未结构化时间戳数据可按时间汇总(例如,成每日时间段单位)以产生时间序列数据,及/或根据一或多个维度(例如,参数、属性及/或变量)阶层式地结构化。例如,数据可存储在阶层式数据结构(例如ROLAP OR MOLAP数据库)中,或可以另一表列形式(例如以平面阶层形式)存储。

[0077] 数据传输网络100也可包含一或多个服务器群106。计算环境114可将选择通信或数据路由到一或多个服务器群106或服务器群内的一或多个服务器。服务器群106可经配置以依预定方式提供信息。例如,服务器群106可响应于通信而存取数据以进行传输。服务器群106可与数据传输网络100内的每一其它装置(例如计算环境114)分开容置,及/或可为装置或系统的部分。

[0078] 服务器群106可代管多种不同类型的数据处理作为数据传输网络100的部分。服务器群106可从网络装置、计算环境114、云网络116或其它来源接收多种不同数据。数据可能已作为来自控制数据库的输入从一或多个传感器获得或收集,或可能已作为来自外部系统或装置的输入而被接收。服务器群106可通过基于由服务器群实施的一或多个规则将原始数据转变成经处理数据来帮助处理数据。例如,可分析传感器数据以确定环境中随时间推移或实时的变化。

[0079] 数据传输网络100也可包含一或多个云网络116。云网络116可包含提供云服务的云基础设施系统。在特定实施例中,由云网络116提供的服务可包含按需可供云基础设施系统的用户使用的众多服务。云网络116在图1中被展示为连接到计算环境114(且因此将计算环境114作为其客户端或用户),但云网络116可连接到图1中的装置中的任何者或由图1中的装置中的任何者利用。由云网络提供的服务可动态地扩展以满足其用户的需求。云网络116可包含一或多个计算机、服务器及/或系统。在一些实施例中,组成云网络116的计算机、服务器及/或系统不同于用户自身的远程部署(on-premises)计算机、服务器及/或系统。例如,云网络116可代管应用程序,且用户可经由通信网络(例如因特网)按需订购及使用应用程序。

[0080] 虽然图1中的每一装置、服务器及系统被展示为单个装置,但将了解,可代替性地使用多个装置。例如,一组网络装置可用于传输来自单个用户的各种通信,或远程服务器140可包含服务器堆叠。作为另一实例,可将数据处理为计算环境114的部分。

[0081] 数据传输网络100内的每一通信(例如,在客户端装置之间,在装置与连接管理系统150之间,在服务器106与计算环境114之间或在服务器与装置之间)可经由一或多个网络108发生。网络108可包含多种不同类型的网络中的一或多者,包含无线网络、有线网络或有

线及无线网络的组合。合适网络的实例包含因特网、个人局域网、局域网 (LAN)、广域网 (WAN) 或无线局域网 (WLAN)。无线网络可包含无线接口或无线接口的组合。作为实例,一或多个网络108中的网络可包含短距离通信通道(例如蓝牙 (Bluetooth) 或蓝牙低能量 (Bluetooth Low Energy) 通道)。有线网络可包含有线接口。有线及/或无线网络可使用路由器、接入点、网桥、网关或类似者实施以连接网络114中的装置,如将关于图2进一步描述。一或多个网络108可整个并入内部网络、商际网络或其的组合内或可包含内部网络、商际网络或其的组合。在一个实施例中,两个或更多个系统及/或装置之间的通信可由安全通信协议(例如安全套接层 (SSL) 或传输层安全性 (TLS)) 实现。另外,可对数据及/或事务细节加密。

[0082] 一些方面可利用物联网 (IoT), 其中事物(例如,机器、装置、电话、传感器)可连接到网络且可在事物内及/或事物外部收集及处理来自这些事物的数据。例如, IoT可包含许多不同装置中的传感器,且可应用高价值分析以识别隐藏的关系且驱使增加的效率。此可适用于大数据分析 (big data analytics) 及实时(例如, ESP) 分析两者。IoT可在各种领域中实施以例如用于存取(获取数据及移动数据的技术)、嵌入能力(具有嵌入式传感器的装置)及服务。IoT空间中的工业可包含汽车(联网 (connected) 汽车)、制造(联网工厂)、智慧城市、能源及零售。此将在下文关于图2进一步描述。

[0083] 如所提及,计算环境114可包含通信网络120及传输网络数据库系统118。通信网络120可为用于处理大量数据的基于网络的计算系统。传输网络数据库系统118可用于管理、存储及检索分配给及存储在一或多个网络附接式数据存储库110或驻留于传输网络数据库系统118内的不同位置处的其它数据存储库的大量数据。基于网络的计算系统120及传输网络数据库系统118中的计算节点可共享相同处理器硬件,例如定位于计算环境114内的处理器。

[0084] 图2说明根据本技术的实施例的包含经由交换系统且经由网络彼此通信的一组实例装置的实例网络。如所提及,数据传输网络100内的每一通信可经由一或多个网络发生。系统200包含经配置以经由多种类型的通信通道与多种类型的客户端装置(例如,客户端装置230)通信的网络装置204。

[0085] 如图2中所展示,网络装置204可经由网络(例如,经由基站210的蜂窝式网络)传输通信。通信可经由基站210路由到另一网络装置,例如网络装置205到209。通信也可经由基站210路由到计算环境214。例如,网络装置204可从其周围环境或从其它网络装置(例如网络装置205到209)收集数据且将所述数据传输到计算环境214。

[0086] 尽管网络装置204到209在图2中分别被展示为移动电话、膝上型计算机、平板计算机、温度传感器、运动传感器及音频传感器,但网络装置可为或包含对检测其环境的方面敏感的传感器。例如,网络装置可尤其包含例如水传感器、功率传感器、电流传感器、化学传感器、光学传感器、压力传感器、地理或位置传感器(例如, GPS)、速度传感器、加速度传感器、流量传感器等的传感器。可感测的特性的实例尤其包含力、扭矩、负荷、应变、位置、温度、气压、流体流量、化学性质、电阻、电磁场、辐射、辐照度、接近性、声学、湿度、距离、速度、振动、加速度、电势、电流。传感器可安装到用作多种不同类型的系统的部分(例如,钻油操作)的各种组件。网络装置可检测及记录与其监测的环境有关的数据,且将所述数据传输到计算环境214。

[0087] 如所提及,根据特定实施例,可包含收集待处理及/或传输到计算环境的数据的各种传感器的一种类型的系统包含钻油系统。例如,一或多个钻探操作传感器可包含测量起吊负荷、流体速率、井孔内及外的温度及密度、竖管压力、表面扭矩、钻杆的旋转速度、钻进率、机械比能等的表面传感器,及测量钻头的旋转速度、流体密度、井下扭矩、井下振动(轴向、切向、横向)、施加于钻头处的重量、环孔压力、压差、方位角、倾角、急弯(dog leg)严重性、经测量深度、垂直深度、井下温度等的井下传感器。除了由传感器直接收集的原始数据外,其它数据可包含由传感器开发或由客户端或其它控制装置指派给系统的参数。例如,一或多个钻探操作控制参数可控制例如泥浆发动机速度对流量比、钻头直径、经预测地层尽头(formation top)、地震数据、天气数据等的设置。可使用物理模型(例如地球模型、天气模型、地震模型、井底总成模型、井规划模型、环孔摩擦力模型等)产生其它数据。除传感器及控制设置外,例如,钻进率、机械比能、起吊负荷、流入流体速率、流出流体速率、泵压力、表面扭矩、钻杆的旋转速度、环孔压力、环孔摩擦压力、环孔温度、循环相当密度等的经预测输出也可存储在数据仓库中。

[0088] 在另一实例中,根据特定实施例,可包含收集待处理及/或传输到计算环境的数据的各种传感器的另一类型的系统包含在不同环境(例如办公空间、学校、公共空间、运动场馆或多种其它位置)中的家庭自动化或类似自动化网络。此自动化网络中的网络装置可包含允许用户存取、控制及/或配置定位于用户家中(例如,电视机、收音机、灯、风扇、加湿器、传感器、微波炉、熨斗及/或类似者)或用户家庭外部(例如,外部运动传感器、外部照明、车库开门器、洒水系统或类似者)的各种家用电器的网络装置。例如,网络装置102可包含可与家用电器耦合的家庭自动化交换机。在另一实施例中,网络装置可允许用户存取、控制及/或配置装置,例如办公相关装置(例如,复印机、打印机或传真机)、音频及/或视频相关装置(例如,接收器、扬声器、投影仪、DVD播放器或电视机)、媒体播放装置(例如,光盘播放器、CD播放器或类似者)、计算装置(例如,家庭计算机、膝上型计算机、平板计算机、个人数字助理(PDA)、计算装置或可穿戴装置)、照明装置(例如,灯或嵌入式灯具)、与安全系统相关联的装置、与警报系统相关联的装置、可在汽车中操作的装置(例如,无线电装置、导航装置)及/或类似者。可从此类各种传感器以原始形式收集数据,或可由传感器处理数据以建立由传感器基于原始数据开发或由客户端或其它控制装置指派给系统的参数或其它数据。

[0089] 在另一实例中,根据特定实施例,可包含收集待处理及/或传输到计算环境的数据的各种传感器的另一类型的系统包含电力或能源网。多种不同网络装置可包含于能源网中,尤其例如在一或多个发电厂、能源农场(尤其例如风力农场、太阳能农场)能源存储设施、工厂、消费者的家庭及企业内的各种装置。此类装置中的一或多个者可包含检测能量增益或损失、电输入或输出或损失及多种其它效率的一或多个传感器。这些传感器可收集数据以告知用户能源网及网格内的个别装置可如何运作及可如何使其更高效。

[0090] 网络装置传感器也可在将其收集的数据传输到计算环境114前,或在决定是否将数据传输到计算环境114前,对数据执行处理。例如,网络装置可例如通过比较数据或从数据计算的值且比较数据与一或多个阈值来确定所收集的数据是否满足特定规则。网络装置可使用此数据及/或比较以确定是否应将数据传输到计算环境214以供进一步使用或处理。

[0091] 计算环境214可包含机器220及240。尽管计算环境214在图2中被展示为具有两个机器220及240,但计算环境214可具有仅一个机器或可具有多于两个机器。组成计算环境

214的机器可包含经配置以个别地及/或共同地处理大量数据的专用计算机、服务器或其它机器。计算环境214也可包含含有经结构化数据(例如在一或多个阶层中组织的数据)或未结构化数据的一或多个数据库的存储装置。数据库可与计算环境214内的处理装置通信以将数据分配给所述处理装置。由于网络装置可将数据传输到计算环境214,所以数据可由计算环境214接收且随后存储在所述存储装置内。由计算环境214使用的数据也可存储在数据存储器235中,数据存储器235也可为计算环境214的一部分或连接到计算环境214。

[0092] 计算环境214可经由一或多个路由器225或其它网络间或网络内连接组件与各种装置通信。例如,计算环境214可经由一或多个路由器225与装置230通信。计算环境214可收集、分析及/或存储来自或关于通信、客户端装置操作、客户端规则及/或用户相关联动作的存储在一或多个数据存储器235处的数据。此数据可影响到计算环境214内的装置的通信路由、如何在计算环境214内存储或处理数据,以及其它动作。

[0093] 注意,各种其它装置可进一步用于影响在计算环境214内的装置之间及与计算环境214外部的装置之间的通信路由及/或处理。例如,如图2中所展示,计算环境214可包含网络服务器240。因此,计算环境214可检索所关注数据,例如客户端信息(例如,产品信息、客户端规则等)、技术产品细节、新闻、当前或经预测天气等。

[0094] 除计算环境214收集数据(例如,如从例如传感器的网络装置及客户端装置或其它来源接收)以作为大数据分析项目的部分进行处理外,其也可作为流传输分析环境的部分实时接收数据。如所提及,可使用如经由不同种类的网络或在本地通信的多种来源来收集数据。可在实时流式传输基础上接收此数据。例如,网络装置可周期性地从网络装置传感器接收数据,因为传感器持续感测、监测及跟踪其环境中的变化。计算环境214内的装置也可对其接收的数据执行预分析以确定是否应将所接收的数据作为正在进行的项目的部分进行处理。由计算环境214接收及收集的数据(无论接收的来源或方法或时序如何)可在时段内针对客户端进行处理以基于客户端的需求及规则确定结果数据。

[0095] 图3说明根据本技术的实施例的通信协议系统的概念模型的代表。更明确来说,图3识别对应于各种连接组件的在开放系统交互模型中的计算环境的操作。例如,模型300展示例如计算环境320(或图2中的计算环境214)的计算环境可如何与其网络中的其它装置通信,及控制计算环境与其它装置之间的通信如何且在什么条件下执行。

[0096] 模型可包含层302到314。所述层布置成堆叠。堆叠中的每一层服务比其高一个层级的层(仅应用层除外,其是最高层),且由在其下方一个层级的层服务(仅物理层除外,其是最低层)。物理层是最低层,因为其接收及传输原始数据且为通信系统中距用户最远的层。另一方面,应用层是最高层,因为其直接与软件应用程序交互。

[0097] 如所提及,模型包含物理层302。物理层302表示物理通信且可定义所述物理通信的参数。例如,此物理通信可以电、光学或电磁信号的形式出现。物理层302也定义可控制数据传输网络内的通信的协议。

[0098] 链路层304定义用于跨网络传输(即,移动)数据的链路及机制。链路层管理例如在网格计算环境内的节点间通信。链路层304可检测及校正错误(例如,物理层302中的传输错误)。链路层304也可包含媒体存取控制(MAC)层及逻辑链路控制(LLC)层。

[0099] 网络层306定义用于在网络内路由的协议。换句话说,网络层协调跨相同网络(举例来说,例如网格计算环境)中的节点传送数据。网络层306也可定义用于结构化网络内的

本地寻址的处理过程。

[0100] 传输层308可管理数据的传输及所述数据的传输及/或接收的质量。传输层308可提供用于传送数据的协议,举例来说,例如传输控制协议(TCP)。传输层308可组装及分解用于传输的数据帧。传输层也可检测在其下方的层中发生的传输错误。

[0101] 会话层310可建立、维持及管理网络上的装置之间的通信连接。换句话说,会话层控制网络上的网络装置之间的通信的对话或本质。会话层也可建立核对点设置(checkpointing)、休会(adjournment)、终止及重启过程。

[0102] 表示层312可提供应用程序与网络层之间的通信的转译。换句话说,此层可基于已知由应用程序或网络层接受的数据类型来加密、解密及/或格式化数据。

[0103] 应用层314直接与软件应用程序及终端用户交互,且管理它们之间的通信。应用层314可使用应用程序识别目的地、本地资源状态或可用性及/或通信内容或格式化。

[0104] 网络内连接组件322及324被展示为分别在例如物理层302及链路层304的较低层级中操作。例如,集线器可在物理层中操作且交换机可在链路层中操作。网络间连接组件326及328被展示为在较高层级(例如层306到314)上操作。例如,路由器可在网络层中操作且网络装置可在输送、会话、表示及应用层中操作。

[0105] 如所提及,在各种实施例中,计算环境320可与各种层中的一者、多者、所有者或任何者交互及/或在其上操作。例如,计算环境320可(例如,经由链路层)与集线器交互以调整集线器与哪些装置通信。物理层可由链路层服务,因此其可实施来自链路层的此数据。例如,计算环境320可控制其将从哪些装置接收数据。例如,如果计算环境320知晓特定网络装置已被关闭、破坏或以其它方式变得不可用或不可靠,那么计算环境320可指示集线器防止任何数据从所述网络装置传输到计算环境320。此处理过程可有益于避免接收不准确或已受不受控制的环境影响的数据。作为另一实例,计算环境320可与网桥、交换机、路由器或网关通信且影响组件选择系统(例如,系统200)内的哪一装置作为目的地。在一些实施例中,计算环境320可通过通过路由或修改现有通信与在特定层上操作的装备交换通信来与各种层交互。在另一实施例中(例如在网格计算环境中),节点可基于由模型内的其它层提供的特定参数或信息来确定应如何路由环境内的数据(例如,哪一节点应接收特定数据)。

[0106] 如所提及,计算环境320可为通信网格环境(其的通信可如图3的协议中所展示般实施)的一部分。例如,返回参考图2,机器220及240中的一或多个可为通信网格计算环境的一部分。可在其中数据驻留于机器或计算节点上的存储器中的具有非交互式工作负载的分布式系统中采用网格化计算环境。在此环境中,分析代码而非数据库管理系统控制由节点执行的处理。数据是通过将其预分配给网格节点来共同定位,且每一节点上的分析代码将本地数据加载到存储器中。每一节点可被指派特定任务,例如处理项目的一部分,或组织或控制网格内的其它节点。

[0107] 图4说明根据本技术的实施例的包含多种控制及工作节点的通信网格计算系统400。通信网格计算系统400包含三个控制节点及一或多个工作节点。通信网格计算系统400包含控制节点402、404及406。控制节点经由通信路径451、453及455通信地连接。因此,控制节点可彼此传输(例如,与通信网格或通知相关的)信息且彼此接收信息。尽管通信网格计算系统400在图4中被展示为包含三个控制节点,但通信网格可包含多于或少于三个控制节点。

[0108] 通信网络计算系统(或仅“通信网络”)400也包含一或多个工作节点。图4中展示六个工作节点410到420。尽管图4展示六个工作节点,但根据本技术的实施例的通信网络可包含多于或少于六个工作节点。包含于通信网络中的工作节点的数目可尤其取决于通信网络正处理的项目或数据集有多大、每一工作节点的容量、对通信网络指定的用以完成项目的时间。通信网络400内的每一工作节点可(有线或无线,且直接地或间接地)连接到控制节点402到406。因此,每一工作节点可从控制节点接收信息(例如,用以对项目执行工作的指令)且可将信息(例如,来自对项目执行的工作的结果)传输到控制节点。此外,工作节点可彼此通信(直接地或间接地)。例如,工作节点可在彼此之间传输与所执行的作业或由所述工作节点执行的作业内的个别任务有关的数据。但是,在特定实施例中,工作节点可能未例如连接(通信地或以其它方式)到某些其它工作节点。在实施例中,工作节点可仅能够与控制其的控制节点通信,且可能无法与通信网络中的其它工作节点通信,无论其是否为由控制工作节点的控制节点控制的其它工作节点,或由通信网络中的其它控制节点控制的工作节点。

[0109] 控制节点可与外部装置连接,控制节点可与所述外部装置通信(举例来说,例如服务器或计算机的网格用户可连接到网格的控制器)。例如,服务器或计算机可连接到控制节点且可将项目或作业传输到节点。项目可包含数据集。数据集可具有任何大小。一旦控制节点接收包含大数据集的此项目,控制节点便可分配数据集或与数据集有关的项目以由工作节点执行。替代地,对于包含大数据集的项目,数据集可由除控制节点以外的机器(例如,Hadoop数据节点)接收或存储。

[0110] 控制节点可维持网格中的节点的状态的知识(即,网格状态信息),接受来自客户端的工作请求,跨工作节点细分工作,协调工作节点,以及其它职责。工作节点可接受来自控制节点的工作请求且向控制节点提供由工作节点执行的工作的结果。网格可从单个节点(例如,机器、计算机、服务器等)启动。此第一节点可被指派或可作为将控制进入网格的任何额外节点的主控制节点启动。

[0111] 当项目经提交以供执行(例如,由客户端或网格的控制器)时,其可被指派给一组节点。在将节点指派给项目后,可建立数据结构(即,通信器)。通信器可由项目使用以在于每一节点上运行的项目代码之间共享信息。可在每一节点上建立通信句柄(handle)。例如,句柄是在单个节点上的单个处理过程内有效的通信器的引用,且可在请求节点之间的通信时使用句柄。

[0112] 控制节点(例如控制节点402)可被指定为主控制节点。服务器、计算机或其它外部装置可连接到主控制节点。一旦控制节点接收项目,主控制节点便可将项目的部分分配给其工作节点以执行。例如,当在通信网络400上起始项目时,主控制节点402控制待针对项目执行的工作以按照请求或指示完成项目。主控制节点可基于各种因素将工作分配给工作节点,例如项目的哪些子集或部分可最高效地且在正确时间量内完成。例如,工作节点可对已在工作节点本地(例如,存储在工作节点上)的数据的一部分执行分析。在每一工作节点执行及完成其作业后,主控制节点也协调及处理由每一工作节点执行的工作的结果。例如,主控制节点可从一或多个工作节点接收结果,且控制节点可组织(例如,收集及组装)所接收的结果且对其进行编译以产生从终端用户接收的项目的完整结果。

[0113] 任何剩余控制节点(例如控制节点404及406)可被指派为项目的备份控制节点。在

实施例中,备份控制节点可能未控制项目的任何部分。代替性地,备份控制节点可用作主控制节点的备份,且如果主控制节点失败则接替作为主控制节点。如果通信网络包含仅单个控制节点,且控制节点失败(例如,控制节点被关闭或中断),那么通信网络整体可能失败且在通信网络上运行的任何项目或作业可能失败且可能未完成。虽然项目可再次运行,但此失败可引起项目完成的延迟(在一些情况中严重延迟,例如整夜延迟)。因此,具有多个控制节点(包含备份控制节点)的网格可为有益的。

[0114] 例如,为了将另一节点或机器添加到网格,主控制节点可打开一对侦听套接字(listening socket)。套接字可用于接受来自客户端的工作请求,且第二套接字可用于接受来自其它网格节点(例如,其它机器、计算机、服务器)的连接。主控制节点可被提供将参与网格的其它节点(例如,其它机器、计算机、服务器)的列表,及每一节点将填写网格的角色。在主控制节点(例如,网格上的第一节点)启动时,主控制节点可使用网络协议以在网格中的每一其它节点上启动服务器处理过程。例如,命令行参数可告知每一节点一或多条信息,尤其例如:节点将在网格中具有的角色、主控制节点的主机名、主控制节点在其上接受来自同级节点的连接的端口号。信息也可以配置文件提供,经由安全外壳隧道传输,从配置服务器复原。虽然网格中的其它机器最初可能不知晓网格的配置,但所述信息也可由主控制节点发送到每一其它节点。网格信息的更新也可随后发送到所述节点。

[0115] 对于除添加到网格的主控制节点外的任何控制节点,控制节点可开启三个套接字。第一套接字可接受来自客户端的工作请求,第二套接字可接受来自其它网格成员的连接,且第三套接字可(例如,永久地)连接到主控制节点。当控制节点(例如,主控制节点)接收来自另一控制节点的连接时,其首先查看同级节点是否在网格中的经配置节点的列表中。如果其未在列表上,那么控制节点可清除连接。如果其在列表上,那么其可尝试认证连接。如果认证成功,那么认证节点可向其同级传输信息,例如节点在其上侦听连接的端口号、节点的主机名、关于如何认证节点的信息,以及其它信息。当节点(例如新控制节点)接收关于另一作用中节点的信息时,其将查看其是否已具有到所述另一节点的连接。如果其不具有到所述节点的连接,那么其可建立到所述控制节点(例如主控制节点)的连接。

[0116] 添加到网格的任何工作节点可建立到主控制节点及网格上的任何其它控制节点(例如,主控制节点)的连接。在建立连接后,其可向网格(例如,任何控制节点,包含主及备份两者,或控制网络的服务器或用户)认证自身。在成功认证后,工作节点可接受来自控制节点的配置信息。

[0117] 当节点加入通信网络时(例如,当节点电源开启或连接到网格上的现有节点,或两者时),节点被指派(例如,由网格的操作系统)通用唯一标识符(UUID)。此唯一标识符可帮助其它节点及外部实体(装置、用户等)识别节点且将其与其它节点区分。当节点连接到网格时,节点可与网格中的其它节点共享其唯一标识符。由于每一节点可共享其唯一标识符,所以每一节点可知晓网格上的每一其它节点的唯一标识符。唯一标识符也可指定网格内的节点(例如,备份控制节点)中的每一者的阶层。例如,备份控制节点中的每一者的唯一标识符可存储在备份控制节点的列表中,以指示备份控制节点将接替失败主控制节点成为新的主控制节点的顺序。但是,也可使用除使用节点的唯一标识符外的方法来确定节点的阶层。例如,可基于其它预定因素来预先确定或指派阶层。

[0118] 网格可在任何时间添加新机器(例如,从任何控制节点起始)。在将新节点添加到网格时,控制节点可首先将新节点添加到其网格节点表中。接着,控制节点也可通知每一其

它控制节点关于新节点。接收通知的节点可认可其已更新其配置信息。

[0119] 例如,主控制节点402可将一或多个通信传输到备份控制节点404及406(及例如到通信网格内的其它控制或工作节点)。除其它协议外,此类通信可在项目执行的已知固定阶段之间以固定时间间隔周期性地发送。由主控制节点402传输的通信可具有变化类型且可包含多种类型的信息。例如,主控制节点402可传输通信网络的快照(例如,状态信息),使得备份控制节点404始终具有通信网络的最近快照。例如,快照或网格状态可包含网格的结构(例如,包含网格中的工作节点、节点的唯一标识符,或其与主控制节点的关系)及项目的状态(例如,包含每一工作节点的项目部分的状态)。快照也可包含从通信网格中的工作节点接收的分析或结果。备份控制节点可接收及存储从主控制节点接收的备份数据。备份控制节点可传输对来自主控制节点的此快照(或其它信息)的请求,或主控制节点可将此信息周期性地发送到备份控制节点。

[0120] 如所提及,如果主控制节点失败,那么备份数据可允许备份控制节点接替作为主控制节点,而无需网格从头开始启动项目。如果主控制节点失败,那么将接替作为主控制节点的备份控制节点可检索从主控制节点接收的快照的最新版本且使用所述快照从由备份数据指示的项目的阶段继续项目。此可防止项目整体失败。

[0121] 备份控制节点可使用各种方法来确定主控制节点已失败。在此方法的一个实例中,主控制节点可向备份控制节点(例如,周期性地)传输指示主控制节点正在工作且尚未失败的通信(例如心跳通信)。如果备份控制节点在特定预定时段内未接收到心跳通信,那么备用控制节点可确定主控制节点已失败。替代地,备份控制节点也可从主控制节点本身(在其失败前)或从工作节点接收主控制节点已失败的通信(例如,因为主控制节点未能与工作节点通信)。

[0122] 可执行不同方法来确定一组备份控制节点(例如,备份控制节点404及406)的哪一备份控制节点将接替失败主控制节点402且成为新的主控制节点。例如,可基于备份控制节点的唯一标识符基于其的排名或“阶层”来选取新的主控制节点。在替代实施例中,备份控制节点可由通信网格中或来自外部装置(例如,控制通信网络的系统基础设施或终端用户,例如服务器或计算机)的另一装置指派为新的主控制节点。在另一替代实施例中,可基于带宽或关于通信网格的其它统计数据来指定接替作为新的主控制节点的备份控制节点。

[0123] 通信网格内的工作节点也可能失败。如果工作节点失败,那么由失败工作节点执行的工作可在操作工作节点当中重新分配。在替代实施例中,主控制节点可向仍在通信网格上的可操作工作节点中的每一者传输工作节点中的每一者也故意失败的通信。在工作节点中的每一者失败后,其可相应检索其最新保存的其状态的核对点且从所述核对点重启项目以最小化所执行的项目的丢失进展。

[0124] 图5说明展示根据本技术的实施例的用于在节点的失败后调整通信网格或通信网格中的工作项目的实例处理过程的流程图。例如,处理过程可包含接收包含由通信网格中的节点执行的项目的一部分的项目状态的网格状态信息,如在操作502中描述。例如,控制节点(例如,连接到通信网格上的主控制节点及工作节点的备份控制节点)可接收网格状态信息,其中所述网格状态信息包含主控制节点的项目状态或工作节点的项目状态。主控制节点的项目状态及工作节点的项目状态可包含由通信网格中的主及工作节点执行的项目的一或多个部分的状态。处理过程也可包含存储网格状态信息,如在操作504中描述。例如,

控制节点(例如,备份控制节点)可将经接收的网格状态信息在本地存储在控制节点内。替代地,可将网格状态信息发送到用于存储的另一装置,控制节点可在所述另一装置处存取信息。

[0125] 处理过程也可包含在操作506中接收对应于通信网络中的节点的失败通信。例如,节点可接收包含主控制节点已失败的指示的失败通信,以提示备份控制节点接替主控制节点。在替代实施例中,节点可接收工作节点已失败的失败,以提示控制节点重新指派由工作节点执行的工作。处理过程也可包含重新指派节点或由失败节点执行的项目的一部分,如在操作508中描述。例如,控制节点可在接收失败通信时基于失败通信将备份控制节点指定为新的主控制节点。如果失败节点是工作节点,那么控制节点可使用通信网络的快照识别失败工作节点的项目状态,其中失败工作节点的项目状态包含在失败时间由失败工作节点执行的项目的一部分的状态。

[0126] 处理过程也可包含基于重新指派接收经更新网格状态信息,如在操作510中描述,及基于经更新网格状态信息将指令集传输到通信网络中的一或多个节点,如在操作512中描述。经更新网格状态信息可包含主控制节点的经更新项目状态或工作节点的经更新项目状态。可将经更新信息传输到网格中的其它节点以更新其的陈旧存储信息。

[0127] 图6说明根据本技术的实施例的包含控制节点及工作节点的通信网络计算系统600的一部分。出于说明目的,通信网络600计算系统包含一个控制节点(控制节点602)及一个工作节点(工作节点610),但可包含更多工作及/或控制节点。控制节点602经由通信路径650通信地连接到工作节点610。因此,控制节点602可经由路径650将(例如,与通信网络或通知有关的)信息传输到工作节点610及从工作节点610接收信息。

[0128] 类似于在图4中,通信网络计算系统(或仅“通信网络”)600包含数据处理节点(控制节点602及工作节点610)。节点602及610包含多核心数据处理器。每一节点602及610包含在与所述节点相关联的数据处理器上执行且与也与所述节点相关联的缓冲存储器622对接的网格化的(grid-enabled)软件组件(GESC)620。每一节点602及610包含在控制节点602处的数据库服务器(未展示)上及在工作节点610处的数据库服务器(未展示)上执行的数据库管理软件(DBMS)628。

[0129] 每一节点也包含数据存储库624。类似于图1中的网络附接式数据存储库110及图2中的数据存储库235,数据存储库624用于存储待由计算环境中的节点处理的数据。数据存储库624也可将由计算系统产生的在处理后的任何中间或最终数据存储例如在非易失性存储器中。但是,在特定实施例中,网格计算环境的配置允许其操作经执行使得中间及最终数据结果可单独存储在易失性存储器(例如,RAM)中,而不需要将中间或最终数据结果存储到非易失性类型的存储器。将此数据存储于易失性存储器中在特定情形中可为有用的,例如当网格接收到来自客户端的查询(例如,特定)时及当需要快速地或实时地产生通过处理大量数据而产生的响应时。在此情形中,网格可经配置以将数据留存于存储器内,使得可以不同细节层次产生响应且使得客户端可交互地查询此信息。

[0130] 每一节点也包含用户定义的功能(UDF)626。UDF提供使DBMS 628将数据传送到存储在由DBMS管理的数据存储库624中的数据库或从所述数据库接收数据的机制。例如,UDF 626可由DBMS调用以将数据提供到GESC以进行处理。UDF 626可与GESC建立套接连接(未展示)以传送数据。替代地,UDF 626可通过将数据写入到可由UDF及GESC两者存取的共享存储

器而将数据传送到GESC。

[0131] 在节点602及620处的GESC 620可经由网络(例如图1中所展示的网络108)连接。因此,节点602及620可经由网络使用预定通信协议(举例来说,例如消息传递接口(MPI))彼此通信。每一GESC 620可参与另一节点处的GESC的点对点通信,或经由网络与多个GESC进行集体通信。每一节点处的GESC 620可含有相同(或几乎相同)软件指令。每一节点可能操作作为控制节点或工作节点。控制节点602处的GESC可经由通信路径652与客户端装置630通信。更明确来说,控制节点602可与由客户端装置630代管的客户端应用程序632通信,以接收查询且在处理大量数据后对所述查询作出响应。

[0132] DMBS 628可控制节点602或610内的数据库或数据结构(未展示)的建立、维护及使用。数据库可组织存储在数据存储库624中的数据。控制节点602处的DMBS 628可接受对数据的请求且针对请求传送适当数据。通过此处理过程,数据集合可跨多个物理位置分布。在此实例中,每一节点602及610将由管理系统管理的总数据的一部分存储在其相关联数据存储装置624中。

[0133] 此外,DBMS可负责使用复制技术进行保护使免受数据丢失。复制包含提供存储在一或多个其它节点上的一个节点上的数据的备份副本。因此,如果一个节点失败,那么可从驻留于另一节点处的经复制副本复原来自失败节点的数据。但是,如本文中关于图4所描述,通信网格中的每一节点的数据或状态信息也可与网格上的每一节点共享。

[0134] 图7说明展示根据本技术的实施例的用于在网格计算系统内执行项目的实例方法的流程图。如关于图6所描述,控制节点处的GESC可与客户端装置(例如,客户端装置630)传输数据以接收用于执行项目的查询且在已处理大量数据后对所述查询作出响应。查询可传输到控制节点,其中查询可包含执行项目的请求,如在操作702中描述。查询可含有关于待执行的数据分析类型及是否应使用基于网格的计算环境来执行项目的指令,如在操作704中展示。

[0135] 为了起始项目,控制节点可确定查询是否请求使用基于网格的计算环境来执行项目。如果确定为否,那么控制节点在单独环境中(例如,在控制节点处)起始项目的执行,如在操作710中描述。如果确定为是,那么控制节点可在基于网格的计算环境中起始项目的执行,如在操作706中描述。在此情形中,请求可包含网格的经请求配置。例如,请求可包含在执行项目时待在网格中使用的控制节点的数目及工作节点的数目。在项目已完成后,控制节点可传输由网格产生的分析结果,如在操作708中描述。无论是否在单独或基于网格的环境中执行项目,控制节点在操作712中提供项目的结果。

[0136] 如关于图2所提及,本文中所描述的计算环境可收集数据(例如,如从例如传感器的网络装置(例如图2中的网络装置204到209)及客户端装置或其它来源接收)以作为数据分析项目的部分进行处理,且可作为流式传输分析环境(例如,ESP)的部分实时接收数据。可使用如经由不同种类的网络或在本地通信的多种来源例如在实时流式传输基础上收集数据。例如,网络装置可周期性地从网络装置传感器接收数据,因为传感器持续感测、监测及跟踪其环境中的变化。更明确来说,越来越多的分布式应用程序通过在将数据分配给地理上分布的接收者前对数据应用查询来开发或产生来自分布式来源的连续流动数据。事件流处理引擎(ESPE)可在接收数据时对数据连续应用查询且确定哪些实体应接收数据。客户端或其它装置也可订阅ESPE或处理ESP数据的其它装置,使得其可在处理后接收数据(例

如,基于由处理引擎确定的实体)。例如,图2中的客户端装置230可订阅计算环境214中的ESPE。在另一实例中,关于图10进一步描述的事件订阅装置1024a到1024c也可订阅ESPE。ESPE可确定或定义来自网络装置或其它发布者(例如,图2中的网络装置204到209)的输入数据或事件流如何被转换成有意义的输出数据以供订阅者(举例来说,例如图2中的客户端装置230)消费。

[0137] 图8说明包含根据本技术的实施例的事件流处理引擎(ESPE)的组件的框图。ESPE 800可包含一或多个项目802。项目可被描述为由ESPE 800管理的引擎模型中的第二层级容器,其中项目的线程集区大小可由用户定义。一或多个项目802的每一项目可包含含有数据流(其是传入事件流的数据变换)的一或多个连续查询804。一或多个连续查询804可包含一或多个源窗口806及一或多个导出窗口808。

[0138] ESPE可在时段内接收与特定事件有关的流式传输数据,例如由一或多个网络装置感测的事件或其它数据。ESPE可执行与处理由一或多个装置建立的数据相关联的操作。例如,ESPE可从图2中所展示的一或多个网络装置204到209接收数据。如所提及,网络装置可包含感测其环境的不同方面的传感器,且可基于所述经感测的观察随时间收集数据。例如,ESPE可在图2中所展示的机器220及240中的一或多个者内实施。ESPE可在此机器内由ESP应用程序实施。ESP应用程序可将具有其自身的一或多个专用线程集区的ESPE嵌入到其应用程序空间中,其中主要应用程序线程可进行应用程序特定工作且ESPE至少通过将模型的例子建立到处理对象中来处理事件流。

[0139] 引擎容器是管理一或多个项目802的资源的模型中的顶层容器。在说明性实施例中,例如,对于ESP应用程序的每一例子,可仅存在一个ESPE 800,且ESPE 800可具有唯一引擎名称。另外,一或多个项目802可各自具有唯一项目名称,且每一查询可具有唯一连续查询名称且以一或多个源窗口806的唯一命名的源窗口开始。ESPE 800可为或可能并非持久的。

[0140] 连续查询建模涉及针对事件流操纵及变换定义窗口的有向图。在事件流操纵及变换的上下文中,窗口是事件流处理模型中的处理节点。连续查询中的窗口可对流经窗口的数据执行汇总、计算、模式匹配及其它操作。连续查询可被描述为源、关系、模式匹配及过程窗口的有向图。一或多个源窗口806及一或多个导出窗口808表示产生对查询结果集的更新作为通过ESPE 800的新的事件块流的连续执行的查询。例如,有向图是由边缘连接的一组节点,其中边缘具有与其相关联的方向。

[0141] 事件对象可被描述为可作为字段集合存取的数据分组,其中字段中的至少一者定义为关键词或唯一标识符(ID)。事件对象可使用包含二进制、字母数字、XML等的多种格式建立。每一事件对象可包含被指定为事件的主标识符(ID)的一或多个字段,因此ESPE 800可支持包含插入、更新、更新插入(upsert)及删除的事件的操作代码(操作码)。如果关键字段已存在,那么更新插入操作码更新事件;否则,插入事件。为说明,事件对象可为一组域值的紧缩二进制表示,且包含与事件相关联的元数据及字段数据两者。元数据可包含指示事件是否表示插入、更新、删除或更新插入的操作码,指示事件是否为正常、部分更新或来自保留策略管理的保留产生的事件的一组旗标,及可用于延时测量的一组微秒时间戳。

[0142] 事件块对象可被描述为事件对象的群组或封装。事件流可被描述为事件块对象流。一或多个连续查询804的连续查询使用一或多个源窗口806及一或多个导出窗口808将

由发布到ESPE 800中的流式传输事件块对象组成的源事件流变换成一或多个输出事件流。连续查询也可被视为数据流建模。

[0143] 一或多个源窗口806在有向图的顶部处不具有馈送到其中的窗口。事件流被发布到一或多个源窗口806中,且事件流可从所述处引导到如由有向图定义的下一组经连接窗口。一或多个导出窗口808全部为并非源窗口且具有将事件流式传输到其中的其它窗口的具现化窗口。一或多个导出窗口808可对传入事件流执行计算或变换。一或多个导出窗口808基于窗口类型(即,例如加入、滤波、计算、汇总、复制、模式匹配、过程、合并等的运算符)及窗口设置来变换事件流。在事件流被发布到ESPE 800中时,其被连续查询,且连续更新这些查询中的所得导出窗口集。

[0144] 图9说明展示根据本技术的一些实施例的包含由事件流处理引擎执行的操作的实例处理过程的流程图。如所提及,ESPE 800(或相关联ESP应用程序)定义如何将输入事件流变换成有意义的输出事件流。更明确来说,ESP应用程序可定义如何将来自发布者(例如,提供经感测数据的网络装置)的输入事件流变换成由订阅者消费的有意义的输出事件流(例如,由机器或机器组执行的数据分析项目)。

[0145] 在应用程序内,用户可与在ESPE的控制下独立地或通过浏览器应用程序以用户可选择的顺序在显示器中呈现给用户的一或多个用户接口窗口交互。例如,用户可执行ESP应用程序,此引起呈现第一用户接口窗口,所述第一用户接口窗口可包含与ESP应用程序相关联的多个菜单及选择器(例如下拉式菜单、按钮、文本框、超链接等),如所属领域的技术人员理解。如所属领域的技术人员进一步理解,可例如使用多个线程并行执行各种操作。

[0146] 在操作900,ESP应用程序可定义及启动ESPE,借此具现化装置(例如机器220及/或240)处的ESPE。在操作902中,建立引擎容器。为了说明,ESPE 800可使用将引擎容器指定为模型的管理器的函数调用来具现化。

[0147] 在操作904中,通过ESPE 800将一或多个连续查询804具现化为模型。一或多个连续查询804可用产生更新作为通过ESPE 800的新事件流的一或多个专用线程集区具现化。为说明,可建立一或多个连续查询804以建模ESPE 800内的业务处理逻辑,预测ESPE 800内的事件,建模ESPE 800内的物理系统,预测ESPE 800内的物理系统状态等。例如,如所提及,ESPE 800可用于支持传感器数据监测及管理(例如,感测可包含力、扭矩、负荷、应变、位置、温度、气压、流体流量、化学性质、电阻、电磁场、辐射、辐照度、接近性、声学、湿度、距离、速度、振动、加速度、电势或电流等)。

[0148] ESPE 800可分析及处理运动中的事件或“事件流”。代替存储数据及对经存储数据运行查询,ESPE 800可存储查询且通过其流式传输数据以允许在接收数据时对数据进行连续分析。可基于将输入事件流变换成输出事件流以基于所定义的连续查询模型及对流式传输的数据的应用来建模、模拟、评分、测试、预测等的关系、模式匹配及过程算法来建立一或多个源窗口806及一或多个导出窗口808。

[0149] 在操作906中,针对ESPE 800初始化发布/订阅(pub/sub)能力。在说明性实施例中,针对一或多个项目802的每一项目初始化pub/sub能力。为了针对ESPE 800初始化及启用pub/sub能力,可提供端口号。pub/sub客户端可使用运行ESPE的ESP装置的主机名及端口号来建立到ESPE 800的pub/sub连接。

[0150] 图10说明根据本技术的实施例的对接于发布装置1022与事件订阅装置1024a到

1024c之间的ESP系统1000。ESP系统1000可包含ESP装置或子系统1001、事件发布装置1022、事件订阅装置A 1024a、事件订阅装置B 1024b及事件订阅装置C 1024c。通过发布装置1022将输入事件流输出到ESP装置1001。在替代实施例中,可通过多个发布装置建立输入事件流。多个发布装置可进一步将事件流发布到其它ESP装置。由ESPE 800具现化的一或多个连续查询可分析及处理输入事件流以形成输出到事件订阅装置A1024a、事件订阅装置B 1024b及事件订阅装置C 1024c的输出事件流。ESP系统1000可包含事件订阅装置的更多或更少数目个事件订阅装置。

[0151] 发布-订阅是基于间接寻址的消息定向交互范式。经处理数据接收者通过订阅特定类别的事件来指定其有兴趣从ESPE 800接收信息,而信息源将事件发布到ESPE 800而不直接寻址接收方。ESPE 800协调交互且处理数据。在一些情况下,数据源接收数据接收者已接收经发布信息的确认。

[0152] 发布/订阅API可被描述为使事件发布者(例如发布装置1022)能够将事件流发布到ESPE 800或事件订阅者(例如事件订阅装置A1024a、事件订阅装置B 1024b及事件订阅装置C 1024c)以订阅来自ESPE 800的事件流的链接库。为说明,可定义一或多个发布/订阅API。使用发布/订阅API,事件发布应用程序可将事件流发布到ESPE 800的正在运行的事件流处理器项目源窗口中,且事件订阅应用程序可订阅ESPE 800的事件流处理器项目源窗口。

[0153] 发布/订阅API提供ESP应用程序与其它网络链接应用程序(例如在发布装置1022处具现化的事件发布应用程序,以及在事件订阅装置A 1024a、事件订阅装置B 1024b及事件订阅装置C 1024c中的一或多个者处具现化的事件订阅应用程序)之间的跨平台连接性及字节序(endianness)兼容性。

[0154] 返回参考图9,操作906初始化ESPE 800的发布/订阅能力。在操作908中,启动一或多个项目802。所述一或多个经启动项目可在ESP装置上的后台运行。在操作910中,从事件发布装置1022的一或多个计算装置接收事件块对象。

[0155] ESP子系统1001可包含发布客户端1002、ESPE 800、订阅客户端A 1004、订阅客户端B 1006及订阅客户端C 1008。发布客户端1002可由在发布装置1022处执行的事件发布应用程序使用发布/订阅API来启动。订阅客户端A 1004可由在事件订阅装置A1024a处执行的事件订阅应用程序A使用发布/订阅API来启动。订阅客户端B 1006可由在事件订阅装置B 1024b处执行的事件订阅应用程序B使用发布/订阅API来启动。订阅客户端C 1008可由在事件订阅装置C 1024c处执行的事件订阅应用程序C使用发布/订阅API来启动。

[0156] 含有一或多个事件对象的事件块对象从事件发布装置1022上的事件发布应用程序的例子注入到一或多个源窗口806的源窗口中。事件块对象可例如由事件发布应用程序产生且可由发布客户端1002接收。在于ESPE 800的一或多个源窗口806及/或一或多个导出窗口808之间,且到订阅客户端A1004、订阅客户端B 1006及订阅客户端C 1008,以及到事件订阅装置A 1024a、事件订阅装置B 1024b及事件订阅装置C 1024c传递事件块对象时,可维持唯一ID。发布客户端1002可进一步在事件块对象由连续查询处理时产生唯一嵌入式事务ID且将其包含于事件块对象中,以及发布装置1022指派给事件块对象的唯一ID。

[0157] 在操作912中,通过一或多个连续查询804处理事件块对象。在操作914中,将经处理事件块对象输出到事件订阅装置1024a到1024c的一或多个计算装置。例如,订阅客户端

A1004、订阅客户端B 1006及订阅客户端C 1008可分别将经接收的事件块对象发送到事件订阅装置A 1024a、事件订阅装置B 1024b及事件订阅装置C 1024c。

[0158] ESPE 800从事件块被发布到源窗口中时维持经接收事件块的事件块容器关系 (containership) 方面,且在输出到订阅者前运用各种事件转译完成由一或多个连续查询 804定义的有向图。订阅者可通过比较发布者(例如发布装置1022)附接到事件块对象的事件块对象的唯一ID与由订阅者接收的事件块ID而使群组经订阅事件往回与群组经发布事件相关。

[0159] 在操作916中,进行关于是否停止处理的确定。如果未停止处理,那么处理在操作 910中继续以继续从例如一或多个网络装置接收含有事件块对象的一或多个事件流。如果停止处理,那么处理在操作918中继续。在操作918中,停止经启动项目。在操作920中,关闭 ESPE。

[0160] 如所提及,在一些实施例中,在接收及存储数据后,处理大数据以进行分析项目。在其它实施例中,分布式应用程序通过在将数据分配给地理上分布的接收者前对数据应用查询来实时地处理来自分布式来源的连续流动数据。如所提及,事件流处理引擎(ESPE)可在接收数据时对数据连续应用查询且确定哪些实体接收经处理数据。此允许实时地处理及分配在多种环境中接收及/或收集的大量数据。例如,如关于图2所展示,可从可包含物联网内的装置(例如家庭自动化网络内的装置)的网络装置收集数据。但是,可从多种不同环境中的多种不同资源收集此数据。在任何此情形中,本技术的实施例允许实时处理此数据。

[0161] 本公开的方面提供技术问题的技术解决方案,例如在ESP装置失败时出现的计算问题,此导致完全服务中断及潜在显著数据丢失。当流式传输的数据正在支持任务关键操作(例如支持正在进行的制造或钻探操作的操作)时,数据丢失可能为灾难性的。ESP系统的实施例实现在多个ESP装置处运行的ESPE的快速且无缝故障转移(failover),而无服务中断或数据丢失,因此显著改进依靠数据流的实况或实时处理的操作系统的可靠性。事件发布系统、事件订阅系统及未在失败ESP装置处执行的每一ESPE都不知晓所述失败ESP装置或不受其影响。ESP系统可包含数千个事件发布系统及事件订阅系统。ESP系统将故障转移逻辑及感知保持于向外消息传递网络连接器及向外消息传递网络装置的边界内。

[0162] 在一个实例实施例中,提供一种用以支持事件流处理(ESP)事件块时的故障转移的系统。系统包含但不限于向外消息传递网络装置及计算装置。计算装置包含但不限于处理器及可操作地耦合到所述处理器的计算机可读媒体。处理器经配置以执行ESP引擎(ESPE)。计算机可读媒体具有存储在其上的指令,所述指令在由处理器执行时引起计算装置支持故障转移。事件块对象是从包含唯一标识符的ESPE接收。确定计算装置的第一状态为作用中或备用。当所述第一状态是在作用中时,确定计算装置的第二状态为最新在作用中或并非最新在作用中。当计算装置从备用状态切换到作用中状态时,确定为最新在作用中。当第二状态是最新在作用中时,确定唯一地识别上一个发布的事件块对象的上一个发布的事件块对象标识符。从可由计算装置存取的非暂时性计算机可读媒体选择下一个事件块对象。下一个事件块对象具有大于经确定的上一个发布的事件块对象标识符的事件块对象标识符。将选定的下一个事件块对象发布到向外消息传递网络装置。当计算装置的第二状态并非最新在作用中时,将经接收事件块对象发布到向外消息传递网络装置。当计算装置的第一状态是备用时,将经接收事件块对象存储在非暂时性计算机可读媒体中。

[0163] 图11是根据一些方面的用于产生及使用机器学习模型的处理过程的实例的流程图。机器学习是与可从数据学习、对数据分类及进行关于数据的预测的数学模型有关的人工智能分支。可被称为机器学习模型的此类数学模型可在两个或更多个类别当中对输入数据进行分类；在两个或更多个群组当中聚类输入数据；基于输入数据预测结果；识别输入数据中的模式或趋势；识别输入数据在空间中的分布；或这些的任何组合。机器学习模型的实例可包含：(i) 神经网络；(ii) 决策树，例如分类树及回归树；(iii) 分类器，例如朴素偏置分类器、逻辑回归分类器、脊回归分类器、随机森林分类器、最小绝对收缩及选择器 (least absolute shrinkage and selector) (LASSO) 分类器及支持向量机；(iv) 聚类器 (clusterer)，例如k均值聚类器、均值位移聚类器及频谱聚类器；(v) 因子分解器 (factorizer)，例如因子分解机、主分量分析器及核心主分量分析器；及(vi) 机器学习模型的集合或其它组合。在一些实例中，神经网络可包含深度神经网络、前馈神经网络、递归神经网络、卷积 (convolutional) 神经网络、径向基底函数 (RBF) 神经网络、回波状态神经网络、长短期记忆神经网络、双向递归神经网络、闸控神经网络、阶层式递归神经网络、随机神经网络、模块化神经网络、脉冲神经网络、动态神经网络、级联神经网络、神经模糊神经网络或这些的任何组合。

[0164] 不同机器学习模型可互换地使用以执行任务。可至少部分使用机器学习模型执行的任务的实例包含各种类型的评分；生物信息学；化学信息学；软件工程学；诈欺检测；客户分段；产生在线推荐；自适应网站；确定客户终身价值；搜索引擎；实时地或近实时地投放广告；对DNA序列进行分类；情感计算；执行自然语言处理及理解；对象辨识及计算机视觉；机器人运动；玩游戏；优化及元启发式；检测网络入侵；医疗诊断及监测；或预测资产 (例如机器) 何时将需要维护。

[0165] 任何数目及组合的工具可用于建立机器学习模型。用于建立及管理机器学习模型的工具的实例可包含 SAS® Enterprise Miner、SAS® 快速预测建模器及 SAS® 模型管理器、SAS云分析服务 (CAS) ®、SAS Viya ®，全部其是由北卡罗莱纳州喀里市 (Cary, North Carolina) 的 SAS Institute 公司提供。

[0166] 机器学习模型可通过被称为训练的至少部分自动化 (例如，很少或没有人类参与) 的处理过程建构。在训练期间，输入数据可迭代地供应到机器学习模型，以使机器学习模型能够识别与输入数据有关的模式或识别输入数据与输出数据之间的关系。运用训练，机器学习模型可从未经训练状态变换为经训练状态。输入数据可分割成一或多个训练集及一或多个验证集，且训练处理过程可重复多次。分割可遵循k折交叉验证规则、留规则、留p规则或箝制规则。下文关于图11的流程图描述训练及使用机器学习模型的概述。

[0167] 在框1104中，接收训练数据。在一些实例中，训练数据是从由各种数据子集建构的远程数据库或本地数据库接收，或由用户输入。训练数据可以其原始形式用于训练机器学习模型，或被预处理成接着可用于训练机器学习模型的另一形式。例如，训练数据的原始形式可被平滑化、截断、汇总、聚类或以其它方式操纵成接着可用于训练机器学习模型的另一形式。

[0168] 在框1106中，使用训练数据训练机器学习模型。可以监督式、无监督式或半监督式方式训练机器学习模型。在监督式训练中，使训练数据中的每一输入与期望输出相关。此期望输出可为标量、向量、或不同类型的数据结构，例如文字或图像。此可使机器学习模型能

够学习输入与期望输出之间的映射。在无监督式训练中,训练数据包含输入但无期望输出,使得机器学习模型必须独自找出输入中的结构。在半监督式训练中,仅使训练数据中的一些输入与期望输出相关。

[0169] 在框1108中,评估机器学习模型。例如,可例如经由用户输入或从数据库获得评估数据集。评估数据集可包含与期望输出相关的输入。可将输入提供到机器学习模型且可比较来自机器学习模型的输出与期望输出。如果来自机器学习模型的输出与期望输出紧密对应,那么机器学习模型可具有高准确度。例如,如果来自机器学习模型的90%或更多的输出与评估数据集中的期望输出相同,那么机器学习模型可具有高准确度。否则,机器学习模型可具有低准确度。90%的数量仅为实例。实际及期望准确度百分比取决于问题及数据。

[0170] 在一些实例中,如果机器学习模型针对特定任务具有不足准确度,那么处理过程可返回到框1106,其中可使用额外训练数据进一步训练或以其它方式修改机器学习模型以改进准确度。如果机器学习模型针对特定任务具有足够准确度,那么处理过程可继续到框1110。

[0171] 在框1110中,接收新数据。在一些实例中,新数据是从由各种数据子集建构的远程数据库或本地数据库接收,或由用户输入。机器学习模型可能不知晓新数据。例如,机器学习模型先前可能未处理或分析新数据。

[0172] 在框1112中,使用经训练的机器学习模型来分析新数据且提供结果。例如,可将新数据作为输入提供到经训练的机器学习模型。经训练的机器学习模型可分析新数据且提供包含新数据到特定类别中的分类、新数据到特定群组中的聚类、基于新数据的预测或这些的任何组合的结果。

[0173] 在框1114中,对结果进行后处理。例如,可作为作业的部分将结果与其它数据相加、相乘或以其它方式组合。作为另一实例,可将结果从第一格式(例如时间序列格式)转换成另一格式(例如计数序列格式)。可在后处理期间对结果执行任何数目及组合的操作。

[0174] 机器学习模型的更特定实例是图12中所展示的神经网络1200。神经网络1200被表示为可在彼此之间交换数据的经互连神经元(例如神经元1208)的多个层。所述层包含用于接收输入数据的输入层1202、隐藏层1204及用于提供结果的输出层1206。隐藏层1204因其在神经网络1200的正常运作期间可能无法直接观察或无法使其输入可直接存取而被称为隐藏。尽管出于示范性目的将神经网络1200展示为具有特定数目个层及神经元,但神经网络1200可具有任何数目及组合的层,且每一层可具有任何数目及组合的神经元。

[0175] 神经元及神经元之间的连接可具有可在训练期间调谐的数值权重。例如,可将训练数据提供到神经网络1200的输入层1202,且神经网络1200可使用训练数据来调谐神经网络1200的一或多个数值权重。在一些实例中,可使用反向传播来训练神经网络1200。反向传播可包含基于神经网络1200的实际输出与神经网络1200的期望输出之间的差异来确定特定数值权重的梯度。基于梯度,可更新神经网络1200的一或多个数值权重以减少差异,借此增加神经网络1200的准确度。此处理过程可重复多次以训练神经网络1200。例如,此处理过程可重复数百或数千次以训练神经网络1200。

[0176] 在一些实例中,神经网络1200是前馈神经网络。在前馈神经网络中,每一神经元仅将输出值传播到神经网络1200的后续层。例如,在前馈神经网络中,数据可仅在从一个神经元到下一神经网络的一个方向(前向)上移动。

[0177] 在其它实例中,神经网络1200是递归神经网络。递归神经网络可包含一或多个反馈回路,而允许数据通过神经网络1200在前向及后向两者上传播。此可允许信息持续存在于递归神经网络内。例如,递归神经网络可至少部分基于递归神经网络前已看到的信息来确定输出,从而给予递归神经网络以使用先前输入来告知输出的能力。

[0178] 在一些实例中,神经网络1200通过以下步骤来操作:从一个层接收数字向量;使用数值权重矩阵、非线性度或两者将数字向量变换成新数字向量;及将新数字向量提供到神经网络1200的后续层。神经网络1200的每一后续层可重复此处理过程,直到神经网络1200在输出层1206处输出最终结果。例如,神经网络1200可在输入层1202处接收数字向量作为输入。神经网络1200可将数字向量与数值权重矩阵相乘以确定经加权向量。可在神经网络1200的训练期间调谐数值权重矩阵。神经网络1200可使用非线性度(例如S型正切或双曲正切)来变换经加权向量。在一些实例中,非线性度可包含可使用以下等式表达的线性整流单元(rectified linear unit):

[0179] $y = \max(x, 0)$

[0180] 其中y是输出且x是来自经加权向量的输入值。经变换输出可供应到神经网络1200的后续层,例如隐藏层1204。神经网络1200的后续层可接收经变换输出,将经变换输出与数值权重矩阵及非线性度相乘,且将结果提供到神经网络1200的又一层。此处理过程继续直到神经网络1200在输出层1206处输出最终结果。

[0181] 本公开的其它实例可包含具有任何数目及组合的特性的任何数目及组合的机器学习模型。所述机器学习模型可以监督式、半监督式或无监督式方式或这些的任何组合来训练。所述机器学习模型可使用单个计算装置或多个计算装置(例如上文所论述的通信网络计算系统400)来实施。

[0182] 至少部分通过使用机器学习模型来实施本公开的一些实例可减少由计算装置在分析数据时消耗的处理迭代的总次数、时间、存储器、电力或这些的任何组合。例如,神经网络可比其它方法更容易识别数据的模式。此可使神经网络能够使用比其它方法更少的处理循环及更少的存储器来分析数据,同时获得类似或更高的准确度。

[0183] 一些机器学习方法可用机器学习特定处理器(例如,并非通用CPU)更高效地且更快速地执行及处理。当相较于通用CPU时,此类处理器也可提供能源节省。例如,这些处理器中的一些可包含图形处理单元(GPU)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、人工智能(AI)加速器、神经计算核心、神经计算引擎、神经处理单元、用于深度学习的特制(purpose-built)芯片架构,及/或使用半导体(例如,硅(Si)、砷化镓(GaAs))装置实施机器学习方法或一或多个神经网络的一些其它机器学习特定处理器。此外,这些处理器也可用于具有若干及多种不同类型的核心、引擎、节点及/或层的异质计算架构中,以在相较于采用CPU用于通用计算的均质计算架构时,在系统的各个部分每一处实现各种能量效率、处理速度改进、数据通信速度改进,及/或数据效率目标及改进。

[0184] 本技术的一或多个实施例对于进行关于用于将材料结合在一起的工艺(例如,接合操作)的预测是有用的。存在用以将材料结合在一起的许多不同方式,包含粘合剂接合、熔接接合、针脚式接合及焊接接合。接合操作用于包含制造及建设的许多不同工业中。接合可用于制造各种产品及其封装,例如车辆、用户装备、移动装置、计算机、医疗装置、服装及化学产品。接合可用于建设中,例如形成建筑物、桥梁及电系统。在接合操作中,可能难以在

不破坏接合件本身的情况下知晓接合件的质量及因此经形成产品的质量。关于接合操作的预测对于在不破坏经制造产品的情况下进行关于接合操作的确定可为有用的。例如,一或多个实施例对于制造集成计算机(IC)芯片是有用的。

[0185] 如所属领域的技术人员所知,用于IC芯片的半导体制造过程通常在常被称为“前段”及“后段”的两个阶段中执行。前段指代用于从空白晶片制造硅晶片的制造过程。经完成,经制造的硅晶片便具有在其上的多个IC芯片。“后段”指代在已在硅晶片上建立所有IC芯片后发生的制造过程。出于描述实例实施例的目的,将更详细描述实例“后段”过程。

[0186] 图13及14说明根据一个实施例的用于制造IC芯片的实例“后段”半导体制造过程的概览。特定来说,图13说明用于执行半导体制造过程的“后段”的实例方法1300。图14进一步用图形说明用于执行后段半导体制造过程的方法1400及与半导体制造过程的后段相关联的一些组件。通过实例,此类组件包含其上具有多个裸片1404的硅晶片1402。每一裸片1404包括功能电路制造于其上的IC芯片内部的半导体材料块。因此,每一裸片1404是尚未封装的全功能IC芯片。

[0187] 如图13中所见,实例方法1300开始于晶片研磨阶段(方框1302)。在此阶段中,通过磨料磨轮将硅晶片(例如,图14的硅晶片1402)的厚度减小到期望厚度(例如,50 μm 到75 μm)。例如,可在多个步骤中实施研磨。例如,在第一步骤中,磨轮(未展示)可使用粗砂首先粗糙地研磨硅晶片(例如,图14的硅晶片1402)且移除其的大部分过量厚度。在第二步骤中,磨轮可使用不同砂(例如,细得多的砂)来完精磨硅晶片。更细砂用于将硅晶片准确地研磨到期望厚度且抛光硅晶片。

[0188] 一旦研磨完成,图13的方法1300便在晶片锯切阶段(方框1304)中将裸片(例如,图14的裸片1404)彼此分离。也被称为“晶片切割”的晶片锯切是通过其在硅晶片(例如,图14的硅晶片1402)上将裸片彼此分离的工艺。一般来说,晶片切割是通过在位于裸片之间的区域中机械地锯切硅晶片来完成。如图14中所见,用于锯切的实例“区域”分别使用垂直及水平线1406、1408来指示,且通常被称为“切割道”或“划线道”。可通过额外锯切或研磨来进一步移除或减少裸片周围的过量区域。

[0189] 接着,在图13的裸片接合阶段(方框1306)中将经分离裸片(例如,图14的裸片1404)附接或“接合”到引线框架(例如,图14的引线框架1410)。“引线框架”可为在芯片封装内部的金属(例如,铜)导电结构,其经配置以将信号从裸片载送到“外部世界”或反之亦然。在图14中,裸片1404是使用环氧树脂粘合剂或焊料接合到引线框架1410。特定来说,将环氧树脂粘合剂或焊料以预定图案放置于引线框架1410上。接着,将在晶片锯切阶段中彼此分离的个别裸片1404a、1404b、1404c拾取及放置于经施配的环氧树脂粘合剂或焊料上且进行固化。固化使环氧树脂或焊料硬化,使得每一裸片1404a、1404b、1404c保持附接到引线框架1410,且使得每一裸片1404a、1404b、1404c实现期望机械及电性质。

[0190] 所属领域的技术人员将容易了解,本公开并不仅限于使用环氧树脂粘合剂或焊料用于将裸片1404接合到引线框架1410。在其它实施例中,例如,裸片1404是使用熟知热压裸片接合技术接合到引线框架1410。热压裸片接合未使用粘合剂材料将裸片1404附接到引线框架1410。实际上,对裸片1404施加热量及力以与下层引线框架1410形成金属接合。

[0191] 图13的后段工艺中的下一阶段是导线接合阶段(方框1308)。在此阶段中,例如,导线接合机(未展示)使用多条导线建立裸片(例如,裸片1404a)与引线框架(例如,引线框架

1410)之间的电连接。如下文更充分描述,每一连接涉及在本文中被称为“球形接合件”及“针脚式接合件”的两种接合件。球形接合件是在导线与裸片之间形成,而针脚式接合件是在引线框架导线之间形成。可在此阶段中使用任何数目条导线来将裸片接合到引线框架。例如,取决于制造商,可使用不同数目条导线(例如,>5、>10、>20或>100)来将裸片电连接到引线框架。

[0192] 一旦导线接合阶段完成,便在图13中执行模制阶段(方框1310)。此阶段有助于在机械上(例如,免受由一或多个物理冲击造成的损坏)以及在环境上(免受颗粒及其它杂质影响)保护裸片(例如,裸片1404)。如图14中所见,模制阶段通过将裸片1404包覆于预定材料(举例来说,例如塑料环氧树脂材料)中来囊封裸片1404。例如,模制工艺可为若干熟知模制方法中的任何者。例如,在第一方法(即,密闭式方法)中,将陶瓷板或金属盖附接到裸片1404,借此形成密封。在第二方法中,将塑料环氧树脂材料熔融到裸片1404上且进行固化以形成密封。但是,与在模制阶段中使用的特定方法无关,模制化合物都使用热量来固化。

[0193] 在图13中的模制阶段后的下一阶段可包含增强阶段(方框1312)。例如,镀膜操作可用于通过在引线框架(例如,引线框架1410)的引线上方施覆薄金属层以将引线框架机械且电连接及耦合到下层衬底(举例来说,例如印刷电路板(PCB))来帮助保护引线框架1410免受腐蚀及磨损。作为另一实例,蚀刻操作可用于增强引线框架的导电性。

[0194] 在增强阶段完成后,在图13中执行标记阶段(方框1314)。在此阶段中,制造商用多种标识符及其它区别信息(例如制造商的名称及标志、IC芯片的名称、日期代码、批次标识符及类似者)来标记IC芯片的外壳。可使用任何数目种方法(举例来说,例如基于油墨及基于激光的方法)来标记外壳。但是,在本技术的至少一个实施例中,基于激光的方法是优选的。但是,与在标记阶段中使用的特定方法无关,标记都能够识别IC芯片及来源且促进可追溯性。

[0195] 后段工艺中的下一阶段是图13中的修整及成型阶段(方框1316)。在此阶段中,切除通常被称为“分隔件(dambar)”的连接从IC芯片的外壳延伸的引线的金属支柱。另外,将引线弯曲或以其它方式形成为适当形状用于放置到下层衬底上。最后,在图13的阶段(方框1318)中,检验、测试及封装经修整且经成型的IC芯片以进行分配。

[0196] 图13及14展示可作为半导体制造过程的部分执行的实例阶段。每一阶段可具有一或多个操作。此外,可执行不同、更多或更少的阶段(例如,阶段可组合或按不同顺序)。因此,图13及14中所说明的方法1300及1400仅为用于描述IC芯片的实例接合操作的实例。无论如何,如先前所描述,导线接合机可在导线接合阶段(方框1308)中使用多条导线以在图14的裸片1404与引线框架1410之间建立电连接。图15A及15B说明根据本技术的至少一个实施例的IC芯片的一些组件及接合件的实例。特定来说,图15A及15B说明接合到下层衬底1500的示范性引线框架1410。引线框架1410的导电“支脚”或“引脚”也附接到衬底1500。在此实施例中,图15A将引线框架1410说明为具有三(3)个导电支脚1410a、1410b、1410c。但是,此仅用于说明性目的。所属领域的技术人员将容易了解,可根据需要或期望存在耦合到引线框架1410的更多或更少支脚。但是,与导电支脚的数目无关,如图15B中所见,引线框架1410的支脚1410a、1410b、1410c可在修整及成型阶段(图13的方框1316)期间弯曲,使得导电支脚与下层衬底(举例来说,例如印刷电路板(PCB))形成适当物理接触。

[0197] 如图15A中所见,金属细线1504延伸于安置在裸片1404a的表面上的接合垫1506与

安置在引线框架1410的每一支脚1410a、1410b、1410c上的目标垫1508之间。在一些实施例中,导电膜层1502可安置在目标垫1508与引线框架1410的每一支脚1410a、1410b、1410c之间。接合垫1506及目标垫1508两者都包括薄的导电金属层或膜,举例来说,例如银。但是,如所属领域的技术人员将理解,接合垫1506及目标垫1508中的一者或两者可包括除银外的导电材料。但是,与其组合物无关,导线1504都经由“球形接合件”1510附接到接合垫1506,且经由“针脚式接合件”1512附接到目标垫1508。导线1504、球形接合件1510及针脚式接合件1512提供在裸片1404a与引线框架1410的支脚之间建立电连接的物理构件。

[0198] 图16A到16D说明根据本公开的一个实施例的用于形成球形接合件1510及针脚式接合件1512的实例方法1600。如这些图中所见,进给导线1604延伸穿过毛细管工具1602的孔。一般来说,进给导线1604包括导电材料,举例来说,例如金、铜或铝。

[0199] 为了产生球形接合件1510,将高压电荷施加到进给导线1604(图16A)。此电荷熔融进给导线1604的末端。但是,表面张力而非滴落或掉落引起经熔融的进给导线1604在毛细管工具1602的尖端处形成为球1606。接着,朝向裸片1404a降低毛细管工具1602,使得将熔融球1606按压到在一些情况下可被加热的接合垫1506上(图16B)。例如,接着连同毛细管工具1602的向下压力一起施加热量(例如,热压)及/或超声波能量(例如,超声波能量)以形成球形接合件1510。

[0200] 接着,将毛细管工具1602提离球形接合件1510且使其朝向引线框架1410的目标垫1508移动。接着,毛细管工具1602下降以接触目标垫1508,借此挤压毛细管工具1602的尖端与目标垫1508的表面之间的进给导线1604(图16C)。由毛细管工具1602施加于进给导线1604上的压力连同热量及/或超声波能量一起形成针脚式接合件1512。如此形成后,接着将毛细管工具1602提离目标垫1508,此引起进给导线1604撕裂,且移动到下一接合垫1506以形成下一球形接合件1510(图16D)。

[0201] 导线接合是非常复杂的过程,其因在IC芯片上可能有多个裸片而变得甚至更加复杂。本公开的实施例读取及记录在接合期间的多个工艺参数。根据本实施例且如下文更充分描述,工艺参数可用于产生指示在此类接合操作期间发生异常的风险的异常指示符。

[0202] 图17是已经历先前描述的半导体制造过程的后段的实例IC芯片1700的透视图。如图17中所见,IC芯片1700包括多个裸片1404a、1404b、1404c及1404d。每一裸片1404a、1404b、1404c、1404d具有其自身到引线框架1410的接合连接,且因此分别经由导线1504以及球形及针脚式接合件1510、1512物理且电连接到引线框架1410,如先前所描述。此实施例说明40条导线1504的导线布局。但是,所属领域的技术人员应容易了解,此仅用于说明性目的。根据需要或期望,此类导线布局可包括更多或更少导线1504。

[0203] 在一个实施例中,裸片1404a、1404b、1404c、1404d全部具有将其连接到引线框架1410的相同数目条导线1504。但是,此并非必需的。裸片1404a、1404b、1404c及1404d可不具有相同数目条导线1504。例如,裸片1404a可具有将其连接到引线框架1410的第一数目条导线1504,且裸片1404c可具有将其连接到引线框架1410的第二、不同数目条导线1504。其它裸片1404b、1404d中的一者或两者也可具有与IC芯片1700上的其它裸片相同或不同数目的导线。

[0204] 如上文所说明,半导体制造过程的后段可包含检验及测试阶段(例如,图13中的方框1318)。在此阶段期间,测试在导线接合阶段(例如,图13中的方框1308)中产生的多个球

形接合件1510及针脚式接合件1512的质量。此测试并非针对所有已完成IC芯片执行,而是针对来自每一批次的IC芯片的代表性样本执行。

[0205] 一般来说,存在用于确定导线接合件的质量的不同类型的测试(非破坏性及破坏性)。非破坏性测试大体上涉及球形接合件1508及针脚式接合件1512的视觉检验,且通过所述接合件的相应直径来推断其的质量。如果其通过所述视觉检验,那么可将IC芯片放回到工作流程中以进行分配。但是,视觉检验可为用以推断接合件的质量的不准确的、易于出错的方式。

[0206] 相比之下,破坏性测试通过损坏球形及针脚式接合件1510、1512及/或导线1504或使其变形而破坏接合件,使得待检验的IC芯片无法复原以进行分配。因此,尽管破坏性测试对于确定给定球形及/或针脚式接合件1510、1512的精确故障点是有用的,但其必然耗时且昂贵,且通常由单独测试机器进行。此外,由于接合件被破坏,所以其仅可用于进行关于未经历破坏性测试的其它经接合芯片的质量的推断。

[0207] 一般来说,半导体制造行业依靠两种特定破坏性测试。这些是“球剪切”测试及“针脚拉拔”测试。对于关于这些测试的细节,感兴趣的读者请参阅微电子装置的政府规定标准(例如,2016年4月25日针对军事及航空航天电子系统发行的标题为“测试方法标准-微电路(Test Method Standard-Microcircuits)”的MIL-STD-883C的修订版K中的美国政府标准,其全文以引用的方式并入本文中)。但是,简单来说,“球剪切”测试用于测试球形接合件1508的质量,而“针脚拉拔”测试用于测试针脚式接合件1512的质量。

[0208] 图18A到18C是说明用于执行球剪切测试的实例方法1800的透视图。如这些图中所见,球形接合件1510接合到接合垫1506(图18A)。工具臂1802定位于接合垫1506上方且横向移动以与球形接合件1510按压接触(图18B)。工具臂1802继续抵着球形接合件1510横向移动,直到由工具臂1802施加的力引起球形接合件1510与接合垫1506分离(即,剪切)(图18C)。剪切破坏接合件。在整个测试中测量由工具臂1802施加的剪切力(例如,以克为单位)。因此,在测试结束时,将球形接合件1510从接合垫1506剪切掉所需的剪切力的精确量是已知的。

[0209] 图19A到19C是说明用于执行针脚拉拔测试的实例方法1900的透视图。如上文,针脚式接合件1512附接到目标垫1508且球形接合件1510附接到接合垫1506。

[0210] 在针脚拉拔测试中,降低钩工具1902且将其放置于导线1504下方(图19A)。如此定位后,接着将钩工具1902升高以与导线1504接触(图19B)。钩工具1902继续向上移动使得其在导线1504上施加向上力。最终,此向上力将针脚式接合件1512拉离目标垫1508,从而有效地将导线1504拉离引线框架1410(图19C)且破坏接合件。在整个测试中测量由钩工具1902施加的向上力(例如,以克为单位)。因而,在测试结束时,可确定将针脚式接合件提离目标垫1508所需的向上力的精确量。

[0211] 为了执行这些破坏性测试,将选定IC芯片从执行导线接合阶段的任何机器物理移动到经配置以执行测试的另一机器。通常,这些动作手动地执行。此外,并未测试给定IC芯片的所有球形接合件1510及针脚式接合件1512。相比之下,仅测试所述接合件的代表性样本。

[0212] 由于这些测试如此有破坏性,所以待测IC芯片被破坏,即使并未测试IC芯片上的所有接合件。因此,芯片制造商通常取来自给定批次的IC芯片的代表性样本且接着对所述

代表性样本执行这些破坏性测试。接着,存储从测试(若干)样本实现的结果且随后将其用于进行关于所述批次中的IC芯片的质量的推断。

[0213] 但是,此类测试非常耗时且可能减慢集成电路制造过程。此外,由于这些测试破坏IC芯片,所以所述IC芯片必须被丢弃而非出售。因此,破坏性测试也为昂贵的。

[0214] 另外,关于应选择给定IC芯片上的哪些导线1504用于测试,出现问题。理想地,具有不良质量导线及/或接合件的IC芯片应在离开厂区前被识别。但是,无法保证IC芯片及其对应导线及接合件的给定随机样本将包含此类IC芯片。此外,IC芯片及导线的随机样本仍可能不正确地证明给定批次具有可接受质量。目前,不存在用于识别具有不良质量接合连接的IC芯片的可靠方法。因此,目前没有用以识别特定导线或导线子集以进行质量测试的方式。

[0215] 在其它类型的接合情形中出现类似问题。因而,接合器将必须选择破坏性测试(其破坏研究中的接合件)或非破坏性测试。用于非破坏性测试的先前常规测试技术有可能为不准确的且耗时的(其需要手动、视觉检验)。此外,先前测试方法通常仅在给定批次中的产品的小子集上进行。因此,确定测试哪些产品通常存在问题。

[0216] 因此,本公开的实施例通过提供专门用于在不破坏接合件的情况下预测可能异常(例如,在不进行质量保证测试的情况下预测从接合垫1506剪切球形接合件1510及将针脚式接合件1512与目标垫1508分离所需的力)的模型来解决这些及其它问题。技术可使用特征工程学来选择及变换从原始数据测量值导出的相关变量以处理模型(例如,使用机器学习或统计建模指定的预测模型)的数据。特征工程学可涉及通过识别将为预测模型的有用预测变量的原始变量及/或由操纵原始变数(例如,加法、减法、乘法及比率)建立导出特征来建立特征。特征工程学可涉及通过进一步操纵经建立的预测量变量以执行模型性能(例如,确保变量在相同尺度上,或在模型的可接受范围内)而进行的变换。特征工程学可涉及通过从原始数据自动建立新变量而进行的特征提取;及通过使用算法对特征进行分析、判断及排名以确定模型的特征而进行的特征选择(例如,用以移除不相关或冗余特征)。

[0217] 特定来说,本文中所公开的实施例利用在接合操作(例如,导线接合)期间测量的工艺数据来开发分析模型。通过特征工程学获得的工艺参数可在本文中被称作“运动特征”。分析模型允许用户准确地估计目前为止仅可通过破坏性测试获得的球剪切及针脚拉力(即,值)。因此,本实施例可基于运动特征数据准确地预测导线接合件的质量。

[0218] 此类准确预测提供常规测试方法未提供的益处及优点。例如,根据本实施例产生的分析模型可减少所执行的检验量且在一些例子中可将其完全消除。甚至在其中出于收缩原因需要定量的检验或破坏性测试的情形中,本文中的方法仍可使制造商能够在产品群体中具有策略性以进行取样。例如,可从用于装运的群体及因此从经取样群体废弃由本实施例预测为具有不良质量接合件的IC芯片。对优选群体进行取样可提供现场质量的总体改进且可改进效率(例如,通过减少召回或拒绝经装运及经取样组的机会)。

[0219] 如先前所陈述,本文中所描述的一或多个实施例仅作为实例在集成电路的接合的上下文中进行说明。所属领域的技术人员将了解,本文中所描述的技术可应用于其它类型的接合操作。

[0220] 图20A说明训练系统2000的实例框图。例如,训练系统对于训练计算机模型以进行关于接合操作的预测可为有用的。例如,接合操作可为本文中所描述的接合操作,例如粘合

剂接合操作(例如,用以将挡风玻璃粘着到车辆或在晶片接合中)或熔接接合(例如,用于导线接合),仅举几例。存在许多类型及子类型的接合。例如,关于熔接接合,存在包含以下的子类型:钨极惰性气体(TIG)熔接,也被称为氩弧及气体钨极电弧熔接(GTAW);包芯焊条电弧熔接(FCAW);棒熔接,也被称为屏蔽金属电弧熔接(SMAW);金属惰性气体(MIG)熔接,也被称为气体金属电弧熔接(GMAW);激光束熔接;电子束熔接;等离子体电弧熔接;原子氢熔接;及电渣熔接。为简洁起见,本文中不同的可能接合操作或混合接合操作简称为接合操作。

[0221] 训练系统2000包含计算装置2002。训练系统2000经配置以在系统中的装置之间交换信息(例如,经由有线及/或无线传输)。例如,网络(未展示)可将训练系统2000的一或多个装置连接到训练系统2000的一或多个其它装置。

[0222] 例如,在一或多个实施例,训练系统2000包含用于经由一或多个输入接口2005接收与接合操作2040A有关的信息(例如,用于训练计算装置2002)的一或多个输入装置2004A。例如,如果训练系统正在训练系统以进行关于接合操作的预测时,接合操作2040A可包含用于接合操作的系统。接合操作在许多不同环境(例如,制造或建设环境)中发生。例如,在集成电路制造环境中,接合可用于将一或多条导线接合到表面(例如,使用球形、楔形、针脚式及/或柔性凶兆)。接合操作2040A可包括将第一组的多条导线的导线接合到第一组表面的对应表面以形成集成电路芯片。接合操作2040A可具有不同类型的接合操作。例如,环境2040A可包含球形接合操作2042A及/或针脚式接合操作2044B。

[0223] 在一或多个实施例,计算装置2002可接收用于建置模型以表示接合操作2040A或后续环境的模型信息2030。作为实例,计算系统可接收训练数据2032用于建置模型。

[0224] 训练数据2032可包含接合操作2040A的测量值2038。额外地或替代地,训练数据2032可包含工艺数据2034,工艺数据可从测量值2038产生。例如,工艺数据可包括测量值2038的测量值或为针对一或多个测量值或测量类型产生的信息。例如,如果接合操作2040A是物理环境的一部分,那么输入装置(例如传感器2020A)可捕捉关于物理环境的测量值2038,例如用于加热经接合表面或接合剂的热量测量值及/或用于接合材料的力的力测量值。替代地或替代地,环境可为经模拟环境或包括用于物理环境的额外经模拟动作或计算(例如,接合操作2040A或来自测量值2038的经推断信息的经模拟方面),且计算系统组件(例如,计算装置2002或计算系统2024)可产生关于物理或经模拟环境的信息(例如,工艺数据2034)。额外地或替代地,计算装置2002本身可导出工艺数据2034(例如,使用工艺数据应用程序2012)。

[0225] 额外地或替代地,训练数据2032包括接合的状态。作为实例,状态的每一状态可包括与检测接合操作2040A中的一或多个异常有关的目标的一或多个候选结果。例如,如果接合操作2040A包括用以将第一组的多条导线接合到第一组表面的操作,那么状态可为多条导线在接合到所述第一组表面后的状态(例如,质量状态或性能状态)。候选选项可包括接合操作2040A中的集成电路芯片的一或多个有缺陷芯片结果。候选选项也可在个别导线级上(例如,个别导线级的质量保证测试结果)。

[0226] 另外,训练系统2000包含用于经由一或多个输出接口2006输出基于接合操作2040A的信息的一或多个输出装置2052。例如,计算装置2002可输出基于与接合操作2040A有关的训练数据2032实施预测模型的机器学习模型2050。例如,计算装置2002可产生工艺数据的一或多个权重,使得输入到机器学习模型的工艺数据预测目标(例如,与检测接合操

作2040A中的一或多个异常有关的目标)的一或多个候选结果。

[0227] 作为实例,可通过监督式学习基于接收训练数据且将其用于产生一或多个权重使得机器学习模型可预测候选结果来训练机器学习模型2050。例如,如果针对导线接收一组测量值且接着测试所述导线且给出正常或异常的状态,那么可对测量值进行加权以预测所述状态为正常或异常,以用于其中状态可能未知(例如,其中测试将破坏或损坏来自接合操作的接合件)的未来测量。

[0228] 计算装置2002具有计算机可读媒体2010及处理器2008。计算机可读媒体2010是信息的电子保存空间或存储器,因此信息可由处理器2008存取。计算机可读媒体2010可包含但不限于任何类型的随机存取存储器(RAM)、任何类型的只读存储器(ROM)、任何类型的快闪存储器等,例如磁性存储装置(例如,硬盘、软盘、磁条)、光学盘片(例如,光盘(CD)、数字多功能光盘(DVD))、智能卡、快闪存储器装置等。

[0229] 处理器2008执行(例如,存储在计算机可读媒体2010处的)指令。指令可由专用计算机、逻辑电路或硬件电路实行。在一或多个实施例中,处理器2008是在硬件及/或固件中实施。处理器2008执行指令,此意谓其执行或控制由所述指令呼叫的操作。术语“执行”是运行应用程序或实行由指令呼叫的操作的处理过程。指令可使用一或多种编程语言、脚本语言、汇编语言等撰写。在一或多个实施例中,处理器2008可从永久性存储器装置检索指令集且将指令以可执行形式复制到例如大体上为某一形式的RAM的暂时性存储器装置。处理器2008与计算装置2002的组件(例如,输入接口2005,与输出接口2006及与计算机可读媒体2010)可操作地耦合以接收、发送及处理信息。

[0230] 在一或多个实施例中,计算机可读媒体2010存储指令以供处理器2008执行。例如,在一或多个实施例中,计算机可读媒体2010包括使工艺数据应用程序2012选择或导出工艺数据2034且使机器学习模型应用程序2014导出机器学习模型2050的指令,如本文中所描述。

[0231] 在一或多个实施例中,存储在计算机可读媒体2010上的一或多个应用程序是在存储在计算机可读媒体2010中且可由处理器2008存取以用于执行指令的软件(例如,计算机可读及/或计算机可执行指令)中实施。一或多个应用程序可与其它分析工具(例如由美国北卡罗莱纳州喀里市的SAS Institute公司提供的分析工具)集成。仅出于说明,应用程序是使用一或多个SAS软件工具实施或与一或多个SAS软件工具集成,所述一或多个SAS软件工具例如JMP®、Base SAS、SAS®Enterprise Miner™、SAS/STAT®、SAS®事件流处理、SAS®高性能分析服务器、SAS®视觉数据挖掘及机器学习、SAS®LASR™SAS®数据库内产品、SAS®可扩展性能数据引擎、SAS®云分析服务、SAS/OR®、SAS/ETS®、SAS®库存优化、SAS®库存最佳化工作台、SAS®视觉分析、SAS®Viya™、用于Hadoop®的SAS存储器内统计、SAS®预报服务器及SAS/IML®,其全部由美国北卡罗莱纳州喀里市的SAS Institute公司开发及提供。

[0232] 在一或多个实施例中,更少、不同及额外组件可并入到计算装置2002或包括计算装置2002的系统中。例如,在一或多个实施例中,存在多个输入装置2004。在相同或不同实施例中,存在多个输出装置2052。作为另一实例,相同接口支持输入接口2005及输出接口2006两者。例如,触控屏幕提供用于用户输入及用于将输出呈现给用户的机制。替代地,输入接口2005具有使用相同或不同接口技术的多于一个输入接口。替代地或额外地,输出接

口2006具有使用相同或不同接口技术的多于一个输出接口。在一或多个实施例中,一或多个输入装置2004或一或多个输出装置2052的功能性集成到计算装置2002中。一或多个应用程序可组合或进一步分解成分开的应用程序。

[0233] 机器学习模型2050可由控制系统使用以进行关于另一环境的预测或控制另一环境。图20B说明用于接合操作2040B的控制系统2060的实例框图。在此实例中,接合操作2040A及接合操作2040B是不同接合操作。例如,接合操作2040A可为包括用以将第一组的多条导线接合到第一组表面的操作的第一组接合操作,且接合操作2040B可为包括用以将第二组的多条导线接合到第二组表面的操作的第二接合操作。第二组的多条导线可不同于第一组的多条导线。例如,第一组导线可能已被破坏以确定与接合操作2040A有关的状态,或接合操作2040A及2040B可在不同制造位置处或在不同时间执行。第二组表面可不同于第一组表面(例如,第二组表面可与不同于第一组表面的批次集成芯片有关)。本文中的实施例提供用于基于在先前接合操作中接收的训练进行关于接合操作预测的框架。使用此框架,无需破坏接合件以进行关于其质量的预测及/或控制接合操作以确保质量接合件。

[0234] 控制系统2060经配置以在控制系统2060中的装置之间交换信息(例如,经由有线及/或无线传输)。例如,网络(未展示)可将控制系统2060的一或多个装置连接到控制系统2060的一或多个其它装置。替代地或额外地,控制系统2060集成到一个装置中(例如,计算装置2002可在制造工厂或建设工地处且包括用以捕捉与接合操作2040A及接合操作2040B两者有关的信息的装备)。在此实例中,控制系统2060包括图20A的计算装置2002。在其它实例中,控制系统2060包括不同计算装置。例如,训练系统2000的计算装置可在接合操作2040A中且控制系统2060的另一计算装置可在另一环境(例如,不同制造工厂或建设工地)中。例如,计算装置2002可代替性地为输出装置2052或从输出装置2052接收机器学习模型2050。无论如何,计算装置2002可存取机器学习模型。在此实例中,计算机可读媒体2010具有在接合操作2040A的训练数据2032上训练的机器学习模型2050。

[0235] 在一或多个实施例中,控制系统2060包含用于经由一或多个输入接口2005接收与接合操作2040B有关的信息(例如,用于进行关于接合操作2040B的预测)的一或多个输入装置2004B。例如,计算装置2002可接收输入数据2070。输入数据2070可指示从接合操作2040B的测量值2072产生的工艺数据2074。

[0236] 输入数据2070可从与训练数据2032相同类型的测量值产生。例如,工艺数据2074可包括测量值2038的测量值或为针对一或多个测量值或测量类型产生的信息。例如,如果接合操作2040B是物理环境的一部分,那么输入装置(例如传感器2020B)可捕捉关于物理环境的测量值2072,例如用于加热经接合表面或接合剂的热量测量值及/或用于接合材料的力的力测量值。替代地或替代地,环境可为经模拟环境或包括用于物理环境的额外经模拟动作或计算(例如,接合操作2040B或来自测量值2072的经推断信息的经模拟方面),且计算系统组件(例如,计算装置2002或计算系统2024)可产生关于物理或经模拟环境的信息(例如,工艺数据2074)。额外地或替代地,计算装置2002本身可导出工艺数据2074。

[0237] 额外地或替代地,输入数据2070可包含更少、更多或与训练数据2032不同的测量类型或从其导出。例如,制造位置可更新机器学习模型2050以考虑更少、更多或不同的测量类型。

[0238] 在一或多个实施例中,控制系统2060包含用于经由一或多个输出接口2006输出基

于接合操作2040B的信息的一或多个输出装置2094。例如,信息可包括用于控制接合操作2040B的接合操作控制信息2090(例如,控制信息2090可归因于接合操作2040B的经预测质量问题而指示停止或调整接合操作)。例如,接合操作控制2090可包括指示接合操作2040B中的异常发生的风险的异常预测量2052。例如,如果接合操作2040B包括用以接合一组多条导线的导线以形成集成电路芯片的接合操作,那么异常预测量可指示接合操作2040B中的集成电路芯片制造过程中的异常的风险。

[0239] 在一或多个实施例中,计算装置具有用于产生异常预测量2052的计算机可读媒体2010及处理器2008。例如,在一或多个实施例中,计算机可读媒体2010包括用于以下的指令:使机器学习模型应用程序(例如,机器学习模型应用程序2014)存取在第一接合操作的训练数据上训练的机器学习模型(例如,在接合操作2040A上训练的机器学习模型2050),根据机器学习模型对输入数据2070进行加权,及基于根据机器学习模型对输入数据2070进行加权来产生异常预测量2052。

[0240] 在一或多个实施例中,计算系统(例如,计算装置2002、训练系统2000及/或控制系统2060)实施如本文中所描述的方法(例如,图21A到21C中所展示的方法)。例如,计算系统可为具有拥有用于观察环境(例如,制造或建设环境)及用于经由因特网与其它装置或系统交换数据(例如,来自环境的反馈)的传感器的装置的物联网(IoT)系统的一部分。

[0241] 图21A说明用于训练机器学习模型(例如,机器学习模型2050)的方法2100的实例流程图。方法2100包括用于接收训练数据(例如,训练数据2032)的操作2101。训练数据包括从第一接合操作的测量值产生的工艺数据。训练数据包括多条导线在接合到第一组表面后的状态。所述状态的每一状态包括与检测第一接合操作中一或多个异常有关的目标的一或多个候选结果。例如,机器学习模型可经训练以预测与在接合操作中产生的接合件的质量有关的方面。在此情况下,一或多个候选结果可包括用于第一接合操作中的个别导线或接合件的一或多个质量保证测试。作为另一实例,机器学习模型可经训练以预测与由于接合操作而产生的正常或有缺陷产品有关的方面。在此情况下,一或多个候选结果可包括第一接合操作中的集成电路芯片的一或多个有缺陷芯片结果。在其中接合操作包括多个接合件(例如,针脚式接合件及球形接合件)的案例中,可针对个别接合件(例如,针脚式接合件)、组件(例如,导线)或产品(例如,芯片)基于针脚式接合件及/或球形接合件的参数的测量值来进行预测。例如,在其中导线在两端处接合以形成单个路径的情形中,将两个接合件的参数包含于训练数据中可存在优点。

[0242] 方法2100包括用于产生工艺数据的一或多个权重,使得输入到机器学习模型的工艺数据预测目标的一或多个候选结果的操作2102。例如,加权可在模型内使用以使特定变量或观察对算法影响更大。例如,工艺数据的一或多个权重可经产生用于训练数据的梯度提升模型(例如,将权重0指派给经排除变量且将最高权重指派给最重要变量)。

[0243] 图21B是说明用于控制接合操作的实例方法2130的流程图。方法2130包括用于存取在第一接合操作的训练数据上训练的机器学习模型(例如,根据方法2100产生的机器学习模型)的操作2131。第一接合操作包括用以将第一组的多条导线接合到第一组表面的操作。通过监督式学习来训练机器学习模型。

[0244] 方法2130包括用于接收指示从第二接合操作的测量值产生的工艺数据的输入数据的操作2132。第二接合操作包括用以将第二组的多条导线接合到第二组表面的操作。第

二组的多条导线不同于第一组的多条导线。第二组表面不同于第一组表面。

[0245] 例如,输入数据可包括在第二接合操作期间接收的(例如,关于集成式计算机芯片的给定导线的)实时传感器测量值。传感器测量值可实现在没有人为监督或破坏接合件的情况下对接合过程质量的监督。但是,输入数据也可包含操作者观察。传感器测量值可包括在第二接合操作中涉及的接合系统的测量值。例如,仅举几例,测量值可包括热量测量值、功率相关测量值、力测量值、放电结球(EFO)测量值、超声波测量值、球针脚相关测量值及球形接合相关测量值。

[0246] 方法2130包括用于根据机器学习模型对输入数据进行加权的操作2133。例如,在训练数据上开发的模型可用于预测后续接合操作中的异常。方法2130包括用于基于根据机器学习模型对输入数据进行加权来产生指示第二接合操作中的异常发生的风险的异常预测的操作2134。例如,异常预测可为接合件(例如,导线与引线框架或裸片之间的有缺陷接合件)的组件的缺陷,或有缺陷产品(例如,由于例如浮动状况、材料污染(例如引线框架污染)及裸片倾斜中的一或多者的异常工艺状况)的预测。方法2130包括用于视需要输出异常预测以控制第二接合操作的操作2135。例如,异常预测可控制第二接合操作以校正第二接合操作中涉及的接合系统中的一或多个异常或减少所述一或多个异常的发生(例如,指示批次产品中的缺陷的可能性以提醒进一步调查所述批次或针对未来批次更改工艺)。

[0247] 在一或多个实施例中,异常预测预测第二接合操作中的个别导线或接合件中的异常,而未对所述个别导线或接合件执行破坏性质量保证测试。这样,在存在关于产品的接合件中的缺陷的某一指示前,无需破坏接合件。例如,本文中所描述的接合操作可包括球形接合操作,且破坏性质量保证测试包括用于测试球形接合件的球剪切测试,其将剪断球形接合件。作为另一实例,本文中所描述的接合操作可包括针脚式接合操作,且其中破坏性质量保证测试包括用于测试针脚式接合件的针脚拉拔测试,其将拉出针脚式接合件。

[0248] 在一或多个实施例中,可在后续接合操作期间改进预言算法。图21C是说明用于更新用以控制接合操作的机器学习模型的实例方法2160的流程图。

[0249] 方法2160包括用于接收指示异常预测正确地或不正确地预测在第二接合操作中制造的特定芯片中的异常的反馈的操作2161。方法2160包括用于基于反馈更新机器学习模型(例如,对输入数据进行不同地加权或选择不同或额外输入数据类型)的操作2162。

[0250] 方法2160视需要包括用于调整接合操作的操作2163。例如,当通过芯片制造系统执行第二接合操作时,操作2163包括基于以下中的一或多者来调整继由芯片制造系统执行的第二接合操作后的接合操作:指示第二接合操作中的异常发生的风险的异常预测;及指示异常预测正确地或不正确地预测在第二接合操作中制造的特定芯片中的异常发生的反馈。例如,可基于指示疑似错误的高百分比停止接合操作(例如,以检查过程组件或经制造产品)。可基于一致指示的问题来使用不同材料或用于接合材料的力。或者,可定位特定产品以用于校正怀疑有问题的产品。例如,在一或多个实施例中,输入数据包括用指示在第二接合操作中涉及的特定导线、裸片或芯片的标识符或位置中的一或多者的原始信息标记的经接收传感器测量值。异常预测识别在第二接合操作中发生的异常,且与原始信息相关以指示异常的位置。此对于校正个别有缺陷芯片或工艺可为有用的。

[0251] 例如,图22A是说明在运动特征模式与对应芯片在引线框架上的位置之间的关系

的图表2200。所展示的波形涉及在接合操作中从0到n的每一导线的连续工艺测量值(例如,力或功率测量值)。图表中的波形的位置与在设施中制造的集成电路的布置相关。图22B说明对应于图22A的图表的芯片在引线框架上的实例位置的图2250。在此实例中,图22B中的引线框架具有16个裸片(每垂直位置标识符有四个)。水平位置标识符及导线编号与图2250中的特定导线位置(例如,HID 1到HID 2,W1到Wn)相关。垂直位置标识符(VID)与图2250中的特定芯片(例如,VID 1到VID 4)相关。图形用户接口可用于显示图表及图(例如,图表2200及2250),以帮助开发模型(例如,接收用于建置模型的重要变量的用户指示)及观察用于校正有缺陷芯片或过程的实时工艺行为。特定导线可具有对应于引线框架标识符、垂直位置标识符、水平标识符及/或导线编号的关键词或标识符。在此实例中,跨导线编号存在显著变动,但针对给定列号,工艺测量值展现类似模式。

[0252] 图23A说明比较用于在接合操作期间形成的接合件的破坏性测试的预测模型的训练数据及测试数据的实例质量保证(QA)数据。

[0253] 在一或多个实施例中,工艺数据(随后,输入数据或训练数据)包含经导出数据,例如指示与在第二接合操作中接合于特定芯片中的多条导线有关的测量值的中值或平均值的经产生值,及包括与所述多条导线中的每一者的值的偏差的一组经产生偏差。

[0254] 图23B说明用于产生经导出处理数据的实例方法。在此实例中,对于用于机器学习模型的经考虑因素中的每一者(例如,导线的最大经测量阻抗),数据表2350展示针对每一选定导线编号2351及对应水平位置标识符编号2352计算一或多个经导出值。可考虑或选择性地考虑接合操作的所有导线。例如,测量值数据可限于具有待考虑目标变量的最大数字(例如最正常操作或最差质量)的导线。在此实例数据表2350中,N是特定芯片上的导线的总数。

[0255] 在此实例中,将第一经导出值计算为待考虑测量值的中值正常值2353。通过针对通过质量测试的导线取给定唯一导线位置的测量值(例如,针对图22A及23B中的水平位置标识符2352及导线编号2351的测量值)的中值来计算中值正常值。通过针对特定导线及位置标识符从实际测量值减去中值来计算第二经导出值(例如,差量值(正常)2354)。可基于包含针对质量测试展现异常行为的导线的中值计算其它经导出值,例如异常差量值2355。

[0256] 额外地或替代地,输入数据包含针对考虑偏差组的产品(或在考虑芯片的经接合导线的偏差的芯片级上)的经导出数据,例如经产生度量。图24是用于导出与芯片级上的接合操作系统相关联的工艺数据的实例图表2400。在此实例中,可计算一或多个度量。例如,当一次取一个变量且使用其接合结果指示通过质量的导线时,使用正总和、负总和及零交叉数目来区分正常与异常状况。通过平均化中值上方的区域来计算正总和值,即, $Pos_sum = \sum_{i=1}^N Del_i$;其中 $Del_i > 0$ 。通过平均化中值下方的区域来计算负总和值,即, $Neg_sum = \sum_{i=1}^N Del_i$;其中 $Del_i < 0$ 。N是经考虑芯片上的导线的总数。零交叉(例如,零交叉2402)的总数是4。可在这些特征上训练模型以识别异常状况。可考虑更少、较少或不同芯片度量(例如,最大、最小或平均值芯片制造工艺测量值)。在此实例中,在芯片级进行分析以识别异常状况的本质。经导出工艺数据整体上与集成电路芯片的特定芯片有关,且从与特定芯片相关联且在第二接合操作中接合的导线的测量数据导出。测量及异常检测可以某一其它粒度(例如在导线或裸片级上)进行。

[0257] 在一或多个实施例中,从经导出数据产生工艺数据,所述经导出数据包括与多种不同类型的测量值相关且与相同导线有关的经产生奇异数据值。例如,接合操作中的测量值可包括导线的电压(V)及电流(I)测量值,但可导出电阻(R)值(例如,根据等式 $V=IR$)。这样,此一个值可考虑不同测量类型之间的关系。

[0258] 因此,在一或多个实施例中,包括工艺数据的训练数据可从包括经产生奇异数据值的经导出数据产生,所述经产生奇异数据值可与多种不同类型的测量值相关且与第一接合操作中的相同导线有关。指示从第二或后续接合操作的测量值产生的工艺数据的输入数据可包括与多种不同类型的测量值相关且与第二接合操作中的相同导线有关的经产生奇异数据值。

[0259] 额外地或替代地,可通过从多种不同测量类型导出考虑第一接合操作中的测量类型之间的关系的的信息来产生包括工艺数据的训练数据。指示从第二接合操作的测量值产生的工艺数据的输入数据可包括从多种测量类型导出考虑第二接合操作中的测量类型之间的关系的的信息。

[0260] 在一或多个实施例中,工艺数据可用于预测关于接合操作的质量预测的信息。例如,第一及第二操作中的测量值可包括与形成针脚式接合件的过程及/或形成球形接合件的工艺相关联的测量值。

[0261] 图25是说明根据本实施例建模的经预测球剪切值与由于破坏性测试而获得的实际球剪切值之间的对应性的图表2500。

[0262] 规格限制用于产品的最终接受或拒绝。在此实例中,分隔线2510表示球剪切测试测量值(例如,以克为单位)的规格限制的下限且分隔线2512表示上限。可使用比规格限制更严格的控制限制。控制限制用于工艺控制目的以确保产品在规格限制内产生。在此实例中,分隔线2520表示球剪切测试测量值的控制限制的下限且分隔线2522表示上限。分隔线具有在经预测球剪切轴2530上的对应值(例如,在图表2500的两条轴上,规格下限可在A处,控制下限可在B处,控制上限可在C处,且规格上限可在D处)。标绘球剪切破坏性测试中的每一者的测试结果对比由包含不同正常状况及异常状况(材料污染(MC)、旧盖(OC)及浮动状况(FC))的关键词识别的不同对象目标的球剪切的经标绘预测值。

[0263] 评估模型的准确度。TP代表真阳性且表示由模型识别为差的较差质量接合件。TN代表真阴性且表示由模型识别为良好的良好质量接合件。如表2550中所展示,图表展示正确的大多数预测(291个TP及11,922个TN)。FN代表伪阴性且表示由模型识别为良好的较差质量接合件。FP代表伪阳性且表示由模型识别为差的良好质量接合件。仅存在少量错误预测(27个FP及40个FN)。

[0264] 一些度量对于不同工业更为重要。例如,伪阴性在集成电路构造中可为重要的,因为这些类型的接合件可能导致现场退货及保修索赔的增加。因此,也考虑良性伪阴性(B_{FN})。这些表示具有介于规格下(上)限与控制下(上)限之间的接合强度的接合件。这些伪阴性关于规格限制不具有不良质量。大多数伪阴性是良性伪阴性(40个中的32个)。可使用像精度、召回率及f1的统计工具来评估模型。例如,精度可被计算为真阳性的数目除以真阳性的数目及伪阳性的数目的总和:

$$[0265] \quad \text{精度} = \frac{\text{真阳性的\#}}{\text{真阳性的\#} + \text{伪阳性的\#}}$$

[0266] 召回率可被计算为真阳性的数目除以真阳性的数目及伪阴性的数目的总和：

$$[0267] \quad \text{精度} = \frac{\text{真阳性的\#}}{\text{真阳性的\#} + \text{伪阴性的\#}}$$

[0268] F1是精度及召回率的调和平均值。例如,其可被计算为：

$$[0269] \quad f1 = 2 * \frac{(\text{精度}) * (\text{召回率})}{\text{精度} + \text{召回率}}$$

[0270] 在此实例中,针对具有0.91509的精确率、0.8795的召回率及0.89676的F1的预测,模型表现良好。在此情况下,通过使用预测,用户可潜在地避免进行破坏性质量测试且密切了解集成电路结果。如果从计算移除32个良性伪阴性,那么可产生较高模型准确度。

[0271] 图26是说明根据本实施例建模的经预测针脚拉拔值与由于破坏性测试而获得的实际针脚拉拔值之间的对应性的图表。在此情况下,仅存在一个目标值。分隔线2610表示针脚拉拔测试测量值(例如,以克为单位)的规格限制的下限。未在图表上展示上限,因为其远高于所标绘的值。分隔线2620表示针脚拉拔测试测量值的控制限制的下限且分隔线2622表示上限。分隔线具有在经预测针脚拉拔轴2530上的对应值(例如,在图表2600的两条轴上,规格下限在E处,控制下限在F处,控制上限在G处,且规格上限超出图表)。对于如表2650中所展示的除了7条离群导线(其归因于在针脚拉拔测试中具有1的精度、召回率及F1为的问题而被排除)外的所有准确预测,此模型也表现良好。除此模型是无需破坏性测试的准确模型外,本文中所描述的模型的预测能力也不像手动测试中所发生那样易受人为错误影响。

[0272] 在图25及26的实例中,使用由多个决策树组成的梯度提升模型。预测模型定义输入变量与目标变量之间的关系。预测模型的目的是从输入预测目标值。模型是通过使用其中目标值已知的训练数据来建立。接着,可将模型应用于其中目标未知的观察。如果预测良好地拟合新数据,那么称模型良好地一般化。良好一般化是预测任务的主要目标。预测模型可能良好地拟合训练数据但不良地一般化。

[0273] 决策树是已在统计学及人工智能社群中独立地开发的类型的预测模型。梯度提升过程可用于通过拟合一组可加性树(例如,使用由北卡罗莱纳州喀里市的SAS Institute公司提供的GRADBOOST相关软件)来建立预测模型。

[0274] 预测模型的质量可取决于控管训练过程的各种选项的值;这些选项被称为超参数。这些超参数的默认值可能不适合于所有应用,且可选择这些超参数的值的唯一组合来最小化目标功能或某一准确度量度。如此,可使用本文中描述的用以建置模型的其它或额外方法。例如,可关于相同类型不同地产生不同模型且使用验证过程(例如,使用k折交叉验证)选择经训练模型中的一者。另外,可使用超出质量保证目标的其它目标值。

[0275] 例如,导线接合机中可存在多种异常状况。这些异常状况可移动经测量特征或质量保证值的分布(例如与球剪切及针脚拉拔有关)或其也可能导致有缺陷导线。早期识别异常状况是必要的,使得可采取适当校正措施且减少较差质量产品的数量。正常目标变量可与正常操作相关,例如稳定工艺、在控制状态中、符合及/或超出规格比率,以及不存在任何系统性干扰原因的指示符。可基于工艺数据中的经检测异常来预测异常状况,例如裸片倾斜(DT)、材料污染(MC)、旧盖(OC)及浮动状况(FC)。

[0276] 例如,在裸片倾斜的情况下,裸片附接工艺可用于电子封装中的系统的裸片、装置及其余者之间的连接。在此过程期间,裸片在由液相线焊料诱发的毛细管力作用下移动。此

裸片倾斜现象经常发生且严重恶化装置的可靠性及性能。预测量可用于了解来自液相线焊料的接合件的质量。在浮动状况下, 归因于裸片浮动, 可能存在不粘接。

[0277] 一种类型的材料污染可为引线框架污染, 其中在引线框架上存在表面有机物、有机化合物及/或残余物。此也可与接合中的质量问题相关。

[0278] 早期动作可导致不具有这些异常的工艺良率的增加。例如, 使用关于图25到26描述的数据, 模型准确度是显著的。下表展示也被称为平均绝对百分比偏差 (MAPD) 的平均绝对百分比错误 (MAPE), 其是对来自5个批次的集成电路芯片生产的观察的预测准确度的量度。

观察	目标变量	MAPE
1	正常	3.64%
2	正常	3.74%
3	Abnormal_FC	3.62%
5	正常	3.72%
6	Abnormal_MC	12.22%

[0279] 表1

[0281] 在表1的实例中, 不同目标变量的总体映射是4.39%。观察6具有导致较大MAPE的四个离群点。如果排除这些点, 那么观察6的MAPE降低到5.67%且总体MAPE下降到3.84%。

[0282] 用以检测异常状况的模型在芯片级特征上使用梯度提升方法。这些特征是使用传感器变量(或运动特征)的芯片级波形来提取。波形具有x轴上的导线编号及y轴上的运动特征值。在此情况下, 梯度提升模型具有标称目标。在其它情况下, 可使用间隔目标值。在此实例中, 异常目标变量与浮动状况(FC)及材料污染(MC)有关。也可预测其它类型的异常状况, 例如裸片倾斜。

[0283] 图27A是说明经配置用于流式传输及边缘分析的事件流处理(ESP)系统2700的堆叠的功能框图。ESP组件2702可包括如本文中所描述(例如, 在图8中)的一或多个ESP引擎。例如, ESP组件可用于提供决定制造执行系统(MES)的数据质量、预处理、AI评分、预测, 以及异常预测量、实时规定警报、流内训练、联合AI、部署客户模型及操作系统中的一或多个者。因此, ESP组件2702可与客户端服务2704装置及系统以及数据管理系统2706交换信息。模型管理器2708可用于基于(例如, 由客户端服务2704及/或ESP组件2702)接收的数据管理、开发及更新本文中所描述的模型。人工智能及机器学习组件2710可用于训练及更新本文中所描述的模型(例如, 如由模型管理器2708指示)。梯度提升是作为用于建构模型的实例方法给出。但是, 可采用额外或不同模型方法(例如, 随机森林、线性回归及神经网络)。模型管理器2708以及人工智能及机器学习组件2710可与客户端服务2704及数据管理2706交换信息。因此, 本文中所描述的实施例(例如, 图27A中的系统)可用于开发分析模型以预测导线接合件质量, 包含从随机测试(其通常可涉及破坏性测试)移动到模型建议的基于风险的标定测试。经开发模型可提供制造过程中存在异常状况(或故障)的早期指示且具有较低拒绝率且增加产出率。边缘处理可提供现场指导。

[0284] 图27B是说明根据本公开的一个实施例的用于产生机器学习模型的实例方法2750的流程图。如图27B中所见, 方法2750包括用于开发提取、变换及加载(ETL)处理过程以读取输入数据及建立用于产生机器学习模型的数据集的操作(方框2752)。在一个实施例中, 例

如,ETL处理过程包括在由计算机的处理电路系统执行时从输入数据提取及组合数据集的一或多个编码脚本文件。由ETL处理过程读取的输入数据可呈所需要或期望的任何形式,但在一个实施例中,以.csv文件进行格式化且包括在导线接合阶段期间测量的数据值(图13中的方框1308)。

[0285] 接着,方法2750对输入数据执行一或多个质量检查(方框2754)以确保输入数据及数据集具有足以产生机器学习模型的质量。例如,如上文所陈述,执行球形及针脚式接合过程(图13中的方框1308)的导线接合机经配置以在执行所述工艺时测量不同操作参数。参数是针对后段半导体制造过程的每一“运行”测量,其中每一运行包括表示“批次”的多个参数。

[0286] 根据本公开,按批次获得输入数据,其中每一批次经分析以确定其数据的质量。例如,本实施例的至少一个方面识别每一批次中的若干不同变量、变量名称及变量分布。接着,比较所述数据与相关联于一或多个先前获得的批次(或具有已知良好数据的基线批次)的数据,以确定参数的值是否存在任何巨大变化(例如,所述值未在预定义容差范围内或明显不同于基线值)。例如,数据的此类巨大变化可指示输入数据不适合用于产生模型。

[0287] 但是,甚至在数据可能不可用的情况下,根据本公开确定参数值之间存在巨大差异仍为有益的。例如,裸片倾斜及机器读取错误是可能导致不可用参数值的测量及收集的事物类型的一些常见实例。但是,知晓存在这些错误可使操作者能够更快地识别及校正导致记录输入数据的问题。

[0288] 作为另一实例,质量检查可指示所测量的值的范围未充分反映可在实际操作期间产生的所有值。拥有此信息,操作者可更改用于导线接合阶段的监测/测量过程以确保充分涵盖适当操作范围。确保监测及测量过程充分涵盖在生产期间实际所经历的操作范围将有助于确保所测量的参数的值是准确的且适合于产生数据集及机器学习模型。

[0289] 根据本公开,可监测及测量与导线接合阶段期间的操作有关的参数中的任何者。但是,在本公开的一个实施例中,测量以下参数中的一或多者以确定与异常状况(例如浮动状况及引线框架污染)中的一或多者有关的异常。

[0290] 如果输入数据被视为具有足够质量,那么方法2750变换及/或建立一或多个变量(方框2756)。例如,构成ETL处理过程的编码脚本文件中的一或多者可经配置以变换输入数据中的变量(例如,在单位之间、按比例放大或缩小值等)。额外地或替代地,构成ETL处理过程的编码脚本文件中的一或多者可从输入数据计算一或多个经导出变量。在一个实例中,经变换变量及/或经计算的经导出变量由本实施例使用以产生一或多个建模数据集。训练数据集及/或测试数据集也可由本实施例使用经变换及/或经计算的经导出变量来产生。

[0291] 额外地或替代地,本公开的实施例对输入数据执行分析以获得预测量及目标变量的分布,以及用以指示各种统计数据的数据的一或多个不同绘图及表。在至少一个实施例中,对输入数据进行相关性分析以确定在先前记录的“基线”参数值与由测量导线接合操作的一或多个传感器产生的参数值之间是否存在任何相关性。此类分析提供对输入数据的更好理解且对于产生机器学习模型非常有益。

[0292] 在适当处理输入数据的情况下,方法2750接着产生用于预测导线质量参数的机器学习模型(方框2758)。这些模型包含用于估计破坏性球剪切测试的结果(即,将球形接合件1410从接合垫1508剪切掉所需的力)及针脚拉拔测试的结果(即,将导线1504与目标垫1512

分离所需的力)的一或多个预测模型。可利用多种技术中的任何者来产生预测模型,包含但不限于线性回归技术、随机森林技术、梯度提升技术及神经网络技术。在一个实例实施例中,使用梯度提升技术产生预测模型提供最准确结果。在此实施例中,使用运用K折交叉验证的自动调谐来调谐模型超参数。

[0293] 方法2750执行模型稳定性检查以帮助确保模型提供跨不同数据条件的准确预测(方框2760)。在此步骤中,在已训练模型后检查模型超参数的稳定性。例如,在一个实例实施例中,将经调谐超参数用于多个随机训练/测试分割(即,“运行”的次数)。超参数值可用于训练训练分割且接着对测试分割进行评分。模型的性能是在每一测试分割上确定。为完成此,本公开的一个实施例评估一组预定模型性能统计数据。性能统计数据的一些实例包括但不限于F1比率及均方错误。

[0294] 此外,在本公开的一些实施例中,产生性能评估汇总表。所述汇总表使用跨所有测试分割的性能统计数据提供性能评估。此表可具有所需要或期望的任何结构,但在至少一个实例中,此表具有以下数据元素(例如,表栏)中的一或多个者:性能统计数据;运行次数;特定性能度量的最小值;最大值及/或平均值。

[0295] 一旦模型经完成,其便可用于估计异常状况的存在或不存在。例如,可基于通过边缘装置的观察评分(方框2762)来确定异常状况的存在或不存在。在一些实施例中,模型将提供指示存在的特定异常状况的信息。图27B展示包括一或多个操作的阶段的实例方法。所属领域的技术人员将了解,方法可具有更少或更多阶段或以不同顺序或递归地执行(例如,以保持改进模型)。

[0296] 在一或多个实施例中,可观察数个不同测量值,且计算系统通过选择性地选取在接合操作中观察的参数类型的子集来接收训练数据及/或输入数据。用于导出模型的测量值可为针对所述参数类型的所述子集的测量值。图28是计算机程序产品的功能框图。在一个实例中,计算机程序产品包括例如由计算装置的处理电路系统执行的控制程序。

[0297] 更详细地,图28说明计算装置的处理电路系统2900,以及其执行的单元/模块。各种单元/模块可由硬件及/或由用处理器或处理电路执行的软件代码来实施。在此实例中,单元/模块包含机器学习模型存取单元/模块2910、输入数据接收单元/模块2920、输入数据加权单元/模块2930、异常指示符产生单元/模块2940及异常指示符输出单元/模块2950。

[0298] 在此实例中,机器学习模型存取单元/模块2910配置处理电路系统2900以存取在第一接合操作的训练数据上训练的一或多个机器学习模型。在至少一个实例中,第一接合操作是其中将第一组的多条导线接合到第一组表面(例如,分别将球形及针脚式接合件接合到接合垫及目标垫)的接合操作。另外,通过监督式学习来训练机器学习模型。例如,在一些实例中,监督式训练包括接收训练数据。训练数据包含从第一接合操作的测量值产生的工艺数据及多条导线在接合到第一组表面后的状态。每一状态包括与检测第一接合操作中的一或多个异常有关的目标的一或多个候选结果。另外,产生工艺数据的一或多个权重,使得输入到机器学习模型的工艺数据预测所述目标的一或多个候选结果。

[0299] 输入数据接收单元/模块2920配置处理电路系统2900以接收指示从其中将第二组的多条导线接合到第二组表面(例如,分别将球形及针脚式接合件接合到接合垫及目标垫)的第二接合操作的测量值产生的工艺数据的输入数据。第二组的多条导线不同于第一组的

多条导线,且第二组表面不同于第一组表面。

[0300] 输入数据加权单元/模块2930配置处理电路系统2900以根据机器学习模型对输入数据进行加权。

[0301] 异常指示符产生单元/模块2940配置处理电路系统2900以基于根据机器学习模型对输入数据进行加权来产生指示第二接合操作中的异常发生的风险的异常预测量。

[0302] 异常指示符输出单元/模块2950配置处理电路系统2900以输出异常预测量以控制第二接合操作。

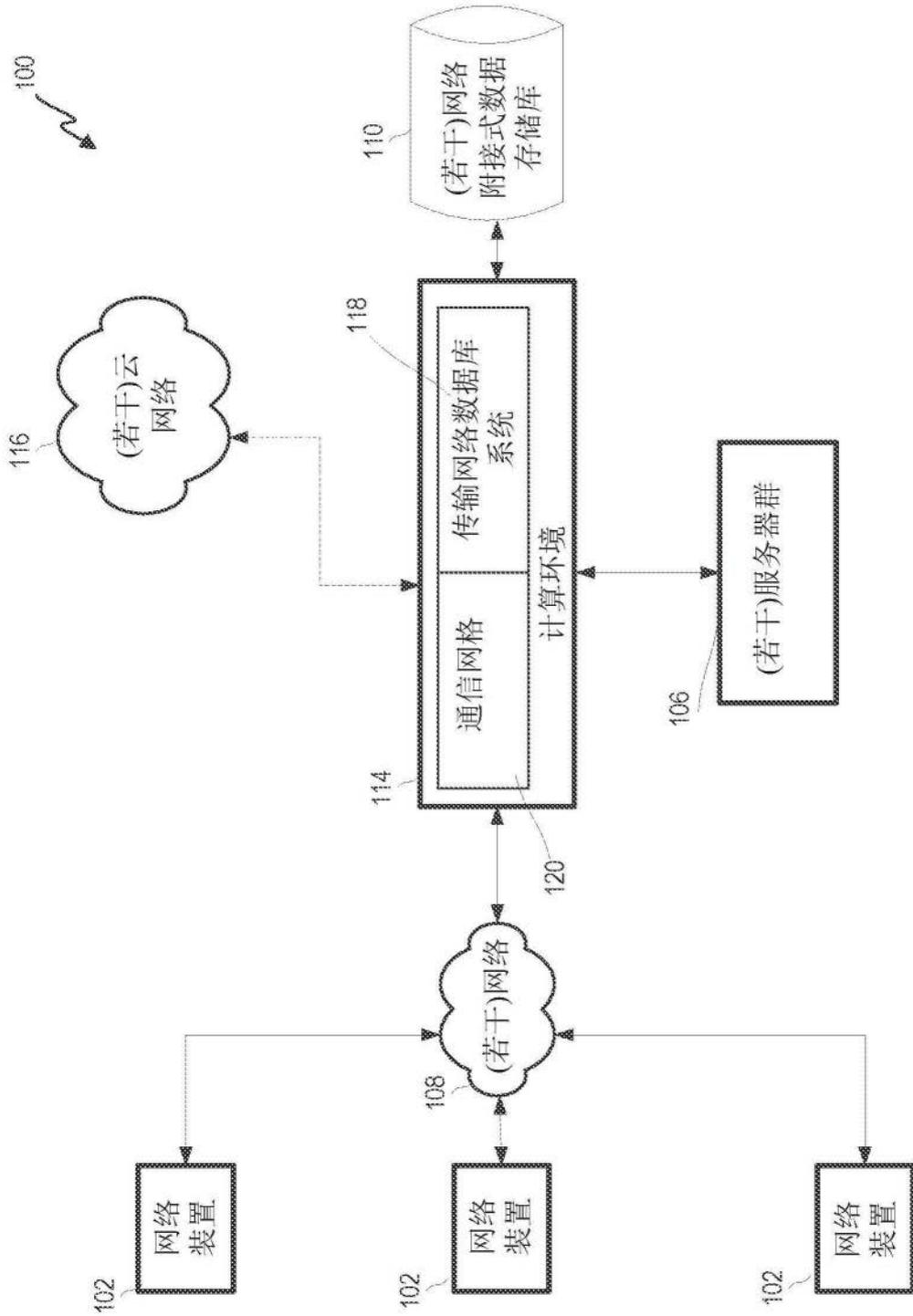


图1

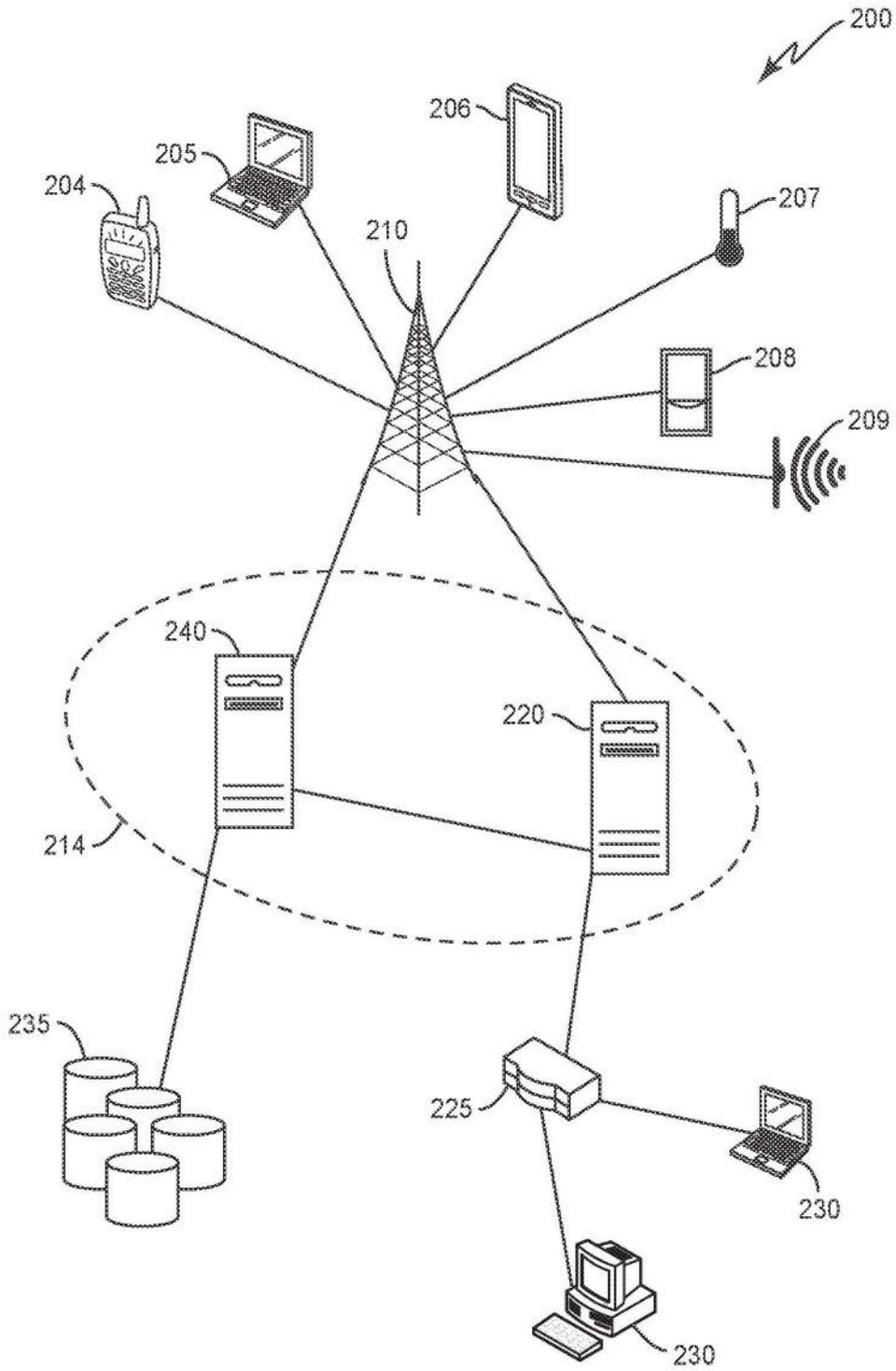


图2

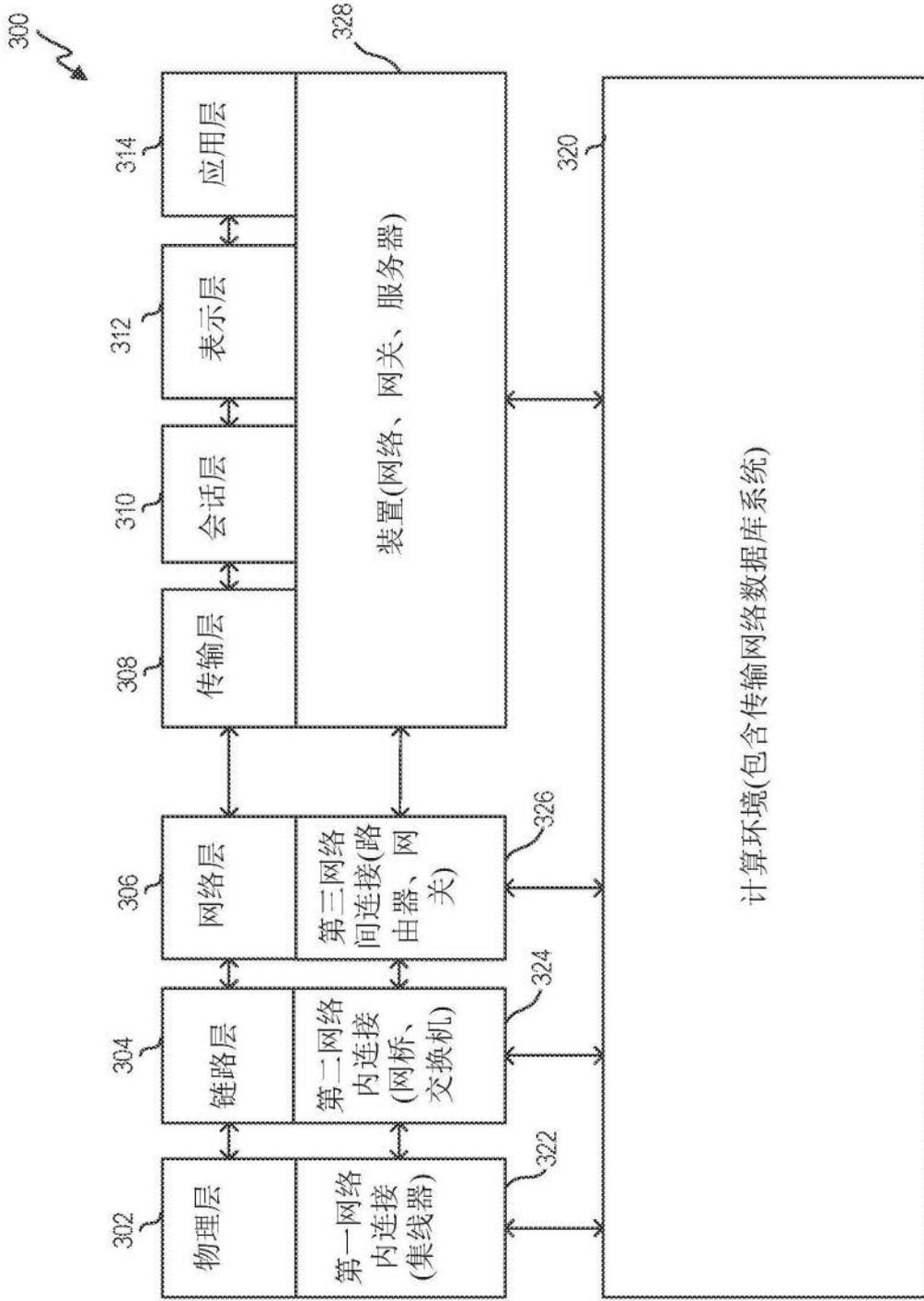


图3

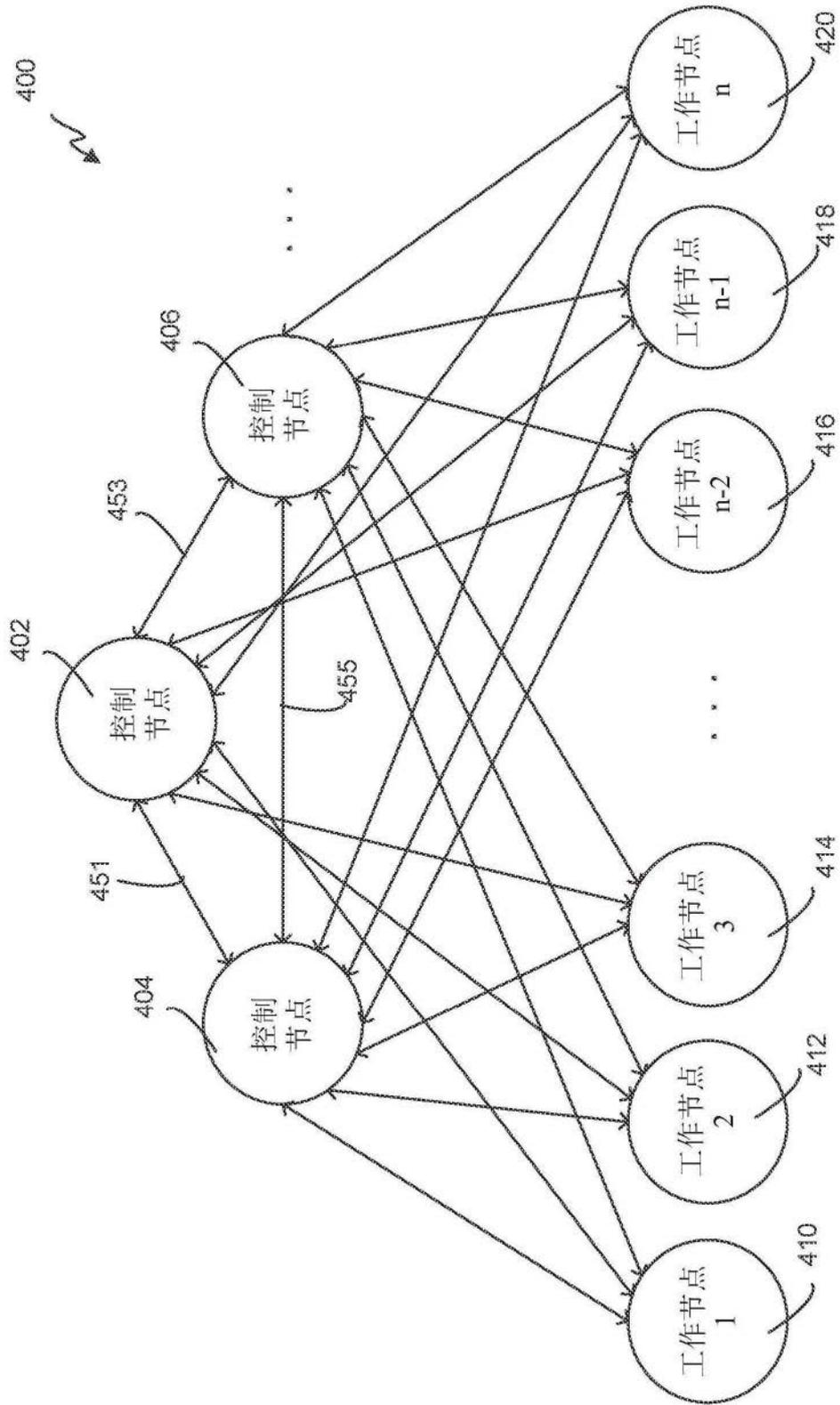


图4

500

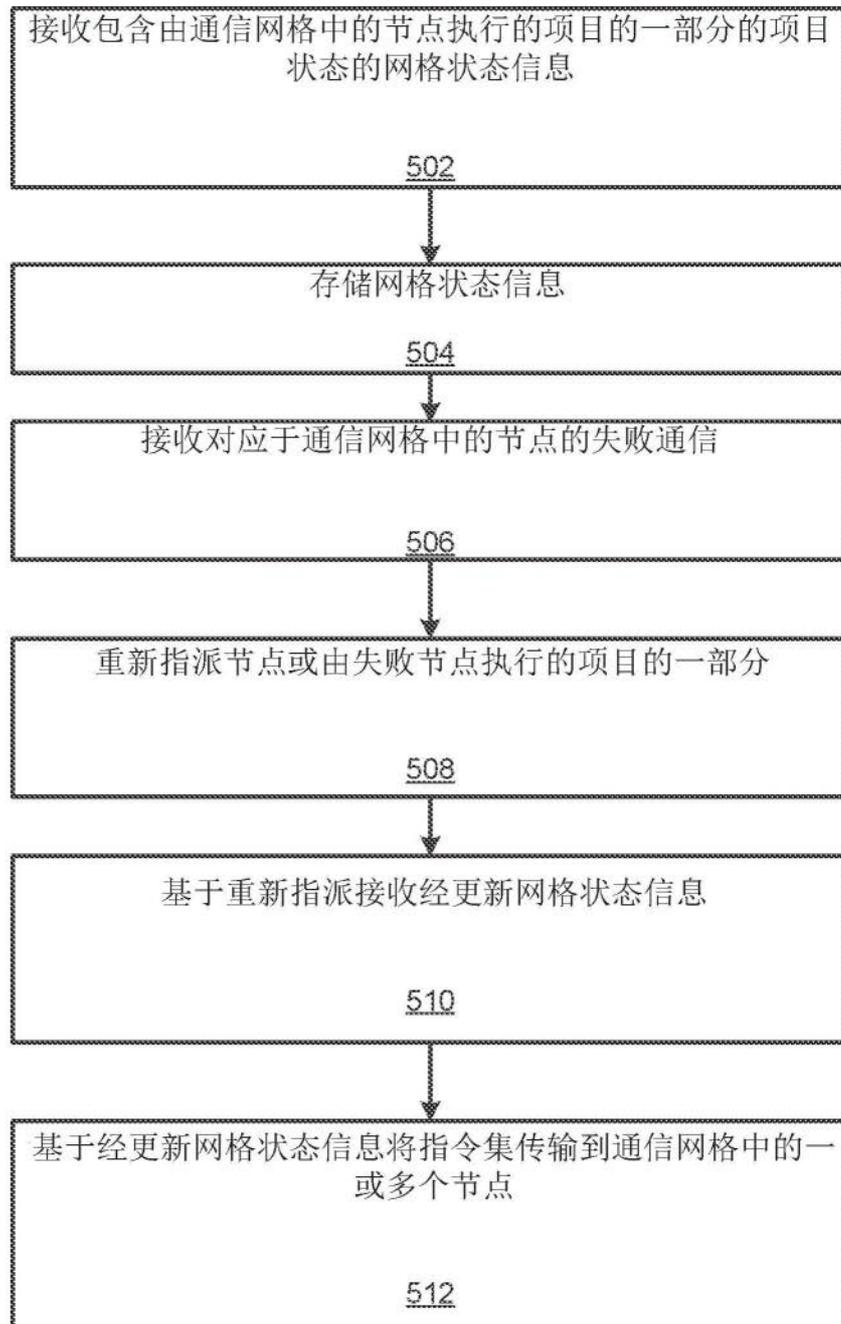


图5

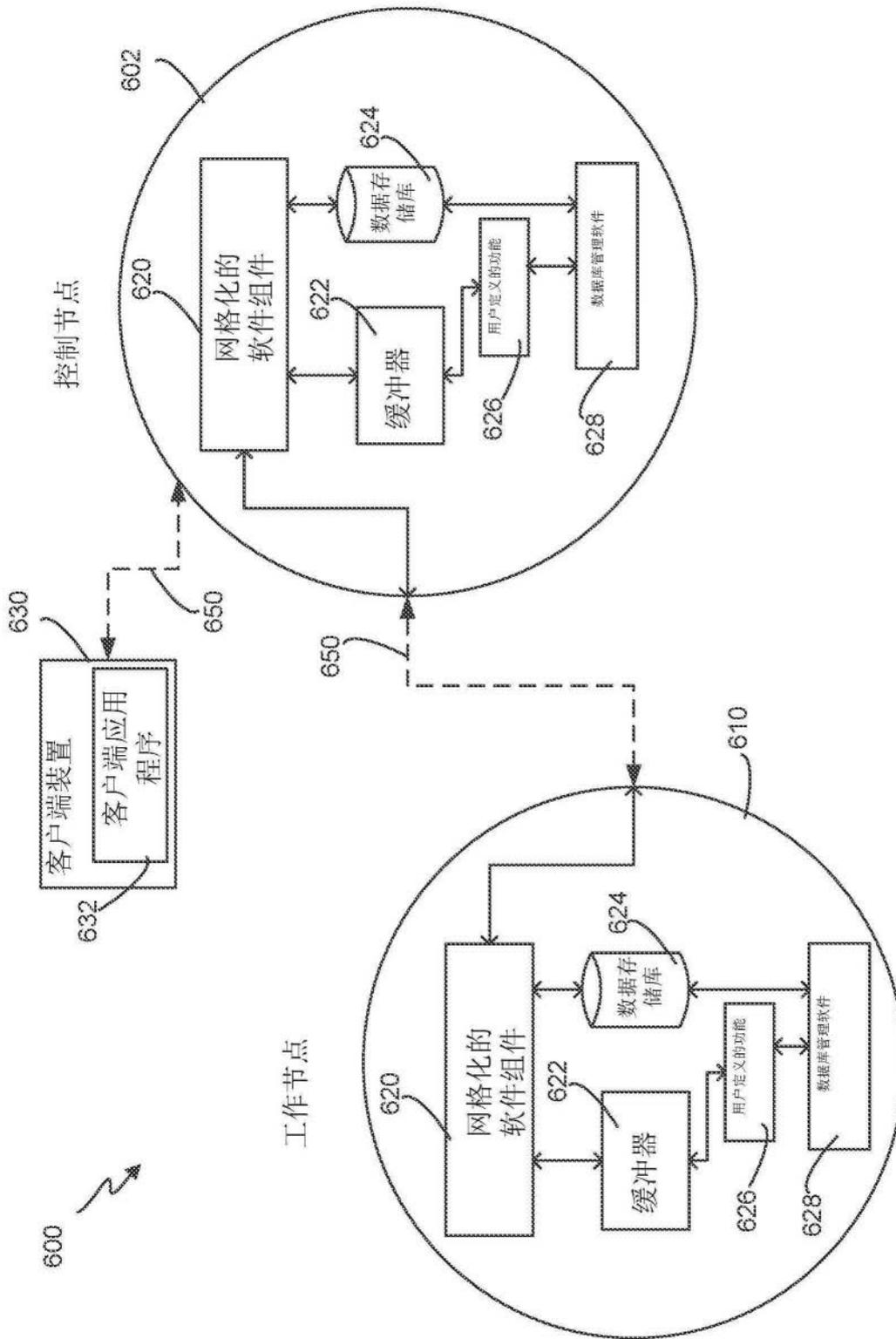


图6

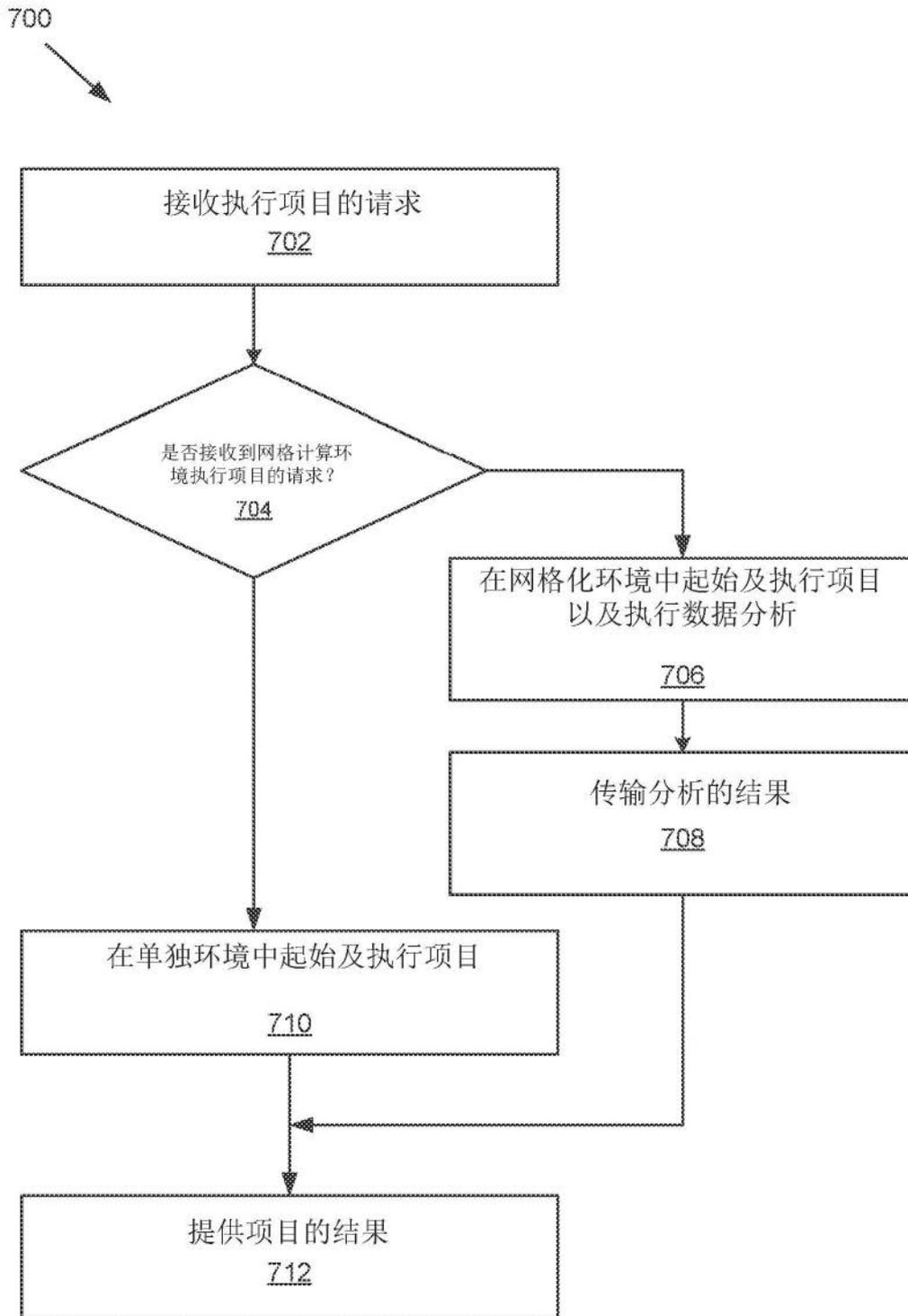


图7

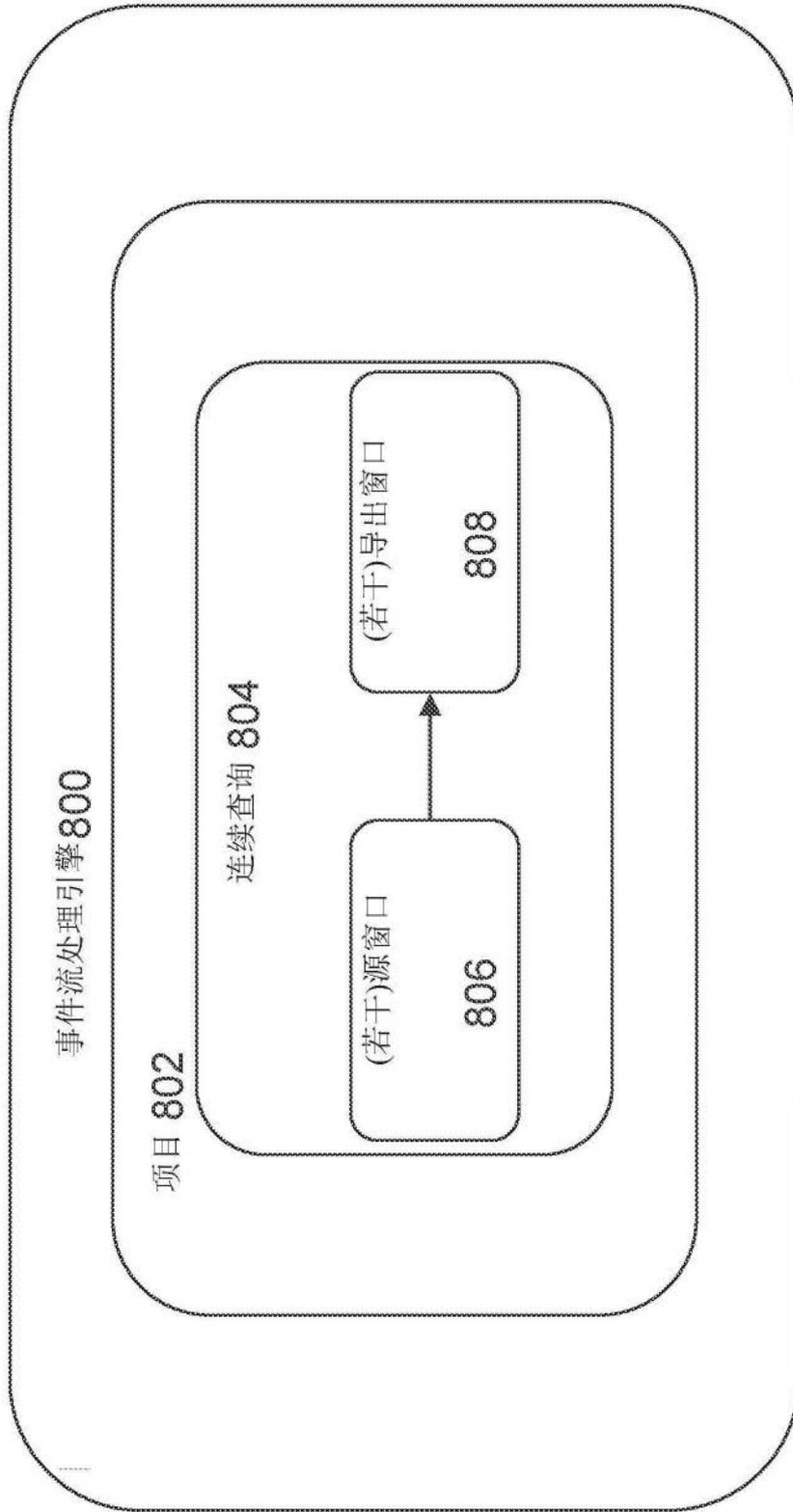


图8

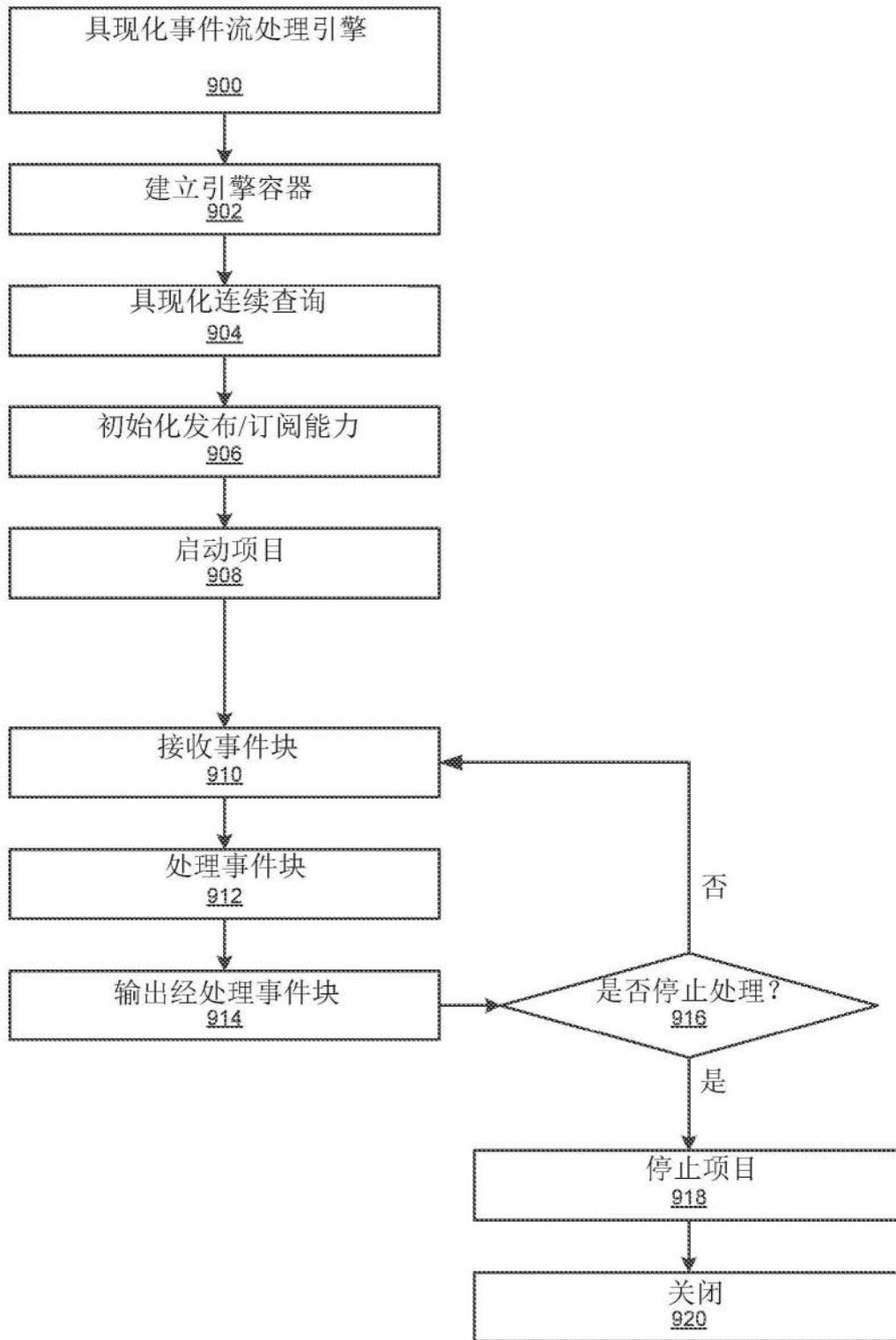


图9

1000 ↗

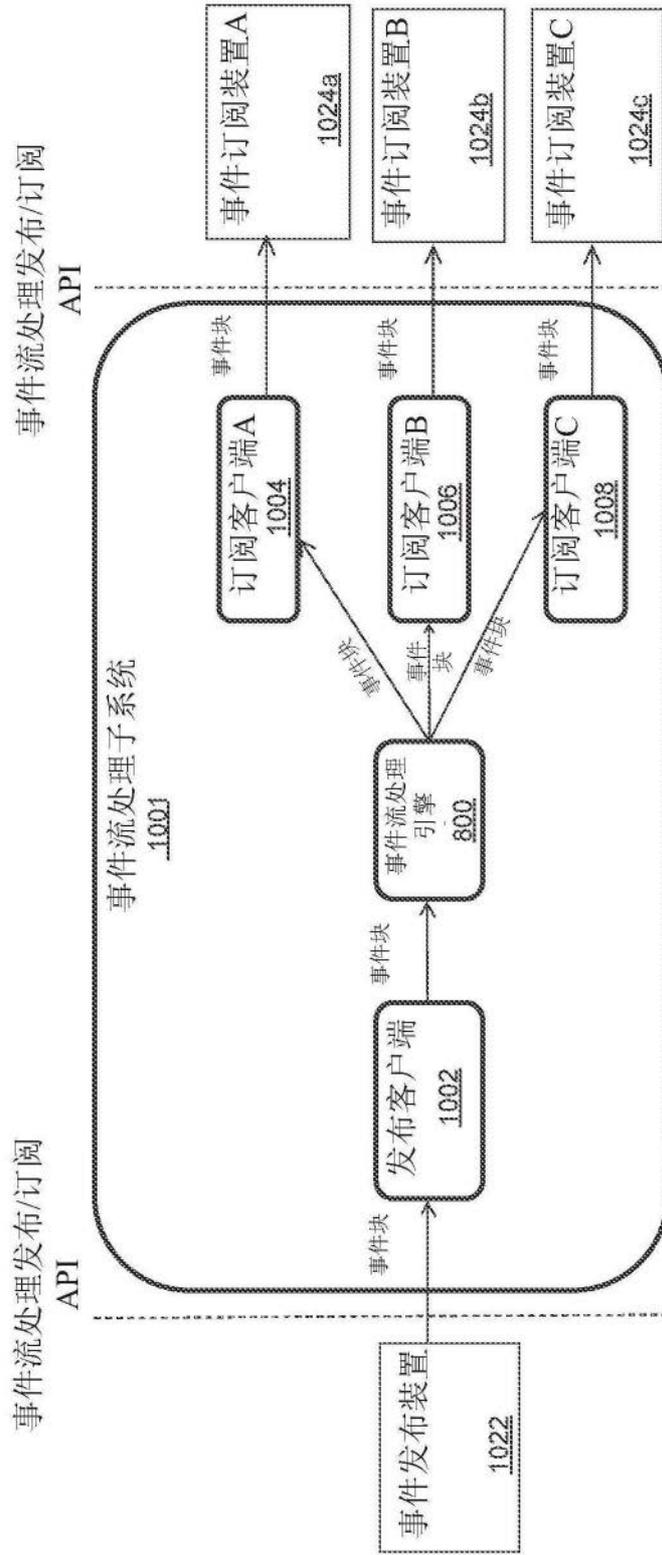


图10

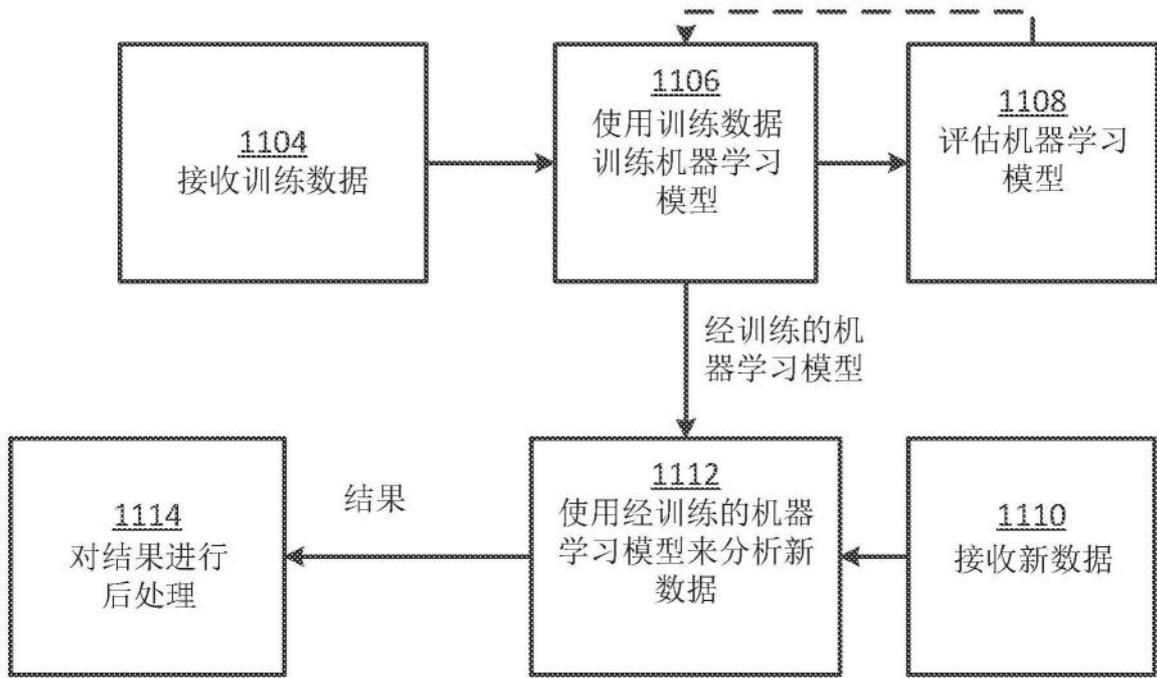


图11

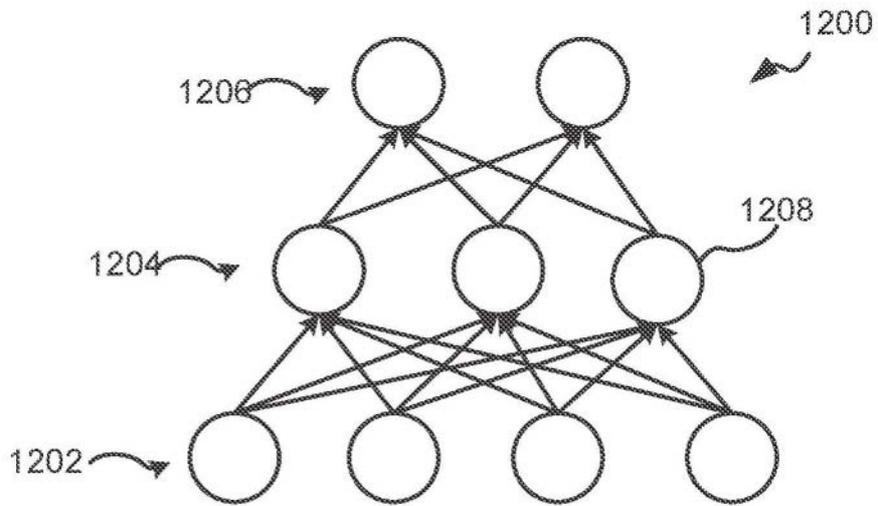


图12

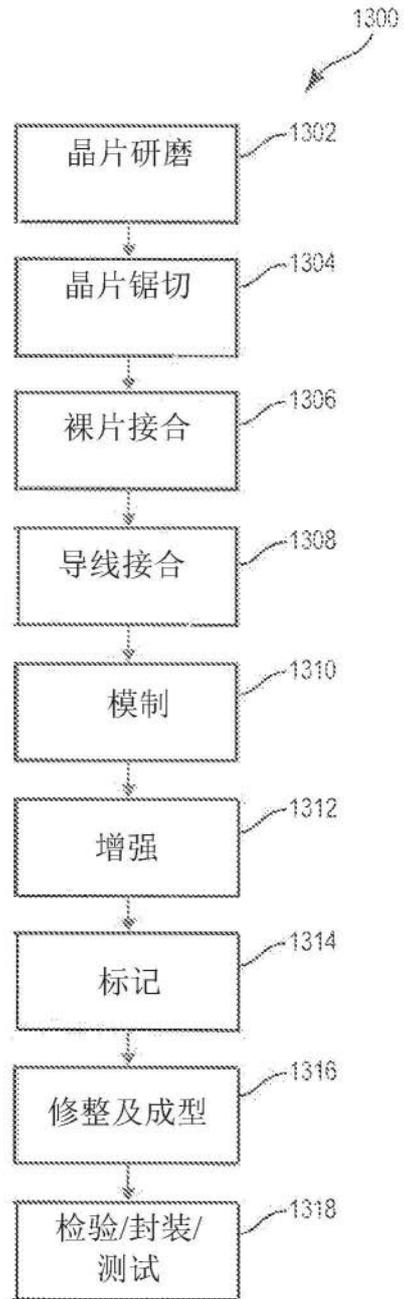


图13

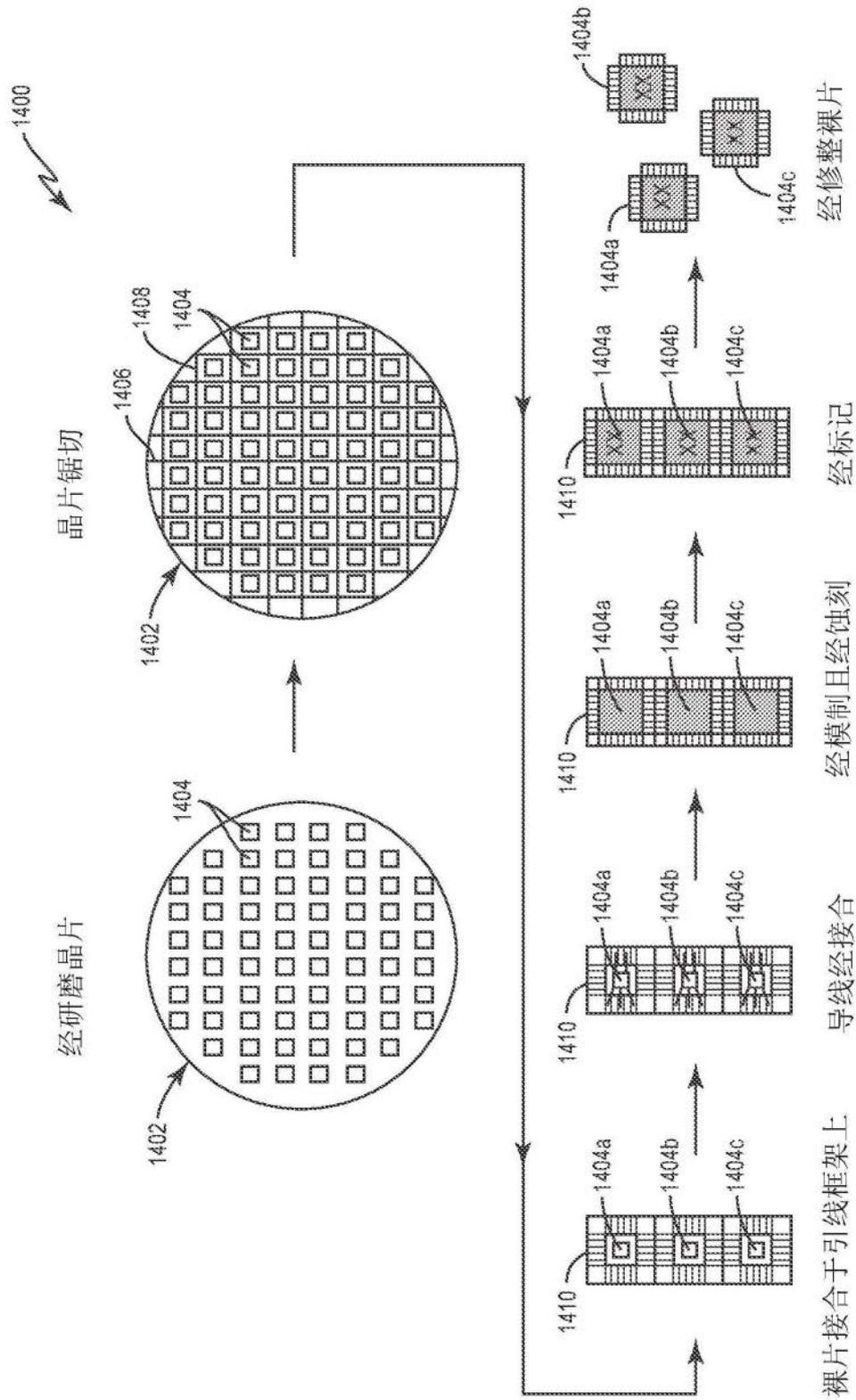


图14

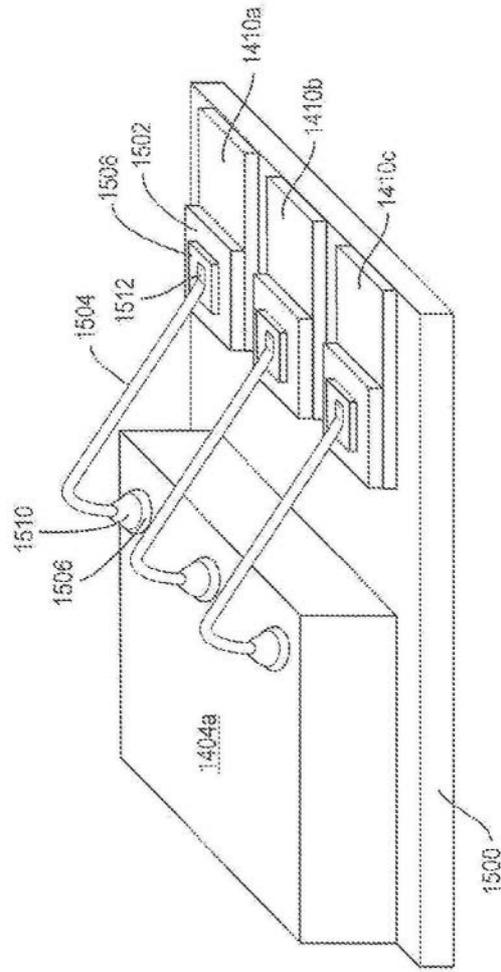


图15A

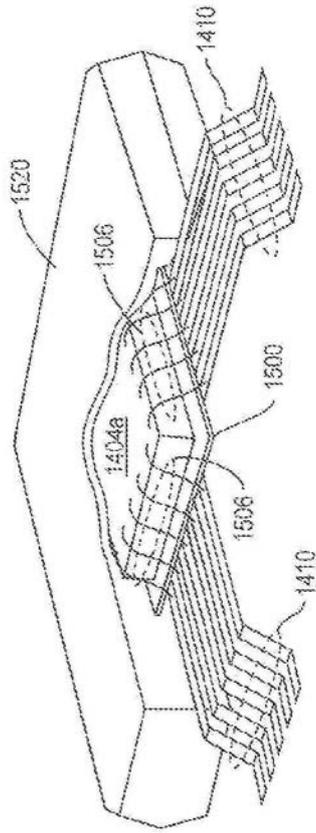


图15B

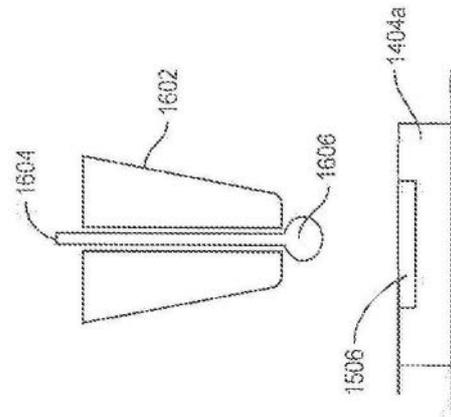


图16A

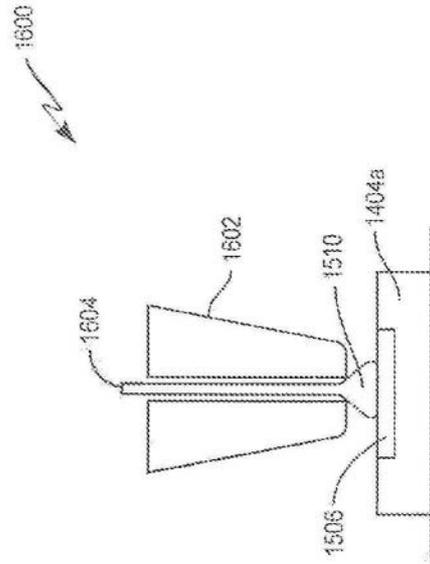


图16B

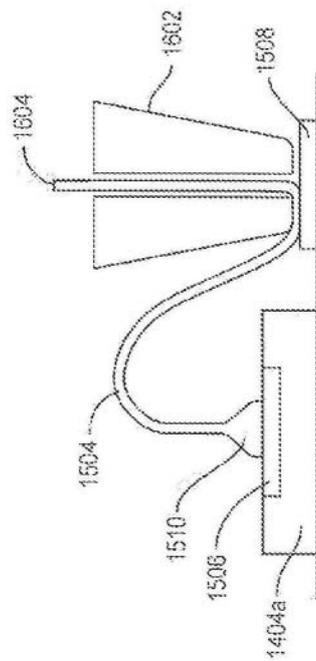


图16C

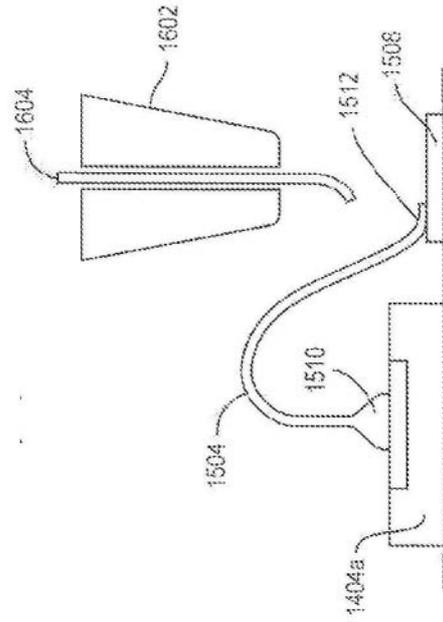


图16D

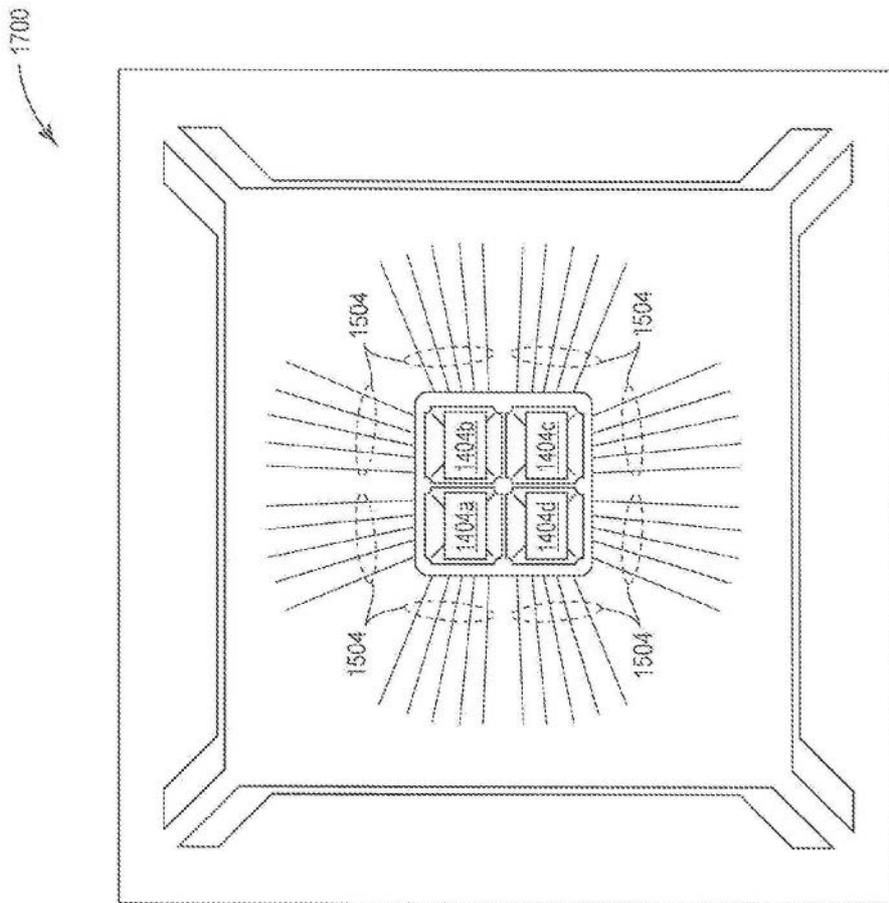


图17

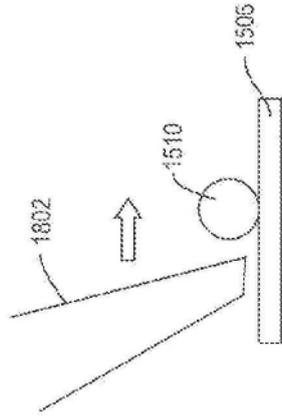


图18A

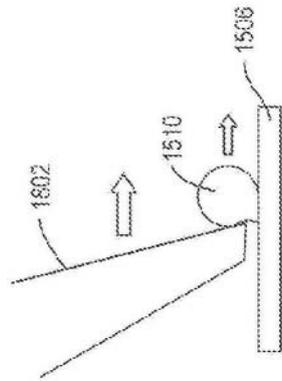


图18B

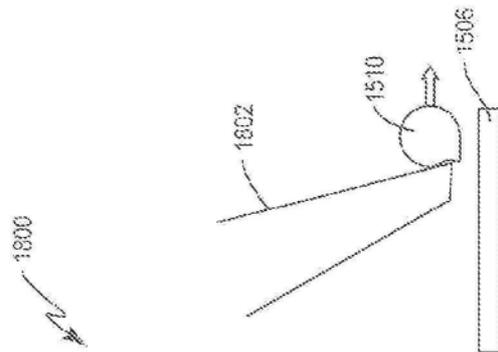


图18C

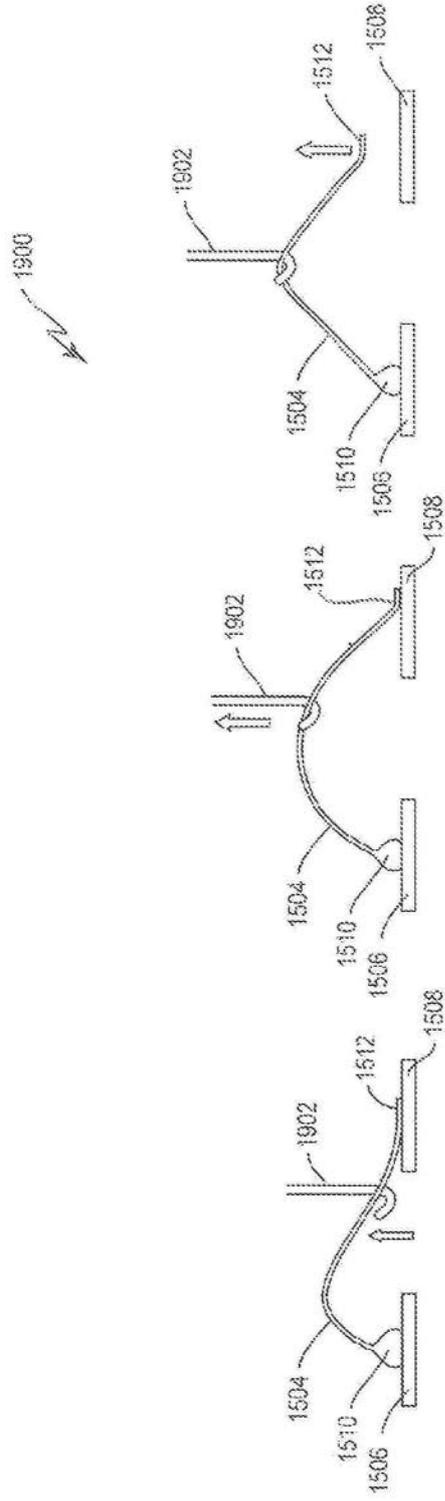


图19A

图19B

图19C

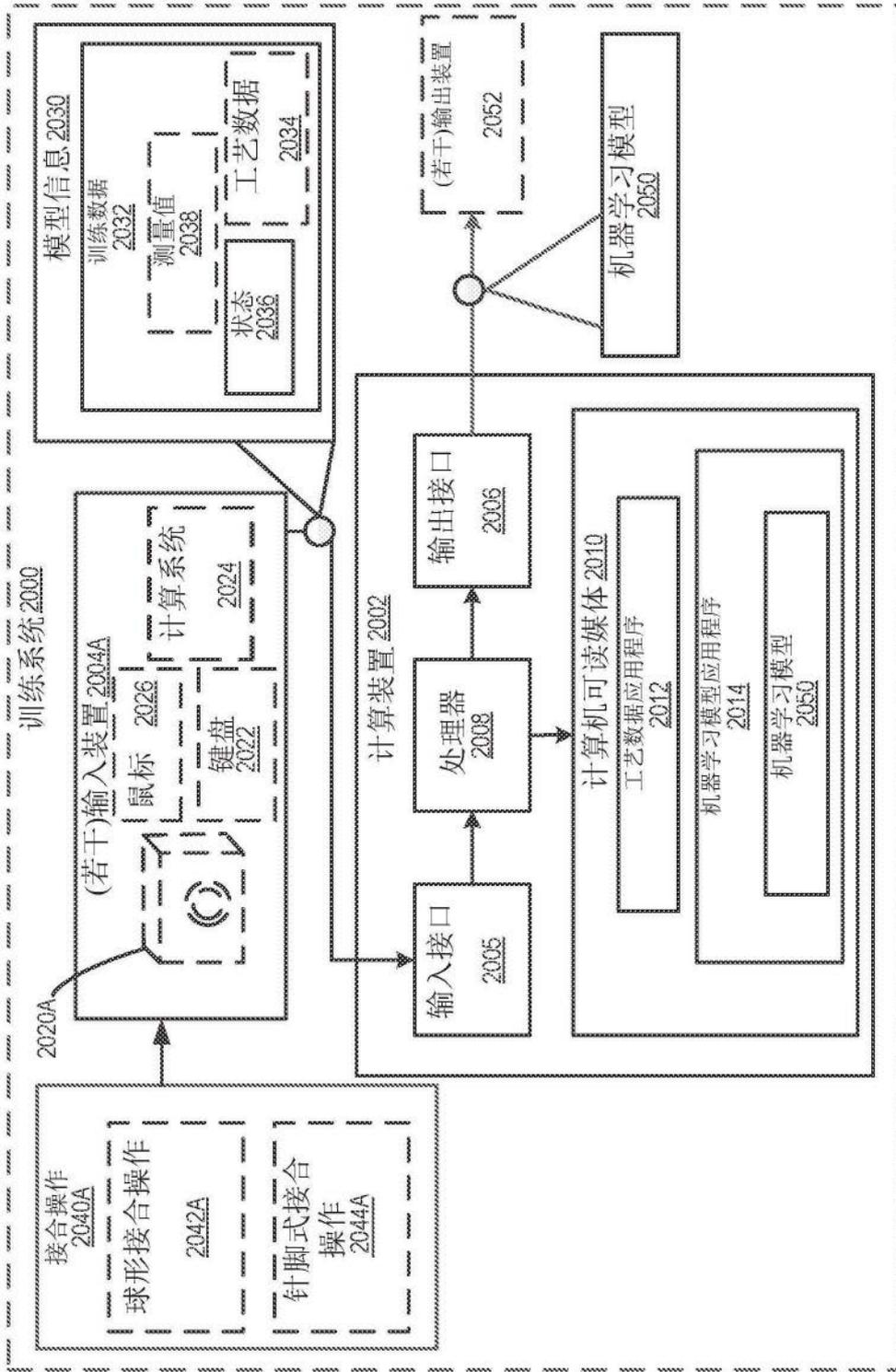


图20A

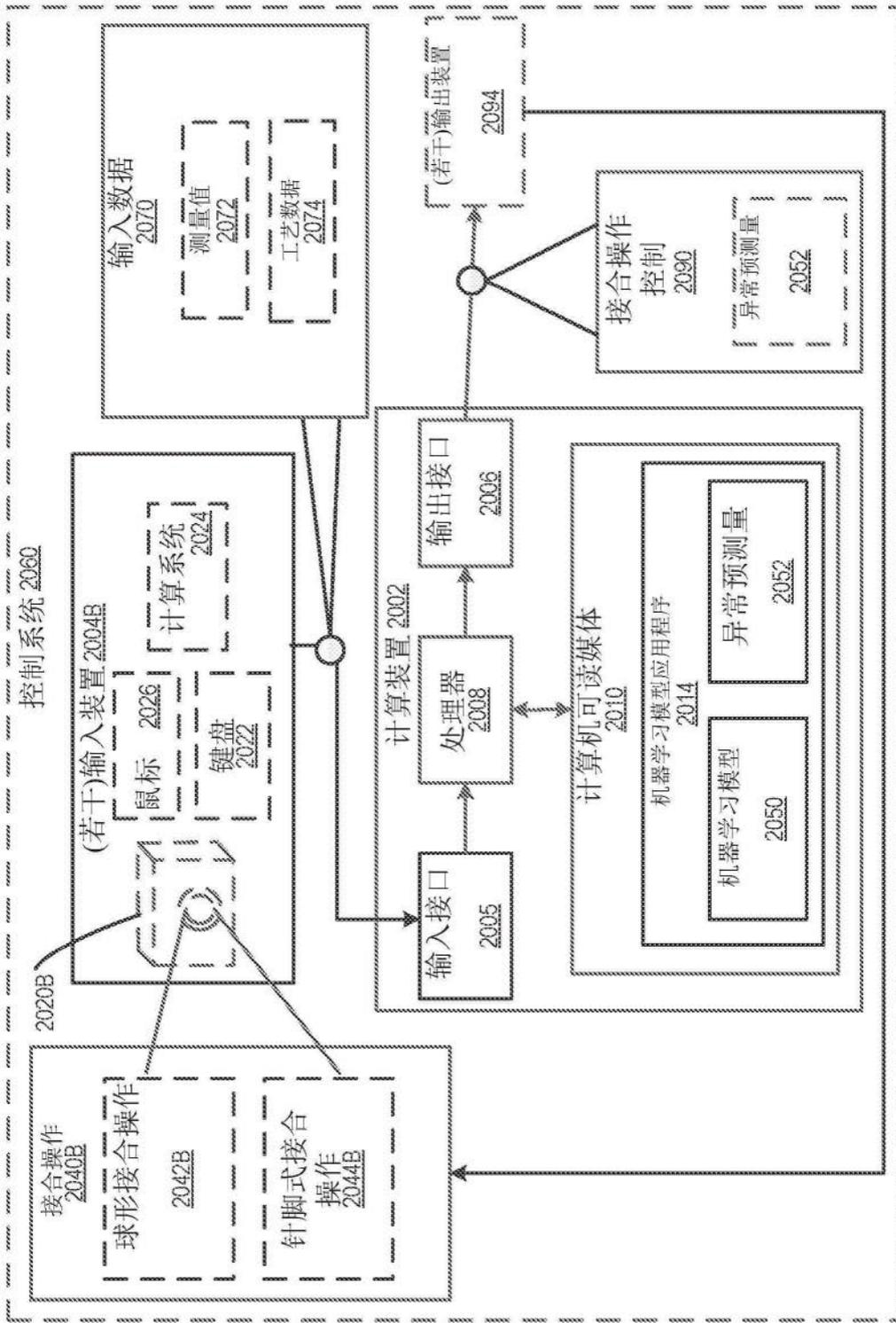


图20B

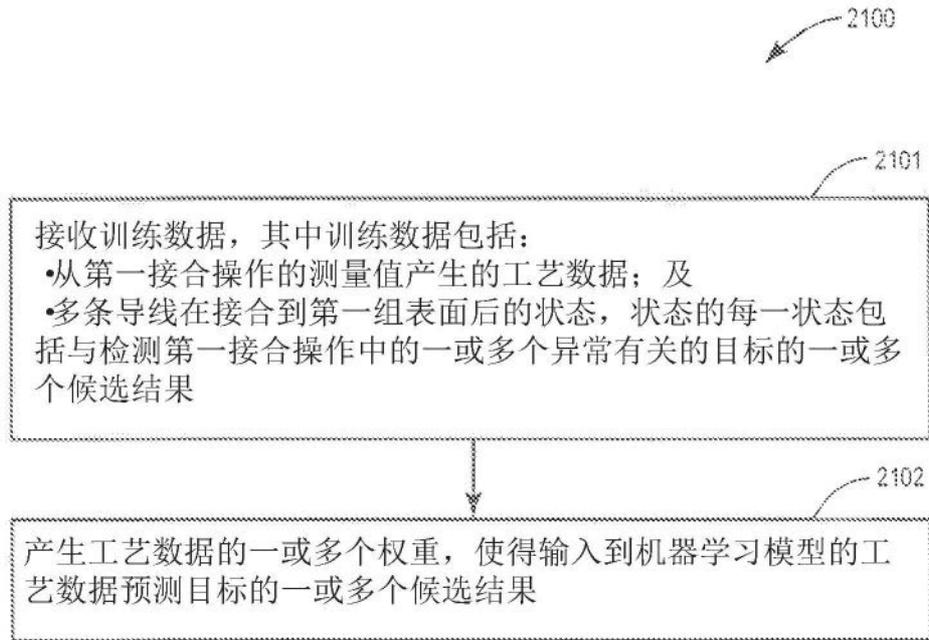


图21A

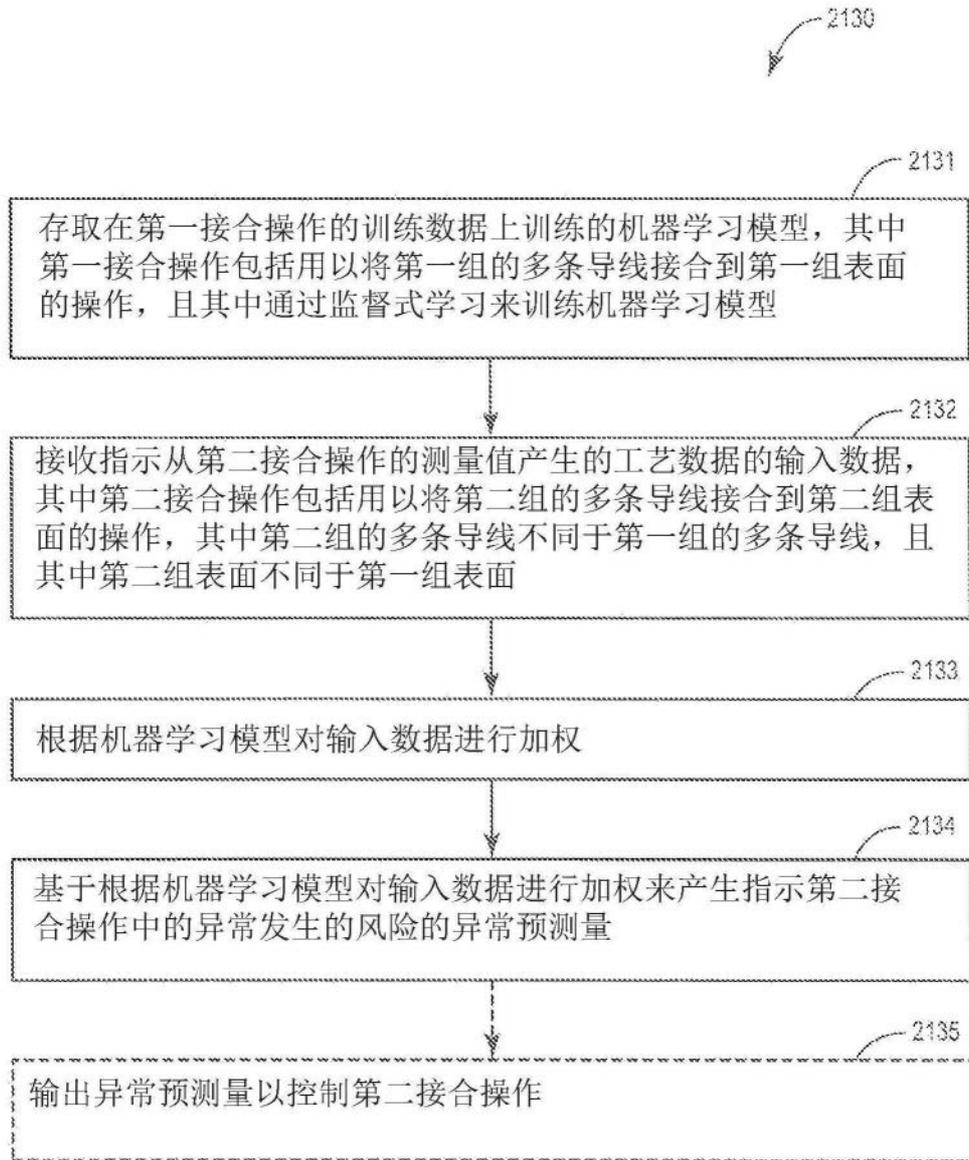


图21B

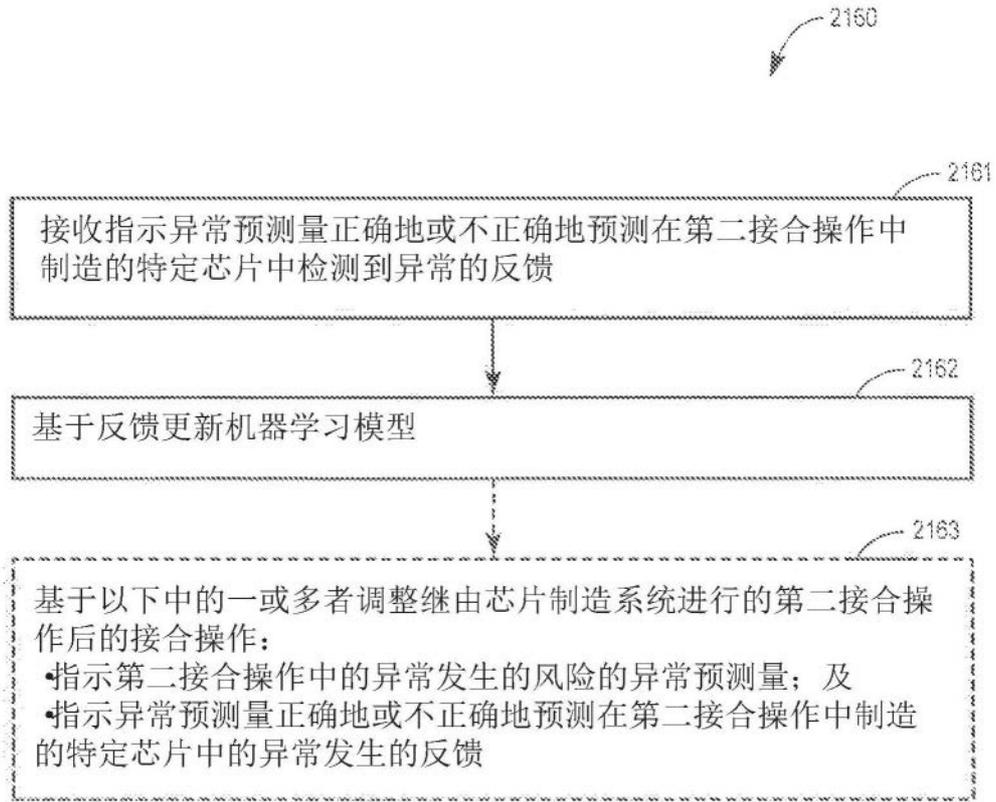


图21C

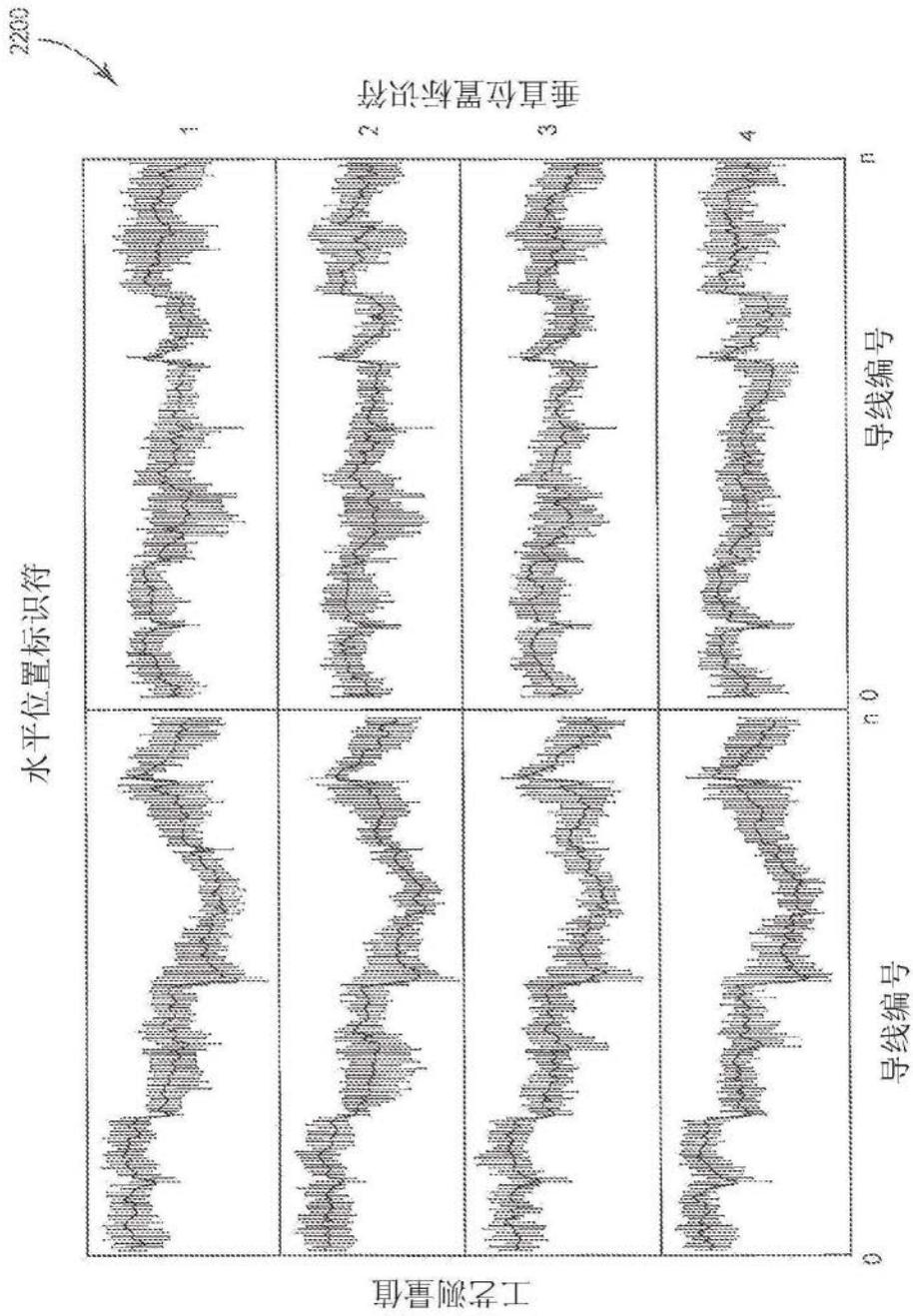


图22A

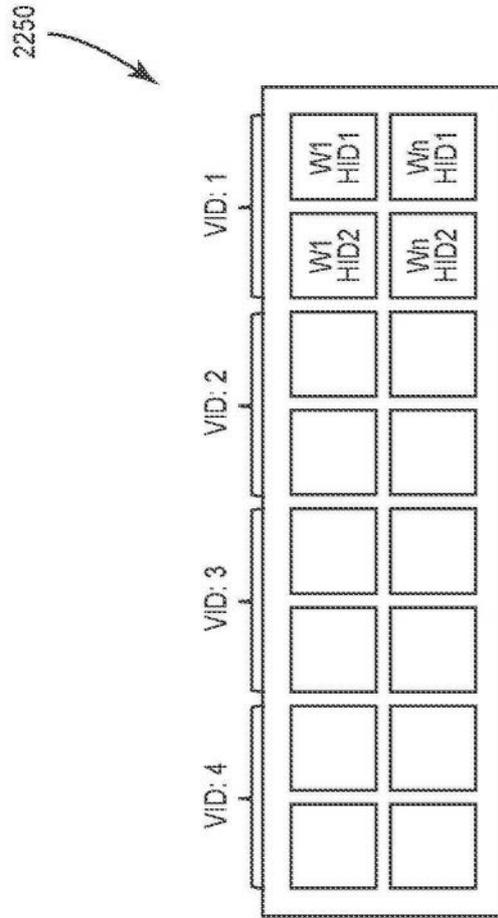


图22B

The table is a 3x3 grid. The first column contains the text '模型目标变量 (质量保证值)', '球剪切', and '针脚拉拔'. The second column contains '49,114' and '48,765'. The third column contains '12,280' and '12,193'. A bracket labeled '2301' spans the first column. A bracket labeled '2302' spans the first two columns. A bracket labeled '2303' spans the second and third columns. An arrow labeled '2300' points to the entire table.

模型目标变量 (质量保证值)	训练数据	测试数据
球剪切	49,114	12,280
针脚拉拔	48,765	12,193

图23A

导线编号	HID 编号	中值(正常)	差量值(正常)	异常差量值
1	1	11.82	0.08	0.2
1	2	11.70	-0.09	0.12
2	1	11.92	-0.05	0.2
2	2	11.88	0.03	-0.1
...		
N	1	12.20	0.20	1.2
N	2	12.28	-0.16	1.09

2351

2352

2353

2354

2355

2350

图23B

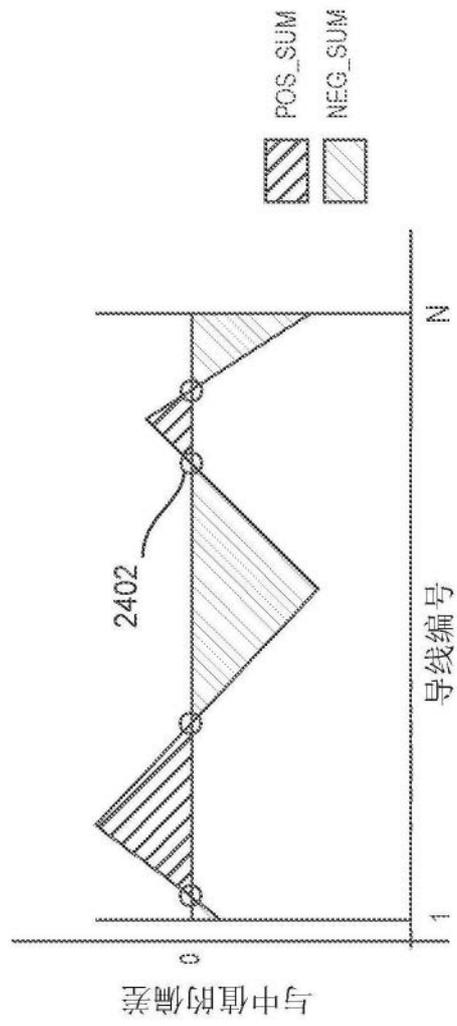


图24

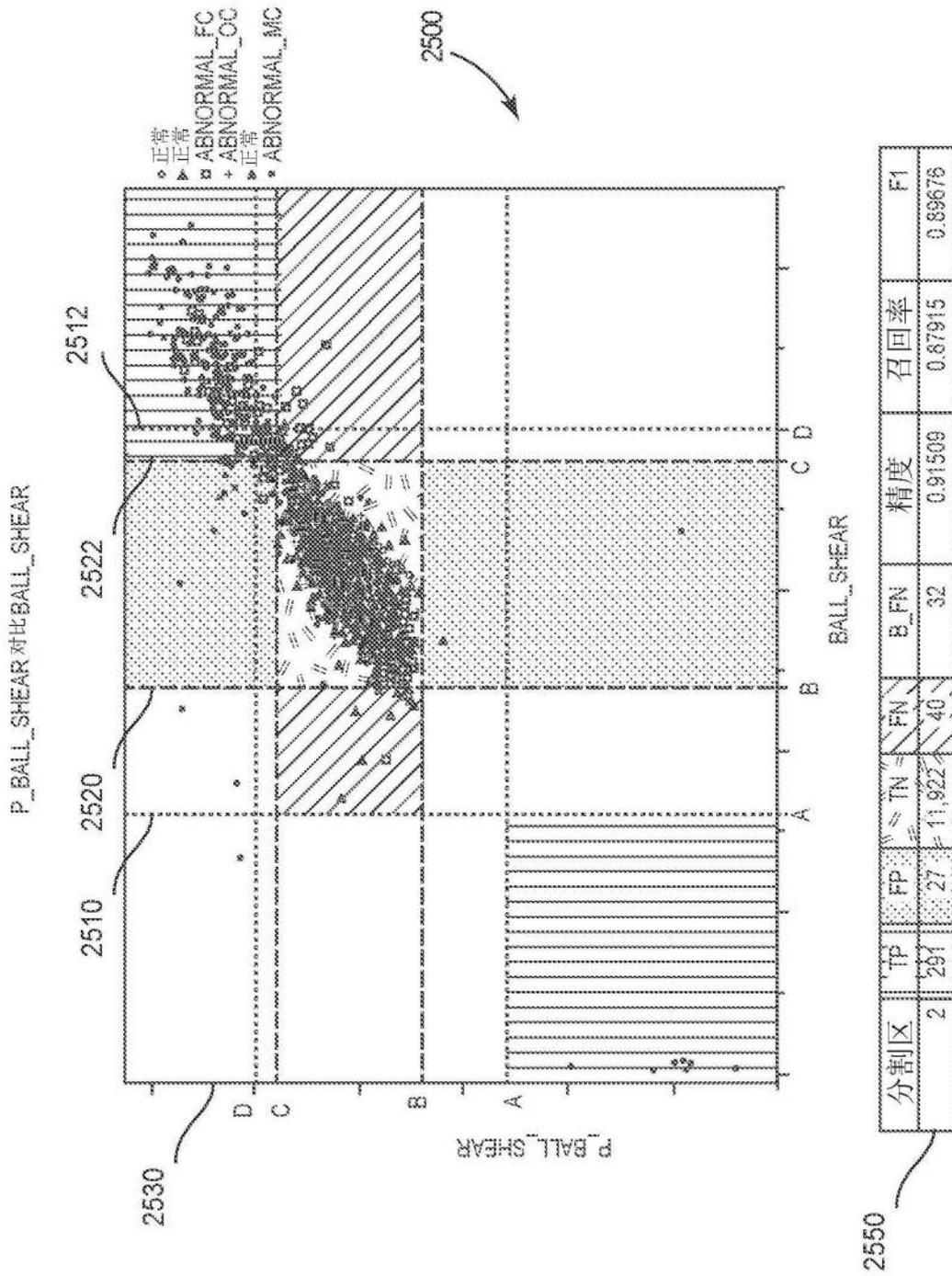


图25

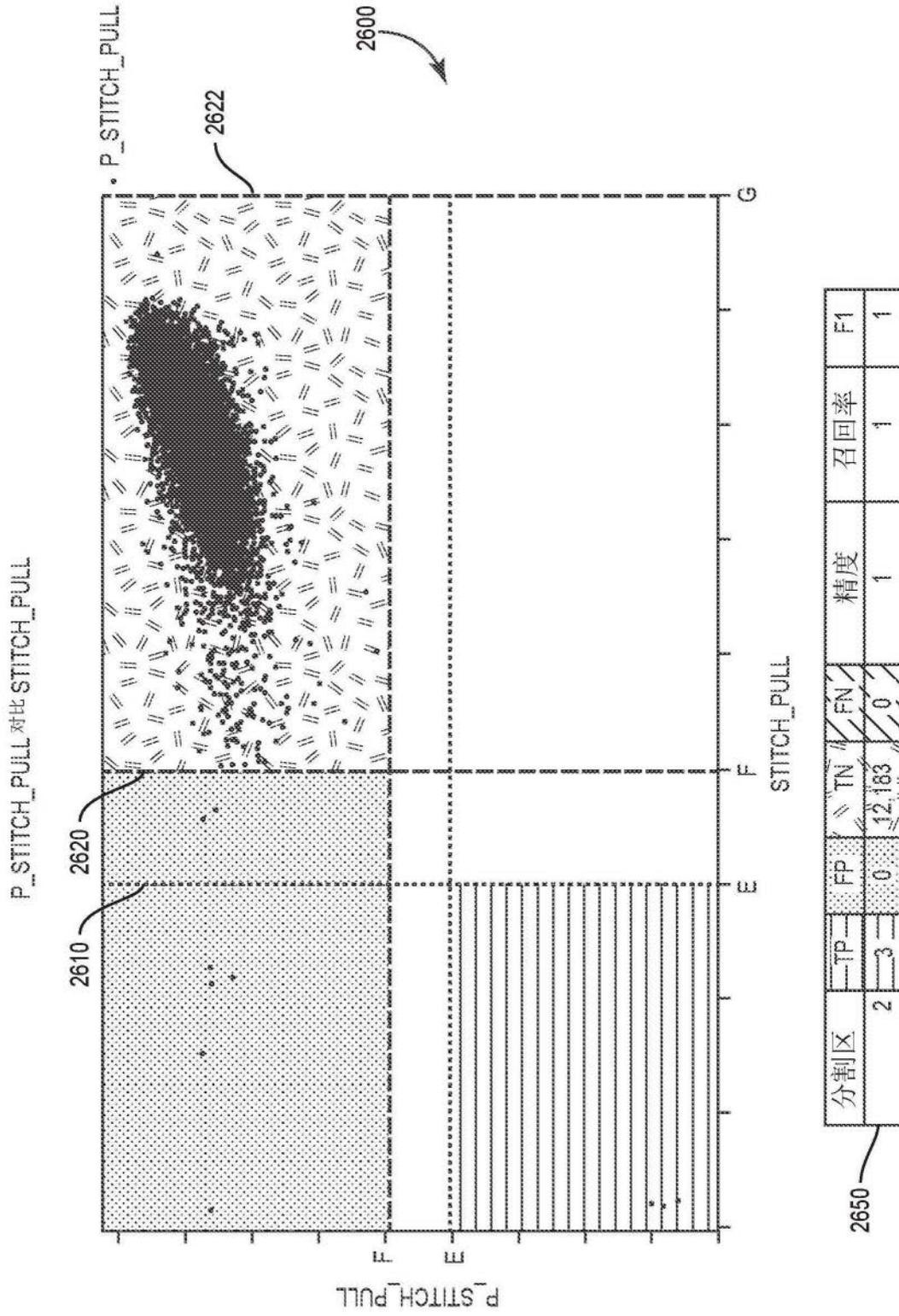


图26

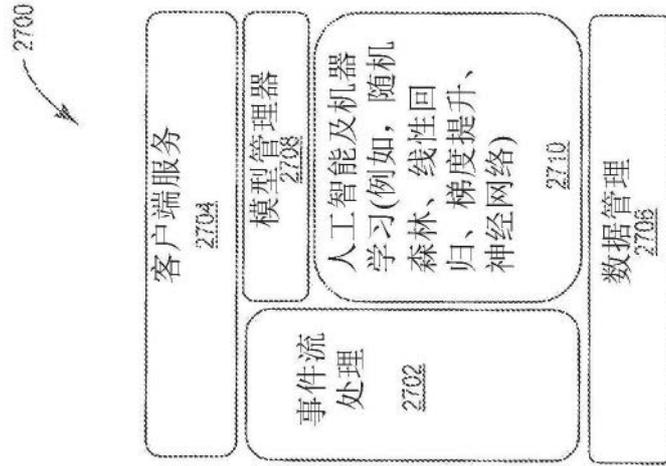


图27A

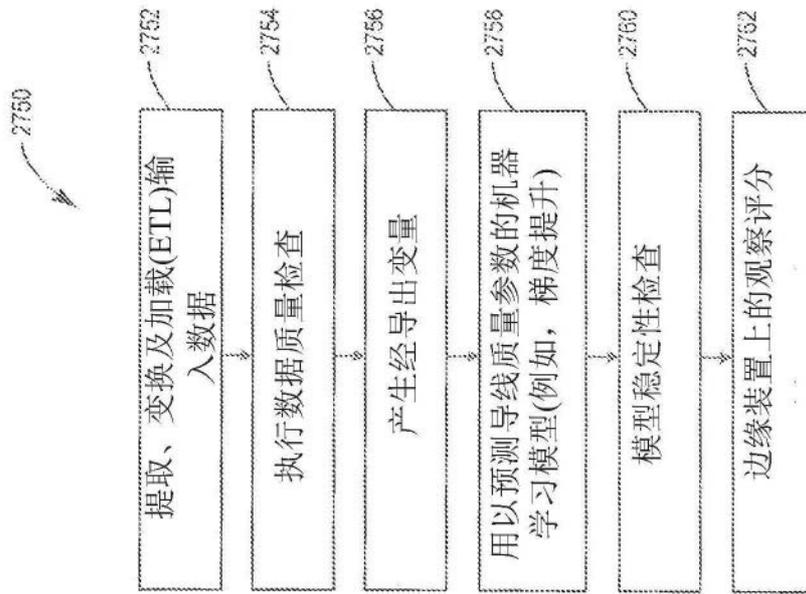


图27B

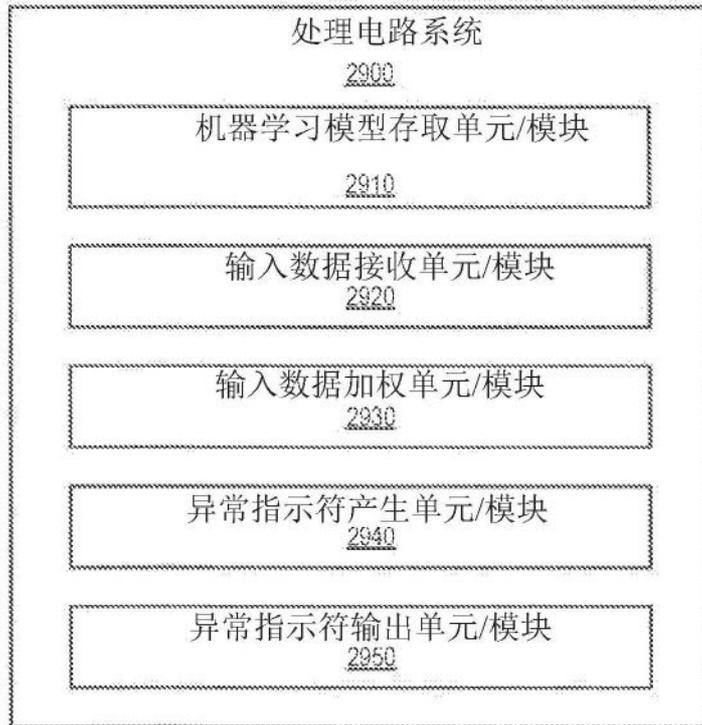


图28