



(12) 发明专利申请

(10) 申请公布号 CN 113272797 A

(43) 申请公布日 2021.08.17

(21) 申请号 201980088405.6

(72) 发明人 埃万盖洛斯·斯莫迪斯

(22) 申请日 2019.11.06

(74) 专利代理机构 北京安信方达知识产权代理有限公司 11262

(30) 优先权数据

代理人 李健

62/757,517 2018.11.08 US

62/799,697 2019.01.31 US

62/852,769 2019.05.24 US

62/875,919 2019.07.18 US

(51) Int.Cl.

G06F 16/22 (2006.01)

G06F 16/21 (2006.01)

G06N 20/00 (2006.01)

(85) PCT国际申请进入国家阶段日

2021.07.08

(86) PCT国际申请的申请数据

PCT/US2019/060094 2019.11.06

(87) PCT国际申请的公布数据

W02020/097221 EN 2020.05.14

(71) 申请人 埃万盖洛斯·斯莫迪斯

地址 美国加利福尼亚州

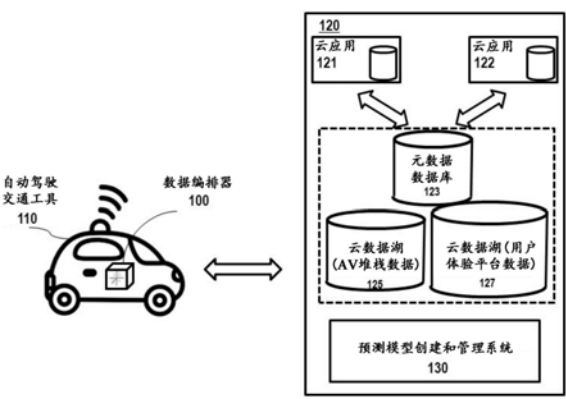
权利要求书4页 说明书29页 附图20页

(54) 发明名称

用于管理交通工具数据的系统和方法

(57) 摘要

本公开提供了用于管理自动驾驶交通工具数据的方法 and 系统。该方法可以包括：(a) 从自动驾驶交通工具收集所述自动驾驶交通工具数据，其中自动驾驶交通工具数据具有至少为1太字节的大小；(b) 处理所述自动驾驶交通工具数据以生成与所述自动驾驶交通工具数据相对应的元数据，其中所述自动驾驶交通工具数据被存储在数据库中；(c) 使用元数据的至少一部分从数据库中检索自动驾驶交通工具数据的子集，该自动驾驶交通工具数据的子集具有小于自动驾驶交通工具数据的大小；并且 (d) 存储或传输自动驾驶交通工具数据的子集。



1. 一种用于管理交通工具的交通工具数据的方法,包括:

(a) 从所述交通工具收集所述交通工具数据,其中所述交通工具数据具有至少为1太字节的大小;

(b) 处理所述交通工具数据以生成与所述交通工具数据相对应的元数据,其中所述交通工具数据被存储在数据库中;

(c) 使用所述元数据的至少一部分从所述数据库中检索所述交通工具数据的子集,所述交通工具数据的子集具有小于所述交通工具数据的大小;以及

(d) 存储或传输所述交通工具数据的所述子集。

2. 根据权利要求1所述的方法,其中所述交通工具数据至少包括由一个或多个传感器所捕获的传感器数据以及由所述交通工具上的一个或多个应用所产生的应用数据。

3. 根据权利要求2所述的方法,其中所述元数据还包括由所述一个或多个传感器中的传感器所生成的第一元数据或由所述一个或多个应用中的应用所生成的第二元数据。

4. 根据权利要求1所述的方法,其中处理所述交通工具数据包括:对准由所述交通工具的一个或多个传感器所收集的传感器数据。

5. 根据权利要求1所述的方法,还包括:将在(b)中处理的所述交通工具数据存储在所述数据库中。

6. 根据权利要求1所述的方法,其中(c)包括:使用所述元数据从所述数据库中检索所述交通工具数据的所述子集以用于训练预测模型,并且其中所述预测模型被用于管理所述交通工具数据。

7. 根据权利要求6所述的方法,其中所述预测模型可用于将所述交通工具数据从所述交通工具传输到远程实体。

8. 根据权利要求7所述的方法,还包括:使用所述预测模型将所述交通工具数据从所述交通工具传输到由所述数据编排器所管理的数据库。

9. 根据权利要求1所述的方法,还包括:接收来自用户的用于访问所述交通工具数据的请求,并且至少部分地基于所述请求来选择所述元数据的所述至少一部分。

10. 根据权利要求1所述的方法,其中所述交通工具是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具。

11. 一种用于管理交通工具的交通工具数据的系统,包括:

数据库;以及

可操作地耦合到所述数据库的一个或多个计算机处理器,其中所述一个或多个计算机处理器被单独地或共同地编程为(i)从所述交通工具收集所述交通工具数据,其中所述交通工具数据具有至少为1太字节的大小;(ii)处理所述交通工具数据以生成与所述交通工具数据相对应的元数据,其中所述交通工具数据被存储在所述数据库中;(iii)使用所述元数据的至少一部分从所述数据库检索所述交通工具数据的子集,所述交通工具数据的子集具有小于所述交通工具数据的大小;以及(iv)存储或传输所述交通工具数据的所述子集。

12. 根据权利要求11所述的系统,其中所述交通工具数据至少包括由一个或多个传感器所捕获的传感器数据以及由所述交通工具上的一个或多个应用所产生的应用数据。

13. 根据权利要求12所述的系统,其中所述元数据还包括由所述一个或多个传感器中的传感器所生成的第一元数据或由所述一个或多个应用中的应用所生成的第二元数据。

14. 根据权利要求11所述的系统,其中所述元数据是通过对准由所述交通工具的一个或多个传感器所收集的传感器数据而生成的。

15. 根据权利要求11所述的系统,其中所述元数据被用来从所述数据库中检索所述交通工具数据的所述子集以用于训练预测模型,并且其中所述预测模型被用于管理所述交通工具数据。

16. 根据权利要求15所述的系统,其中所述预测模型可用于将所述交通工具数据从所述交通工具传输到远程实体。

17. 根据权利要求16所述的系统,其中所述预测模型可用于将所述交通工具数据从所述交通工具传输到由所述系统所管理的所述数据库。

18. 根据权利要求11所述的系统,其中所述交通工具是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具。

19. 一种用于管理交通工具数据的数据编排器,包括:

数据存储库,所述数据存储库被配置为存储:(i)与请求所述交通工具数据的一个或多个子集的一个或多个远程实体相关的数据,以及(ii)与生成所述交通工具数据的所述一个或多个子集的一个或多个应用相关的数据,其中所述数据存储库位于收集或生成所述交通工具数据的所述交通工具的本地;

知识库,所述知识库被配置为存储基于机器学习的预测模型和用户定义的规则,用于确定数据传输规则,所述数据传输规则包括:(i)所述交通工具数据中要被传输的被选部分;(ii)何时传输所述交通工具数据的所述被选部分;以及(iii)用于接收所述交通工具数据的所述被选部分的所述一个或多个远程实体中的远程实体;以及

传输模块,所述传输模块被配置为基于存储在所述存储库中的所述数据和所述传输规则来传输所述交通工具数据的一部分。

20. 根据权利要求19所述的数据编排器,其中所述存储库、知识库和所述传输模块在所述交通工具上。

21. 根据权利要求19所述的数据编排器,其中所述一个或多个远程实体包括云应用、数据中心、第三方服务器或另一交通工具。

22. 根据权利要求19所述的数据编排器,其中所述数据存储库存储指示所述交通工具数据的所述一个或多个子集的可用性、传输定时延迟、所述关联的数据子集的数据类型、或者传输协议的数据。

23. 根据权利要求19所述的数据编排器,其中所述基于机器学习的预测模型被存储在模型树结构中。

24. 根据权利要求23所述的数据编排器,其中所述模型树结构表示基于机器学习的预测模型之间的关系。

25. 根据权利要求23所述的数据编排器,其中所述模型树结构的节点表示基于机器学习的预测模型,并且所述节点包括模型架构、模型参数、训练数据集或测试数据集中的至少一个。

26. 根据权利要求19所述的数据编排器,其中所述基于机器学习的预测模型由位于数据中心中的模型创建器生成。

27. 根据权利要求26所述的数据编排器,其中使用元数据和所述交通工具数据来训练

和测试所述基于机器学习的预测模型。

28. 根据权利要求26所述的数据编排器,其中所述模型创建器被配置为生成可用于所述交通工具的预测模型。

29. 根据权利要求19所述的数据编排器,其中所述知识存储可用于所述交通工具的预测模型。

30. 根据权利要求19所述的数据编排器,其中所述交通工具数据的所述被选部分包括交通工具数据的所述子集中的一个或多个的聚合。

31. 根据权利要求19所述的数据编排器,其中所述交通工具是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具。

32. 一种用于管理交通工具数据的方法,包括:

(a) 在云中,接收从交通工具传输的交通工具数据,其中所述交通工具数据至少包括传感器数据;

(b) 处理所述交通工具数据以生成与所述交通工具数据相对应的元数据,其中所述元数据包括由捕获所述传感器数据的传感器所生成的数据;以及

(c) 将所述元数据存储于元数据数据库中。

33. 根据权利要求32所述的方法,其中所述交通工具数据包括流数据和批处理数据。

34. 根据权利要求32所述的方法,其中所述交通工具数据包括应用数据。

35. 根据权利要求34所述的方法,其中所述元数据还包括与产生所述应用数据的应用相关的元数据。

36. 根据权利要求32所述的方法,其中所述交通工具数据由流水线引擎处理。

37. 根据权利要求36所述的方法,其中所述流水线引擎包括一个或多个功能组件。

38. 根据权利要求37所述的方法,其中所述一个或多个功能组件中的至少一个是经由用户界面从一组功能中选择的。

39. 根据权利要求37所述的方法,其中所述一个或多个功能组件中的至少一个被配置为创建场景数据对象,其中所述场景数据对象用于指定使用特定元数据的场景。

40. 根据权利要求32所述的方法,其中在(b)中处理的所述交通工具数据被存储在作为所述云的一部分的一个或多个数据库中。

41. 根据权利要求40所述的方法,还包括:使用存储在所述一个或多个数据库中的所述交通工具数据来训练预测模型。

42. 根据权利要求41所述的方法,其中所述预测模型被用于从所述交通工具中检索所述交通工具数据的至少一个子集。

43. 根据权利要求41所述的方法,其中所述元数据被用来从所述一个或多个数据库中检索所述交通工具数据的子集以用于训练所述预测模型。

44. 根据权利要求43所述的方法,还包括:根据所述预测模型的目标对所述交通工具数据的所述子集执行适当性分析。

45. 根据权利要求44所述的方法,还包括:基于所述适当性分析的结果来校正所述交通工具数据的所述子集。

46. 根据权利要求32所述的方法,其中所述元数据还包括与(b)中处理所述交通工具数据相关的元数据。

47. 根据权利要求32所述的方法, 其中所述元数据可用于检索所述交通工具数据的一个或多个子集。

48. 根据权利要求32所述的方法, 其中基于传输方案来传输所述交通工具数据的至少一部分, 并且其中基于来自所述云的请求来确定所述传输方案。

49. 根据权利要求32所述的方法, 其中所述交通工具是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具。

50. 一种用于管理交通工具的交通工具数据的系统, 包括:

数据库; 以及

可操作地耦合到所述数据库的一个或多个计算机处理器, 其中所述一个或多个计算机处理器被单独地或共同地编程为 (i) 接收从交通工具传输的交通工具数据, 其中所述交通工具数据至少包括传感器数据; (ii) 处理所述交通工具数据以生成与所述交通工具数据相对应的元数据, 其中所述元数据包括由捕获所述传感器数据的传感器所生成的数据; (iii) 将所述元数据存储于所述数据库中。

51. 根据权利要求50所述的系统, 其中所述交通工具数据包括流数据和批处理数据。

52. 根据权利要求50所述的系统, 其中所述交通工具数据包括应用数据。

53. 根据权利要求52所述的系统, 其中所述元数据还包括与产生所述应用数据的应用相关的元数据。

54. 根据权利要求50所述的系统, 其中所述交通工具数据由流水线引擎处理。

55. 根据权利要求54所述的系统, 其中所述流水线引擎包括一个或多个功能组件。

56. 根据权利要求55所述的系统, 其中所述一个或多个功能组件中的至少一个是经由用户界面从一组功能中选择的。

57. 根据权利要求55所述的系统, 其中所述一个或多个功能组件中的至少一个被配置为创建场景数据对象, 其中所述场景数据对象用于指定使用特定元数据的场景。

58. 根据权利要求50所述的系统, 其中所述一个或多个处理器被编程为使用存储在所述数据库中的所述交通工具数据来进一步训练预测模型。

59. 根据权利要求58所述的系统, 其中所述预测模型被用于从所述交通工具中检索所述交通工具数据的至少一个子集。

60. 根据权利要求50所述的系统, 其中所述元数据还包括与处理所述交通工具数据相关的元数据。

61. 根据权利要求50所述的系统, 其中所述元数据可用于检索所述交通工具数据的一个或多个子集。

62. 根据权利要求50所述的系统, 其中基于传输方案来传输所述交通工具数据的至少一部分, 并且其中基于请求来确定所述传输方案。

63. 根据权利要求50所述的系统, 其中所述交通工具是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具。

用于管理交通工具数据的系统和方法

交叉引用

[0001] 本申请要求2018年11月8日提交的美国临时专利申请第62/757,517号、2019年1月31日提交的美国临时专利申请第62/799,697号、2019年5月24日提交的美国临时专利申请第62/852,769号以及2019年7月18日提交的美国临时申请第62/875,919号的优先权,其全部内容均通过引用并入本文。

背景技术

[0002] 自动驾驶交通工具是可能能够在很少或没有用户输入的情况下感测其环境并进行导航的交通工具。自动驾驶交通工具系统可以使用诸如雷达、激光成像检测和测距(激光雷达)、图像传感器等感测设备来感测其环境。自动驾驶交通工具系统可以进一步使用来自全球定位系统(GPS)技术、导航系统、交通工具对交通工具通信、交通工具对基础设施技术和/或线控驱动系统的信息来对交通工具进行导航。

[0003] 单个高度自动化的交通工具或自动驾驶交通工具每小时可以生成1到5太字节(1-5TB)的原始数据。每天运行14到16个小时可能意味着生成多达每天每辆汽车50太字节(terabyte),或者每年每辆汽车20拍字节。一支由5,000辆高度自动化的交通工具组成的适度车队(仅纽约市就有14,000辆出租车)每年可能会生成超过100艾字节的原始数据。这样的数据可以例如由自动驾驶交通工具堆栈或自动化的交通工具堆栈生成,其可以包括所有支持任务,诸如通信、数据管理、故障安全以及中间件和软件应用。这样的数据还可以包括从交通工具之间的通信或从运输基础设施生成的数据。自动驾驶交通工具堆栈或自动化的交通工具堆栈可以将多个领域(诸如,感知、数据融合、云/空中(OTA)、本地化行为(又称为驾驶策略)、控制和安全性)整合到可以处理端到端自动化的平台中。例如,自动驾驶交通工具堆栈或自动化的交通工具堆栈可以包括各种运行时软件组件或基本软件服务,诸如感知(例如,专用集成电路(ASIC)、现场可编程门阵列(FPGA)、图形处理单元(GPU)加速器、单指令多数据(SIMD)存储器、传感器/检测器(诸如,相机、激光雷达、雷达、GPS等))、本地化和规划(例如,数据路径处理、双倍数据速率(DDR)存储器、本地化数据集、惯性测量、全球导航卫星系统(GNSS)、决策或行为(例如,运动引擎、纠错码(ECC)存储器、行为模块、仲裁、预测器)、控制(例如,锁步处理器、DDR存储器、安全性监视、故障安全后备、线控控制器)、连接性和输入/输出(I/O)(例如,射频(RF)处理器、网络交换机、确定性总线、数据记录)以及各种其他。这样的数据可以由一个或多个传感器和/或作为自动驾驶交通工具堆栈或自动化的交通工具堆栈的一部分的各种其他模块来生成。

发明内容

[0004] 大量的自动驾驶或自动化的交通工具数据可能是有价值的,并且可能需要针对成本、定时和隐私的不同优先级而在交通工具、边缘基础设施和云上下文中进行识别、选择、处理、传输和存储。

[0005] 在本文中认识到需要用于以安全、成本有效、可扩展并促进开放应用的方式来管

理自动驾驶或自动化的交通工具数据的方法和系统。

[0006] 本公开提供了用于管理交通工具数据的系统和方法。特别地,所提供的数据管理系统和方法可以被应用于与汽车价值链的各个方面相关的数据,包括例如交通工具设计、测试和制造(例如,自动驾驶交通工具的小批处理制造和产品化),车队的创建(涉及配置、订购服务、融资、保险和租赁车队),运营可能涉及服务、个性化、乘车管理和交通工具管理的车队,对交通工具进行维护、修理、加油和维修,以及处理这些交通工具或车队发生的事故和其他事件。如本文中所使用的,除非上下文另有说明,否则术语“交通工具数据”通常是指由任何类型的交通工具(诸如,互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具)生成的数据。如本文中所利用的术语“自动驾驶交通工具数据”通常是指由自动驾驶交通工具生成的数据。尽管已经关于自动驾驶交通工具描述了本公开的实施方式,但是应当了解,这些实施方式可以适用于或适于自动化的交通工具。

[0007] 在一些实施方式中,所提供的数据管理系统可以包括自动驾驶交通工具或自动化交通工具上的数据编排器(data orchestrator)。数据编排器可能能够编排和管理交通工具数据。在一些情况下,自动驾驶交通工具数据可以包括由自动驾驶交通工具堆栈所生成的数据(例如,由自动驾驶交通工具的传感器所捕获的数据)以及驾驶员和乘客数据。数据编排器可以被配置为确定交通工具数据的哪些(哪个部分)要被传送给哪个数据中心或第三方实体,以及何时传输此类数据。例如,一些自动驾驶交通工具数据可能需要立即或在自动驾驶交通工具运动时被传送,而其他数据可以在自动驾驶交通工具静止时(在等待下一个指派/任务或正被维护时)被传送。

[0008] 在一方面,提供了一种用于管理交通工具的交通工具数据的方法。该方法可以包括:(a)从交通工具收集交通工具数据,其中交通工具数据具有至少1TB的大小;(b)处理交通工具数据以生成与交通工具数据相对应的元数据(metadata),其中交通工具数据被存储在数据库中;(c)使用元数据的至少一部分从数据库中检索交通工具数据的子集,该交通工具数据的子集具有小于交通工具数据的大小;并且(d)存储或传输交通工具数据的子集。

[0009] 在一些实施方式中,该方法还包括:将在(b)中处理的交通工具数据存储在数据库中。在一些实施方式中,步骤(c)包括:使用元数据从数据库中检索交通工具数据的子集以用于训练预测模型,并且其中该预测模型被用于管理交通工具数据。在一些情况下,预测模型可用于将交通工具数据从交通工具传输到远程实体。例如,该方法还包括使用预测模型将交通工具数据从交通工具传输到由数据编排器所管理的数据库。

[0010] 在一些实施方式中,该方法还包括:接收来自用户的用于访问交通工具数据的请求;并且至少部分地基于该请求来选择元数据的至少一部分。在一些实施方式中,交通工具是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具。

[0011] 在另一方面,提供一种用于管理交通工具的交通工具数据的系统。该系统包括:数据库;以及可操作地耦合到数据库的一个或多个计算机处理器,其中该一个或多个计算机处理器被单独地或共同地编程为(i)从交通工具收集交通工具数据,其中交通工具数据具有至少为1TB的大小;(ii)处理交通工具数据以生成与交通工具数据相对应的元数据,其中交通工具数据被存储在数据库中;(iii)使用元数据的至少一部分从数据库中检索交通工具数据的子集,该交通工具数据的子集具有小于交通工具数据的大小;并且(iv)存储或传输交通工具数据的子集。

[0012] 在一些实施方式中,交通工具数据至少包括由一个或多个传感器所捕获的传感器数据和由交通工具上的一个或多个应用所产生的应用数据。在一些情况下,元数据还包括由一个或多个传感器中的传感器所生成的第一元数据或由一个或多个应用中的应用所生成的第二元数据。在一些实施方式中,通过对准由交通工具的一个或多个传感器所收集的传感器数据来生成元数据。在一些实施方式中,元数据被用来从数据库中检索交通工具数据的子集以用于训练预测模型,并且其中预测模型被用于管理交通工具数据。在一些情况下,预测模型可用于将交通工具数据从交通工具传输到远程实体。在一些情况下,预测模型可用于将交通工具数据从交通工具传输到由系统所管理的数据库。在一些实施方式中,交通工具是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具。

[0013] 本公开的另一相关但独立的方面提供了一种用于管理交通工具数据的数据编排器。数据编排器可以在自动驾驶交通工具上。数据编排器可以包括:数据存储库(data repository),被配置为存储:(i)与请求交通工具数据的一个或多个子集的一个或多个远程实体相关的数据,以及(ii)与生成交通工具数据的一个或多个子集的一个或多个应用相关的数据,其中数据存储库位于收集或生成交通工具数据的交通工具的本地;知识库(knowledge base),被配置为存储基于机器学习的预测模型和用户定义的规则,用于确定数据传输规则,该数据传输规则包括:(i)交通工具数据中要被传输的被选部分;(ii)何时传输交通工具数据的被选部分;以及(iii)用于接收交通工具数据的被选部分的一个或多个远程实体中的远程实体;以及传输模块,被配置为基于存储在存储库中的数据和传输规则来传输交通工具数据的一部分。

[0014] 在一些实施方式中,存储库、知识库和传输模块在交通工具上。在一些实施方式中,一个或多个远程实体包括云应用、数据中心、第三方服务器或另一交通工具。在一些实施方式中,数据存储库存储指示交通工具数据的一个或多个子集的可用性、传输定时延迟、关联的数据子集的数据类型或传输协议的数据。

[0015] 在一些实施方式中,基于机器学习的预测模型被存储在模型树结构中。在一些情况下,模型树结构表示基于机器学习的预测模型之间的关系。在一些情况下,模型树结构的节点表示基于机器学习的预测模型,并且该节点包括模型架构、模型参数、训练数据集或测试数据集中的至少一个。

[0016] 在一些实施方式中,基于机器学习的预测模型由位于数据中心的模型创建器生成。在一些情况下,使用元数据和交通工具数据来训练和测试基于机器学习的预测模型。在一些情况下,模型创建器被配置为生成可用于交通工具的预测模型。

[0017] 在一些实施方式中,知识存储可用于交通工具的预测模型。在一些实施方式中,交通工具数据的被选部分包括交通工具数据的子集中的一个或多个的聚合。在一些实施方式中,交通工具是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具。

[0018] 本公开的另一方面提供了一种用于管理交通工具数据的方法。该方法包括:(a)在云中,接收从交通工具传输的交通工具数据,其中交通工具数据至少包括传感器数据;(b)处理交通工具数据以生成与交通工具数据相对应的元数据,其中元数据包括由捕获传感器数据的传感器所生成的数据;并且(d)将元数据存储于元数据数据库中。

[0019] 在一些实施方式中,交通工具数据包括流数据和批处理数据。在一些实施方式中,交通工具数据包括应用数据。在一些情况下,元数据还包括与产生应用数据的应用相关的

元数据。

[0020] 在一些实施方式中,交通工具数据由流水线引擎处理。在一些情况下,流水线引擎包括一个或多个功能组件。例如,一个或多个功能组件中的至少一个是经由用户界面从一组功能中选择的。在一些情况下,一个或多个功能组件中的至少一个被配置为创建场景数据对象,其中该场景数据对象用于指定使用特定元数据的场景。

[0021] 在一些实施方式中,在 (b) 中处理的交通工具数据被存储在作为云的一部分的一个或多个数据库中。在一些情况下,该方法还包括:使用存储在一个或多个数据库中的交通工具数据来训练预测模型。在一些情况下,预测模型被用于从交通工具中检索交通工具数据的至少一个子集。在一些情况下,元数据被用来从一个或多个数据库中检索交通工具数据的子集以用于训练预测模型。该方法还包括:根据预测模型的目标对交通工具数据的子集执行适当性分析,并基于适当性分析的结果来校正交通工具数据的子集。

[0022] 在一些实施方式中,元数据还包括与处理 (b) 中的交通工具数据相关的元数据。在一些实施方式中,元数据可用于检索交通工具数据的一个或多个子集。在一些实施方式中,基于传输方案来传输交通工具数据的至少一部分,并且其中基于来自云的请求来确定传输方案。在一些实施方式中,交通工具是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具。

[0023] 本公开的另一方面提供一种用于管理交通工具的交通工具数据的系统。该系统包括:数据库;以及可操作地耦合到数据库的一个或多个计算机处理器,其中一个或多个计算机处理器被单独地或共同地编程为 (i) 接收从交通工具传输的交通工具数据,其中交通工具数据至少包括传感器数据;(ii) 处理交通工具数据以生成与交通工具数据相对应的元数据,其中元数据包括由捕获传感器数据的传感器所生成的数据;(iii) 将元数据存储于数据库中。

[0024] 在一些实施方式中,交通工具数据包括流数据和批处理数据。在一些实施方式中,交通工具数据包括应用数据。在一些情况下,元数据还包括与产生应用数据的应用相关的元数据。

[0025] 在一些实施方式中,交通工具数据由流水线引擎处理。在一些情况下,流水线引擎包括一个或多个功能组件。在一些情况下,一个或多个功能组件中的至少一个是经由用户界面从一组功能中选择的。在一些实例中,一个或多个功能组件中的至少一个被配置为创建场景数据对象,其中该场景数据对象用于指定使用特定元数据的场景。

[0026] 在一些实施方式中,一个或多个处理器被编程为使用存储在数据库中的交通工具数据来进一步训练预测模型。在一些情况下,预测模型被用于从交通工具中检索交通工具数据的至少一个子集。在一些实施方式中,元数据还包括与处理交通工具数据相关的元数据。在一些实施方式中,元数据可用于检索交通工具数据的一个或多个子集。在一些实施方式中,基于传输方案来传输交通工具数据的至少一部分,并且其中基于请求来确定传输方案。在一些实施方式中,交通工具是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具。

[0027] 本公开的另一方面提供了一种包括机器可执行代码的非暂时性计算机可读介质,该机器可执行代码在由一个或多个计算机处理器执行时,实现本文上面或本文其他地方的任何方法。

[0028] 本公开的另一方面提供了一种系统,该系统包括一个或多个计算机处理器和与其耦合的计算机存储器。该计算机存储器包括机器可执行代码,该机器可执行代码在由一个或多个计算机处理器执行时实现本文上面或本文其他地方的任何方法。

[0029] 通过下面的详细描述,本公开的其他方面和优点对于本领域技术人员将变得容易理解,其中,仅示出和描述了本公开的说明性实施方式。将会认识到,本公开内容能够具有其他和不同的实施方式,并且其若干细节能够在各种明显的方面进行修改,而所有这些都脱离本公开内容。因此,附图和描述本质上应被认为是说明性的,而不是限制性的。

援引并入

[0030] 本说明书中提到的所有出版物,专利和专利申请都以相同的程度通过引用并入本文,就好像每个单独的出版物,专利或专利申请被明确地并单独地指出通过引用并入一样。在通过引用并入的出版物和专利或专利申请与说明书中包含的公开内容相抵触的程度上,该说明书旨在取代和/或优先于任何此类矛盾的材料。

附图说明

[0031] 在所附权利要求中具体阐述了本公开的新颖特征。通过参考以下阐述了说明性实施方式的详细描述,将更好地理解本公开的特征和优点,其中利用了本公开的各种原理以及附图(在本文中也称为“图”),

其中:

[0032] 图1示意性地图示出了数据编排器和数据中心之间的数据流。

[0033] 图2示出了跨各种数据中心、自动驾驶交通工具和远程实体的数据流的示例。

[0034] 图3示出了应用表的示例。

[0035] 图4示意性地示出了与一个或多个远程实体通信的数据编排器的示例。

[0036] 图5图示出了预测模型知识库的示例。

[0037] 图6示出了模型创建器与元数据数据库和云数据湖进行交互以训练和开发预测模型的示例。

[0038] 图7图示出了创建预测模型的方法。

[0039] 图8示出了数据管理系统的示例组件。

[0040] 图9示出了数据摄取流水线和由流水线引擎所执行的功能的示例。

[0041] 图10示出了示例数据摄取过程。

[0042] 图11图示出了由诸如对准之类的数据处理所生成的元数据的示例,由应用和/或传感器生成的元数据的示例。

[0043] 图12示出了场景元数据的示例。

[0044] 图13示出了被编程或以其他方式被配置为实现数据管理系统的计算机系统。

[0045] 图14示出了自动化和自动驾驶交通工具的生命周期中的各种应用的示例。

[0046] 图15图示出了动态更新交通工具中的预测模型的示例。

[0047] 图16示意性地示出了根据本发明的一些实施方式的借助于OEM来管理的数据传输。

[0048] 图17示出了数据编排器与一个或多个云应用之间的数据传输过程。

[0049] 图18示意性地图示出了多层数据架构。

[0050] 图19示意性地示出了用于管理交通工具层与雾层之间以及雾层与云层之间的数据传输的数据编排器的示例。

[0051] 图20示出了可以在其中实现数据编排器的环境。

具体实施方式

[0052] 虽然本文已经示出和描述了各种实施方式,但是对于本领域技术人员而言容易理解的是,仅通过示例的方式提供了这样的实施方式。在不脱离本公开的情况下,本领域技术人员可以想到许多变化,改变和替换。应当理解,可以采用本文描述的实施方案的各种替代方案。

[0053] 如本文中所使用的,当被用于描述交通工具时,术语“自主控制”、“自动驾驶”、“自动驾驶”和“无人驾驶”通常是指本身能够执行至少一些或全部驾驶任务和/或沿着至少一部分路线监视驾驶环境的交通工具。自动驾驶交通工具可以是自动化的交通工具。这样的自动化的交通工具可以是至少部分或完全自动化的。自动驾驶交通工具可以被配置为在驾驶员或乘客的一些干预或不干预的情况下驾驶。自动驾驶交通工具可以从一个点行进到另一点,而无需来自该自动驾驶交通工具上的人员的任何干预。在一些情况下,自动驾驶交通工具可能是指具有国家公路交通安全管理局(NHTSA)定义中规定的交通工具自动化功能的交通工具,尤其是NHTSA定义中的第4级“交通工具上的自动驾驶系统(ADS)本身可以执行所有驾驶任务,并可以在特定情况下监控驾驶环境-本质上,可以进行所有驾驶。在这种情况下,人们无需关注”或NHTSA定义的第5级“交通工具上的自动驾驶系统(ADS)可以在所有情况下进行所有驾驶。乘员只是乘客,不需要参与驾驶。”自动化的交通工具还可以包括具有2级以上自动化驾驶能力的那些交通工具,其中AI被用来改善2级ADAS,而仍然需要一致的驾驶员控制。自动驾驶交通工具数据还可以包括由自动化的交通工具所生成的数据。

[0054] 自动驾驶交通工具可以被称为无人驾驶交通工具。自动驾驶交通工具可以是飞行器、陆地交通工具或横穿水体的交通工具。自动驾驶交通工具可以被配置为在任何合适的环境内移动,诸如在空中(例如,固定翼飞机、旋转翼飞机或者既没有固定翼也没有旋转翼的飞机)、在水中(例如,船舶或潜水艇)、在地上(例如,机动交通工具,诸如汽车、卡车、公共汽车、面包车、摩托车或火车),在地下(例如,地铁)、在太空中(例如,航天飞机、卫星或探测器)或这些环境的任何组合。

[0055] 每当术语“至少”,“大于”或“大于或等于”在一系列两个或更多个数值中的第一数值之前时,术语“至少”,“大于”或“大于或等于”适用于该系列数值中的每个数值。例如,大于或等于1、2或3等于大于或等于1,大于或等于2或大于或等于3。

[0056] 每当术语“不大于”,“小于”或“小于或等于”在一系列两个或更多个数值中的第一数值之前时,术语“不大于”,“小于”或“小于或等于”适用于该系列数值中的每个数值。例如,小于或等于3、2或1等于小于或等于3,小于或等于2或小于或等于1。

[0057] 如本文中所使用的,术语“实时”通常是指诸如计算机处理器的小于1秒、十分之一秒、百分之一秒、毫秒或更少的响应时间。实时还可以是指相对于第二事件的发生,第一事件的同时发生或基本同时发生。

[0058] 本公开提供了用于数据和知识管理(包括数据处理和存储)的方法和系统。本公开的方法和系统可以被应用于各种类型的交通工具,诸如互联交通工具、互联且自动化交通

工具、或者互联且自动驾驶交通工具。互联交通工具可以是指使用许多不同通信技术中的任何计数来与驾驶员、道路上的其他交通工具(交通工具对交通工具[V2V])、路边基础设施(交通工具对基础设施[V2I])以及“云”[V2C]进行通信的交通工具。本公开提供了可以在各种上下文中使用的数据编排器,包括交通工具(例如,自动驾驶交通工具)和非交通工具上下文。本公开的数据编排器可以被用于管理来自各种来源或用于诸如物联网(IoT)平台、网络物理软件应用和业务流程之类的各种用途的数据以及用于能源、制造、航空、汽车、化工、制药、电信、零售、保险、医疗保健、金融服务、公共部门等等中的组织。

数据管理系统

[0059] 本公开提供了用于管理诸如自动驾驶交通工具数据或自动化交通工具数据之类的交通工具数据的系统和方法。特别地,所提供的数据管理系统和方法可以被应用于与汽车价值链各个方面相关的数据,包括例如交通工具设计、测试和制造(例如自动驾驶交通工具的小批处理制造和产品化),车队的创建(涉及配置、订购服务、融资、保险和租赁车队),运营可能涉及服务、个性化、乘车管理和交通工具管理的车队,对交通工具进行维护、修理、加油和维修,以及处理这些交通工具或车队发生的事故和其他事件。数据管理系统可能能够以每小时至少约0.1太字节(TB)、0.5TB、1TB、2TB、3TB、4TB、5TB或更大规模的原始数据来管理和编排由车队所生成的数据。在一些实例中,数据管理可能能够以每小时至少约50TB、60TB、70TB、80TB、90TB、100TB的原始数据的规模来管理和编排由车队所生成的数据。在一些实例中,数据管理可能能够以每小时至少1吉字节(GB)、2GB、3GB、4GB、5GB或更大规模的原始数据来管理和编排由车队所生成的数据。数据管理系统可能能够管理和编排每小时高达0.5TB、1TB、2TB、3TB、4TB、5TB、50TB、60TB、70TB、80TB、90TB、100TB或更多数据的任意体量的数据。

[0060] 在一些实施方式中,数据和知识管理系统可以与驻留在自动驾驶或自动化交通工具上的数据编排器通信。数据编排器可能能够管理交通工具数据。数据编排器可以是数据路由器。数据编排器可以被配置为以智能方式将交通工具数据路由到数据和知识管理系统。数据编排器可以被配置为确定要将哪个自动驾驶交通工具数据或自动驾驶交通工具数据的哪个部分传送给哪个数据中心或第三方实体的数据和知识管理系统,以及何时传输自动驾驶交通工具数据的这部分。例如,一些自动驾驶交通工具数据可能需要立即被传送或在自动驾驶交通工具行驶时被传送,而其他数据可能在自动驾驶交通工具静止时(在等待下一个指派/任务或正被维护时)被传送。所提供的数据管理系统还可以包括预测模型创建和管理系统,其被配置为训练或开发预测模型,以及将模型部署到数据编排器和/或自动驾驶交通工具堆栈的组件或者自动化交通工具堆栈的组件。在一些情况下,预测模型创建和管理系统可以驻留在远程实体(例如,数据中心)上。所提供的数据管理系统还可以包括数据和元数据管理系统,其被配置为存储和管理由自动驾驶交通工具所生成的数据和关联的元数据,并且处理针对该数据和元数据所发出的查询和API调用。数据编排器或数据和知识管理系统可以作为独立系统来实现或提供。

[0061] 图1示意性地图示出了数据编排器100与数据中心120之间的数据流。在一些情况下,数据编排器100可以被配置为使数据管理过程自动化,包括例如数据创建、数据清理、数据充实、以及跨数据中心、系统和第三方实体而递送数据。从自动驾驶交通工具110收集的数据可以包括由自动驾驶交通工具的传感器所捕获的数据。此类传感器可以包括例如导航

系统、交通工具上的传感器诸如激光成像检测和测距(激光雷达)、雷达、声纳、差分全球定位系统(DGPS)、惯性测量单元(IMU)、陀螺仪、磁力计、加速度计、超声波传感器、图像传感器(例如,可见光、红外)、热传感器、音频传感器、振动传感器、电导率传感器、化学传感器、生物传感器、辐射传感器、电导率传感器、接近度传感器或任何其他类型的传感器或其组合。从自动驾驶交通工具收集的数据还可以包括车队数据(例如来自交通工具操作系统的数据)、驾驶员数据(例如驾驶员情绪、驾驶员机敏性、驾驶风格等)、乘客数据(例如来自用户体验平台的数据诸如访问音乐、游戏、来自诸如移动应用之类的用户设备的数据)以及各种其他数据。由交通工具的数据编排器所发送的数据由驻留在数据中心中的数据和知识管理系统接收。数据可以包括来自一个或多个传感器的数据流(例如,摄像机的输出、传感器融合数据)、批处理数据和/或个体记录(例如,乘客在被运输时进行的购买交易、由交通工具子系统(例如,监视交通工具引擎健康的系统或监视轮胎状况的子系统)所产生的个体记录或一系列记录,这是用户与交通工具上运行的应用之一进行交互的结果)。

[0062] 在一些实施方式中,数据编排器100可以是边缘智能平台。例如,数据编排器可以是基于雾或边缘计算概念的基于软件的解决方案,其将数据处理和编排扩展为更靠近边缘(例如,自动驾驶交通工具)。尽管边缘计算可以指的是将服务实例化的位置,但是雾计算可能意味着在最终用户或最终节点控制下的通信、计算以及存储资源和服务在设备和系统上或附近(例如,在5米之内或1米之内)的分布。保持与边缘设备(例如,自动驾驶交通工具、传感器)的距离非常近,而不是将所有数据发送到遥远的集中式云中,从而将时延最小化,实现了最大化的性能、更快的响应时间以及更有效的维护和操作策略。它还显著地降低了总体带宽需求和管理广泛分布的网络的成本。所提供的数据管理系统可以采用边缘智能范式,其可以在边缘处执行至少一部分数据处理。在一些实例中,机器学习模型可以在云上构建和训练,并在边缘设备或边缘系统(例如,硬件加速器)上运行。本公开的系统和方法可以提供一种高效且高度可缩放的边缘数据编排平台,其实现了实时、现场交通工具数据编排。

[0063] 数据管理系统的软件堆栈可以是在边缘和云上运行的服务的组合。在边缘上运行的软件或服务可以采用预测模型以进行数据编排。在云上运行的软件或服务可以提供用于训练、开发和管理预测模型的预测模型创建和管理系统130。在一些情况下,数据编排器可以支持传感器数据到本地存储库(例如,本地时间序列数据库)中的摄取、数据清理、数据丰富(例如,将第三方数据与已处理数据合并)、数据对准、数据注释、数据标注或数据聚合。原始数据可以在持续时间(例如,大约1、2、3、4、5、6、7、8、9、10秒,大约1、2、3、4、5、6、7、8、9、10分钟,大约1、2、3、4、5、6、7、8、9、10小时等)上被聚合。可替代地或另外,原始数据可以跨数据类型或数据源而被聚合并作为包而被发送到远程实体。

[0064] 数据编排器可以跨数据中心、云应用或驻留在数据中心中的任何组件(例如,与第三方实体相关联)而递送数据。数据编排器可以确定要向哪个数据中心和/或实体传输哪个数据或数据的哪个部分以及何时传输数据的这部分。例如,一些自动驾驶交通工具数据(例如,数据包或数据的第一部分)可能需要立即被传送或在自动驾驶交通工具行驶时被传送,而其他数据(例如,数据包或数据的第二部分)可能在自动驾驶交通工具静止时(在等待下一个指派/任务或正被维护时)被传送。在另一示例中,数据的第一部分可以被传输到托管用于基于实时数据提供实时反馈和控制的车队管理器应用的数据中心,而数据的第二部分(例如,批处理数据)可以被传输到保险公司服务器以基于批处理数据来计算保险范围。在

一些实施方式中,可以至少部分地基于预测模型和/或手工制定的规则来确定数据递送或数据传输。在一些实施方式中,可以基于预测模型、手工制定的规则以及存储关于目的地和传输协议的数据的存储库来发起数据传输。在示例中,数据编排器100可以支持用于数据聚合的服务,以及用于将聚合的数据发送到云、不同的数据中心或实体以进行进一步分析的数据发布的服务。稍后将在本文中描述有关数据编排器和预测模型的详细信息。

[0065] 预测模型创建和管理系统130可以包括在云或本地部署 (on-premises) 环境中运行以远程配置和管理数据编排器100的服务或应用。该环境可以运行在一个或多个公共云中 (例如,Amazon Web Services (AWS)、Azure等) 和/或在混合云配置中,在混合云配置中系统的一个或多个部分运行在私有云中,而其他部分运行在一个或多个公共云中。例如,预测模型创建和管理系统130可以被配置为训练和开发预测模型。在一些情况下,可以通过预测模型更新模块将经训练的预测模型部署到数据编排器或边缘基础设施。关于预测模型更新模块的细节相关于图15进行了描述。在一些情况下,预测模型创建和管理系统130还可能能够在云中开发的机器学习模型翻译成可以在边缘处执行的传感器表达式。预测模型创建和管理系统130还可以支持将从自动驾驶交通工具传输的数据摄取到一个或多个数据库或云存储123、125、127中。预测模型创建和管理系统130可以包括允许集成的管制和管理的应用,包括监视或存储在云中或在私有数据中心处的数据。在一些实施方式中,预测模型创建和管理系统130可以包括用于查看分析、传感器数据 (例如,视频) 的用户界面 (UI) 模块,或者包括用于开发和部署分析表达式、将数据编排应用部署到边缘 (例如,自动驾驶交通工具操作系统、边缘网关、边缘基础设施、数据编排器)、监视预测模型性能以及配置预测模型的管理UI。注意,尽管将预测模型创建和管理系统示为数据中心的组件,但是预测模型创建和管理系统可以是独立系统。

[0066] 所提供的数据管理系统可以采用任何合适的技术,诸如容器和/或微服务。例如,数据编排器的应用可以是容器化应用。数据管理系统可以在边缘处的软件基础设施中部署基于微服务的架构,诸如在容器中实现应用或服务。在另一示例中,云应用和/或预测模型创建和管理系统130可以提供由微服务支持的管理控制台或云分析。

[0067] 容器技术使由操作系统 (OS) 管理的计算机服务器资源 (如存储器、中央处理器 (CPU) 和存储) 虚拟化,其开销可忽略不计,而无需为每个租户复制整个OS内核 (因此与hypervisor技术不同)。容器是作为流行的Linux开源操作系统的一部分而被开发的,并且随着像Docker和CoreOS之类的高级管理框架的可用,容器在软件开发和数据中心操作 (“DevOps”) 中获得了巨大的吸引力。诸如Kubernetes之类的其他容器编排框架也可以被使用。Kubernetes提供了称为“pod”的高级抽象层,其使得多个容器能够在主机上运行并共享资源,而没有冲突的风险。Pod可以被用来定义共享服务 (如目录或存储),并将其暴露给Pod中的所有容器。在物联网 (IoT) 用例 (其包括工厂、仓库、零售店和其他设施等物理位置) 中,在与物理传感器网络非常靠近的近线计算基础设施上消耗软件和分析以用于处理传感器数据的需求与日俱增。这些计算节点例如包括从中型 (例如,双核处理器和4GB存储器) 到小型 (例如,具有小于1GB存储器的单核处理器核心) 的服务器,其连接到互联网并且可以访问在操作中部署的各种异构传感器设备和控制系统。数据管理系统提供了用于在这些边缘计算基础设施设置中智能地部署和管理容器技术的方法。

[0068] 数据中心或远程实体120可以包括用于存储自动驾驶交通工具数据和元数据的一

个或多个存储库或云存储。例如,数据中心120可以包括元数据数据库123、用于存储自动驾驶交通工具堆栈数据125的云数据湖、以及用于存储用户体验平台数据127的云数据湖。本文所述的用户体验平台可以包括硬件和/或在交通工具的车厢内操作的软件组件。用户体验平台可以被配置为管理车厢的环境和乘员的交互,例如车厢温度、每个乘员的娱乐选择、每个乘员的生命体征、情绪和机敏性等。在一些情况下,元数据数据库123和/或云数据湖可以是云存储对象。

[0069] 自动驾驶交通工具堆栈可以将多个领域(诸如感知、数据融合、云/OTA、本地化、行为(又称为驾驶策略)、控制 and 安全性)整合到可以处理端到端自动化的平台中。例如,自动驾驶交通工具堆栈可以包括各种运行时软件组件或基本软件服务,诸如感知(例如,ASIC、FPGA、GPU加速器、SIMD存储器、传感器/检测器(诸如相机、激光雷达、雷达、GPS等))、本地化和规划(例如数据路径处理、DDR存储器、本地化数据集、惯性测量、GNSS)、决策或行为(例如运动引擎、ECC存储器、行为模块、仲裁、预测器)、控制(例如锁步处理器、DDR存储器、安全性监视、故障安全后备、线控控制器)、连接性和I/O(例如RF处理器、网络交换机、确定性总线、数据记录)。自动驾驶交通工具堆栈数据可以包括如上所述由自动驾驶堆栈所生成的数据。用户体验平台数据127可以包括与诸如数字服务(例如,对音乐、视频或游戏的访问)、交易以及乘客商务或服务之类的用户体验应用相关的数据。例如,用户体验平台数据可以包括与以下相关的数据:订阅访问内容,例如音乐流媒体服务,新闻服务,礼宾服务等;在运输过程中以及交通工具间歇停车时(例如在加油站,饭店,咖啡店等) (例如,充电站运营商,例如能源公司,可以与一家咖啡连锁店合作,向在加油时购买的乘客提供咖啡饮料折扣) 商品,服务和内容的基于交易的购买;以及忠诚度积分的兑换,例如,交通工具制造商和车队运营商可以使用类似于航空公司或酒店连锁店的系统来奖励客户的忠诚度,在该系统中,使用此类程序可以用与这些行业和其他行业非常相似的方式兑换忠诚度积分。在一些情况下,用户体验平台数据127还可以包括第三方合作伙伴数据,诸如由用户移动应用所生成的数据。用户可以是车队运营商或乘客。

[0070] 云应用121、122可以针对各种用例进一步处理或分析从自动驾驶交通工具传输的数据。云应用可以允许在诸如原始装备制造商(OEM)、酒店和招待所、餐厅和餐饮、旅游和娱乐、医疗保健、服务交付和各种其他之类的行业中使用无人操纵/无人驾驶交通工具的一系列用例。特别地,所提供的数据管理系统和方法可以被应用于与汽车价值链的各个方面相关的数据,包括例如交通工具设计、测试和制造(例如自动驾驶交通工具的小批处理制造和产品化),车队的创建(涉及配置、订购服务、融资、保险和租赁车队),运营可能涉及服务、个性化、乘车管理和交通工具管理的车队,对交通工具进行维护、修理、加油和维修,以及处理这些交通工具或车队发生的事故和其他事件。

[0071] 图2示出了跨各种数据中心、自动驾驶交通工具和实体的数据流的示例。如示例中所示,由使用自动驾驶交通工具车队的消费者202在自动驾驶交通工具车队(AV车队) 201期间所生成的数据可以借助于数据编排器220被传输到各种远程实体。各种远程实体可以包括例如政府211、车队租赁公司209、保险公司207、车队管理器205、车队运营商203、数字服务217、其他运输服务219(例如诸如火车和班车、拼车、即时用车服务、共享旅程或私人旅行、步行、自行车、电动踏板车、出租车等)、平台提供商215和原始装备制造商(OEM) 213。传输给各个实体的数据可以相同也可以不同。例如,传输到数字服务217的数据(例如,包括更

多的消费者相关数据或乘客相关数据)可以不同于传输到车队运营商203的数据(例如,包括更多的AV车队数据或传感器数据)。数据可以在不同的时间点和/或频率而被传输到各个实体。例如,传感器数据流可以实时地或在交通工具行驶时被发送到车队管理器205或平台提供商215,而包括批处理数据的消息包可以在交通工具处于静止或处于较低频率时被发送到政府211或车队租赁公司209。在一些实施方式中,应用存储库或应用表可以被用来存储与交通工具/交通工具应用与远程实体/云应用之间的数据传输相关的信息。应用表可以是本文其他地方描述的数据编排器的组件。在一些情况下,在云或远程实体上运行的一个或多个云应用(例如,云应用121、122或租户应用)可以在应用表中注册。云应用可以被链接到在自动驾驶交通工具或自动驾驶交通工具的操作系统上运行的一个或多个应用(例如,边缘应用、本地应用)。应用表可以存储与如下内容相关的数据:云应用感兴趣的特定交通工具数据(例如,数据类型、指向要传输的数据的指针),生成特定交通工具数据的应用(例如,在自动驾驶交通工具上运行的应用),特定数据将被传输到的应用和/或数据中心(例如,云应用121、122,应用在服务器上的位置),数据传输方案或协议(例如,传输的定时诸如延迟时间或频率、通信协议、被用于传输的压缩或加密方法)以及各种其他内容(例如,关于数据被传输之前的隐私的监管规则)。

[0072] 数据编排器220也可以是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具的一部分。图20示出了在其中可以实现数据编排器的环境。交通工具2001-1、2001-2可以是私有的,也可以是车队的一部分。交通工具可以被用于乘客运输、长途或短途物流、最后一英里递送(例如,在5英里、4英里、3英里、2英里或1英里之内的递送)或者具有混合使用(例如乘客和包裹)。注意,上面提及的数据可以作为任何其他合适的数据结构而被存储。在一些情况下,应用表可以被存储在本地存储中,并由数据编排器进行管理。另外或可替代地,应用表可以由数据编排器以及预测模型创建和管理系统两者来管理。

[0073] 在一些情况下,在云或远程实体(例如,诸如Amazon Web Services (AWS) 和Azure之类的公共云或私有云)上运行的应用可以通过发布/订阅方案在特定交通工具的数据编排器(或车队的编排器)的应用表中注册。在一些情况下,在雾/边缘服务器或远程实体上运行的应用程序可以通过发布/订阅方案在“应用表”中注册。在一些实施方式中,注册应用可以指定其需要从其中接收数据的交通工具ID和/或在其需要从其中接收数据的对应交通工具上运行的(一个或多个)特定交通工具应用。

[0074] 在一些实施方式中,由注册应用生成的数据请求可以由云的订阅模块来组织和管理,并且数据请求可以经由消息通过空中(OTA)而被传送到一个或多个相关交通工具。消息可以包括对一个或多个交通工具应用的一个或多个请求。在一些情况下,包括在由交通工具所接收的消息中的请求可以被注册在应用表中。订阅模块可以被配置为管理数据请求或注册应用请求。例如,订阅模块可能能够聚合多个注册应用请求,从而有益地减少通信带宽消耗。例如,关于从同一交通工具应用(例如,Pothole检测器应用)请求数据的多个注册应用请求可以被聚合。在其他示例中,关于从特定交通工具组上运行的不同交通工具应用请求数据的多个注册应用请求(例如,在2010-2015年间制造的所有BMW Model 3交通工具)可以被聚合并打包成单个消息。

[0075] 图3示出了应用表300的示例。应用表300可以是数据编排器的一部分(例如,参见图4)。如上所述,应用表300的条目可以存储数据。例如,应用表的行可以包括在自动驾驶交

通工具中运行的应用(即,交通工具应用)的名称(例如,Pothole检测器)。该应用可以生成要被传送到远程实体、数据中心或云应用的数据。应用表的行还可以包括:指示新数据是否可用于传输到在特定数据中心中运行的一个或多个应用的标志(例如,传输标志),在其中生成数据的交通工具的标识符(例如,交通工具ID),云应用(例如,应用名称、应用在云上的位置)或从此应用请求数据的数据中心以及要将数据发送到的位置(例如,app_name,loc),要被传输的数据的类型(例如,视频流、CAN数据),指向要从交通工具传输的实际数据(例如,Stream1)的指针,传输时间(例如,传输定时延迟),压缩类型(例如,无损),加密类型(例如RSA)和监管规则。注意,所图示的应用表仅仅是示例。与数据传输相关的任何其他数据都可以被包括在“应用表”中。

[0076] 在一些情况下,可以由本地/交通工具应用设置一个或多个条目。例如,可以由本地/交通工具应用设置指示被请求的数据是否可用于传输的传输标志。在一些情况下,可以由数据编排器设置一个或多个条目。例如,可以由数据编排器设置交通工具ID或监管规则。

[0077] 在一些实施方式中,云数据湖可以组织车队中的每个交通工具周围的数据。例如,来自特定的AV堆栈和特定的用户体验平台的数据可以与对应的交通工具(例如,交通工具ID)相关联地组织和存储。如上所述,交通工具可以注册在云数据湖中,并可以通过如下信息来识别:其交通工具ID、其所使用的各种数据采集应用、由每个数据采集应用所访问的传感器、每个传感器的性能(例如,传感器可以每5秒捕获一次数据,或者传感器可以捕获720p分辨率的视频)等等。在一些情况下,用户、网络中的实体或注册在系统中的一方可以被允许使用交通工具ID来自动导出附加的信息,诸如交通工具、交通工具的品牌、型号和制造年份。在一些情况下,交通工具可以是向数据管理系统注册的车队的一部分(例如,公司车队、车队、租车公司、由特定OEM制造的私人拥有的交通工具集合)。

数据编排器

[0078] 数据编排器可以在自动驾驶交通工具本地或在自动驾驶交通工具上。在一些示例中,数据编排器驻留于自动驾驶交通工具上。如上所述,数据编排器也可以是互联交通工具、互联且自动化交通工具、或者互联且自动驾驶交通工具的一部分。所提供的数据管理系统可以采用在边缘或边缘网关处执行数据编排的边缘智能范式。在一些实例中,一个或多个机器学习模型可以在云/数据中心上建立和训练,并在交通工具或边缘系统(例如,硬件加速器)上运行。

[0079] 在一些情况下,可以部分地使用边缘计算平台或边缘基础设施/系统来实现数据编排器。边缘计算平台可以以软件、硬件、固件、嵌入式硬件、独立硬件、专用硬件或其任何组合来实现。本文所描述的数据编排器及其组件、边缘计算平台和技术可以以数字电子电路系统、集成电路系统、专门设计的ASIC(专用集成电路)、计算机硬件、固件、软件和/或其组合来实现。这些系统、设备和技术可以包括在一个或多个计算机程序中的实现,该一个或多个计算机程序可在包括至少一个可编程处理器的可编程系统上执行和/或解释,该至少一个可编程处理器可以是专用或通用的,并被耦合以从中接收数据和指令以及将数据和指令传输到存储系统、至少一个输入设备和至少一个输出设备。这些计算机程序(也称为程序、软件、软件应用或代码)可以包括用于可编程处理器的机器指令,并且可以以高级过程和/或面向对象的编程语言和/或以汇编/机器语言来实现。如本文中所使用的,术语“机器可读介质”和“计算机可读介质”是指被用来向可编程处理器提供机器指令和/或数据的任

何计算机程序产品、装置和/或设备(诸如磁盘、光盘、存储器或可编程逻辑设备(PLD))。

[0080] 在一些实施方式中,所提供的数据编排器可能能够确定哪个自动驾驶交通工具数据或自动驾驶交通工具数据的哪个部分将被传送给哪个数据中心或第三方实体,以及何时传输数据的这部分。可以使用应用表、规则和预测模型来确定数据传输或数据递送。预测模型可以是基于机器学习的模型。

[0081] 机器学习已经演变成为一种自动发现数据中的模式并使用所构建的模型来在各种应用中进行智能预测的关键计算结构。诸如机器学习算法之类的人工智能可以被用来训练用于数据编排的预测模型。机器学习算法例如可以是神经网络。神经网络的示例包括深度神经网络、卷积神经网络(CNN)和递归神经网络(RNN)。机器学习算法可以包括以下中的一个或多个:支持向量机(SVM)、朴素贝叶斯分类、线性回归、分位数回归、逻辑回归、随机森林、神经网络、CNN、RNN、梯度提升的分类器或抑制器、或另一种有监督或无监督的机器学习算法。

[0082] 图4示意性地示出了与一个或多个远程实体420、430通信的数据编排器410的示例。在一些实施方式中,数据编排器410可以包括决策引擎413和数据通信模块415。在一些情况下,数据编排器410可以可选地包括数据处理模块411。在一些情况下,数据处理模块411可以提供流数据和批处理数据的预处理。在替代实施方式中,数据处理模块411可以驻留于云420上,诸如数据和元数据管理模块425的组件。数据编排器410可以耦合到或具有一个或多个本地数据库,诸如应用存储库405和/或预测模型知识库407。应用存储库405可以存储如上所述的应用表。预测模型知识库405可以被配置为存储用于确定数据传输(方案)的机器学习模型和/或手工制定的规则。数据传输方案可以指定哪个自动驾驶交通工具数据要被传送到哪个数据中心或第三方实体,以及何时传输此类数据。预测模型知识库405可以存储除了由数据编排器所使用的机器学习模型之外的其他模型。例如,预测模型知识库405可以存储被用于交通工具的自动驾驶移动性的模型,被用于交通工具的(一个或多个)车厢的个性化以及在(一个或多个)车厢内部执行的其他功能的模型和/或被用于车队的安全和优化操作的模型。

[0083] 数据编排器410可以与预测模型管理模块421通信。预测模型管理模块421可以与如图1中所示的预测模型创建和管理系统130相同。预测模型管理模块421可以驻留在诸如数据中心、云、服务器等的远程实体420上。在一些情况下,预测模型管理模块421可以包括在云或本地部署环境中运行以通过网络远程配置和管理数据编排器410的服务或应用。在一些情况下,预测模型管理模块421是独立系统。在一些情况下,预测模型管理模块421可以是数据中心的组件,并且数据中心可以托管利用自动驾驶交通工具数据的一个或多个应用423。

[0084] 上面提及的应用存储库405可以与如上所述的应用表相同或包括应用表。

[0085] 预测模型知识库407可以存储机器学习模型和/或手工制定的规则。在基于知识的环境中,信息的可用性和利用与关联的人类专业知识结合,对于流程、实现和利用效率的改善是至关重要的。知识库在可以利用互联网访问或其他相关技术从全球位置进行访问的多个数据源中提供了有关特定主题的大量信息。

[0086] 应用存储库405、预测模型知识库407、一个或多个数据库107,109可以利用任何合适的数据库技术。例如,结构化查询语言(SQL)或“NoSQL”数据库可用于存储用户简档数据,

历史数据,预测模型或用于生成个性化交通计划,地图或其他数据的算法。一些数据库可以使用各种标准数据结构来实现,例如数组,哈希,(链接的)列表,结构,结构化文本文件(例如XML),表,JavaScript Object Notation(JSON),NOSQL等。这样的数据结构可以存储在存储器和/或(结构化)文件中。在另一替代方案中,可以使用面向对象的数据库。对象数据库可以包含多个对象集合,这些对象集合通过通用属性分组和/或链接在一起;它们可能通过一些公共属性与其他对象集合相关。面向对象的数据库的性能类似于关系数据库,不同的是对象不仅是数据片段,而且还可以具有封装在给定对象中的其他类型的功能。在一些实施方式中,数据库可以包括图形数据库,该图形数据库使用用于带有节点,边和属性的语义查询的图形结构来表示和存储数据。如果本公开的数据库被实现为数据结构,则本公开的数据库的使用可以被集成到另一组件中,诸如本公开的组件。而且,数据库可以实现为数据结构,对象和关系结构的混合。数据库可以通过标准数据处理技术进行整合和/或分布变化。数据库的部分,例如表,可以被导出和/或导入,从而分散和/或集成。

[0087] 在一些实施方式中,数据管理系统可以构造数据库以用于快速和有效的数据检索、查询和递送。例如,个人交通管理系统101可以提供定制的算法以提取,变换和加载(ETL)数据。在一些实施方式中,个人交通管理系统101可以使用专有数据库架构或数据结构来构建数据库,以提供适合于大规模数据库,易于扩展,查询和数据检索效率高,或与使用其他数据结构相比已经减少存储器需求的有效数据库模型。例如,可以使用树数据结构来存储模型树,其中节点表示模型的不同版本,并且节点参数表示模型的目标、性能特性以及各种其他参数。

[0088] 在一些实施方式中,数据编排器可以被应用于多层数据架构。图18示意性地图示出了多层数据架构1800。在所图示的示例中,数据编排器可以是基于如上所述的雾或边缘计算概念的基于软件的解决方案。在一些情况下,多层数据架构可以包括交通工具层(例如,车内数据1810)、雾层(例如,雾/边缘数据1820)和云层(例如,云数据1830)。多层数据架构可以包括任何数量的层。例如,雾层可以包括一层或多层。交通工具层处的数据可以包括由用户体验平台401和/或交通工具堆栈403、交通工具上的传感器以及本文其他地方所描述的各种其他源生成的车内数据1810。交通工具层处的数据可以与如上所述的自动驾驶交通工具数据相同。雾层处的数据(例如,雾/边缘数据1820)可以由数据编排器生成、管理和直接访问。雾/边缘数据1820可以包括由数据处理模块411处理之后的数据。数据处理模块411可以支持传感器数据到本地存储库(例如,本地时间序列数据库)的摄取、数据清理,数据丰富(例如,用元数据装饰数据)、数据对准、数据注释、数据标注、数据聚合以及各种其他数据处理。雾/边缘数据1820还可以包括根据传输方案要被传输到云的中间数据。

[0089] 数据编排器可以被配置为或能够确定哪个交通工具数据或交通工具数据的哪个部分保留在车内数据库中,要被移动/传输到雾层数据库(例如,雾/边缘数据库),以及哪个雾/边缘数据或雾/边缘数据中的哪个部分将被传送到哪个数据中心或第三方实体,数据的这部分何时以什么频率被传输。在一些情况下,可以从车内数据库中删除被卸载或移动到边缘/雾数据库的数据,以提高存储效率。可替代地,可以在将车内数据库中的数据卸载到边缘/雾数据库之后将其保留预定的时间段。

[0090] 图19示意性地示出了用于管理交通工具层与雾层之间以及雾层与云层之间的数据传输的数据编排器1910的示例。可以使用应用表、规则和预测模型来确定多个层之间的

数据传输或数据递送。预测模型可以是如上所述的基于机器学习的模型。

[0091] 数据编排器1910可以与如上所述的数据编排器410相同。例如,数据编排器1910可以包括决策引擎1913和数据通信模块1915。在一些情况下,数据编排器1910可以可选地包括数据处理模块(未示出)。在一些情况下,数据处理模块可以提供对从车内数据库1920传输的流数据和批处理数据的预处理。车内数据库1920可以在交通工具上并且存储交通工具数据(例如,车内数据1810)。数据编排器可以管理车内数据库1920和雾/边缘数据库1930之间以及雾/边缘数据库1930和云数据库之间的数据传输。

[0092] 数据编排器1910可以耦合到或具有一个或多个本地数据库,诸如如上所述的应用存储库405和/或预测模型知识库407。应用存储库405可以存储如上所述的应用表。预测模型知识库405可以被配置为存储用于确定数据传输(方案)的机器学习模型和/或手工制定的规则。数据传输方案可以指定哪个交通工具数据或交通工具数据的哪个部分保留在车内数据库1920中,要被移动/传输到雾层数据库(例如,雾/边缘数据库1930),以及此类数据何时和/或以什么频率被传输。数据传输方案还可以指定哪个雾/边缘数据或雾/边缘数据中的哪个部分将被传送到哪个数据中心或第三方实体,数据的这部分何时以什么频率被传输。预测模型知识库405可以存储除了由数据编排器所使用的机器学习模型之外的其他模型。在一些情况下,预测模型知识库405可能与该系统的知识库不同,或者不需要与该系统的知识库相同,该系统的知识库可以存储被用于交通工具的自动驾驶移动性的模型,被用于交通工具的(一个或多个)车厢的个性化以及在(一个或多个)车厢内部执行的其他功能的模型和/或被用于车队的安全和优化操作的模型。

[0093] 图5图示出了预测模型知识库500的示例。在一些实施方式中,预测模型知识库500可以包括汽车本体501和一个或多个模型树503。在一些实施方式中,预测模型知识库可以既包括手工制定的规则也包括基于机器学习的预测模型。手工制定的规则和基于机器学习的预测模型可以独立或共同确定监管数据传输的规则或协议。例如,规则可以指定要向其传输数据的给定聚合(例如,Message_package)的应用和/或数据中心,要传输的数据的聚合,数据传输方案(例如,诸如延迟时间之类的传输定时或频率、通信协议、被用于传输的压缩或加密方法)以及各种其他参数(例如,有关数据传输之前的隐私的监管规则)。

[0094] 手工制定的规则可以从外部源被导入或由一个或多个用户定义(例如,手工制定的规则可以是用户定义的规则)。在一些情况下,可以由请求来自交通工具的数据的远程应用来提供手工制定的规则。在一些情况下,可以基于来自远程应用的请求来确定数据传输方案。在一些情况下,请求可以是来自远程第三方应用(例如,应用423、430)发送到中间组件(例如,原始装备制造商(OEM))的请求。例如,保险应用可以以预定频率(例如,一周、两周、一个月、两个月等等)从与交通工具相关联的OEM系统请求某些类型的数据(例如,由OEM嵌入式设备所收集的数据),以了解驾驶员与他所支付的保险费率相比是否可能过度驾驶、创建新的保险产品、为驾驶员提供安全特征折扣、评估风险、事故现场管理、第一损失通知、增强索赔过程等。

[0095] 该请求可以包含有关应用所需的数据的类型、所需数据的频率、要传输这种类型的数据的时间段的信息或其他信息。在一些情形中,当数据传输不频繁或要传输的数据量相对较小时,可以基于上述请求生成数据传输方案,而无需使用诸如可以使用机器学习模型创建的智能传输方案。例如,请求应用(例如,保险应用)可以向与目标交通工具(或交通

工具组) 相关联的OEM系统发送指示需要来自目标交通工具的数据类型和此类数据的频率的请求。在一些情况下, 请求可以指定一组交通工具。例如, 请求可以指定特定型号 (例如, 奥迪A8)、型号年份 (例如, 2017)、具有特定驾驶自动化特征的型号 (例如, 具有车道变更监控器的A8) 等。OEM系统可以将请求传递给相应目标交通工具的数据编排器 (例如, 发送请求消息以中继该请求)。在接收到请求后, 数据编排器可以将请求推送到队列, 然后将响应消息发送回OEM系统以确认请求的接收。然后, OEM系统可以向请求应用发送消息, 指示该请求已被日志记录。

[0096] 接下来, 数据编排器可以基于请求中包含的信息来传输被请求的数据。数据编排器可以将被请求的数据直接发送到请求应用。在这样的情况下, 可以通过中间实体 (例如, OEM系统) 来传送与从数据编排器传输到远程应用 (例如, 请求应用) 的数据相关的信息。例如, 除了传递请求/响应消息之外, OEM系统/应用还可以向数据编排器发送消息, 该消息命令数据编排器在传输时段完成时 (例如, 在从数据编排器接收到完成消息后) 从队列中删除传输请求。然后, 数据编排器可以从队列中删除该条目, 并向OEM系统发送消息, 该消息指示该条目被删除。OEM系统可以向请求应用发送消息, 该消息指示请求已完成。

[0097] 预测模型知识库500可以存储除了由数据编排器所使用的机器学习模型之外的其他模型。例如, 预测模型知识库500可以存储被用于交通工具的自动驾驶移动性的模型, 被用于 (一个或多个) 车厢的个性化以及在交通工具和/或交通工具的 (一个或多个) 车厢内部执行的其他功能的模型和/或被用于车队的安全和优化操作的模型。存储在预测模型知识库500中的模型可以包括由数据编排器所使用的预测模型、由自动驾驶交通工具堆栈正在使用的预测模型、由用户体验平台所使用的模型或车队管理系统。可替代地, 由自动驾驶交通工具堆栈正在使用的预测模型和由用户体验平台所使用的模型可以分别被存储在由相应自动驾驶交通工具堆栈或用户体验平台所管理的预测模型知识库中。

[0098] 汽车本体501可以由一个或多个个人、组织手动开发, 从外部系统或资源中导入, 或者可以使用与用户协作的机器学习系统来部分学习 (例如, 从自然语言文本中提取汽车术语)。在一些情况下, 汽车本体的一部分可以基于来自模型树的数据。例如, 模型的目标和/或洞察力的描述可以被存储在模型树的节点中, 而目标和/或洞察力的描述也可以是汽车本体的一部分。

[0099] 预测模型知识库500可以存储其他本体或模型。在一些情况下, 可以创建场景元数据以使用特定的元数据来指定场景的特性, 特定的元数据然后被用来从数据库中检索适当的交通工具数据。预测模型知识库可以包括分层场景本体, 其可以被用来创建新场景以及以各种细节级别创建场景。例如, 以较高细节级别 (即, 有关场景的较高级别的信息) 描述的场景可以被用来创建低保真度模拟或预测模型, 而以较低细节级别 (即, 有关场景的更详细的较低级别的信息) 描述的相同场景可以被用来生成高保真度模拟或预测模型。

[0100] 一个或多个模型树503可以是树结构的集合。树结构可以包括一个或多个节点507, 其中每个节点包括预测模型的特性和指向被用来生成预测模型的数据 (例如, 训练数据、测试数据) 的指针。实际数据 (例如, 训练数据、测试数据) 可以被存储在云数据库429中。云数据库429可以与云数据湖125、127相同, 或者包括云数据湖125、127中的一个或两个。节点在给定模型树中的分层结构可以表示特定预测模型的版本以及模型之间的关系。预测模型的特性可以包括例如预测模型的目标/功能、模型性能特性以及各种其他特征。节点507

还可以存储模型参数(例如权重、超参数等)、关于模型参数的元数据、模型的性能统计信息或模型架构(例如层数、层中的节点数、CNN、RNN)。在一些情况下,节点507还可以包括关于执行模型所需的(一个或多个)计算资源(例如,一个图形处理单元(GPU)、两个GPU、三个CPU等)的信息。节点可以包括如上所述的数据的全部或任何组合。

[0101] 在一些情况下,可以使用不同的模型树结构来存储各种预测模型。知识库可以具有不同的模型树结构,这取决于例如预测模型正被使用的位置。例如,用于存储由用户体验平台所使用的预测模型的模型树结构可以不同于存储由数据编排器所使用的预测模型的模型树结构。

[0102] 模型树可以是动态的。例如,可以响应于模型的原始架构的更改、模型的性能特性的更改或训练数据或测试数据的更改来创建新节点。

[0103] 在一些情况下,预测模型知识库也可以存储手工制定的规则。手工制定的规则可以由一个或多个个人、组织手动开发,也可以从外部系统或资源中导入。手工制定的规则和预测模型可以独立地、顺序地或同时地被应用。

[0104] 下面示出了根据图3中描述的应用表的关于数据传输的规则示例:

如果设置了Pothole_Detector_App的Transmission_Flag

那么

确定何时必须传输数据

对于带有Transmission_Flag=SET的每个Vehicle_ID

接收来自交通工具的视频流

应用所有可应用的监管规则

在应用可应用的监管规则并创建Encrypted_Data_File之后,对产生的数据集加密
压缩Encrypted_Data_File,创建Compressed_Data_File

创建由[Compressed_Data_File,File_Size,Transmission_Delay,Data_Center_Address_List]组成的Message_Package

将Message_Package发送到通信模块队列

[0105] 在一些实施方式中,数据传输方案还可以指定如何传输数据。例如,数据传输方案可以指定被用于传输的压缩方法(例如,无损压缩算法、有损压缩算法、编码等)或加密方法(例如,RSA、三重DES、Blowfish、Twofish、AES等)。在一些情况下,可以基于规则来确定用于传输的数据压缩方法和/或加密方法。例如,规则可以根据给定的数据类型、使用该数据的应用、该数据的目的地等等来确定压缩方法和/或加密方法。用于确定数据压缩方法和/或加密方法的规则可以被存储在数据编排器可访问的数据库中,诸如如上所述的预测模型知识库。在一些情况下,用于确定数据压缩方法和/或加密方法的规则可以是用于确定数据传输的规则的一部分。例如,用于确定加密方法或压缩方法的规则集(例如,通过规则集标识符)可以被调用以用于确定数据传输方案。

[0106] 用于确定压缩方法和/或加密方法的规则可以是手工制定的规则。例如,关于压缩方法和/或加密方法的预定或手工制定的规则可以在接收到指定数据的类型、与应用相关的数据、数据的目的地等等的传输请求后被应用。这样的手工制定的规则可以被存储在数据编排器可访问的数据库中,诸如如上所述的预测模型知识库。在一些情况下,可以通过机器学习算法训练的模型来确定压缩方法和/或加密方法。例如,当用于数据压缩或加密的预

定规则集不可用时(例如,规则集标识符不可用,之前看不到数据集的类型等),经训练的模型可以被应用于要被传输的数据集并生成用于对数据集进行压缩或加密的规则。在一些情况下,由经训练的模型所生成的规则集可以被存储在预测模型知识库中,以用于将来的数据传输(方案)。

[0107] 返回参见图4,数据处理模块411可以支持传感器数据到本地存储库(例如,本地时间序列数据库)的摄取、数据清理、数据丰富(例如,用元数据装饰数据)、数据对准、数据注释、数据标注、数据聚合和各种其他数据处理。来自用户体验平台401和/或交通工具堆栈403、交通工具上的传感器以及本文其他地方所描述的各种其他源的数据可以由数据处理模块摄取和处理。例如,数据处理模块可以经由一个或多个协议(例如,MQ遥测传送、OPC统一架构、Modbus和DDS)从传感器收集或摄取数据。由传感器提供或输出的数据可以是二进制数据流。该数据从传感器到数据处理模块的传输或递送可以是推或拉方法。在一些情况下,数据处理模块可以通过将原始二进制数据解码为可消耗的数据格式(诸如JavaScript Object Notation)或与其他必要的和有用的元数据合并来丰富来自传感器的传入数据。在一些实施方式中,元数据可以涉及捕获感觉数据(例如GPS、激光雷达、相机等)、对数据进行预处理(例如,对准和创建时间序列)的传感器以及针对特定用例或应用(例如,避开行人、模式识别、避开障碍物等)对数据进行操作的各种应用和/或预测模型。可替代地,这种数据处理可以由云上的应用来执行。例如,数据处理模块411可以驻留在云420上而不是数据编排器上。稍后将在本文中描述有关数据处理方法和元数据创建的细节。

[0108] 决策引擎413可以被配置为执行预测模型知识库407中的规则。例如,决策引擎可以不断地在预测模型知识库407中寻找合格或准备好执行的规则,然后执行动作与符合条件的规则相关联并调用数据通信模块415以将结果(例如,聚合数据,Message_Package)传输到目的地(例如,被请求的数据中心420、应用431、远程实体、第三方实体431等)。

[0109] 数据通信模块415可以依照规则将自动驾驶交通工具数据的选定部分或处理后的数据发送到目的地。下面示出了由数据通信模块所执行的过程的示例:

步骤1检查Queue_Status

步骤2如果Queue_Status=1,那么

对于每个Message_Package

审查File_Size和Transmission_Delay以确定Transmission_Channel_Type

选择Transmission_Channel_Type的Transmission_Channel

通过选定的Transmission_Channel将Compressed_Data_File发送到Data_Center_Address_List中的每个地址

如果Transmission_Success=1,那么从队列Else GoTo中删除Message_Package,否则转到发送步骤3转到步骤1

[0110] 各种通信协议可以被用来促进数据编排器与云或远程实体之间的通信。这些通信协议可以包括VLAN、MPLS、TCP/IP、隧道、HTTP协议、无线应用协议(WAP)、特定于供应商的协议、定制协议以及各种其他协议。虽然在一个实施方式中,通信网络是互联网,但是在其他实施方式中,通信网络可以是任何合适的通信网络,包括局域网(LAN)、广域网(WAN)、无线网络、内联网、专用网络、公共网络、交换网络以及它们的组合等等。该网络可以包括使用无线和/或有线通信系统的局域网和/或广域网的任何组合。例如,网络110可以包括因特网以

及移动电话网络。在一实施方式中,网络110使用标准通信技术和/或协议。因此,网络110可以包括使用诸如以太网,802.11,全球微波接入互通性(WiMAX),2G/3G/4G或长期演进(LTE)移动通信协议,红外(IR)通信技术,和/或Wi-Fi的技术的链路,并且可以是无线,有线,异步传输模式(ATM),无线宽带,PCI快速高级交换或它们的组合。网络110上使用的其他联网协议可以包括多协议标签交换(MPLS),传输控制协议/因特网协议(TCP/IP),用户数据报协议(UDP),超文本传输协议(HTTP),简单邮件传输协议(SMTP),文件传输协议(FTP)等。可以使用包括二进制形式的图像数据(例如,便携式网络图形(PNG)),超文本标记语言(HTML),可扩展标记语言(XML)等的技术和/或格式来表示通过网络交换的数据。另外,可以使用诸如安全套接字层(SSL),传输层安全性(TLS),互联网协议安全性(IPsec)等的常规加密技术来加密所有或一些链路。在另一实施方式中,网络上的实体可以替代上述方法或除了上述方法之外,使用定制和/或专用数据通信技术。该网络可以是无线的,有线的或其组合。

预测模型创建和管理系统

[0111] 预测模型管理模块421可以与图1中描述的预测模型创建和管理系统相同。预测模型管理模块421可以包括在云或本地部署环境中运行以远程配置和管理数据编排器410或数据编排器的一个或多个组件(例如,预测模型知识库)的服务或应用。

[0112] 在一些实施方式中,预测模型管理模块421可以包括模型创建器和模型管理器。在一些情况下,模型创建器可以被配置为使用来自云数据湖和元数据数据库的数据来训练、开发或测试预测模型。模型管理器可以被配置为管理各个组件(例如,云数据湖、元数据数据库、数据编排器、模型创建器)之间的数据流,提供精确、复杂和快速的查询(例如模型查询、元数据查询)、模型部署、维护、监视、模型更新、模型版本化、模型共享以及各种其他功能。例如,部署上下文可以取决于边缘基础设施而有所不同,并且模型管理器可以考虑应用清单,诸如边缘硬件规格,部署位置,有关兼容系统的信息,用于安全性和隐私的数据访问清单,用于在模型部署和维护期间对给定部署和版本管理中不可用的数据字段进行建模的模拟器。

[0113] 由预测模型管理模块所提供的数据管理可以被应用于自动化交通工具和自动驾驶交通工具的整个生命周期。例如,可以在交通工具设计阶段、交通工具/车队验证阶段或交通工具/车队部署阶段中的各种应用中应用数据管理。图14示出了自动化和自动驾驶交通工具的生命周期中的各种应用的示例。例如,数据管理可以被用于在交通工具设计阶段、交通工具/车队验证阶段或交通工具/车队部署阶段中创建新模型或更新现有模型。

[0114] 图6示出了模型创建器600与元数据数据库427和云数据湖429交互以用于训练和开发预测模型的示例。可以测试经训练的预测模型的性能,满足性能要求的预测模型然后可以被插入到预测模型知识库407中。在一些实施方式中,云数据库429可以包括如图1中所示的两个云数据湖125、127。

[0115] 模型创建器可以被配置为开发由数据编排器所使用的预测模型,由自动驾驶交通工具堆栈所使用的预测模型,由用户体验平台、由车队管理系统和各种其他系统所使用的预测模型。除了数据管理和编排之外,模型创建器还可以训练和开发被用于交通工具的自动驾驶移动性、用于(一个或多个)车厢的个性化以及在交通工具和/或(一个或多个)车厢内部执行的其他功能和/或用于车队的安全和优化操作和/或各种其他应用的模型。

[0116] 图7图示出了创建预测模型的方法700。该方法或过程可以由如上所述的模型创建

器执行。为了生成预测模型,模型目标和性能特性(例如,准确性)可以被确定(操作701)。另外,期望的数据特性(例如,完整性、有效性、准确性、一致性、可用性和及时性)可以被确定(操作702)。接下来,可以从数据库(例如,云数据湖)中选择被标记的数据或数据集以用于对模型进行训练(操作703)。在一些情况下,从数据库中检索数据可以包括:利用数据特性查询元数据数据库内的元数据,然后基于元数据查询结果从云数据湖中检索数据。如果没有从云数据湖返回数据,则可以调整数据特性(即,重复操作702)。在一些情况下,可以在下一步之前对所返回的数据或数据集进行采样。

[0117] 在一些情况下,可以鉴于模型目标来分析被标记的数据或数据集的适当性(操作704)。例如,可以确定被标记的数据集对于预测目标是否足够,例如,开发使得自动驾驶交通工具能够自动进行右转弯的预测模型。可以利用各种合适的方法来确定被标记的数据集的适当性。例如,可以计算统计能力并将其用于分析。统计能力是一项研究在有效果要被检测时将检测出效果的可能性。如果统计能力为高,则发生II型错误或当事实上有效果时得出没有效果的结论的可能性不大。统计能力主要受效果的大小和用于检测该效果的样本大小的影响。较大的效果比较小的效果更容易检测,而较大的样本比较小的样本具有更高的测试灵敏度。

[0118] 在操作704处产生的分析结果可以确定数据集是否需要被校正。适当性分析的结果可以指示数据集是否满足适当的要求,适当性的级别或者是否需要被校正。例如,当被标记的数据集的适当性被计算并且低于预定阈值时,可以确定该数据集不满足适当性要求并且可能需要校正。在确定数据集不需要校正后,数据集可以被用于训练预测模型(操作706)。在一些情况下,对模型进行训练可以涉及:选择模型类型(例如CNN、RNN、梯度提升的分类器或抑制器等)、选择模型的架构(例如层数、节点、ReLU层等)、设置参数、创建训练数据(例如配对数据、生成输入数据向量)以及处理训练数据以创建模型。在一些情况下,如果数据集被分析并确定为需要数据校正,则可以执行校正(操作705)。在数据集不能被校正的情况下,可以从数据库中选择新的或不同的数据集(即,重复操作703)。

[0119] 可以使用从预测模型知识库407中检索的测试数据来测试和优化经训练的模型(操作707)。接下来,可以将测试结果与性能特性进行比较,以确定预测模型是否满足性能要求(操作708)。如果性能良好,即满足性能要求,则可以将模型插入到预测模型知识库407中(操作709)。

[0120] 在一些情况下,将新模型插入到预测模型知识库中可以包括:确定将新模型插入到模型树中的位置(例如,作为新节点添加到现有模型树中或新模型树中)。与新模型一起,其他数据(诸如模型目标、模型架构、模型参数、训练数据、测试数据、模型性能统计信息)也可以被归档在模型树结构中。接下来,可以由模型创建器或模型管理器不断地监视预测模型性能(操作710)。如果经训练的模型没有通过性能测试,则过程可以继续以确定较差的性能是由数据特性还是由模型特性引起的。遵循决策,可以重复操作701(例如,调整性能特性)和/或操作702(例如,调整数据特性)。

[0121] 在一些情况下,在创建新的预测模型或对现有的预测模型进行更新/更改后,预测模型可能对被选交通工具可用。例如,一旦预测模型被更新并被存储在预测模型知识库中,就可以将预测模型下载到车队中的一个或多个交通工具上。可用预测模型可以被下载或更新在以动态方式选择的一个或多个中。图15图示出了动态更新交通工具中的预测模型的示

例。

[0122] 如上所述,除了数据管理和数据编排之外,预测模型还可以包括被用于交通工具的自动驾驶移动性、用于(一个或多个)车厢的个性化以及在交通工具和/或(一个或多个)车厢内部执行的其他功能和/或用于车队的安全和优化操作和/或各种其他应用的模型。可以创建新模型以便使得交通工具能够应对新的情形。可以更新模型以便基于已收集并存储在云数据湖中的新数据来提高整体性能。在一些情况下,系统可访问的由一组交通工具或车队中的特定交通工具所使用的预测模型的列表被保持在交通工具数据库中。

[0123] 在一些情况下,新模型的这种更新、更改或创建可以被预测模型管理模块的组件自动检测到。例如,参考图15,当新模型被创建或现有模型已经被更新时,预测模型更新模块1501可以被预测模型知识库1503通知。预测模型更新模块1501然后可以选择一个或多个交通工具接收已更新模型的副本。可以基于订阅、模型的利用或其他标准来选择或确定该一个或多个交通工具。预测模型更新组件还可以确定何时在被选交通工具中更新模型。例如,预测模型更新组件可以立即、在交通工具静止时(例如,在维护、清洁、维修等期间)或根据需要确定模型需要被更新。例如,在交通工具是即时用车车队的一部分的情况下,只有在该交通工具是被指派完成涉及经过受影响区域的旅程(例如,接载乘客,放下乘客,或在其他地方接载或放下乘客的过程中通过该区域)时,才可能需要一种用于在旧金山棒球场中进行比赛时在夜间进行右转弯的预测模型。预测模型更新模块1501可以是如上所述的云订阅模块的一部分。

[0124] 返回参见图4,云或数据中心420或所提供的交通工具数据管理系统还可以包括数据和元数据管理模块425。数据和元数据管理模块425可以执行各种功能,包括由数据处理模块411实施的数据处理,以及元数据创建和管理。数据和元数据管理模块可以被配置为存储和管理由自动驾驶交通工具所生成的数据和关联的元数据,并处理针对数据和元数据所发出的查询和API调用。例如,结合图8至图12讨论关于数据和元数据管理模块的细节。

[0125] 云或数据中心420还可以包括云应用423和用于查看分析、传感器数据(例如,视频)和/或已处理数据的用户界面(UI)模块425。UI还可以包括管理UI,用于开发和部署分析表达式、将数据编排应用部署到边缘(例如,自动驾驶交通工具操作系统、边缘网关、边缘基础设施、数据编排器)以及配置和监视数据编排。

[0126] 图8示出了交通工具数据管理系统800的示例组件,尤其是驻留在远程实体(例如,数据中心)上的交通工具数据管理系统的组件。在一些实施方式中,交通工具数据管理系统800可以包括数据和元数据管理系统以及预测模型创建和管理系统。数据和元数据管理系统可以与如上所述的数据和元数据管理模块相同。数据和元数据管理系统可以被配置为存储和管理由自动驾驶交通工具所生成的数据和关联的元数据,并且处理针对数据和元数据所发出的查询和API调用。在一些实施方式中,交通工具数据管理系统800可以包括数据和元数据管理系统,该数据和元数据管理系统至少包括流水线引擎801和预测模型创建和管理系统803。在一些实施方式中,数据和元数据管理系统可以包括其他功能组件,诸如数据库查询引擎805、元数据查询引擎807、数据系统管理815、数据系统归档规则817、数据系统安全性819、数据库API 821、监管规则823和云-云通信825。例如,元数据数据库809可以通过数据库查询引擎805使用元数据查询语言而被访问。在另一示例中,数据湖811、813中的数据可以作为元数据查询的结果来访问,或者可以直接使用数据库查询引擎805来访问。

云-云通信825可以包括各种通信协议,诸如VLAN、MPLS、TCP/IP、隧道、HTTP协议、无线应用协议(WAP)、特定于供应商的协议、定制协议以及各种其他协议。云-云通信可以包括可以由系统(例如基于云的系统)和/或设备使用的接口并且可以利用各种API(例如使用REST架构的API、SOAP、Web服务、企业服务总线协议或任何被设计来为系统提供方法给系统和/或过程以处理通信的数据交换协议)。在一些实施方式中,交通工具数据管理系统800还可以包括元数据数据库809、用于存储自动驾驶交通工具堆栈数据811的云数据湖和用于存储用户体验平台数据813的云数据湖。

[0127] 在一些实施方式中,如上所述的一个或多个组件可以与一个或多个云应用或企业应用交互(例如,维护车队831、管理车队833、地图更新835、配置车队837)。云应用可以被托管在远程实体上,并且可以利用由数据管理系统管理的交通工具数据。在一些情况下,云应用可以具有由预测模型创建和管理系统803创建的数据库或知识库832、834、836、838。在一些情况下,云应用可以具有访问和操纵以下数据的权限:存储在用于存储自动驾驶交通工具堆栈数据的云数据湖811中的数据,用于存储用户体验平台数据813的云数据湖,或者元数据数据库809中存储的元数据。在一些情况下,数据可以被调遣到云应用,并且为了将数据调遣到对应的云应用(如在元数据或应用表中所标识的),预测模型创建和管理系统可以将云上所有资源(即应用)的地址本地列出在表中,以用于快速查找。

[0128] 流水线引擎801可以被配置为对从数据编排器传输的原始数据或批处理数据的连续流进行预处理。例如,可以对数据进行处理,以便将其馈送到机器学习分析中。数据处理例如可以包括数据标准化、用元数据标记数据、标注、数据对准、数据分段以及各种其他处理。在一些情况下,处理方法可由构建机器学习分析的开发人员通过API来编程。

[0129] 图9示出了数据摄取流水线900和由流水线引擎执行的功能的示例。流水线900可以包括用于处理以流或以批处理方式摄取的数据的多个功能。数据可以包括例如模拟数据、来自交通工具的数据(诸如车队数据)、操作环境数据、运输数据、交通工具遥测、由AV堆栈和用户体验平台生成的数据、第三方合作伙伴数据(例如来自用户移动应用的数据)、传感器数据(例如GPS、IMU、相机、激光雷达、红外传感器、热传感器、超声传感器等)、地理位置数据以及本文其他地方所述的各种其他数据。流数据可以包括各种类型的数据,包括但不限于:时间序列数据(诸如环境的时空点测量)、多维数据(诸如来自雷达、激光雷达、卫星、声纳的网格化测量和/或以面向阵列的数据格式或任何其他适合表示多维数据、可视化数据(诸如地图图块)、元数据、原始数据(诸如来自传感器、文档、数字服务的原始输入数据)的数据格式所形成的模拟过程的输出、数据收集、集成、处理分析的源代码和各种其他数据。各批处理可以是特定于租户的,特定于应用的,并且可以被归组为上下文感知的子组以进行并行处理。如本文其他地方所述,可以从数据编排器生成并传输各批处理。

[0130] 流水线900可以是可定制的。例如,流水线900的一个或多个功能可以由用户创建。可替代地或另外,一个或多个功能可以由管理系统创建,也可以从其他系统或第三方源中导入。在一些情况下,可以允许用户从功能集(例如,可用功能920)中进行选择,并将被选功能添加到流水线中。在一些情况下,可以经由由用户界面模块(例如,图4中的用户界面模块425)提供的图形用户界面(GUI)来执行创建或修改流水线。例如,可以在GUI内显示一组可用功能。用户可以在GUI内通过点击功能来选择表示CREATE SCENARIOS功能的图形元素,也可以通过拖放将功能添加到当前流水线中。

[0131] 在一些情况下,可以在显示器上提供图形用户界面(GUI)或用户界面。显示器可以是触摸屏,也可以不是。显示器可以是发光二极管(LED)屏幕、有机发光二极管(OLED)屏幕、液晶显示器(LCD)屏幕、等离子屏幕或任何其他类型的屏幕。显示器可以被配置为示出通过应用(例如,经由在用户设备上、在云上或在数据编排器上执行的应用编程接口(API))呈现的用户界面(UI)或图形用户界面(GUI)。

[0132] 在一些实施方式中,多个功能可以包括第三方功能,诸如摄取901、过滤905、清理907、标注909、增强911、注释913、匿名化915以及各种其他功能(例如,模拟)。例如,数据清理907可以包括从数据中消除噪音(例如,图像处理中的降噪)、校正错误数据(例如,一个相机出现故障并且没有显示光、但是白天)、建立通用数据格式(例如,使用度量系统,所有数字到小数点后三位等)、或准备数据以使得预期的数据消费者或应用可以经由API快速、轻松地访问它。在另一示例中,数据增强911可以包括:将合成数据与真实数据结合为更完整的数据集,以测试自动驾驶交通工具模型;利用来自合作伙伴的数据来增强所捕获的数据以启用某些类型的预测;将交通拥堵数据与天气数据相结合以预测行驶时间;组合若干数据集以创建信息丰富的数据(例如,将交通工具操作数据与城市交通基础设施数据和拥堵数据相结合,以预测一天中的特定时间的交通工具到达时间)。在另一示例中,数据标注909或注释913可以包括在元数据的每个级别和创建处发生的多媒体数据(例如,图像、激光雷达、音频)的注释。可以在数据管理环境中的数据移动期间创建元数据。例如,可能需要检索图像,对图像进行注释(最可能是通过一些手动干预),然后重新索引。所创建的元数据可以被合并到元数据目录中。其他元数据(诸如,手动或自动生成的各种类型的元数据)也可以被插入在元数据目录中。多个功能还可以包括专有功能,诸如数据对准903和创建场景921。

[0133] 图10示出了示例数据摄取过程1000。在一些情况下,流数据和/或批处理数据可以在流水线引擎中被摄取。在一些情况下,被摄取的流数据可以被递送到流处理系统1001,并且被摄取的批处理数据可以被递送到提取-变换-加载(ETL)系统1003。ETL系统1003可以执行传统的ETL功能性或定制的功能性。例如,ETL系统可以将被摄取的批处理数据变换为对用户更有用的格式。例如,数据变换可以包括:仅选择某些列以加载为格式,翻译编码值,导出新的计算值,对数据进行分类,对数据进行聚合,转置或旋转数据,将一系列拆分为多列以及其他处理。

[0134] 尽管本文讨论了流处理系统1001和ETL系统1003,但是可以使用其他模块或替代模块来实现本文所描述的功能性。流处理系统和ETL系统仅是可以实现的许多可执行模块的示例。

[0135] 在一些情况下,数据对准可以由ETL系统或流处理系统执行。在一些情况下,由不同传感器捕获的数据(例如,传感器可以以不同的频率捕获数据)或来自不同源的数据(例如,第三方应用数据)可以被对准。例如,由相机、激光雷达所捕获的数据和遥测数据(例如,温度、交通工具状态、电池电量等)可以相对于时间被对准。在一些情况下,可以自动执行数据对准。可替代地或另外,用户可以指定要从其对准传感器或源的所收集的数据和/或要在其期间对准数据的时间窗口。在示例中,结果数据可以是关于时间对准的时间序列数据。应该注意,数据可以沿其他维度对准,诸如应用、数据结构等。

元数据

[0136] 交通工具数据管理系统可以提供元数据管理。在一些情况下,如上所述,元数据创

建和管理可以由数据和元数据管理系统提供。在一些情况下,元数据可以允许基于元数据选择数据的子集或自动驾驶交通工具数据的一部分。在一些实施方式中,元数据可以提供关于,该传感器捕获感觉数据(例如GPS、激光雷达、相机等)的传感器、对数据进行预处理(例如对准和创建时间序列)以及针对特定用例或应用(例如避开行人、模式识别、避开障碍物等)对数据进行操作的各种应用和/或预测模型的信息。元数据可以在交通工具上被创建。例如,元数据可以由在交通工具上运行的传感器或应用来生成。在另一示例中,元数据可以由交通工具上的数据编排器来生成。元数据可以在远离交通工具的地方或由远程实体生成。例如,关于数据处理(例如,对准)的元数据可以在数据中心中或由云应用生成。在一些情况下,元数据的至少一部分在交通工具上生成并传输到远程实体。在一些情况下,元数据的至少一部分由在远程实体上提供的组件(例如,云应用或流水线引擎)生成。所创建的元数据可以被存储在由数据管理系统管理的元数据数据库中。作为替代或补充,元数据可以被存储在具有被用来生成元数据的数据中的至少一些或全部的数据库中。

[0137] 图11图示出了由对准、应用和传感器所生成的元数据的示例。例如,当不同的传感器数据1111、1113被对准时,可以创建元数据(例如,对准创建的元数据1103)以提供对准信息(例如,结构填充、频率、时间窗口等)。在一些情况下,可以创建关于产生数据的源或传感器的元数据(例如,传感器创建的元数据1105)。例如,传感器创建的元数据可以包括关于传感器的信息、传感器的标识符、数据类型以及其他。在一些情况下,可以通过处理和/或生成数据的应用来创建关于应用的元数据(例如,应用创建的元数据1101)。例如,应用程序创建的元数据1101可以提供关于应用的名称、应用的开发者、应用版本的信息以及各种其他信息。

[0138] 在一些实施方式中,数据管理系统可以生成元数据的元数据,以用于从数据库中快速检索或查询数据。例如,可以创建场景元数据以使用特定的元数据来指定场景的特性,然后将其用来从数据库中检索适当的交通工具数据。图12示出了场景元数据1200的示例。在以下示例中,描述其中交通工具在多云的早晨进行右转弯以从城市街道汇入到高速公路上的场景的场景数据对象可以被定义为:

Scenario_Name:例如,右侧汇入

Scenario_Type:例如,右侧汇入到高速公路中

Static_Objects:例如,树

Dynamic_Objects:例如,moving_vehicle

环境:例如,cloudy_morning

场景描述:例如,从城市地面街道进入高速公路的交通工具

Trigger_Rules:例如,在交通工具接近交叉路口之前10秒钟开始选择,并在汇入完成之后20秒停止

数据:例如,time-series-segment1

[0139] 以下是创建新方案对象的示例过程:

步骤1指定Scenario_Type

步骤2使用MetadataCatalog中的元数据指定每个Static_Object、每个Dynamic_Object和Scene_Description

步骤3使用元数据查询语言发出查询,使用被选元数据以创建ResultSet

步骤4如果ResultSet为空,那么确定NewTimeSeriesData是否可以被物理收集或模拟,否则转到步骤7。

步骤5如果需要模拟NewTimeSeriesData,那么确定是否需要模拟整个数据集,或者是否仅需要模拟一部分数据集,然后将其与物理收集的数据合并。转到步骤6。

步骤6使用Data_Ingestion_Pipeline来处理NewTimeSeriesData (收集的、模拟的或其组合)

步骤7对于NewTimeSeriesDataSet中的每个TimeSeries

a.使用Triger_Rules来选择满足步骤2中指定的元数据的TimeSeries分段,以创建TimeSeriesSegment。

b.创建新ScenarioObject

c.将名称与被选分段相关联

d.填充步骤2中的元数据

ed.在ScenarioObject中插入TimeSeriesSegment

计算机系统

[0140] 本文所描述的交通工具数据管理系统、数据编排器或过程可以由一个或多个处理器来实现。在一些实施方式中,一个或多个处理器可以是以细粒度的空间架构的形式可编程处理器(例如,中央处理单元(CPU)、图形处理单元(GPU)、通用处理单元或微控制器),诸如现场可编程门阵列(FPGA)、专用集成电路(ASIC)和/或一个或多个高级RISC机器(ARM)处理器。图13示出了计算机系统1301,其被编程或以其他方式配置为实现个人交通管理系统。计算机系统1301可以调节本公开的各个方面。计算机系统1301可以是用户的电子设备或相对于电子设备位于远程的计算机系统。该电子设备可以是移动电子设备。

[0141] 计算机系统1301包括中央处理单元(CPU,在本文中也称为“处理器”和“计算机处理器”)1305,其可以是单核或多核处理器,或者是用于并行处理的多个处理器。计算机系统1301还包括存储器或存储器位置1310(例如,随机存取存储器,只读存储器,闪存),电子存储单元1315(例如,硬盘),用于与一个或多个其他系统通信的通信接口1320(例如,网络适配器)以及外围设备1325,例如高速缓存,其他存储器,数据存储和/或电子显示适配器。存储器1310,存储单元1315,接口1320和外围设备1325通过诸如主板的通信总线(实线)与CPU 1305通信。存储单元1315可以是用于存储数据的数据存储单元(或数据存储库)。计算机系统1301可以借助于通信接口1320可操作地耦合到计算机网络(“网络”)1030。网络1030可以是因特网,内联网和/或外联网,或与因特网通信的内部网和/或外部网。在一些情况下,网络1030是电信和/或数据网络。网络1030可以包括一个或多个计算机服务器,其可以启用分布式计算,例如云计算。在一些情况下,网络1030可以在计算机系统1301的帮助下实现对等网络,该对等网络可以使耦合到计算机系统1301的设备能够充当客户端或服务器。

[0142] CPU 1305可以执行一系列机器可读指令,其可以体现在程序或软件中。指令可以存储在诸如存储器1310的存储器位置中。指令可以被定向到CPU 1305,CPU 1305可以随后对CPU 1305进行编程或以其他方式配置CPU 1305以实现本公开的方法。CPU 1305执行的操作的示例可以包括获取,解码,执行和写回。

[0143] CPU 1305可以是诸如集成电路的电路的一部分。系统1301的一个或多个其他组件可以被包括在电路中。在一些情况下,该电路是专用集成电路(ASIC)。

[0144] 存储单元1315可以存储文件,例如驱动程序,库和保存的程序。存储单元1315可以存储用户数据,例如,用户偏好和用户程序。在一些情况下,计算机系统1301可以包括计算机系统1301外部的一个或多个其他数据存储单元,例如位于通过内部网或因特网与计算机系统1301通信的远程服务器上。

[0145] 计算机系统1301可通过网络1030与一个或多个远程计算机系统通信。例如,计算机系统1301可以与用户(例如,用户设备)的远程计算机系统通信。远程计算机系统的示例包括个人计算机(例如,便携式PC),平板或平板计算机(例如,Apple®iPad, Samsung®Galaxy Tab),电话,智能电话(例如,Apple®iPhone,支持Android的设备, Blackberry®),智能手表,智能扬声器或个人数字助理。用户可以通过网络1030访问计算机系统1301。

[0146] 本文所述的方法可以通过存储在计算机系统1301的电子存储位置(例如,存储器1310或电子存储单元1315上)的机器(例如,计算机处理器)可执行代码来实现。机器可执行或机器可读代码可以以软件的形式提供。在使用期间,代码可以由处理器1305执行。在一些情况下,可以从存储单元1315检索代码并将其存储在存储器1310中,以供处理器1305随时访问。在一些情况下,可以不包括电子存储单元1315,并且将机器可执行指令存储在存储器1310中。

[0147] 该代码可以被预编译并配置为与具有适于执行该代码的处理器一起使用,或者可以在运行时期被编译。可以以可以选择的编程语言来提供代码,以使代码能够以预编译或编译时的方式执行。

[0148] 本文提供的系统和方法的方面,例如计算机系统1301,可以在编程中体现。可以将技术的各个方面视为通常以机器可读介质的类型承载或体现的机器(或处理器)可执行代码和/或关联数据的形式“产品”或“制品”。机器可执行代码可以存储在电子存储单元上,例如存储器(例如,只读存储器,随机存取存储器,闪存)或硬盘。“存储”类型的介质可以包括计算机,处理器等的任何或所有有形存储器,或其相关模块,例如各种半导体存储器,磁带驱动器,磁盘驱动器等,它们可以随时提供非暂时性存储进行软件编程。软件的全部或部分有时可以通过因特网或各种其他电信网络进行通信。例如,这样的通信可以使得能够将软件从一个计算机或处理器加载到另一计算机或处理器,例如从管理服务器或主机加载到应用服务器的计算机平台。因此,可以承载软件元件的另一种类型的介质包括光波,电波和电磁波,例如通过有线和光学座机网络以及通过各种空中链路在本地设备之间的物理接口上使用。诸如有线或无线链路,光链路等的携带这种波的物理元件也可以被视为承载软件的介质。如本文所使用,除非限于非暂时性有形“存储”介质,否则诸如计算机或机器“可读介质”的术语是指参与向处理器提供指令以供执行的任何介质。

[0149] 因此,诸如计算机可执行代码的机器可读介质可以采取许多形式,包括但不限于有形存储介质,载波介质或物理传输介质。非易失性存储介质包括例如光盘或磁盘,诸如任何计算机中的任何存储设备等,诸如可用于实现附图中所示的数据库等。易失性存储介质包括动态存储器,例如这种计算机平台的主存储器。有形的传输介质包括同轴电缆;铜线和光纤,包括构成计算机系统内总线的电线。载波传输介质可以采用电或电磁信号或声波或光波的形式,例如在射频(RF)和红外(IR)数据通信期间生成的那些。因此,计算机可读介质

的常见形式例如包括：软盘，软盘，硬盘，磁带，任何其他磁介质，CD-ROM，DVD或DVD-ROM，任何其他光学介质，打孔卡纸磁带，带孔图案的任何其他物理存储介质，RAM，ROM，PROM和EPROM，FLASH-EPROM，任何其他存储芯片或盒带，用于传输数据或指令的载波，用于传输此类载波的电缆或链接或计算机可以从中读取编程代码和/或数据的任何其他介质。这些形式的计算机可读介质中的许多可能涉及将一个或多个指令的一个或多个序列传送给处理器以供执行。

[0150] 计算机系统1301可以包括电子显示器1335或与其通信，该电子显示器1335包括用户界面(UI) 1340，该用户界面(UI) 1340提供例如本文其他地方所述的图形用户界面。UI的示例包括但不限于图形用户界面(GUI) 和基于Web的用户界面。

[0151] 本公开的方法和系统可以通过一种或多种算法来实现。可以由中央处理单元1305在执行时通过软件来实现算法。该算法可以例如是训练后的模型，诸如预测模型。

[0152] 在一些实施方式中，可以根据不是使用AI算法生成的预定数据传输方案将交通工具数据的至少一部分传输到远程实体(例如，云应用)。例如，在一些情形中，当数据传输不频繁或要传输的数据量相对较小时，可以基于来自云应用的请求来生成数据传输方案，而无需使用机器学习模型。在这种情形中，数据传输可以由中间实体(例如，原始装备制造厂商(OEM))管理，该中间实体在远程实体与驻留在目标交通工具上的数据编排器之间处理/传递请求和响应。中间实体可以充当代理，以在远程实体和数据编排器之间传递未修改的或处理后的数据传输请求/响应。在一些情况下，中间实体可以基于请求来确定一个或多个目标交通工具以传输交通工具数据。在一些情况下，中间实体还可以聚合或组装交通工具数据的至少一部分，并将其发送到请求应用。在一些情况下，中间实体可以生成描述交通工具数据和/或关于传输的信息(例如，数据源、数据处理方法等)的元数据，并将该元数据传输到请求应用。图16示意性地示出了根据本发明的一些实施方式的借助于OEM 1630管理的数据传输。

[0153] OEM 1630可以管理基本的交通工具数据和功能性。OEM 1630可以直接与远程实体(诸如一个或多个云应用、企业云或其他第三方实体1640-1、1640-2)进行通信，如本文其他地方所述。OEM可以提供运行时软件组件或基本软件服务，诸如感知(例如，ASIC、FPGA、GPU加速器、SIMD存储器、传感器/检测器(诸如相机、激光雷达、雷达、GPS等))、本地化和规划(例如，数据路径处理、DDR存储器、本地化数据集、惯性测量、GNSS)、决策或行为(例如运动引擎、ECC存储器、行为模块、仲裁、预测器)、控制(例如锁步处理器、DDR存储器、安全性监视、故障安全后备、线控控制器)、连接性和I/O(例如RF处理器、网络交换机、确定性总线、数据记录)。OEM可以收集或管理由上述软件服务或传感器生成的远程信息处理数据。远程信息处理数据可以包括例如速度相关的数据(例如，剧烈加速、提速、频繁加速)、停车相关的数据(例如，剧烈制动、频繁停止、频繁制动)、转弯相关的数据(例如，剧烈转弯、转弯之前的加速、驶出之前的过刹车、急转弯)、与正常驾驶的路线相关的数据(例如，高速公路与本地道路、已知交通拥堵的区域、事故发生率高/低的区域)或其他数据(例如，疲劳的转弯、通常行驶在快速车道上、转弯指示器的使用)。OEM 1630可以与一个或多个交通工具1610-1、1610-2和/或一个或多个数据编排器1620-1、1620-2通信。

[0154] 在一些实施方式中，诸如OEM之类的中间实体可以管理数据和知识管理系统，该数据和知识管理系统被配置为确定来自预测模型管理模块的哪个(哪些)预测模型发送给被

选交通工具或交通工具车队,以及哪个(哪些)组件可以接收这些模型。在一些情况下,可以通过云订阅模块将(一个或多个)模型OTA传输到(一个或多个)相关交通工具。在一些情况下,远程应用可以通过向OEM发送请求来从一个或多个交通工具请求数据。例如,为了发现欺诈行为、创建新的保险产品、为驾驶员提供安全性特征折扣、评估风险、事故现场管理、第一损失通知、增强索赔流程等,保险应用可以以预定频率(例如一周、两周、一个月、两个月等)从与目标交通工具相关联的OEM系统请求某些类型的数据(例如,由OEM嵌入式设备所收集的数据)。然后,OEM可以将请求传递给与目标交通工具相关联的数据编排器,以协调数据传输。被请求的数据类型可以直接从数据编排器传输到请求应用1640-1、1640-2。

[0155] 图17示出了数据编排器1720与一个或多个云应用之间的数据传输过程。在一些情况下,可以借助于诸如交通工具OEM 1730之类的中间实体来协调和管理一个或多个数据编排器与一个或多个云应用之间的数据传输。如本文中其他地方所述,数据编排器可以驻留在交通工具本地。

[0156] 在一些实施方式中,一个或多个云应用可以将请求某些类型的交通工具数据的请求1710发送到交通工具OEM 1730。例如,请求1710可以包含有关应用(例如,App 1)所需的数据类型、所需数据的频率、要传输此类数据的时间段的信息或其他信息,诸如目标交通工具识别号。例如,请求应用App 1(例如,保险应用)可以向与目标交通工具相关联的交通工具OEM 1730发送指示需要来自目标交通工具的数据类型和此类数据的频率的请求。

[0157] 交通工具OEM 1730可以将请求1711(例如,发送请求消息以中继该请求)传递给目标交通工具的数据编排器。传递给数据编排器的请求1711可以是与原始请求1710相同的未修改的请求。可替代地或另外,交通工具OEM 1730可以处理从云应用App 1接收到的请求1710并确定哪些交通工具/数据编排器是接收请求1711的目标交通工具/数据编排器。例如,原始请求1710可以从一种类型的交通工具请求远程信息处理数据以用于增强索赔过程,而无需指定目标交通工具(例如,不知道交通工具ID),然后交通工具OEM 1730可以识别满足交通工具类型要求的目标交通工具,并将请求1711发送给所识别的目标交通工具/数据编排器。在一些情况下,请求可以指定一组交通工具。例如,请求可以指定特定型号(例如,奥迪A8)、型号年份(例如,2017)、具有特定驾驶自动化特征的型号(例如,具有车道变更监控器的A8)等。OEM系统可以将请求传递给相应目标交通工具的数据编排器(例如,发送请求消息以中继该请求)。如上所述,交通工具OEM可以充当代理,以在数据编排器和发出请求的应用程序之间传递请求和响应。由于交通工具ID或其他交通工具信息可能不暴露给第三方(例如,云应用),因此这可以有利的增加一层安全性。

[0158] 在接收到请求1711时,数据编排器可以将请求推送到队列,并将消息发送回交通工具OEM 1730以确认请求的接收。交通工具OEM1730然后可以向请求应用发送消息(即,响应),指示该请求已被日志记录。

[0159] 与目标交通工具相关联的一个或多个数据编排器可以基于包含在请求1711中的信息来将被请求的交通工具数据传输到请求应用。例如,一个或多个数据编排器可以将被请求的数据(例如,数据分组)直接发送到请求应用。

[0160] 在一些情况下,除了传递和中继请求/响应消息外,交通工具OEM还可以发送指令以协调数据传输。例如,交通工具OEM可以向数据编排器发送消息,该消息命令数据编排器在传输时段完成时(例如,在从数据编排器接收到完成消息后)从队列中删除传输请求。然

后,数据编排器可以从队列中删除该条目,并向交通工具OEM发送消息,该消息指示该条目被删除。OEM系统可以向请求应用发送消息,该消息指示请求已完成。

[0161] 尽管已经在本文中示出和描述了本发明的优选实施方式,但是对于本领域技术人员而言容易理解的是,仅通过示例的方式提供了这样的实施方式。并非意图通过说明书中提供的特定示例来限制本发明。尽管已经参考前述说明书描述了本发明,但是本文中的实施方式的描述和图示并不意味着以限制性的意义来解释。在不脱离本发明的情况下,本领域技术人员现在将想到许多变化,改变和替代。此外,应理解,本发明的所有方面不限于本文所阐述的具体描述,构造或相对比例,其取决于各种条件和变量。应当理解,本文所述的本发明的实施方案的各种替代方案可以用于实施本发明。因此,可以预期的是,本发明也将涵盖任何这样的替代,修改,变化或等同形式。旨在由以下权利要求书限定本发明的范围,并且由此涵盖这些权利要求书范围内的方法和结构及其等同物。

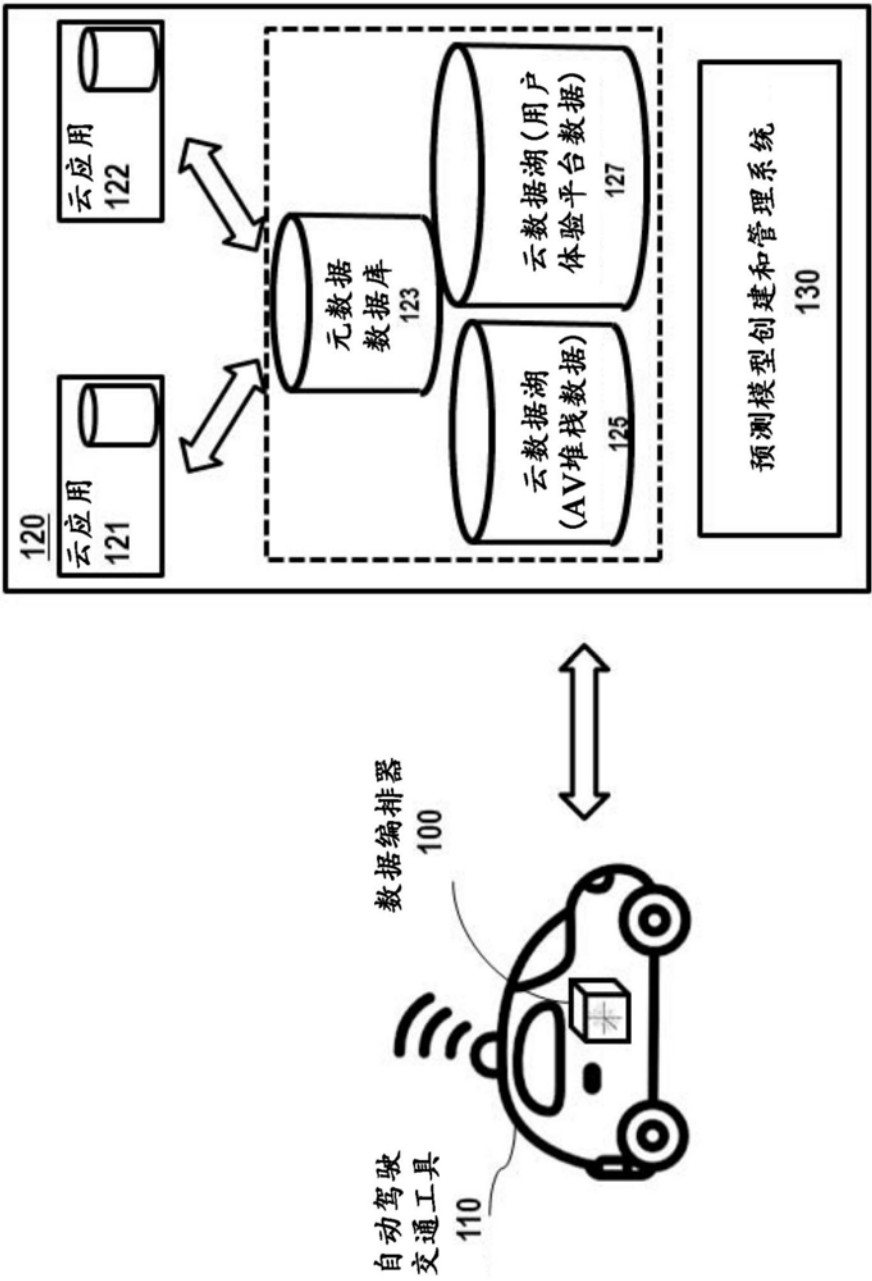


图1

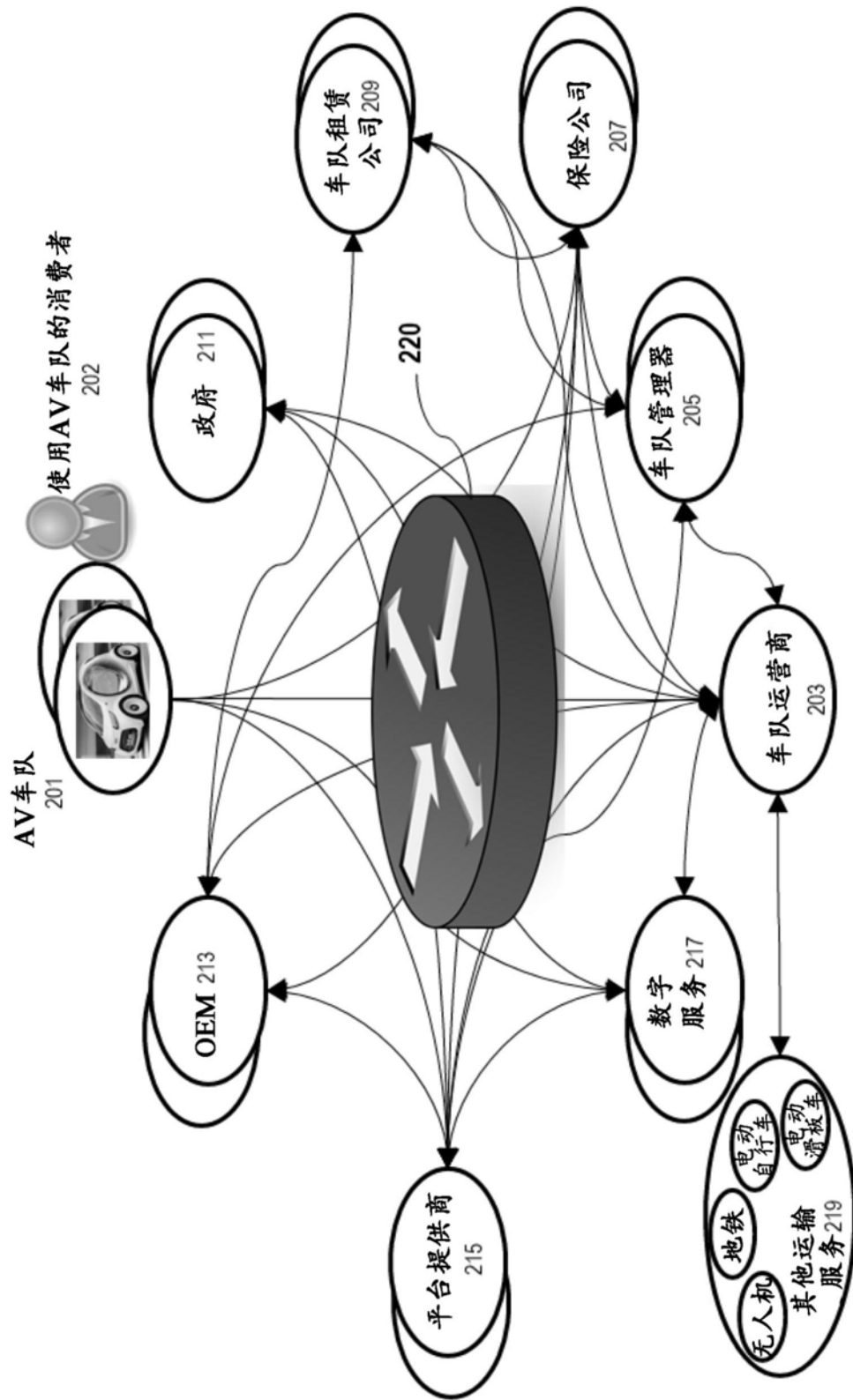


图2

300

应用名称	传输标志	交通工具ID	请求数据 中心应用	数据类型	所需数据	传输延迟	压缩类型	加密类型	监管规则
Pothole 检测器	设置	字符串	(app_name , loc)	视频流	流1	0	无损	RSA	空

图3

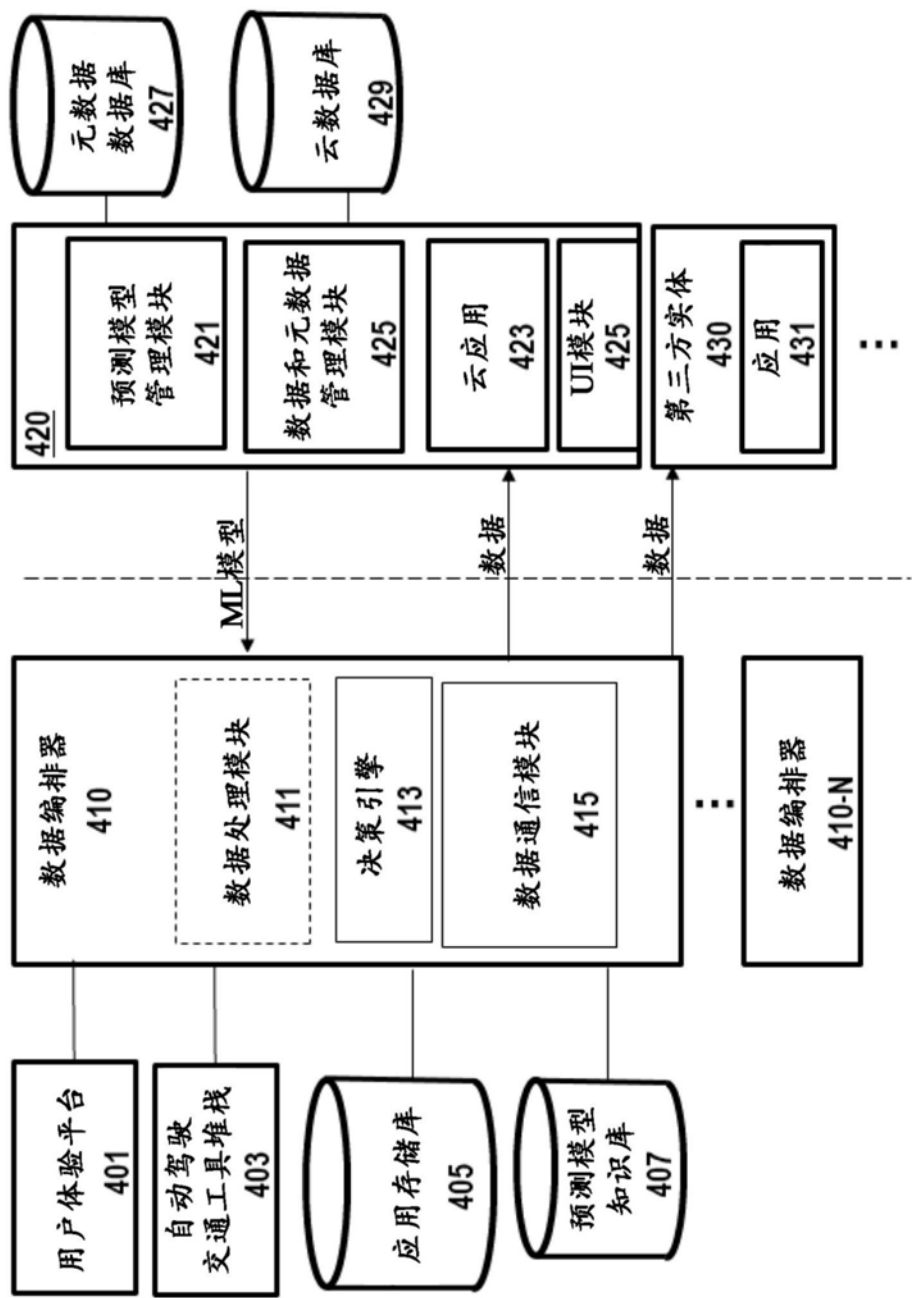


图4

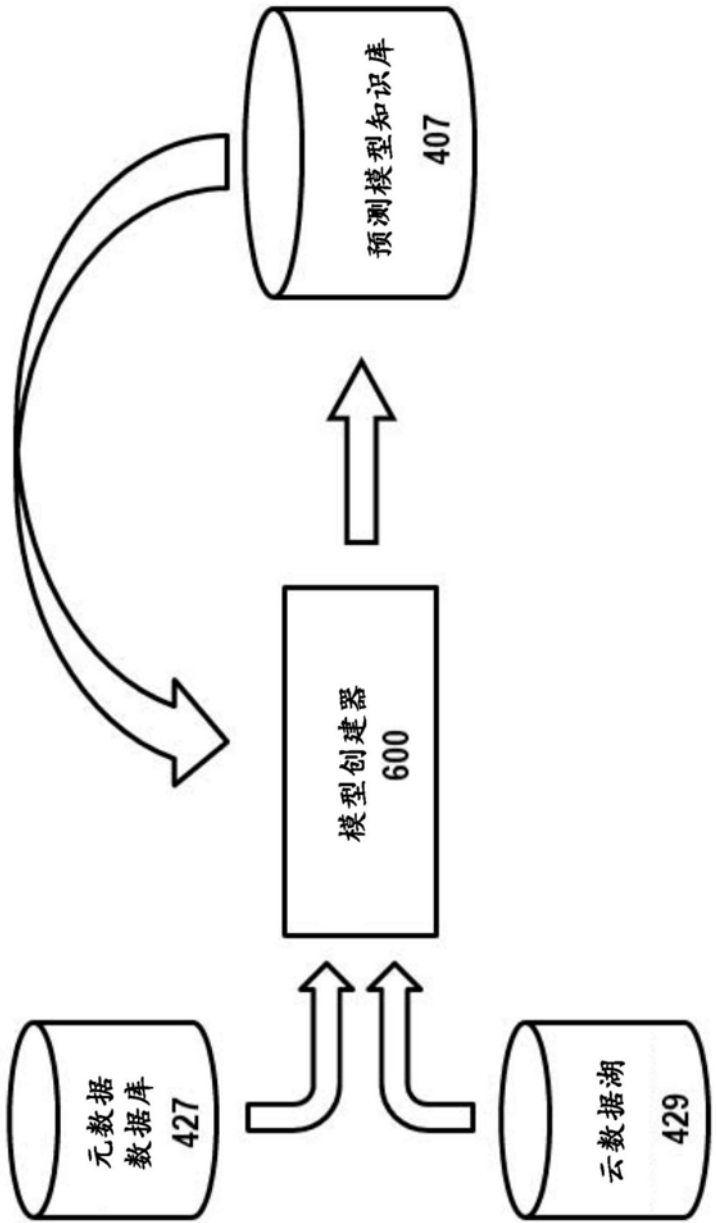


图6

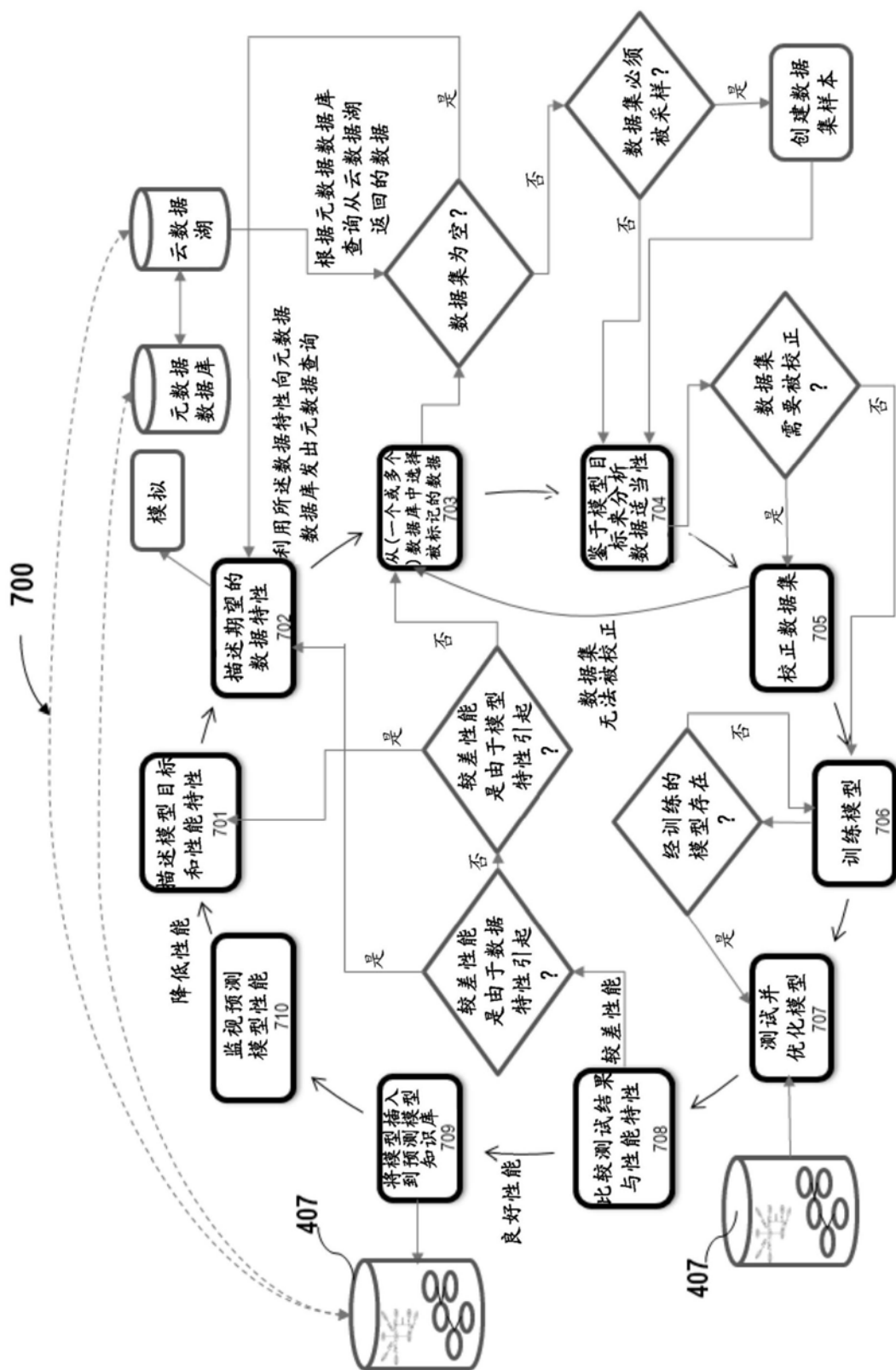


图7

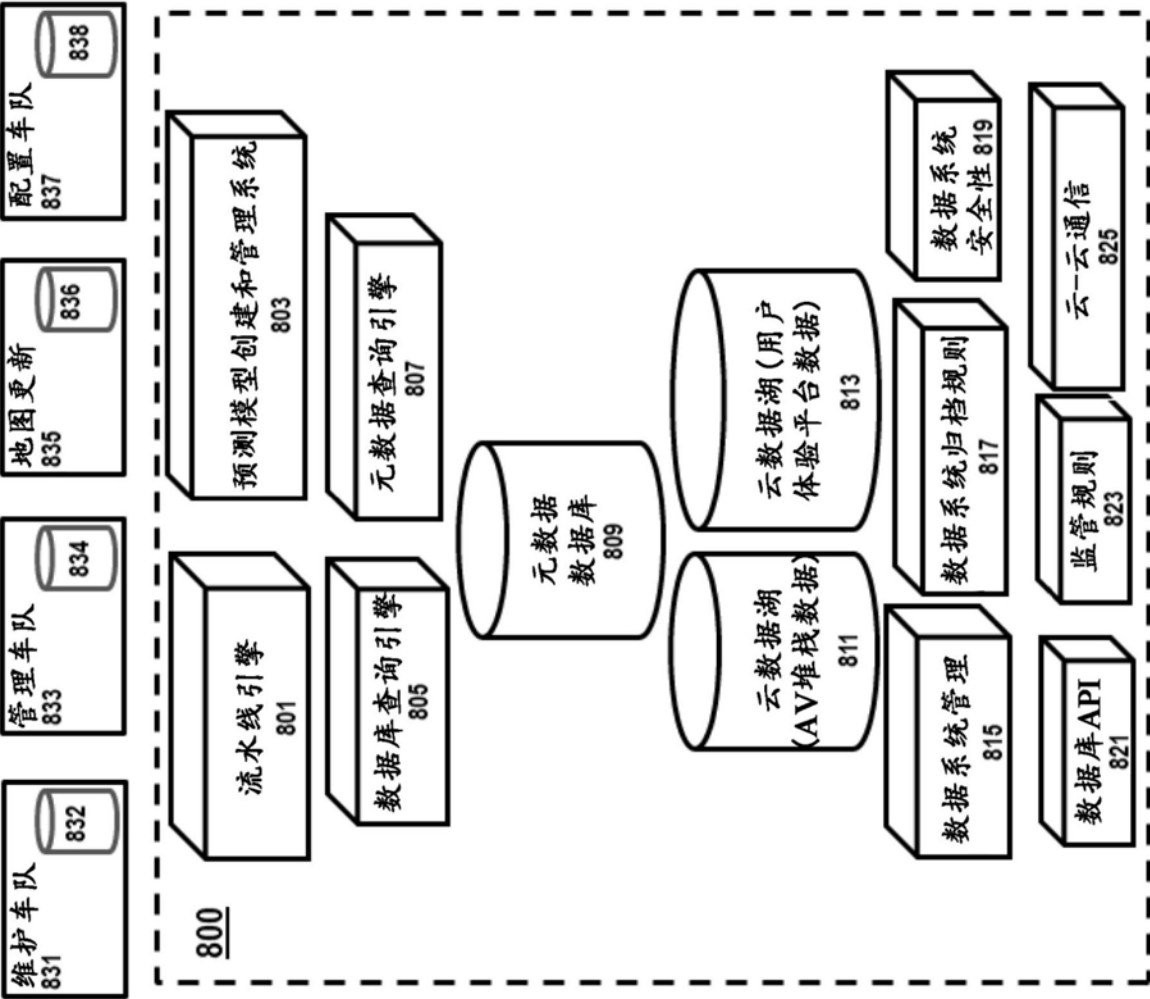


图8

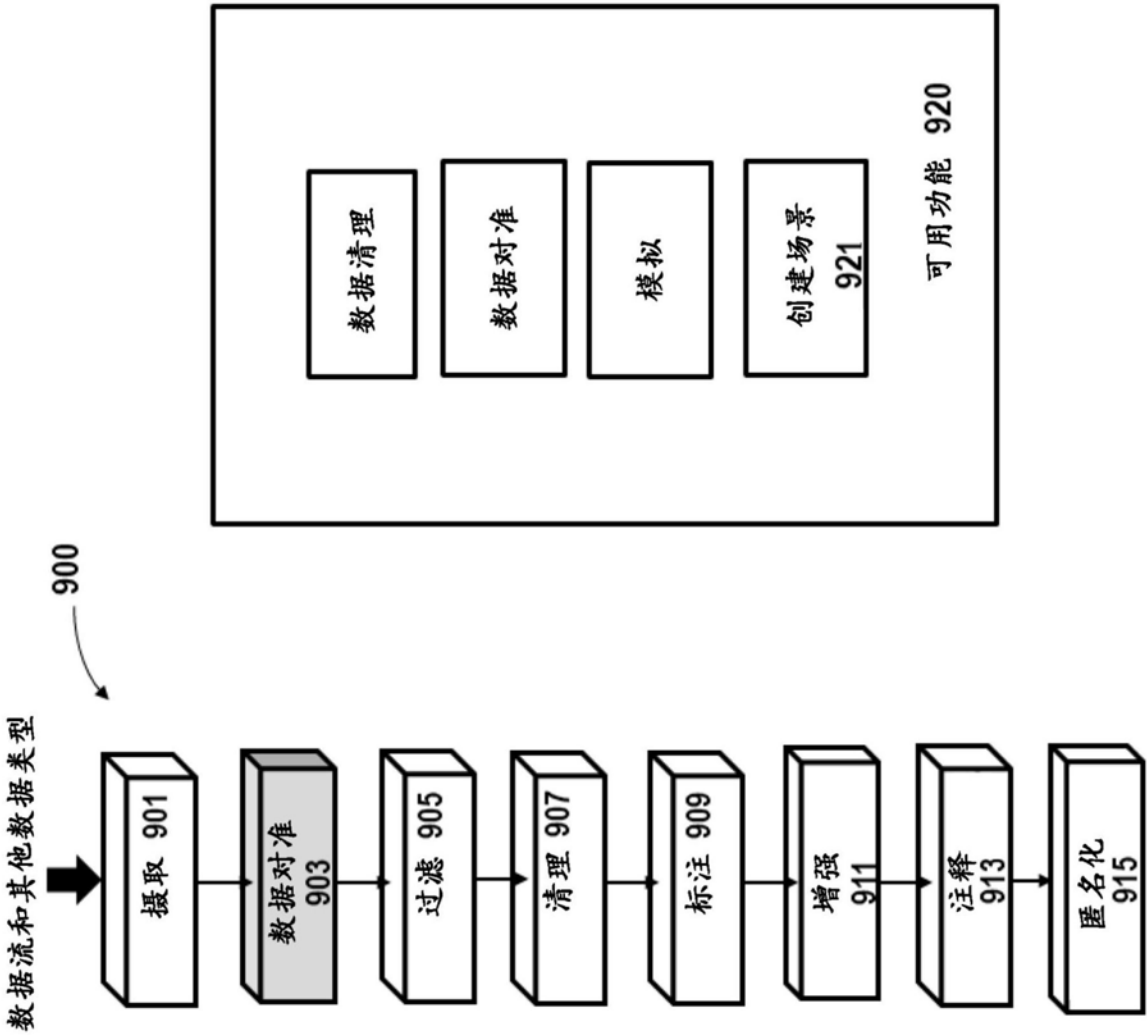


图9

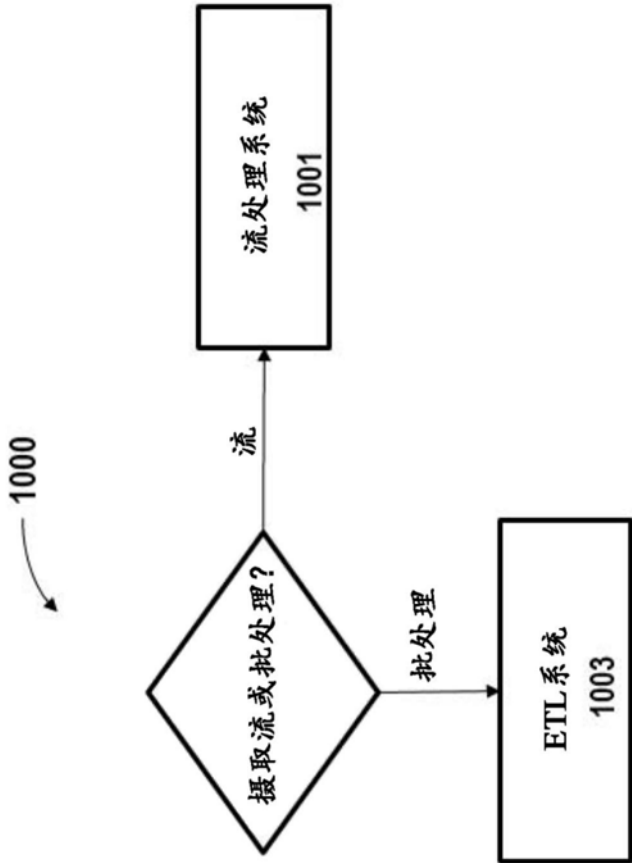


图10

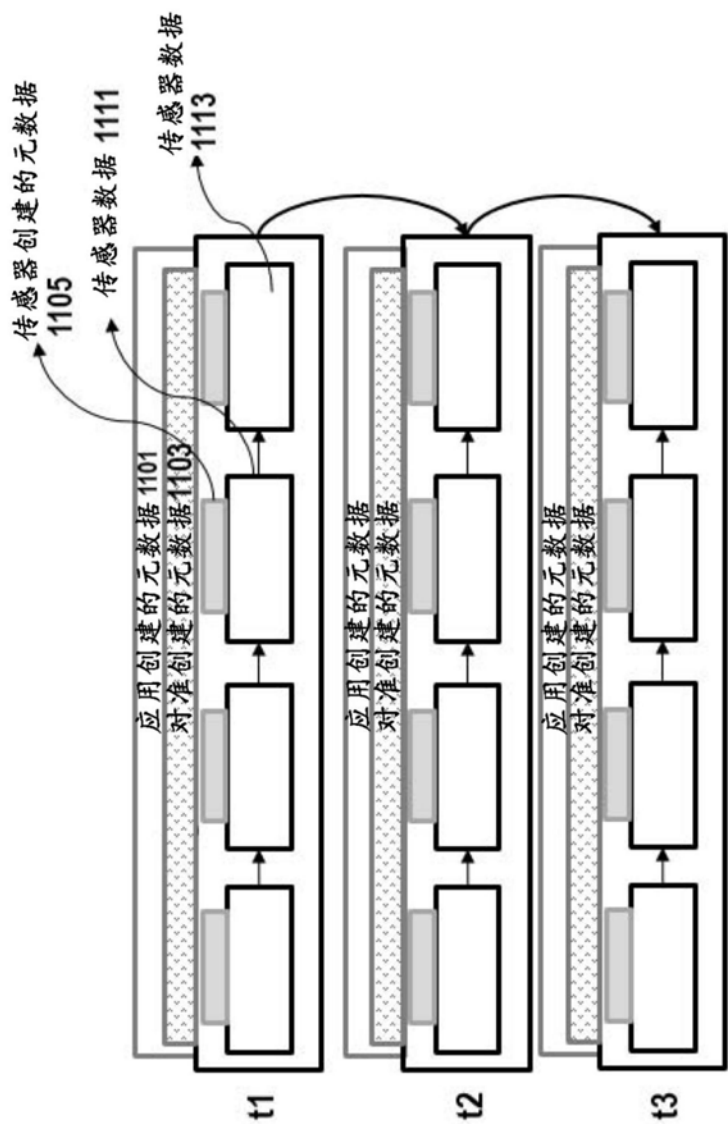


图11

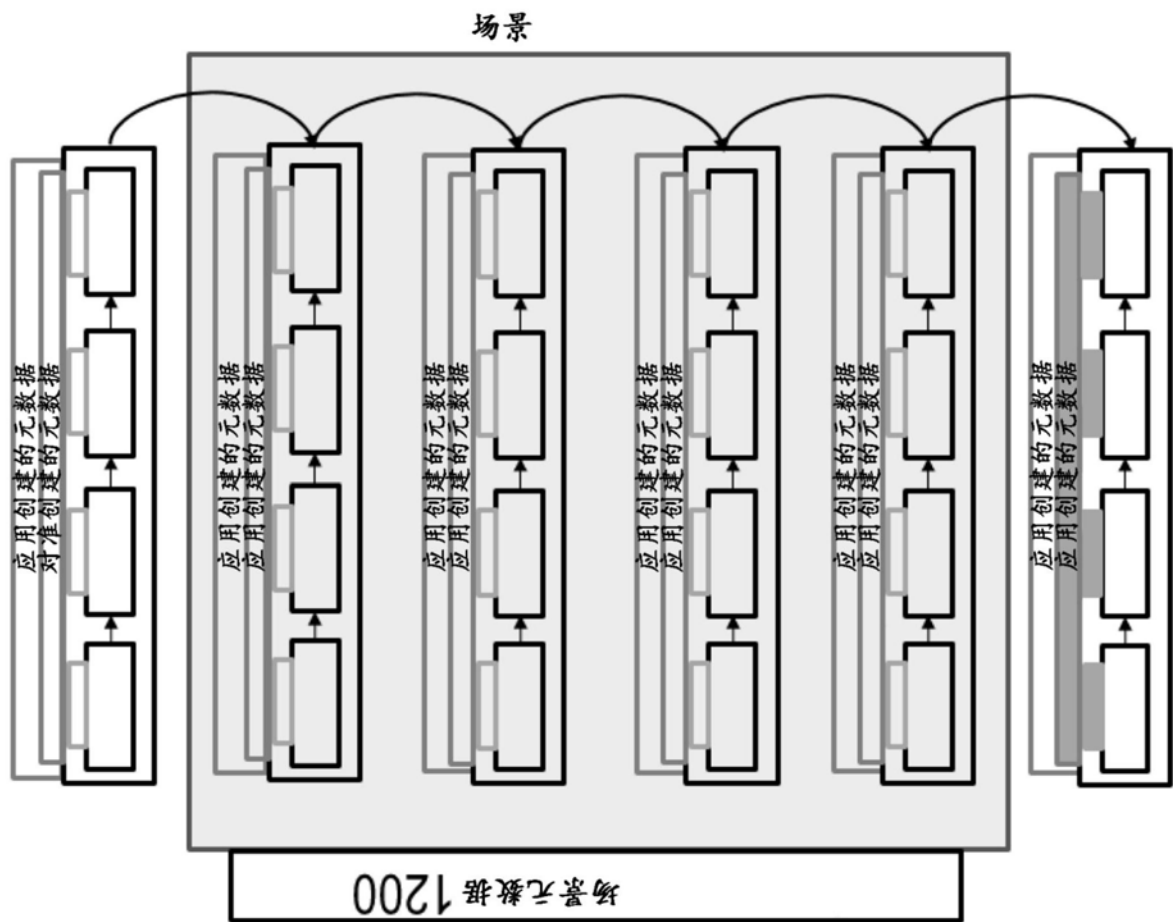


图12

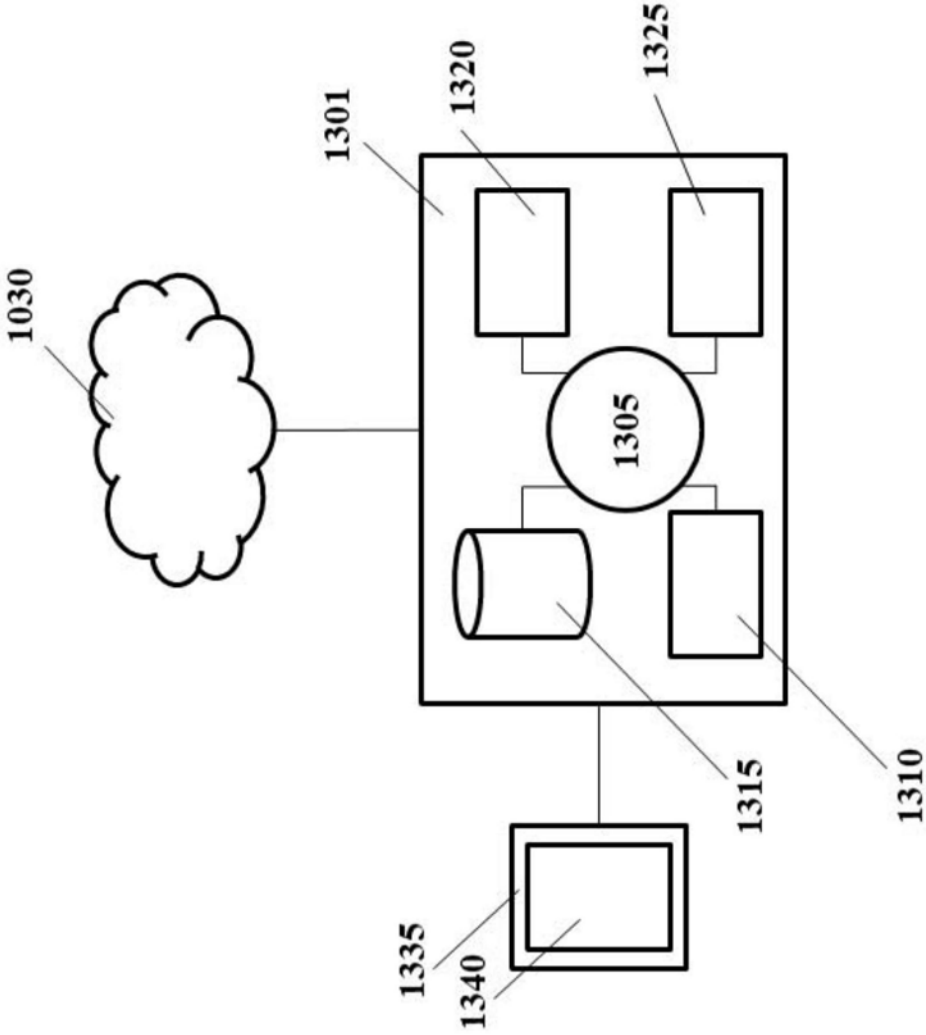


图13

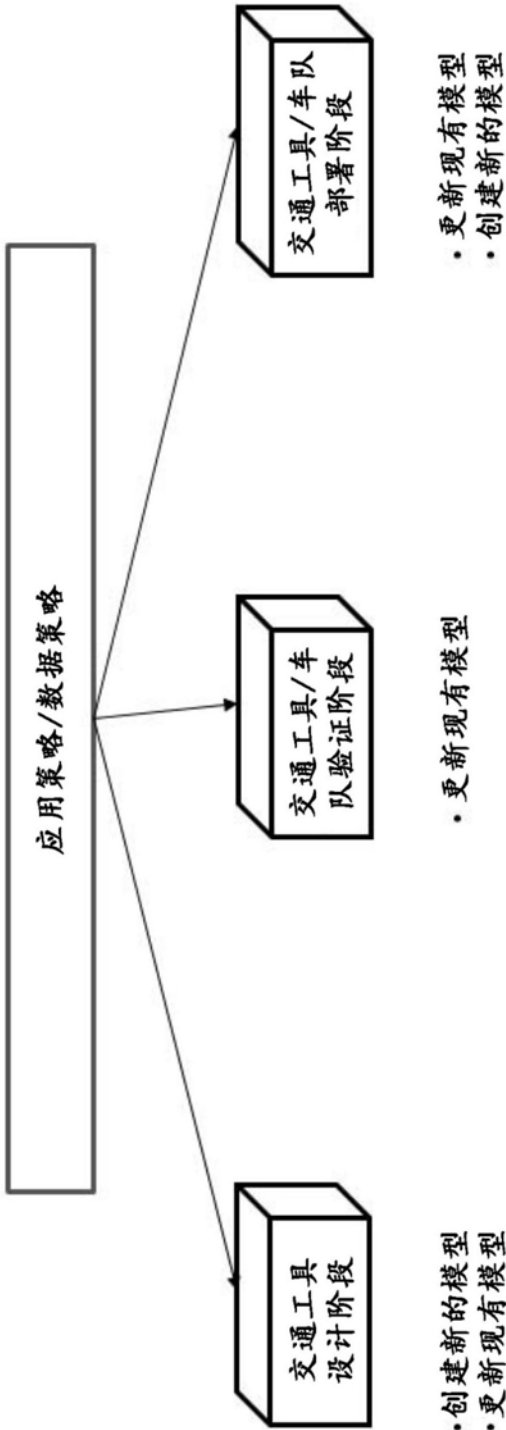


图14

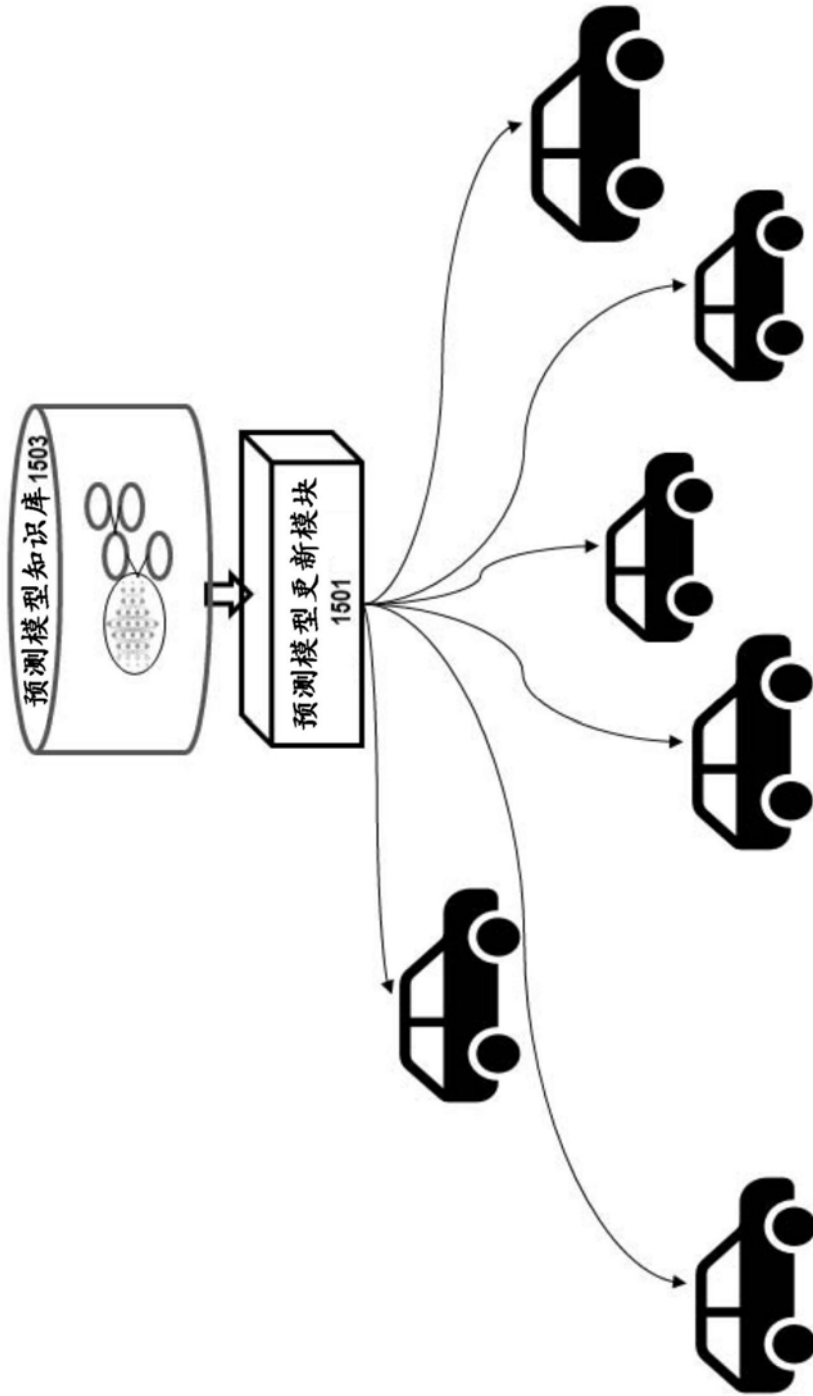


图15

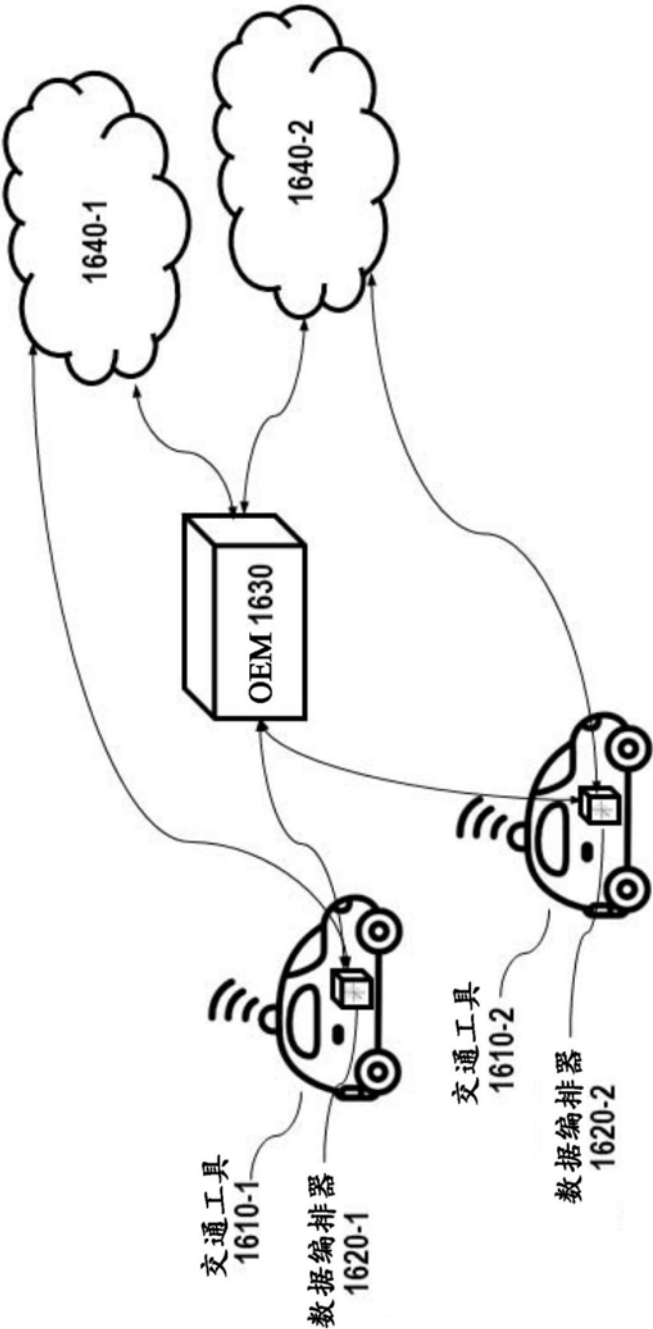


图16

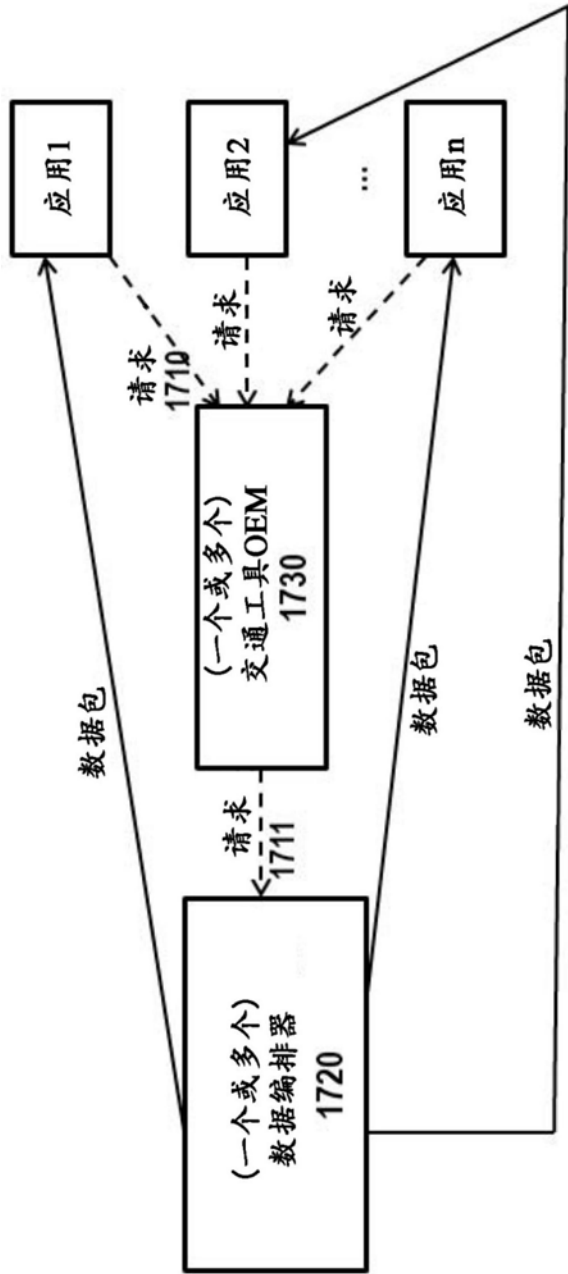


图17

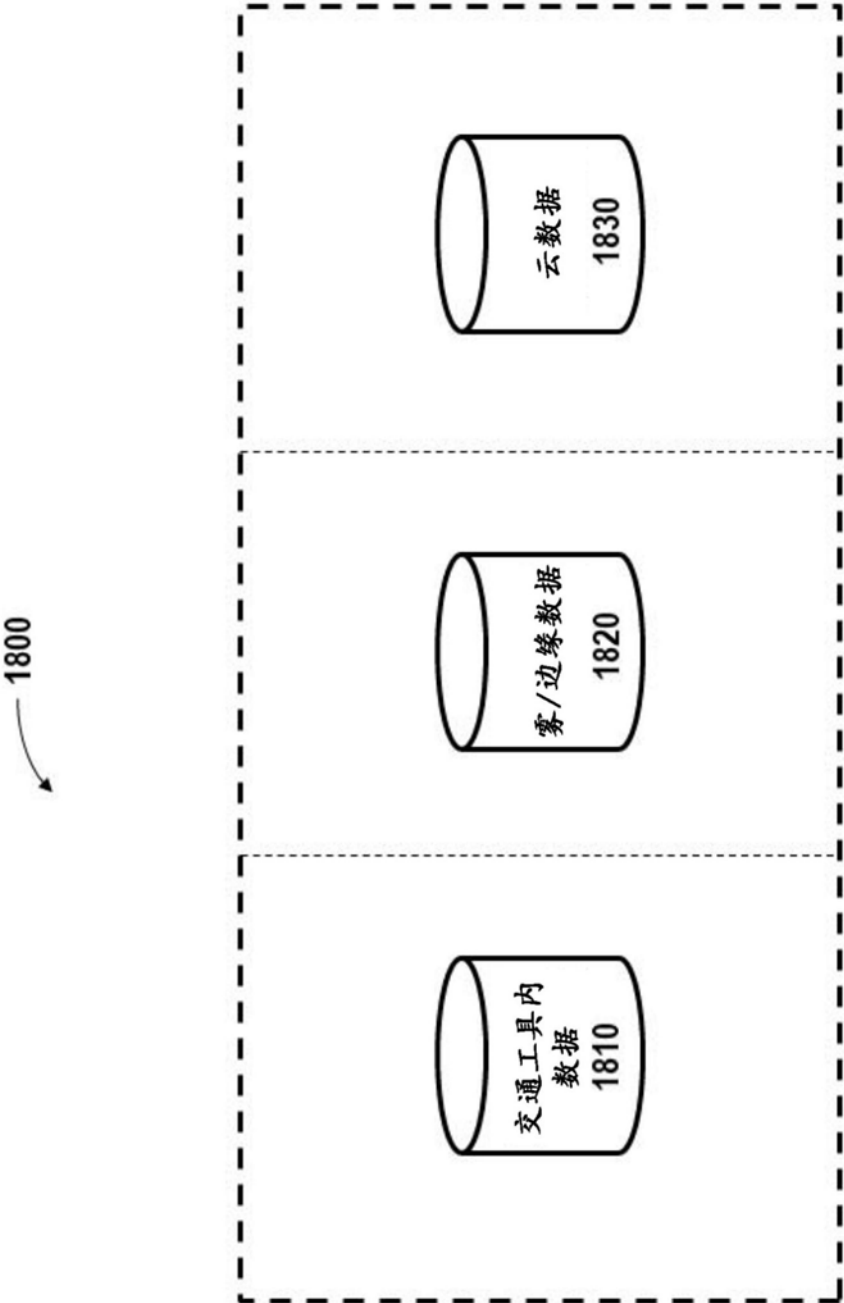


图18

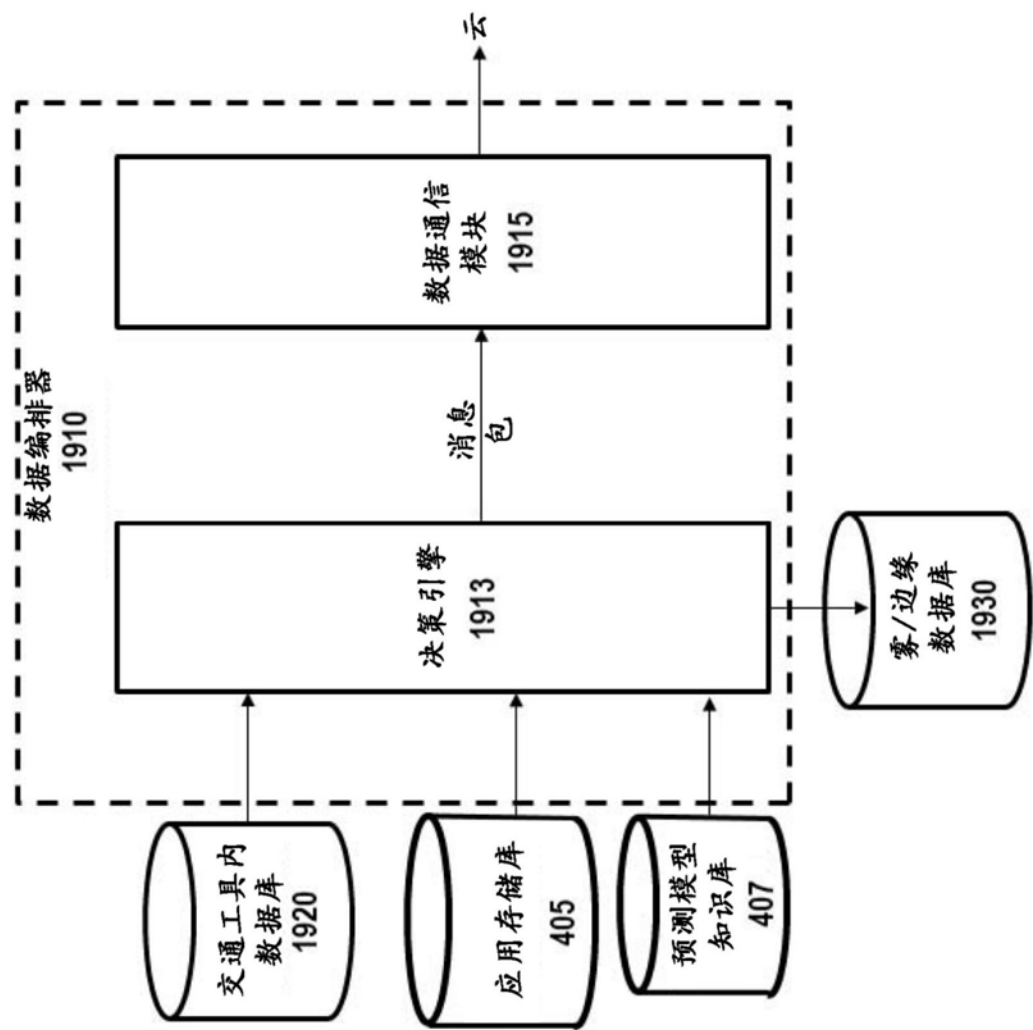


图19

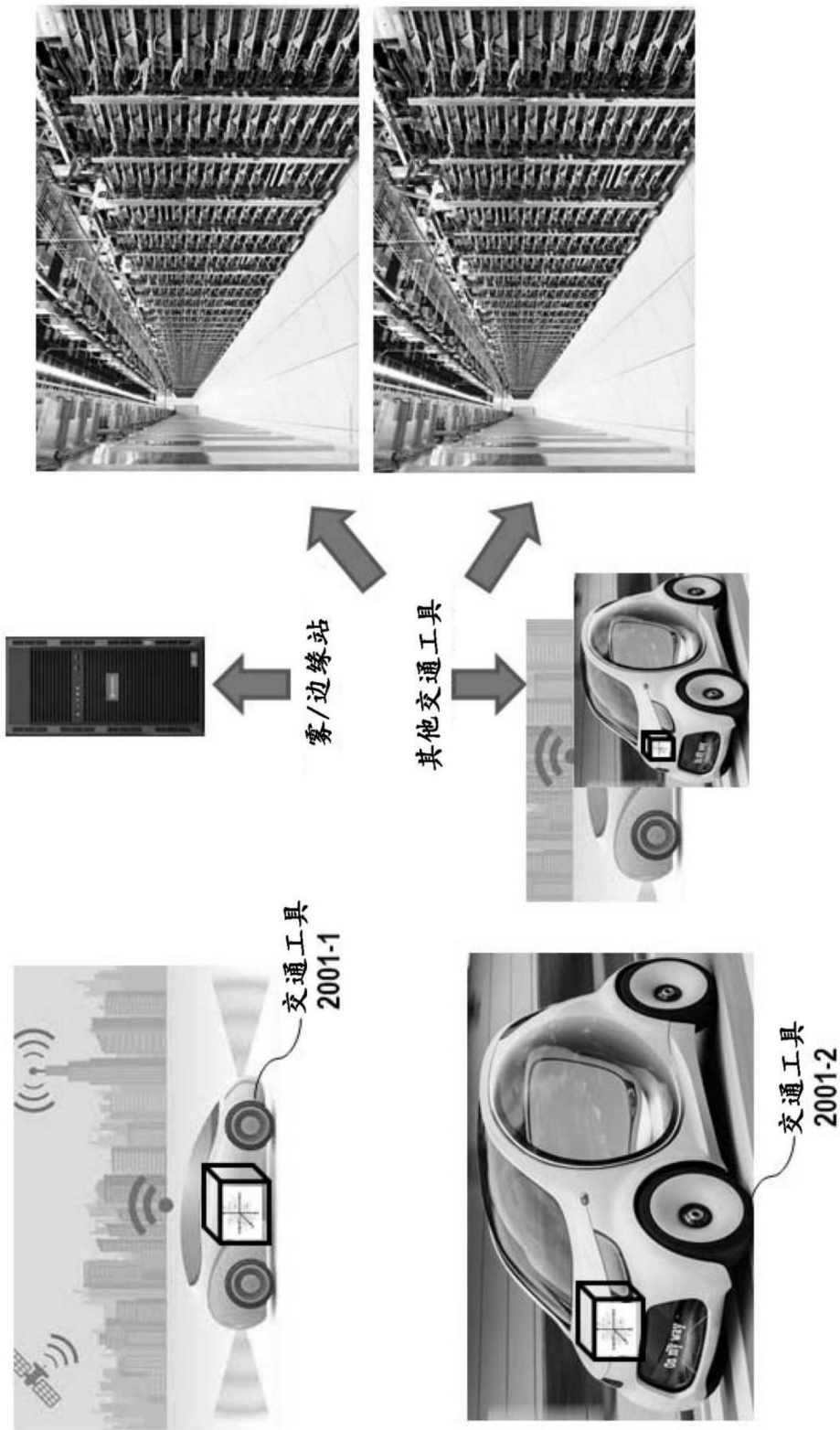


图20