



(12)发明专利

(10)授权公告号 CN 106796419 B

(45)授权公告日 2020.09.11

(21)申请号 201580054441.2

(22)申请日 2015.10.05

(65)同一申请的已公布的文献号

申请公布号 CN 106796419 A

(43)申请公布日 2017.05.31

(30)优先权数据

62/060,408 2014.10.06 US

14/873,721 2015.10.02 US

(85)PCT国际申请进入国家阶段日

2017.04.06

(86)PCT国际申请的申请数据

PCT/US2015/053931 2015.10.05

(87)PCT国际申请的公布数据

W02016/057365 EN 2016.04.14

(73)专利权人 费希尔-罗斯蒙特系统公司

地址 美国德克萨斯州

(72)发明人 P·佐尔尼欧 M·J·尼克松

W·K·沃杰茨尼斯 M·J·卢卡斯

E·D·罗特沃尔德 P·R·穆斯顿

T·L·布莱文斯 G·K·劳

(74)专利代理机构 永新专利商标代理有限公司
72002

代理人 曹雯

(51)Int.Cl.

G05B 19/042(2006.01)

G05B 19/418(2006.01)

(56)对比文件

CN 1170464 A,1998.01.14

CN 104035392 A,2014.09.10

GB 2485092 A,2012.05.02

CN 104049575 A,2014.09.17

JP 2001036542 A,2001.02.09

审查员 马兵

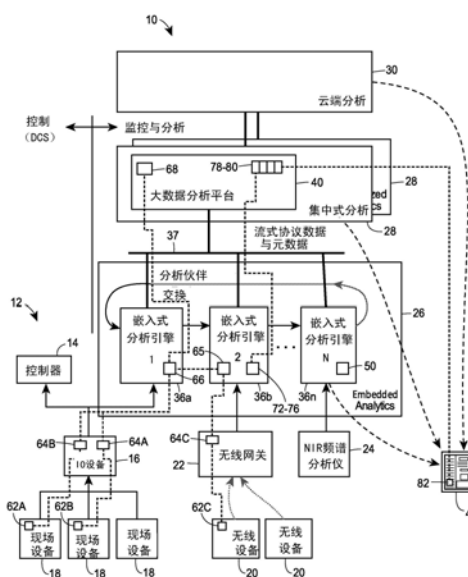
权利要求书6页 说明书41页 附图11页

(54)发明名称

用于过程控制系统分析的数据管道

(57)摘要

一种被用作基本处理元件以用于实现在过程工厂或监控系统中自动或自主地执行基于信号处理的学的技术的数据管道。每个数据管道包括一组通信互连的数据处理块,其以预定顺序对一个或多个数据源执行处理,以便例如清除数据、滤波数据、选择进行进一步处理的数据、对数据执行监督或无监督学习等。在数据管道内的各个处理块或模块可以在工厂网络中的不同设备处被存储和执行以进行分布式数据处理。此外,每个数据管道可以被集成到一个或多个较高级别的分析模块中,该分析模块对经处理的数据执行诸如质量预测、故障检测等较高级别的分析。在工厂网络内使用数据管道使得工厂控制或监控系统中收集的数据在工厂的持续运行期间能够被自动处理并在工厂内的各种较高级别的分析模块中使用。



1. 一种用于在实现过程的过程工厂环境内使用的工厂通信系统,包括:

过程控制网络,所述过程控制网络包括设置在所述过程工厂内的多个过程控制设备以控制所述过程,以及通信地耦合到所述多个过程控制设备的过程控制通信网络,其中,所述多个过程控制设备中的一个或多个过程控制设备收集或生成过程控制数据;

另外的工厂通信网络,所述另外的工厂通信网络包括通过另外的工厂通信网络链路通信地耦合在一起的多个处理设备,其中,所述另外的工厂通信网络上的所述多个处理设备中的至少一个处理设备通信地连接到所述过程控制通信网络;以及

分布式数据管道,所述分布式数据管道在配置环境中被定义为通过使用跨所述过程工厂环境的多个不同设备以特定通信顺序通信地连接在一起的一系列数据处理模块来对过程控制数据执行特定处理操作,并且所述分布式数据管道的不同部分从所述配置环境被下载到所述过程工厂环境的所述多个不同设备中以用于对所述特定处理操作的运行时间执行,所述分布式数据管道包括耦合到一组下游数据处理模块的一个或多个上游数据处理模块,其中:

所述分布式数据管道内的所述一组下游数据处理模块中包括的每一个数据处理模块处理从所述分布式数据管道内的相应的一个或多个上游数据处理模块接收的数据,

所述分布式数据管道的所述一个或多个上游数据处理模块中的至少一个设置在连接到所述过程控制通信网络的所述多个过程控制设备中的一个中,并且

所述分布式数据管道的所述下游数据处理模块中的至少一个设置在所述另外的工厂通信网络上的所述多个处理设备中的一个中。

2. 根据权利要求1所述的工厂通信系统,还包括通信地耦合到所述分布式数据管道的所述下游数据处理模块中的一个下游数据处理模块的数据库,用于储存由所述分布式数据管道的所述下游数据处理模块中的所述一个下游数据处理模块生成的数据,并使所储存的数据经由所述另外的工厂通信网络可用于其它应用,所述其它应用包括至少一个分析模块或应用。

3. 根据权利要求1所述的工厂通信系统,其中,所述上游数据处理模块中的一个上游数据处理模块设置在所述过程控制网络中并且收集用于由所述分布式数据管道处理的原始过程控制数据,并且还包括通信地耦合到所述过程控制通信网络的数据库,所述数据库用于独立地储存所收集的原始过程控制数据。

4. 根据权利要求3所述的工厂通信系统,其中,所述数据库被设置在耦合到所述另外的工厂通信网络的所述处理设备中的一个处理设备中。

5. 根据权利要求1所述的工厂通信系统,其中,所述另外的工厂通信网络是大数据通信网络。

6. 根据权利要求1所述的工厂通信系统,其中,所述另外的工厂通信网络包括耦合在一起的多个不同的通信网络。

7. 根据权利要求6所述的工厂通信系统,其中,所述多个不同的通信网络包括第一基于工厂的通信网络和第二基于云端的通信网络。

8. 根据权利要求6所述的工厂通信系统,其中,所述分布式数据管道包括储存在所述多个不同的通信网络中的每一个通信网络内的处理设备中的至少一个不同的数据处理模块。

9. 根据权利要求6所述的工厂通信系统,其中,所述多个不同的通信网络中的至少一个

是无线通信网络。

10. 根据权利要求1所述的工厂通信系统,其中,所述过程控制通信网络包括通信地耦合在一起的多个不同的控制通信网络。

11. 根据权利要求1所述的工厂通信系统,其中,所述数据处理模块中的一个数据收集模块,所述数据收集模块从所述过程控制网络内的数据源获得数据。

12. 根据权利要求11所述的工厂通信系统,其中,所述数据处理模块中的一个数据清除模块,所述数据清除模块清除从所述过程控制网络内的所述数据源获得的数据。

13. 根据权利要求11所述的工厂通信系统,其中,所述数据处理模块中的一个数据对齐模块,所述数据对齐模块对来自所述过程控制网络内的多个数据源的数据进行时间对齐。

14. 根据权利要求11所述的工厂通信系统,其中,所述数据处理模块中的一个敏感性数据分析模块,所述敏感性数据分析模块确定数据对所述工厂内的特定事件的敏感性。

15. 根据权利要求11所述的工厂通信系统,其中,所述数据处理模块中的一个因果关系分析模块,所述因果关系分析模块确定来自数据源的数据与所述工厂内的事件之间的因果关系。

16. 根据权利要求11所述的工厂通信系统,其中,所述数据处理模块中的一个对所述数据执行监督学习的学习模块。

17. 根据权利要求11所述的工厂通信系统,其中,所述数据处理模块中的一个对所述数据执行无监督学习的学习模块。

18. 根据权利要求11所述的工厂通信系统,其中,所述数据处理模块中的一个数据呈现模块,所述数据呈现模块向用户呈现在所述数据管道内生成的数据。

19. 根据权利要求1所述的工厂通信系统,其中,所述分布式数据管道包括一个或多个数据收集模块、一个或多个中间数据处理模块以及一个或多个最终用途数据处理模块,所述一个或多个数据收集模块从所述过程控制网络收集过程数据,所述一个或多个中间数据处理模块处理由另一个中间数据处理模块或所述数据收集模块中的一个或多个收集的数据,所述一个或多个最终用途数据处理模块以能够由一个或多个其它应用访问的方式储存由所述一个或多个中间数据处理模块或所述一个或多个最终用途数据处理模块生成的输出数据。

20. 根据权利要求19所述的工厂通信系统,其中,所述分布式数据管道的所述数据处理模块中的一个或多个在所述分布式数据管道内间歇地连接。

21. 根据权利要求20所述的工厂通信系统,其中,所述一个或多个间歇地连接的数据处理模块中的一个数据 display 模块,所述数据 display 模块向用户显示所述分布式数据管道内产生的数据。

22. 根据权利要求20所述的工厂通信系统,其中,所述一个或多个间歇地连接的数据处理模块中的一个被设置在间歇地连接到所述另外的工厂通信网络的手持设备中。

23. 根据权利要求19所述的工厂通信系统,其中,所述输出数据可由两个或更多个其它应用同时访问。

24. 根据权利要求1所述的工厂通信系统,其中,所述多个过程控制设备中的一个是以控制器执行速率来执行一个或多个控制例程的过程控制器,并且其中,所述过程控制数据

由设置在所述多个过程控制设备中的一个中的所述分布式数据管道的所述一个或多个上游数据处理模块收集,所述一个或多个上游数据处理模块以大于所述控制器执行速率的速率来收集过程控制数据。

25. 根据权利要求24所述的工厂通信系统,其中,所述分布式数据管道内的所述数据处理模块中的两个或更多个数据处理模块以大于所述控制器执行速率的速率来处理收集的数据。

26. 一种用于在实现过程的过程工厂环境内使用的过程工厂通信系统,包括:

过程控制网络,所述过程控制网络包括设置在所述工厂内的多个过程控制设备以控制所述过程,以及通信地耦合在所述多个过程控制设备之间的过程控制通信网络,其中,所述多个过程控制设备中的一个或多个过程控制设备收集或生成过程控制数据;

另外的工厂通信网络,所述另外的工厂通信网络包括通过另外的工厂通信网络链路通信地耦合在一起的多个处理设备,其中,所述另外的工厂通信网络上的所述多个处理设备中的至少一个处理设备通信地连接到所述过程控制通信网络;

多个分布式数据管道,每个分布式数据管道在配置环境中被定义为通过使用跨所述过程工厂环境中包括的相应的多个不同设备以相应的特定次序通信连接在一起的相应一系列不同的数据处理模块来对过程控制数据执行相应的处理操作,从而在运行时间期间对过程控制数据执行所述相应的处理操作,所述不同的数据处理模块从所述配置环境被下载到所述相应的多个不同设备中,其中,每个分布式数据管道包括一个或多个数据收集模块、一个或多个中间数据处理模块以及一个或多个结束数据处理模块,所述一个或多个数据收集模块从所述过程控制网络收集过程数据,所述一个或多个中间数据处理模块处理由另一个中间数据处理模块或所述数据收集模块中的一个或多个收集的数据,所述一个或多个结束数据处理模块使用由所述一个或多个中间数据处理模块生成的数据来产生数据管道输出数据,并且其中,每个分布式数据管道的至少两个数据处理模块分别位于所述过程控制通信网络和所述另外的工厂通信网络内的不同设备中;以及

一组较高级别的分析应用,所述一组较高级别的分析应用通信地连接在所述另外的工厂通信网络内,其中,所述一组较高级别的分析应用中的每一个订阅由所述分布式数据管道中的一个或多个分布式数据管道的所述结束数据处理模块生成的数据。

27. 根据权利要求26所述的过程工厂通信系统,其中,所述一组较高级别的分析应用中的一个间歇地连接到所述分布式数据管道中的一个或多个分布式数据管道。

28. 根据权利要求26所述的过程工厂通信系统,其中,所述分布式数据管道中的一个分布式数据管道的所述数据处理模块在所述过程的在线操作期间在所述分布式数据管道中的所述一个分布式数据管道内处理数据并传送数据。

29. 根据权利要求26所述的过程工厂通信系统,其中,所述较高级别的分析应用中的一个订阅来自所述分布式数据管道中的两个或更多个分布式数据管道的数据。

30. 根据权利要求26所述的过程工厂通信系统,其中,所述较高级别的分析应用中的一个被设置在经由无线通信连接连接到所述另外的工厂网络的手持设备中。

31. 根据权利要求26所述的过程工厂通信系统,还包括通信地耦合到所述分布式数据管道中的一个分布式数据管道的数据库。

32. 根据权利要求31所述的过程工厂通信系统,其中,所述数据库是大数据机器数据

库。

33. 根据权利要求26所述的过程工厂通信系统, 其中, 所述过程控制网络包括将所述多个过程控制设备通信地耦合在一起的两个或更多个通信网络, 并且其中, 所述分布式数据管道中的一个分布式数据管道包括跨所述两个或更多个通信网络中的每一个通信网络分布的相应设备中的数据处理模块。

34. 根据权利要求26所述的过程工厂通信系统, 其中, 所述另外的工厂通信网络包括将所述多个数据处理设备通信地耦合在一起的两个或更多个通信网络, 并且其中, 所述分布式数据管道中的一个分布式数据管道包括跨所述两个或更多个通信网络中的每一个通信网络分布的相应设备中的数据处理模块。

35. 根据权利要求26所述的过程工厂通信系统, 其中, 所述另外的工厂通信网络包括大数据通信网络。

36. 根据权利要求26所述的过程工厂通信系统, 其中, 所述另外的工厂通信网络包括耦合到基于云端的通信网络的区域大数据通信网络。

37. 根据权利要求26所述的过程工厂通信系统, 其中, 所述分布式数据管道中的一个分布式数据管道的所述数据处理模块中的一个数据清除模块, 所述数据清除模块清除从所述过程控制网络内的数据源获得的数据。

38. 根据权利要求37所述的过程工厂通信系统, 其中, 所述分布式数据管道中的所述一个分布式数据管道中的所述数据处理模块中的另一个是数据对齐模块, 所述数据对齐模块对来自所述过程控制网络内的多个数据源的数据进行时间对齐。

39. 根据权利要求37所述的过程工厂通信系统, 其中, 所述分布式数据管道中的所述一个分布式数据管道中的所述数据处理模块中的另一个是敏感性数据分析模块, 所述敏感性数据分析模块确定数据对所述工厂内的特定事件的敏感性。

40. 根据权利要求37所述的过程工厂通信系统, 其中, 所述分布式数据管道中的所述一个分布式数据管道中的所述数据处理模块中的另一个是因果关系分析模块, 所述因果关系分析模块确定来自数据源的数据与所述工厂内的事件之间的因果关系。

41. 根据权利要求37所述的过程工厂通信系统, 其中, 所述分布式数据管道中的所述一个分布式数据管道中的所述数据处理模块中的另一个是执行监督或无监督数据学习的学习模块。

42. 根据权利要求37所述的过程工厂通信系统, 其中, 所述分布式数据管道中的所述一个分布式数据管道中的所述数据处理模块中的另一个是数据呈现模块, 所述数据呈现模块向用户呈现在所述分布式数据管道中的所述一个分布式数据管道内生成的数据。

43. 一种在具有过程控制网络的工厂中处理数据的方法, 所述过程控制网络包括设置在所述工厂内的多个过程控制设备以控制过程, 以及通信地耦合在所述多个过程控制设备之间的过程控制通信网络, 其中, 所述多个过程控制设备中的一个或多个过程控制设备收集或生成过程控制数据, 并且包括另外的工厂网络, 所述另外的工厂网络包括通过另外的工厂通信网络链路通信地耦合在一起的多个处理设备, 所述方法包括:

将分布式数据管道储存在所述过程控制网络内以及所述另外的工厂网络中, 所述分布式数据管道在配置环境中被定义为一组不同的数据处理模块, 所述一组不同的数据处理模块被配置为跨过程工厂环境的多个不同设备以特定顺序通信地连接在一起, 并且所述

分布式数据管道的不同部分从所述配置环境被下载到所述多个不同设备中,从而在运行期间对包括所述过程控制设备和处理设备中的各种不同设备的所述多个不同设备在过程控制数据上执行处理操作;

在所述过程控制网络的在线操作期间,在所述分布式数据管道的所述数据处理模块中的一个或多个数据处理模块处收集数据;

在所述过程控制网络的在线操作期间,处理所述分布式数据管道的所述数据处理模块中的所收集的过程数据;

将通过所述分布式数据管道的所述数据处理模块而得到的经处理的数据储存在计算机存储器设备中;以及

在另一处理设备上执行的另一应用中使用通过所述分布式数据管道的所述数据处理模块而得到的所述数据。

44. 根据权利要求43所述的处理数据的方法,其中,将通过所述分布式数据管道的所述数据处理模块而得到的经处理的数据储存在计算机存储器设备中包括:将所述经处理的数据储存在大数据机器中。

45. 根据权利要求43所述的处理数据的方法,其中,将所述分布式数据管道储存在所述过程控制网络内以及在所述另外的工厂网络中包括:将所述数据处理模块中的至少一个数据处理模块储存在所述多个过程控制设备中的一个过程控制设备内,并且包括:将所述数据处理模块中的另一个数据处理模块储存在所述另外的工厂网络的所述处理设备中的一个处理设备中。

46. 根据权利要求43所述的处理数据的方法,其中,将所述分布式数据管道储存在所述过程控制网络内以及在所述另外的工厂网络中包括:将所述数据处理模块中的至少一个数据处理模块储存在所述另外的工厂网络的第一工厂子网络内的设备中,并且将所述数据处理模块中的另一个数据处理模块储存在所述另外的工厂网络的第二工厂子网络中。

47. 根据权利要求46所述的处理数据的方法,其中,所述第一工厂子网是所述工厂中的大数据网络并且所述第二工厂子网络是基于云端的网络。

48. 根据权利要求43所述的处理数据的方法,其中,在所述过程控制网络的在线操作期间在所述分布式数据管道的所述数据处理模块中的一个或多个数据处理模块处收集数据包括:以比在所述过程控制网络内使用的控制器的执行速率更快的速率来收集数据以执行在线过程控制。

49. 根据权利要求43所述的处理数据的方法,其中,所述过程控制网络或所述另外的工厂网络中的至少一个包括无线通信网络。

50. 根据权利要求43所述的处理数据的方法,其中,所述数据处理模块中的一个数据清除模块,所述数据清除模块执行以清除从所述过程控制网络内的数据源获得的数据。

51. 根据权利要求43所述的处理数据的方法,其中,所述数据处理模块中的一个数据对齐模块,所述数据对齐模块执行以对来自所述过程控制网络内的多个数据源的数据进行时间对齐。

52. 根据权利要求43所述的处理数据的方法,其中,所述数据处理模块中的一个敏感性数据分析模块,所述敏感性数据分析模块执行以确定数据对所述工厂内的特定事件的敏感性。

53. 根据权利要求43所述的处理数据的方法, 其中, 所述数据处理模块中的一个因果关系分析模块, 所述因果关系分析模块执行以确定来自数据源的数据与所述工厂内的事件之间的因果关系。

54. 根据权利要求43所述的处理数据的方法, 其中, 所述数据处理模块中的一个执行以执行监督或无监督数据学习的学习模块。

55. 根据权利要求43所述的处理数据的方法, 其中, 所述数据处理模块中的一个数据呈现模块, 所述数据呈现模块执行以向用户呈现在所述分布式数据管道内生成的数据。

56. 根据权利要求43所述的处理数据的方法, 其中, 所述分布式数据管道包括一个或多个数据收集模块、一个或多个中间数据处理模块以及一个或多个最终用途数据处理模块, 所述一个或多个数据收集模块从所述过程控制网络收集过程数据, 所述一个或多个中间数据处理模块处理由另一个中间数据处理模块或所述数据收集模块中的一个或多个收集的数据, 所述一个或多个最终用途数据处理模块以能够由一个或多个其它应用访问的方式储存由所述一个或多个中间数据处理模块或所述一个或多个最终用途数据处理模块生成的输出数据。

57. 根据权利要求56所述的处理数据的方法, 其中, 所述分布式数据管道的所述数据处理模块中的一个数据处理模块在所述分布式数据管道内间歇地连接。

58. 根据权利要求57所述的处理数据的方法, 其中, 所述间歇地连接的数据处理模块被设置在间歇地连接到所述另外的工厂通信网络的手持设备中。

用于过程控制系统分析的数据管道

[0001] 相关申请的交叉引用

[0002] 这是常规提交的申请,其要求2014年10月6日提交的题为“Data Pipeline for Process Control System Analysis”的美国临时专利申请序列号No.62/060,408的优先权和权益,其全部公开内容在此通过引用明确地并入本文。

[0003] 本申请还涉及2013年3月3日提交的题为“BIG DATA IN PROCESS CONTROL SYSTEMS”的美国专利申请No.13/784,041;2013年9月17日提交的题为“METHOD AND APPARATUS FOR CONTROLLING A PROCESS PLANT WITH LOCATION AWARE MOBILE CONTROL DEVICES”的美国专利申请No.14/028,785;2014年2月6日提交的题为“COLLECTING AND DELIVERING DATA TO A BIG DATA MACHINE IN A PROCESS CONTROL SYSTEM”的美国专利申请No.14/174,413;2014年3月14日提交的题为“DISTRIBUTED BIG DATA IN A PROCESS CONTROL SYSTEM”的美国专利申请No.14/212,493;2014年3月14日提交的题目为“DETERMINING ASSOCIATIONS AND ALIGNMENTS OF PROCESS ELEMENTS AND MEASUREMENTS IN A PROCESS”美国专利申请No.14/212,411;以及2014年3月17日提交的题为“DATA MODELING STUDIO”的美国专利申请No.14/216,823,其要求2013年3月15日提交的题目为“DATA MODELING STUDIO”的美国临时专利申请序列号No.61/798,820的优先权和权益,这些申请中的每一个申请的全部公开内容在此通过引用明确地并入本文。

[0004] 另外,本申请涉及与本申请同时提交的题为“REGIONAL BIG DATA IN PROCESS CONTROL SYSTEMS”的美国专利申请No.14/507,188;与本申请同时提交的题为“AUTOMATIC SIGNAL PROCESSING-BASED LEARNING IN A PROCESS PLANT”的美国专利申请No.14/507,252;以及与本申请同时提交的题为“STREAMING DATA FOR ANALYTICS IN PROCESS CONTROL SYSTEMS”的美国专利申请No.14/506,863,这些申请中的每一个申请的全部公开内容在此通过引用明确地并入本文。

技术领域

[0005] 概括地,本专利涉及控制和监控系统(诸如在过程工厂中使用的系统),更具体而言,涉及安装和使用数据管道(data pipeline)以便以分布式方式在监控和控制系统中处理数据。

背景技术

[0006] 如在化学、石油、工业或其它过程工厂中用于制造、精炼、转换、生成或生产物理材料或产品的过程控制系统通常包括一个或多个过程控制器,其经由模拟、数字或组合的模拟/数字总线或经由无线通信链路或网络通信地耦合到一个或多个现场设备。现场设备(其可以是例如阀、阀定位器、开关和变送器(例如,温度、压力、液位和流动速率传感器))位于过程环境内,并且通常执行物理或过程控制功能(诸如打开或关闭阀、测量诸如温度或压力等的过程和/或环境参数)以控制在过程工厂或系统内执行的一个或多个过程。智能现场设备(诸如符合公知的现场总线协议的现场设备)还可以执行控制计算、报警功能和通常在控

制器内执行的其它控制功能。过程控制器(其通常也位于工厂环境内)接收表示由现场设备进行的过程测量的信号和/或与现场设备相关的其它信息,并且执行控制器应用,该控制器应用运行例如不同的控制模块,这些控制模块做出过程控制决策、基于所接收的信息生成控制信号,并且与在现场设备(诸如HART[®]、WirelessHART[®]和FOUNDATION[®] Fieldbus现场设备)中执行的控制模块或块进行协调。控制器中的控制模块在通信线路或链路上将控制信号发送给现场设备,由此控制过程工厂或系统的至少一部分的操作。例如,控制器和现场设备控制由过程工厂或系统控制的过程的至少一部分。

[0007] 来自现场设备和控制器的信息通常可以通过数据高速通道或通信网络而可用于一个或多个其它硬件设备(诸如操作员工作站、个人计算机或计算设备、数据历史库、报告生成器、集中式数据库或其它集中式管理计算设备),其通常被放置在控制室或远离较苛刻的设备环境的其它位置。这些硬件设备中的每一个通常跨过程工厂或跨过程工厂的一部分而集中。这些硬件设备运行应用,这些应用例如可以使操作员能够执行关于控制过程和/或操作过程工厂的功能(诸如更改过程控制例程的设置、修改控制器或现场设备内的控制模块的操作、查看过程的当前状态、查看由现场设备和控制器生成的报警、为了训练人员或测试过程控制软件的目的仿真过程的操作、保持和更新配置数据库等)。由硬件设备、控制器和现场设备使用的数据高速通道可以包括有线通信路径、无线通信路径或有线和无线通信路径的组合。

[0008] 作为示例,由艾默生过程管理公司出售的DeltaV[™]控制系统包括储存在位于过程工厂内的不同位置的不同设备并由其执行的多个应用。位于一个或多个工作站或计算设备中的配置应用使用户能够创建或更改过程控制模块,并经由数据高速通道将这些过程控制模块下载到专用分布式控制器。通常,这些控制模块由通信地互连的功能块组成,这些功能块是面向对象的编程协议中的对象,这些对象基于对其的输入执行控制方案内的功能并且向控制方案内的其它功能块提供输出。配置应用还可以允许配置设计者创建或更改操作员界面,所述操作员界面由查看应用使用以向操作员显示数据并且使操作员能够在过程控制例程内更改设置(诸如设定点)。每个专用控制器以及在某些情况下一个或多个现场设备储存和执行相应的控制器应用,所述控制器应用运行向其分配和下载的控制模块以实现实际的过程控制功能。可以在一个或多个操作员工作站上(或在与操作员工作站和数据高速通道通信连接的一个或多个远程计算设备上)执行的查看应用经由数据高速通道从控制器应用接收数据并使用用户界面向过程控制系统设计者、操作员或用户显示该数据,并且可以提供多个不同视图中的任意个视图(诸如操作员视图、工程师视图、技术人员视图之类的)。数据历史库应用通常储存在数据历史库设备中并且由其执行,该数据历史库设备收集并储存跨数据高速通道提供的一些或全部数据,而配置数据库应用可以在附接到数据高速通道的另外的计算机中运行以储存当前过程控制例程配置和与其相关联的数据。替代地,配置数据库可以位于与配置应用相同的工作站中。

[0009] 在过程工厂或过程控制系统或其它监控系统中,当发生异常状况或故障的迹象时(例如,当生成报警时,或当发现过程测量或致动设备有过度变化时),操作员、仪器技术人员或过程工程师通常使用分析工具结合他或她对由系统控制的过程的知识及其通过系统的流动路径,来尝试确定可能对异常状况或故障的迹象的产生有贡献的上游测量值和过程变量。例如,操作员可以将时间上从过程控制设备(例如,现场设备、控制器等)的输出捕

获了的数据的历史日志馈入DeltaV™批量分析产品或连续数据分析工具,以尝试确定各种过程变量和/或测量值对异常或故障状况的贡献。通常,用户基于他或她的过程知识来决定将哪些历史数据日志馈入分析工具并识别候选上游因素(例如,测量值、过程变量等)。随后,这些数据分析工具利用例如主成分分析(PCA)来确定哪些候选上游因素影响下游预测的质量参数。因此,由分析工具提供的输出的准确性和有效性是基于或限于用户的知识,并且因此可能不能对异常状况或故障的来源提供完整或正确的洞察。

[0010] 另外,当前已知的过程控制设备和过程控制系统的架构受到有限的控制器和设备存储器、通信带宽以及控制器和设备处理器能力的强烈影响。例如,在当前已知的过程控制系统架构中,控制器中动态和静态非易失性存储器的使用通常被最小化,或者至少被小心地管理。因此,在系统配置(例如,先验)期间,用户通常必须选择控制器中的哪些数据被归档或保存,将被保存的频率,以及是否使用压缩,并且控制器被相应地配置有这种有限的数据规则集。因此,在故障排除和过程分析中可能有用的数据通常不被归档,并且如果收集到了该数据,则有用的信息可能由于数据压缩而已丢失。

[0011] 上文所讨论的当前已知的过程工厂和过程控制系统的限制以及其它限制可能在过程工厂或过程控制系统的操作和优化中(例如在工厂操作、故障排除和/或预测建模期间)不合期望地显示它们自身。例如,这种限制强制必须被执行的繁琐且冗长的工作流程,以便获得用于故障排除的历史数据,手动地将数据馈送到独立的离线信号处理工具中,并且手动地监督对信号处理工具的输出的分析以生成经更新的模型。即使这样,故障排除结果和模型可能不完整或不完全代表实际系统,因为对其生成的输入依赖于特定操作员的经验和知识。

[0012] “大数据”通常是指一个或多个数据集的集合,这些数据集如此大或复杂以致传统的数据库管理工具和/或数据处理应用(例如,关系型数据库和桌面统计包)在可容忍的时间内不能管理这些数据集。通常,使用大数据的应用是事务性和最终用户定向的或聚焦的。例如,网页搜索引擎、社交媒体应用、营销应用和零售应用可以使用和操纵大数据。大数据可以由分布式数据库支持,该分布式数据库允许现代化的多进程、多核服务器的并行处理能力被充分利用。

发明内容

[0013] 一种用于提供基于分布式信号处理的数据处理和分析以确定例如潜在的故障源、异常操作和/或通过控制过程工厂中的过程而生成的信号的行为的变化的系统和方法包括:在过程工厂或过程控制系统的各种互连的设备内形成和安装一个或多个数据管道,其中,数据管道以预定方式自动收集和处理数据以执行部分或完整的数据分析。数据管道可以用于实现基于信号处理的学习,其可以是例如一种基于大数据的学习。一般而言,使用数据管道的系统和方法可以自动执行实时数据处理,或者对在过程工厂中实时收集的数据,或者对储存在诸如大数据机器之类的数据库中的数据执行数据处理。在一些情况下,使用数据管道架构来执行的一些信号处理被集成到过程工厂中,并且当数据由信号源产生时被实时地执行为接近于一个或多个信号的源。由数据管道架构执行信号处理的数据可以是例如指示过程控制设备的输出信号的在时间上的值、过程变量、诸如温度、压力、流动速率等的物理参数的测量值、能量平衡、质量平衡、性能参数、分析功能的输出和/或基于在过程工

厂中被控制的过程而生成的任何其它值。此外,数据管道架构可以处理除了过程控制数据之外的其它类型的数据,诸如监控数据(由不执行控制的监控系统产生)、指示设备操作参数的设备数据、由维护人员用于维修、修理和更换网络中的设备的维护数据、通信数据等。

[0014] 此外,数据管道架构可以用在例如过程工厂网络中执行基于大数据的学习的系统中,并且数据处理管道可以包括多个数据处理模块,其通信地连接在一起以特定的顺序对数据执行预配置的处理操作以产生经处理的数据。数据处理管道可以执行数据处理操作,其包括例如数据收集、数据选择和清除、数据对齐、敏感性分析、因果关系分析、监督学习和/或无监督学习。这些单独的模块或块可以位于工厂网络内的相同或不同的设备中,并且可以通过工厂网络内的一个或多个不同的通信网络被通信地连接在一起。

[0015] 数据管道可以在数据建模工作室(studio)或配置系统中生成,其中,用户选择将被应用于特定过程数据的各种数据处理模块,并将模块连接在一起以形成具有以特定顺序连接在一起的多个处理模块的数据管道。用户可以对储存的历史数据(诸如储存在大数据设备中的数据)测试数据管道,并且在测试完成后,可以将数据管道下载到工厂网络的各种处理器、节点或设备,以实现实时或非实时数据中的数据处理,因为该数据是由工厂设备收集或生成的。此外,数据建模工作室配置系统可以使用户能够在管道被下载到工厂网络时指定每个数据处理管道元件的节点、设备或其它位置。

[0016] 在一个实施例中,数据管道运行或延伸通过过程工厂内或与其相关联的各种(例如,两个或更多个)通信网络,该过程工厂包括过程控制网络和另外的通信网络。更具体地,一种用于在实现过程的过程工厂环境中使用的工厂通信系统包括:过程控制网络,所述过程控制网络包括设置在过程工厂内的多个过程控制设备以控制过程,以及通信地耦合到所述多个过程控制设备的过程控制通信网络,其中,所述多个过程控制设备中的一个或多个过程控制设备收集或生成过程控制数据。该系统还包括另外的工厂通信网络,其包括通过另外的工厂通信网络链路通信地耦合在一起的多个处理设备,其中,所述另外的工厂通信网络上的所述多个处理设备中的至少一个处理设备通信地连接到所述过程控制通信网络。在该系统中,在所述工厂通信系统和所述另外的工厂通信网络内设置分布式数据管道,并且所述分布式数据管道包括以特定通信顺序通信地连接在一起的一系列数据处理模块,用于以该特定通信顺序对过程控制数据执行处理,以使得所述数据管道包括耦合到一组下游数据处理模块的一个或多个上游数据处理模块。所述数据管道内的所述一组下游数据处理模块中的每一个处理从所述数据管道内的一个或多个上游数据处理模块接收的数据,并且所述数据管道的所述一个或多个上游数据处理模块中的至少一个设置在连接到所述过程控制通信网络的所述多个过程控制设备中的一个中,并且所述数据管道的所述下游数据处理模块中的至少一个设置在所述另外的工厂通信网络上的所述多个处理设备中的一个中。

[0017] 该系统还可以包括通信地耦合到所述数据管道的所述下游数据处理模块中的一个下游数据处理模块的数据库,用于储存由所述数据管道的所述下游数据处理模块中的所述一个下游数据处理模块生成的数据,并使所储存的数据经由所述另外的工厂通信网络可用于其它应用。如果需要,所述上游数据处理模块中的一个上游数据处理模块设置在所述过程控制网络中并且收集用于由所述数据管道处理的原始过程控制数据,并且数据库通信地耦合到所述过程控制通信网络,以用于独立地储存所收集的原始过程控制数据。所述数据库可以被设置在耦合到所述另外的工厂通信网络的所述处理设备中的一个处理设备中,

所述另外的工厂通信网络可以是大数据通信网络。所述另外的工厂通信网络可以将多个不同的通信网络耦合在一起,所述多个不同的通信网络包括第一基于工厂的通信网络和第二基于云端的通信网络。在这种情况下,所述数据管道可以包括储存在所述多个不同通信网络中的每一个通信网络内的处理设备中的至少一个不同的数据处理模块。所述不同的通信网络可以包括无线和/或有线通信网络,并且所述过程控制通信网络可以包括通信地耦合在一起的多个不同的控制通信网络。

[0018] 如果需要,所述数据处理模块可以是或包括从所述过程控制网络内的数据源获得数据的数据收集模块、清除从所述过程控制网络内的所述数据源获得的数据的数据清除模块、对来自所述过程控制网络内的多个数据源的数据进行时间对齐的数据对齐模块、确定数据对工厂内的特定事件的敏感性的敏感性数据分析模块、确定来自所述工厂内的数据源和事件的因果关系分析模块、对过程数据执行监督或无监督学习的学习模块或者向用户呈现在所述数据管道内生成的数据的数据呈现模块中的任何一个。

[0019] 另外,所述工厂通信系统可以被构造成使得所述数据管道包括一个或多个数据收集模块,其从所述过程控制网络收集过程数据,一个或多个中间数据处理模块,其处理由另一个中间数据处理模块或所述数据收集模块中的一个或多个收集的数据,以及一个或多个最终用途(end-use)数据处理模块,其以能够由一个或多个其它应用访问的方式储存由所述一个或多个中间数据处理模块或所述一个或多个最终用途数据处理模块生成的输出数据。此外,所述数据管道的所述数据处理模块的一个或多个数据处理模块可以在所述数据管道内间歇地连接到所述数据管道内的其它数据处理模块或者从所述数据管道订阅数据的应用。例如,一个或多个间歇地连接的数据处理模块可以是向用户显示在所述数据管道内产生的数据的数据显示模块,这些模块可以被设置在间歇地连接到所述另外的工厂通信网络等的手持设备中。

[0020] 另外,所述数据管道的输出数据可以被两个或更多个其它应用同时访问。此外,所述数据管道模块可以以任何期望的执行速率执行,该速率对于所述数据管道内的不同模块可能是不同的。同样地,任何所述数据管道模块可以以快于、慢于、等于正在控制在所述数据管道内收集和分析过程数据的过程的过程控制器的执行速率的速率执行。这里,所述多个过程控制设备中的一个是以控制器执行率执行一个或多个控制例程的过程控制器,以及可以以大于所述控制器执行速率的速率来收集由设置在所述多个过程控制设备中的一个过程控制设备中的所述数据管道的所述一个或多个上游数据处理模块收集的所述过程控制数据。此外,所述数据管道内的任何数量(例如两个或更多个)数据处理模块可以以大于所述控制器执行率的速率来处理收集的数据。

[0021] 在另一个实施例中,数据管道被设置在过程工厂的一个或多个通信网络中,该数据管道以比所述过程工厂的所述过程控制系统的所述控制器执行速率更快的速率处理数据。具体地,一种用于控制工厂内的过程的过程工厂控制和通信系统包括:设置在所述工厂内的多个过程控制设备以控制过程,所述多个过程控制设备包括过程控制器,所述过程控制器以执行速率来执行以控制所述过程。所述工厂控制和通信系统还包括设置在所述工厂内的多个数据处理设备以及通信地耦合所述多个过程控制设备和所述多个数据处理设备的一个或多个通信网络。这里,一个或多个数据源被设置在所述工厂内,并且这些数据源分布在所述多个过程控制设备之间,用于以比所述控制器的所述执行速率大的收集数据速率从所

述工厂内的所述过程控制设备收集原始数据。此外,数据库通信地耦合到所述一个或多个通信网络,用于储存以所述收集数据速率从所述一个或多个数据源所收集的原始数据,并且数据管道设置在所述工厂内的所述多个处理设备中的一个或多个处理设备中,所述数据管道包括以特定通信顺序通信地连接在一起的一系列数据处理模块,其中,所述数据管道的所述数据处理模块中的至少一个数据处理模块对来自所述一个或多个数据源的原始数据执行数据处理操作,并且所述数据处理模块中的至少另一个数据处理模块对从所述数据管道中的所述数据处理模块中的一个数据处理模块接收的数据执行数据处理。这里,所述数据处理模块中的至少两个数据处理模块以比所述过程控制器的所述执行速率更快的速率进行操作。

[0022] 在该系统中,所述数据处理模块中的所述至少两个数据处理模块可以通过以比所述过程控制器的所述执行速率更快的速率在所述两个数据处理模块之间传送数据和/或通过以比所述过程控制器的所述执行速率更快的速率对数据执行处理,来以比所述过程控制器的所述执行速率更快的速率进行操作。

[0023] 在另一个实施例中,多个分布式数据管道被设置在过程工厂中以处理所述工厂内的数据并且将经处理的数据馈送到较高级别的分析应用。在一个示例中,一种用于在实现过程的过程工厂环境内使用的过程工厂通信系统包括:过程控制网络,其具有设置在所述工厂内的多个过程控制设备以控制过程,以及通信地耦合在所述多个过程控制设备之间的过程控制通信网络,其中,所述多个过程控制设备中的一个或多个过程控制设备收集或生成过程控制数据。此外,另外的工厂通信网络设置在所述工厂中,并且包括通过另外的工厂通信网络链路通信地耦合在一起的多个处理设备,其中,所述另外的工厂通信网络上的所述多个处理设备中的至少一个处理设备通信地连接到所述过程控制通信网络。多个分布式数据管道被设置在所述系统中,每个数据管道包括通信连接在一起的一系列不同的数据处理模块,用于以特定顺序对过程控制数据进行处理,其中,每个数据管道包括一个或多个数据收集模块,其从所述过程控制网络收集过程数据,一个或多个中间数据处理模块,其处理由另一个中间数据处理模块或所述数据收集模块中的一个或多个收集的数据,以及一个或多个结束数据处理模块,其使用由所述一个或多个中间数据处理模块生成的数据来产生数据管道输出数据。这里,每个数据管道的至少两个数据处理模块位于所述过程控制通信网络和所述另外的工厂通信网络内的不同设备中。所述系统还包括在所述另外的工厂通信网络内通信连接的一组较高级别的分析应用,其中,所述一组较高级别的分析应用中的每一个订阅由所述分布式数据管道中的一个或多个分布式数据管道的结束数据处理模块生成的数据。

[0024] 如果需要,所述一组较高级别的分析应用中的一个或多个较高级别的分析应用可以间歇地连接到所述分布式数据管道中的一个或多个分布式数据管道,并且所述分布式数据管道中的一个分布式数据管道的数据处理模块在所述过程的在线操作期间在所述分布式数据管道中的所述一个分布式数据管道内处理数据并传送数据。如果需要,所述较高级别的分析应用中的一个或多个较高级别的分析应用可以从所述分布式数据管道的两个或更多个分布式数据管道订阅数据(输出数据或最终用途数据),并且所述较高级别的分析应用中的一个或多个较高级别的分析应用被设置在经由无线通信连接连接到所述另外的工厂网络的手持设备中。

[0025] 在另一个实施例中,一种在工厂中处理数据的方法可以在具有过程控制网络的工厂中使用,所述过程控制网络包括设置在所述工厂内的多个过程控制设备以控制过程,以及通信地耦合在所述多个过程控制设备之间的过程控制通信网络,其中,所述多个过程控制设备中的一个或多个过程控制设备收集或生成过程控制数据,并且所述工厂控制网络包括另外的工厂网络,其包括通过另外的工厂通信网络链路通信地耦合在一起的多个处理设备。所述方法可以包括:将分布式数据管道储存在所述过程控制网络内以及所述另外的工厂网络中,其中,储存所述分布式数据管道包括储存一系列数据处理模块,所述一系列数据处理模块被配置为以特定顺序通信地连接在一起,以对过程控制设备和处理设备中各种不同设备中的过程控制数据执行处理。所述方法还可以包括:在所述过程控制网络的在线操作期间在所述数据管道的所述数据处理模块的一个或多个数据处理模块处收集数据,并且在所述过程控制网络的在线操作期间处理在所述分布式数据管道的所述数据处理模块中所收集的过程数据。此外,所述方法可以包括:将通过所述分布式数据管道的所述数据处理模块而得到(develope)的经处理的数据储存在计算机存储器设备中,并且在另一处理设备上执行的另一应用中使用通过所述分布式数据管道的所述数据处理模块而得到的所述数据。

[0026] 如果需要,将通过所述分布式数据管道的所述数据处理模块而得到的经处理的数据储存在计算机存储器设备中包括:将经处理的数据储存在大数据机器中。此外,将所述分布式数据管道储存在所述过程控制网络内以及储存在所述另外的工厂网络中包括:将所述数据处理模块中的至少一个数据处理模块储存在所述多个过程控制设备中的一个过程控制设备内,并将所述数据处理模块中的另一个数据处理模块储存在所述另外的工厂网络中的所述处理设备中的一个处理设备中。此外,将所述分布式数据管道储存在所述过程控制网络内以及储存在所述另外的工厂网络中包括:将所述数据处理模块中的至少一个数据处理模块储存在所述另外的工厂网络中的第一工厂子网络内的设备中,并且将所述数据处理模块中的另一个数据处理模块储存在所述另外的工厂网络中的第二工厂子网络中。这里,所述第一工厂子网络可以是例如所述工厂中的大数据网络,并且所述第二工厂子网可以是例如基于云端的网络。

[0027] 再次,如果需要,在所述过程控制网络的在线操作期间在所述数据管道的所述数据处理模块的一个或多个数据处理模块处收集数据可以包括:以比在所述过程控制网络内使用的控制器的执行速率更快的速率来收集数据以执行在线过程控制。

[0028] 如果需要,所述数据管道可以包括一个或多个数据收集模块,其从所述过程控制网络收集过程数据,一个或多个中间数据处理模块,其处理由另一个中间数据处理模块或所述数据收集模块中的一个或多个收集的数据,以及一个或多个最终用途数据处理模块,其以可由一个或多个其它应用访问的方式储存由所述一个或多个中间数据处理模块或所述一个或多个最终用途数据处理模块生成的输出数据。例如,所述数据管道的所述数据处理模块中的一个数据处理模块可以在所述数据管道内间歇地连接,和/或所述间歇地连接的数据处理模块可以被设置在间歇地连接到所述另外的工厂通信网络的手持设备中。

[0029] 在另一个实施例中,一种在过程工厂使用的配置数据处理管道的方法,所述过程工厂具有过程控制网络,所述过程控制网络操作以控制所述过程的多个过程控制设备,并具有过程通信网络,所述过程通信网络具有通信地耦合在一起的多个处理设备。所述配置

方法包括:将一组数据处理模块储存在所述多个过程控制设备和所述处理设备中的不同设备的内,以固定的方式通信地互连所述数据处理模块以创建串行地通信连接的一组三个或更多个数据处理模块,以及将串行地连接的所述一组三个或更多个数据处理模块配置为对在所述工厂中收集的数据执行特定类型的数据处理。配置步骤可以包括:将所述过程控制网络内的过程控制设备内的一系列数据处理模块中的一个或多个数据处理模块配置成数据收集模块以收集原始过程数据;将所述一系列数据处理模块中的一个或多个数据处理模块配置成中间过程模块以与一个或多个数据收集模块进行通信,并处理由所述数据收集模块收集的所述原始过程数据以创建中间过程数据;以及将所述一系列数据处理模块中的一个或多个数据处理模块配置成最终用途过程模块,以与一个或多个中间数据处理模块进行通信以便处理所述中间过程数据从而生成最终用途过程数据。此外,所述配置方法可以包括:将所述最终用途数据处理模块中的一个或多个最终用途数据处理模块配置为向通信地连接到所述过程通信网络的一个或多个订阅应用提供所述最终用途过程数据。

[0030] 如果需要,所述配置数据处理管道的方法可以包括:通过将所述一系列数据收集模块中的一个或多个数据收集模块配置为以比在所述过程控制网络中执行过程控制的过程控制器的控制器执行速率更快的速率收集所述原始过程数据,来将储存在所述过程控制网络内的过程控制设备内的所述一系列数据处理模块中的所述一个或多个数据处理模块配置成数据收集模块。此外,将所述一系列数据处理模块中的一个或多个数据处理模块配置成中间数据处理模块可以包括:将所述一个或多个中间数据处理模块配置为以比所述过程控制器的所述执行速率更快的速率来处理所述原始过程数据。此外,将所述一系列数据处理模块中的一个或多个数据处理模块配置成最终用途过程模块可以包括:将所述最终用途过程模块中的一个或多个最终用途过程模块配置为以比所述过程控制器的所述执行速率更快的速率来处理所述中间过程数据。

[0031] 在过程工厂内配置数据处理管道的方法可以包括:通过将所述数据处理模块中的一些数据处理模块储存在所述过程控制网络设备中以及将所述数据处理模块的其它数据处理模块储存在所述过程通信网络设备中,来将一组数据处理模块储存在所述多个过程控制设备和所述处理设备中的不同设备内。在过程工厂内配置数据处理管道的方法还可以包括:通过储存不同的数据处理模块以便以特定顺序对过程数据执行处理,来将一组数据处理模块储存在所述多个过程控制设备和所述处理设备中的不同设备内,其中,所储存的一组数据处理模块包括一个或多个上游数据收集模块,所述一个或多个上游数据收集模块从所述过程控制网络收集过程数据,一个或多个中间过程模块,所述一个或多个中间过程模块通信地连接在所述一个或多个数据收集模块或一个或多个其它的中间过程模块的下游,所述一个或多个中间过程模块处理由另一个中间过程模块或所述数据收集模块中的一个或多个收集的数据,以及一个或多个最终用途过程模块,所述一个或多个最终用途过程模块通信地耦合在所述一个或多个中间过程模块的下游,所述一个或多个最终用途过程模块使用来自所述一个或多个中间过程模块的所述中间数据来产生所述最终用途数据。

[0032] 同样地,将所述一系列数据处理模块中的一个或多个数据处理模块配置为处理所述过程数据以生成中间过程数据可以包括:将所述一个或多个中间数据处理模块配置为从两个或更多个上游数据处理模块接收数据,并且组合来自所述两个或更多个上游数据处理模块的数据以创建所述中间过程数据,并且可以包括:将所述一个或多个中间过程模块配

置为向两个或更多个下游数据处理模块提供所述中间过程数据。将所述最终用途过程模块中的一个或多个最终用途过程模块配置为处理所述中间过程数据以生成最终用途过程数据可以包括:将所述一个或多个最终用途过程模块配置为向多个不同的应用提供最终用途数据流,和/或向一个或多个间歇地连接的应用提供最终用途数据。将串行地连接的所述一组三个或更多个数据处理模块配置为对在所述工厂中收集的数据执行特定类型的数据处理包括:将所述数据处理模块中的两个数据处理模块配置为间歇地通信连接在一起。

[0033] 作为示例,在过程工厂内配置数据处理管道的方法可以包括:通过将所述通信地耦合的数据处理模块中的不同的数据处理模块储存在包括所述过程控制网络和过程通信网络的不同通信网络中,来将一组数据处理模块储存在所述多个过程控制设备和所述处理设备中的不同设备内。该储存步骤还可以包括以下步骤中的任何一个或全部:将所述数据处理管道的所述数据处理模块中的一个数据处理模块储存在过程控制现场设备中;将所述数据处理模块中的一个数据处理模块储存在过程控制器中;将所述数据处理模块中的一个数据处理模块储存在区域大数据分析机器中;将所述数据处理模块中的一个数据处理模块储存在基于云端的数据分析机器中;和/或将所述数据处理模块中的一个数据处理模块储存在手持设备中。

附图说明

[0034] 图1是包括过程控制设备、监控设备和大数据设备(appliance)装置的工厂网络的框图,其中在该工厂网络中,可以创建和安装数据管道。

[0035] 图2A-2C是各种不同示例性数据管道的框图,该各种不同示例性数据管道可以在图1的工厂网络中被创建和安装。

[0036] 图3是例示一组数据处理功能的框图,该一组数据处理功能可以以各种方式组合以在工厂网络中创建一个或多个数据管道。

[0037] 图4是用于过程工厂或过程控制以及监控系统的示例性大数据网络的框图,该示例性大数据网络可以支持使用数据管道进行自动或自主的基于信号处理的学习。

[0038] 图5是图4的示例性大数据设备的框图。

[0039] 图6例示了示例性过程工厂或过程控制系统的另一框图,其中可以执行基于数据管道的自动或自主的基于信号处理的学习。

[0040] 图7例示了过程工厂或过程控制系统的示例的框图,该过程工厂或过程控制系统包括收集和存储来自工厂的数据的大数据机器,以及可用于创建和执行数据管道的数据工作室,该数据管道实现一个或多个数据模型,该一个或多个数据模型使用或分析从工厂收集的数据。

[0041] 图8是在通用的工厂环境内连接的图6的数据建模工作室的示例性实施例的框图。

[0042] 图9是图7和图8的数据建模工作室的图形用户界面的屏幕呈现,其可以用于创建用于在实时工厂网络中测试和实现的一个或多个数据管道。

具体实施方式

[0043] 用于在过程工厂或过程控制系统中自动或自主执行基于信号处理的学习的技术、系统和方法包括创建和使用一个或多个数据管道作为基础结构,在其上实时地或在线工厂

网络中执行基于信号处理的学习。通常,这些技术、系统和方法使得能够自动或自主地对由于过程工厂内的控制系统或监控系统的操作以及基于维护活动或维护数据、监控数据、设备数据、业务或规划数据等而生成的数据(例如,信号)执行信号处理或向其应用信号处理。通常,数据管道操作的数据或信号表示基于过程的控制、维护、监控等随时间而变化的参数值。因此,如本文所使用的,术语“信号”通常是指参数值的时间序列(例如控制器的输出的时间序列、扰动变量的时间序列、过程变量测量值的时间序列、监控或测量变量的时间序列等)。与其中用户后验的从过程工厂获得历史信号馈送并将历史信号馈送提供给离线的单独信号处理器的已知过程控制信号处理技术形成相比,本文描述的技术、系统和方法使用数据管道来使一个或多个信号处理功能能够被集成到过程工厂、维护系统、监控系统或过程控制系统中,以便当一个或多个信号源实时生成信号时,实时自主地执行信号处理(例如,作为后台进程)。因此,使用本文公开的技术、系统和方法,信号处理被并入或集成到操作的在线过程工厂或系统中,例如通过将各种信号处理功能放置在或定位在靠近处理工厂或监控系统内的各种信号源,并使这些信号处理活动的结果能够被其它分析模块或工厂内较高的分析应用所使用。的确,在一些实施例,一个或多个信号处理功能可以与信号源成为一体或相集成,而在其它情况下,信号处理功能可以位于更集中式的设备中。无论如何,集成或并入的信号处理功能可以被配置为自主或自动执行由于在线过程工厂或监控系统的实时操作而由信号源生成的实时信号。

[0044] 注意,本文所讨论的技术、系统和方法可以分别应用于单个信号、多个信号,和/或作为一组或整体的多个信号。然而,为了便于本文讨论而不是为了限制的目的,使用单数术语“信号”。

[0045] 使用一个或多个数据管道结构来执行一个或多个信号处理功能(例如,将一个或多个信号处理功能应用于实时信号)可以提供指示信号的一个或多个特征,指示关于信号或一组信号等的其它信息的结果。在另一种情况下,除了数据处理或信号调节之外,可以实现数据管道的使用以执行数据挖掘和数据学习。例如,信号中存在的一个或多个主要频率或谐波可以通过在数据管道中应用一个或多个信号处理功能来确定。在一些过程工厂或过程控制系统中,通过由数据管道执行的数据处理确定的信号的特征被提供给在该过程工厂实时操作期间也自动或自主操作的一个或多个分析功能。分析功能可以确定信号的特征的一个或多个源,例如,基于元件对齐图或类似资源,其指示或标识在信号源的过程上游的元件(例如,过程元件)。在这些或其它情况下,可以使用数据管道来执行数据学习,以识别关于过程工厂、控制系统、监控或监控系统等的其它类型的知识。

[0046] 一般而言,数据管道被用作在工厂内实现的分布式处理系统中的基本结构,以对来自工厂或其它监控系统中各种来源的数据进行清除、滤波、处理以及执行学习和分析。具体地,可以建立一个或多个数据管道以定义对从各种来源(诸如从控制系统资源、维护系统资源、现场设备、控制器、监控系统、实验室分析等)接收的一个或多个数据流执行的特定处理。每个数据管道由多个不同的处理块或数据处理模块组成,这些处理模块或数据处理模块在提供给其的某些信号或数据(例如,数据流)上进行操作,并产生随后提供给数据管道中的下游块或模块的一个或多个数据流。一般而言,数据管道的初始输入将是一组或多组原始数据(诸如控制信号、过程变量测量信号(例如,压力、温度、流量等数据)、监控信号、设备数据、维护数据等)。此外,数据管道中的每个块将以某种方式处理输入到其的数据(例

如,从数据管道中的一个或多个其它数据处理模块接收),以便例如清除数据,组合数据,填充缺失数据,对数据进行滤波,对数据执行学习或分析等。当数据流动通过数据管道的各种不同的数据处理模块时,数据管道创建新数据(可能是中间数据或最终用途数据),其可以被暴露(expose)并且可用于其它数据消费者,包括数据管道内的其它数据处理模块、其它数据管道、诸如较高级别分析应用之类的其它分析模块。此外,数据管道(本文中称为最终用途数据)的输出可以表示或提供可以用于例如设置警报,执行更好的维护或控制,向用户提供图形和图表的分析结果,以示出工厂或监控系统的某些操作或从数据中收集的任何其它知识,如在数据管道内所处理的。此外,可以在工厂网络或其它网络内的不同设备中实现数据管道的各种不同的模块或块,并且数据可以以任何期望的方式并以任何期望的速度或速率在数据管道的不同块之间流动。一般而言,数据管道内的数据处理模块可以采取以下形式:(1)数据收集模块,其从工厂或过程收集诸如原始过程数据、控制器数据、I/O数据、维护数据等的某种数据,(2)中间过程模块,其通信地连接到一个或多个数据收集模块或其它中间通信模块(也就是说,在其下游),并且操作以分析(例如,组合、处理、滤波等)收集的数据或从其它中间过程模块提供到其中的其它数据以产生中间数据,以及(3)最终用途过程模块,其可以通信地连接到一个或多个数据收集模块或中间过程模块(即,在其下游),并且使用来自这些其它模块的数据来创建被提供给一个或多个数据消费者(诸如用户接口、较高级别的分析应用、数据库等)的经处理的数据(最终用途数据)。如果需要,数据管道可以在过程操作期间操作以实时生成最终用途数据,并且可以恒定地、周期性地或间歇地向数据消费者提供最终用途数据。也就是说,最终用途应用可以连接到并保持连接到数据管道,在一致或恒定的基础上从数据管道接收最终用途数据,或者可以周期性地或间歇地连接到数据管道,以在这些应用或用户需要该数据时接收最终用途数据。此外,数据管道内的数据可以以任何期望的速率(诸如比对过程进行控制的一个或多个过程控制器的执行速率更快的速率)收集和处理。在这种情况下,过程工厂内的数据管道独立于过程控制器进行操作(即使数据管道可以使用或收集由过程控制器生成的数据)。

[0047] 因此,如将理解的,数据处理和知识学习技术可以在工厂网络内使用一个或多个数据管道自动或自主地实现,该一个或多个数据管道通过其操作实现对在工厂内收集、生成或测量的数据的特定动作,从而使得能够执行特定类型的较高级别的数据处理和学习。每个数据管道定义了将以特定的(例如,固定的)顺序对一组特定数据或数据集执行的一系列操作或功能,从而执行一些整体处理,其导致关于工厂的知识或工厂内的学习。换言之,每个数据管道以特定的预定义顺序实现一系列预定的数据处理功能,以执行或实现一个或多个数据模型和/或在工厂内实现数据学习。通常,尽管不一定,但是使用至少部分数据管道的过程工厂包括过程控制大数据网络和过程控制大数据网络节点或设备。例如,本文所公开的系统、方法和技术(诸如在上述美国申请No.13/784,041中,上述美国专利申请No.14/174,413号中,上述美国专利申请No.14/212,493中和/或上述题为“REGIONAL BIG DATA IN PROCESS CONTROL SYSTEMS”的美国专利申请No.14/507,188中描述的)中的至少一些可以在支持本地化的、区域的和/或集中式的大数据的过程工厂中实现。因此,使用数据管道执行自动或自主的基于信号处理的学习的过程工厂可以包括一个或多个大数据设备,其中至少一些包括相应的分布式或嵌入式大数据设备以对由大数据提供者节点生成的大数据进行操作。

[0048] 在一个示例性系统中,单独的数据管道可以包围(wrap)在分析模块中,分析模块可以被标记、分布并用于生成工厂控制或监控设置内的预测、专门的可视化、报警和警报。也就是说,分析模块可以组合或使用一个或多个数据管道,以使用数据管道的输出或由数据管道产生的暴露数据来对工厂数据进行分析。然后,分析块的输出可用于执行预测,监控故障并实时在线提供建议(例如,随着工厂内事物的演变而改变,向操作员和维护人员提供建议)。如将理解的,数据管道可以用于以任何期望的方式收集、储存、处理和组合时间序列、资产、机器健康数据、监控数据等,以产生任何期望或有用的输出。

[0049] 作为示例,数据管道架构可以在诸如图1所例示的工厂网络10的工厂网络中使用。图1的工厂网络10被例示为包括控制系统12,控制系统12具有控制器14,控制器14经由输入/输出设备16耦合到各种现场设备18,其可以是阀、传感器、变送器或测量过程、设备或工厂参数和/或在工厂内执行一些物理控制动作的任何其它类型的设备。虽然现场设备18在图1中被例示为有线现场设备,但是它们也可以或替代地是无线设备,并且可以以任何期望的方式配置以执行对设备资产、系统、流体流动等的控制。此外,工厂网络10包括监控设备或其它现场设备或测量设备20,其被例示为经由无线网关设备22连接在工厂网络10内的无线设备。设备20可以是任何类型的设备(诸如维护设备、监控设备、传感器等)。此外,设备20可以是无线或有线设备,并且可以使用任何期望的通信协议和/或通信网络在工厂网络10内进行通信。此外,图1例示了耦合到工厂网络10的具有频谱分析仪24形式的数据处理设备。这里,频谱分析仪24可以周期性地对工厂的样本或提供给其的其它数据进行分析,以提供或生成关于工厂操作的某些方面的信息或数据。当然,频谱分析仪24仅是测试或分析设备的示例,该测试或分析设备可以在工厂网络10中使用以提供关于工厂、控制系统、监控系统等的操作的信息或数据。此外,应当理解,任何其它类型的数据提供者可以耦合到工厂网络10,这包括提供离线实验室测量或分析的设备、其它类型的分析仪,诸如任何类型的信号分析仪。

[0050] 如图1所例示,工厂网络10还包括各种级别的分析结构,其包括一个或多个嵌入式分析层26(本文中也称为区域大数据节点)、集中式分析层28和基于云端的分析层30。区域、集中式和基于云端的分析层26、28、30中的每一个包括一个或多个分析引擎或信号处理引擎,其可以在工厂网络10中的这些层处的一个或多个处理器、计算机、服务器等上执行。如图1所例示,区域嵌入式分析层或块26可以包括一个或多个分析引擎36a-36n,其可以嵌入在工厂10中的设备内更接近(以逻辑或通信连接方式)控制系统12和监控系统设备(例如,20、22、24等)。虽然分析引擎36a-36n可以被实现在控制器、现场设备、网关设备等中时,但是这些引擎也可以或替代地在独立处理设备中实现(例如,实现大数据机器)。分析引擎36a-36n可以经由流式协议总线37连接在一起,并且可以经由可以使用相同或不同通信协议的各种不同总线或通信网络连接到控制系统12的控制设备14、16、18中的各个控制设备,连接到监控设备20、24或连接到诸如的频谱分析仪24的系统内的其它设备。在一些情况下,嵌入式分析引擎36a-36n可以嵌入在控制和监控设备或其它设备内,或者可以是执行数据分析的设备内的独立设备或芯片。

[0051] 如图1所例示,嵌入式分析引擎36a-36n连接在一起,并且经由分析总线37连接到一个或多个集中式分析层或块28中的分析平台40上。集中式层或集中式大数据节点28中的分析平台40可以是具有大数据设备的大数据平台(如本文将更详细描述)。分析引擎36a-

36n和大数据平台40可以储存来自图1的工厂网络10的较低级别的控制、监控和分析机器中的设备的数据。另外,集中式层28的大数据设备可以连接到基于云端的分析块30内的数据分析。虽然在图1中未例示,但是基于云端的分析块30可以包括任何数量的分析设备和大数据设备,其通过图1的其它级别或块对提供给它或其可用的数据执行数据分析。

[0052] 此外,如图1所例示,一个或多个手持设备42可以无线地连接到任何分析机器(例如分析块26、28和30中的任何设备),以执行数据分析。因此,例如,一个或多个手持设备42可以与分析引擎36a-36n相连接,与大数据设备40相连接和/或与云端分析块30中的服务器或其它机器相连接,以使用数据管道来执行数据分析。尽管在图1中没有明确例示,但是手持设备42可以连接到其它设备(诸如现场设备18、监控设备20、控制器14、I/O设备16、网关22、数据处理设备24等)以获取数据。

[0053] 应当理解,图1的较低层例示了从几个数据源(例如,I/O设备、I/O总线、无线和有现场设备、近红外光谱设备和设备监控系统)收集数据。尽管一部分数据可以被发送到传统控制器14,但大部分(bulk)数据被流式传输到网络10的区域大数据节点(即,嵌入式分析层26)中的嵌入式分析引擎36a-36n。嵌入式分析引擎36a-36n缓存(cache)数据并实时开始本地处理,并且嵌入式分析引擎36a-36n可以包括学习和运行时间执行能力。因为可能有许多类型的分析、故障检测和建议系统,所以分析引擎36a-36n可能能够执行范围广泛的算法。作为该功能的一部分,分析引擎36a-36n可以包括专门的信号处理和学习芯片(例如,TI信号处理和IBM神经学习芯片)。可以利用硬件和软件的这种组合来执行预测、故障检测和提供建议。作为该系统的一部分,嵌入式分析可以开始确定其监控设备的主要指标(leading indicators),然后与其它嵌入式分析引擎一起确定对这些主要指标影响最大的参数。一旦经过训练,嵌入式分析模块36a-36n然后可以被下载到然后其将被用于监控过程、资产和设备的地方。一个独特的特征是分析引擎36a-36n可以以数据管道所定义的方式连续学习,并且这种学习可以被缓存以供以后在任何大数据设备或节点处考虑。

[0054] 此外,如图1所例示,可以在手持设备42中执行或运行数据分析。实际上,随着具有内置分析能力的新芯片组的发展,手持单元42可以是实际运行分析的最佳场所中的一个场所。考虑到这个因素,手持设备42可以从诸如控制器14、现场设备18、监控设备20之类的下级设备,以及从诸如嵌入式分析引擎36a-36n、集中式大数据服务器40和云端30之类的上级设备拉取数据,以执行专门的处理。在这种情况下,可以针对特定装备、设备、过程、公司等开发专门的小应用程序(applet),并且这些小应用程序可以在手持设备42上执行。

[0055] 如上所述,例如,发送到嵌入式分析引擎36a-36n的策略可以包括一起链接到数据管道中的一个或多个分析块。例如,数据管道中的第一块可以用于滤波某些数据(例如,控制器数据、现场设备数据、测量数据、监控数据、设备数据等),数据管道中的第二块可以用于处理经滤波的数据集内丢失的数据(例如,找到丢失的数据,基于插值或其它技术填充数据等),数据管道中的第三个块可以用于对来自第二块的数据进行时间对齐,数据管道中的第四块可以用于从经对齐的数据的阵列或矩阵中选择参数,数据管道中的第五块可以用于从所选择的参数执行故障检测,并且在数据管道的最终块可以用于对故障检测数据执行预测。例如,每个数据管道可以由包括报警/警报的模块托管,并且可以手动调节报警限制。

[0056] 当然,每个数据管道可以包括可以用于对在工厂10内收集的数据执行处理的各种不同的数据处理块或各种不同类型的数据处理块。例如,数据管道内的块或模块可以包括

或执行1) 数据清除 (以去除噪声和不一致的数据), 2) 数据集成 (其中可以组合多个数据源), 3) 数据选择 (其中从数据库获取与分析任务相关的数据), 4) 数据转换 (其中数据被转换和合并为下游算法或块可用的形式), 5) 数据挖掘 (应用智能方法提取数据内的模式的基本过程), 6) 模式评估 (以基于兴趣度测量来识别代表知识的真正有兴趣的模式) 等。此外, 知识呈现和预测可以由一个或多个数据管道或数据管道内的块生成。在某些情况下, 例如, 数据管道可以包括呈现发现的可视化技术, 并且可以包括可以用于在控制或监控中提供预测的在线模型。此外, 在某些情况下, 预测也可能被束缚在 (tied back into) 控制策略中, 从而基于数据管道分析来关闭控制回路。

[0057] 仅作为示例, 图2A-2C例示了可以在诸如图1的工厂网络的工厂网络10中实现的各种不同的数据管道。具体地, 图2A例示了具有通信耦合在一起的五个数据分析块的数据管道50 (如箭头所例示), 以在单个数据流上执行不同的数据处理任务。在图2A的示例性数据管道50中, 第一数据处理块52收集特定类型的过程数据 (例如监控或阀数据), 第二块54滤波由块52收集的数据, 第三数据处理块56通过填充丢失的数据或处理经滤波的数据以消除异常值等来执行数据平滑。块56被例示为提供或使某些计算的数据可见或暴露以供其它块 (作为不同数据管道或分析模块的一部分, 其可以连接到或获得暴露的数据) 使用。在块56的情况下, 提供经滤波和平滑的数据的平均值 (均值) 和标准偏差 (SD) 作为暴露的参数。在任何情况下, 数据管道50的第四块58对来自块56的经滤波和平滑的数据执行傅里叶变换, 并且可以在其输出处提供数据的频率表示。然后, 数据管道50中的块59对块58的输出处的频率数据进行分析, 以确定该信号的一个或多个主频率 (dominant frequency)。主频率数据可以是数据管道50的最终输出, 并且可以储存在数据库中或暴露以供其它分析块使用。

[0058] 以类似的方式, 图2B例示了具有各种数据处理块的数据管道60, 该各种数据处理块会聚数据管道60内的数据流。在该示例中, 数据管道60具有三个数据收集块62A-62C、滤波块64A-64C和数据处理块65。数据收集块62A-62C被连接以从各种不同的设备或数据源接收不同的数据流, 滤波块64A-64C滤波来自块62A-62C的数据流, 并且数据处理块65可以处理来自滤波块64A中的一个滤波块的数据以去除异常值。块66然后对来自块64B和64C和65的数据进行时间对齐, 以产生时间对齐数据的矩阵, 而块68选择由块66产生的数据矩阵内的数据以执行进一步的处理。滤波块64A、64B和64C被例示为关于经滤波的数据的暴露参数 (诸如经滤波或输入的数据的任何统计测量、经滤波的系数等)。当然, 其它数据可以由图2B的数据管道60中的这些或任何其它块计算和暴露其它数据。

[0059] 另外, 图2C例示了包括对发散数据流进行操作的块的数据管道70。在这种情况下, 第一数据处理块72 (例如, 数据收集模块) 收集特定类型的过程数据 (例如, 监控或阀数据), 第二块74 (例如, 中间过程模块) 滤波由块72收集的数据, 以及第三数据处理块76 (例如, 另一个中间过程模块) 通过填充丢失的数据或处理经滤波的数据以消除异常值等来平滑经滤波的数据。此后, 块78A-78D (也是中间过程模块) 对来自块76的经滤波和平滑的数据以不同的采样率进行傅里叶变换, 并且块79A-79E对来自傅立叶变换块78A-78D的数据进行滤波。在这种情况下, 块79D和79E可以使用不同的滤波系数 (例如, 通过范围) 滤波来自块78D的相同的数据, 以产生不同的滤波数据集。在任何情况下, 来自块79A-79E的数据可以通过用于其它分析的块80 (其可以是最终用途过程模块) 储存在数据库中, 可以通过块82 (其可以是最终用途过程模块) 等作为图形或图表呈现给用户。

[0060] 应当理解,图2A-2C的各种数据管道中的每个块可以在图1的工厂网络10内的不同的或相同的设备、处理器等中被储存和执行。作为示例,图2B的数据管道60在图1中被例示为具有设置在工厂网络10内的不同设备中的各种数据处理块。在这种情况下,数据收集块62A-62C被设置在图1的现场设备18和20的各个设备中,滤波块64A-64C在I/O设备16和网关设备22中被储存并执行。平滑器块65在分析引擎36b中的一个分析引擎中被储存并执行,而时间对齐块66在分析引擎36a中的另一个分析引擎被储存并执行。选择块68在大数据分析平台40中被储存并执行。当然,数据管道60的每个单独的块在耦合在执行这些块的设备之间的不同通信网络上通信地连接,并且数据管道60的块被配置为使用针对该网络的适当通信或经由网络间通信来彼此通信(由数据管道60定义)。因此,可以在使用不同通信网络和协议并且仍然无缝地操作的设备之间分布特定数据管道。当然,由数据管道60的任何块暴露或输出的数据可以被提供给工厂网络10中的任何其它设备或应用,或者由工厂网络10中的任何其它设备访问,这些设备包括手持设备42、云端30等等。

[0061] 作为另一示例,图2A的整个数据管道50被例示为设置在分析引擎36n中并在分析引擎36n中执行。在另一示例中,图2C的数据管道70的前三个块被例示为在分析引擎36b中执行,而图2C的数据管道70的其余数据处理块的大部分被例示为在具有图表块80的大数据平台40中被执行,图表块80被例示为在手持设备42中运行。

[0062] 当然,图2C的数据管道的各种块的配置和位置仅是示例性的,并且数据管道的任何特定块或模块可以基于例如,设备的处理能力、设备的数据储存能力、设备与数据源的接近度等被储存在并执行在工厂网络10内的任何适当的设备中。此外,相同数据管道的不同数据处理模块可以被储存在不同的网络中(诸如在过程控制网络(例如,与图1的现场设备18和/或20以及控制器14相关联的任何一个过程控制网络)中和/或工厂通信网络(诸如图1的通信网络26、27、28中的任何一个通信网络))。需要指出的是,一旦创建并安装在工厂网络10中的各种设备中,数据管道就对提供给其的数据进行操作,以产生可以储存或提供给各种分析模块、例程、其它数据管道等中的任何一个的新数据,该各种分析模块、例程、其它数据管道等可以使用该数据来运行模型、执行较高级别的分析、向用户呈现数据等。在一些情况下,数据管道本身可以是分析模块,其通常分布在工厂10中以已知或预定的方式来处理数据流,以使得新数据或新类型的数据可供其它分析模块使用或处理或由用户消费。当然,如果需要,数据管道可以用于实现整个分析模块(从开始到完成),或者可以并入或用作分析模块的一部分。因此,一旦安装在工厂中,特定数据管道就可能被多于一个分析模块用于在同一数据上执行不同类型的分析或使用相同的数据来执行不同类型的分析。

[0063] 此外,应当理解,图2A-2C的数据管道仅仅是示例,并且可以设计和使用限定以任何特定顺序通信地耦合在一起的任何一组处理块的任何其它类型的数据管道。此外,数据管道可以具有会聚或发散的流或两者,并且可能比图2A-2C所例示的简单数据管道复杂得多,具有任何数量的发散和/或会聚分支、以任何方式连接的任何数量的模块等。

[0064] 如将理解的,在监控和分析系统中使用数据管道作为基本数据处理架构概念能够实现对于分析模块所使用的实时或时间段中的数据的高效处理。这样的数据管道包括的一组算法,其已经以特定顺序布置并被配置以便开发可以由其它分析模块使用的模型。所得到的模型可以在运行时间环境中下载并使用以执行预测、故障检测和建议。可能包含在数据管道中的模块包括例如数据选择和/或数据清除模块、数据对齐模块、敏感性分析模块、因

果关系分析模块、监督学习模块和无监督学习模块。

[0065] 此外,数据管道或任何数据管道内的任何处理模块可以以任何期望的通信和数据处理速率执行数据收集和处理。作为示例,数据管道或其处理模块可以以比设置在生成原始数据的过程控制网络内的过程控制器的执行速率更快、相同或更低的数据速率收集原始过程数据。一般而言,通常期望以比可以使用这样的数据来执行工厂内的过程控制活动的过程控制器的执行更快的数据速率收集原始过程数据。此外,为了实时处理这样的数据,通常期望将数据管道中的至少一些数据处理模块配置为以比过程控制器的执行速率更快的执行速率操作,并且可能以比过程控制器的执行速率更快的速率在数据管道的不同处理模块之间发送数据。这种更高的数据速率使数据管道能够以比控制器执行速率更快的数据速率处理和分析原始数据,这通常能够实现更好或更完整的过程分析。

[0066] 在更一般的意义上,图3描绘了可以连接在一起以在工厂网络内形成数据管道的各种类型的数据处理块或功能的框图85。具体地,块86表示工厂内的一个或多个数据源,其包括结构化和非结构化数据,诸如过程数据、谱数据、振动数据、图像数据等。由块86表示的数据源可以是任何数据源(诸如测量或原始数据、经处理的数据、控制数据、监控数据、维护数据、设备数据、由其它数据管道或其它分析模块生成的数据等的来源)。该数据可以从在线来源,从数据库或历史库或从外部来源(例如,互联网、工厂网络的外部来源等)获得。

[0067] 图85还包括表示可以由数据管道中的一个或多个数据处理块执行的各种类型的处理的块,其包括数据选择和清除块87、数据对齐块88、敏感性分析块89、因果关系分析块90、监督学习块91和无监督学习块92。块87-92包括可以或将在数据管道中执行的主要类型的数据处理功能,并且每个数据管道将典型地包括由以某种方式通信地连接在一起的块87-92表示的类型的一个或多个处理块。如图3中的箭头所例示,数据管道中的块87-92将从块86获得数据(例如,一些指定的数据源),并且可以以任何顺序或方式连接在一起以执行各种数据处理功能。

[0068] 此外,如图3中的呈现块93和探索性分析块94所例示,从块87-92中的一个或多个的组合创建的数据管道可以在工厂网络中向各种用户提供输出或处理的数据。呈现块93例如提供了由数据管道产生的数据的某种类型的呈现。这样的呈现可以是例如,图形、图表、视图框、动画等,其提供由数据管道的一个或多个块产生的数据的一些表示、说明、描绘或其它查看(例如,从因果关系块90、敏感性块89、监督学习块91或无监督学习块等)。同样地,探索性分析块94可以向用户提供到数据管道中的接口以进行改变、添加新的或更改数据源、调试数据管道等等,全是为了力图创建新的或更有用的数据管道,以在数据管道中尝试新的数据处理或新数据源等。

[0069] 当然,有可以用于执行块87-92中的任何块中的数据处理的许多类型的算法或技术,而下文仅讨论这些技术的几个示例。具体地,块87的数据选择和清除功能可以使用各种技术来实现。例如,这些块可以选择或指定要清除的参数或数据,以及使用任何期望的技术清除或使用的数据的时间段。这些选择可以由块创建者指定的绝对的或固定的,或者可以基于一些正在进行的标准或计算来确定。例如,当指定的参数显示高可变性(在线或在数据库中储存的历史数据)时,数据选择可以自动启动。数据选择块87可以探索关注参数的可变性高的时间段,并且如果这样的数据可用,可以基于过程响应时间例如,将该时间段延长到过去和未来。在另一种情况下,数据选择块87可以在定义的时间段内检查下游和上游参

数的可变性,并且可以选择超过最小变化的参数(假设上游/下游参数在配置或其它情况下是预定义的)。如果没有定义工厂内的上游/下游路径,则数据选择块87可以选择处理单元的参数,在选定的时间段内开发与数据的主成分分析(PCA),可以计算所有参数的振荡显著性指数(OSI),然后可以基于该指数值选择参数。当然,也可以使用选择数据的许多其它方式。

[0070] 在另一种情况下,数据选择和清除块87可以检测丢失或不良(bad)的数据,并对该数据执行数据校正或排除。例如,块87可以给出该变量的最后已知的良好值的不良状态数据值(例如,在线)。在另一种情况下,块87可以使用插值技术来填充缺失的数据。例如,如果数据的最后一个良好部分的数据趋势的斜率与缺失数据期间之后的数据部分的斜率大致相同,则可以根据斜率填充缺失的参数值。该规则可以被应用到等于过程响应时间的数据长度,或者如果过程响应时间是未知的,则可以被应用到预定数量的样本(例如10)。如果丢失数据之前和之后的数据斜率不相同,则可以应用数据的变量平均值以生成丢失的数据值。同样地,如果缺少大量的数据值(例如是可用数据的两倍),或者如果在同一时间段内有几个变量是不良的,则可以从数据集中去掉数据不良的时间段。可以使用用于排除或分离用于多状态模型的数据或用于执行数据切片的任何已知的过程。在这种情况下,可以在定义相对于前导变量的参数延迟之后执行数据切片。对于无监督学习,可以应用简单的数据切割。此外,如果在选定期间超过20%的值是不良的,则该变量可能被完全排除在建模/学习之外,除非这个变量是在分析中绝对需要的关键变量。

[0071] 在另一种情况下,块87可以通过应用例如 3σ (标准偏差规则)或应用汉默尔(Hammel)分析(即,中间绝对偏差规则),或应用基于模型的规则(例如,AR/ARIMA)或应用基于卡尔曼(Kalman)的异常值检测规则(仅举几例)来执行异常值消除。

[0072] 另外,数据选择和清除块87可以执行数据滤波。例如,块87可以仅在特定应用需要滤波时应用滤波,或者可以提供任何期望或已知的滤波技术(诸如Savitzky-Golay滤波、指数滤波、平均滤波、可选的小波滤波等)的选择。

[0073] 以类似的方式,块88可以使用任何期望的技术来执行数据对齐。例如,可以通过选择用于监督学习的下游质量参数或者具有用于无监督学习的显著变化的下游前导(leading)参数的过程来执行数据对齐。数据对齐可以逐个应用每个上游参数的互相关算法来定义上游参数的时间延迟,或者可以针对每个上游参数应用PCA技术来定义时间延迟。块88可以基于时间延迟来开发对齐数据的新文件,并且当在实验室中确定质量参数时,可以仅在改变之后使用第一值。

[0074] 在另一种情况下,块89可以使用例如时间对齐文件(file)(诸如产生块88)来执行敏感性分析并且开发偏最小二乘(PLS)模型。块89可以通过实验式地使用每个顺序输入上的阶跃变化来确定每个模型输入的增益(以确定变量对其它变量的变化的敏感性)。块89可以从模型配置排除具有小于最小增益的输入,并且可以可选择地应用PCA并开发用于输入参数的OSI(振荡显著性指数)。

[0075] 同样地,块90可以使用非对齐和/或对齐的文件来执行因果关系分析,并且可以从这些文件中开发PCA模型。块90还可以定义用于输入参数的OSI指数(显著性指数),并且可以从模型配置排除具有小于所需最小指数值的输入。

[0076] 另外,块91可以使用任何期望的非线性模型(诸如神经网络模型)或诸如PLS或MLR

模型的线性模型来执行监督学习。块92可以通过应用诸如PCA和SVD建模的分析建模,通过执行数据聚类、K均值、k-NN、决策树、装袋(bag)、提升(boost)、随机森林、条件贝叶斯概率分析等来执行无监督学习。当然,这些只是在数据管道中可能发生或可能在块87-92中的每个块中实现的处理类型的几个示例,并且也可以或替代地应用许多其它类型的数据处理。

[0077] 作为使用数据管道的另一示例,图4描绘了用于控制一个或多个过程并且支持过程控制大数据的过程工厂或过程控制系统的示例性大数据网络100的简化框图,并且特别地,其使用一个或多个数据管道来实现基于自动或自主的信号处理的学习的一个或多个实例。过程控制大数据网络100包括一个或多个过程控制大数据节点102-110,每个过程控制大数据节点102-110对过程控制大数据进行收集、观测、生成、储存、分析、访问、发送、接收和/或操作。本文中可互换使用的术语“过程控制大数据”、“过程大数据”和“大数据”通常是指包含在过程控制系统或工厂中并与其相关的设备生成、接收和/或观测的所有(或几乎所有)数据。在一些情况下,由包含在过程工厂中并与过程工厂中相关联的所有设备生成、创建、在其处接收或以其它方式观测的所有数据被收集并储存为过程控制大数据网络100内的大数据。

[0078] 示例性过程控制大数据网络100包括一个或多个不同类型的过程控制大数据节点或设备102-110,每个大数据节点或设备102-110对过程控制大数据进行收集、观测、生成、储存、分析、访问、发送、接收和/或操作,该过程大数据依据或基于由过程工厂或过程控制系统对一个或多个过程的处理的控制而生成。每个过程控制大数据节点或设备102-110连接到过程控制系统大数据网络主干(未示出),并且可以使用主干与一个或多个其它过程控制大数据节点进行通信。因此,过程控制大数据网络100包括过程控制系统大数据网络骨干和与其通信连接的过程控制大数据节点102-110。在一示例中,过程控制大数据网络100包括多个联网的计算设备或交换机,其被配置为经由骨干将分组路由到网络100的各种其它设备、交换机或节点,或将路由来自网络100的各种其它设备、交换机或节点的分组。如将要更详细地描述的,可以在大数据网络100的各种设备、节点等内实现一个或多个数据管道,以对大数据网络100的各个节点102-110内生成的,收集的,或以其它方式在大数据网络100的各个节点102-110处获得的数据执行数据处理以及监督或无监督学习。

[0079] 过程控制大数据网络主干可以包括任何数量的有线通信链路和任何数量的无线通信链路,其支持一个或多个适当的例如,包括在互联网协议(IP)套件中的协议(例如,UDP(用户数据报协议)、TCP(传输控制协议)、以太网等)的路由协议,或其它适当的路由协议。主干可以支持诸如流控制传输协议(SCTP)的流协议和/或另一适当的流协议,以在过程控制大数据节点之间流式传输(例如,传输)数据。例如,题为“STREAMING DATA FOR ANALYTICS IN PROCESS CONTROL SYSTEMS”的美国专利申请No.14/506,863描述了用于过程控制大数据的流协议和技术的示例,其中的一个或多个可以被网络100中的过程控制大数据网络主干使用。通常,包括在过程数据大数据网络100中的每个节点可以支持至少由主干支持的路由协议的应用层(以及一些节点、附加层)。在一些情况下,每个过程控制大数据节点102-110在过程控制系统大数据网络100内被唯一地识别(例如通过唯一的网络地址)。

[0080] 过程控制系统大数据网络100的至少一部分可以是自组织网络。因此,节点102-110(和/或一个或多个诸如用户接口设备130的其它节点)中的至少一些节点可以以特别的方式连接到网络主干(或网络100的另一节点)。

[0081] 因为图4是描绘过程控制大数据网络100中的各个大数据节点102-110之间的通信连接的简化图,所以过程控制网络主干在图4中未明确例示。然而,上述美国专利申请No.13/784,041中描述了可以与本文所述的任何或所有技术一起使用的这种骨干的示例。当然,本文描述的技术中的任何或全部技术不限于与美国专利申请No.13/784,041中描述的主干一起使用,而是可以与任何适当的通信网络主干一起使用。

[0082] 现在转向不同类型的过程控制大数据节点或设备102-110,通常,网络100的过程控制大数据节点可以是“大数据提供者”和/或可以包括“大数据设备”,如下文所讨论。

[0083] 在本文中可互换使用的术语“大数据提供者”、“大数据提供者节点”或“提供者节点”通常是指使用过程控制大数据网络100来收集、生成、观测和/或转发过程控制相关的大数据的过程控制大数据节点。由提供者节点生成、收集、观测和/或转发的过程控制大数据可以包括已经在工厂内的控制或监控过程中直接使用,或从工厂内的控制或监控过程生成的数据(例如,一阶实时),以及由诸如控制器、输入/输出(I/O)设备的过程控制设备和诸如变送器、传感器等的现场设备生成或使用的配置数据。另外或者替代地,过程控制大数据提供者节点可以生成、收集、观测和/或转发与在过程工厂内传送和路由这样的一阶过程控制数据和其它数据有关的数据,(例如,与大数据网络100的网络控制和/或工厂中其它通信网络的网络控制相关的数据、指示带宽的数据、网络访问尝试、诊断数据、监控数据等)。此外,一些过程控制大数据提供者节点可以通过分析过程控制大数据网络100已经收集的过程控制数据、监控数据、维护数据等来生成、收集、观测和/或转发指示过程控制大数据网络100中学习到的结果、学习和/或信息的数据。通常,这样的分析结果、学习和/或学习到的信息是通过一个或多个过程控制大数据节点使用如本文所描述的一个或多个数据管道而执行的自动、自主分析来生成的。

[0084] 在一些情况下,大数据提供者节点可以包括用于实时(例如,流式地)发送和接收大数据的多核硬件(例如,多核处理器),并且在一些实施例,用于缓存实时大数据准备在过程控制大数据网络100上流式传输或其它传送。在一些实施例中,大数据提供者节点还可以包括用于缓存实时大数据的高密度存储器。可以由大数据提供者节点发送、接收、流式传输,缓存、收集和/或以其它方式观测的实时数据的示例可以包括诸如测量数据、配置数据、批量数据、事件数据、监控数据、维护数据(例如由维护设备收集或生成的维护数据、算法、程序等)和/或连续数据的过程控制数据。例如,与配置、批量接收、设定点、输出、速率、控制动作、诊断、报警、事件和/或其之改变相对应的实时数据可以被收集。实时数据的其它示例可以包括过程模型、统计、状态数据以及网络和工厂管理数据。在一些实施例中,大数据提供者节点并不对其观测到的实时大数据中的至少一些进行缓存,而是随着该节点处观测、接收或者产生该数据,将未缓存的数据流式传输到一个或更多个其它大数据节点。可以在前述美国专利申请No.13/784,041、14/174,413和14/212,493中,找到可以结合本文所描述的技术中的任何一种或者全部来使用的大数据提供者节点的示例。当然,替代地或者另外地,也可以结合不同于美国申请No.13/784,041、14/174,413和14/212,493中所描述的那些的大数据提供者节点来使用本文所描述的技术中的任何一种或者全部。

[0085] 在本文中可互换使用的术语“大数据设备”、“大数据设备节点”或“设备节点”通常是指接收、储存、获取和分析过程控制大数据的过程控制大数据节点。因此,过程控制大数据设备(或“BDA”)通常对由一个或多个过程控制大数据提供者节点生成或提供的大数据进

行操作。在一些情况下,大数据设备包含在大数据提供者节点中,或者与大数据提供者整体共同驻留在同一个节点或设备内。在这种情况下,大数据设备被称为“嵌入式大数据设备”,这是由于该设备嵌入在提供者节点或设备中,并且对由共同驻留的大数据提供者所接收、收集或生成的大数据进行操作。在一个示例中,嵌入式大数据设备对该嵌入式大数据设备所位于的大数据提供者节点本地生成和/或提供的大数据进行分析,以发现或学习知识。这种所学到的知识可以储存在该嵌入式大数据设备中,由该嵌入式大数据设备进行在本地操作,和/或作为大数据提供给其它大数据节点。本文所描述的技术中的任何一种或全部可以结合诸如前述美国专利申请No.14/212,493和/或题为“REGIONAL BIG DATA IN PROCESS CONTROL SYSTEMS”的美国专利申请No.14/507,188中描述的嵌入式大数据设备来使用,例如,尽管可以另外或替代地使用其它适当的嵌入式大数据设备。此外,应注意,在大数据提供者节点包括嵌入式大数据设备的实施例,大数据提供者节点的缓存可以在尺寸上被减小或者省略,因为嵌入式大数据设备提供本地数据储存能力。

[0086] 在一些情况下,大数据设备可以是过程控制大数据网络100的独立大数据节点。也就是说,在这些情况下,大数据设备不嵌入大数据提供者节点或与大数据提供者节点共同驻留。因此,包括大数据设备的过程控制大数据节点可能不一定自身是大数据的提供者。

[0087] 图5示出了示例性大数据设备214的简化框图,其实例可包括在图4的过程控制大数据网络100的任何节点102-110或图1的类似节点中。示例性大数据设备214包括用于历史化或存储接收的大数据的大数据储存区220、一个或多个大数据设备接收器222和一个或多个大数据设备请求服务器(servicer) 224。每个大数据设备接收器222被配置为接收大数据分组(其可以从另一个节点流式传输和/或可以由设备214所在的大数据提供者节点生成),处理数据分组以获取其中携带的实质性数据和时间戳,并将实质性数据和时间戳存储在设备214的大数据储存区220中,例如作为时间序列数据,并且还可任选地作为元数据存储。例如,数据点可被标记并存储为元数据。大数据储存区220可以包括多个本地和/或远程物理数据驱动器或储存实体,诸如RAID(独立磁盘冗余阵列) 储存器、固态储存器、云端储存器、高密度数据储存器和/或适用于数据库或数据中心储存的任何其它合适的数据储存技术,并且对于其它节点具有单个或单一的逻辑数据储存区域或实体的外观。此外,每个大数据设备请求服务器224被配置为例如按照请求实体或应用的请求来访问存储在大数据设备储存区220中的时间序列数据和/或元数据。

[0088] 在一些实例中,大数据设备214包括一个或多个大数据分析器226,在不使用任何用户输入以启动和/或执行学习分析的情况下,通常以自动和/或自主的方式对存储的大数据的至少一部分执行相应的数据分析和/或学习。大数据分析器226可以使用一个或多个数据管道单独地和/或共同地对存储的数据执行大规模数据分析(例如,数据挖掘、数据发现等),以发现、检测或学习新的信息或知识。例如,数据挖掘通常涉及检查大量数据以提取新的或先前未知的有趣数据或模式(例如异常记录或多组数据记录)的过程。大数据分析器226可以另外或可替代地对所存储的数据执行大规模数据分析(例如,机器学习分析、数据建模、模式识别、预测分析、相关性分析等)以预测、计算或识别存储数据内的隐含关系或推论。在一些情况下,多个大数据分析器226(和/或至少一个大数据分析器226的多个实例)并行和/或协作地操作以分析存储在设备214的大数据储存区220中的数据。在上述的题为“REGIONAL BIG DATA IN PROCESS CONTROL SYSTEMS”的美国专利申请No.14/507,188中可

以找到可以与本文所描述的任何或所有技术一起使用的协同数据分析的示例,但任何合适的一种或多种协同数据分析技术都可以与本公开内容的任何或所有方面一起使用。

[0089] 通常,大数据设备接收器222、大数据设备请求服务器224和大数据分析器226中的每一个包括存储在一个或多个非暂时实体存储器或数据储存设备上的各自的计算机可执行指令,并且可由一个或多个处理器执行以执行一个或多个它们各自的大数据功能。在一些情况下,大数据分析器226不包括在大数据设备214中,而是与大数据设备214通信连接。例如,包括储存区220、接收器222和服务器224的大数据设备214可以通过第一计算机可执行指令集来实施,大数据分析器226可以通过第二计算机可执行指令集(其可以或不存储在与第一计算机可执行指令集相同的非暂时的实体存储器或数据储存设备上)来实施。可以在上述美国专利申请No.13/784,041、No.14/174,413和No.14/212,493中找到可以与本文描述的任何或所有技术一起使用的各种类型的示例性大数据设备及其组件的说明,但可以理解,本文描述的任何或所有技术可以与其它合适的大数据设备一起使用。

[0090] 再次参考图4,过程控制大数据网络100可以包括在相对于由诸如控制器、I/O设备、现场设备等的过程控制设备直接生成、路由和/或使用的的第一级或主过程相关数据的各种级别、层或级操作的过程控制大数据提供者节点102-110。最低级、层或级别,“本地”大数据提供者节点或设备102a-102n在最靠近过程处操作,用以收集、生成、观察和/或转发与过程工厂中的过程设备和装置的输入、操作和输出相关的主过程大数据。因此,“本地大数据提供者节点或设备”102a-102n通常是生成、路由和/或接收主过程控制数据或监控数据(来自过程监控系统)的节点和/或设备,以实现在过程工厂中实时控制或监控一个或多个过程。本地大数据提供者节点102a-102n的示例包括其主要功能被指示为生成和/或操作过程控制数据以控制过程(例如有线和无线现场设备、控制器和I/O设备)的设备。这些过程控制设备可以以分布式的方式通信地彼此连接和/或连接到一个或多个过程控制通信网络。例如,一个或多个现场设备通信地连接到一个或多个I/O设备,I/O设备又通信地连接到一个或多个控制器,控制器又可通信地耦合到一个或多个过程控制通信网络(例如,HART®、WirelessHART®、过程控制大数据、FOUNDATION®Fieldbus等)。

[0091] 本地大数据提供者节点102a-102n的其它示例包括其主要功能是通过过程控制系统的一个或多个通信网络(其可以包括过程控制大数据网络100和/或其它通信网络)提供对主过程数据的访问或路由的设备。这种类型的本地大数据提供者节点102a-102n的示例包括接入点、路由器、到有线控制总线的接口、到无线通信网络的网关、到外部网络或系统的网关、以及其它此类路由和联网设备。本地大数据提供者节点102a-102n的其它示例包括被配置为临时存储整个过程控制系统的大数据的设备(例如,作为溢出缓存、路站或路由队列),诸如历史库设备。

[0092] 在一些情况下,本地大数据提供者节点包括相应的本地大数据设备,如图4的节点102a、102n所示,它们分别包括嵌入式大数据设备112a、112n。每个本地嵌入式大数据设备112a、112n接收并存储由其各自提供者102a、……、102n提供的相应本地大数据。此外,在一些本地大数据提供者节点中,例如在节点102a中,可以对存储在设备112a中的至少一些本地大数据执行一个或多个分析功能、例程、操作或进程(由带圆圈的A₁表示)。分析A₁可以由图5的一个或多个大数据分析器226执行,分析A₁可以用一个或多个数据管道执行或实施。

一个或多个分析 A_1 的学习到的信息、学习和/或结果也可以存储在本地大数据设备112a中,并且学习到的信息或结果中的至少一些可以被提供给另一个大数据节点106a。

[0093] 例如,由节点102n所示的一些本地提供者节点包括用于本地大数据收集和历史库的相应的本地嵌入式大数据设备112n,然而,驻留设备112n执行最少的分析或不执行分析。因此,节点102n仅将本地存储的大数据流式传输(或以其它方式(例如基于请求或在适当的时间)传送)到另一节点106b,例如,用于分析处理或进一步转发。一些本地大数据节点(例如,节点102b)根本不包括任何大数据设备。这样的节点102b可以实时地或者借助于缓存将本地观察到的大数据流式传输到一个或多个其它大数据节点102a、106b。

[0094] 诸如过程相关数据、工厂相关数据和其它类型的数据之类的各种类型的实时数据可以由大数据提供者节点或设备102a-102n作为大数据被缓存、收集、存储、传送和/或流式传输。过程相关数据的示例包括在过程工厂中控制过程时生成的连续、批量、测量和事件数据(并且在某些情况下,表示该过程的实时执行的效果)。此外,过程相关数据可以包括过程定义、布置或设置数据,例如配置数据和/或批量配方数据、对应于过程诊断的配置、执行和结果的数据等。

[0095] 工厂相关数据,例如与过程工厂有关但是可能不由直接配置、控制或诊断过程工厂中的过程的应用生成的数据,也可以由大数据提供者节点102a-102n作为大数据进行缓存、收集、存储、传送和/或流式传输。工厂相关数据的示例包括振动数据、疏水阀数据、指示对应于工厂安全性的参数值(例如,腐蚀数据、气体检测数据等)的数据、指示与工厂安全相对应的事件的数据、对应于机器、工厂装置和/或设备健康的数据、对应于装置、机器和/或设备诊断的配置、执行和结果的数据以及对诊断和预测有用的数据。

[0096] 此外,包括与过程控制大数据网络主干网有关的和过程工厂的各种通信网络的数据高速公路流量和网络管理数据的其它类型的数据、诸如与用户流量相关的数据的用户相关数据、登录尝试、查询和指令、文本数据(例如日志、操作过程、手册等)、空间数据(例如,基于位置的数据)和多媒体数据(例如闭路电视、视频剪辑等)可以由大数据提供者节点102a-102n作为大数据进行缓存、收集、存储、传送和/或流式传输。本文公开的数据管道可以对任何这种数据操作或使用任何这种数据。

[0097] 在一些实施例中,动态测量和控制数据可以由大数据提供者节点102a-102n作为大数据进行自动缓存、收集、存储、传送和/或流式传输。动态测量和控制数据的示例包括指定处理操作中的变化的数据、指定操作参数的变化的数据,操作参数例如设定点、过程和硬件警报以及诸如下载或通信故障的事件的记录等。此外,诸如控制器配置、批量配方、警报和事件的静态数据可以在检测到变化或控制器或其它实体最初添加到大数据网络100时默认地自动收集。当然,本文中公开的数据管道可以对任何这种数据操作或使用任何这种数据。

[0098] 此外,在某些场景中,当检测到元数据的变化时,在大数据提供者节点102a-102n中捕获至少一些描述或识别动态控制和测量数据的静态元数据。例如,如果在控制器配置中做出影响必须由控制器发送的模块或单元中的测量和控制数据的变化时,则相关元数据的更新由大数据提供者节点102a-102n自动地捕获。另外或可替换地,与用于缓冲来自外部系统或来源(例如,天气预报、公共事件、公司决策等)的数据、监控数据和/或其它类型的监控数据的特殊模块相关联的参数可以由大数据提供者节点102a-102n自动捕获。本文公开

的数据管道可以对任何这种数据操作或使用任何这种数据。

[0099] 在一些情况下,最终用户创建的附加参数被自动地捕获在大数据提供者节点102a-102n中。例如,最终用户可以在模块中创建特殊计算,或者可以向单元添加需要收集的参数,或者最终用户可能希望收集默认情况下不传送的标准控制器诊断参数。最终用户可任选地配置的参数可以以与默认参数相同的方式传送。

[0100] 再次参考图4的网络100,在本地大数据节点102a-102n上方的一个或多个等级或层级上,过程控制大数据网络100可以包括一个或多个区域大数据节点106a-106m。为了实施区域大数据,过程工厂或过程控制系统可以被视为具有可以根据任何期望的方式(例如地理、物理、功能、逻辑等)来描述的多个不同范围或区域。在说明性但是非限制性示例,过程工厂可以具有接收原材料并产生第一中间产物的第一区域,接收其它原料并产生第二中间产物的第二区域,以及接收第一中间产物和第二中间产物以产生输出产品的第三区域。这三个不同示例性区域中的每一个可以由相应的“区域”大数据节点106a、106b、106m服务,以对由其相应区域产生的大数据进行操作。因此,“区域大数据节点”为由本地大数据提供者节点102的相应归组(grouping)或区域并且在某些情况下由其它大数据提供者节点104生成和/或提供的数据提供大数据支持和服务。其它大数据提供者节点104可以包括例如在工厂区域外部的大数据节点(例如,便携式诊断设备或离线仿真器)、用户接口设备130或完全在过程工厂外部的数据源(例如,材料供应商的计算设备、提供天气预报的馈送等)。如将理解的,区域大数据节点106a-106m可对应于图1的区域层或块26中的任何一个。

[0101] 如图4所示,由区域大数据节点106a-106m服务的相应归组或区域可以包括按照一些地理、物理、功能、逻辑或其它所期望的方式相关的一个或多个大数据提供者节点102a-102n和/或其它大数据节点104。例如,区域大数据节点106a服务包括本地大数据提供者节点102a和102b的区域,区域大数据节点106b服务包括本地大数据节点102b和102n以及另一大数据节点104的区域。为了区域大数据储存、访问和/或分析的目的,包括在特定区域中的特定节点102、104可以将数据流式传输或传送到其相应的区域大数据节点106。此外,任何大数据提供者节点102a-102n和/或其它大数据节点104可以与特定区域大数据节点106a-106m通信,以请求区域可用服务和/或访问存储在其中的区域大数据和元数据,无论这样的请求节点是否包括在特定区域大数据节点106a-106m的特定区域中。此外,如本文所述的每个数据管道可以在特定区域大数据节点106a-106m内或跨各个区域大数据节点106a-106m实施。

[0102] 因此,每个区域大数据节点106a-106m包括相应的区域大数据设备116a-116m,大数据通过该区域大数据设备被接收,存储为区域大数据,并被访问或请求。此外,每个区域大数据节点106a-106m通常包括可以对至少一些区域大数据单独地和/或协作地操作的一个或多个分析功能、例程、操作或过程(例如,A₂-A_w)。例如,区域大数据设备116b可以从本地提供者节点102b、102n接收本地大数据,本地提供者节点102b、102n被配置为协作地控制通过过程工厂的一部分或区域的液体流,节点106b可以对至少一些所接收的数据执行分析过程A₄以确定过程工厂的一部分或区域内的液体的平均输送延迟。分析过程A₄可以例如连接并使用由一个或多个数据管道生成或开发的数据,或者代之以可以由一个或多个数据管道部分地或全部地实施。然后,分析A₄的结果可以作为附加区域大数据被存储或历史化在区域大数据设备116b内。在一个实施例中,分析A₂-A_w中的每一个由图5的一个或多个大数据分

析器226执行,它们存在于它们各自的大数据节点上,这些分析可以使用下载到大数据节点106a-106m的各种设备中并在各种设备实施或执行的一个或多个数据管道来实施。

[0103] 在一些情况下,区域大数据节点或设备106a-106m与另一区域大数据节点106a-106m(例如对等点)传送接收或生成的大数据、学习到的知识或信息和/或分析结果。在某些情况下,可以在多个对等节点之间实施数据管道。为了通过继续上述示例来说明,区域大数据节点116a接收由区域大数据节点106b执行的分析学分析A₄生成的学习信息。随后,区域大数据节点106a于是可以结合从在其自己的区域内的本地大数据节点102a、102b接收的本地大数据,对来自节点106b的对等提供的学习信息的至少部分执行一个或多个相应的区域分析A₂、A₃。这些分析可以使用彼此和/或与分析A₄相同或不同的数据管道来实施。分析A₂、A₃又可以使用另外的数据管道,生成用于在区域大数据设备116a的历史化和/或用于提供给其它大数据节点106b、106c、108的附加区域大数据。因此,由于区域大数据节点106a-106m可以在某些场景中(例如,基于由此执行的任何驻留分析的结果或学习)来产生区域大数据,区域大数据节点106a-106m也可以作为区域大数据提供者节点来操作。

[0104] 大数据提供者节点102a-102n在相应的区域大数据节点下的归组可以根据任何期望的方式进行,例如地理、物理、功能、逻辑等。例如,在说明性但非限制性的场景中,过程工厂中的过程可以基于两个中间产物产生特定产品。这样,本地大数据提供者节点102a可以表示产生第一中间产物的第一控制回路,本地大数据提供者节点102b可以表示产生第二中间产物的第二控制回路。因此,由两个控制回路102a、102b生成、收集、接收或以其它方式观察的所有过程控制数据可以被传送到区域大数据节点106a用于历史化、储存和分析。

[0105] 以类似的方式,区域大数据节点106b可以从其相应的大数据提供者节点组接收和分析数据。例如,区域大数据节点106b可以负责结合由其它源104提供的大数据,分析对应于基于来自每个大数据提供者节点102b、102n的中间产物的另一产物的产生的大数据。

[0106] 在区域大数据节点106a处,可以分析所接收的大数据(例如,通过使用一个或多个分析功能或过程A₂、A₃)来跨时间和/或跨各种数据集中的至少一些创建或生成描述有意义的关系、模式、相关性、趋势等的学习知识。例如,当最终产生了特定产品时,两个控制回路102a、102b中的事件的某种组合可能导致较差的产品质量。为了确定产品质量较差的根本原因,区域大数据节点106a分析在事件发生时或紧随其后(例如,当在区域大数据节点106a处接收到对应于事件发生的数据时)的事件组合生成的数据,并且可以使用为此目的而创建并在节点106a中的各种设备中执行的一个或多个数据管道来这样做。区域大数据节点106a可以生成基于这些事件的发生来预测较差产品质量的学习知识,和/或可以实时地自动调整或改变一个或多个参数以减轻事件组合的影响(在将来如果它们发生时)。例如,区域大数据节点106a可以确定修正的设定点或修正的参数值,以更好地调节和管理两个控制回路102a、102b。

[0107] 通常,每个区域大数据节点106a-106m(或其相应的大数据设备116a-116m)分析来自其相应的大数据提供者节点的组或区域的数据,以确定有意义的模式、相关性、趋势等,并且可以使用一个或多个数据管道来执行这些分析。然后将学习的模式、相关性、趋势等作为学习知识存储在相应的区域大数据设备116a-116m中。如本文所使用的,术语“学习到的知识”或“学习”通常是指作为对大数据执行一个或多个分析的结果而生成的数据、服务、功能、例程和/或应用。此外,每个区域大数据节点106a-106m(或其相应的大数据设备116a-

116m) 可以基于最初学习到的知识确定或定义新的服务、功能、例程或应用 (和/或修改现有的服务、功能、例程或应用), 其转而作为进一步的学习知识进行存储。

[0108] 区域大数据节点106a-106m可以用于分层或分级学习。例如, 一个或多个区域大数据节点可以将其学习的知识和/或存储的数据传送到监督多个区域的上游大数据节点。如图4所示, 区域大数据节点106c从区域大数据节点106a和106b接收学习知识和/或数据, 并且节点106c将其接收的大数据历史化在其相应的嵌入式设备116c中。区域大数据节点106c可以对所接收的学习知识和/或数据中的至少一些执行进一步的分析或学习 (例如, 通过使用分析功能A₈-A_w中的一个或多个) 来生成额外的学习知识 (例如, 数据模式、趋势、相关性等、服务、功能、例程和/或应用), 其又可以作为额外的区域大数据存储在嵌入式设备116c内和/或提供给其它大数据节点106a、106b、108。当然, 数据管道可以用于实施分析功能A₈-A_w中的任何一个或分析功能A₁-A_w的任意组合。

[0109] 在一些情况下, 使用数据管道在自下而上的基础上进行分层或分级学习。在说明性但非限制性的示例中, 区域大数据节点106a分析从其相应的本地大数据提供者节点102a、102b的组接收的数据, 以确定其“区域”是否正确地操作。区域大数据节点106a从其分析中学习的知识可以导致区域大数据节点106a生成新的诊断参数 (例如, 学习的参数)。区域大数据节点106a可以将按照由数据管道定义的所生成的诊断参数传送到上游大数据节点106c用于储存、使用和/或访问。区域大数据节点106a可以使用相同或不同的数据管道独立地启动与上游区域大数据节点106c (例如, 随生成而自动地或周期性地) 共享新的诊断参数, 或当上游区域大数据节点106c请求区域大数据节点106a共享一种或多种类型的新学习知识时, 区域大数据节点106a可以致使新的诊断参数被传送。

[0110] 当然, 使用数据管道的分层或分级学习可以在自上而下的基础上进行。为了通过继续上述示例来说明, 上游区域大数据节点106c可以分析所接收的诊断例程, 并确定诊断例程对于其它区域大数据节点 (例如, 区域大数据节点106b) 是有用的或可应用的。因此, 上游区域大数据节点106c可以使用一个或多个其它数据管道将诊断参数分配给区域大数据节点106b, 使得区域大数据节点106b和/或其区域中包括的本地提供者节点102a、102n、104的任意一个能够将诊断参数用于其各自的诊断目的。上游区域大数据节点106c可以独立地启动与区域大数据节点106b共享新的诊断参数, 或者上游区域大数据节点106c可以根据区域大数据节点106b做出的请求致使新的诊断参数被传送。可替换地或另外, 上游区域大数据节点106c可以通过聚合和分析从其监督或连接到的所有区域大数据节点接收的学习知识来生成通用诊断参数。在这种场景中, 当上游区域大数据节点106c从区域大数据节点接收指示通用诊断可能有用的数据时, 或者由于其它原因 (例如由不同的数据管道定义的), 上游区域大数据节点106c根据特定区域大数据节点的请求, 将通用诊断参数分配给其区域大数据节点中的任何一个或全部, 例如, 随生成而自动地或在周期性基础上。随后, 以类似的方式, 区域大数据节点106c下游的每个区域大数据节点可以将通用诊断例程分配给其各自区域中的任何数量的本地大数据提供者。

[0111] 在其它情况下, 区域大数据节点 (例如, 节点106a和106b) 可以使用一个或多个数据管道 (例如, 以对等方式) 彼此共享学习知识。例如, 区域大数据节点106a将新的或学习的分析例程或参数直接传送到区域大数据节点106b, 使得区域大数据节点106b可以将新的分析例程由于其自己的目的。

[0112] 注意,在图4中,描绘了仅一个上游区域大数据节点106c。然而,关于图4讨论的技术和概念可以应用于支持多层或多级的大数据历史库、储存和学习的任何数量的上游区域大数据节点。

[0113] 此外,由于区域大数据设备和本地化大数据设备都服务于过程工厂的相应的大数据节点和/或不同的相应组或区域,但不服务于整个过程工厂或多于一个其区域,区域大数据设备和本地化大数据设备在此通常且归类地称为“分布式大数据设备”。一般来说,分布式大数据设备与多个其它大数据设备进行大数据通信。例如,包括在特定大数据提供者节点中的本地大数据设备可以将学习知识和/或大数据传送到其它大数据提供者节点中包括的其它本地化大数据设备、一个或多个区域大数据设备和/或集中式大数据设备(下面将更详细地描述)。类似地,区域大数据设备可以从一个或多个本地化大数据设备和/或大数据提供者节点接收大数据。区域大数据设备可以将学习知识和/或大数据传送到其它区域大数据设备和/或集中式大数据设备。

[0114] 如上所述,在过程控制大数据网络100的一些配置中,区域大数据节点或设备106a-106m、本地大数据节点或设备102a-102n和/或其它大数据节点或设备104中的至少一些将相应的大数据、分析结果和/或学习信息传送给集中式大数据节点108,并且可以根据一个或多个预定义的数据管道这样做。“集中式大数据节点”(对应于图1的层28)通常服务于过程工厂的多个区域,在某些情况下,服务于过程工厂的大部分或全部。因此,集中式大数据节点108包括一个或多个集中式嵌入式大数据设备118,用以接收、存储过程工厂大数据和提供对过程工厂大数据的访问。例如,集中式大数据设备118可以提供由过程工厂生成的大部分或所有大数据的全面的长期历史化,和/或集中式大数据设备118可以将过程工厂范围的可用性的数据发布到其它大数据节点,甚至是处理工厂内部或外部的不是过程控制大数据节点的计算设备。

[0115] 在一些配置中,单个集中式大数据节点108或设备118可能不服务于整个过程控制系统或工厂,而是服务于过程控制系统或工厂的多于一个区域。例如,出于安全和访问目的,在单个工厂或系统中可以使用不同的集中式大数据节点108或设备118来划分大数据的不同类型或区域。在一些配置中,单个集中式大数据节点108或设备118服务于整个过程工厂。

[0116] 在过程工厂中,区域大数据节点106a-106m中的一个或多个可以使得其生成或接收的学习知识和/或数据的一些或全部被流式传输或以其它方式传送到集中式大数据节点108。例如,区域大数据节点106a-106m中的一个或多个将其各自存储的学习知识和/或数据中的至少一些传送到集中式大数据节点108。在一些实施例中,区域大数据节点106a-106m中的一个或多个以一个或多个适当的数据管道定义的方式,以周期性间隔将其各自存储的学习知识和/或数据中的至少一些推送到集中式大数据节点108。在一些实施例中,响应于来自集中式大数据节点108的请求,区域大数据节点106a-106m中的一个或多个提供其各自存储的学习知识和/或数据中的至少一部分。在这种情况下,数据管道块可以被设置为将数据自动发布或发送到管道中的下一个(上游)块或者从数据管道中的前一个(下游)块请求数据。

[0117] 集中式大数据节点108和/或其嵌入式设备118可以被配置为进一步分析从区域大数据节点106a-106m接收到的所接收的学习知识和/或数据中的任何一个或全部,例如通过

利用一个或更多的分析功能 A_x-A_y 。在一个实施例中,分析 A_x-A_y 中的每一个由图5的一个或多个大数据分析器226执行,其驻留在相应的大数据节点上。一个或多个分析功能 A_x-A_y 可以根据一个或多个数据管道所定义的对所接收的学习知识和/或数据操作,以生成额外的知识并确定过程工厂内部和外部的各个实体和提供者之间的关系。例如,额外的知识和确定的关系可以作为在嵌入式设备118处的额外的集中式大数据被存储和以其它方式利用。在某些情况下,集中式大数据节点108或设备118利用生成的知识和关系来相应地控制工厂的一个或多个过程。

[0118] 实际上,大数据网络100的任何节点102-106可以向集中式大数据设备118流式传输或以其它方式提供大数据,例如用于历史记录或长期储存。例如,本地大数据提供者节点102可以将其大数据直接流式传输到集中式大数据节点108。类似地,大数据网络的任何节点102-106可以请求由嵌入式集中式设备118提供的服务,和/或可以请求对存储在其中的数据和元数据访问。此外,在多个集中式大数据节点108或设备118服务于单个过程工厂的实施例中,多个集中式大数据节点108或设备118可以以对等方式进行通信,类似于针对区域大数据节点106a-106m所述的。

[0119] 同样类似于区域大数据节点106a-106m,在某些情况下,集中式大数据节点108本身可以是大数据的生产者或提供者,诸如当由集中式大数据节点108执行的分析(例如,一个或多个分析功能 A_x-A_y)导致存储在集中式大数据设备118处并使得其它大数据节点102-106可访问的额外发现或学习的信息时。然而,通常,从其它大数据节点102-106接收由集中式大数据设备118处理和/或服务的大量数据的大部分。在上述美国专利申请No.13/784,041中可以找到可与本文所描述的技术一起使用的示例性集中式大数据节点108和示例性集中式大数据设备118的说明。然而,应当理解,本文所描述的任何或全部技术可以可替换地或附加地与除了在美国专利申请No.13/784,041中说明的那些以外的集中式大数据设备一起使用。

[0120] 在一些配置中,集中式大数据节点108将数据和学习信息传送到远程大数据节点(例如,相对于过程工厂较远的大数据节点),用于大数据历史记录、储存、访问和/或分析。这个大数据节点(本文称为“云端大数据节点110”)可以向多个不同的过程工厂或过程控制系统10提供服务。例如,运营多个不同炼油厂的公司可以提供云端大数据节点110和云端大数据设备121,以服务于与其所有炼油厂相关的大数据。如本文所描述的数据管道可以部分地或完全在云端大数据节点110中实施。例如,通过云端大数据节点110和驻留的云端大数据设备121,特定炼油厂的集中式大数据节点可以获得由过程工厂的过程控制大数据网络100生成的发布大数据,可以使用相同的数据管道将获得的发布大数据用于特定炼油厂的运行。在一些实施例中,任何大数据节点102-106可以根据一个或多个数据管道将数据直接流式传输或提供到云端大数据节点110。类似地,任何大数据节点102-106可以根据一个或多个数据管道请求云端大数据节点110的嵌入式设备121提供的服务和/或访问存储在其中的数据和元数据。尽管图4中未示出,但云端大数据节点110可以在其中包括一个或多个相应的分析例程、功能或过程,例如,可以由图5的大数据分析器226提供的,其可以根据如本文所述的一个或多个数据管道来实施。

[0121] 此外,应当注意,并非所有类型的大数据节点都包括在所有过程工厂中。例如,特定过程工厂的最高级别的大数据处理可能在区域级,因此特定过程工厂可能不包括任何集

中式大数据节点108,并且可能不连接到云端大数据节点110。尽管如此,通常为了有利于或支持过程控制大数据,过程工厂包括至少一个本地大数据提供者节点102和至少一个大数据设备112、116、118。

[0122] 另外,在一些实施例中,过程工厂包括不固有地包括任何大数据支持的一个或多个传统过程控制设备(未示出)。在这些实施例中,工厂中的网关节点或直接耦合到传统设备的附属设备可以在由传统设备所使用的协议和过程控制大数据网络主干所使用的协议之间变换或转换数据消息,从而可通信地连接传统设备和过程控制大数据网络100。然而,数据管道可以通过这些节点,以在数据管道的各种数据处理块之间进行通信。

[0123] 此外,通常,大数据节点或设备102-110不具有集成用户接口,尽管一些大数据节点或设备102-110中的一些可以具有与一个或多个用户接口设备130通信连接的能力,例如通过有线或无线通信链路进行通信,或者通过将用户接口设备130插入到大数据节点或设备102-110的端口中。在图如图4所示,用户接口设备130被示出为无线连接到过程控制大数据网络100的大数据节点。

[0124] 用户接口设备130是包括一个或多个集成用户接口的设备(例如,移动或固定计算设备、工作站、手持设备、桌面计算设备、平板电脑等),通过其用户或操作员可以与设备和过程控制系统或过程工厂交互以执行与过程工厂相关的活动(例如,配置、查看、监控、测试、诊断、排序、计划、单独、注释和/或其它活动)。集成用户接口可以包括屏幕、键盘、辅助键盘、鼠标、按钮、触摸屏、触摸板、生物测量接口、扬声器和麦克风、照相机和/或任何其它用户接口技术。用户接口设备130可以包括到过程控制系统大数据网络主干的直接有线和/或无线连接,或者可以包括例如经由接入点或网关到主干的间接连接。用户接口设备130可以被实施为无线手持设备,例如图1的设备42,并且可以连接到一个或多个数据管道以从一个或多个数据管道接收最终用途数据。此外,这些设备或最终用途设备可以间歇地连接到数据管道,或者可以长时间保持连接到数据管道。

[0125] 在一些实施例中,用户接口设备130可以具有一个或多个内置的分析能力(在图4中带圆圈的 A_z 表示)。即,用户接口设备130可以与任何数量的大数据节点和/或大数据设备进行通信,以下载和/或接收数据,并对下载/接收的数据执行本地分析 A_z ,以发现或学习知识,并且可以作为数据管道的一部分或通过连接到数据管道来这样做。实际上,在一些配置中,用户接口设备130本身可以是大数据提供者节点,并且本身可以将其分析 A_z 的结果中的至少一些作为大数据提供给一个或多个其它本地、区域、集中式或云端大数据节点102-110。例如,可以在例如前述的美国专利申请No. 14/028,785中找到在过程控制大数据网络中使用用户接口设备(其可以与本文描述的任何或全部技术一起使用)的示例。但当然,使用与过程控制大数据网络100的用户接口设备的其它技术可以可替换地或另外地与本文所描述的任何或所有技术结合使用。

[0126] 图6示出了包括与过程工厂或过程控制系统401相关联的各种设备的另一示例性工厂网络400的方框图,其中使用本文公开的一个或多个数据管道可以实施和包括基于自动信号处理的学习技术、方法,系统和装置。例如,基于自动信号处理的学习系统的至少一部分可以使用一个或多个数据管道在过程工厂网络400中实施。过程工厂网络400的至少一部分可以由诸如图4的过程控制大数据网络100的过程控制大数据网络来支持。然而,虽然为了便于讨论而非限制目的下面关于图1、4和5所示的特征来说明过程工厂网络400,但可

以理解,图6的处理厂网络400可以利用除图1、4和5中描述的以外的过程控制大数据网络,或过程工厂网络400可以省略任何过程控制大数据网络、节点和/或设备。

[0127] 如图6所示,过程控制大数据节点或设备由“BD”参考标记指示,表示该节点是过程控制大数据提供者节点、大数据设备或两者。例如,由图6中的“BD”参考标记指示的节点或设备可以参考图4,本地大数据提供者节点和/或设备102a-102n、112a-112n,区域大数据提供者节点和/或设备106a-106m、116a-116m,集中式大数据提供者节点108和/或设备118,或另一种类型的大数据节点104,或者可以参考图1中的类似节点。

[0128] 在图6中,过程控制大数据节点BD是过程控制大数据网络400的节点。如果需要,过程控制大数据网络400可以是图4的过程控制大数据网络100,节点BD可以是网络100的节点102-108。在图6中,节点BD经由过程控制系统大数据网络主干405在网络400上通信地连接。主干405包括多个联网的计算设备或交换机,其被配置为往/来于各个过程控制大数据节点BD路由分组。主干405的多个联网计算设备可以通过任何数量的无线和/或有线链路相互连接,并且大数据网络主干405可以支持一个或多个合适的路由协议,诸如过程控制大数据流式传输协议,如先前关于网络100所讨论的。

[0129] 如图6所示,过程控制大数据网络400包括集中式大数据设备408和多个其它大数据提供者节点411、415、416、418、419、420、421、426、428、435、442a、442b、444、452a、455a、458、472、475、478。示例性大数据提供者节点之一是大数据处理控制器设备411,其本地收集、分析和存储过程控制网络或工厂40的大数据。控制器411经由输入/输出(I/O)卡426和428通信地连接到有线现场设备415-422,并且经由无线网关435和过程控制大数据网络主干405通信地连接到无线现场设备440-446。但在另一实施例中,控制器411可以使用除大数据主干网405之外的通信网络来通信地连接到无线网关435,诸如通过使用包括任何数量的有线和/或无线通信链路的过程控制通信网络,通过使用公共或基于互联网的通信网络,使用局域网等。在图6中,控制器411是过程控制系统大数据网络400的大数据提供者节点BD,并且直接连接到过程控制大数据网络主干405。

[0130] 控制器411(举例而言,其可以是由艾默生过程管理公司出售的DeltaV™控制器)可以操作以使用至少一些现场设备415-422和440-446实施批量过程或连续过程。除了通信地连接到过程控制大数据网络主干405之外,控制器411还可以使用任何期望的硬件、软件和/或与例如标准4-20mA设备、I/O卡426、428和/或任何智能通信协议(例如FOUNDATION® Fieldbus协议、HART®协议、WirelessHART®协议等)相关联的通信链路或网络通信地连接到至少一些现场设备415-422和440-446。控制器411可以使用过程控制大数据网络主干405与至少一些现场设备415-422和440-446通信地连接。在图6中,控制器411、现场设备415-422和I/O卡426、428被示出为有线设备,现场设备440-446被示出为无线现场设备。当然,有线现场设备415-422和无线现场设备440-446可以符合任何其它期望的标准或协议,例如任何有线或无线协议,包括将来开发的任何标准或协议。

[0131] 过程控制器设备411包括实施或监督一个或多个过程控制例程(例如,存储在存储器432中的过程控制例程)的处理器430,其可以包括控制回路。处理器430被配置为与现场设备415-422以及通信地连接到主干405的其它过程控制大数据节点BD通信。应当注意,如果希望的话,本文所述的任何控制例程或模块(包括质量预测和故障检测模块或功能块)可以具有由不同控制器或其它设备实施或执行的其部分。同样地,要在过程控制系统40内实

施的本文所述的控制例程或模块可以采取任何形式,包括软件、固件、硬件等。控制例程可以以任何期望的软件格式来实施,例如使用面向对象编程、梯形逻辑、顺序功能图、功能框图、或使用任何其它软件编程语言或设计范例。控制例程可以存储在任何期望类型的存储器中,诸如随机存取存储器(RAM)或只读存储器(ROM)。类似地,控制例程可以硬编码到例如一个或多个EPROM、EEPROM、专用集成电路(ASIC)或任何其它硬件或固件元件中。因此,控制器411可以被配置为以任何期望的方式实施控制策略或控制例程。

[0132] 在一些实施例中,控制器411使用通常称为功能块的来实施控制策略,其中,每个功能块是整个控制例程的对象或其它部分(例如,子例程),并且结合其它功能块(通过称为链路的通信)操作以实施过程控制系统401内的过程控制回路。基于控制的功能块通常执行诸如与发射器、传感器或其它过程参数测量设备相关联的输入功能、诸如与执行PID、模糊逻辑等控制的控制例程相关联的控制功能、或控制诸如阀门的某些设备的操作以在过程控制系统401内执行一些物理功能的输出功能中的一个。当然,存在混合和其它类型的功能块。功能块可以存储在控制器411中并由控制器411执行,这通常是这些功能块用于或关联于标准4-20mA设备和某些类型的智能现场设备(例如HART设备)的情况,或者可以存储在现场设备自身中并由现场设备实施,这可以是Fieldbus设备的情况。控制器411可以包括可以实施一个或多个控制回路的一个或多个控制例程438。每个控制回路通常被称为控制模块,并且可以通过执行功能块中的一个或多个来执行。

[0133] 图6所示的有线设备411-422包括大数据有线过程控制设备415、416和418-421以及I/O卡426、428。图6还示出了可以结合过程工厂内的有线大数据设备415、418-421、426、428一起操作的有线传统设备417和422。有线现场设备415-422可以是任何类型的设备,例如传感器、阀门、发射器、定位器等,而I/O卡426和428可以是符合任何期望的通信或控制器协议的任何类型的I/O设备。在图6中,现场设备415-418是通过模拟线路或组合的模拟和数字线路与I/O卡426通信的标准4-20mA设备或HART设备,而现场设备419-422是智能设备,例如FOUNDATION[®]Fieldbus现场设备,其使用Fieldbus通信协议通过数字总线与I/O卡428通信。然而,在一些实施例中,大数据有线现场设备415、416和418-421中的至少一些和/或大数据I/O卡426、428中的至少一些另外或可替代地使用大数据网络主干405与控制器411通信。

[0134] 网络400还被示为包括监控系统(其可以是或可以不是过程控制系统401的一部分)。监控系统包括图6所示的无线现场设备440-446,这些还包括无线大数据节点或设备BD(例如,设备442a、442b、444)的示例。图6还包括传统无线设备(例如,设备446)的示例。无线现场设备440-446使用诸如WirelessHART协议的无线协议在无线网络470中进行通信。这种无线现场设备440-446可以直接与同样被配置为进行无线通信(例如使用无线流式传输协议)的过程控制大数据网络400的一个或多个其它大数据设备或节点BD进行通信。为了与未被配置为进行无线通信的一个或多个其它大数据节点进行通信,无线现场设备440-446可以利用连接到主干405或另一个过程控制通信网络的无线网关435。任何数量的支持大数据的无线现场设备都可以在过程工厂40中使用。

[0135] 如图6所示,无线网关435是包括在过程控制工厂或系统40中的大数据节点BD的另一示例,并且提供往/来于无线通信网络470的各种无线设备440-458的访问。具体地,无线网关435提供无线设备440-458、有线设备411-428和/或过程控制大数据网络400的其它节

点或设备(包括图6的控制器411)之间的通信耦合。例如,无线网关435可以通过使用大数据网络主干405和/或通过使用过程工厂40的一个或多个其它通信网络来提供通信耦合。

[0136] 无线网关435在一些情况下通过到有线和无线协议栈的较低层的路由、缓冲和定时服务(例如,地址转换、路由、分组分段、优先化等)同时隧道传输有线和无线协议栈的一个或多个共享层来提供通信耦合。在其它情况下,无线网关435可以在不共享任何协议层的有线协议和无线协议之间转换命令。除了协议和命令转换之外,无线网关435可以提供与在无线网络470中实施的无线协议相关联的调度方案的时隙和超帧(时间间隔相等的通信时隙的集合)所使用的同步时钟控制。此外,无线网关435可以为无线网络470提供网络经营和管理功能,例如资源管理、性能调整、网络故障缓解、监控流量、安全性等。

[0137] 类似于有线现场设备415-422,无线网络470的无线设备440-446可以执行过程工厂40内的监控功能和/或物理控制功能,例如打开或关闭阀门或进行过程参数测量。然而,无线现场设备440-446被配置为使用网络470的无线协议进行通信。因此,无线网络470的无线现场设备440-446、无线网关435和其它无线节点452-458是无线通信分组的生产者和消费者。

[0138] 在某些场景中,无线网络470可以包括非无线设备,其可以是或可以不是大数据设备。例如,图6的现场设备448可以是传统的4-20mA设备,现场设备450可以是传统的有线HART设备。为了在网络470内通信,现场设备448和450可以经由无线适配器(WA)452a或452b连接到无线通信网络470。在图6中,无线适配器452b被示出为使用无线协议进行通信的传统无线适配器,无线适配器452a被示出为支持大数据,因此可通信地连接到大数据网络主干405。另外,无线适配器452a、452b可以支持诸如**Foundation[®]** Fieldbus、PROFIBUS、DeviceNet等的其它通信协议。此外,无线网络470可以包括一个或多个网络接入点455a、455b,其可以是与无线网关435有线通信的分离的物理设备或可以作为整体设备被提供有无线网关435。在图6中,网络接入点455a被示出为大数据设备BD,而网络接入点455b是传统接入点。无线网络470还可以包括一个或多个路由器458,以将分组从一个无线设备转发到无线通信网络470内的另一个无线设备,每个无线设备可以或可以不支持过程控制系统40中的分布式大数据。如果无线设备是分布式和/或集中式大数据设备,则无线设备440-446和452-458可以通过无线通信网络470的无线链路460和/或经由大数据网络主干网405彼此及与无线网关435通信。

[0139] 因此,图6的系统包括节点BD的大数据设备的几个示例,其主要用于向过程控制系统的各种网络提供网络路由功能和管理,并且可以在其中或通过其可以实施数据管道。例如,无线网关435、接入点455a和路由器458各自包括在无线通信网络470中路由无线分组的功能,并且这些节点或设备可以作为实施一个或多个数据管道的一部分来实施这种通信。无线网关435为无线网络470执行流量经营和管理功能,并且往来于与无线网络470通信连接的有线网络路由流量。无线网络470可以利用特定支持诸如WirelessHART的过程控制消息和功能,并且也可以用于数据管道通信的无线过程控制协议。如图6所示,无线网络470的设备435、455a、452a、442a、442b和458支持过程控制工厂40中的大数据,然而,无线网络470的任何数量的任何类型的节点可以支持过程工厂40中的分布式大数据。

[0140] 使用其它无线协议进行通信的其它设备可以是过程控制大数据网络400的大数据节点或设备BD。在图6中,一个或多个无线接入点472是大数据设备BD,其可以利用其它无线

协议实施数据管道或数据管道的部分,诸如Wi-Fi或其它IEEE 802.11兼容无线局域网协议、诸如WiMAX(全球微波接入互操作性)、LTE(长期演进)或其它ITU-R(国际电信联盟无线电通信部门)兼容协议的移动通信协议、诸如近场通信(NFC)和蓝牙的短波无线电通信或其它无线通信协议。通常,这样的无线接入点472允许手持或其它便携式计算设备(例如,用户接口设备)通过不同于无线网络470并支持与无线网络470不同的无线协议的相应无线网络进行通信。除了便携式计算设备之外,在一些场景中,一个或多个过程控制设备(例如,控制器411、现场设备415-422或无线设备435、440-458)也可以使用由接入点472支持的无线协议进行通信。

[0141] 另外在图6的系统中,到直接过程控制系统40外部的系统的一个或多个网关475、478是过程控制大数据网络400的大数据节点或设备BD,其可以用于实施数据管道或使用数据管道的分析模块。通常,这样的系统是由过程控制系统40生成或在过程控制系统40上操作的信息的客户或提供者。例如,工厂网关节点475可以将直接过程工厂40(具有其各自的过程控制大数据网络主干405)与具有其各自的过程控制大数据网络主干的另一个过程工厂通信地连接。在另一示例中,单个过程控制大数据网络主干405可以服务于多个过程工厂或过程控制环境。在又一示例中,工厂网关节点475将直接过程工厂40与云端大数据节点110和/或云端大数据设备121通信地连接。

[0142] 图6所示的工厂网关节点475将直接过程工厂40通信地连接到不包括过程控制大数据网络400或主干405的传统或现有技术过程工厂。在该示例中,工厂网关节点475可以在由工厂40的过程控制大数据主干405使用的协议与由作为数据管道的一部分的传统系统(例如,以太网、Profibus、Fieldbus、DeviceNet等)使用的不同协议之间变换或转换消息。一个或多个外部系统网关节点478将过程控制大数据网络400与外部公共或私有系统的网络通信地连接,外部公共或私有系统例如实验室系统(例如,实验室信息管理系统或LIMS)、操作员回合数据库、材料处理系统、维护管理系统、产品库存控制系统、生产调度系统、气象数据系统、运输和处理系统、包装系统、互联网、另一提供者的过程控制系统或其它外部系统。

[0143] 尽管图6仅示出了具有有限数量的现场设备415-422和可能的440-446的单个控制器411,但该示例仅是说明性和非限制性的实施例。任何数量的控制器411可以支持大数据,并且任何控制器411可以与任何数量的有线或无线现场设备415-422、440-446进行通信,以控制工厂40中的过程。此外,过程工厂40可以还包括任何数量的无线网关435、路由器458、接入点455、无线过程控制通信网络470、接入点472和/或网关475、478。此外,图6的系统可以包括任何数量的集中式大数据设备408,其可以接收和存储从过程工厂40中的任何或所有设备收集的数据和/或生成的学习数据或知识。在一些实施例中,过程控制大数据网络400可以包括任何数量的区域大数据设备和节点(图6中未示出)。

[0144] 此外,包括在如图6所示的示例性过程工厂40中的各方面、设备和组件的组合仅是示例性的。本文公开的数据管道技术、系统、方法和装置可以在具有零个或多个图6所示任何方面的过程工厂中使用。例如,本文公开的数据管道技术、系统、方法和装置可以在没有集中式大数据设备408的过程工厂中或在具有一个或多个区域大数据设备和/或节点的过程工厂中使用。在另一个示例中,本文公开的技术、系统、方法和装置可以在仅具有传统设备的过程工厂中使用。

[0145] 现在同时参考图1-6,过程控制大数据网络的任何数量的任何节点102、106、108、110,大数据设备112、116、118、121,大数据节点411、415、416、418、419、420、421、426、428、435、442a、442b、444、452a、455a、458、472、475、478和/或传统设备417、422、440、446、448、450、452b、455b可以包括数据管道的一个实例或一部分。即,实施为数据管道的分析模块的实例可以与生成其值随时间变化的信号的过程工厂或过程控制系统中包含或相关联的任何节点或一个或多个设备(例如过程控制设备、大数据节点、分析器等)集成。实施为数据管道的信号处理模块的至少一部分被实施为存储在图1-6所示的任何一个或多个节点或设备的存储器上并且可以由共同驻留在节点或设备上的处理器执行的软件或计算机可执行指令。

[0146] 在一些情况下,被实施为数据管道的信号处理模块的至少一部分被实施为芯片、集成电路、半导体电路或包括在图1和4-6所示的任何一个节点或设备中的其它合适的硬件。例如,信号处理模块的至少一部分可以在一个或多个集成数字信号处理(DSP)芯片上实施,例如Texas Instruments KeyStone多核DSP和ARM™处理器或其它合适的DSP芯片。一个或多个集成DSP芯片可以嵌入生成其值随时间变化的信号的过程工厂或过程控制系统中包含或相关联的任何节点或设备(例如过程控制设备、大数据节点、分析器等)中或与之集成。例如,集成DSP芯片中的一个或多个可以包括在阀门、振动测量设备、发射器等中。可以根据所期望的增加芯片的数量以加速、增强和/或添加信号处理功能/能力。

[0147] 在一些情况下,实施为数据管道的信号处理模块不与生成其值随时间变化的信号的(过程工厂或过程控制系统的)节点或设备集成,而是信号处理模块通信地耦合到节点或设备。例如,以数据管道形式实施的信号处理模块可以物理地附接到生成其值随时间变化的信号的节点或设备,使得信号源的输出在通过过程控制网络传送之前遍历信号处理模块的各个块或组件,或者信号处理模块可以监控传送节点或设备的输出的网络链路。

[0148] 可以通过数据工作室配置应用来执行创建或配置数据管道,数据工作室配置应用通常提供用户可以使用的配置环境,用于轻松地定义数据管道中的块和各块之间的互连,以及将在工厂网络内下载并运行或执行每个块的位置。数据工作室通常可以包括存储可用于创建数据管道的数据处理块的模板的配置区域和使用户能够指定在数据管道中使用哪些处理块以及这些块之间的通信互连的配置区域。一旦在配置区域中创建了,就可以使用大数据机器中的历史数据来测试数据管道。在测试完成后,创建的数据管道可以下载到工厂网络中的各种设备,并且可以随着在工厂内收集数据而在数据上自动运行,以在工厂内不断提供新的学习。

[0149] 一般来说,数据管道中的块集合可以用于读取数据、滤波、形成、执行分析、可视化结果及对数据执行其它功能。这些块用于构建“数据管道”,将数据源组合和形成为可以由关于图1-6描述的分析块处理和测试的形式。

[0150] 在一般意义上,可以在两个级别上观察数据工作室。首先,数据工作室为用户提供结构化的环境,用以创建执行数据处理以实际上生成预测、执行故障检测和生成预测的处理和计算实体。数据工作室用于形成数据、生成模型、测试模型、比较模型,并执行数据科学家通常执行的功能。一旦开发了模型,可以将相同的模型集合下载到工厂内的计算引擎中,例如图1-图6的嵌入式分析引擎或大数据设备。这些模型在本文中也称为分析模块。

[0151] 图7是包括大数据网络500的过程工厂的示例性过程控制或监控系统(例如图1、图

4或图6的)的框图,所述大数据网络500又进一步收集过程工厂内的数据。图7的过程控制系统还包括在例如过程工厂环境中提供的用户接口站503,其执行数据工作室509,以使用户能够在由大数据网络收集的数据上创建和运行各种数据模型。用户接口站503还执行数据知识应用514(其可以与图3的框94中的功能相关联或被实施为所述功能),其执行进行系统或全面诊断、预后、分析、识别关系等的模型,以发现和评估从过程工厂获取的数据。虽然图7示出了只有单个用户接口站503托管数据工作室509和数据知识应用514,但在一些实施例中,数据工作室509和数据知识应用514可以托管在不同用户接口站503上、多个用户接口站503上或可以分布在过程工厂509中的一个或多个设备中。更具体地,在工厂或过程控制系统内实施的数据工作室509为用户或模型开发者提供了一种工具来帮助系统地 and 有效地定义或创建在工厂环境中运行的、以前从工厂收集的历史数据、目前在工厂内收集的实时数据或两者上的数据管道。数据工作室509和数据知识应用514能够利用大数据网络500对在工厂或过程控制系统内生成或收集的所有或大部分数据的收集、存储和提供有效访问的事实。此外,数据建模工作室509使用该数据为模型开发者提供在工厂或过程控制系统中设计、测试和实施模型的更有效的方式。数据知识应用514提供用于用工厂或过程控制系统植入模型以及查看和探究数据之间的关系以及所执行的模型的结果的附加环境。如图7所示,数据工作室509可以是由一个或多个设备的一个或多个处理器执行的指令集,诸如在用户接口站503上。当然,数据工作室509可以在各种计算设备的任意一个上执行,例如平板电脑、移动电话、笔记本计算机、台式计算机、服务器等,及在一个设备内执行,或者可以分布在任何数量的这种设备中。类似地,数据知识应用514可以是由一个或多个设备的一个或多个处理器执行并且在各种计算设备中的任意一个上执行的指令集。

[0152] 在一般意义上,示例性过程控制系统大数据网络500包括过程控制系统大数据装置或设备502、过程控制系统大数据网络主干505和多个节点508,其通信地连接到主干505。节点508另外连接到过程工厂设备,如控制器、现场设备、工厂设备、监控设备等。可以在多个节点508处收集并缓存与过程有关的数据、与工厂有关的数据、与通信相关的数据和其它类型的数据,并且该数据可以经由网络主干网505传送到过程控制系统大数据装置或设备502以用于长期储存(例如,“历史化”)和处理。至少一些数据可以在网络500的节点508之间传送,例如用以实时地控制过程。

[0153] 与过程工厂系统相关的任何类型的数据可以被收集并存储在过程控制系统大数据设备502上。例如,可以收集和存储实时过程数据,例如连续的、批量、测量和事件数据,其过程或工厂网络中控制过程的同时生成(在某些情况下,指示过程的实时执行的效果)。此外,可以收集和存储过程定义、布置或设置数据,例如过程工厂配置数据和/或批量配方数据。此外,可以收集和存储与过程诊断的配置、执行和结果相对应的数据。当然,也可以收集和存储其它类型的过程数据,例如警报或报警。以类似的方式,可以收集和存储过程工厂10的主干505和各种其它通信网络的数据高速公路流量和网络管理数据。类似地,可以收集和存储诸如与用户流量相关的数据的用户相关数据、登录尝试、查询和指令。此外,可以收集和存储文本数据(例如,日志、操作过程、手册等)、空间数据(例如,基于位置的数据)和多媒体数据(例如闭路电视、视频剪辑等)

[0154] 过程工厂大数据网络主干505可以包括多个联网的计算设备,诸如用户接口设备503(托管数据工作室509和知识发现应用514)或被配置为往/来于过程控制系统大数据网

络500的各个节点508和往/来于过程控制大数据设备502(其本身是过程控制系统大数据网络500的节点)路由分组的交换机。主干505的多个网络计算设备可以通过任何数量的无线和/或有线链路相互连接。此外,工厂网络大数据网络主干505可以包括一个或多个防火墙设备。

[0155] 举例而言,大数据网络主干505可以支持一个或多个合适的路由协议,例如包括在网际协议(IP)套件中的协议(例如,UPD(用户数据报协议)、TCP(传输控制协议)、以太网等)或其它合适的路由协议。如果期望的话,节点508中的至少一些节点可以利用诸如流控制传输协议(SCTP)之类的流式传输协议来经由网络主干505从节点508向过程控制大数据设备502流式传输缓存的数据。通常,包括在过程数据大数据网络500中的每个节点508可以至少支持由主干505支持的路由协议的应用层(对于一些节点,附加层)。每个节点508可以在过程控制系统大数据网络500内,例如通过唯一的网络地址来唯一地识别。此外,过程控制系统大数据网络500的至少一部分可以是自组网。因此,节点508中的至少一些可以以自组织的方式连接到网络主干505(或网络500的另一个节点)。

[0156] 此外,与过程工厂网络内的设备(例如,包括在过程工厂中的物理设备,例如机器和设备)相关但是可能不由直接配置、控制或诊断过程的应用生成的数据也可以被收集并存储在大数据机器502中。例如,可以存储振动数据和疏水阀数据、工厂安全数据和指示对应于工厂安全性的参数的值的数据(例如,腐蚀数据、气体检测数据,等等)。类似地,指示与工厂安全相对应的事件的数据可以被收集并存储在大数据机器502中。同样,可以收集并存储对应于机器、工厂设备和/或设备的健康的数据、设备数据(例如,基于振动数据和其它数据确定的泵健康数据)、对应于装置、机器和/或设备诊断的配置、执行和结果的数据。

[0157] 在某些情况下,由过程或工厂外部的实体生成或传输的数据可以被收集并存储在大数据机器502中,例如与原材料的成本、部件或设备的预期到达时间有关的数据、天气数据和其它外部数据。当然,由通信地连接到网络主干505的任何节点508生成、接收或观察到的所有数据可以被收集并且导致被存储在过程控制系统大数据设备502上。当然,任何或者所有这些类型的数据可以由一个或多个数据管道的一个或多个数据收集模块来收集,或者可以由任何数据管道的任何处理模块使用。

[0158] 如图7所示,数据工作室509可以被配置为提供到过程控制系统大数据网络500中的主接口,用于例如在用户接口或其它接口设备上的配置和数据探索,以供用户或其它应用使用。数据工作室509可以经由过程控制系统大数据网络主干505连接到大数据设备502,或者直接连接到过程控制系统大数据设备502,或者可以以任何其它方式耦合到大数据设备502。

[0159] 在图7的网络中,过程控制大数据装置或设备502集中在网络500内,并且被配置为从网络500的节点508接收数据(例如,经由流式传输和/或经由一些其它协议),并且存储所接收的数据。因此,过程控制大数据装置或设备502可以包括用于历史化或存储从节点508接收的数据的数据储存区520,可以收集用于储存在大数据机器502中的数据的多个设备数据接收器522以及从大数据储存器520请求和访问数据的多个设备请求服务器535。在任何情况下,有用的是如果在过程控制网络中收集并存储在数据储存区520中的所有或大部分数据对于数据建模工作室509和知识发现应用514的用户可用。

[0160] 过程控制系统大数据储存区520可以包括多个物理数据驱动器或储存实体,诸如

RAID (独立磁盘冗余阵列) 储存器、云端储存器或适用于数据库或数据中心储存的任何其它合适的数据储存技术。然而,对于网络500的节点508,数据储存区520具有单个或单一的逻辑数据储存区或实体的外观。因此,数据存储520可以被视为用于过程控制大数据网络500或过程工厂的集中式大数据储存区520。如果期望的话,单个逻辑集中式数据储存区520可以服务于多个过程工厂(例如,过程工厂和另一个过程工厂)。例如,集中式数据储存区520可以服务于能源公司的几个炼油厂。此外,集中式数据储存区520可以经由例如至少一个高带宽通信链路直接连接到主干505。此外,集中式数据储存区520可以包括集成防火墙。

[0161] 单一逻辑数据储存区520的结构可以支持所有或大多数过程控制系统相关数据的储存。例如,数据储存实体的每个条目、数据点或观察可以包括数据的身份的指示(例如,源、设备、标记、位置等)、数据的内容(例如,测量、值等),以及指示收集、生成、接收或观察数据的时间的时间戳。因此,这些条目、数据点或观察在本文中称为“时间序列数据”。例如数据可以使用包括支持可扩展储存、流式传输数据和低等待时间查询的模式的公共格式存储在数据储存区520中。

[0162] 在一种情况下,模式可以包括在每行中存储多个观察值,并且使用具有自定义哈希函数的行密钥来过滤该行中的数据。哈希函数可以基于时间戳和标记。例如,哈希函数可以是时间戳的舍入值,标记可以对应于过程控制系统的或相关的事件或实体。对应于每行或一组行的元数据也可以与时间序列数据整体地或者与时间序列数据分离地存储在数据储存区520中。例如,元数据可以与时间序列数据分离地以无模式的方式存储。

[0163] 如果期望的话,用于在设备数据储存器520处存储数据的模式也可用于将数据存储在至少一个节点508的缓存中。因此,在这种情况下,在将数据通过主干505从节点508的本地储存区传送到过程控制系统大数据设备数据储存器520时维持该模式。

[0164] 除了数据储存器520之外,过程控制系统大数据设备502还可以包括一个或多个设备数据接收器522,每个设备数据接收器522被配置为从主干505接收数据分组,处理数据分组以取得其中携带的实质性数据和时间戳,并将实质性数据和时间戳存储在数据储存区520中。设备数据接收器522可以驻留在例如多个计算设备或交换机上。多个设备数据接收器522(和/或至少一个数据接收器522的多个实例)可以对多个数据分组并行操作。

[0165] 在接收到的数据分组包括由过程控制大数据设备数据储存区520使用的模式的情况下,设备数据接收器522仅使用模式信息填充数据储存区520的附加条目或观察(如果期望的话,可以可任选地存储对应的元数据)。在接收的数据分组不包括由过程控制大数据设备数据储存区520使用的模式的情况下,设备数据接收器522可以对分组进行解码并填充过程控制大数据设备数据储存区520的时间序列数据观察或数据点(以及可任选地相应的元数据)。

[0166] 另外,过程控制系统大数据设备502可以包括一个或多个设备请求服务器535,每个设备请求服务器被配置为访问存储在过程控制系统大数据设备储存器520中的时间序列数据和/或元数据,例如,根据请求实体或应用的请求。设备请求服务器535可以驻留在例如多个计算设备或交换机上。设备请求服务器535和设备数据接收器522中的至少一些可以驻留在相同的一个或多个计算设备(例如,集成设备)上,或者可以被包括在集成应用中。

[0167] 在一些情况下,多个设备请求服务器535(和/或至少一个设备请求服务器535的多个实例)可以并行地对来自多个请求实体或应用的多个请求操作。当然,单个设备请求服务

器535可以服务于多个请求,诸如来自单个实体或应用的多个请求,或来自应用的不同实例的多个请求。在任何情况下,图7的数据建模工作室509和知识发现应用514可以经由一个或多个请求服务器535访问存储在数据储存区520中的数据。

[0168] 图8是经由图7的过程控制大数据主干505耦合到大数据设备502和过程工厂网络的数据工作室509的示例性实施例的框图。如图8进一步所示,一个更多的另外的用户接口设备513可以经由过程控制大数据主干505访问数据工作室509。另外,知识发现应用514耦合到过程控制大数据主干505。一般来说,数据工作室509提供用于图形地创建和编程数据管道的结构化环境,数据管道被配置为对从大数据机器或其它数据源收集的数据执行操作。在此背景下,数据管道是对数据执行的一系列相互连接的数学或分析例程,以产生输出,该输出提供有关过程工厂(如该过程工厂)操作的一些知识。数据管道由一个或多个数据输入、对数据输入执行的一个或多个功能和作为功能结果的一个或多个输出构成。

[0169] 如图8所示,数据工作室509包括用于生成有利于图形数据管道创建的结构化环境的配置引擎523和用于下载或执行创建的管道的运行时间引擎524。更具体地,配置引擎523包括用于生成用于结构化环境的图形用户界面的元素的接口例程525,用作数据管道的构件块的多个模板529和将模型转换为可由运行时间引擎524执行的数据格式的编译器527。运行时间引擎524包括用于获取在执行数据管道中使用的数据的数据获取例程526和用于运行可执行数据管道的执行例程528。

[0170] 接口例程525包括存储在存储器中的指令集,当由处理器执行时,生成拖放图形用户界面的一组用户界面元素,以便于创建包括显示模板529的库区域525a的数据管道,和用作创建模型的主呈现窗口的画布区域525b。模板529用作数据管道的构件块,并且可以被选择并移动到画布区域525b以创建数据管道。包括在数据工作室509的配置引擎523中的模板529包括数据源或读模板529a(其指定在管道中使用的数据源),数据写模板529b,其指定在何处写入数据或输出数据或以其它方式从管道发送数据,滤波器模板529c,其指定要在数据管道中的数据流上运行的滤波例程,合并模板529d,其指定要对一个或多个数据流执行的数据合并操作,状态模板529e,其指定状态检测和相关的处理活动,数据分析模板529f,其指定要对数据执行的分析,以及可视化模板529g,例如其指定要对数据执行的各种图形化或其它可视化功能,以向用户显示数据。当然,图8所示的模板仅是示例性的,也可以创建和存储其它类型的模板。

[0171] 在一些情况下,数据源或读模板529a指示数据源,其收集在过程工厂其传送、生成、接收和/或观察的数据的各种数据类型(结构化和/或非结构化)、背景和/或边界条件。数据源或读模板529a可以涉及数据库数据、流式传输数据、事务数据和/或经由过程控制系统大数据网络传送并被存储或历史化在过程控制系统大数据存储器520中的任何其它类型的数据。数据源或读模板529a还可以包括存储在大数据设备502中的数据源,其收集和存储作为过程工厂的大数据架构的一部分而收集的所有(或几乎所有)过程数据和工厂数据。数据源529a可以定义存储在大数据设备502中的全面的高保真度数据,定义实施大数据架构的过程工厂的任何或全部的各种参数。

[0172] 示例性数据源模板529a包括压力读数、阀门读数、温度读数、振动读数。用户可以设置历史数据或实时数据的参数,以便仅选择数据的子集。例如,数据源模板529a可以指示收集通常预期处于温度A到温度B范围内的摄氏度的温度数据的数据源。数据源模板529a还

可以表示在过程工厂的外部的数据源,例如外部工厂现场或外部数据库,如天气或环境数据库。以类似的方式,数据写模板529b可以指定任何数据库、机器或其它数据管道或分析模块,在该数据库、机器或其它数据管道或分析模块上写入数据或使来自管道中的数据可用。

[0173] 其它模板529c-529g可以实施或定义要对数据执行的任何基本或单一功能。例如,分析模板529f可以包括诸如神经网络例程、相关性例程、阈值例程或诸如平均值例程、最大值例程、最小值例程、滤波器例程等的统计过程之类的数学函数。模板529e可以包括诸如随机森林算法或偏最小二乘回归的分类技术。此外,模板529c-f可以包括用于准备用于分析的数据的数据处理技术,诸如去除高于某个阈值水平的数据的滤波例程或者用于去除异常值数据的剪切例程。用户还能够改变模板529的一个或多个参数。例如,用户可以选择滤波器例程并定义低点滤波器(特定斜率)、限幅滤波器(阈值水平)等的特性。

[0174] 可视化模板529g定义如何解释和/或呈现由数据管道对数据执行的功能的结果。例如,如果将管道的输出图形地呈现给用户,则模板529g可以表示图形、图表或显示界面。输出模板529g还可以定义一个或多个数据消息收发技术以准备好数据用于由输出模板指定的图形呈现。例如,条形图可能需要舍入到两个有效数字的数值,输出模板还可以包括用于准备数据的舍入例程。输出或可视化模板529g还可以指示应当发送数据管道的输出的特定位置。例如,模板529g可以指示由数据管道产生的值集合将被发送到过程工厂(例如该过程工厂)中的一个或多个控制器。示例值包括设定点值、控制例程等。然而,输出可以是任何值,例如表示关系的连接、特定值、指示数据输出是否满足阈值的二进制值、对应于效力的百分比等。

[0175] 然而,用户不限于使用包括在数据工作室509中的预定义模板。用户可以编辑现有模板,添加新模板和/或将编辑的/新模板保存到数据工作室509以备将来使用。例如,如果预先存在的平均值例程接受两个数据源输入,则用户可以编辑平均值例程模板以接受三个输入。用户还可以从头开始创建新模板以向数据工作室509添加额外功能。用户还可以将数据工作室509创建的数据管道或数据管道的部分保存为模板以备将来使用。

[0176] 编译器527可以包括存储在存储器中的编译器例程527a,其在处理器上执行,以将使用用户界面创建的图形数据管道转换为由运行时间引擎524支持的可执行数据格式的编译数据管道。

[0177] 执行编译数据管道以产生期望输出(由数据管道定义)的运行时间引擎524包括用于从存储器520和/或从工厂设备获取数据的数据获取例程526,以及用于运行和/或执行可执行模型的执行例程528。更具体地,处理器可以执行获取例程526以从数据源获取数据并从存储器获取对应于数据管道中使用的图形模板529的功能。在一些实施例中,获取例程526可以经由大数据网络主干505从大数据设备520和/或从工厂自身获取数据。执行例程528可以包括存储在存储器中的指令集,其在处理器上执行编译数据管道以产生一个或多个输出。

[0178] 执行例程528可以在与工厂的操作隔离的离线环境中运行数据管道,其中,数据管道可以在存储在存储器中的历史数据上运行,或者可以在在线环境中运行数据管道,其中,获取例程526可以从工厂或大数据设备中的一个或多个设备获取实时或接近实时的数据流,并产生可以提供给用户和/或可用于影响工厂运行的输出。

[0179] 作为更具体的示例,数据源模板529a(定义要输入到数据管道的数据)可以连接到

定义滤波功能的第一滤波器模板529c(例如,低通、高通、平滑滤波器等)。滤波器功能模板529c可以连接到定义剪切例程的第二滤波器模板529c,第二滤波器模板529c可以连接到定义神经网络技术的分析模板529f,分析模板529f可以随后连接到定义另一种滤波技术的滤波器模板529c,该滤波器模板529c可以连接到定义求平均值技术的模板529f,模板529f可以连接到定义显示输出的方式的可视化模板529g。

[0180] 图9示出了用于有利于图形创建数据管道的数据工作室509的示例性图形用户界面600。如上所述,配置引擎523的接口例程525生成包括显示模板529的库区域525a和用作创建数据管道的主呈现窗口的画布区域525b的拖放图形用户界面500。如图9所示,库区域525a显示用作数据管道的图形构件块的多个模板529。模板529可以包括各种数据模板529,包括读模板529a、写模板529b、滤波器模板529c、合并模板529d、状态模板529e、分析模板529f和可视化模板529g。

[0181] 用户可以构建从库区域525a中选择不同的模板529并将它们拖动到画布区域525b的数据管道。用户可以将模板529布置在画布区域525b中,并且使用通信互连来定义模板之间的流,以便定义数据管道内的数据流。尽管图9示出了拖放图形用户界面,但数据工作室509的其它实施例可以包括其它图形或非图形界面。

[0182] 配置引擎523进一步使得用户能够通过定义要在模板529中使用的特定参数和数据模板之间的互连来修改画布区域525b中示出的数据管道模板529,以创建形成数据管道的互连块,定义通信连接、执行速率等。数据管道可以包括例如定义要为数据管道获取的数据的特定数据源的输入块,定义要对管道内的数据执行的数据处理过程的一个或多个功能块和定义与一个或多个数据输出的输出相关联的操作一个或多个输出块。

[0183] 在图9的示例中,创建数据管道,其包括四个读取块602(定义从其获得用于管道的数据的不同数据源)、连接以滤波每个获取的数据流的滤波器块604以及连接到每个滤波器块604的时间偏移块606,其可以操作以将来自各种源的每一个的数据与公共或对齐的时间偏移或时间对齐。类似地,由块606中的每一个产生的数据被提供给块608,块608合并时间序列数据,并以线形图的形式向可视化块610提供合并数据或数据的矩阵。可视化块610可以产生以所期望格式显示合并的时间序列数据的线形图。

[0184] 当然,当创建图9的数据管道时,用户可以使用各种参数(例如,滤波器系数、采样率、时间偏移、可视化参数等)来配置每个块。此外,当数据管道作为数据管道下载到工厂并在工厂中执行时,用户可以指定每个块在数据管道中的位置或地方以及数据管道内块的执行速率和通信速率。例如,用户可以指定在要在工厂运行或运行时间期间下载和运行每个块的设备、节点、大数据机器等。

[0185] 此外,用户可以在存储在数据库中的历史数据上测试创建的数据管道,当满足数据管道的操作时,可以将数据管道下载到工厂中,使得将数据管道的各个块下载到指定的节点、机器或设备中并在其中执行。此外,此时,配置引擎523或其编译器527可以通过安装在存储和执行块的各种设备之间的通信网络来配置数据管道的块以彼此通信。在一些情况下,相同数据管道的不同块可以使用不同的通信网络甚至不同的通信协议彼此通信。

[0186] 一旦下载了,数据管道在工厂的运行时间期间在连续或持续的基础上在工厂中运行或执行,以对由数据管道指定的数据提供数据处理。该数据随后可供任何用户、任何机器或工厂内任何分析模块查看和使用。此外,分析模块可以使用一个或多个数据管道或可以

由一个或多个数据管道组成,单个数据管道可以在多个其它分析模块中使用。

[0187] 如将理解的,本文描述的监控和分析引擎、算法和数据管道可以在分布式控制系统(DCS)、资产管理系统以及典型的过程控制工厂或者网络的现有机器健康系统的外部运行。这些系统中的每一个被认为对其各自组织(操作、维护和可靠性工程设计)至关重要,因此运行这些模块在不同机器上运行的分析模块和数据管道可以防止这些系统、试图支持它们的运营公司、以及试图支持它们的各种组织和实体的每一个压力过大。

[0188] 此外,本文描述的DCS系统与监控和分析系统之间的区别在于监控和分析系统不是实时地消耗数据。应用被设计为在系统的时间常数内接收数据、执行分析、作出预测、执行故障检测、生成可用于人和其它系统可用的建议(使用数据管道)。例如,在批量操作中,更新的预测可以是按照分钟的。大多数连续操作也是如此。这种操作与控制操作极为不同,控制动作对过程的操作有直接的影响。当应用实际需要数据时,才需要使用流式传输数据,因此本文所述的数据管道(以及在数据管道上运行的分析模块)可以以不同的速率运行。

[0189] 此外,如将理解的,本文所述的数据管道结构能够用于识别前导指标,然后使用这些前导指标来识别对前导指标具有最大影响的参数。为了从该分析中获得最佳值,可以使用关于每个参数与每个其它参数(例如上游、下游)的关系的信息以及提供给数据管道或由数据管道使用的数据的一部分。在提供此信息的典型工厂中有几个控制和设备层级可用。例如,可以使用从工厂的诸如DeltaV和AMS及PI&Ds的系统可用的设备层级来定义设备的物理关系(例如,原材料槽在脱盐器的上游,脱盐器在加热器的上游,加热器在大气塔的上游)。

[0190] 分析模块可以借助由数据管道执行的处理使用这些数据中的任何一个来执行持续或按需分析。因此,本文描述的分析模块可以提供以具有可执行上下文的数据工作室定义的管道。分析模块可以具有标记、描述、执行速率,并且可以支持数据管道块和参数。分析模块可以下载到系统中的任何计算机或执行引擎中,并且可以自动地解析对可以经由监控和分析总线发送的数据的引用。分析模块还可以在计算元件之间移动,配置系统可以在此情况下自动重新绑定数据链路,以确保数据管道保持完整。

[0191] 另外,一旦调试了分析模块,它们就可以被存储为可以重新用于类似应用的模板。以此方式,可以创建分析模板库并重复使用多次。如果需要,这些模块也可以被许可使用和转售。

[0192] 此外,如果期望的话,可以使用例如数据工作室509调试分析模块和/或在分析模块内使用或支持分析模块的数据管道。如果分析预测或数据处理关闭,则用户可以使用数据工作室509在线操作或运行模块(或数据管道),并在数据管道或模块中的每个块处查看数据值。用户然后可以使用数据工作室509来查看数据值并校正导致错误的块、块内的算法等。

[0193] 本文公开的任何或全部系统、方法和技术可用于被配置为实时控制或监控过程的任何过程工厂或过程控制系统中。此外,如本文所使用的,“一件设备”、“设备件”或“设备”通常是指在控制过程或其部分期间可能或可能不直接利用的物理元件或组件,但是仍然可以相关于过程的控制或流程与其它设备件和/或过程元件一起来配有。

[0194] 当以软件实施时,本文所述的任何应用、块、服务和引擎可以存储在任何实体非暂时的计算机可读存储器中,例如磁盘、激光盘、固态存储器设备、分子存储器储存设备或其

它储存介质,存储在计算机或处理器的RAM或ROM中等。虽然本文公开的示例性系统被公开为除了其它组件以外,包括在硬件上执行的软件和/或固件,但是应当是指出这样的系统仅仅是说明性的,不应该被认为是限制性的。例如,可以想到,这些硬件、软件和固件组件中的任何一个或全部可以专门以硬件、专门以软件或以硬件和软件的任何组合来体现。因此,虽然本文所述的示例性系统被说明为在一个或多个计算机设备的处理器上执行的软件中实施,但是本领域普通技术人员将容易理解,所提供的示例不是实施这样的系统的唯一方式。

[0195] 因此,虽然已经参考具体示例说明了本发明,但这些具体示例仅仅旨在是说明性的而不是对本发明的限制,对于本领域普通技术人员显而易见的是在不脱离本发明的精神和范围的情况下可以对所公开的实施例进行改变、添加或者删除。

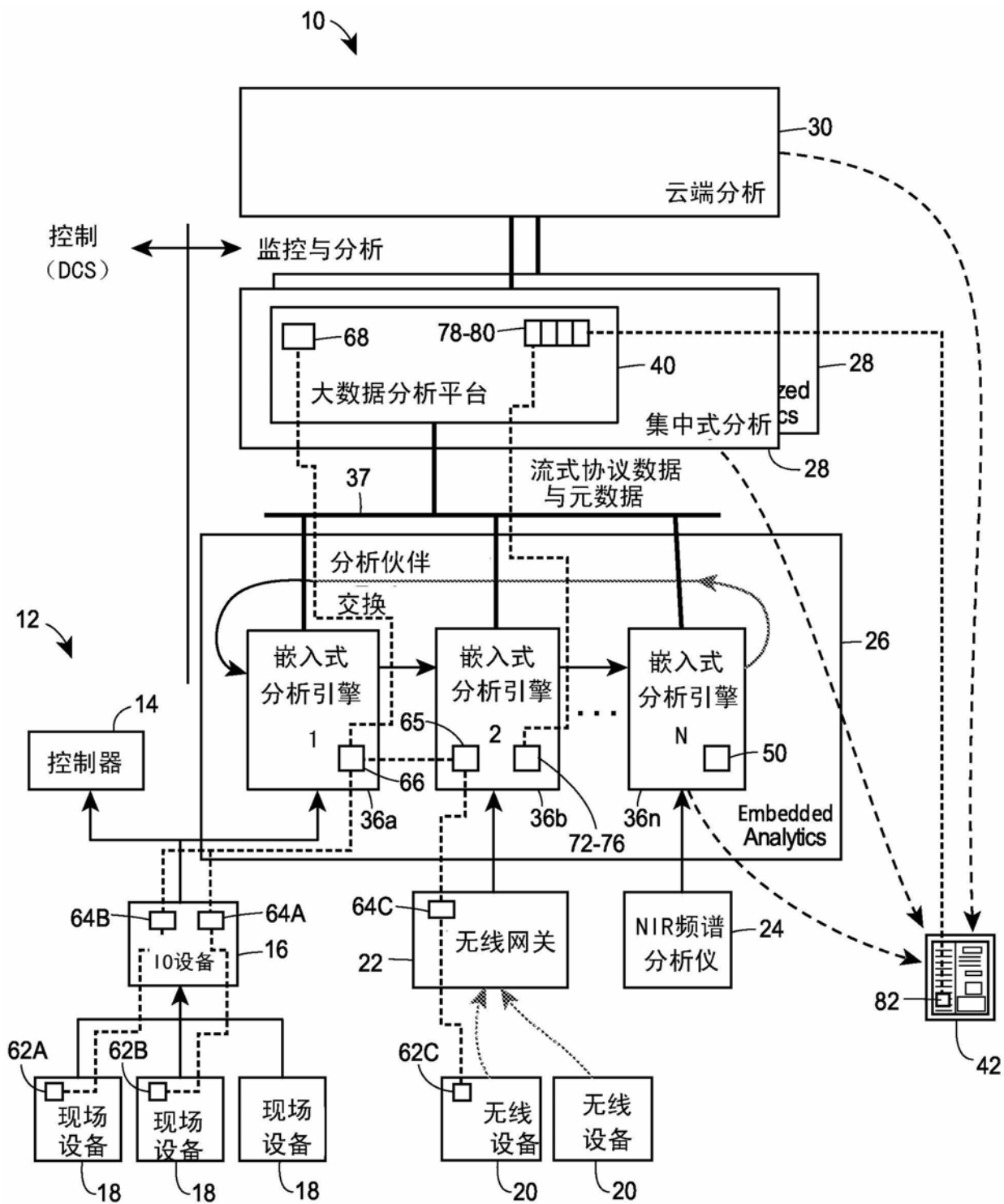


图1

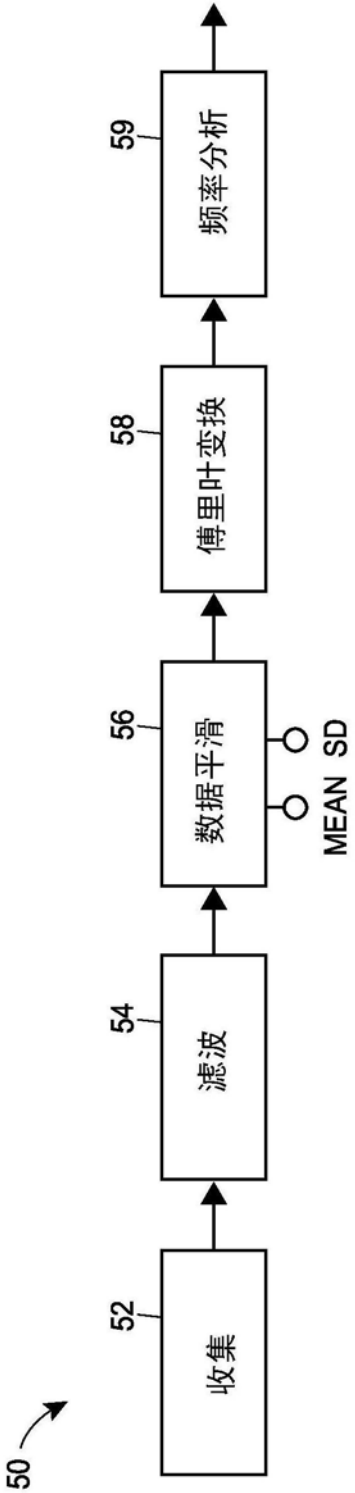


图2A

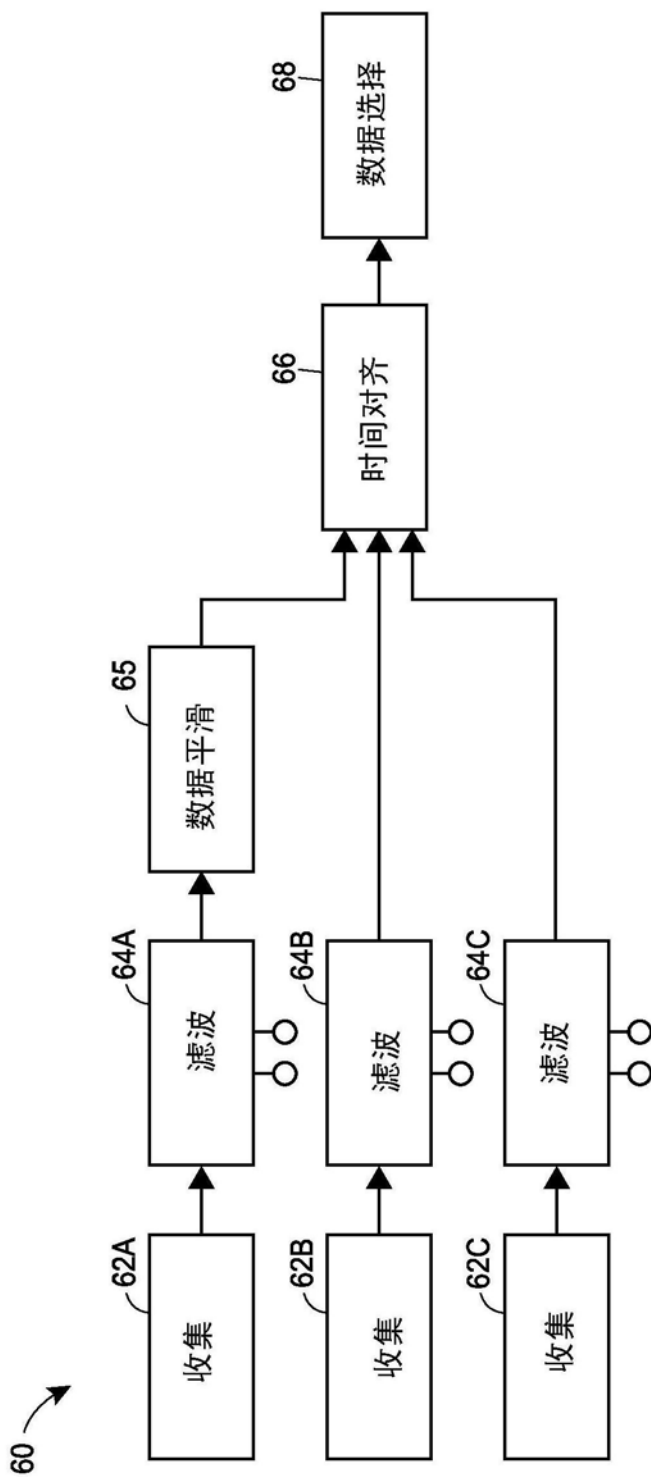


图2B

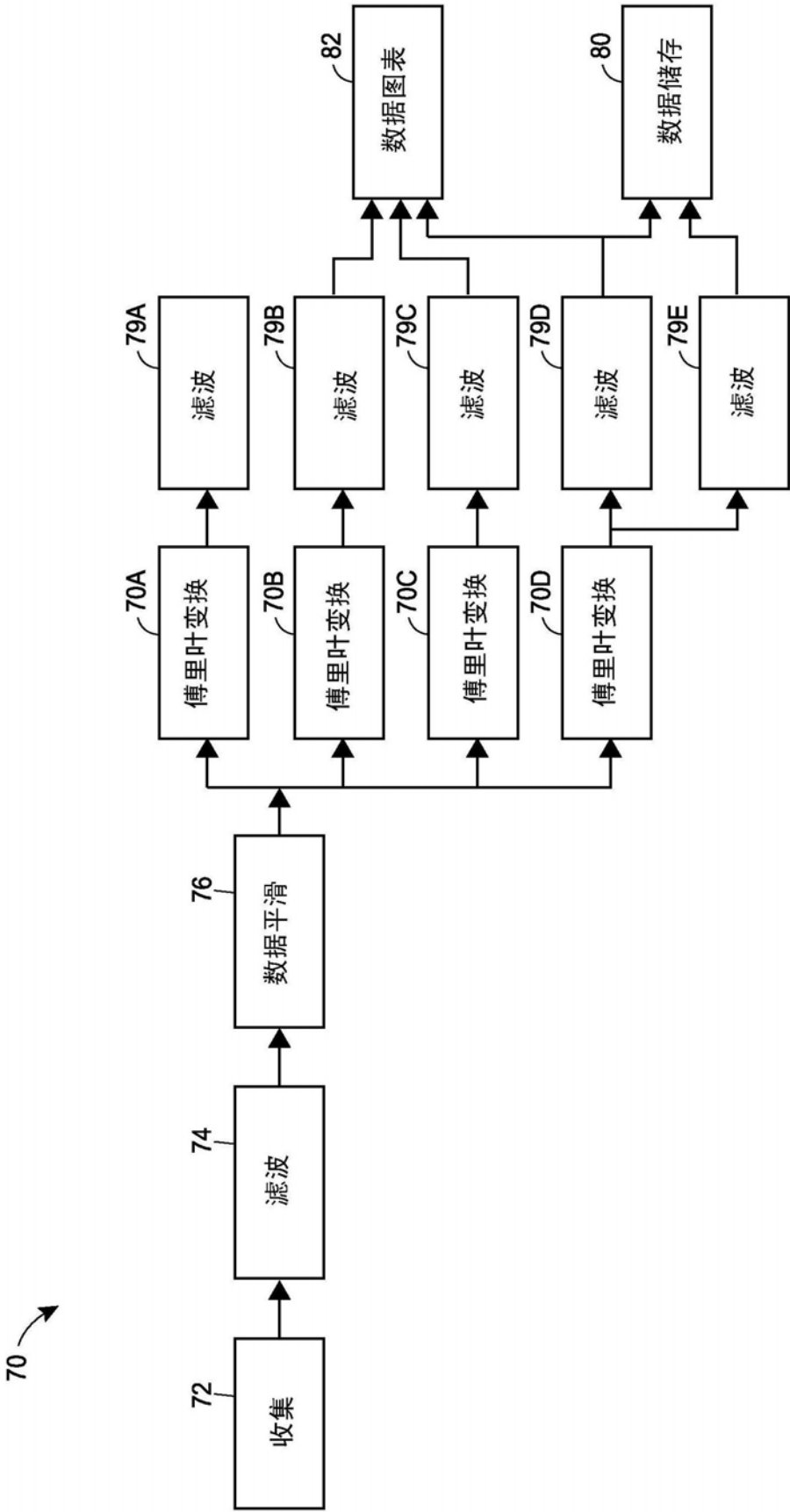


图2C

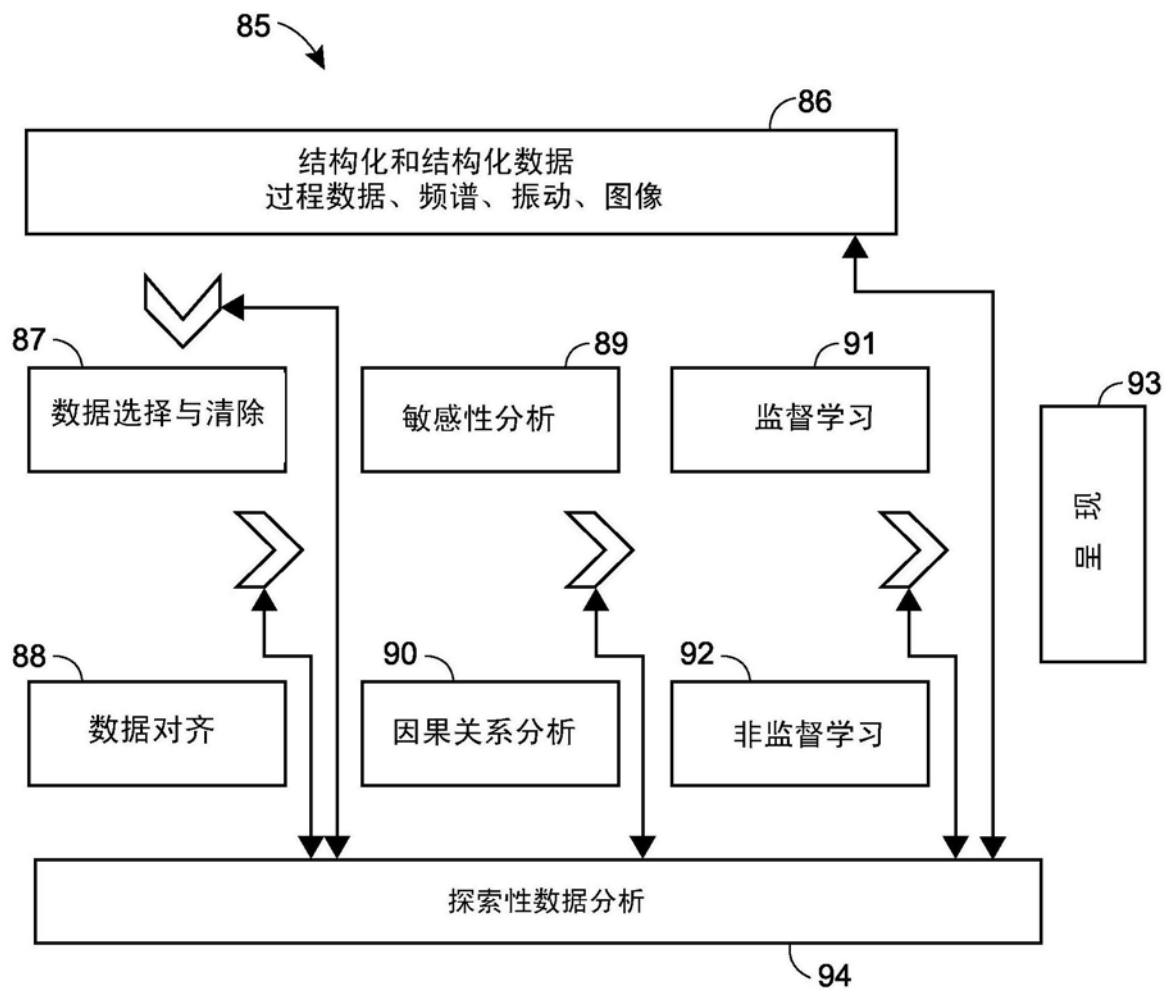


图3

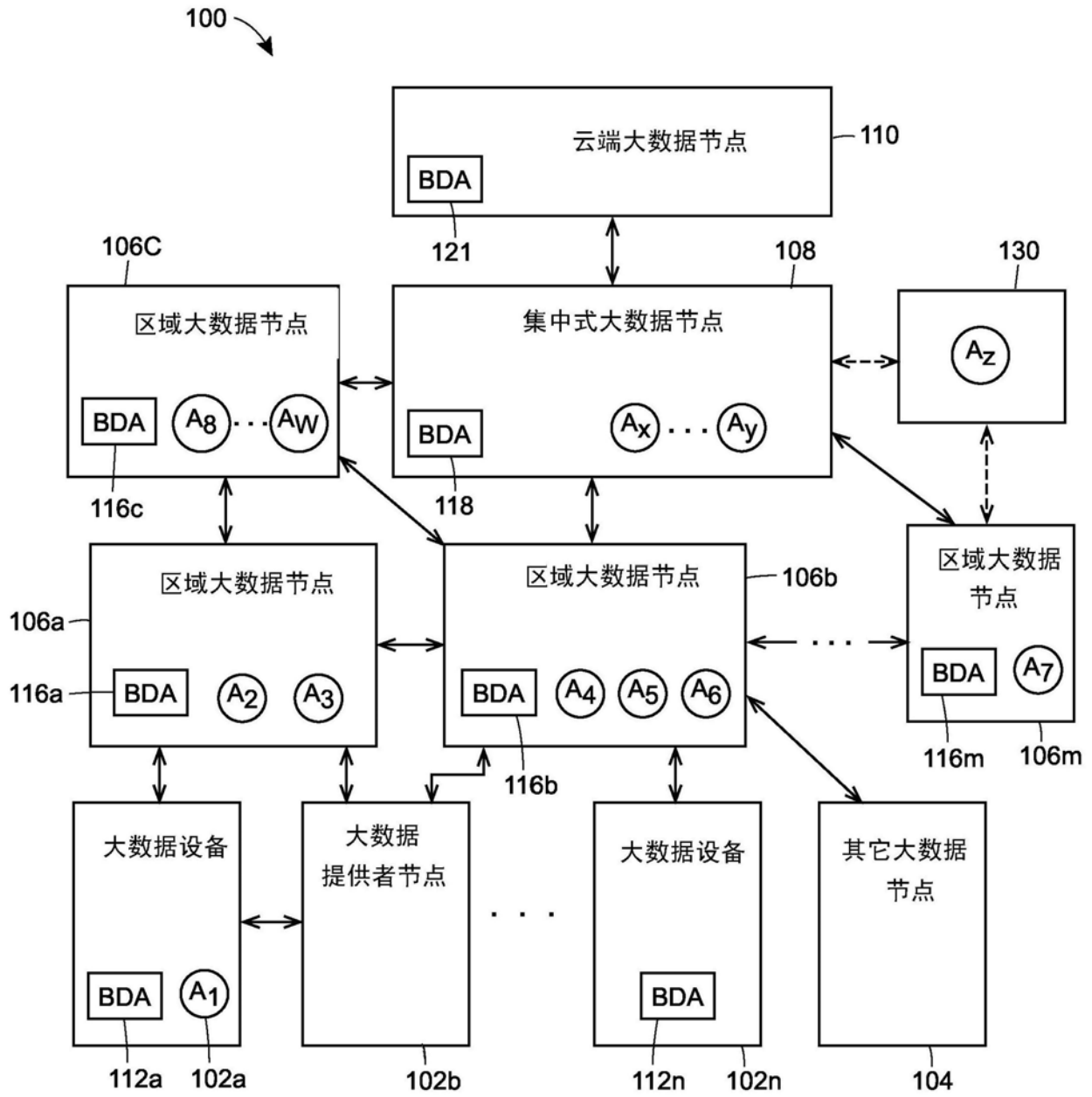


图4

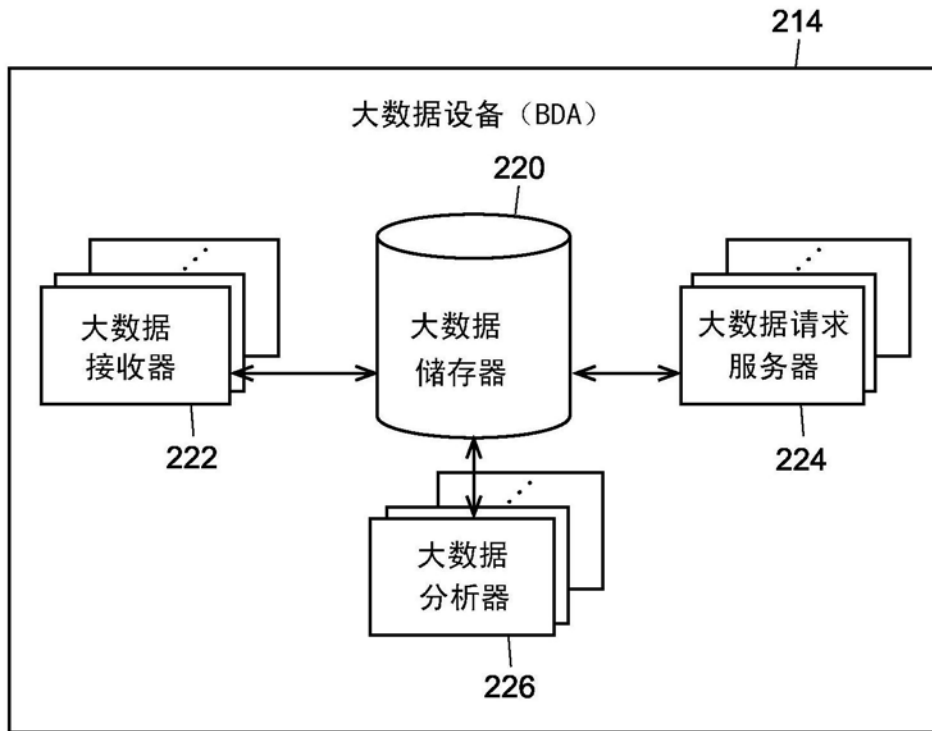


图5

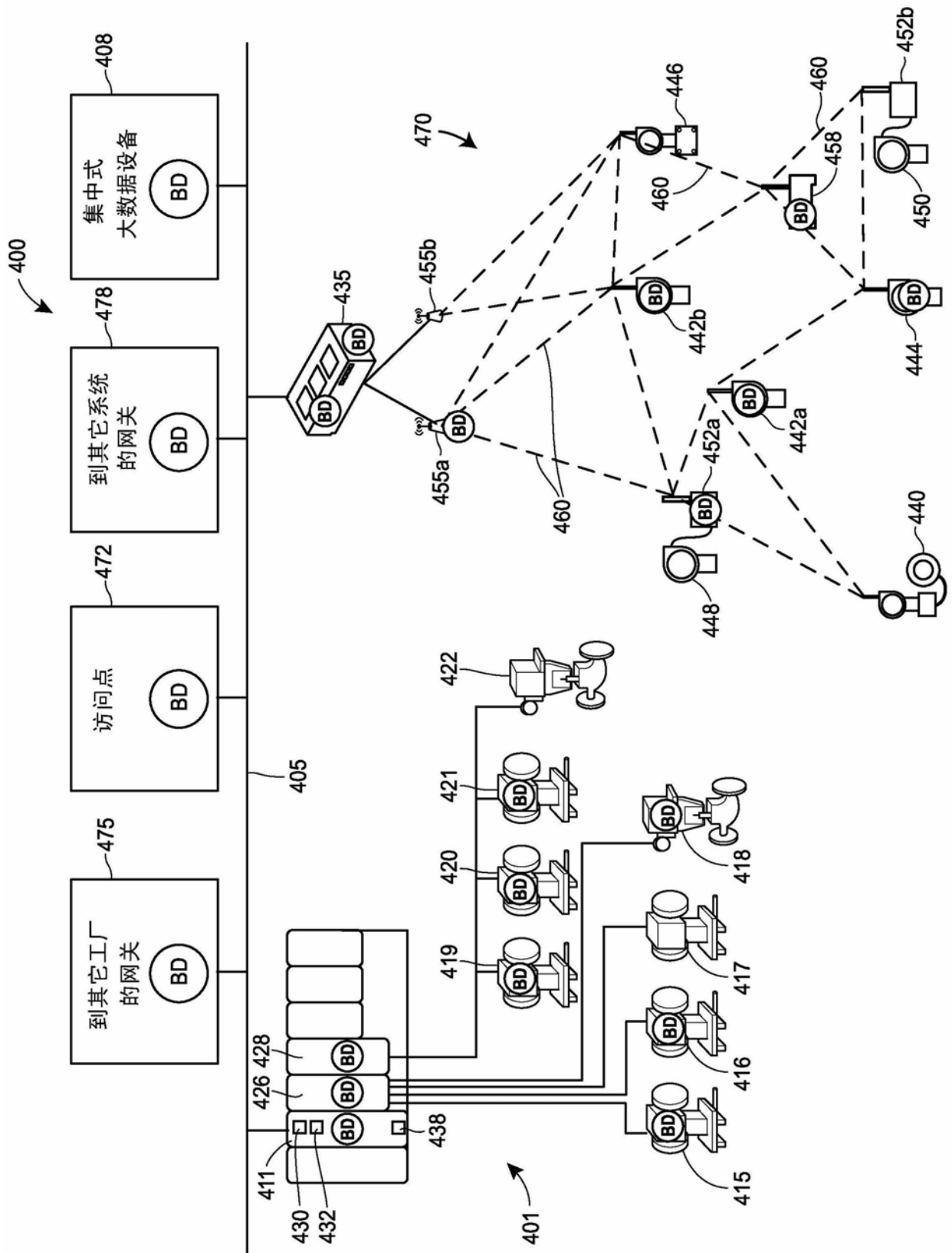


图6

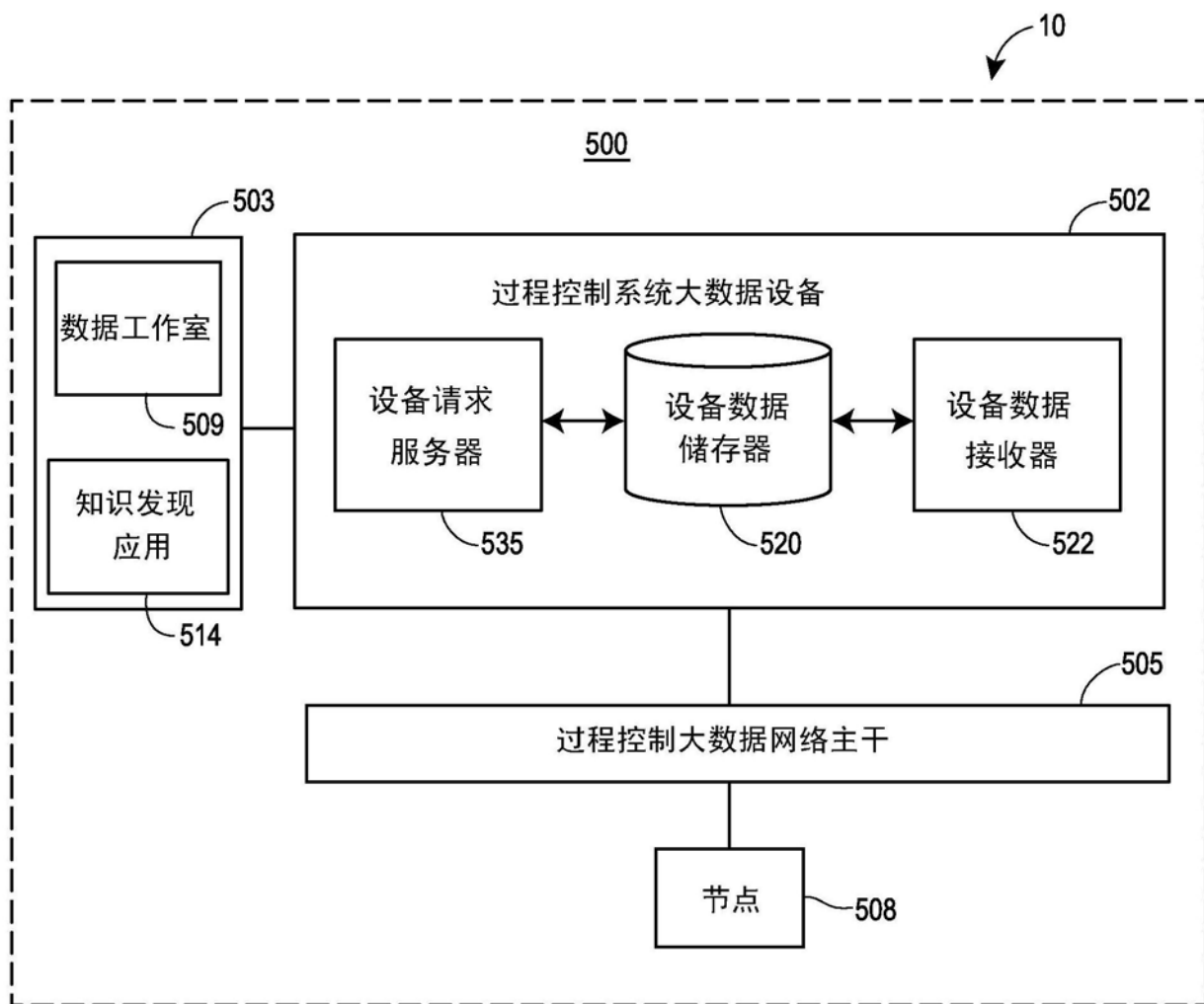


图7

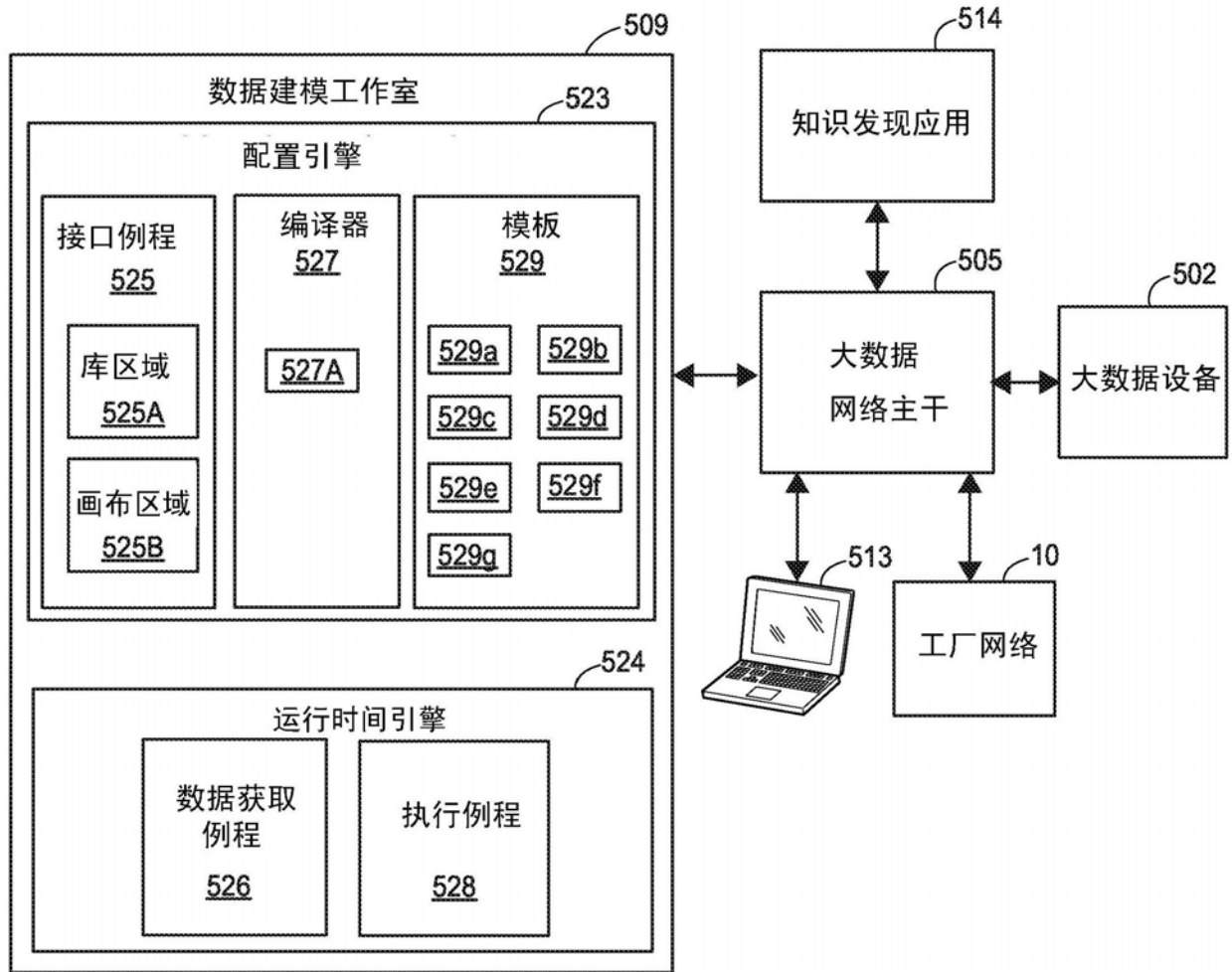


图8

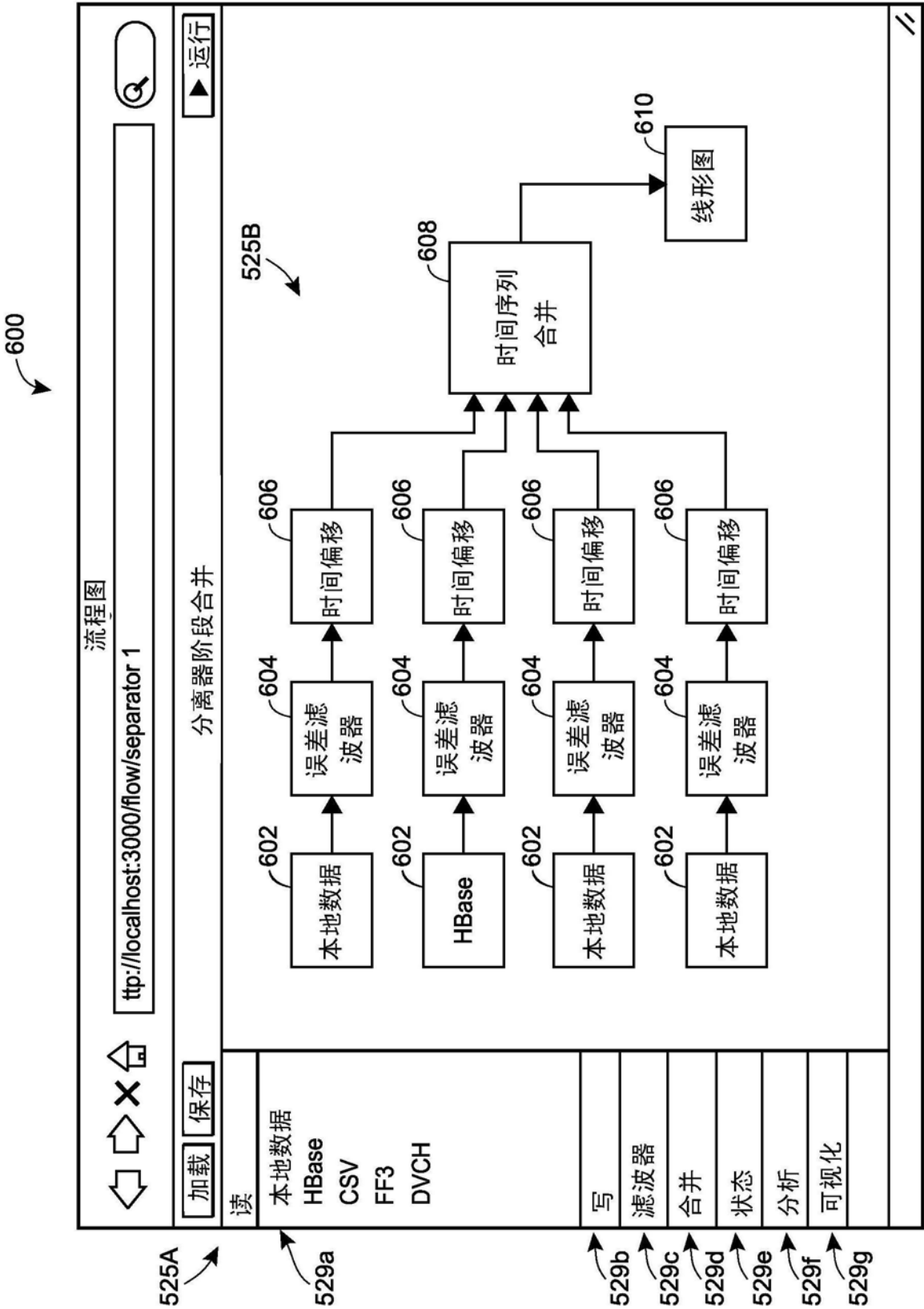


图9