(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2021 年 9 月 30 日 (30.09.2021)

WIPO | PCT

(10) 国际公布号
WO 2021/189362 A1

(51) 国际专利分类号:
G06F 16/215 (2019.01)

(21) 国际申请号: PCT/CN2020/081440

(22) 国际申请日: 2020 年 3 月 26 日 (26.03.2020)

(25) 申请语言: 中文

(26) 公布语言: 中文

(71) 申请人: 深圳先进技术研究院 (**SHENZHEN INSTITUTES OF ADVANCED TECHNOLOGY**) [CN/CN]; 中国广东省深圳市南山区西丽大学城学苑大道1068号, Guangdong 518055 (CN)。

(72) 发明人: 彭磊 (**PENG, Lei**); 中国广东省深圳市南山区西丽大学城学苑大道1068号, Guangdong 518055 (CN)。 张俊楠 (**ZHANG, Junnan**); 中国广东省深圳市南山区西丽大学城学苑大道1068号, Guangdong 518055 (CN)。 李慧云 (**LI, Huiyun**); 中国广东省深圳市南山区西丽大学城学苑大道1068号, Guangdong 518055 (CN)。

(74) 代理人: 广州三环专利商标代理有限公司 (**SCIHEAD IP LAW FIRM**); 中国广东省广州市越秀区先烈中路80号汇华商贸大厦1508室, Guangdong 510070 (CN)。

(81) 指定国(除另有指明，要求每一种可提供的国家保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) **Title:** TIME SERIES DATA GENERATION METHOD AND DEVICE BASED ON MULTI-CONDITION CONSTRAINTS, AND MEDIUM

(54) 发明名称: 基于多条件约束的时间序列数据生成方法、装置及介质

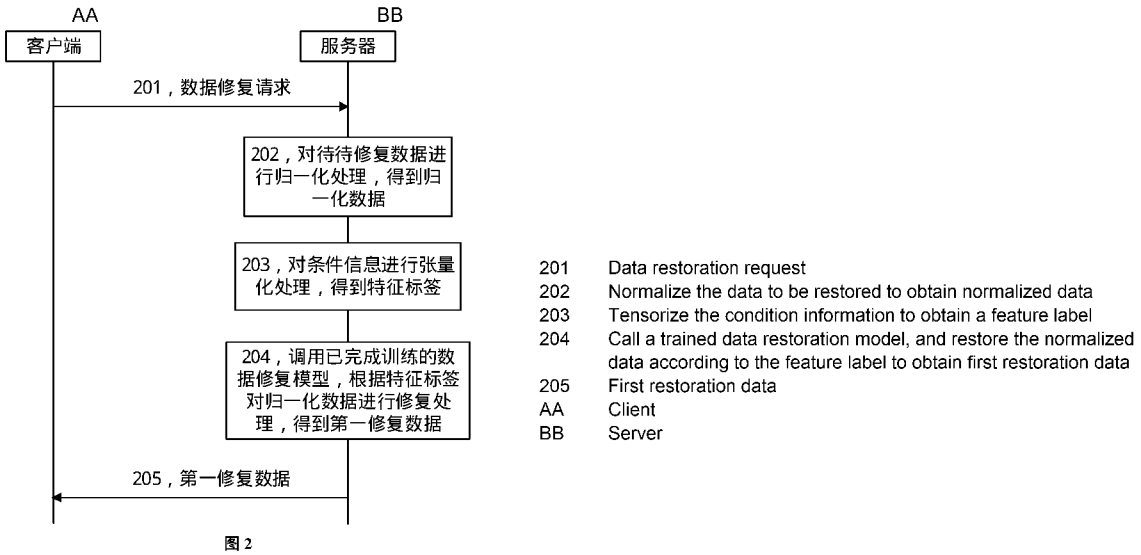


图 2

(57) **Abstract:** A time series data generation method and device based on multi-condition constraints, and a medium. The method comprises: a data restoration request (201), the data restoration request comprising data to be restored and condition information; normalize the data to be restored to obtain normalized data (202); tensorize the condition information to obtain a feature label (203); call a trained data restoration model, and restore the normalized data according to the feature label to obtain first restoration data (204), the data restoration model being obtained by training according to sample data, a first sample condition, real sample data, and a second sample condition, and the sample data being noise data; and a server sends the first restoration data to a client (205). By using the method, abundant features of the data can be obtained without taking a large amount of historical data or sample data having high matching degree as a training basis, the accuracy and timing of restored data are ensured, and the restoration efficiency and quality are improved.

WO 2021/189362 A1

[见续页]

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW。

(84) 指定国(除另有指明，要求每一种可提供的地区保护)：ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(57) 摘要：一种基于多条件约束的时间序列数据生成方法、装置及介质，其中，所述方法包括：数据修复请求（201）；数据修复请求包括待修复数据及条件信息；对待修复数据进行归一化处理，得到归一化数据（202）；并对条件信息进行张量化处理，得到特征标签（203）；调用已完成训练的数据修复模型，根据特征标签对归一化数据进行修复处理，得到第一修复数据（204），数据修复模型是根据样本数据、第一样本条件、真实样本数据及第二样本条件进行训练得到的，样本数据为噪声数据；服务器发送第一修复数据至客户端（205）。采用所述方法，无需大量的历史数据或匹配度高的样本数据作为训练基础，便可获取数据的丰富特征，保证了修复数据的准确性及时序性，提高了修复效率及质量。

基于多条件约束的时间序列数据生成方法、装置及介质

技术领域

本申请涉及大数据技术领域，尤其涉及一种基于多条件约束的时间序列数据生成方法、装置及介质。

背景技术

时间序列数据是指在不同时间收集到的数据，此类数据在生产和生活中用于描述某一事物或现象等随时间的变化情况。但是，由于此类数据的数据点较为密集且抗干扰性差，在数据的采集、应用或传递等过程中，容易造成数据的缺失。目前，针对缺失数据的修复方法主要包括如下两种：第一种是基于先验知识的插值方法；第二种是获取与已缺失数据最匹配的样本数据，利用该样本数据训练生成对抗网络，以修复已缺失数据。但是第一种方法需要有大量的历史数据作为基础，不适用于海量数据的修复；第二种方法较难获取匹配度高的样本数据且难以学习到数据的有效特征，修复的数据准确性差且不具有时序性。

发明内容

本申请实施例提供一种基于多条件约束的时间序列数据生成方法、装置及介质，无需大量的历史数据或匹配度高的样本数据作为训练基础，便可获取待修复数据的丰富特征，保证了修复数据的准确性及时序性，提高了修复效率及质量。

第一方面，本申请实施例提供一种基于多条件约束的时间序列数据生成方法，包括:

接收客户端的数据修复请求，所述数据修复请求包括待修复数据及条件信息，所述数据修复请求用于请求根据所述条件信息对所述待修复数据进行数据修复，所述条件信息为与所述待修复数据相匹配的特征条件；

对所述待修复数据进行归一化处理，得到所述待修复数据的归一化数据，并对所述条件信息进行张量化处理，得到所述条件信息的特征标签；

调用已完成训练的数据修复模型，根据所述特征标签对所述归一化数据进行修复处理，得到第一修复数据，所述数据修复模型是根据样本数据、第一样本条件、真实样本数据及第二样本条件对所述数据修复模型进行训练得到的，

所述样本数据为噪声数据；

发送所述第一修复数据至所述客户端。

在该技术方案中，客户端发送包括待处理数据及条件信息的数据修复请求至服务器，以使服务器对待修复数据进行归一化处理得到归一化数据，并对条件信息进行张量化处理，得到特征标签，该条件信息为与待修复数据相匹配的特征条件，调用已完成训练数据修复模型，根据特征标签对归一化数据进行修复处理，得到第一修复数据，并将该第一修复数据发送至客户端。通过这种方法，基于条件信息的输入，可获取待修复数据的丰富特征，以使生成的第一修复数据更接近真实数据的分布特征，保证了修复数据的准确性及时序性，提高了修复效率及质量。

第二方面，本申请实施例提供一种基于多条件约束的时间序列数据生成装置，包括：

收发单元，用于接收客户端的数据修复请求，所述数据修复请求包括待修复数据及条件信息，所述数据修复请求用于请求根据所述条件信息对所述待修复数据进行数据修复，所述条件信息为与所述待修复数据相匹配的特征条件；

处理单元，用于对所述待修复数据进行归一化处理，得到所述待修复数据的归一化数据，并对所述条件信息进行张量化处理，得到所述条件信息的特征标签；调用已完成训练的数据修复模型，根据所述特征标签对所述归一化数据进行修复处理，得到第一修复数据，所述数据修复模型是根据样本数据、第一样本条件、真实样本数据及第二样本条件对所述数据修复模型进行训练得到的，所述样本数据为噪声数据；

所述收发单元，还用于发送所述第一修复数据至所述客户端。

第三方面，本申请实施例提供一种基于多条件约束的时间序列数据生成装置，包括处理器、存储器和通信接口，所述处理器、所述存储器和所述通信接口相互连接，其中，所述存储器用于存储计算机程序，所述计算机程序包括程序指令，所述处理器被配置用于调用所述程序指令，执行如第一方面所描述的方法。该处理设备解决问题的实施方式以及有益效果可以参见上述第一方面所描述的方法以及有益效果，重复之处不再赘述。

第四方面，本申请实施例提供一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有一条或多条第一指令，所述一条或多条第一指令适于由处理器加载并执行如第一方面所描述的方法。

本申请实施例中，客户端发送数据修复请求至服务器，该数据修复请求包括待修复数据集条件信息，服务器对待修复数据进行归一化处理得到归一化数据，并对条件信息进行张量化处理得到特征标签，该条件信息为与待修复数据相匹配的特征条件，此处基于条件信息进行数据修复可以充分考虑待修复数据特征的多样性，以便得到更准确的修复数据。调用已完成训练数据修复模型，根据特征标签对归一化数据进行修复处理，得到第一修复数据，并将该第一修复数据发送至客户端。其中，数据修复模型的训练方法为：通过多条件约束的生成对抗网络对输入的至少一组样本数据、第一样本条件、真实样本数据及第二样本条件进行迭代监督训练，其中，样本数据为噪声数据，真实样本数据为真实的时间序列数据，通过本实施例的方法，可以对存在缺失的时间序列数据进行修复，无需大量的历史数据或手动获取与缺失数据匹配度高的样本数据作为训练基础，便可实现对模型的训练，并且，通过多个特征条件信息的引入，可以获取到待修复数据的丰富特征，采用长短神经记忆 (Long Short-Term Memory, LSTM)网络作为多条件约束的生成对抗网络中生成器及判别器的内置网络，保证了修复数据的准确性及时序性，提高了修复效率及质量。

附图说明

为了更清楚地说明本申请实施例中的技术方案，下面将对实施例中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本申请的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

图1是本申请实施例提供的一种基于多条件约束的时间序列数据生成系统的架构图；

图2是本申请实施例提供的一种基于多条件约束的时间序列数据生成方法的流程图；

图3是本申请实施例提供的另一种基于多条件约束的时间序列数据生成方法的流程图；

图4是本申请实施例提供的一种多条件生成对抗网络的框架图；

图5是本申请实施例提供的一种平均余弦相似度计算结果的示意图；

图6是本申请实施例提供的又一种基于多条件约束的时间序列数据生成方法的流程图；

图7(a)是本申请实施例提供的一种无条件约束生成数据与真实数据的对比示意图；

图7(b)是本申请实施例提供的一种单条件约束生成数据与真实数据的对比示意图；

图7(c)是本申请实施例提供的一种多条件约束生成数据与真实数据的对比示意图；

图8(a)是本申请实施例提供的一种无条件约束的残差分析结果示意图；

图8(b)是本申请实施例提供的一种单条件约束的残差分析结果示意图；

图8(c)是本申请实施例提供的一种多条件约束的残差分析结果示意图；

图9是本申请实施例提供的一种基于多条件约束的时间序列数据生成装置的结构示意图；

图10是本申请实施例提供的另一种基于多条件约束的时间序列数据生成装置的结构示意图。

具体实施方式

为了使本申请的目的、技术方案和优点更加清楚，下面将结合附图对本申请实施例的技术方案进行描述。显然，所描述的实施例是本申请一部分实施例，而不是全部的实施例。基于本申请中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本申请保护的范围。

本申请的说明书和权利要求书及上述附图中的术语“第一”、“第二”、“第三”等是用于区别不同对象，而非用于描述特定顺序。此外，术语“包括”以及它们任何变形，意图在于覆盖不排他的包含。例如包含了一系列步骤或模块的过程、方法、系统、产品或装置没有限定于已列出的步骤或模块，而是可选地还包括没有列出的步骤或模块，或可选地还包括对于这些过程、方法、产品或装置固有的其它步骤或模块。

时间序列数据是一类在不同时间收集到的具有时间信息的一维数据，例如：交通监控数据、停车场的停车情况数据，等等。该类数据在生产和生活中用于描述某一事物或现象等随时间的变化情况，但是，由于此类数据的数据点较为密集且抗干扰性差，在数据的采集、应用或传递等过程中，容易造成数据的缺失，从而对生产及生活带来巨大影响。目前，针对缺失数据的修复方法主要包括如下两种：第一种是基于先验知识的插值方法，该方法需要有大量的历史数

据作为基础，无法修复缺失较多数据，不适用于海量数据的修复；另一种方法需要获取与已缺失数据最匹配的样本数据，利用该样本数据训练生成对抗网络，得到生成数据，将生成数据填补至缺失数据中。其中，生成对抗网络可以包括：深度卷积生成对抗网络（Deep Convolution Generative Adversarial Networks，DCGAN）。但是该方法较难获取匹配度高的样本数据，且生成数据具有无序性，例如：在生成一周的停车场数据时，无法确定生成的数据是哪一天的数据。同时，利用生成对抗网络进行训练时需要提取样本的特征，由于，样本特征的多样化，使得不能从一个样本中学习到全部的特征，从而影响数据修复的准确性。

为解决上述问题，本申请实施例提供一种基于多条件约束的时间序列数据生成方法，该时间序列数据生成方法基于多条件生成对抗网络（Multi-condition Generative Adversarial Networks，MCGAN）进行海量数据的修复。该多条件生成对抗网络包括生成器网络及判别器网络，通过上述多条件生成对抗网络构建的数据修复模型，可以基于待修复数据及待修复数据对应的条件信息对待修复数据进行数据修复，得到第一修复数据。其中，条件信息为待修复数据的特征条件，例如：时间、空间及气候，等等。通过本实施方式，无需大量的历史数据或特定的样本数据作为基础，便可获取丰富的样本特征，以使数据修复模型可以生成更接近真实数据的修复数据，保证了海量修复数据的准确性及时序性，提高了数据修复质量。

可选的，本实施方式可以应用于停车场数据修复场景，具体的，可以获取有缺失的停车场数据作为待修复数据，并获取影响车位分布的条件信息，例如：时间、空间、气候等信息。将上述待修复数据经过归一化处理得到归一化数据，并将条件信息进行张量化处理得到特征标签。则可以调用基于多条件生成对抗网络的数据修复模型，根据特征标签对归一化数据进行数据修复，得到第一修复数据，该第一修复数据可以理解为符合停车场真实情况的数据。

上述提及的基于多条件约束的时间序列数据生成方法可应用于如图1所示的基于多条件约束的时间序列数据生成系统中，该多条件约束的时间序列数据生成系统可包括客户端101及服务器102。该客户端101的形态和数量用于举例，并不构成对本申请实施例的限定。例如，可以包括两个客户端101。

其中，客户端101可以为向服务器102发送数据修复请求的客户端，也可以为在数据修复模型训练时，用于为服务器102提供样本数据、第一样本条件、真

实样本数据及第二样本条件的客户端，该客户端可以为以下任一种：终端、独立的应用程序、应用程序编程接口（Application Programming Interface，API）或者软件开发工具包（Software Development Kit，SDK）。其中，终端可以包括但不限于：智能手机（如Android手机、IOS手机等）、平板电脑、便携式个人计算机、移动互联网设备（Mobile Internet Devices，MID）等设备，本申请实施例不做限定。服务器102可以包括但不限于集群服务器。

在本申请的实施例中，客户端101向服务器102发送数据修复请求，服务器102根据该数据修复请求所包括的待修复数据及条件信息，获取该待修复数据的第一修复数据。具体的，对待修复数据进行归一化处理，得到归一化数据，并对条件信息进行张量化处理，得到特征标签，通过预先训练好的数据修复模型结合特征标签对归一化数据进行修复处理，得到第一修复数据，将该第一修复数据发送至客户端101，以使客户端101的操作用户103可以根据该第一修复数据进行某一事物或现象等随时间的变化情况的分析。

请参见图2，图2是本申请实施例提供的一种基于多条件约束的时间序列数据生成方法的流程示意图，如图2所示，该时间序列数据生成方法可以包括201~205部分，其中：

201、客户端101发送数据修复请求至服务器102。

具体的，客户端101发送数据修复请求至服务器102，相应的，服务器102接收来自客户端101的数据修复请求，该数据修复请求包括待修复数据及条件信息，该数据修复请求用于请求根据条件信息对待修复数据进行数据修复，其中，待修复数据为已存在缺失情况的数据，具体的，可以为已存在缺失情况的时间序列数据。条件信息为与待修复数据相匹配的特征条件，例如：若待修复数据为停车场数据，则条件信息可以包括时间、空间及气候，等等。

202、服务器102对待修复数据进行归一化处理，得到待修复数据的归一化数据。

具体的，服务器102对待修复数据进行归一化处理，得到待修复数据的归一化数据。该待处理数据为时间序列数据，例如：为时间序列数据M，则该待处理数据可以表示为

$$M = \{m_{t_1}, m_{t_2} \dots m_{t_k} | k = 1,2 \dots l\}$$

其中m_{tk}表示t_k时刻对应的待修复数据的数据值，l为待修复数据的长度。

则归一化处理方法可以为，获取序列的最大值M_{max}=max（M），将上述待修复数据中各个数据与该最大值分别作商，即可得到归一化数据，该归一化数据可以表示为：

$$\{r_{t_1}, r_{t_2} \dots r_{t_l}\} = \frac{M}{M_{max}} = \left\{\frac{m_{t_1}}{M_{max}}, \frac{m_{t_2}}{M_{max}} \dots \frac{m_{t_l}}{M_{max}}\right\}$$

其中r_{t_l}表示t_l时刻对应的归一化数据的数据值。

可选的，在对待修复数据进行归一化处理之前，可以对待修复数据进行数据清洗。通过该可选的实施方式，有利于归一化数据在数据修复模型中更容易收敛。

203、服务器102对条件信息进行张量化处理，得到条件信息的特征标签。

具体的，服务器102对条件信息进行张量化处理，得到条件信息的特征标签。该条件信息为与上述待修复数据相匹配的特征条件，其中，条件信息可以包括但不限于：静态条件信息，例如：当待修复数据为停车场数据时，静态条件信息可以为停车场周围建筑分布；动态连续性条件信息，例如：时间序列标签；离散条件信息，例如：一周有7天，则7天是不同的离散特性，或天气、节假日等社会事件。则在输入至服务器102时。需要对获取的条件信息进行张量化处理，得到条件信息的特征标签。

进一步的，对于静态条件信息的张量化方法，以建筑物为例，不同的建筑物对应的静态条件信息不同，则需要获取不同分布情况的建筑物，即静态条件信息的多个影响因子$l_1 \dots l_n$，其中，n为影响因子的数量，则对静态条件信息的张量化处理可以采用归一化的方法，得到的特征标签可以表示为：

$$C = \left(\frac{|l_1|}{\sum_{i=1}^{n}|l_i|}, \frac{|l_2|}{\sum_{i=1}^{n}|l_i|} \dots \frac{|l_n|}{\sum_{i=1}^{n}|l_i|}\right)$$

进一步的，对于动态连续性条件信息的张量化方法，可以采用归一化处理方法，获取动态连续性条件信息的条件序列，该条件序列可以理解为采用时间顺序排列的条件标签，该条件序列可以表示为:

$$L = \{l_{t_1}, l_{t_2} \dots l_{t_k} | k = 1,2 \dots l\}$$

其中，l_{t_k}表示在时刻t_k时刻的条件标签值。

则归一化处理方法可以为，获取条件序列的最大值L_{max}=max（L），将上述待修复数据中各个数据与该最大值分别作商，即可得到特征标签，该特征标签可以表示为：

$$C = \{l_{t_1}, l_{t_2} \dots l_{t_l}\} = \frac{L}{L_{max}} = \left\{\frac{l_{t_1}}{L_{max}}, \frac{l_{t_2}}{L_{max}} \dots \frac{l_{t_l}}{L_{max}}\right\}$$

进一步的，对于离散条件信息的张量化处理，可以使用独热编码(one-hot)模式，该特征标签可以表示为：

$$C = onehot(n)$$

其中，n为事件出现的可能性数目，例如：若事件出现的可能性数目为2，则事件1和事件2的表示方法可以分别为{1,0}及{0,1}.

204、服务器102调用已完成训练的数据修复模型，根据特征标签对归一化数据进行修复处理，得到第一修复数据。

具体的，服务器102调用已完成训练的数据修复模型，根据特征标签对归一化数据进行修复处理，得到第一修复数据。其中，数据修复模型是根据样本数据、第一样本条件、真实样本数据及第二样本条件对数据修复模型进行训练得到的，该样本数据为噪声数据。该数据修复模型为生成器网络及判别器网络利用样本数据、第一样本条件、真实样本数据及第二样本条件进行反复迭代训练构建的模型。

进一步的，上述待修复数据包括时间点序列，则根据特征标签对归一化数据进行修复处理，得到第一修复数据，可以为根据时间点序列，对归一化数据中各个数据进行排序，根据特征标签对已完成排序的归一化数据进行数据修复处理，得到第一修复数据。其中，时间点序列为待修复数据中各个数据的生成时间点所组成的序列，即上述待处理数据中$\{t_1, t_2 \dots t_l\}$，则将待修复数据经由归一化处理得到得归一化数据也携带有该时间点序列，即上述归一化数据中的$\{t_1, t_2 \dots t_l\}$。在对归一化数据中的各个数据按照时间点完成排序的情况下，将已完成排序的归一化数据输入至数据修复模型的生成器网络中，该生成器的内置网络为长短神经记忆 (Long Short-Term Memory, LSTM)网络，采用LSTM网络，可以提升该数据修复模型对时间序列的处理能力。具体的，将已完成排序的归一化数据中的各个数据按照时间顺序输入至LSTM网络的各个对应细胞接口中，其中，LSTM网络的细胞接口数据与归一化数据的长度相同。并且将特征标签分别输入至每个细胞接口，以使生成器网络可以根据该特征标签对归一化数据进行数据修复，得到第一修复数据。

通过执行本实施方式，可以基于待修复数据的丰富特征，得到更准确的修复数据，采用LSTM网络可以保证生成数据的时序性，提高了数据修复的质量。

205、服务器102发送第一修复数据至客户端101。

具体的，服务器102将第一修复数据发送至客户端101。相应的，客户端101接收该第一修复数据，以使客户端101的操作用户103可以根据该第一修复数据进行某一事物或现象等随时间的变化情况的分析。该第一修复数据为完成修复的接近真实情况的数据

可见，通过实施图2所描述的方法，客户端101在发送数据修复请求后，服务器102对数据修复请求中的待修复数据进行归一化处理得到归一化数据，并对数据修复请求中的条件信息进行张量化处理，得到特征标签，该条件信息为与待修复数据相匹配的特征条件，此处基于条件信息进行数据修复可以充分考虑待修复数据特征的多样性，以便得到更准确的修复数据。调用已完成训练数据修复模型，根据特征标签对归一化数据进行修复处理，得到第一修复数据，并将该第一修复数据发送至客户端101。通过执行本实施例的方法，无需大量的历史数据或匹配度高的样本数据作为训练基础，便可获取待修复数据的丰富特征，保证了修复数据的准确性及时序性，提高了修复效率及质量。

请参见图3，图3是本申请实施例提供的一种基于多条件约束的时间序列数据生成方法的流程示意图，如图3所示，该时间序列数据生成方法可以包括301~308部分，其中：

301、服务器102获取样本数据及第一样本条件。

具体的，服务器102可以从客户端101或其他数据平台获取样本数据，及与该样本数据相匹配的第一样本条件。其中，第一样本条件的相关描述可以参见步骤201中条件信息的相关描述，此处不赘述。样本数据可以为在噪声空间采样得到的噪声样本序列数据，该样本数据可以表示为：

$$Z = \{z^{(1)}, z^{(2)}, z^{(3)} \ldots z^{(n)}\}$$

302、服务器102对样本数据进行归一化处理，得到第一处理数据。

具体的，服务器102在获取到样本数据的情况下，对样本数据进行归一化处理，得到样本数据的第一处理数据。此处的归一化处理的方法可以参见步骤202中对待修复数据的归一化处理的相关描述，此处不赘述。

303、服务器102对第一样本条件进行张量化处理，得到第一样本标签。

具体的，服务器102对第一样本条件进行张量化处理，得到第一样本标签。此处的张量处理的方法可以参见步骤203中对条件信息的张量化处理的相关描

述，此处不赘述。

进一步的，在获取到第一处理数据及第一样本标签后，可以根据第一处理数据及与该第一处理数据相对应的第一样本标签，构建样本监督数据集，该样本监督数据集用于输入至MCGAN中进行网络训练，构建数据修复模型。其中，第一样本标签Y_p可以表示为：

$$Y_p = \{y_{p,t_1}, y_{p,t_2} \dots y_{p,t_k} | k = 1,2 \dots l; p = 1,2 \dots n\}$$

其中n为第一样本标签的个数，则上述监督数据集可以表示为：

$$S = \{(r_{t_k}, y_{1,t_k}, y_{2,t_k} y_{3,t_k} \dots y_{n,t_k})\}_{k=1,2\dots l}$$

304、服务器102获取真实样本数据及第二样本条件。

具体的，服务器102可以从客户端101或其他数据平台获取真实样本数据，及与真实样本数据相匹配的第二样本条件。其中，第二样本条件的相关描述可以参见步骤201中条件信息的相关描述，此处不赘述，可选的，第二样本条件可以为与第一样本条件相同的特征条件。真实样本数据为存在缺失的真实数据，例如：存在缺失的真实停车场数据，该真实样本数据X可以表示为：

$$X = \{x^{(1)}, x^{(2)}, x^{(3)} \dots x^{(n)}\}$$

其中，n为真实样本数据的个数。

305、服务器102对真实样本数据进行归一化处理，得到第二处理数据。

具体的，服务器102对真实样本数据进行归一化处理，得到真实样本数据的第二处理数据。此处的归一化处理的方法可以参见步骤202中对待修复数据的归一化处理的相关描述，此处不赘述。

306、服务器102对第二样本条件进行张量化处理，得到第二样本标签。

具体的，服务器102对第一样本条件进行张量化处理，得到第一样本标签。此处的张量处理的方法可以参见步骤203中对条件信息的张量化处理的相关描述，此处不赘述。

进一步的，可以根据第二处理数据及与该第二处理数据相对应的第二样本标签，构建真实监督数据集，该真实监督数据集构成方法的相关描述可以步骤303中样本监督数据集的构成，此处不赘述。

307、服务器102根据第一处理数据、第一样本标签、第二处理数据及第二样本标签对数据修复模型进行监督训练，确定模型函数。

具体的，服务器102根据第一样本数据、第一样本标签、第二处理数据及第

二样本标签对数据修复模型进行监督训练，确定模型函数，以使可以进一步根据该模型的函数，优化网络参数，构建数据修复模型，即执行步骤308。

进一步的，训练该数据修复模型所涉及的网络为MCGAN网络，主要包括生成器及判别器，该网络的框架图可参见图4所示，利用该MCGAN网络进行监督训练的过程为：从噪声空间采样得到样本数据，对样本数据Z进行归一化处理得到第一处理数据，并获取与样本数据Z相匹配的第一样本条件，将第一样本条件进行张量化处理得到第一样本标签。其中，第一样本标签可以包括多个张量化特征条件C ,将第一处理数据输入至生成器的内置LSTM网络的每个数据细胞接口中，并将第一样本标签的各个张量化特征条件C通过条件通道输入至每个细胞接口中，每个条件通道可以传输一个张量化特征条件C，则上述生成器的输入数据可以表示为$\{Z, C_1, C_2 \dots C_n\}$。进一步可得到第二修复数据，此处关于LSTM网络的相关描述可以参见步骤204中对应的描述，此处不赘述。在得到第二修复数据的情况下，可以将第二修复数据F及对应的第一样本标签输入至判别器中进行判别处理，可以得到第一判别结果；并且将完成归一化处理的真实样本数据R及与该真实样本数据R对应的经过张量化处理的第二样本标签输入至判别器中，该第二样本标签包括多个张量化特征条件C，此时判别器的输入数据可以表示为$\{(F\ or\ R), C_1, C_2 \dots C_n\}$，输出第二判别结果。其中，判别器的内置网络为LSTM网络，该判别器与生成器相同，均包含有条件通道及数据通道。可选的，在生成器及判别器中均可以配置有状态转换向量，该状态转换向量可以控制上述条件通道的开通或关闭，从而调整训练所需的特征条件。

进一步的，上述第一样本条件包括n个特征条件，n为正整数，则在获取样本数据及第一样本条件之前，客户端101还可以发送条件指令至服务器102，该条件指令用于指示从n个特征条件中获取x个特征条件，x为小于或等于n的非负整数。则服务器102可以根据条件指令从包括n个特征条件的第一样本条件信息中获取x个特征条件。具体的，可以包括如下三种情况：当客户端101不发送该条件指令时，则生成器及判别器的条件通道为全部开启状态；若接收到客户端101发送的指令，当条件指令指示关闭部分条件通道时，则生成器及判别器关闭指定的条件通道；当条件指令指示关闭全部，则生成器及判别器根据关闭全部的条件通道，此时生成修复数据的过程不需要考虑特征条件。可选的，该条件指令可以包括状态转换向量，该状态转换向量可以嵌入到条件通道中，以便控制各个条件通道的开关。则在增加状态转换向量后的生成器LSTM网络的输入

数据G'可以表示为：

$$G' = \{Z, S_1 * C_1, S_2 * C_2 \dots S_n * C_n\}$$

其中，S=1代表通路，即条件通道的状态为开启；S=0代表闭路，即条件通道的状态为关闭。进一步的，在增加了状态转换向量后的判别器LSTM网络的输入数据D'可以表示为：

$$D' = \{(F\ or\ R), S_1 * C_1, S_2 * C_2 \dots S_n * C_n\}$$

通过该实施方式，可以提高网络的应用范围，对于不同的特征条件情况，该网络结构可以作出适应性的调整 ,以获取不同特征条件下 ,生成的修复数据。

进一步的，模型函数可以包括生成损失函数、判别损失函数及目标函数，则确定模型函数的过程可以为：根据第一样本标签对第一处理数据进行修复处理，得到第二修复数据，并对第二修复数据及第一样本标签进行判别处理，得到第一判别结果；对第二处理数据及第二样本标签进行判别处理，得到第二判别结果；根据第一判别结果及第二判别结果，确定判别损失函数；根据第一判别结果，确定生成损失函数；对判别损失函数及生成损失函数进行优化，确定目标函数。

具体的，生成器根据第一样本标签Y_p对第一处理数据进行修复处理，得到携带有第一样本标签Y_p的第二修复数据$G(Z|Y_p)$，其中，第一处理数据为完成归一化处理的样本数据Z。在得到携带有第一样本标签Y_p的第二修复数据$G(Z|Y_p)$的情况下，将携带有第一样本标签Y_p的第二修复数据$G(Z|Y_p)$输入至判别器中，以使判别器对第二修复数据进行判别，得到第一判别结果$D(G(Z|Y_p))$。判别器一方面需要判断生成的数据是否满足真实的样本分布，另一方面，也需要判断生成的数据是否满足相应的特征条件。若判别结果为是，则说明生成的第二修复数据为满足真实样本特征的数据；若判断结果为否，则需要网络参数，继续迭代训练生成满足真实样本特征的修复数据，使判别器的输出尽可能全为真。其中判别器的诊断网络J可以表示为：

$$J = J_{real-sample-distribustion}(D')\&J_{condition-1}(D')\&J_{condition-1}(D')\& \dots$$
$$\dots J_{condition-n}(D')$$

其中，$J_{real-sample-distribustion}(D')$表示对生成的数据是否满足真实的样本分布的判断结果，$J_{condition-n}(D')$表示对生成的数据是否满足相应的特征条件的判断结果，D'表示判别器输出的判别结果，$D' = \{d_1, d_2 \dots d_n\}$。

基于上述判别器的诊断网络，进一步的，可以对第二处理数据及第二样本

标签Y_p进行判别处理，得到第二判别结果$D(X|Y_p)$，其中，第二处理数据为完成归一化处理的真实样本数据X。在获取到上述第一判别结果及第二判别结果的情况下，可以根据第一判别结果及第二判别结果，确定判别损失函数，该判别损失函数为判别器的损失函数，该判别器的损失函数可以表示为：

$$\mathrm{Loss} = -\left(\log \mathrm{D}\left(\mathrm{x}|_{\mathrm{y}}^{\mathrm{r}}\right) + \log\left(1 - \mathrm{D}\left(\mathrm{G}\left(\mathrm{z}|_{\mathrm{y}}^{\mathrm{r}}\right)\right)\right)\right)$$

可以根据第一判别结果，确定生成损失函数，该生成损失函数为生成器的损失函数，该生成器的损失函数可以表示为：

$$\mathrm{Loss} = \log\left(1 - \mathrm{D}\left(\mathrm{G}\left(\mathrm{z}|_{\mathrm{y}}^{\mathrm{r}}\right)\right)\right)$$

则在获取上述判别损失函数及生成损失函数的情况下，对判别损失函数及生成损失函数进行优化，确定目标函数。其中，对于判别器的优化目标是通过优化使得目标函数可以取得最大值。本实施方式中判别器的损失函数为负数形式，则以优化求得判别损失函数的最小值为目标。对于生成器的优化目标是通过优化使得目标函数取得最小值。则目标函数可以表示为：

$$\min_G \max_D V\left(D, G, {}_{y}^{r}\right) = E_{x\sim P_{data}} E^{r}_{y\sim P_{cond}\epsilon y_1 * \ldots y_n}\left[logD\left(x|_{y}^{r}\right)\right] + E_{z\sim P_z(z)}$$

$$E^{r}_{y\sim P_{cond}\epsilon y_1 * \ldots y_n}\left[\log\left(1 - \mathrm{D}\left(\mathrm{G}\left(\mathrm{z}|_{\mathrm{y}}^{\mathrm{r}}\right)\right)\right)\right]$$

其中，P_{cond}为样本条件所包括的各个特征条件的联合分布概率，$y_1 * \ldots y_n$为n个特征条件y所组成的条件概率空间y_{cond}，则服从条件概率空间的联合分布概率为：

$$P_{cond}: y_1 * \ldots y_n$$

若输入数据为真实样本数据，则联合分布概率P_{cond}、噪声空间$p_z(z)$为定量及概率空间y_{cond}均为定量。

308、服务器102根据模型函数优化网络参数，构建数据修复模型。

具体的，服务器102在确定上述生成损失函数、判别损失函数及目标函数的过程中，经过反复迭代训练，根据损失函数，优化生成器网络及判别器网络的网络参数，根据优化完成的网络参数，构建数据修复模型。其中，优化的过程可以为：根据判别损失函数的结果，使用自适应时刻估计（Adaptive Moment Estimation，Adam）优化算法对判别器进行优化，在对判别器优化完成的情况

下，根据优化完成优化的判别器对生成器进行优化，根据生成损失函数的结果，使用Adam算法对生成器进行优化，通过生成器及判别器的不断迭代对抗训练，使得损失函数收敛，此处损失函数收敛的过程及目标可以参见步骤307，此处不赘述。进一步的，在损失函数完成收敛及对网络参数的优化过程后，根据优化完成的网络参数，构建数据修复模型。

可选的，在优化网络参数之前，还可以对第二修复数据及第二处理数据进行平均余弦相似度计算，得到相似度结果。根据相似度结果，优化数据修复模型的网络参数。其中，第二处理数据为与第二修复数据相匹配的真实数据。具体的，生成的第二修复数据的数据序列可以表示为：

$$R_m = \{r_{m,t_1}, r_{m,t_2} \dots r_{m,t_l}\}$$

其中，m为生成的数据序列的迭代次数，l为数据序列的长度，

第二处理数据的数据序列可以表示为：

$$R_{groudtruth} = \{r'_{t_1}, r'_{t_2} \dots r'_{t_k}\}$$

其中，k为第二处理数据的长度，该第二处理数据与第二修复数据相对应的原始真实样本数据。则对第二处理数据及第二修复数据进行平均余弦相似度计算，该相似度的计算方法可以表示为：

$$S = \frac{1}{m}\sum_{j=0}^{m}\frac{\sum_1^k r_{j,t_i} * r'_{t_i}}{\sqrt{\sum_{i=1}^k (r_{j,t_i})^2} * \sqrt{\sum_{i=1}^k (r'_{t_i})^2}}$$

进一步的，可以在每轮迭代训练时，计算本次训练的平均余弦相似度，根据平均余弦相似度结果，对网络参数进行优化，以使生成器可以生成更接近真实样本分布的修复数据。

举例来说，可以根据不同的条件通道的开关模式，进行基于平均余弦相似度的网络参数优化过程。本实施方式以三个特征条件为例，则四种不同的条件通道的开关模式可以为：全闭合（无特征条件）、部分闭合（一个特征条件及两个特征条件）及全开（三个特征条件），此处对条件通道的开关模式调控可以参见步骤307中的相关描述，此处不赘述。则在不同条件通道开关模式下，余弦相似度的结果随着迭代训练的变化情况示意图可以参见图5，如图5所示，在条件通道全开状态下，即考虑全部特征条件时，生成的修复数据与真实样本数据的相似度最高，生成的修复数据更接近真实样本的分布。且随着对网络参数的优化，生成的修复数据越接近真实样本数据。

通过执行本实施方式，可以更直观的展示修复数据的生成质量及对网络的训练情况。多条件信息的引入有助于学习到样本的丰富特征，以使数据修复模型可以生成更接近真实分布的修复数据，提高了生成数据的质量及效率。

可见，通过实施图3所描述的方法，服务器102在获取到样本数据及对应的第一样本条件的情况下，对样本数据经过归一化处理，得到第一处理数据，对第一样本条件进行张量化处理，得到第一样本标签。并且服务器102在获取到真实样本数据及第二样本条件的情况下，对真实样本数据进行归一化处理，对第二样本条件进行张量化处理，得到第二处理数据及第二样本标签。则可以根据第一处理数据、第一样本标签、第二处理数据及第二样本标签对数据修复模型进行监督训练，确定包括生成损失函数、判别损失函数及目标函数的模型函数，并根据模型函数优化网络参数，根据优化完成的网络参数，构建数据修复模型。通过执行本实施方式，根据已知的时间序列样本数据进行监督训练，即可得到数据修复模型，无需大量的历史数据或手动获取匹配度高的样本数据作为训练基础，解决了实验成本高，样本获取困难的问题。通过引入与数据相对应的特征条件，可以学习到样本的丰富特征，使得生成的修复数据可以更接近真实的样本分布，采用LSTM作为生成器及判别器的内置网络，保证了生成的修复数据的时序性，提高了数据修复过程的效率及质量。

请参见图6，图6是本申请实施例提供的一种基于多条件约束的时间序列数据生成方法的流程示意图，如图6所示，该时间序列数据生成方法可以包括601~606部分，其中：

601、服务器102获取验证数据，并对验证数据进行归一化处理，得到第三处理数据。

具体的，服务器102可以从客户端101或其他数据平台获取验证数据，该验证数据可以理解为一种样本数据，可选的，样本数据可以包括训练数据、验证数据及测试数据。则对验证数据进行归一化处理得到第三处理数据的相关过程可以参见步骤302中对样本数据进行归一化处理的相关描述，此处不赘述。

602、服务器102获取验证条件，并对验证条件进行张量化处理，得到验证标签。

具体的，服务器102获取与步骤401中的验证数据相匹配的验证条件，并对该验证条件进行张量化处理，得到验证标签。此处关于验证条件及验证标签获

取方法的相关描述可以参见步骤601及步骤303中关于第一样本条件及其获取方法的相关描述，此处不赘述。

603、服务器102调用已完成训练的数据修复模型，根据验证标签对第三处理数据进行修复处理，得到第三修复数据。

具体的，服务器102调用已完成训练的数据修复模型，根据验证标签对第三处理数据进行修复处理，得到第三修复数据的过程，可以参见步骤204中第一修复数据的生成过程，此处不赘述。

604、服务器102获取真实验证数据，并对真实验证数据进行归一化处理，得到真实验证数据的第四处理数据。

具体的，服务器102获取真实验证数据，并对真实验证数据进行归一化处理的相关描述，可以参见步骤304及步骤305中关于真实样本数据的获取及归一化处理方法的相关描述，此处不赘述。该第四处理数据为与第三修复数据相匹配的真实数据。

605、服务器102对第三修复数据及第四处理数据进行残差分析，得到残差分析结果。

具体的，服务器102对第三修复数据及第四处理数据进行残差分析，得到残差分析结果。该残差分析结果可以为残差分析结果图，通过该残差分析结果图可以更直观的展示第三修复数据的生成质量。

举例来说，可以基于不同条件通道的开关模式，进行残差分析过程。本实施方式以三个特征条件为例，则三种不同的条件通道的开关模式可以为：全闭合（无特征条件）、单通道开启（单特征条件）及多通道开启（多特征条件），此处对条件通道的开关模式调控可以参见步骤307中的相关描述，此处不赘述。则在不同条件通道开关模式下，生成的第三修复数据与真实的第四处理数据对比示意图可以参见图7(a)、图7(b)及图7(c)，如图7(a)、图7(b)及图7(c)所示，引入的特征条件越多，生成的第三修复数据接近第四处理处理数据，即生成的修复数据越接近真实样本的分布。进一步的，对图7(a)、图7(b)及图7(c)中相应的第三修复数据与真实的第四处理数据分别进行残差分析，残差分析结果示意图可以分别参见图8(a)、图8(b)及图8(c)，其中，如图8(a)中区域1所显示的深灰色的部分，为残差分析不接受的部分，即得到的第三修复数据偏离真实的第四处理数据较大。如图8(a)、图8(b)及图8(c)所示，多特征的引入使得生成的第三修复数据接近第四处理处理数据，即生成的修复数据越接近真实样本的分布。所

以，基于多条件约束的MCGAN在修复时间序列数据时，可以得到更准确、更接近真实分布情况的修复数据，提高了数据修复的效率及质量。

606、服务器102发送残差分析结果至客户端101。

具体的，服务器102将该残差分析结果发送至客户端101，相应的，客户端101接收该残差分析结果，以使客户端101将该残差分析结果展示给客户端101的操作用户103，操作用户103可以根据该残差分析结果直观的评估生成的修复数据的质量及该数据修复模型的训练情况。

可见，通过实施图6所描述的方法，服务器102在获取到验证数据及验证条件的情况下，对验证数据进行归一化处理，得到第三处理数据，并对验证条件进行张量化处理，得到验证标签。调用已完成训练的数据修复模型，根据验证标签对第三处理数据进行修复处理，得到第三修复数据，根据获取到的真实的第四处理数据对第三修复数据进行残差分析，得到残差分析结果，并将该残差分析结果发送给客户端。通过执行本实施方式，可以更直观、准确的评估生成的修复数据的质量及该数据修复模型的训练情况，并且，还可以确定在多特征条件约束下生成的修复数据可以更接近真实样本的分布情况，多特征条件的引入可以获取到修复数据更丰富的特征，提高了数据修复的效率及质量。

基于上述方法实施例的描述，本申请实施例还提出一种基于多条件约束的时间序列数据生成装置。该基于多条件约束的时间序列数据生成装置可以是运行于处理设备中的计算机程序（包括程序代码）；请参见图9所示，该图像可视化处理装置可以运行如下单元：

收发单元901，用于接收客户端的数据修复请求，所述数据修复请求包括待修复数据及条件信息，所述数据修复请求用于请求根据所述条件信息对所述待修复数据进行数据修复，所述条件信息为与所述待修复数据相匹配的特征条件；

处理单元902，用于对所述待修复数据进行归一化处理，得到所述待修复数据的归一化数据，并对所述条件信息进行张量化处理，得到所述条件信息的特征标签；调用已完成训练的数据修复模型，根据所述特征标签对所述归一化数据进行修复处理，得到第一修复数据，所述数据修复模型是根据样本数据、第一样本条件、真实样本数据及第二样本条件对所述数据修复模型进行训练得到的，所述样本数据为噪声数据；

所述收发单元901，还用于发送所述第一修复数据至所述客户端。

在一种实施方式中，所述待修复数据包括时间点序列；

所述根据所述特征标签对所述归一化数据进行修复处理，得到第一修复数据，处理单元901，还可用于根据所述时间点序列，对所述归一化数据中各个数据进行排序，所述时间点序列为所述待修复数据中各个数据的生成时间点所组成的序列，所述归一化数据中各个数据为所述待修复数据中各个数据经过归一化处理后得到的；

根据所述特征标签，对所述已完成排序的所述归一化数据进行数据修复处理，得到第一修复数据。

再一种实施方式中，所述调用已完成训练的数据修复模型之前，处理单元901，还可用于获取所述样本数据及所述第一样本条件，并对所述样本数据进行归一化处理，得到所述样本数据的第一处理数据，对所述第一样本条件进行张量化处理，得到第一样本标签；

获取所述真实样本数据及所述第二样本条件，并对所述真实样本数据进行归一化处理，得到所述真实样本数据的第二处理数据，对所述第二样本条件进行张量化处理，得到第二样本标签；

根据所述第一处理数据、所述第一样本标签、所述第二处理数据及所述第二样本标签对所述数据修复模型进行监督训练，确定模型函数；

根据所述模型函数优化网络参数，构建所述数据修复模型。

再一种实施方式中，所述第一样本条件包括n个特征条件，n为正整数；

所述获取所述样本数据及所述第一样本条件之前，处理单元901，还可用于接收所述客户端发送的条件指令，所述条件指令用于指示从n个所述特征条件中获取x个所述特征条件，x为小于或等于n的非负整数；

根据所述条件指令从所述包括n个特征条件的第一样本条件信息中获取x个所述特征条件。

再一种实施方式中，所述模型函数包括生成损失函数、判别损失函数及目标函数；

所述根据所述第一处理数据、所述第一样本标签、所述第二处理数据及所述第二样本标签对所述数据修复模型进行监督训练，确定模型函数，处理单元901，还可用于根据所述第一样本标签对所述第一处理数据进行修复处理，得到第二修复数据，并对所述第二修复数据及所述第一样本标签进行判别处理，得到第一判别结果；

对所述第二处理数据及所述第二样本标签进行判别处理，得到第二判别结果；

根据所述第一判别结果及所述第二判别结果，确定所述判别损失函数，所述判别损失函数为判别器的损失函数；

根据所述第一判别结果，确定所述生成损失函数，所述生成损失函数为生成器的损失函数；

对所述判别损失函数及所述生成损失函数进行优化，确定所述目标函数。

再一种实施方式中，处理单元901，还可用于对所述第二修复数据及所述第二处理数据进行平均余弦相似度计算，得到相似度结果，所述第二处理数据为与所述第二修复数据相匹配的真实数据；

根据所述相似度结果，优化所述数据修复模型的网络参数。

再一种实施方式中，所述样本数据包括验证数据，处理单元901，还可用于获取所述验证数据，并对所述验证数据进行归一化处理，得到所述验证数据的第三处理数据；

获取验证条件，并对所述验证条件进行张量化处理，得到所述验证条件的验证标签；

调用已完成训练的所述数据修复模型，根据所述验证标签对所述第三处理数据进行修复处理，得到第三修复数据；

获取真实验证数据，并对所述真实验证数据进行归一化处理，得到所述真实验证数据的第四处理数据，所述第四处理数据为与所述第三修复数据相匹配的真实数据；

对所述第三修复数据及所述第四处理数据进行残差分析，得到残差分析结果。

收发单元901，还可用于将所述残差分析结果发送至所述客户端。

根据本申请的一个实施例，图2、图3及图6所示的基于多条件约束的时间序列数据生成方法所涉及的部分步骤可由基于多条件约束的时间序列数据生成装置中的处理单元来执行。例如，图2中所示的步骤201和205可由收发单元901执行；又如，图2所示的步骤203可由处理单元902执行。根据本申请的另一个实施例，基于多条件约束的时间序列数据生成装置中的各个单元可以分别或全部合并为一个或若干个另外的单元来构成，或者其中的某个（些）单元还可以再拆分为功能上更小的多个单元来构成，这可以实现同样的操作，而不影响本申请

的实施例的技术效果的实现。

请参见图10，是本申请实施例提供的一种基于多条件约束的时间序列数据生成装置的结构示意图，该数据生成装置包括处理器1001、存储器1002及通信接口1003，处理器1001、存储器1002及通信接口1003通过至少一条通信总线连接，处理器1001被配置为支持处理设备执行图2、图3及图6方法中处理设备相应的功能。

存储器1002用于存放有适于被处理器加载并执行的至少一条指令，这些指令可以是一个或一个以上的计算机程序（包括程序代码）。

通信接口1003用于接收数据和用于发送数据。例如，通信接口1003用于发送数据修复请求等。

在本申请实施例中，该处理器1001可以调用存储器1002中存储的程序代码以执行以下操作：

通过通信接口1003接收客户端的数据修复请求，所述数据修复请求包括待修复数据及条件信息，所述数据修复请求用于请求根据所述条件信息对所述待修复数据进行数据修复，所述条件信息为与所述待修复数据相匹配的特征条件；

对所述待修复数据进行归一化处理，得到所述待修复数据的归一化数据，并对所述条件信息进行张量化处理，得到所述条件信息的特征标签；

调用已完成训练的数据修复模型，根据所述特征标签对所述归一化数据进行修复处理，得到第一修复数据，所述数据修复模型是根据样本数据、第一样本条件、真实样本数据及第二样本条件对所述数据修复模型进行训练得到的，所述样本数据为噪声数据；

通过通信接口1003发送所述第一修复数据至所述客户端。

作为一种可选的实施方式，所述待修复数据包括时间点序列；

所述根据所述特征标签对所述归一化数据进行修复处理，得到第一修复数据，该处理器1001可以调用存储器1002中存储的程序代码以执行以下操作：

根据所述时间点序列，对所述归一化数据中各个数据进行排序，所述时间点序列为所述待修复数据中各个数据的生成时间点所组成的序列，所述归一化数据中各个数据为所述待修复数据中各个数据经过归一化处理后得到的；

根据所述特征标签，对所述已完成排序的所述归一化数据进行数据修复处理，得到第一修复数据。

作为一种可选的实施方式，所述调用已完成训练的数据修复模型之前，该

处理器1001可以调用存储器1002中存储的程序代码以执行以下操作：

获取所述样本数据及所述第一样本条件，并对所述样本数据进行归一化处理，得到所述样本数据的第一处理数据，对所述第一样本条件进行张量化处理，得到第一样本标签；

获取所述真实样本数据及所述第二样本条件，并对所述真实样本数据进行归一化处理，得到所述真实样本数据的第二处理数据，对所述第二样本条件进行张量化处理，得到第二样本标签；

根据所述第一处理数据、所述第一样本标签、所述第二处理数据及所述第二样本标签对所述数据修复模型进行监督训练，确定模型函数；

根据所述模型函数优化网络参数，构建所述数据修复模型。

作为一种可选的实施方式，所述第一样本条件包括n个特征条件，n为正整数；

所述获取所述样本数据及所述第一样本条件之前，该处理器1001可以调用存储器1002中存储的程序代码以执行以下操作：

接收所述客户端发送的条件指令，所述条件指令用于指示从n个所述特征条件中获取x个所述特征条件，x为小于或等于n的非负整数；

根据所述条件指令从所述包括n个特征条件的第一样本条件信息中获取x个所述特征条件。

作为一种可选的实施方式，所述模型函数包括生成损失函数、判别损失函数及目标函数；

所述根据所述第一处理数据、所述第一样本标签、所述第二处理数据及所述第二样本标签对所述数据修复模型进行监督训练，确定模型函数，该处理器1001可以调用存储器1002中存储的程序代码以执行以下操作：

根据所述第一样本标签对所述第一处理数据进行修复处理，得到第二修复数据，并对所述第二修复数据及所述第一样本标签进行判别处理，得到第一判别结果；

对所述第二处理数据及所述第二样本标签进行判别处理，得到第二判别结果；

根据所述第一判别结果及所述第二判别结果，确定所述判别损失函数，所述判别损失函数为判别器的损失函数；

根据所述第一判别结果，确定所述生成损失函数，所述生成损失函数为生

成器的损失函数；

对所述判别损失函数及所述生成损失函数进行优化，确定所述目标函数。

作为一种可选的实施方式，该处理器1001可以调用存储器1002中存储的程序代码以执行以下操作：

对所述第二修复数据及所述第二处理数据进行平均余弦相似度计算，得到相似度结果，所述第二处理数据为与所述第二修复数据相匹配的真实数据；

根据所述相似度结果，优化所述数据修复模型的网络参数。

作为一种可选的实施方式，所述样本数据包括验证数据，该处理器1001可以调用存储器1002中存储的程序代码以执行以下操作：

获取所述验证数据，并对所述验证数据进行归一化处理，得到所述验证数据的第三处理数据；

获取验证条件，并对所述验证条件进行张量化处理，得到所述验证条件的验证标签；

调用已完成训练的所述数据修复模型，根据所述验证标签对所述第三处理数据进行修复处理，得到第三修复数据；

获取真实验证数据，并对所述真实验证数据进行归一化处理，得到所述真实验证数据的第四处理数据，所述第四处理数据为与所述第三修复数据相匹配的真实数据；

对所述第三修复数据及所述第四处理数据进行残差分析，得到残差分析结果，并通过通信接口1003将所述残差分析结果发送至所述客户端。

本申请实施例还提供了一种计算机可读存储介质（Memory），可以用于存储图2及图3中所示实施例中处理设备所用的计算机软件指令，在该存储空间中还存放了适于被处理器加载并执行的至少一条指令，这些指令可以是一个或一个以上的计算机程序（包括程序代码）。

上述计算机可读存储介质包括但不限于快闪存储器、硬盘、固态硬盘。

本领域普通技术人员可以意识到，结合本申请中所公开的实施例描述的各示例的单元及算法步骤，能够以电子硬件、或者计算机软件和电子硬件的结合来实现。这些功能究竟以硬件还是软件方式来执行，取决于技术方案的特定应用和设计约束条件。专业技术人员可以对每个特定的应用来使用不同方法来实现所描述的功能，但是这种实现不应认为超出本申请的范围。

在上述实施例中，可以全部或部分地通过软件、硬件、固件或者其任意组合来实现。当使用软件实现时，可以全部或部分地以计算机程序产品的形式实现。计算机程序产品包括一个或多个计算机指令。在计算机上加载和执行计算机程序指令时，全部或部分地产生按照本申请实施例的流程或功能。计算机可以是通用计算机、专用计算机、计算机网络、或者其他可编程装置。计算机指令可以存储在计算机可读存储介质中 ,或者通过计算机可读存储介质进行传输。计算机指令可以从一个网站站点、计算机、服务器或数据中心通过有线（例如同轴电缆、光纤、数字用户线（DSL））或无线（例如红外、无线、微波等）方式向另一个网站站点、计算机、服务器或数据中心进行传输。计算机可读存储介质可以是计算机能够存取的任何可用介质或者是包含一个或多个可用介质集成的服务器、数据中心等数据存储设备。可用介质可以是磁性介质，(例如，软盘、硬盘、磁带)、光介质(例如，DVD)、或者半导体介质（例如固态硬盘(Solid State Disk，SSD)）等。

以上所述的具体实施方式，对本申请的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上所述仅为本申请的具体实施方式而已，并不用于限定本申请的保护范围，凡在本申请的技术方案的基础之上，所做的任何修改、等同替换、改进等，均应包括在本申请的保护范围之内。

权利要求

1、一种基于多条件约束的时间序列数据生成方法，其特征在于，所述方法包括：
接收客户端的数据修复请求，所述数据修复请求包括待修复数据及条件信息，所述数据修复请求用于请求根据所述条件信息对所述待修复数据进行数据修复，所述条件信息为与所述待修复数据相匹配的特征条件；
对所述待修复数据进行归一化处理，得到所述待修复数据的归一化数据，并对所述条件信息进行张量化处理，得到所述条件信息的特征标签；
调用已完成训练的数据修复模型，根据所述特征标签对所述归一化数据进行修复处理，得到第一修复数据，所述数据修复模型是根据样本数据、第一样本条件、真实样本数据及第二样本条件对所述数据修复模型进行训练得到的，所述样本数据为噪声数据；
发送所述第一修复数据至所述客户端。

2、根据权利要求 1 所述的方法，其特征在于，所述待修复数据包括时间点序列；所述根据所述特征标签对所述归一化数据进行修复处理，得到第一修复数据，包括：
根据所述时间点序列，对所述归一化数据中各个数据进行排序，所述时间点序列为所述待修复数据中各个数据的生成时间点所组成的序列，所述归一化数据中各个数据为所述待修复数据中各个数据经过归一化处理后得到的；
根据所述特征标签，对所述已完成排序的所述归一化数据进行数据修复处理，得到第一修复数据。

3、根据权利要求 1 所述的方法，其特征在于，所述调用已完成训练的数据修复模型之前，所述方法还包括：
获取所述样本数据及所述第一样本条件，并对所述样本数据进行归一化处理，得到所述样本数据的第一处理数据，对所述第一样本条件进行张量化处理，得到第一样本标签；
获取所述真实样本数据及所述第二样本条件，并对所述真实样本数据进行归一化处理，得到所述真实样本数据的第二处理数据，对所述第二样本条件进行张量化处理，得到第二样本标签；
根据所述第一处理数据、所述第一样本标签、所述第二处理数据及所述第二样本标签对所述数据修复模型进行监督训练，确定模型函数；
根据所述模型函数优化网络参数，构建所述数据修复模型。

4、根据权利要求 3 所述的方法，其特征在于，所述第一样本条件包括 n 个特征条件，n 为正整数；

所述获取所述样本数据及所述第一样本条件之前，所述方法还包括：

接收所述客户端发送的条件指令，所述条件指令用于指示从 n 个所述特征条件中获取 x 个所述特征条件，x 为小于或等于 n 的非负整数；

根据所述条件指令从所述包括 n 个特征条件的第一样本条件信息中获取 x 个所述特征条件。

5、根据权利要求 3 所述的方法，其特征在于，所述模型函数包括生成损失函数、判别损失函数及目标函数；

所述根据所述第一处理数据、所述第一样本标签、所述第二处理数据及所述第二样本标签对所述数据修复模型进行监督训练，确定模型函数，包括：

根据所述第一样本标签对所述第一处理数据进行修复处理，得到第二修复数据，并对所述第二修复数据及所述第一样本标签进行判别处理，得到第一判别结果；

对所述第二处理数据及所述第二样本标签进行判别处理，得到第二判别结果；

根据所述第一判别结果及所述第二判别结果，确定所述判别损失函数，所述判别损失函数为判别器的损失函数；

根据所述第一判别结果，确定所述生成损失函数，所述生成损失函数为生成器的损失函数；

对所述判别损失函数及所述生成损失函数进行优化，确定所述目标函数。

6、根据权利要求 5 所述的方法，其特征在于，所述方法还包括：

对所述第二修复数据及所述第二处理数据进行平均余弦相似度计算，得到相似度结果，所述第二处理数据为与所述第二修复数据相匹配的真实数据；

根据所述相似度结果，优化所述数据修复模型的网络参数。

7、根据权利要求 3 所述的方法，其特征在于，所述样本数据包括验证数据，所述方法还包括：

获取所述验证数据，并对所述验证数据进行归一化处理，得到所述验证数据的第三处理数据；

获取验证条件，并对所述验证条件进行张量化处理，得到所述验证条件的验证标签；

调用已完成训练的所述数据修复模型，根据所述验证标签对所述第三处理数据进行修复处理，得到第三修复数据；

获取真实验证数据，并对所述真实验证数据进行归一化处理，得到所述真实验证数据的第四处理数据，所述第四处理数据为与所述第三修复数据相匹配的真实数据；

对所述第三修复数据及所述第四处理数据进行残差分析，得到残差分析结果，并将所述残差分析结果发送至所述客户端。

8、一种基于多条件约束的时间序列数据生成装置，其特征在于，包括：

收发单元，用于接收客户端的数据修复请求，所述数据修复请求包括待修复数据及条件信息，所述数据修复请求用于请求根据所述条件信息对所述待修复数据进行数据修复，所述条件信息为与所述待修复数据相匹配的特征条件；

处理单元，用于对所述待修复数据进行归一化处理，得到所述待修复数据的归一化数据，并对所述条件信息进行张量化处理，得到所述条件信息的特征标签；

调用已完成训练的数据修复模型，根据所述特征标签对所述归一化数据进行修复处理，得到第一修复数据，所述数据修复模型是根据样本数据、第一样本条件、真实样本数据及第二样本条件对所述数据修复模型进行训练得到的，所述样本数据为噪声数据；

所述收发单元，还用于发送所述第一修复数据至所述客户端。

9、一种基于多条件约束的时间序列数据生成装置，其特征在于，包括处理器、存储器和通信接口，所述处理器、所述存储器和所述通信接口相互连接，其中，所述存储器用于存储计算机程序，所述计算机程序包括程序指令，所述处理器被配置用于调用所述程序指令，执行如权利要求 1-7 中任一项所述的方法。

10、一种计算机可读存储介质，其特征在于，所述计算机可读存储介质存储有一条或多条指令，所述一条或多条指令适于由处理器加载并执行如权利要求 1-7 任一项所述的方法。

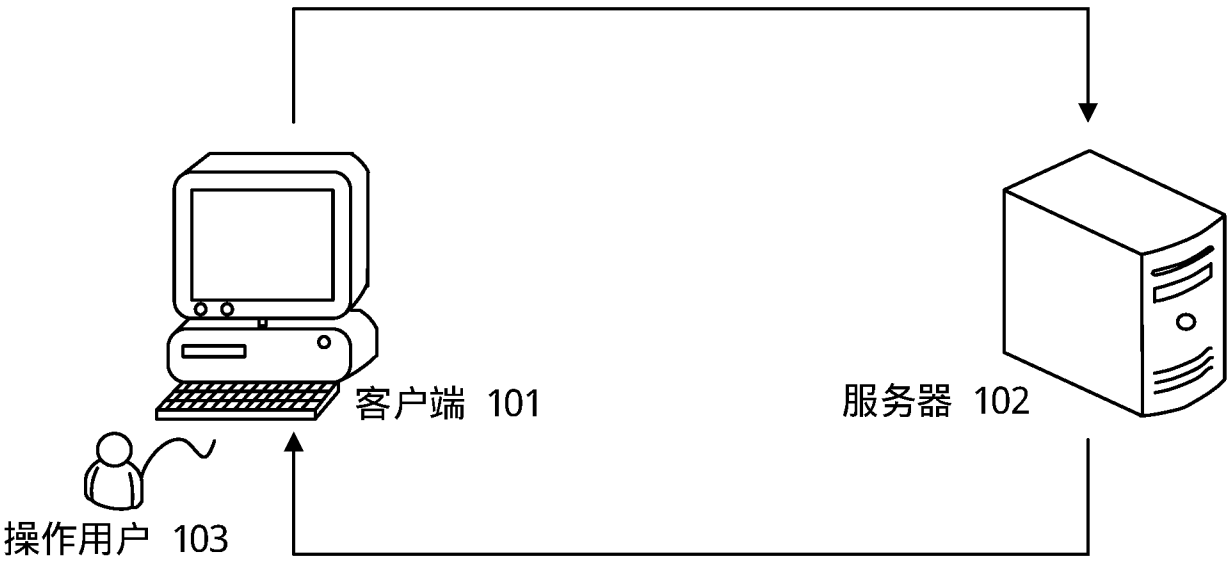


图 1

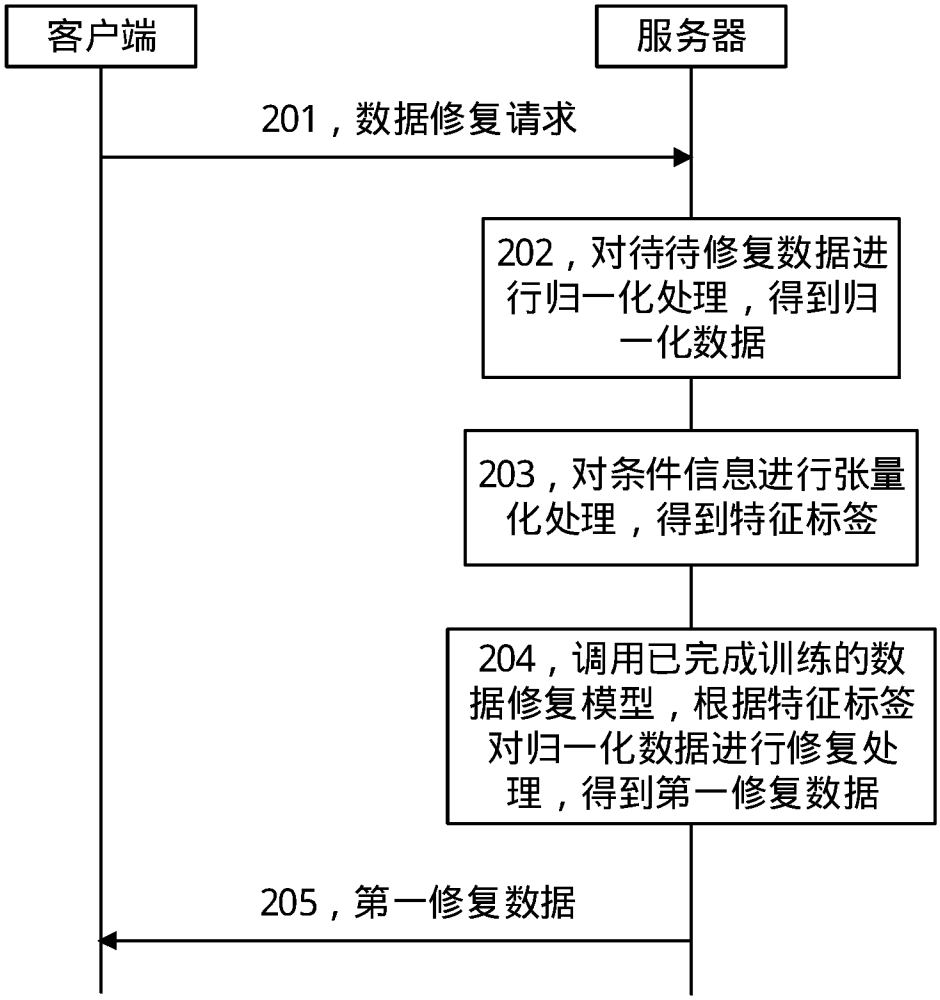

客户端
服务器
201，数据修复请求
202，对待待修复数据进行归一化处理，得到归一化数据
203，对条件信息进行张量化处理，得到特征标签
204，调用已完成训练的数据修复模型，根据特征标签对归一化数据进行修复处理，得到第一修复数据
205，第一修复数据

图 2

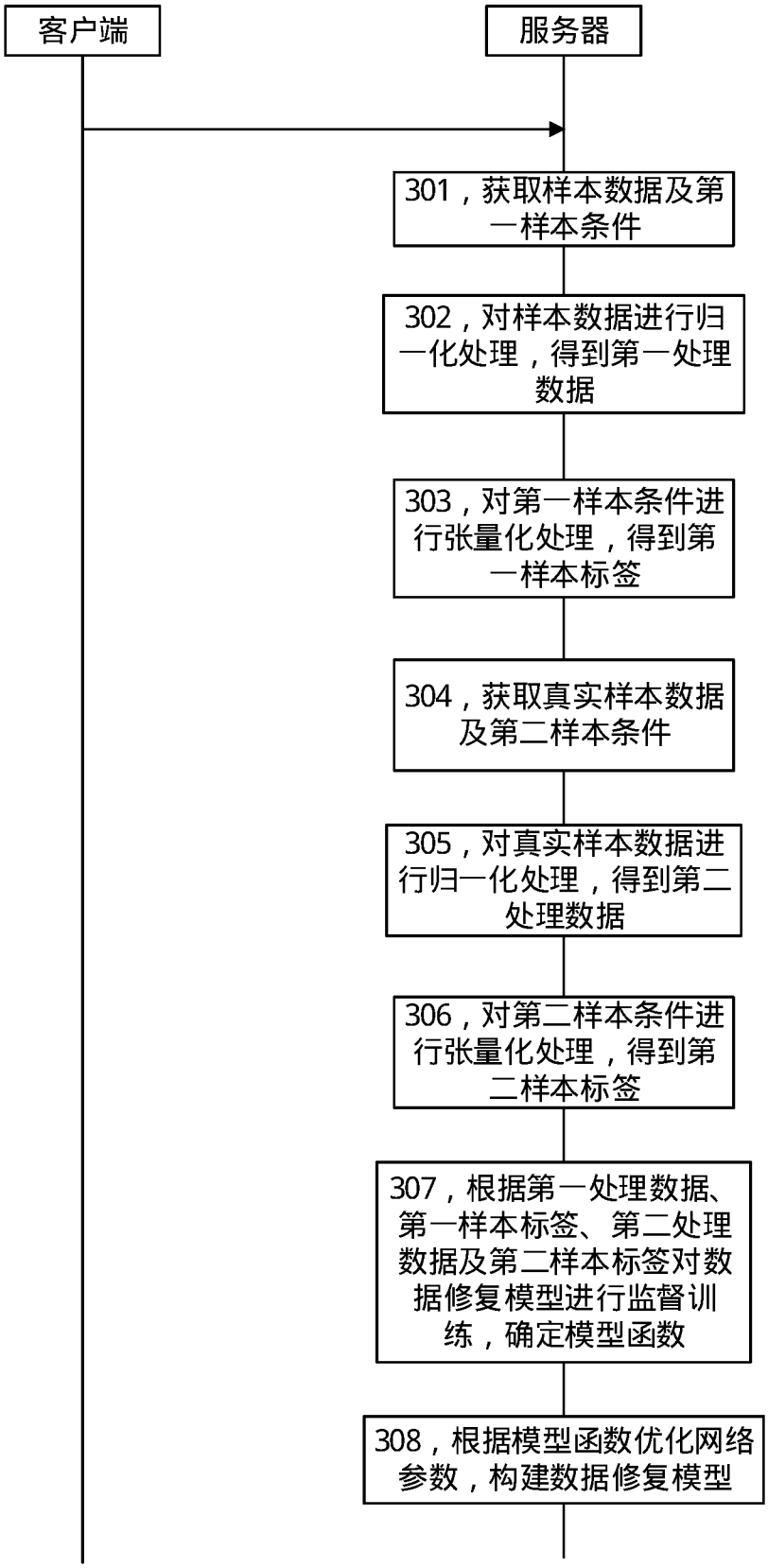

客户端
服务器
301，获取样本数据及第一样本条件
302，对样本数据进行归一化处理，得到第一处理数据
303，对第一样本条件进行张量化处理，得到第一样本标签
304，获取真实样本数据及第二样本条件
305，对真实样本数据进行归一化处理，得到第二处理数据
306，对第二样本条件进行张量化处理，得到第二样本标签
307，根据第一处理数据、第一样本标签、第二处理数据及第二样本标签对数据修复模型进行监督训练，确定模型函数
308，根据模型函数优化网络参数，构建数据修复模型

图 3

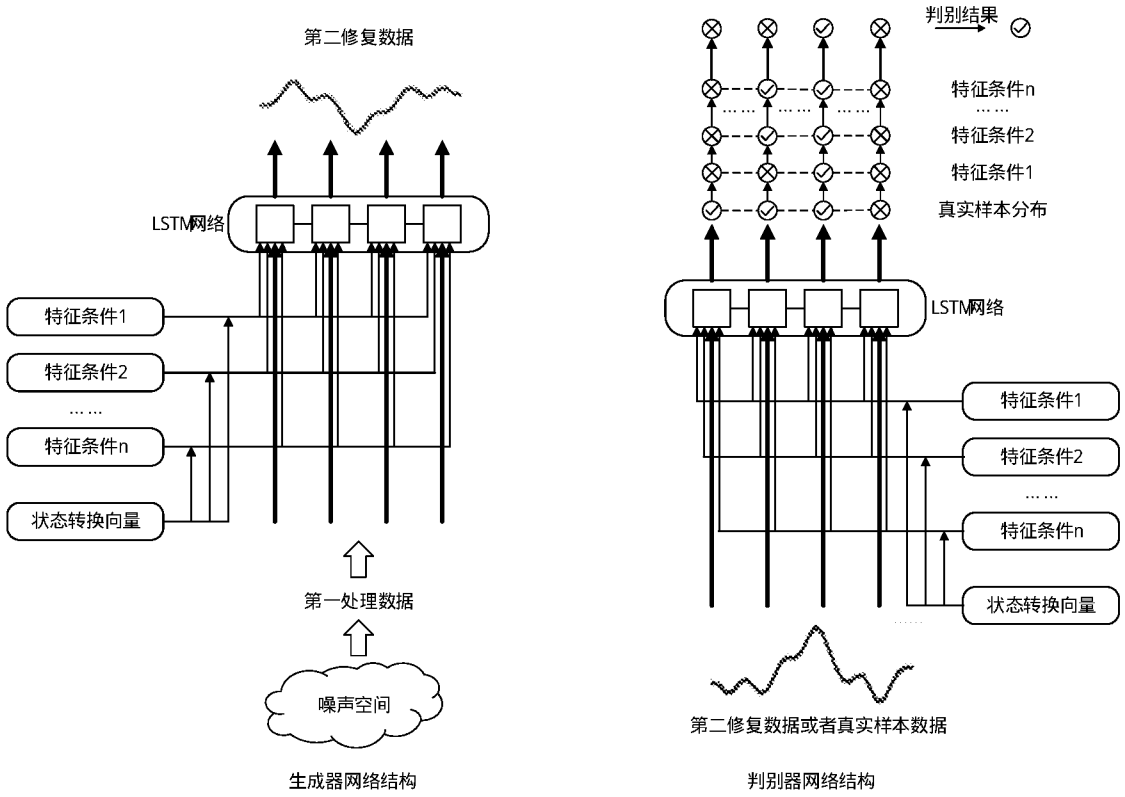

第二修复数据
LSTM网络
特征条件1
特征条件2
……
特征条件n
状态转换向量
第一处理数据
噪声空间
生成器网络结构
判别结果
特征条件n
……
特征条件2
特征条件1
真实样本分布
LSTM网络
特征条件1
特征条件2
……
特征条件n
状态转换向量
第二修复数据或者真实样本数据
判别器网络结构

图 4

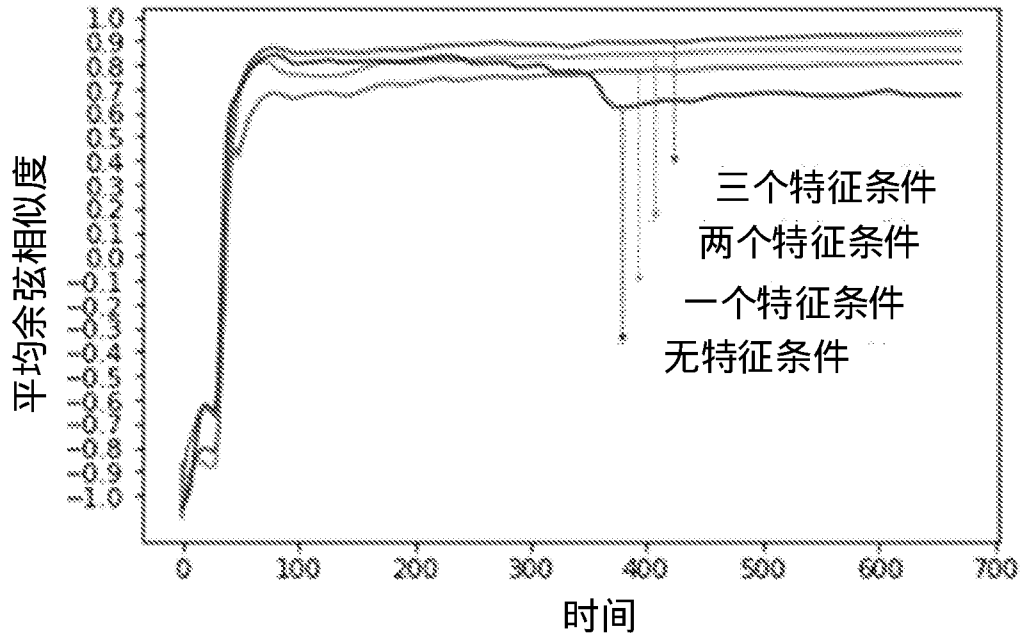

平均余弦相似度
时间
三个特征条件
两个特征条件
一个特征条件
无特征条件
0
100
200
300
400
500
600
700

图 5

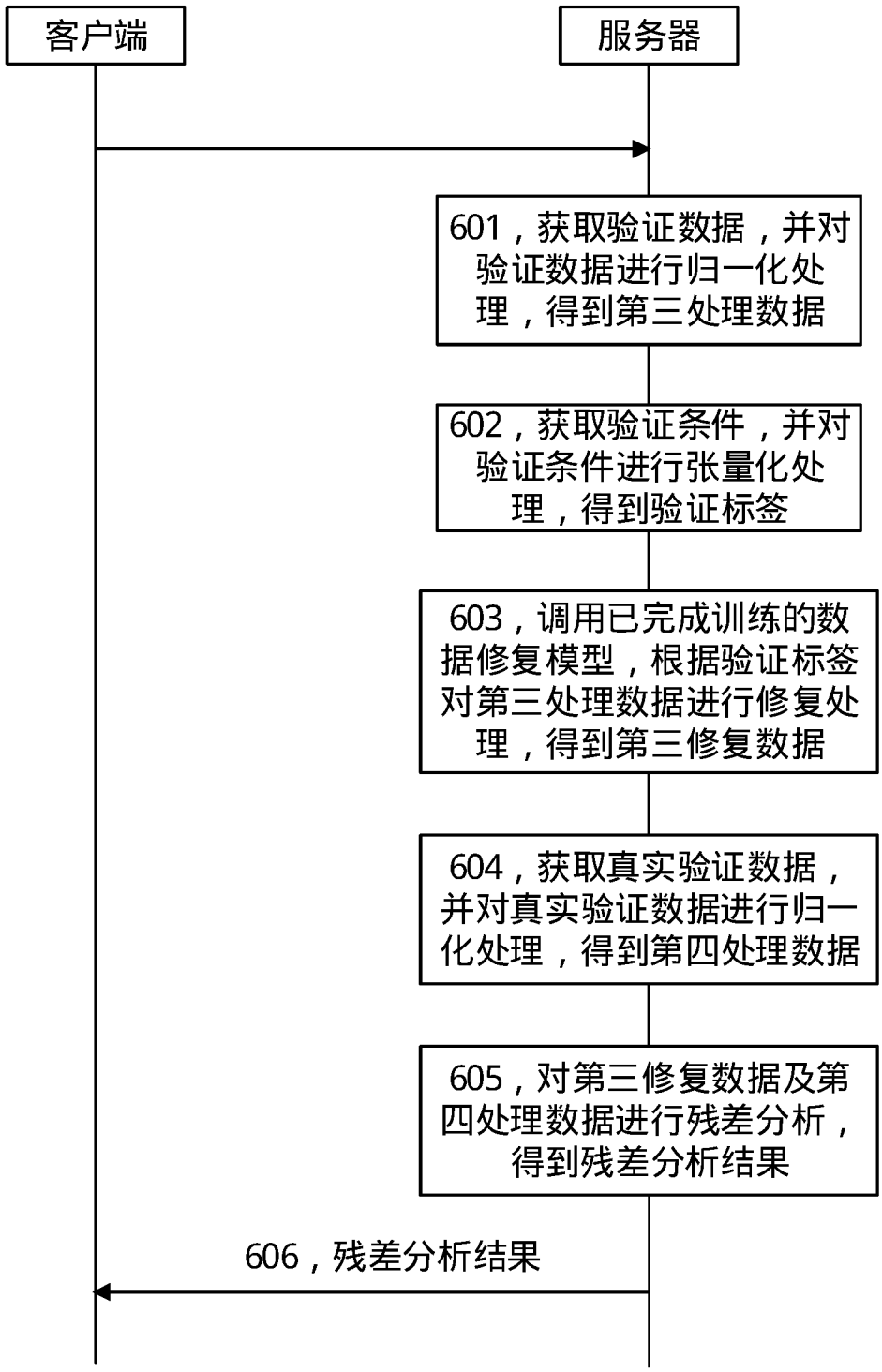

客户端
服务器
601，获取验证数据，并对验证数据进行归一化处理，得到第三处理数据
602，获取验证条件，并对验证条件进行张量化处理，得到验证标签
603，调用已完成训练的数据修复模型，根据验证标签对第三处理数据进行修复处理，得到第三修复数据
604，获取真实验证数据，并对真实验证数据进行归一化处理，得到第四处理数据
605，对第三修复数据及第四处理数据进行残差分析，得到残差分析结果
606，残差分析结果

图 6

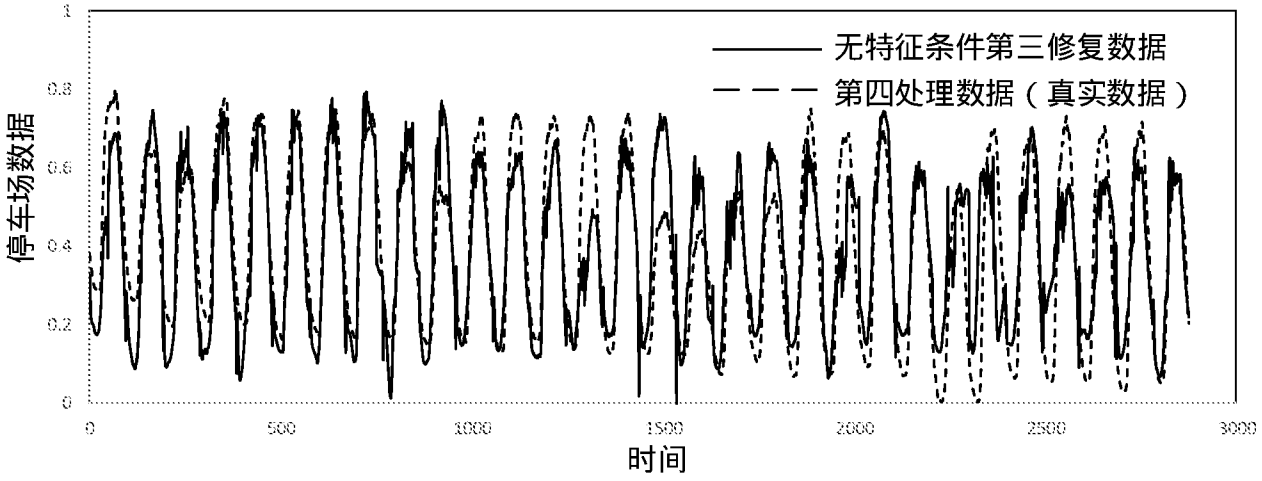

无特征条件第三修复数据
第四处理数据（真实数据）
停车场数据
时间
0
0.2
0.4
0.6
0.8
1
0
500
1000
1500
2000
2500
3000

图 7（a）

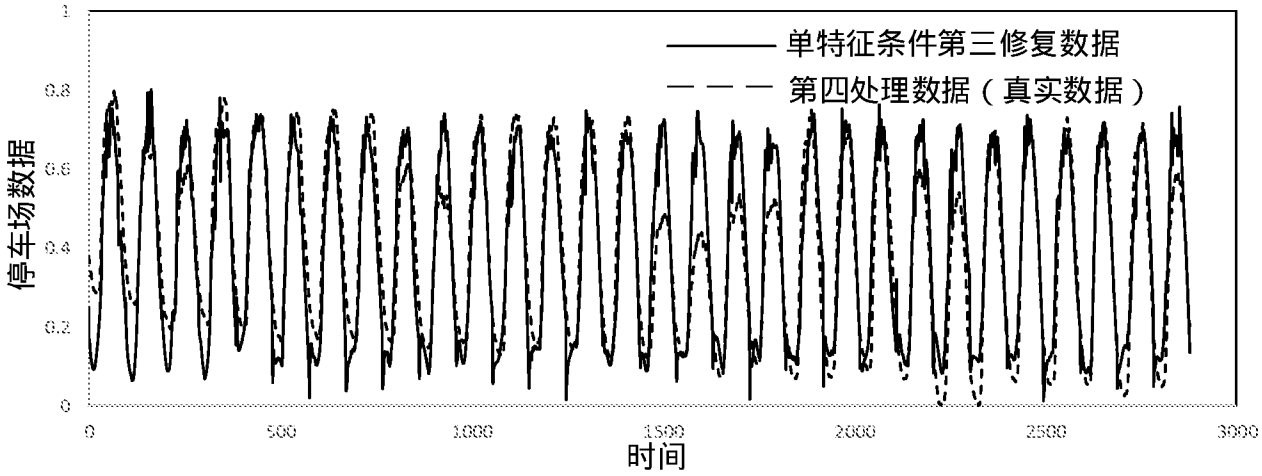

单特征条件第三修复数据
第四处理数据（真实数据）
停车场数据
时间
0
0.2
0.4
0.6
0.8
1
0
500
1000
1500
2000
2500
3000

图 7（b）

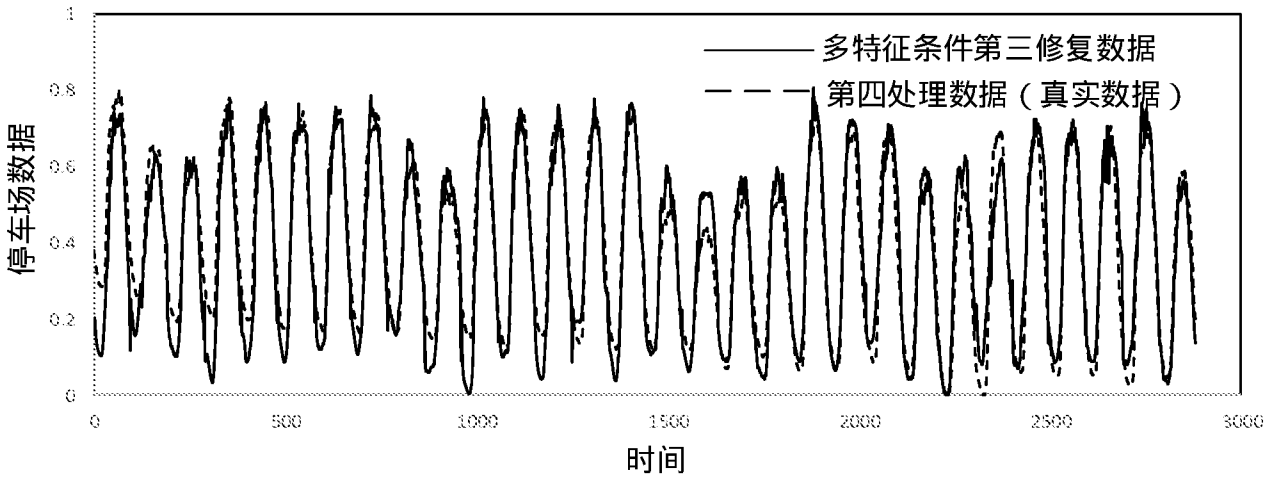

多特征条件第三修复数据
第四处理数据（真实数据）
停车场数据
时间
0
0.2
0.4
0.6
0.8
1
0
500
1000
1500
2000
2500
3000

图 7（c）

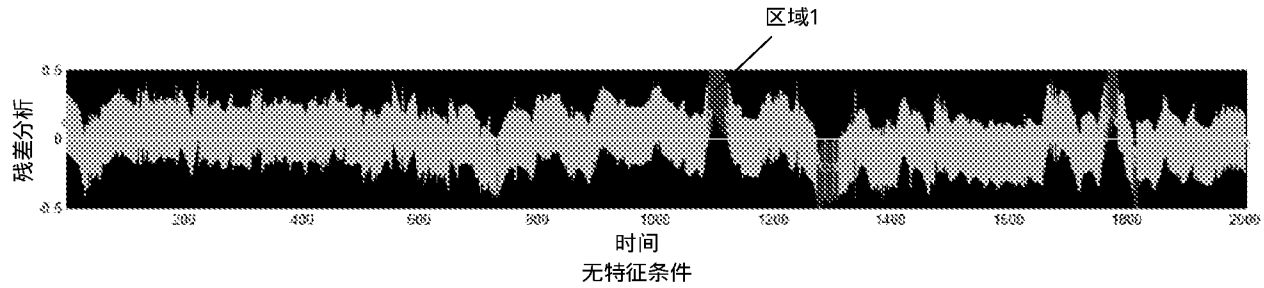


图 8（a）

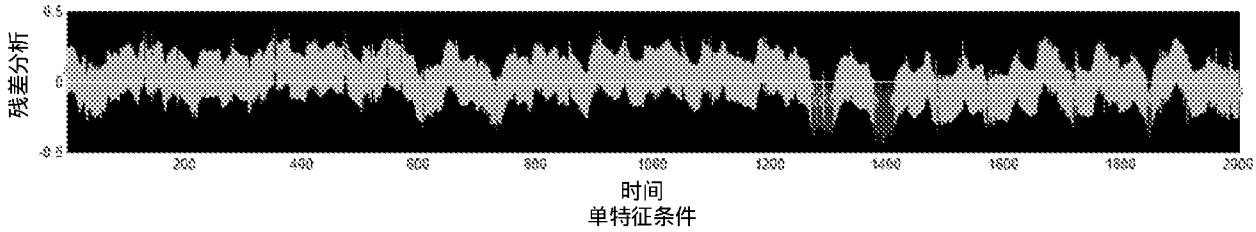


图 8（b）

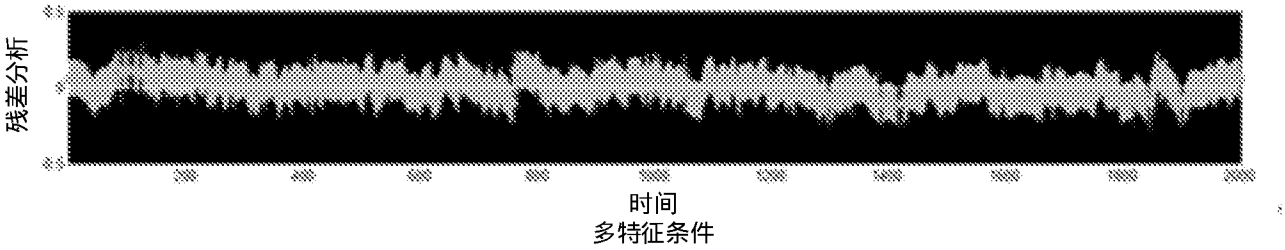


图 8（c）

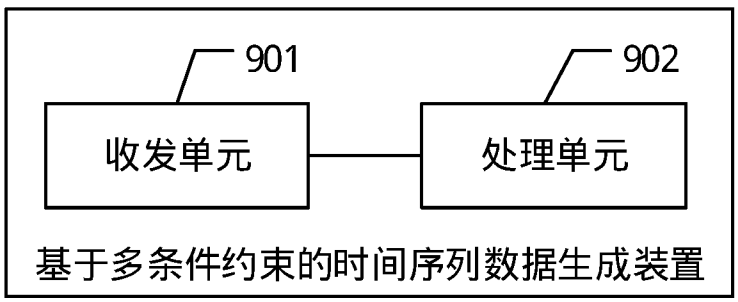


图 9

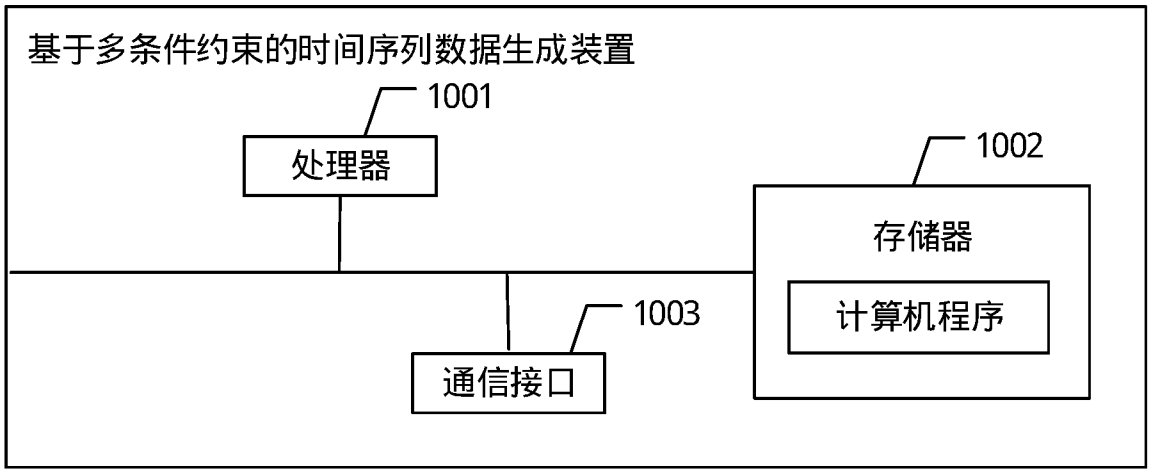


图 10

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2020/081440

A. CLASSIFICATION OF SUBJECT MATTER

G06F 16/215(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT, WPI, EPODOC, CNKI, IEEE, BAIDU: 条件, 约束, 时间序列, 数据, 修复, 特征, 标签, 张量, 训练, 模型, condition, restrict, time sequence, recover, data, character, label, vector, training, model

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 109670580 A (ZHEJIANG UNIVERSITY OF TECHNOLOGY) 23 April 2019 (2019-04-23) description, paragraphs [0006]-[0073]	1-10
A	CN 108009632 A (TSINGHUA UNIVERSITY) 08 May 2018 (2018-05-08) entire document	1-10
A	CN 103577694 A (GUANGDONG OCEAN UNIVERSITY) 12 February 2014 (2014-02-12) entire document	1-10
A	US 2018365521 A1 (ALIBABA GROUP HOLDING LIMITED) 20 December 2018 (2018-12-20) entire document	1-10

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
15 December 2020	**30 December 2020**

Name and mailing address of the ISA/CN	Authorized officer
China National Intellectual Property Administration (ISA/CN) **No. 6, Xitucheng Road, Jimenqiao, Haidian District, Beijing 100088** **China**	
Facsimile No. **(86-10)62019451**	Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2020/081440

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN 109670580 A	23 April 2019	None	
CN 108009632 A	08 May 2018	None	
CN 103577694 A	12 February 2014	None	
US 2018365521 A1	20 December 2018	WO 2017143914 A1	31 August 2017
		TW 201738780 A	01 November 2017
		CN 107122327 A	01 September 2017
		HK 1243517 A0	13 July 2018

国际检索报告

国际申请号
PCT/CN2020/081440

A. 主题的分类

G06F 16/215(2019.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNPAT, WPI, EPODOC, CNKI, IEEE, BAIDU：条件，约束，时间序列，数据，修复，特征，标签，张量，训练，模型，condition, restrict, time sequence, recover, data, character, label, vector, training, model

C. 相关文件

类 型*	引用文件，必要时，指明相关段落	相关的权利要求
A	CN 109670580 A (浙江工业大学) 2019年 4月 23日 (2019-04-23) 说明书第[0006]-[0073]段	1-10
A	CN 108009632 A (清华大学) 2018年 5月 8日 (2018-05-08) 全文	1-10
A	CN 103577694 A (广东海洋大学) 2014年 2月 12日 (2014-02-12) 全文	1-10
A	US 2018365521 A1 (ALIBABA GROUP HOLDING LIMITED) 2018年 12月 20日 (2018-12-20) 全文	1-10

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：
“A” 认为不特别相关的表示了现有技术一般状态的文件
“E” 在国际申请日的当天或之后公布的在先申请或专利
“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）
“O” 涉及口头公开、使用、展览或其他方式公开的文件
“P” 公布日先于国际申请日但迟于所要求的优先权日的文件
“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件
“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性
“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性
“&” 同族专利的文件

国际检索实际完成的日期	国际检索报告邮寄日期
2020年 12月 15日	2020年 12月 30日
ISA/CN的名称和邮寄地址 中国国家知识产权局(ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10)62019451	受权官员 胡百乐 电话号码 86-(10)-53961519

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2020/081440

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	109670580	A	2019年 4月 23日	无			
CN	108009632	A	2018年 5月 8日	无			
CN	103577694	A	2014年 2月 12日	无			
US	2018365521	A1	2018年 12月 20日	WO	2017143914	A1	2017年 8月 31日
				TW	201738780	A	2017年 11月 1日
				CN	107122327	A	2017年 9月 1日
				HK	1243517	A0	2018年 7月 13日