



(12) 发明专利

(10) 授权公告号 CN 116894055 B

(45) 授权公告日 2023. 11. 17

(21) 申请号 202311165793.8

(22) 申请日 2023.09.11

(65) 同一申请的已公布的文献号
申请公布号 CN 116894055 A

(43) 申请公布日 2023.10.17

(73) 专利权人 中国铁道科学研究院集团有限公
司电子计算技术研究所

地址 100000 北京市海淀区大柳树路2号一
区12幢2层201-203、206

专利权人 北京经纬信息技术有限公司

(72) 发明人 贾晓非 阎胜勇 惠伟 甄津
常灿 王少华 郑慧亚 凡凯乐

(74) 专利代理机构 成都创典汇智知识产权代理
事务所(普通合伙) 51383

专利代理师 张一帆

(51) Int. Cl.

G06F 16/2455 (2019.01)

G06F 16/29 (2019.01)

G06Q 10/0637 (2023.01)

G06Q 50/30 (2012.01)

(56) 对比文件

CN 116703282 A, 2023.09.05

CN 116151556 A, 2023.05.23

CN 115384586 A, 2022.11.25

CN 114889678 A, 2022.08.12

US 2020312147 A1, 2020.10.01

审查员 李萌

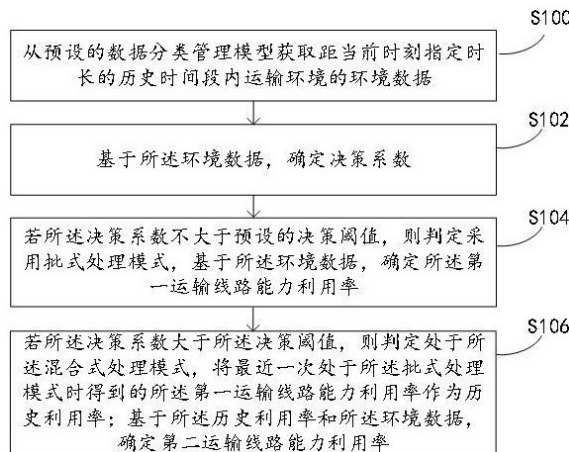
权利要求书2页 说明书11页 附图2页

(54) 发明名称

基于大数据批流一体的铁路线路能力利用率计算方法

(57) 摘要

本申请公开了基于大数据批流一体的铁路线路能力利用率计算方法,以及执行该方法的系统,结合了批处理和流处理的优点,实现基于运输线路能力利用率的运输环境调整。本说明书中的方法及系统,针对铁路运输环境的多层级性、复杂性和多变性,没有仅仅采用单独的某一种数据处理方式,也没有通过复杂的接口设计将不同的数据处理方式进行嫁接,而是结合铁路运输环境的实际情况,在适于进行批处理时,采用批处理模式进行数据处理。在不适于进行批处理时,采用了混合式处理模式,以历史上进行批处理时得到的运输线路能力利用率作为考察的一个方面,使得历史上的批处理的结果能够一定程度的体现在针对当前的实时运输环境的评价上。



1. 基于大数据批流一体的铁路线路能力利用率计算方法,其特征在于,所述方法由基于大数据批流一体的铁路线路能力利用率计算系统执行,所述方法包括:

从预设的数据分类管理模型获取距当前时刻指定时长的历史时间段内运输环境的环境数据;其中,所述运输环境包含:运输线路和相邻的两条所述运输线路之间的站点;所述环境数据包含:所述运输线路单向通车数量、所述运输线路双向通车对数、所述运输线路客运密度、所述运输线路货运密度、所述运输线路相邻列车最长间隔、所述运输线路相邻列车最短间隔、所述站点发车历史客座率以及所述站点最大历史客流量;

基于所述环境数据,确定决策系数;其中,所述决策系数与所述运输线路双向通车对数和所述运输线路单向通车数量的比值负相关、与所述运输线路货运密度和所述运输线路客运密度的比值负相关、与所述运输线路相邻列车最长间隔和所述运输线路单向通车数量的最大值的乘积正相关、与所述运输线路相邻列车最短间隔正相关;

若所述决策系数不大于预设的决策阈值,则判定采用批式处理模式基于所述环境数据,确定第一运输线路能力利用率;其中,所述第一运输线路能力利用率用于表征所述运输环境相较于满负荷状态的使用程度;所述第一运输线路能力利用率与所述决策系数正相关,且与所述运输环境的综合相对占用率负相关;所述综合相对占用率与所述站点最大历史客流量正相关,且与最近一次处于混合式处理模式持续的时长正相关;将所述第一运输线路能力利用率,发送至所述站点,使得所述站点基于所述第一运输线路能力利用率对未来时刻中的、所述站点所属的所述运输环境进行管理;

若所述决策系数大于所述决策阈值,则判定处于所述混合式处理模式;将最近一次处于所述批式处理模式时得到的所述第一运输线路能力利用率作为历史利用率;基于所述历史利用率和所述环境数据,确定第二运输线路能力利用率;其中,所述第二运输线路能力利用率与所述历史利用率正相关、与所述站点的发车实时客座率和所述站点发车历史客座率的比值正相关、且与所述发车实时客座率和所述站点的实时客流量的比值正相关;将所述第二运输线路能力利用率,发送至所述站点,使得所述站点基于所述第二运输线路能力利用率对未来时刻中的、所述站点所属的所述运输环境进行管理。

2. 如权利要求1所述方法,其特征在于,所述方法还包括:

若监测到所述运输线路上行驶的列车的数量大于预设的数量阈值,则从所述数据分类管理模型获取所述环境数据。

3. 如权利要求1所述方法,其特征在于,所述方法还包括:

所述决策阈值与所述环境数据表示出的所述运输线路上行驶的列车平均轴重负相关。

4. 如权利要求1所述方法,其特征在于,所述方法还包括:

所述决策阈值与参照值负相关;所述参照值与所述环境数据表示出的所述运输线路上行驶的列车的最大行驶速度和处于所述最大行驶速度所持续的时长的乘积正相关,且与所述列车处于所述最大行驶速度时,与所述列车处于同向行驶的其他列车的数量与所述运输线路的长度的比值正相关。

5. 如权利要求1所述方法,其特征在于,所述方法还包括:

所述指定时长与最近一次处于所述混合式处理模式持续的时长正相关。

6. 如权利要求1所述方法,其特征在于,所述方法还包括:

所述运输环境是客货混用运输环境。

7. 基于大数据批流一体的铁路线路能力利用率计算系统,其特征在于,所述系统包括:

数据获取模块,配置为:从预设的数据分类管理模型获取距当前时刻指定时长的历史时间段内运输环境的环境数据;其中,所述运输环境包含:运输线路和相邻的两条所述运输线路之间的站点;所述环境数据包含:所述运输线路单向通车数量、所述运输线路双向通车对数、所述运输线路客运密度、所述运输线路货运密度、所述运输线路相邻列车最长时间间隔、所述运输线路相邻列车最短时间间隔、所述站点发车历史客座率以及所述站点最大历史客流量;

决策系数确定模块,配置为:基于所述环境数据,确定决策系数;其中,所述决策系数与所述运输线路双向通车对数和所述运输线路单向通车数量的比值负相关、与所述运输线路货运密度和所述运输线路客运密度的比值负相关、与所述运输线路相邻列车最长时间间隔和所述运输线路单向通车数量的最大值的乘积正相关、与所述运输线路相邻列车最短时间间隔正相关;

第一运输线路能力利用率确定模块,配置为:若所述决策系数不大于预设的决策阈值,则判定采用批式处理模式基于所述环境数据,确定第一运输线路能力利用率;其中,所述第一运输线路能力利用率用于表征所述运输环境相较于满负荷状态的使用程度;所述第一运输线路能力利用率与所述决策系数正相关,且与所述运输环境的综合相对占用率负相关;所述综合相对占用率与所述站点最大历史客流量正相关,且与最近一次处于混合式处理模式持续的时长正相关;将所述第一运输线路能力利用率,发送至所述站点,使得所述站点基于所述第一运输线路能力利用率对未来时刻中的、所述站点所属的所述运输环境进行管理;

第二运输线路能力利用率确定模块,配置为:若所述决策系数大于所述决策阈值,则判定处于所述混合式处理模式;将最近一次处于所述批式处理模式时得到的所述第一运输线路能力利用率作为历史利用率;基于所述历史利用率和所述环境数据,确定第二运输线路能力利用率;其中,所述第二运输线路能力利用率与所述历史利用率正相关、与所述站点的发车实时客座率和所述站点发车历史客座率的比值正相关、且与所述发车实时客座率和所述站点的实时客流量的比值正相关;将所述第二运输线路能力利用率,发送至所述站点,使得所述站点基于所述第二运输线路能力利用率对未来时刻中的、所述站点所属的所述运输环境进行管理。

8. 一种电子设备,包括:

处理器;以及

被安排成存储计算机可执行指令的存储器,所述可执行指令在被执行时使所述处理器执行所述权利要求1~6之任一所述方法。

9. 一种计算机可读存储介质,所述计算机可读存储介质存储一个或多个程序,所述一个或多个程序当被包括多个应用程序的电子设备执行时,使得所述电子设备执行所述权利要求1~6之任一所述方法。

基于大数据批流一体的铁路线路能力利用率计算方法

技术领域

[0001] 本申请涉及大数据处理技术领域,尤其涉及基于大数据批流一体的铁路线路能力利用率计算方法。

背景技术

[0002] 我国铁路运输飞速发展,无论是线路铺设、还是运输量等方面,都有所提高。一方面,一定程度的满足了人们对铁路运输的需求;另一方面,也增加了运输环境的复杂程度,为铁路运输的管理带来了一定的难度。该难度具体可以体现为:需要处理的数据量较大。在相关技术领域中,针对大数据可采用批处理,也可以采用流处理,两者的数据处理效果各有千秋。然而,我国幅员辽阔,铁路分布甚广,不同地区的铁路分布情况也不同,此外,不同地区的站点、运输线路的管理方式也不尽相同,难以统一的执行批处理或者流处理。

[0003] 可见,如何结合铁路运输的场景,较为合理地进行大数据处理,成为亟待解决的问题。

发明内容

[0004] 本申请实施例提供了基于大数据批流一体的铁路线路能力利用率计算方法,以至少部分的解决上述技术问题。

[0005] 本申请实施例采用下述技术方案:

[0006] 第一方面,本申请实施例提供基于大数据批流一体的铁路线路能力利用率计算方法,所述方法由基于大数据批流一体的铁路线路能力利用率计算系统执行,所述方法包括:

[0007] 从预设的数据分类管理模型获取距当前时刻指定时长的历史时间段内运输环境的环境数据;其中,所述运输环境包含:运输线路和相邻的两条所述运输线路之间的站点;所述环境数据包含:所述运输线路单向通车数量、所述运输线路双向通车对数、所述运输线路客运密度、所述运输线路货运密度、所述运输线路相邻列车最长时间间隔、所述运输线路相邻列车最短时间间隔、所述站点发车历史客座率以及所述站点最大历史客流量;

[0008] 基于所述环境数据,确定决策系数;其中,所述决策系数与所述运输线路双向通车对数和所述运输线路单向通车数量的比值负相关、与所述运输线路货运密度和所述运输线路客运密度的比值负相关、与所述运输线路相邻列车最长时间间隔和所述运输线路单向通车数量的最大值的乘积正相关、与所述运输线路相邻列车最短时间间隔正相关;

[0009] 若所述决策系数不大于预设的决策阈值,则判定采用批式处理模式基于所述环境数据,确定第一运输线路能力利用率;其中,所述第一运输线路能力利用率用于表征所述运输环境相较于满负荷状态的使用程度;所述第一运输线路能力利用率与所述决策系数正相关,且与所述运输环境的综合相对占用率负相关;所述综合相对占用率与所述站点最大历史客流量正相关,且与最近一次处于混合式处理模式持续的时长正相关;将所述第一运输线路能力利用率,发送至所述站点,使得所述站点基于所述第一运输线路能力利用率对未来时刻中的、所述站点所属的所述运输环境进行管理;

[0010] 若所述决策系数大于所述决策阈值,则判定处于所述混合式处理模式;将最近一次处于所述批式处理模式时得到的所述第一运输线路能力利用率作为历史利用率;基于所述历史利用率和所述环境数据,确定第二运输线路能力利用率;其中,所述第二运输线路能力利用率与所述历史利用率正相关、与所述站点的发车实时客座率和所述站点发车历史客座率的比值正相关、且与所述发车实时客座率和所述站点的实时客流量的比值正相关;将所述第二运输线路能力利用率,发送至所述站点,使得所述站点基于所述第二运输线路能力利用率对未来时刻中的、所述站点所属的所述运输环境进行管理。

[0011] 在本说明书一个可选的实施例中,所述方法还包括:

[0012] 若监测到所述运输线路上行驶的列车的数量大于预设的数量阈值,则从所述数据分类管理模型获取所述环境数据。

[0013] 在本说明书一个可选的实施例中,所述方法还包括:

[0014] 所述决策阈值与所述环境数据表示出的所述运输线路上行驶的列车平均轴重负相关。

[0015] 在本说明书一个可选的实施例中,所述方法还包括:

[0016] 所述决策阈值与参照值负相关;所述参照值与所述环境数据表示出的所述运输线路上行驶的列车的最大行驶速度和处于所述最大行驶速度所持续的时长的乘积正相关,且与所述列车处于所述最大行驶速度时,与所述列车处于同向行驶的其他列车的数量与所述运输线路的长度的比值正相关。

[0017] 在本说明书一个可选的实施例中,所述方法还包括:

[0018] 所述指定时长与最近一次处于所述混合式处理模式持续的时长正相关。

[0019] 在本说明书一个可选的实施例中,所述方法还包括:

[0020] 所述运输环境是客货混用运输环境。

[0021] 第二方面,本申请实施例还提供基于大数据批流一体的铁路线路能力利用率计算系统,所述系统包括:

[0022] 数据获取模块,配置为:从预设的数据分类管理模型获取距当前时刻指定时长的历史时间段内运输环境的环境数据;其中,所述运输环境包含:运输线路和相邻的两条所述运输线路之间的站点;所述环境数据包含:所述运输线路单向通车数量、所述运输线路双向通车对数、所述运输线路客运密度、所述运输线路货运密度、所述运输线路相邻列车最长时间间隔、所述运输线路相邻列车最短时间间隔、所述站点发车历史客座率以及所述站点最大历史客流量;

[0023] 决策系数确定模块,配置为:基于所述环境数据,确定决策系数;其中,所述决策系数与所述运输线路双向通车对数和所述运输线路单向通车数量的比值负相关、与所述运输线路货运密度和所述运输线路客运密度的比值负相关、与所述运输线路相邻列车最长时间间隔和所述运输线路单向通车数量的最大值的乘积正相关、与所述运输线路相邻列车最短时间间隔正相关;

[0024] 第一运输线路能力利用率确定模块,配置为:若所述决策系数不大于预设的决策阈值,则判定采用批式处理模式基于所述环境数据,确定第一运输线路能力利用率;其中,所述第一运输线路能力利用率用于表征所述运输环境相较于满负荷状态的使用程度;所述第一运输线路能力利用率与所述决策系数正相关,且与所述运输环境的综合相对占用率负

相关;所述综合相对占用率与所述站点最大历史客流量正相关,且与最近一次处于混合式处理模式持续的时长正相关;将所述第一运输线路能力利用率,发送至所述站点,使得所述站点基于所述第一运输线路能力利用率对未来时刻中的、所述站点所属的所述运输环境进行管理;

[0025] 第二运输线路能力利用率确定模块,配置为:若所述决策系数大于所述决策阈值,则判定处于所述混合式处理模式;将最近一次处于所述批式处理模式时得到的所述第一运输线路能力利用率作为历史利用率;基于所述历史利用率和所述环境数据,确定第二运输线路能力利用率;其中,所述第二运输线路能力利用率与所述历史利用率正相关、与所述站点的发车实时客座率和所述站点发车历史客座率的比值正相关、且与所述发车实时客座率和所述站点的实时客流量的比值正相关;将所述第二运输线路能力利用率,发送至所述站点,使得所述站点基于所述第二运输线路能力利用率对未来时刻中的、所述站点所属的所述运输环境进行管理。

[0026] 第三方面,本申请实施例还提供一种电子设备,包括:

[0027] 处理器;以及

[0028] 被安排成存储计算机可执行指令的存储器,所述可执行指令在被执行时使所述处理器执行第一方面所述的方法步骤。

[0029] 第四方面,本申请实施例还提供一种计算机可读存储介质,所述计算机可读存储介质存储一个或多个程序,所述一个或多个程序当被包括多个应用程序的电子设备执行时,使得所述电子设备执行第一方面所述的方法步骤。

[0030] 本申请实施例采用的上述至少一个技术方案能够达到以下有益效果:

[0031] 本申请提供一种基于大数据批流一体的铁路线路能力利用率计算方法,以及执行该方法的系统,结合了批处理和流处理的优点,实现基于运输线路能力利用率的运输环境调整。本说明书中的方法及系统,针对铁路运输环境的多层级性、复杂性和多变性,没有仅仅采用单独的某一种数据处理方式,也没有通过复杂的接口设计将不同的数据处理方式进行嫁接,而是结合铁路运输环境的实际情况,在适于进行批处理时,采用批处理模式进行数据处理。在不适于进行批处理时,采用了一种全新的数据处理模式,即混合式处理模式,以历史上进行批处理时得到的运输线路能力利用率(即,第一运输线路能力利用率)作为考察的一个方面,使得历史上的批处理的结果能够一定程度的体现在针对当前的实时运输环境的评价上,无需复杂的接口设计,不必考虑批处理和流处理之间的不相容性,有利于提高运输环境的管理效率。

附图说明

[0032] 此处所说明的附图用来提供对本申请的进一步理解,构成本申请的一部分,本申请的示意性实施例及其说明用于解释本申请,并不构成对本申请的不当限定。在附图中:

[0033] 图1为本说明书实施例提供的基于大数据批流一体的铁路线路能力利用率计算方法的过程示意图;

[0034] 图2为本说明书实施例中一种电子设备的结构示意图。

具体实施方式

[0035] 下面通过具体实施方式结合附图对本发明作进一步详细说明。其中不同实施方式中类似元件采用了相关联的类似的元件标号。在以下的实施方式中,很多细节描述是为了使得本申请能被更好的理解。然而,本领域技术人员可以毫不费力的认识到,其中部分特征在不同情况下是可以省略的,或者可以由其它元件、材料、方法所替代。在某些情况下,本申请相关的一些操作并没有在说明书中显示或者描述,这是为了避免本申请的核心部分被过多的描述所淹没,而对于本领域技术人员而言,详细描述这些相关操作并不是必要的,他们根据说明书中的描述以及本领域的一般技术知识即可完整了解相关操作。

[0036] 另外,说明书中所描述的特点、操作或者特征可以以任意适当的方式结合形成各种实施方式。同时,方法描述中的各步骤或者动作也可以按照本领域技术人员所能显而易见的方式进行顺序调换或调整。因此,说明书和附图中的各种顺序只是为了清楚描述某一个实施例,并不意味着是必须的顺序,除非另有说明其中某个顺序是必须遵循的。

[0037] 本文中为部件所编序号本身,例如“第一”、“第二”等,仅用于区分所描述的对象,不具有任何顺序或技术含义。而本申请所说“连接”、“联接”,如无特别说明,均包括直接和间接连接(联接)。

[0038] 以下结合附图,详细说明本申请各实施例提供的技术方案。

[0039] 批处理(Batch),也称为批处理脚本。批处理是将一定量的数据集合在一起,形成一个数据批次,然后对这个批次中的数据进行处理。Spark和Flink都支持批处理,其中Spark使用的是批处理模型,即将一批数据一次性读入内存,然后对其进行处理,处理完成后再将结果写入磁盘。Flink也支持批处理,但使用的是基于流处理的批处理模式,即将一批数据分成多个数据流进行处理,可以实现更高效的内存管理和更低的延迟。

[0040] 流处理是一种将数据流式地处理的方式,即数据一条一条地进入系统,然后被实时处理。Spark和Flink都支持流处理。在流处理中,Flink使用的是事件驱动模型,即对于每个输入事件进行实时处理。Spark则使用基于微批处理的模型,即将数据流划分成小批次进行处理。由于基于微批处理的模型需要缓存一定量的数据才能开始处理,因此在实时性上略逊于基于事件驱动模型。

[0041] 铁路运输环境狭义上说,不同的站点各自对其所属的运输环境进行较为分散的管理;广义上说,在更加宏观的维度上对不同的站点各自所属的运输环境进行较为同一的管理。这就使得在不同的层级之间存在一定的数据处理内容、以及数据处理方式的隔离。若仅仅是从较为宏观的层面着手,对运输环境进行管理,则有可能忽视了站点所属的运输环境的个体情况,导致该站点出现管理不力。若仅仅从较为微观的层面着手,对运输环境进行管理,则有可能导致不同站点各自所属的运输环境之间的衔接出现问题,导致宏观层面上的管理不力。

[0042] 此外,在技术发展史上,批处理和流处理是针对不同的使用环境、不同的数据处理需求,分别地设计出的数据处理方式,这两种技术完全的融合的难度较大,即便是一定程度的实现了两者的融合,在数据处理效果上也未必能够满足实际需求,特别是在强调铁路运输这种运输环境复杂、数据层级性明显、数据存在交叉、数据变数较大的应用场景时,如何使得批处理和流处理两者的优点均能体现在数据处理效果上成为亟待解决的问题。

[0043] 在本说明书一个可选的实施例中,本说明书中的运输环境是客货混用运输环境。

[0044] 本说明书中的方法由基于大数据批流一体的铁路线路能力利用率计算执行。如图1所示,本说明书中的基于大数据批流一体的铁路线路能力利用率计算方法包含以下步骤:

[0045] S100:从预设的数据分类管理模型获取距当前时刻指定时长的历史时间段内运输环境的环境数据。

[0046] 本说明书中的方法旨在:针对运输环境进行分析,进而基于分析得到的结果,指导站点对其所属的运输环境进行管理。

[0047] 本说明书中的运输环境包含运输线路和相邻的两条运输线路之间的站点。运输线路的端点是站点。针对某一个站点,以该站点为始点、或者以该站点为终点的各条运输线路和该站点共同构成的运输环境,即为该站点所属的运输环境。也就是说,本说明书中的运输环境是针对某一站点的定义,某一运输环境中包含的运输线路可能是一条,也可能是一条以上。针对不同的站点,其各自所属的运输环境之间可能存在一定的重叠,则需针对每个站点各自所属的运输环境分别地进行分析。

[0048] 本说明书的数据分类管理模型用于对环境数据进行分类存储。具体地分类方式为,针对每条线路,分别获取该线路对应的数据,并且记录线路和站点之间的对应关系。此外,还获取站点对应的数据。由此,在针对某一站点进行运输环境分析时,仅需查询该站点,即可得到该站点对应的数据,也可以查询到该站点所属的运输环境包含的各线路对应的数据。可见,本说明书中的数据分类管理模型不仅仅具备存储功能,还具备数据查询功能。相对应地,各站点、运输线路将其采集到的各自的数据上传至数据分类管理模型中。

[0049] 本说明书中的环境数据包含运输线路数据和站点数据。其中,运输线路数据包含:所述运输线路单向通车数量(即,在运输线路上行驶的列车的总数量)、所述运输线路双向通车对数(即,在指定时长中的某一个时刻,在所述运输线路的双向同时行驶有列车的对数)、所述运输线路客运密度(即,指定时长内,客运列车(本说明书中的列车不仅仅包含客运列车,还包含货运列车)的数量分布)、所述运输线路货运密度、所述运输线路相邻列车最长时间间隔、所述运输线路相邻列车最短时间间隔。站点数据包含:所述站点发车历史客座率(通常,即便是列车满座,站点也能够承受人流压力,此处客座率可用于表征由于未满座而产生的站点资源闲置情况)以及所述站点最大历史客流量(旅客乘坐列车的车次不同,不同乘客在站点滞留的时长不同,则不同时刻中站点的乘客数量不同)。

[0050] 可见,本说明书中的环境数据均为历史数据,但并非历史上生成的所有数据,而是距当前时刻指定时长的历史时间段内运输环境的环境数据。由此,一方面能够实现对历史数据的筛选,减小需要处理的数据的量;另一方面,能够考察距当前时刻较近的某一个历史时间段内对运输环境提供的运输资源的占用情况。铁路运输的资源的可利用性有所不同,例如,某一运输线路上行驶有列车1,即便是该列车1驶离了该运输线路上的A位置,但由于列车1处于A位置的下游,距A位置尚且较近,该A位置貌似已经处于空闲状态可供其他列车1行驶,但是为了保障列车1的安全,避免被追尾,该A位置仍然处于不可用状态。通过本说明书的方法,则能够察觉出这种较为隐蔽的不可用状态。此外,针对站点也是相似的道理,但站点的资源占用情况还有其自身的特点。

[0051] S102:基于所述环境数据,确定决策系数。

[0052] 本说明书中的决策系数用于表征:在指定时长内的历史的运输环境对资源的占用造成的资源可用状态滞后性,导致的当前时刻运输环境提供的资源可用性的紧张程度。紧

张程度越高,则表明即便是资源处于空闲状态,该资源不可用的程度越高。

[0053] 具体地,决策系数与所述运输线路双向通车对数和所述运输线路单向通车数量的比值负相关(该比值越高,表明单向的运输线路的占用程度越低。而对于站点,车辆有来有往,能够及时地将滞留旅客带走,有利于释放站点资源)、与所述运输线路货运密度和所述运输线路客运密度的比值负相关(货运的可调度性相较于客运较高。例如,若货运列车不能及时地取得线路的使用权,则可以让货运列车在旁路的轨道上暂停,等待线路资源释放。但是客运必须满足时效性,旅客滞留危害较大。而且,客运对站点的管理压力较大,货运对站点的压力则较小)、与所述运输线路相邻列车最长间隔和所述运输线路单向通车数量的最大值的乘积正相关(表明无论在时间上,还是在地理上,列车在运输环境中的分布不均匀性较高)、与所述运输线路相邻列车最短间隔正相关。

[0054] S104:若所述决策系数不大于预设的决策阈值,则判定采用批式处理模式,基于所述环境数据,确定所述第一运输线路能力利用率。

[0055] 批式处理模式即为采用批处理的方式对环境数据进行处理。决策系数不大于预设的决策阈值,表明由于历史上对运输环境的使用导致的资源占用延续至当前时刻的程度较轻,当前的运输环境不是十分紧张。批处理仅针对历史数据进行处理,一方面不用获取实时数据,也就无需处理实时数据;另一方面,批处理对大数据的处理效率较高,有利于提高数据处理效率。

[0056] 在本说明书一个可选的实施例中,决策阈值是经验值。在本说明书另一个可选的实施例中,所述决策阈值与所述环境数据表示出的所述运输线路上行驶的列车平均轴重负相关。列车的轴重越高,则表明在处于紧急制动的情况下出现高风险的事件的概率越大,为了安全考虑,应降低决策阈值。此外,在其他可选的实施例中,所述决策阈值与参照值负相关,所述参照值与所述运输线路上行驶的列车的最大行驶速度和处于最大行驶速度所持续的时长的乘积正相关,且与所述列车处于所述最大行驶速度时,与所述列车处于同向行驶的其他列车的数量与所述运输线路的长度的比值正相关。

[0057] 本说明书中的第一运输线路能力利用率用于表征所述运输环境相较于满负荷状态的使用程度。此处的“满负荷状态”是在历史上的运输情况对运输环境提供的资源的占用的基础之上的“满负荷”。例如,运输环境提供的资源理论最大值是100,由历史上的运输情况对当前的资源的占用是20,则此时的“满负荷状态”是80。

[0058] 其中,第一运输线路能力利用率与所述决策系数正相关,且与所述运输环境的综合相对占用率负相关。所述综合相对占用率与所述站点最大历史客流量正相关,且与最近一次处于混合式处理模式持续的时长正相关。

[0059] 在确定出第一运输线路能力利用率之前,被利用到的数据主要是运输线路数据,也就是说,确定决策系数的过程主要考虑了运输线路的情况。可以理解的是,运输线路大多数是铺设在较为荒凉的地区,一方面,这样的地区能够提供足够的场地用于线路铺设,另一方面,也能够避免列车的行驶对市区的居民造成干扰。这就使得,若在运输线路上出现故障、事故、自然灾害等问题,则解决该问题的难度较大。而站点距市区较近,除铁路以外的运输方式也较为便捷,对这种问题的处理难度会较低。故此,优先考虑运输线路的情况。以运输线路的情况进行决策系数的确定,能够有效地区分出运输线路和节点之间的状况的差异,实现对问题分级的识别。

[0060] 但是,对站点的情况也不容忽视,在定位出决策系数较低、运输线路提供的资源的占用延迟作用较低、由运输线路导致的问题的可能性较低的情况下,在基于决策系数计算第一运输线路能力利用率时,引入了综合相对占用率。综合相对占用率是基于站点最大历史客流量得到的(由于货运对站点的资源占用较少,货运有自己单独的装卸空间,则客流量是保障运输安全的较为主要的决定因素之一),能够体现出站点的资源占用延迟程度。此外,综合相对占用率还与最近一次处于混合式处理模式持续的时长正相关。本说明书中的方法启动混合式处理模式,则表明运输环境除存在资源占用延迟以外,还存在资源占用,是资源更加紧张的情形。处于混合式处理模式持续的时长越长,则表明在未来时刻中,由被占用的资源导致的资源占用延迟的问题还会延续,则应关注实施情况。综合相对占用率则能够在批处理模式中一定程度的实现流处理所带来的有益效果。

[0061] 此后,将所述第一运输线路能力利用率,发送至所述站点,使得所述站点基于所述第一运输线路能力利用率对未来时刻中的、所述站点所属的所述运输环境进行管理。

[0062] 由于,铁路运输环境中,资源释放具有一定的滞后性,则即便是基于历史数据得到的第一运输线路能力利用率,也不仅仅能够表征历史的状况,还能够表征当前时刻的实时状况以及未来时刻的预测状况。

[0063] 第一运输线路能力利用率表达了站点和站点所属运输环境提供的运输资源的占用情况,第一运输线路能力利用率越高,表明处于可用状态的资源、以及短时间内可释放的资源越少。站点获得第一运输线路能力利用率之后,可以结合其自身的调控能力,对站点或者运输线路的资源进行重新分配,以实现管理。具体地管理方式,可以由站点自行控制。例如,增加站点执勤人员、对站前广场进行布置,以将站前广场作为候车室的一部分、控制部分列车在某些运输线路上的某些路段的行驶速度降低,以避免出现高风险事件,等。

[0064] S106:若所述决策系数大于所述决策阈值,则判定处于所述混合式处理模式,将最近一次处于所述批式处理模式时得到的所述第一运输线路能力利用率作为历史利用率;基于所述历史利用率和所述环境数据,确定第二运输线路能力利用率。

[0065] 由于运输环境资源可用度较高的情况下,是采用批处理模式进行数据处理的,则调整至混合式处理模式意味着数据处理模式的切换,该切换引起的影响、以及由历史利用率对当前的运输环境和未来时刻的运输环境造成的资源占用滞后也应当予以考虑。

[0066] 本说明书中的第二运输线路能力利用率与所述历史利用率正相关、与所述站点发车实时客座率和所述站点发车历史客座率的比值正相关、且与所述站点发车实时客座率和所述站点实时客流量的比值正相关。

[0067] 历史利用率是基于策略系数得到的,策略系数着重考虑了运输线路的情况,则第二运输线路能力利用率能够体现出运输线路的情况。此外,第二运输线路能力利用率的确定过程还着重考虑了站点数据,在运输资源释放能力不足以环节运输压力时,由于旅客的流动性较大,问题在站点最先爆发的可能性较大,座椅应当着重考虑站点的情况。

[0068] 需要说明的是,本说明书中的混合式处理模式本质上是结合了历史利用率的流式处理,具体地,基于所述站点发车实时客座率和所述站点发车历史客座率的比值、所述站点发车实时客座率和所述站点实时客流量的比值得到待定利用率。在站点实时客流量不大于预设的流量阈值(经验值)时,采用历史利用率对待定利用率进行加权,得到第二运输线路能力利用率。在站点实时客流量大于预设的流量阈值时,直接将待定利用率作为第二运输

线路能力利用率。

[0069] 本申请提供一种基于大数据批流一体的铁路线路能力利用率计算方法,以及执行该方法的系统,结合了批处理和流处理的优点,实现基于运输线路能力利用率的运输环境调整。本说明书中的方法及系统,针对铁路运输环境的多层级性、复杂性和多变性,没有仅仅采用单独的某一种数据处理方式,也没有通过复杂的接口设计将不同的数据处理方式进行嫁接,而是结合铁路运输环境的实际情况,在适于进行批处理时,采用批处理模式进行数据处理。在不适于进行批处理时,采用了一种全新的数据处理模式,即混合式处理模式,以历史上进行批处理时得到的运输线路能力利用率(即,第一运输线路能力利用率)作为考察的一个方面,使得历史上的批处理的结果能够一定程度的体现在针对当前的实时运输环境的评价上,无需复杂的接口设计,不必考虑批处理和流处理之间的不相容性,有利于提高运输环境的管理效率。

[0070] 在说明书一个可选的实施例中,若监测到所述运输线路上行驶的列车的数量大于预设的数量阈值(经验值),则从所述数据分类管理模型获取所述环境数据。

[0071] 由前述分析可知,在运输线路上发生的灾害往往是抢救难度较大的,为能够机敏地判断线路的状况,则从运输线路着手,判断是否采取本说明书中的方法。

[0072] 在本说明书进一步可选的实施例中,所述指定时长与最近一次处于所述混合式处理模式持续的时长正相关。混合式处理模式是资源释放压力较大的情况下采取的措施,处于混合式处理模式的持续时长越长,则表明这种压力长期存在,则为后续运输环境资源的释放带来了更大的压力,此时应当增加指定时长,避免这种较为严重的延时被忽略。

[0073] 进一步地,本说明书还提供基于大数据批流一体的铁路线路能力利用率计算系统,所述系统包括:

[0074] 数据获取模块,配置为:从预设的数据分类管理模型获取距当前时刻指定时长的历史时间段内运输环境的环境数据;其中,所述运输环境包含:运输线路和相邻的两条所述运输线路之间的站点;所述环境数据包含:所述运输线路单向通车数量、所述运输线路双向通车对数、所述运输线路客运密度、所述运输线路货运密度、所述运输线路相邻列车最长间隔、所述运输线路相邻列车最短时间间隔、所述站点发车历史客座率以及所述站点最大历史客流量;

[0075] 决策系数确定模块,配置为:基于所述环境数据,确定决策系数;其中,所述决策系数与所述运输线路双向通车对数和所述运输线路单向通车数量的比值负相关、与所述运输线路货运密度和所述运输线路客运密度的比值负相关、与所述运输线路相邻列车最长间隔和所述运输线路单向通车数量的最大值的乘积正相关、与所述运输线路相邻列车最短时间间隔正相关;

[0076] 第一运输线路能力利用率确定模块,配置为:若所述决策系数不大于预设的决策阈值,则判定采用批式处理模式基于所述环境数据,确定第一运输线路能力利用率;其中,所述第一运输线路能力利用率用于表征所述运输环境相较于满负荷状态的使用程度;所述第一运输线路能力利用率与所述决策系数正相关,且与所述运输环境的综合相对占用率负相关;所述综合相对占用率与所述站点最大历史客流量正相关,且与最近一次处于混合式处理模式持续的时长正相关;将所述第一运输线路能力利用率,发送至所述站点,使得所述站点基于所述第一运输线路能力利用率对未来时刻中的、所述站点所属的所述运输环境进

行管理；

[0077] 第二运输线路能力利用率确定模块,配置为:若所述决策系数大于所述决策阈值,则判定处于所述混合式处理模式;将最近一次处于所述批式处理模式时得到的所述第一运输线路能力利用率作为历史利用率;基于所述历史利用率和所述环境数据,确定第二运输线路能力利用率;其中,所述第二运输线路能力利用率与所述历史利用率正相关、与所述站点的发车实时客座率和所述站点发车历史客座率的比值正相关、且与所述发车实时客座率和所述站点的实时客流量的比值正相关;将所述第二运输线路能力利用率,发送至所述站点,使得所述站点基于所述第二运输线路能力利用率对未来时刻中的、所述站点所属的所述运输环境进行管理。

[0078] 该系统能够执行前述任一实施例中的方法,并能够获得相同或相似的技术效果,此处不再赘述。

[0079] 图2是本申请的一个实施例电子设备的结构示意图。请参考图2,在硬件层面,该电子设备包括处理器,可选地还包括内部总线、网络接口、存储器。其中,存储器可能包含内存,例如高速随机存取存储器(Random-Access Memory,RAM),也可能还包括非易失性存储器(non-volatile memory),例如至少1个磁盘存储器等。当然,该电子设备还可能包括其他业务所需要的硬件。

[0080] 处理器、网络接口和存储器可以通过内部总线相互连接,该内部总线可以是ISA(Industry Standard Architecture,工业标准体系结构)总线、PCI(Peripheral Component Interconnect,外设部件互连标准)总线或EISA(Extended Industry Standard Architecture,扩展工业标准结构)总线等。所述总线可以分为地址总线、数据总线、控制总线等。为便于表示,图2中仅用一个双向箭头表示,但并不表示仅有一根总线或一种类型的总线。

[0081] 存储器,用于存放程序。具体地,程序可以包括程序代码,所述程序代码包括计算机操作指令。存储器可以包括内存和非易失性存储器,并向处理器提供指令和数据。

[0082] 处理器从非易失性存储器中读取对应的计算机程序到内存中然后运行,在逻辑层面上形成基于大数据批流一体的铁路线路能力利用率计算装置。处理器,执行存储器所存放的程序,并具体用于执行前述任意基于大数据批流一体的铁路线路能力利用率计算方法。

[0083] 上述如本申请图1所示实施例揭示的基于大数据批流一体的铁路线路能力利用率计算方法可以应用于处理器中,或者由处理器实现。处理器可能是一种集成电路芯片,具有信号的处理能力。在实现过程中,上述方法的各步骤可以通过处理器中的硬件的集成逻辑电路或者软件形式的指令完成。上述的处理器可以是通用处理器,包括中央处理器(Central Processing Unit,CPU)、网络处理器(Network Processor,NP)等;还可以是数字信号处理器(Digital Signal Processor,DSP)、专用集成电路(Application Specific Integrated Circuit,ASIC)、现场可编程门阵列(Field-Programmable Gate Array,FPGA)或者其他可编程逻辑器件、分立门或者晶体管逻辑器件、分立硬件组件。可以实现或者执行本申请实施例中的公开的各方法、步骤及逻辑框图。通用处理器可以是微处理器或者该处理器也可以是任何常规的处理器等。结合本申请实施例所公开的方法的步骤可以直接体现为硬件译码处理器执行完成,或者用译码处理器中的硬件及软件模块组合执行完

成。软件模块可以位于随机存储器,闪存、只读存储器,可编程只读存储器或者电可擦写可编程存储器、寄存器等本领域成熟的存储介质中。该存储介质位于存储器,处理器读取存储器中的信息,结合其硬件完成上述方法的步骤。

[0084] 该电子设备还可执行图1中基于大数据批流一体的铁路线路能力利用率计算方法,并实现图1所示实施例的功能,本申请实施例在此不再赘述。

[0085] 本申请实施例还提出了一种计算机可读存储介质,该计算机可读存储介质存储一个或多个程序,该一个或多个程序包括指令,该指令当被包括多个应用程序的电子设备执行时,执行前述的任意基于大数据批流一体的铁路线路能力利用率计算方法。

[0086] 本领域内的技术人员应明白,本申请的实施例可提供为方法、系统、或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0087] 本申请是参照根据本申请实施例的方法、设备(系统)、和计算机程序产品的流程图和/或方框图来描述的。应理解可由计算机程序指令实现流程图和/或方框图中的每一流程和/或方框、以及流程图和/或方框图中的流程和/或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器,使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的装置。

[0088] 这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中,使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品,该指令装置实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能。

[0089] 这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上,使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理,从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和/或方框图一个方框或多个方框中指定的功能的步骤。

[0090] 在一个典型的配置中,计算设备包括一个或多个处理器(CPU)、输入/输出接口、网络接口和内存。

[0091] 内存可能包括计算机可读介质中的非永久性存储器,随机存取存储器(RAM)和/或非易失性内存等形式,如只读存储器(ROM)或闪存(flash RAM)。内存是计算机可读介质的示例。

[0092] 计算机可读介质包括永久性和非永久性、可移动和非可移动媒体可以由任何方法或技术来实现信息存储。信息可以是计算机可读指令、数据结构、程序的模块或其他数据。计算机的存储介质的例子包括,但不限于相变内存(PRAM)、静态随机存取存储器(SRAM)、动态随机存取存储器(DRAM)、其他类型的随机存取存储器(RAM)、只读存储器(ROM)、电可擦除可编程只读存储器(EEPROM)、快闪记忆体或其他内存技术、只读光盘只读存储器(CD-ROM)、数字多功能光盘(DVD)或其他光学存储、磁盒式磁带,磁带磁磁盘存储或其他磁性存储设备或任何其他非传输介质,可用于存储可以被计算设备访问的信息。按照本文

中的界定,计算机可读介质不包括暂存电脑可读媒体 (transitory media),如调制的数据信号和载波。

[0093] 还需要说明的是,术语“包括”、“包含”或者其任何其他变体意在涵盖非排他性的包含,从而使得包括一系列要素的过程、方法、商品或者设备不仅包括那些要素,而且还包括没有明确列出的其他要素,或者是还包括为这种过程、方法、商品或者设备所固有的要素。在没有更多限制的情况下,由语句“包括一个……”限定的要素,并不排除在包括所述要素的过程、方法、商品或者设备中还存在另外的相同要素。

[0094] 本领域技术人员应明白,本申请的实施例可提供为方法、系统或计算机程序产品。因此,本申请可采用完全硬件实施例、完全软件实施例或结合软件和硬件方面的实施例的形式。而且,本申请可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质(包括但不限于磁盘存储器、CD-ROM、光学存储器等)上实施的计算机程序产品的形式。

[0095] 以上所述仅为本申请的实施例而已,并不用于限制本申请。对于本领域技术人员来说,本申请可以有各种更改和变化。凡在本申请的精神和原理之内所作的任何修改、等同替换、改进等,均应包含在本申请的权利要求范围之内。

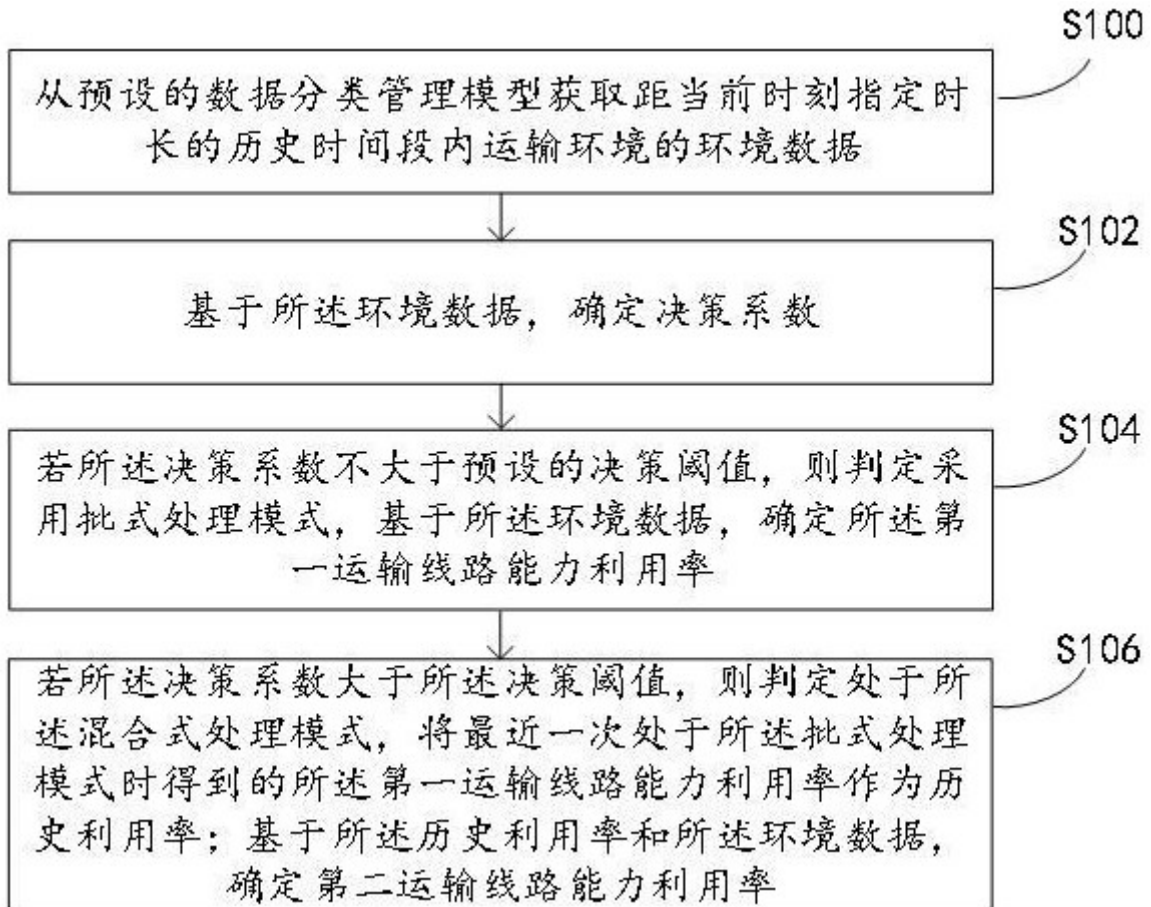


图 1

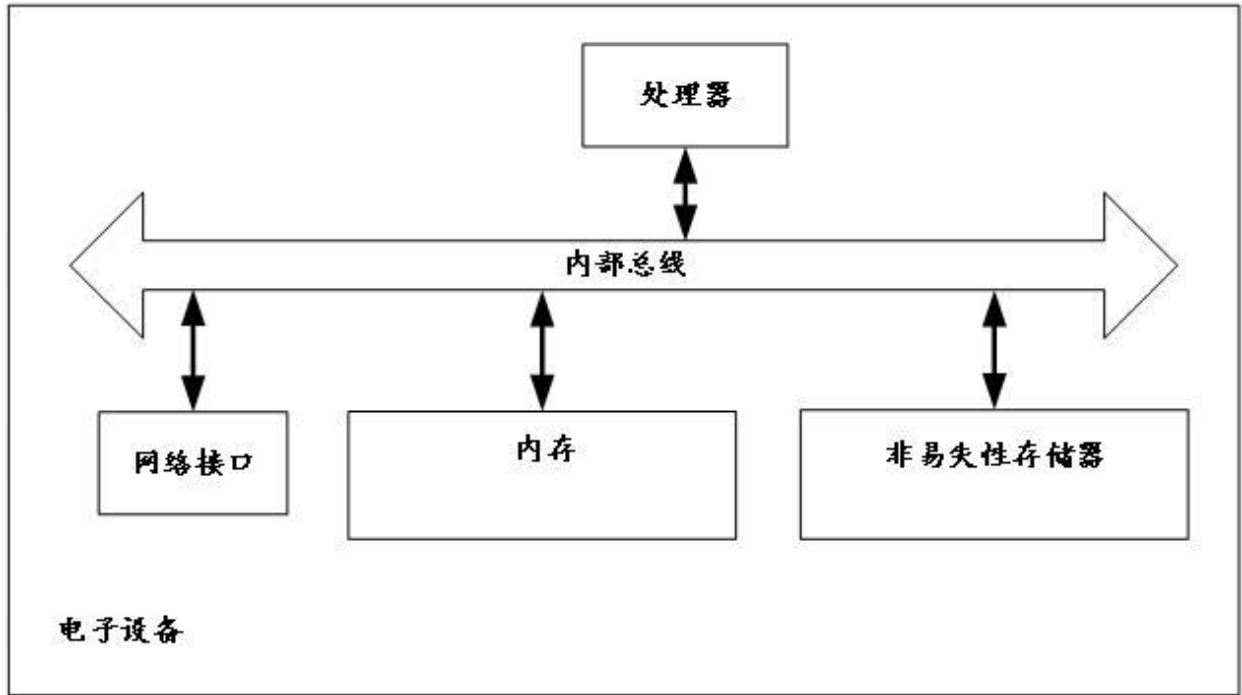


图 2