

(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
2017 年 7 月 20 日 (20.07.2017)



(10) 国际公布号
WO 2017/121285 A1

- (51) 国际专利分类号:
G06Q 50/02 (2012.01)
- (21) 国际申请号: PCT/CN2017/070356
- (22) 国际申请日: 2017 年 1 月 6 日 (06.01.2017)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201610024102.6 2016 年 1 月 14 日 (14.01.2016) CN
- (71) 申请人: 阿里巴巴集团控股有限公司 (ALIBABA GROUP HOLDING LIMITED) [—/CN]; 开曼群岛大开曼资本大厦一座四层 847 号邮箱, Grand Cayman (KY)。
- (72) 发明人: 王瑜 (WANG, Yu); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。叶舟 (YE, Zhou); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。王吉能 (WANG, Jineng); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。杨洋 (YANG, Yang); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集

团法务部, Zhejiang 311121 (CN)。董昭萍 (DONG, Zhaoping); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。陈凡 (CHEN, Fan); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。钱倩 (QIAN, Qian); 中国浙江省杭州市余杭区文一西路 969 号 3 号楼 5 楼阿里巴巴集团法务部, Zhejiang 311121 (CN)。

(74) 代理人: 北京三友知识产权代理有限公司 (BEIJING SANYOU INTELLECTUAL PROPERTY AGENCY LTD.); 中国北京市金融街 35 号国际企业大厦 A 座 16 层, Beijing 100033 (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, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

[见续页]

(54) Title: TIME SERIES-BASED DATA PREDICTION METHOD AND DEVICE

(54) 发明名称: 一种基于时间序列的数据预测方法和装置

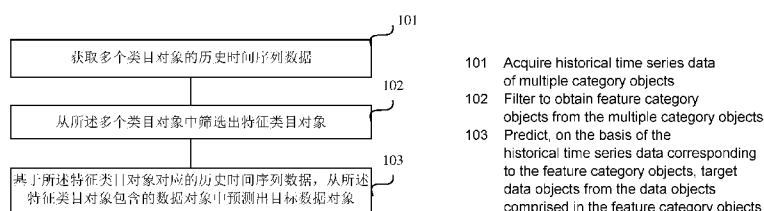


图 1

(57) Abstract: Provided are a time series-based data prediction method and device. The method comprises: acquiring historical time series data of multiple category objects (101), wherein the category object comprises one or more data objects; filtering to obtain feature category objects from the multiple category objects (102), wherein the feature category objects are category objects comprising feature data objects, and the feature data objects are data objects with life cycles less than a preset time threshold; and predicting, on the basis of the historical time series data corresponding to the feature category objects, target data objects from the data objects comprising a preset growth trend and are to be generated within a first preset time period in the future. The method can predict, on the basis of time series data, a target data object having a great market potential in the near future, such that a prediction result is more consistent with reality and more accurate.

(57) 摘要: 一种基于时间序列的数据预测方法和装置, 其中所述方法包括: 获取多个类目对象的历史时间序列数据 (101), 其中, 所述类目对象包括一个或多个数据对象; 从所述多个类目对象中筛选出特征类目对象 (102), 其中, 所述特征类目对象为包含特征数据对象的类目对象, 所述特征数据对象为生命周期小于预设时间阈值的数据对象; 基于所述特征类目对象对应的历史时间序列数据, 从所述特征类目对象包含的数据对象中预测出目标数据对象 (103), 所述目标数据对象为未来第一预设时间段内将要产生的未来时间序列数据满足预设增长趋势的数据对象。本方法可以根据时间序列数据的原理, 预测出近期具有爆发力的目标数据对象, 使得预测结果与实际更加吻合, 准确率更高。



WO 2017/121285 A1



(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))。

一种基于时间序列的数据预测方法和装置

本申请要求 2016 年 01 月 14 日递交的申请号为 201610024102.6、发明名称为“一种基于时间序列的数据预测方法和装置”的中国专利申请的优先权，其全部内容通过引用结合在本申请中。

5

技术领域

本申请涉及数据处理技术领域，特别是涉及一种基于时间序列的数据预测方法和一种基于时间序列的数据预测装置。

10 背景技术

随着信息科技的发展，农村布局成为了越来越多电子商务平台战略布局的一个非常重要的方面：让商品通过电商平台走出去和让外面的商品走进农村去。在农村产品中，很大的部分是一些时效性或者季节性要求较高的商品，甚至于保质期也相当地短暂，如海鲜、河鲜以及新鲜蔬菜水果等。这类商品可以称为时效性商品，时效性商品是指具有

15 一定消费时效特性，且保质期非常短暂的商品。

在实际中，时效性商品的需求虽然庞大，但是对于电商平台及其物流系统的挑战也是巨大的，这体现在两个方面：

（1）如果仓储过多，则会造成物流压力过大、也因本类商品的保质期短，容易造成巨大的浪费；

20 （2）如果错误估计造成仓储不足，则会使得巨大的市场浪费。

因此，对时效性商品等时效性数据对象的识别和预测显得尤其重要。

发明内容

鉴于上述问题，提出了本申请实施例以便提供一种克服上述问题或者至少部分地解决上述问题的一种基于时间序列的数据预测方法和相应的一种基于时间序列的数据预测装置。

25

为了解决上述问题，本申请公开了一种基于时间序列的数据预测方法，所述的方法包括：

获取多个类目对象的历史时间序列数据，其中，所述类目对象包括一个或多个数据

30 对象；

从所述多个类目对象中筛选出特征类目对象，其中，所述特征类目对象为包含特征数据对象的类目对象，所述特征数据对象为生命周期小于预设时间阈值的数据对象；

基于所述特征类目对象对应的历史时间序列数据，从所述特征类目对象包含的数据对象中预测出目标数据对象，所述目标数据对象为未来第一预设时间段内将要产生的未来时间序列数据满足预设增长趋势的数据对象。

优选地，所述方法还包括：

预测所述目标数据对象在所述未来第一预设时间段内的未来时间序列数据。

优选地，所述获取多个类目对象的历史时间序列数据的步骤包括：

针对预设的多个时间区间，计算每个时间区间内预设数据库中存储的，所述数据对象对应的指定特征数据的数量，作为所述数据对象在所述时间区间内的历史特征数据；

组织所述数据对象在所有时间区间的历史特征数据，得到所述数据对象的历史时间序列数据；

按照所述时间区间，统计每个类目对象中包含的数据对象在所述时间区间的历史特征数据的总和；

将所有时间区间的历史特征数据的总和组织成所述类目对象的历史时间序列数据。

优选地，所述从所述多个类目对象中筛选出特征类目对象的步骤包括：

基于所述类目对象的历史时间序列数据，从所述多个类目对象中筛选出第一特征类目对象；

获取预设的第二特征类目对象；

将所述第一特征类目对象以及所述第二特征类目对象组织成特征类目对象。

优选地，所述基于所述类目对象的历史时间序列数据，从所述多个类目对象中筛选出第一特征类目对象的步骤包括：

计算在过去第一预设时间段内每个类目对象的历史时间序列数据的中值 M ；

计算历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量；

若所述历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量在预设范围内，则判定所述类目对象为第一特征类目对象。

优选地，所述基于所述特征类目对象对应的历史时间序列数据，从所述特征类目对象包含的数据对象中预测出目标数据对象的步骤包括：

基于所述特征类目对象对应的历史时间序列数据，对所述特征类目对象进行归一化处理；

将所有归一化处理后的特征类目对象中包含的数据对象进行聚类，得到类簇对象；
从所述类簇对象中预测出目标类簇对象；

将所述目标类簇对象中包含的数据对象，作为目标数据对象。

优选地，所述从所述类簇对象中预测出目标类簇对象的步骤包括：

- 5 基于所述类簇对象中的数据对象在过去一个月内的历史时间序列数据，计算所述类簇对象的第一平均历史时间序列数据；

基于所述类簇对象中的数据对象在过去第十三个月的历史时间序列数据，计算所述类簇对象的第二平均历史时间序列数据；

- 10 基于所述类簇对象中的数据对象在过去第十二个月的历史时间序列数据，计算所述类簇对象的第三平均历史时间序列数据；

根据所述第一平均历史时间序列数据、所述第二平均历史时间序列数据以及所述第三平均历史时间序列数据，预估所述类簇对象在未来第一预设时间段内的未来平均时间序列数据；

- 15 计算所述未来平均时间序列数据与所述第一平均历史时间序列数据的差值，得到所述类簇对象的指标数据；

将指标数据大于预设阈值的类簇对象作为目标类簇对象。

优选地，所述预测所述目标数据对象在所述未来第一预设时间段内的未来时间序列数据的步骤包括：

- 20 对所述类簇对象在未来第一预设时间段内的未来平均时间序列数据进行反归一化处理，得到所述类簇对象中每个数据对象的基准平均时间序列数据；

对所述每个数据对象的基准平均时间序列数据进行修正，得到对应数据对象在未来第一预设时间段内的未来时间序列数据。

- 25 优选地，所述数据对象为商品数据，所述类目对象为商品类目，所述特征类目对象为时效性商品类目，所述生命周期为商品的时效，所述时间序列数据为所述商品的日销量。

本申请还公开了一种基于时间序列的数据预测装置，所述的装置包括：

历史时序数据获取模块，用于获取多个类目对象的历史时间序列数据，其中，所述类目对象包括一个或多个数据对象；

- 30 特征类目对象筛选模块，用于从所述多个类目对象中筛选出特征类目对象，其中，所述特征类目对象为包含特征数据对象的类目对象，所述特征数据对象为生命周期小于

预设时间阈值的数据对象；

目标数据对象预测模块，用于基于所述特征类目对象对应的历史时间序列数据，从所述特征类目对象包含的数据对象中预测出目标数据对象，所述目标数据对象为未来第一预设时间段内将要产生的未来时间序列数据满足预设增长趋势的数据对象。

5 优选地，所述装置还包括：

未来时序数据预测模块，用于预测所述目标数据对象在所述未来第一预设时间段内的未来时间序列数据。

优选地，所述历史时序数据获取模块包括：

10 历史特征数据计算子模块，用于针对预设的多个时间区间，计算每个时间区间内预设数据库中存储的，所述数据对象对应的指定特征数据的数量，作为所述数据对象在所述时间区间内的历史特征数据；

历史特征数据组织子模块，用于组织所述数据对象在所有时间区间的历史特征数据，得到所述数据对象的历史时间序列数据；

15 历史特征数据统计子模块，用于按照所述时间区间，统计每个类目对象中包含的数据对象在所述时间区间的历史特征数据的总和；

历史时序数据组织子模块，用于将所有时间区间的历史特征数据的总和和组织成所述类目对象的历史时间序列数据。

优选地，所述特征类目对象筛选模块包括：

20 第一特征类目对象筛选子模块，用于基于所述类目对象的历史时间序列数据，从所述多个类目对象中筛选出第一特征类目对象；

第二特征类目对象获取子模块，用于获取预设的第二特征类目对象；

组织子模块，用于将所述第一特征类目对象以及所述第二特征类目对象组织成特征类目对象。

优选地，所述第一特征类目对象筛选子模块还用于：

25 计算在过去第一预设时间段内每个类目对象的历史时间序列数据的中值 M ；

计算历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量；

若所述历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量在预设范围内，则判定所述类目对象为第一特征类目对象。

优选地，所述目标数据对象预测模块包括：

30 归一化子模块，用于基于所述特征类目对象对应的历史时间序列数据，对所述特征

类目对象进行归一化处理；

聚类子模块，用于将所有归一化处理后的特征类目对象中包含的数据对象进行聚类，得到类簇对象；

预测子模块，用于从所述类簇对象中预测出目标类簇对象；

5 目标数据对象获取子模块，用于将所述目标类簇对象中包含的数据对象，作为目标数据对象。

优选地，所述预测子模块还用于：

基于所述类簇对象中的数据对象在过去一个月内的历史时间序列数据，计算所述类簇对象的第一平均历史时间序列数据；

10 基于所述类簇对象中的数据对象在过去第十三个月的历史时间序列数据，计算所述类簇对象的第二平均历史时间序列数据；

基于所述类簇对象中的数据对象在过去第十二个月的历史时间序列数据，计算所述类簇对象的第三平均历史时间序列数据；

15 根据所述第一平均历史时间序列数据、所述第二平均历史时间序列数据以及所述第三平均历史时间序列数据，预估所述类簇对象在未来第一预设时间段内的未来平均时间序列数据；

计算所述未来平均时间序列数据与所述第一平均历史时间序列数据的差值，得到所述类簇对象的指标数据；

将指标数据大于预设阈值的类簇对象作为目标类簇对象。

20 优选地，所述未来时序数据预测模块包括：

基准数据获取子模块，用于对所述类簇对象在未来第一预设时间段内的未来平均时间序列数据进行反归一化处理，得到所述类簇对象中每个数据对象的基准平均时间序列数据；

25 修正子模块，用于对所述每个数据对象的基准平均时间序列数据进行修正，得到对应数据对象在未来第一预设时间段内的未来时间序列数据。

优选地，所述数据对象为商品数据，所述类目对象为商品类目，所述特征类目对象为时效性商品类目，所述生命周期为商品的时效，所述时间序列数据为所述商品的日销量。

本申请实施例包括以下优点：

30 在本申请实施例中，能够从多个类目对象中筛选出具有时效特性以及季节特性的特

征类目对象，并基于该特征类目对象的历史时间序列数据，从特征类目对象包含的数据对象中预测出近期将要产生的未来时间序列数据满足预设增长趋势的数据对象，即近期将要爆发的目标数据对象，本申请实施例根据时间序列数据的原理，预测出近期具有爆发力的目标数据对象，使得预测结果与实际更加吻合，准确率更高。

5

附图说明

图 1 是本申请的一种基于时间序列的数据预测方法实施例一的步骤流程图；

图 2 是本申请的一种基于时间序列的数据预测方法实施例一中的类目树示意图；

10 图 3 是本申请的一种基于时间序列的数据预测方法实施例二的步骤流程图；

图 4 是本申请的一种基于时间序列的数据预测装置实施例的结构框图。

具体实施方式

15 为使本申请的上述目的、特征和优点能够更加明显易懂，下面结合附图和具体实施方式对本申请作进一步详细的说明。

参照图 1，示出了本申请的一种基于时间序列的数据预测方法实施例一的步骤流程图，本申请实施例可以应用于电商平台等具有树形类目体系的平台中，树形类目体系可以为按照树状分类法对数据进行分类，得到类目的方法，其中，树状分类法是一种形象
20 的分类法，按照层次，一层一层来分，就像一棵大树，有叶、枝、杆、根。

例如，在电商平台中，为适应当今时代的消费人群在网上商店有针对性的选购各种各样的商品，可以采用树状分类法对商品做出的归类，得到商品类目，例如，服装、配饰、美容、数码、家居、母婴、食品、文体、服务和保险等。

如图 1 所示，本申请实施例可以包括如下步骤：

25 步骤 101，获取多个类目对象的历史时间序列数据；

应用于本申请实施例，一个类目对象可以包括一个或多个数据对象，例如，在电商平台中，如图 2 的类目树示意图所示，在商品类目如“海鲜”类目下，可以包括“大闸蟹”、“八爪鱼”、“瑶柱”等商品数据。

进一步地，每个数据对象具有对应的多个指定特征数据，所述指定特征数据为在先生成的，检测到对所述数据对象发生指定行为时生成的记录。例如，在电商平台中，所
30

述指定行为可以包括销售行为，所述指定特征数据为对某个商品产生销售行为时生成的销售记录。

在具体实现中，数据对象的指定特征数据可以从预设数据库中获取，该预设数据库可以为预先生成的数据库。例如，该预设数据库可以为商品数据库，该商品数据库中存

5 储有多条针对一个或多个商品的销售记录。

在实际中，预设数据库中还可以存储数据对象的数据属性信息，作为一种示例，该数据属性信息可以包括时间属性信息、标识属性信息、特征属性信息等。例如，在商品数据库中，还可以存储每个商品的商品属性信息，该商品属性信息可以包括商品的基本属性、时间属性、交易属性、信用属性及营销属性等。其中，该商品的基本属性可以包

10 括商品的名称、所属商家 ID、价格、上架时长、所属类目等；时间属性可以包括发生购买行为、评论行为、上架行为等行为的时间信息；该商品的交易属性可以包括商品收藏、加购、购买等；该商品的信用属性可以包括商家星级、差评数、差评率、物流评分等；该商品的营销属性可以包括是否为抢购商品、是否为促销商品等。

在本申请实施例的一种优选实施例中，步骤 101 可以包括如下子步骤：

15 子步骤 S11，针对预设的多个时间区间，计算每个时间区间内预设数据库中存储的，所述数据对象对应的指定特征数据的数量，作为所述数据对象在所述时间区间内的历史特征数据；

在具体实现中，时间区间可以为按照时间间隔设定的区间，例如，该时间间隔可以为一天、半天、一周、一个月等，若时间间隔为一天，则时间区间可以为每天的[00:00，

20 23:59]，当然该时间区间还可以添加日期信息，例如 2015 年 11 月 18 日的时间区间为 [2015-11-18-00:00，2015-11-18-23:59]。该预设的时间区间可以为开发人员预先设定的时间区间。

获得多个预设的时间区间以后，可以进一步计算该数据对象在每个时间区间内（例如每天）的指定特征数据的数量，得到该时间区间的历史特征数据。例如，计算某一商

25 品每一天的销售记录的数量，得到日销量。

子步骤 S12，组织所述数据对象在所有时间区间的历史特征数据，得到所述数据对象的历史时间序列数据；

得到数据对象在每个时间区间中的历史特征数据以后，组织所有时间区间的历史特征数据，可以得到该数据对象的历史时间序列数据。其中，时间序列数据是指不同时间

30 点上收集到的数据，这类数据反映了某一事物、现象等随时间的变化状态或程度。时间

序列数据是数据存在的特殊形式，序列的过去值会影响到将来值，这种影响的大小以及影响的方式可由时间序列数据中的趋势周期及非平稳等行为来刻画。时间序列挖掘其本质是根据数据随时间变化的趋势预测将来的值。重点要考虑的是时间的特殊性质，像一些周期性的时间定义如星期、月、季节、年等，不同的日子如节假日可能造成的影响，日期本身的计算方法，还有一些需要特殊考虑的地方如时间前后的相关性(过去的事情对将来有多大的影响力)等。只有充分考虑时间因素，利用现有数据随时间变化的一系列的值，才能更好地预测将来的值。

例如，得到一个商品的日销量以后，组织每天的日销量，得到该商品的历史销量。

一个数据对象的历史时间序列数据可以反映该数据对象在过去某个时间段中的走势。

子步骤 S13，按照所述时间区间，统计每个类目对象中包含的数据对象在所述时间区间的历史特征数据的总和；

由于一个类目对象可以包括一个或多个数据对象，当得到该类目对象下每个数据对象的历史特征数据以后，可以以时间区间为单位，计算该类目对象下所有数据对象在该时间区间的历史特征数据总和。

例如，在某一天中，在“海鲜”类目下，“大闸蟹”的日销量为 1000 斤、“八爪鱼”的日销量为 500 斤、“瑶柱”的日销量为 300 斤，则该“海鲜”类目下在该日期中日销量总和为 1800 斤。

子步骤 S14，将所有时间区间的历史特征数据的总和组织成所述类目对象的历史时间序列数据。

组织所有时间区间的历史特征数据的总和，可以得到该类目对象的历史时间序列数据。

例如，计算“海鲜”类目在近一个月内每天的日销量总和以后，将该一个月的所有天数的日销量总和组织起来，可以得到“海鲜”类目在该月的历史时间序列数据。

一个类目对象的历史时间序列数据可以反映该类目对象在过去某个时间段中的走势。

在具体实现中，步骤 101 可以通过一类目数据生成器完成，该生成器根据当前平台的树形类目体系，生成各类目对象的历史时间序列数据，经过步骤 101 以后，原来海量的数据对象的历史时间序列数据可以归并为各个类目对象的历史时间序列数据，为后续操作提供了有力的数据支撑。

步骤 102, 从所述多个类目对象中筛选出特征类目对象;

在本申请实施例中, 当获得每个类目对象的历史时间序列数据以后, 可以进一步从多个类目对象中筛选出特征类目对象, 其中, 特征类目对象可以为包含特征数据对象的类目对象, 而特征数据对象可以为生命周期小于预设时间阈值的数据对象, 即具有时效性的数据对象。例如, 当类目对象为商品类目时, 该特征类目对象可以为时效性商品类目, 时效性商品类目可以为具有时效性商品的类目对象, 时效性商品是指具有一定消费时效特性, 且保质期非常短暂的商品, 例如: 月饼、大闸蟹等, 而时效性商品类目可以包括蔬菜、水果、海鲜、生肉、熟食等生鲜类目。

在本申请实施例的一种优选实施例中, 步骤 102 可以包括如下子步骤:

子步骤 S21, 基于所述类目对象的历史时间序列数据, 从所述多个类目对象中筛选出第一特征类目对象;

获得当前平台的所有类目对象的历史时间序列数据以后, 可以进一步基于类目对象的历史时间序列数据, 从多个类目对象中自动筛选出第一特征类目对象。

在本申请实施例的一种优选实施例中, 子步骤 S21 进一步可以包括如下子步骤:

子步骤 S211, 计算在过去第一预设时间段内每个类目对象的历史时间序列数据的中值 M ;

具体来说, 中值也称中位数, 是在一组数据中居于中间的数(特别注意的地方是: 这组数据之前已经经过升序或者降序排列), 即在这组数据中, 有一半的数据比它大, 有一半的数据比它小。如果这组数据包含偶数个数字, 中值是位于中间的两个数的平均值, 假如有 n 个数据, 当 n 为偶数时, 中位数为第 $n/2$ 位数和第 $(n+2)/2$ 位数的平均数; 如果 n 为奇数, 那么中位数为第 $(n+1)/2$ 位数的值。

在具体实现中, 可以将每个类目对象的历史时间序列数据的时间范围定义为过去第一预设时间段, 例如, 可以将过去第一时间段设定为过去一年。针对每个类目对象, 可以将其历史时间序列数据按照升序或降序排序, 即将过去一年内该类目对象中所有时间区间对应的历史时间序列数据的总和进行排序, 排序后获得该类目对象的中值 M , 如将过去一年中每个商品类目的每天的日销量总和进行排序后, 获得排序在中间的日销量总和作为该商品类目在过去一年中的中值 M 。

需要说明的是, 此处计算中值而不是计算平均值, 是由于在一组数据中, 平均值易受极端值的影响, 而中值则不会受到极端值的影响, 从而作出与实际情况更吻合的预测。

子步骤 S212, 计算历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量;

得到中值 M 以后, 可以将 M 放大 n 倍, 例如 1.5 倍 (可以表示为 $1.5M$), 并将该类目对象在每个时间区间的历史特征数据的总和与 $1.5M$ 比较, 获得历史特征数据的总和大于 $1.5M$ 的时间区间的数量。例如, 计算商品类目中日销量总和大于 $1.5M$ 的天数。

子步骤 S213, 若所述历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量在预设范围内, 则判定所述类目对象为第一特征类目对象。

若 M 放大 1.5 倍时, 该类目对象的历史特征数据的总和大于 $1.5M$ 的时间区间的数量在预设范围内时, 可以判定该类目对象为第一特征类目对象。

例如, 将预设范围取值为 10-45, 若商品类目中日销量总和大于 $1.5M$ 的天数在这个范围内, 则可以判定该商品类目为时效性商品类目。

10 子步骤 S22, 获取预设的第二特征类目对象;

应用于本申请实施例, 预设的第二特征类目对象可以为白名单中的类目对象, 该白名单可以通过人工的方式预先选定, 例如, 时效性商品类目可以为运营预先选出的商品类目, 并将该选出的商品类目加入白名单中。

15 子步骤 S23, 将所述第一特征类目对象以及所述第二特征类目对象组织成特征类目对象。

得到第一特征类目对象以及第二特征类目对象以后, 可以将第一特征类目对象以及第二特征类目对象组织成特征类目对象, 其中, 组织的方式可以包括去重方式, 即将第一特征类目对象以及第二特征类目对象中重复的特征类目对象去除, 最后输出所有的特征类目对象。

20 在本申请实施例中, 可以通过自动和人工的方式进行特征类目对象的筛选, 使得筛选结果更加符合用户需求, 也更加完善, 智能化程度高。

步骤 103, 基于所述特征类目对象对应的历史时间序列数据, 从所述特征类目对象包含的数据对象中预测出目标数据对象。

25 确定特征类目对象以后, 可以从特征类目对象包含的数据对象中筛选出目标数据对象, 其中, 该目标数据对象为未来第一预设时间段内将要产生的未来时间序列数据满足预设增长趋势的数据对象, 即近期将要产生数量爆发的数据对象。

在具体实现中, 为了提高预测结果的可靠性, 未来第一预设时间段可以为近期的一个时间段, 例如可以包括未来的一中期时间段或者一短期时间段。作为一种示例, 该中期时间段可以为一个月的时间, 即未来第一预设时间段为从当前时间开始接下来的一个月时间; 该短期时间段可以为半个月、一周等短期内的时间, 即未来第一预设时间段为

从当前时间开始接下来的半个月时间或一周时间等。

该目标数据对象可以为未来第一预设时间段内，将要产生的未来时间序列数据满足预设增长趋势的数据对象，即产生的数量具有异常点或爆发点的数据对象。例如，在中秋节前，月饼的销售数量将会爆发性增长，则月饼可以为目标数据对象。

- 5 应用于本申请实施例，确定特征类目对象以后，可以从特征类目对象中包含的数据对象中进一步筛选出目标数据对象。例如，确定时效性商品类目以后，可以进一步从该时效性商品类目中包含的时效性商品中筛选出近期将会热卖（产生爆发点或异常点）的目标时效性商品。

在本申请实施例的一种优选实施例中，步骤 103 可以包括如下子步骤：

- 10 子步骤 S31，基于所述特征类目对象对应的历史时间序列数据，对所述特征类目对象进行归一化处理；

确定特征类目对象以后，为了消除特征类目对象中各个数据对象之间的差别，得到更准确的预测结果，可以对该特征类目对象进行归一化处理。其中，归一化是一种简化计算的方式，即将有量纲的表达式，经过变换，化为无量纲的表达式，成为标量。

- 15 在一种实施方式中，可以采用如下方式对特征类目对象进行归一化处理：

根据上述子步骤 S211 获得的在过去第一预设时间段内特征类目对象的历史时间序列数据的中值 M ；分别计算该历史时间序列数据中的每个历史特征数据的总和与中值 M 的比值，得到归一化后的历史特征数据的总和，将所有归一化后的历史特征数据的总和组织成该特征类目对象的归一化的历史时间序列数据。

- 20 当然，本申请实施例并不限于上述归一化的方式，本领域技术人员采用其他归一化的方式均是可行的。

子步骤 S32，将所有归一化处理后的特征类目对象中包含的数据对象进行聚类，得到类簇对象；

- 25 应用于本申请实施例，将特征类目对象的历史时间序列数据进行归一化处理以后，进一步可以将所有特征类目对象进行聚类，在实际中，该聚类可以为将所有特征类目对象中包含的所有的数据对象进行聚类，将历史时间序列数据具有类似趋势的数据对象（例如，具有类似爆发力的数据对象）聚合在一起，得到一个或多个类簇对象。

- 30 具体的，将物理或抽象对象的集合分成由类似的对象组成的多个类的过程被称为聚类，由聚类所生成的类簇是一组对象的集合，这些对象与同一个簇中的对象彼此相似，与其他簇中的对象相异。在具体实现中，可以采用多种聚类方式进行聚类，例如层次聚

类、划分聚类、基于密度的聚类、基于网格的聚类、基于模型的聚类等，本申请实施例对具体的聚类方法不作限制。

例如，得到的特征类目对象为水果类目、海鲜类目、熟食类目等，可以将这三个类目对象分别进行归一化处理，并将归一化处理后的类目对象中包含的商品进行聚类，把有类似爆发力的商品聚合在一起，得到一个或多个类簇，例如，大闸蟹由于在中秋期间到了多膏美味的顶峰，可以与月饼一起在中秋时节期间同时迎来爆发高峰，两者的历史时间序列数据的走势类似，则可以将大闸蟹与月饼放入同一类簇中。

子步骤 S33，从所述类簇对象中预测出目标类簇对象；

得到类簇对象以后，可以从该类簇对象中筛选出近期（未来第一预设时间段内）将要爆发的类簇对象，作为目标类簇对象。例如，从多个类簇对象中筛选出将要热卖的类簇对象作为目标类簇对象。

在本申请实施例的一种优选实施例中，子步骤 S33 进一步可以包括如下子步骤：

子步骤 S331，基于所述类簇对象中的数据对象在过去一个月内的历史时间序列数据，计算所述类簇对象的第一平均历史时间序列数据；

在具体实现中，可以根据类簇对象中每个数据对象的在过去一个月内（最近一个月）的归一化后的历史时间序列数据，计算该类簇下所有数据对象的历史时间序列数据的平均值，即以时间区间为单位（例如以天为单位），计算该类簇下该时间区间所有数据对象的归一化后的历史特征数据之和除以该时间区间下所有数据对象的数量，得到该时间区间下的平均值；所有时间区间的平均值组成该类簇的第一平均历史时间序列数据。

子步骤 S332，基于所述类簇对象中的数据对象在过去第十三个月的历史时间序列数据，计算所述类簇对象的第二平均历史时间序列数据；

在具体实现中，可以根据类簇对象中每个数据对象的在过去第十三个月（最近一个月对应的去年的日期）的归一化后的历史时间序列数据，计算该类簇下所有数据对象的历史时间序列数据的平均值，即以时间区间为单位（例如以天为单位），计算该类簇下该时间区间所有数据对象的归一化后的历史特征数据之和除以该时间区间下所有数据对象的数量，得到该时间区间下的平均值；所有时间区间的平均值组成所述类簇的第二平均历史时间序列数据。

子步骤 S333，基于所述类簇对象中的目标数据对象在过去第十二个月的历史时间序列数据，计算所述类簇对象的第三平均历史时间序列数据；

采用与上述子步骤 S332 的方法，计算类簇对象的第三平均历史时间序列数据，即计

算去年当前日期的平均归一化数据。

子步骤 S334, 根据所述第一平均历史时间序列数据、所述第二平均历史时间序列数据以及所述第三平均历史时间序列数据, 预估所述类簇对象在未来第一预设时间段内的未来平均时间序列数据;

5 在具体实现中, 得到第一平均历史时间序列数据以后, 可以进一步计算该第一平均历史时间序列数据的第一平均值(类簇的每个时间区间下的平均值之和除以时间区间的数量), 以及, 得到第二平均历史时间序列数据以后, 可以进一步计算该第二平均历史时间序列数据的第二平均值(类簇的每个时间区间下的平均值之和除以时间区间的数量)。

10 然后计算第一平均值与第二平均值的比值, 得到比值 A。

然后将第三平均历史时间序列数据分别乘以比值 A, 得到所述特征类目对象在未来第一预设时间段内的未来平均时间序列数据。

需要说明的是, 该未来第一预设时间段可以为农历基准的时间段, 若在该第一预设时间段内若某个时间区间内出现重大公历节日(如国庆, 元旦等), 则进行公历日假期的相应修正, 即在该节假日中, 将农历基准变成对应的公历基准, 其他非重大公历节日不变。

子步骤 S335, 计算所述未来平均时间序列数据与所述第一平均历史时间序列数据的差值, 得到所述类簇对象的指标数据;

20 得到未来第一预设时间段内的未来平均时间序列数据以后, 可以进一步计算所述未来平均时间序列数据的第一总和(每个时间区间下类簇的平均值之和), 以及, 所述第一平均历史时间序列数据的第二总和。

然后计算第一总和所述第二总和的差值, 可以得到该类簇对象的指标数据。

子步骤 S336, 将指标数据大于预设阈值的类簇对象作为目标类簇对象。

25 获得类簇对象的指标数据以后, 可以筛选出指标数据较大的类簇对象作为目标类簇对象, 在一种实施方式中, 可以筛选出指标数据大于预设阈值的类簇对象作为目标类簇对象。

例如, 得到的两个类簇的指标数据分别如下(M 为归一化前的历史序列数据的中值):

大闸蟹+月饼(第一类簇): $1.1M$

八爪鱼(第二类簇): $-0.01M$

30 经过排序之后, 可以很容易判定未来半个月之内第一类簇, 即大闸蟹和月饼的销量

将会爆发，而八爪鱼则会趋于平稳。

在本申请实施例中，可以根据类簇的爆发力指标数据判定其短期和中期爆发的可能性。

子步骤 S34，将所述目标类簇中包含的数据对象，作为目标数据对象。

5 确定目标类簇对象以后，可以将该目标类簇对象中包含的数据对象，作为目标数据对象。

在本申请实施例中，能够从多个类目对象中筛选出具有时效特性以及季节特性的特征类目对象，并基于该特征类目对象的历史时间序列数据，从特征类目对象包含的数据对象中预测出近期将要爆发的目标数据对象，本申请实施例根据时间序列数据的原理，

10 预测出近期具有爆发力的目标数据对象，使得预测结果与实际更加吻合，准确率更高。

参照图 3，示出了本申请的一种基于时间序列的数据预测方法实施例二的步骤流程图，可以包括如下步骤：

步骤 301，获取多个类目对象的历史时间序列数据；

15 应用于本申请实施例，一个类目对象可以包括一个或多个数据对象。

在本申请实施例的一种优选实施例中，步骤 301 可以包括如下子步骤：

子步骤 S41，针对预设的多个时间区间，计算每个时间区间内预设数据库中存储的，所述数据对象对应的指定特征数据的数量，作为所述数据对象在所述时间区间内的历史特征数据；

20 子步骤 S42，组织所述数据对象在所有时间区间的历史特征数据，得到所述数据对象的历史时间序列数据；

子步骤 S43，按照所述时间区间，统计每个类目对象中包含的数据对象在所述时间区间的历史特征数据的总和；

25 子步骤 S44，将所有时间区间的历史特征数据的总和组织成所述类目对象的历史时间序列数据。

步骤 302，从所述多个类目对象中筛选出特征类目对象；

在本申请实施例中，当获得每个类目对象的历史时间序列数据以后，可以进一步从多个类目对象中筛选出特征类目对象，其中，特征类目对象可以为包含特征数据对象的类目对象，而特征数据对象可以为生命周期小于预设时间阈值的数据对象，即具有时效性的数据对象。

30

在本申请实施例的一种优选实施例中，步骤 302 可以包括如下子步骤：

子步骤 S51，基于所述类目对象的历史时间序列数据，从所述多个类目对象中筛选出第一特征类目对象；

在本申请实施例的一种优选实施例中，子步骤 S51 进一步可以包括如下子步骤：

5 子步骤 S511，计算在过去第一预设时间段内每个类目对象的历史时间序列数据的中间值 M ；

子步骤 S512，计算历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量；

子步骤 S513，若所述历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量在预设范围内，则判定所述类目对象为第一特征类目对象。

10 子步骤 S52，获取预设的第二特征类目对象；

子步骤 S53，将所述第一特征类目对象以及所述第二特征类目对象组织成特征类目对象。

步骤 303，基于所述特征类目对象对应的历史时间序列数据，从所述特征类目对象包含的数据对象中预测出目标数据对象；

15 确定特征类目对象以后，可以从特征类目对象包含的数据对象中筛选出目标数据对象，其中，该目标数据对象可以为未来第一预设时间段内将要产生的未来时间序列数据满足预设增长趋势的数据对象。

在本申请实施例的一种优选实施例中，步骤 303 可以包括如下子步骤：

20 子步骤 S61，基于所述特征类目对象对应的历史时间序列数据，对所述特征类目对象进行归一化处理；

子步骤 S62，将所有归一化处理后的特征类目对象中包含的数据对象进行聚类，得到类簇对象；

子步骤 S63，从所述类簇对象中预测出目标类簇对象；

在本申请实施例的一种优选实施例中，子步骤 S63 进一步可以包括如下子步骤：

25 子步骤 S631，基于所述类簇对象中的数据对象在过去一个月内的历史时间序列数据，计算所述类簇对象的第一平均历史时间序列数据；

子步骤 S632，基于所述类簇对象中的数据对象在过去第十三个月的历史时间序列数据，计算所述类簇对象的第二平均历史时间序列数据；

30 子步骤 S633，基于所述类簇对象中的数据对象在过去第十二个月的历史时间序列数据，计算所述类簇对象的第三平均历史时间序列数据；

子步骤 S634, 根据所述第一平均历史时间序列数据、所述第二平均历史时间序列数据以及所述第三平均历史时间序列数据, 预估所述类簇对象在未来第一预设时间段内的未来平均时间序列数据;

子步骤 S635, 计算所述未来平均时间序列数据与所述第一平均历史时间序列数据的差值, 得到所述类簇对象的指标数据;

子步骤 S636, 将指标数据大于预设阈值的类簇对象作为目标类簇对象。

子步骤 S64, 将所述目标类簇对象中包含的数据对象, 作为目标数据对象。

步骤 304, 预测所述目标数据对象在所述未来第一预设时间段内的未来时间序列数据。

10 在本申请实施例的一种优选实施例中, 步骤 304 可以包括如下子步骤:

子步骤 S71, 对所述类簇对象在未来第一预设时间段内的未来平均时间序列数据进行反归一化处理, 得到所述类簇对象中每个数据对象的基准平均时间序列数据;

由于根据子步骤 S634 预估的所述类簇对象的未来平均时间序列数据是一种归一化后的值, 因此可以首先对该归一化后的值进行反归一化处理, 即将该未来平均时间序列数据乘以中值 M , 可以得到该类簇对象中每个数据对象的基准平均时间序列数据。

子步骤 S72, 对所述每个数据对象的基准平均时间序列数据进行修正, 得到对应数据对象在未来第一预设时间段内的未来时间序列数据。

获得每个数据对象的基准平均时间序列数据以后, 可以对该基准平均时间序列数据进行修正, 得到该数据对象在未来第一预设时间段内的未来时间序列数据。在一种实施方式中, 所述修正可以包括依据预设参考参数进行放大或缩小的补偿修正。

预设参考参数可以为其他数据库中的补偿参数, 例如, 在电商平台中, 为了对抗平台商家数量变化带来的影响, 该预设参考参数可以为商家数据库中的数据, 该商家数据库记录了平台的各个商家及其主要的特征, 包括商家的基本属性、交易属性及信用属性等特征。可以以当前商家数和去年对应时期商家数相比进行基准平均时间序列数据的放大(或缩小)等修正, 得到该商品类目的未来时间序列数据。

例如, 去年与今年同期相比, 商家数据库中保存的商家数量从 100 家增加到 1000 家, 商家数量增加了 10 倍, 而销量增加了 20 倍, 则可以将基准平均时间序列数据放大两倍, 得到未来时间序列数据。

作为本申请实施例的一种优选示例, 若将本申请实施例应用于电商平台中, 则所述数据对象可以为商品数据, 所述类目对象可以为商品类目, 所述特征类目对象可以为时

效性商品类目，所述生命周期可以为商品的时效，所述时间序列数据可以为所述商品的日销量。

在本申请实施例中，能够从多个类目对象中筛选出具有时效特性以及季节特性的特征类目对象，并基于该特征类目对象的历史时间序列数据，从特征类目对象包含的数据对象中预测出近期将要爆发的目标数据对象，并预测该目标数据对象近期的未来时间序列数据，本申请实施例根据时间序列数据的原理，预测出近期具有爆发力的目标数据对象以及该目标数据对象的未来时间序列数据，使得预测结果与实际更加吻合，准确率更高。

对于图 3 的方法实施例而言，由于其与图 1 的方法实施例基本相似，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

需要说明的是，对于方法实施例，为了简单描述，故将其都表述为一系列的动作组合，但是本领域技术人员应该知悉，本申请实施例并不受所描述的动作顺序的限制，因为依据本申请实施例，某些步骤可以采用其他顺序或者同时进行。其次，本领域技术人员也应该知悉，说明书中所描述的实施例均属于优选实施例，所涉及的动作并不一定是本申请实施例所必须的。

参照图 4，示出了本申请的一种基于时间序列的数据预测装置实施例的结构框图，具体可以包括如下模块：

历史时序数据获取模块 401，用于获取多个类目对象的历史时间序列数据，其中，所述类目对象包括一个或多个数据对象；

特征类目对象筛选模块 402，用于从所述多个类目对象中筛选出特征类目对象，其中，所述特征类目对象为包含特征数据对象的类目对象，所述特征数据对象为生命周期小于预设时间阈值的数据对象；

目标数据对象预测模块 403，用于基于所述特征类目对象对应的历史时间序列数据，从所述特征类目对象包含的数据对象中预测出目标数据对象，所述目标数据对象为未来第一预设时间段内将要产生的未来时间序列数据满足预设增长趋势的数据对象。

在本申请实施例的一种优选实施例中，所述装置还可以包括：

未来时序数据预测模块，用于预测所述目标数据对象在所述未来第一预设时间段内的未来时间序列数据。

在本申请实施例的一种优选实施例中，所述历史时序数据获取模块 401 包括：

历史特征数据计算子模块，用于针对预设的多个时间区间，计算每个时间区间内预

设数据库中存储的，所述数据对象对应的指定特征数据的数量，作为所述数据对象在所述时间区间内的历史特征数据；

历史特征数据组织子模块，用于组织所述数据对象在所有时间区间的历史特征数据，得到所述数据对象的历史时间序列数据；

- 5 历史特征数据统计子模块，用于按照所述时间区间，统计每个类目对象中包含的数据对象在所述时间区间的历史特征数据的总和；

历史时序数据组织子模块，用于将所有时间区间的历史特征数据的总和组织成所述类目对象的历史时间序列数据。

在本申请实施例的一种优选实施例中，所述特征类目对象筛选模块 402 包括：

- 10 第一特征类目对象筛选子模块，用于基于所述类目对象的历史时间序列数据，从所述多个类目对象中筛选出第一特征类目对象；

第二特征类目对象获取子模块，用于获取预设的第二特征类目对象；

组织子模块，用于将所述第一特征类目对象以及所述第二特征类目对象组织成特征类目对象。

- 15 在本申请实施例的一种优选实施例中，所述第一特征类目对象筛选子模块还用于：
计算在过去第一预设时间段内每个类目对象的历史时间序列数据的中值 M ；

计算历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量；

若所述历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量在预设范围内，则判定所述类目对象为第一特征类目对象。

- 20 在本申请实施例的一种优选实施例中，所述目标数据对象预测模块 403 包括：

归一化子模块，用于基于所述特征类目对象对应的历史时间序列数据，对所述特征类目对象进行归一化处理；

聚类子模块，用于将所有归一化处理后的特征类目对象中包含的数据对象进行聚类，得到类簇对象；

- 25 预测子模块，用于从所述类簇对象中预测出目标类簇对象；

目标数据对象获取子模块，用于将所述目标类簇对象中包含的数据对象，作为目标数据对象。

在本申请实施例的一种优选实施例中，所述预测子模块还用于：

- 30 基于所述类簇对象中的数据对象在过去一个月内的历史时间序列数据，计算所述类簇对象的第一平均历史时间序列数据；

基于所述类簇对象中的数据对象在过去第十三个月的历史时间序列数据，计算所述类簇对象的第二平均历史时间序列数据；

基于所述类簇对象中的数据对象在过去第十二个月的历史时间序列数据，计算所述类簇对象的第三平均历史时间序列数据；

- 5 根据所述第一平均历史时间序列数据、所述第二平均历史时间序列数据以及所述第三平均历史时间序列数据，预估所述类簇对象在未来第一预设时间段内的未来平均时间序列数据；

计算所述未来平均时间序列数据与所述第一平均历史时间序列数据的差值，得到所述类簇对象的指标数据；

- 10 将指标数据大于预设阈值的类簇对象作为目标类簇对象。

在本申请实施例的一种优选实施例中，所述未来时序数据预测模块包括：

基准数据获取子模块，用于对所述类簇对象在未来第一预设时间段内的未来平均时间序列数据进行反归一化处理，得到所述类簇对象中每个数据对象的基准平均时间序列数据；

- 15 修正子模块，用于对所述每个数据对象的基准平均时间序列数据进行修正，得到对应数据对象在未来第一预设时间段内的未来时间序列数据。

在本申请实施例的一种优选实施例中，所述数据对象为商品数据，所述类目对象为商品类目，所述特征类目对象为时效性商品类目，所述生命周期为商品的时效，所述时间序列数据为所述商品的日销量。

- 20 对于装置实施例而言，由于其与方法实施例基本相似，所以描述的比较简单，相关之处参见方法实施例的部分说明即可。

本说明书中的各个实施例均采用递进的方式描述，每个实施例重点说明的都是与其他实施例的不同之处，各个实施例之间相同相似的部分互相参见即可。

- 25 本领域内的技术人员应明白，本申请实施例的实施例可提供为方法、装置、或计算机程序产品。因此，本申请实施例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，本申请实施例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。

- 30 在一个典型的配置中，所述计算机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。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

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。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非持续性的电脑可读媒体(transitory media), 如调制的数据信号和载波。

本申请实施例是参照根据本申请实施例的方法、终端设备(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终端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终端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。

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终端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。

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终端设备上, 使得在计算机或其他可编程终端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终端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。

尽管已描述了本申请实施例的优选实施例, 但本领域内的技术人员一旦得知了基本创造性概念, 则可对这些实施例做出另外的变更和修改。所以, 所附权利要求意欲解释为包括优选实施例以及落入本申请实施例范围的所有变更和修改。

最后, 还需要说明的是, 在本文中, 诸如第一和第二等之类的关系术语仅仅用来将一个实体或者操作与另一个实体或操作区分开来, 而不一定要求或者暗示这些实体或操

作之间存在任何这种实际的关系或者顺序。而且，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、物品或者终端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、物品或者终端设备所固有的要素。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、物品或者终端设备中还存在另外的相同要素。

以上对本申请所提供的一种基于时间序列的数据预测方法和一种基于时间序列的数据预测装置，进行了详细介绍，本文中应用了具体个例对本申请的原理及实施方式进行了阐述，以上实施例的说明只是用于帮助理解本申请的方法及其核心思想；同时，对于本领域的一般技术人员，依据本申请的思想，在具体实施方式及应用范围上均会有改变之处，综上所述，本说明书内容不应理解为对本申请的限制。

权 利 要 求 书

1、一种基于时间序列的数据预测方法，其特征在于，所述的方法包括：

获取多个类目对象的历史时间序列数据，其中，所述类目对象包括一个或多个数据对象；

5 从所述多个类目对象中筛选出特征类目对象，其中，所述特征类目对象为包含特征数据对象的类目对象，所述特征数据对象为生命周期小于预设时间阈值的数据对象；

 基于所述特征类目对象对应的历史时间序列数据，从所述特征类目对象包含的数据对象中预测出目标数据对象，所述目标数据对象为未来第一预设时间段内将要产生的未来时间序列数据满足预设增长趋势的数据对象。

10 2、根据权利要求 1 所述的方法，其特征在于，还包括：

 预测所述目标数据对象在所述未来第一预设时间段内的未来时间序列数据。

 3、根据权利要求 1 或 2 所述的方法，其特征在于，所述获取多个类目对象的历史时间序列数据的步骤包括：

15 针对预设的多个时间区间，计算每个时间区间内预设数据库中存储的，所述数据对象对应的指定特征数据的数量，作为所述数据对象在所述时间区间内的历史特征数据；

 组织所述数据对象在所有时间区间的历史特征数据，得到所述数据对象的历史时间序列数据；

 按照所述时间区间，统计每个类目对象中包含的数据对象在所述时间区间的历史特征数据的总和；

20 将所有时间区间的历史特征数据的总和组织成所述类目对象的历史时间序列数据。

 4、根据权利要求 3 所述的方法，其特征在于，所述从所述多个类目对象中筛选出特征类目对象的步骤包括：

 基于所述类目对象的历史时间序列数据，从所述多个类目对象中筛选出第一特征类目对象；

25 获取预设的第二特征类目对象；

 将所述第一特征类目对象以及所述第二特征类目对象组织成特征类目对象。

 5、根据权利要求 4 所述的方法，其特征在于，所述基于所述类目对象的历史时间序列数据，从所述多个类目对象中筛选出第一特征类目对象的步骤包括：

 计算在过去第一预设时间段内每个类目对象的历史时间序列数据的中值 **M**；

30 计算历史特征数据的总和大于所述 **M** 的预设倍数的时间区间的数量；

若所述历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量在预设范围内，则判定所述类目对象为第一特征类目对象。

6、根据权利要求 1 或 2 所述的方法，其特征不在于，所述基于所述特征类目对象对应的历史时间序列数据，从所述特征类目对象包含的数据对象中预测出目标数据对象的

5 步骤包括：

基于所述特征类目对象对应的历史时间序列数据，对所述特征类目对象进行归一化处理；

将所有归一化处理后的特征类目对象中包含的数据对象进行聚类，得到类簇对象；

从所述类簇对象中预测出目标类簇对象；

10 将所述目标类簇对象中包含的数据对象，作为目标数据对象。

7、根据权利要求 6 所述的方法，其特征不在于，所述从所述类簇对象中预测出目标类簇对象的步骤包括：

基于所述类簇对象中的数据对象在过去一个月内的历史时间序列数据，计算所述类簇对象的第一平均历史时间序列数据；

15 基于所述类簇对象中的数据对象在过去第十三个月的历史时间序列数据，计算所述类簇对象的第二平均历史时间序列数据；

基于所述类簇对象中的数据对象在过去第十二个月的历史时间序列数据，计算所述类簇对象的第三平均历史时间序列数据；

20 根据所述第一平均历史时间序列数据、所述第二平均历史时间序列数据以及所述第三平均历史时间序列数据，预估所述类簇对象在未来第一预设时间段内的未来平均时间序列数据；

计算所述未来平均时间序列数据与所述第一平均历史时间序列数据的差值，得到所述类簇对象的指标数据；

将指标数据大于预设阈值的类簇对象作为目标类簇对象。

25 8、根据权利要求 7 所述的方法，其特征不在于，所述预测所述目标数据对象在所述未来第一预设时间段内的未来时间序列数据的步骤包括：

对所述类簇对象在未来第一预设时间段内的未来平均时间序列数据进行反归一化处理，得到所述类簇对象中每个数据对象的基准平均时间序列数据；

30 对所述每个数据对象的基准平均时间序列数据进行修正，得到对应数据对象在未来第一预设时间段内的未来时间序列数据。

9、根据权利要求 1 或 2 或 4 或 5 或 7 或 8 所述的方法，其特征在于，所述数据对象为商品数据，所述类目对象为商品类目，所述特征类目对象为时效性商品类目，所述生命周期为商品的时效，所述时间序列数据为所述商品的日销量。

10、一种基于时间序列的数据预测装置，其特征在于，所述的装置包括：

5 历史时序数据获取模块，用于获取多个类目对象的历史时间序列数据，其中，所述类目对象包括一个或多个数据对象；

特征类目对象筛选模块，用于从所述多个类目对象中筛选出特征类目对象，其中，所述特征类目对象为包含特征数据对象的类目对象，所述特征数据对象为生命周期小于预设时间阈值的数据对象；

10 目标数据对象预测模块，用于基于所述特征类目对象对应的历史时间序列数据，从所述特征类目对象包含的数据对象中预测出目标数据对象，所述目标数据对象为未来第一预设时间段内将要产生的未来时间序列数据满足预设增长趋势的数据对象。

11、根据权利要求 10 所述的装置，其特征在于，还包括：

15 未来时序数据预测模块，用于预测所述目标数据对象在所述未来第一预设时间段内的未来时间序列数据。

12、根据权利要求 10 或 11 所述的装置，其特征在于，所述历史时序数据获取模块包括：

20 历史特征数据计算子模块，用于针对预设的多个时间区间，计算每个时间区间内预设数据库中存储的，所述数据对象对应的指定特征数据的数量，作为所述数据对象在所述时间区间内的历史特征数据；

历史特征数据组织子模块，用于组织所述数据对象在所有时间区间的历史特征数据，得到所述数据对象的历史时间序列数据；

历史特征数据统计子模块，用于按照所述时间区间，统计每个类目对象中包含的数据对象在所述时间区间的历史特征数据的总和；

25 历史时序数据组织子模块，用于将所有时间区间的历史特征数据的总和和组织成所述类目对象的历史时间序列数据。

13、根据权利要求 12 所述的装置，其特征在于，所述特征类目对象筛选模块包括：

第一特征类目对象筛选子模块，用于基于所述类目对象的历史时间序列数据，从所述多个类目对象中筛选出第一特征类目对象；

30 第二特征类目对象获取子模块，用于获取预设的第二特征类目对象；

组织子模块，用于将所述第一特征类目对象以及所述第二特征类目对象组织成特征类目对象。

14、根据权利要求 13 所述的装置，其特征在于，所述第一特征类目对象筛选子模块还用于：

5 计算在过去第一预设时间段内每个类目对象的历史时间序列数据的中值 M ；

 计算历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量；

 若所述历史特征数据的总和大于所述 M 的预设倍数的时间区间的数量在预设范围内，则判定所述类目对象为第一特征类目对象。

15、根据权利要求 10 或 11 所述的装置，其特征在于，所述目标数据对象预测模块
10 包括：

 归一化子模块，用于基于所述特征类目对象对应的历史时间序列数据，对所述特征类目对象进行归一化处理；

 聚类子模块，用于将所有归一化处理后的特征类目对象中包含的数据对象进行聚类，得到类簇对象；

15 预测子模块，用于从所述类簇对象中预测出目标类簇对象；

 目标数据对象获取子模块，用于将所述目标类簇对象中包含的数据对象，作为目标数据对象。

16、根据权利要求 15 所述的装置，其特征在于，所述预测子模块还用于：

 基于所述类簇对象中的数据对象在过去一个月内的历史时间序列数据，计算所述类
20 簇对象的第一平均历史时间序列数据；

 基于所述类簇对象中的数据对象在过去第十三个月的历史时间序列数据，计算所述类簇对象的第二平均历史时间序列数据；

 基于所述类簇对象中的数据对象在过去第十二个月的历史时间序列数据，计算所述类簇对象的第三平均历史时间序列数据；

25 根据所述第一平均历史时间序列数据、所述第二平均历史时间序列数据以及所述第三平均历史时间序列数据，预估所述类簇对象在未来第一预设时间段内的未来平均时间序列数据；

 计算所述未来平均时间序列数据与所述第一平均历史时间序列数据的差值，得到所述类簇对象的指标数据；

30 将指标数据大于预设阈值的类簇对象作为目标类簇对象。

17、根据权利要求 16 所述的装置，其特征在于，所述未来时序数据预测模块包括：
基准数据获取子模块，用于对所述类簇对象在未来第一预设时间段内的未来平均时间序列数据进行反归一化处理，得到所述类簇对象中每个数据对象的基准平均时间序列数据；

- 5 修正子模块，用于对所述每个数据对象的基准平均时间序列数据进行修正，得到对应数据对象在未来第一预设时间段内的未来时间序列数据。

18、根据权利要求 10 或 11 或 13 或 14 或 16 或 17 所述的装置，其特征在于，所述数据对象为商品数据，所述类目对象为商品类目，所述特征类目对象为时效性商品类目，所述生命周期为商品的时效，所述时间序列数据为所述商品的日销量。

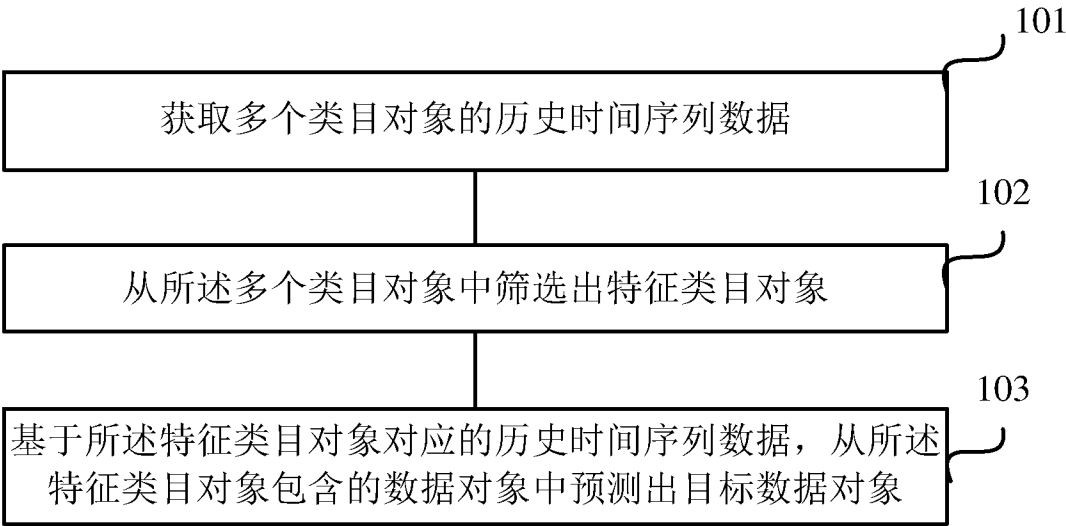


图 1

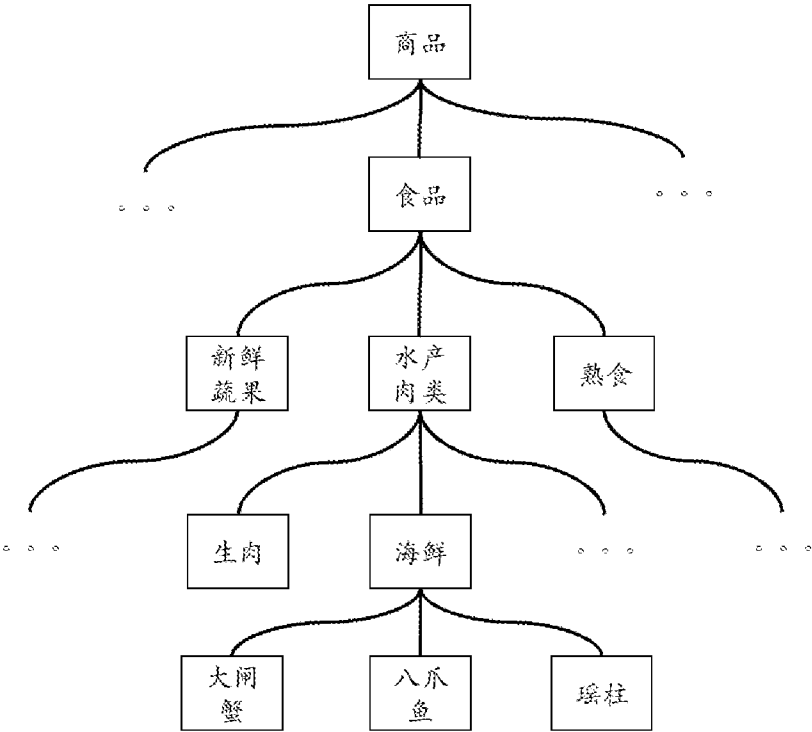


图 2

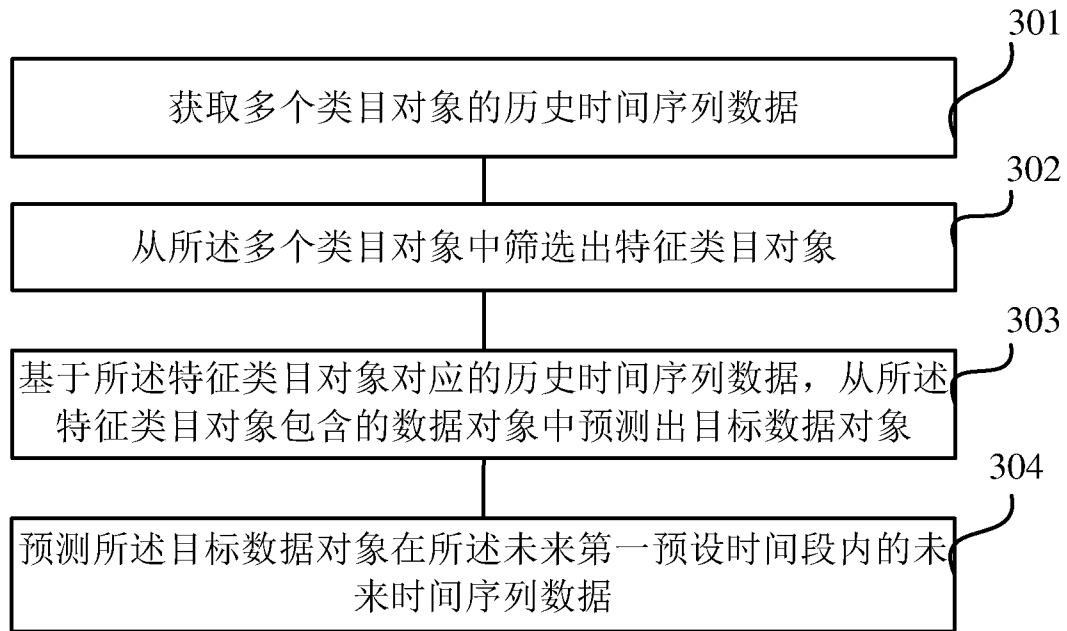


图 3

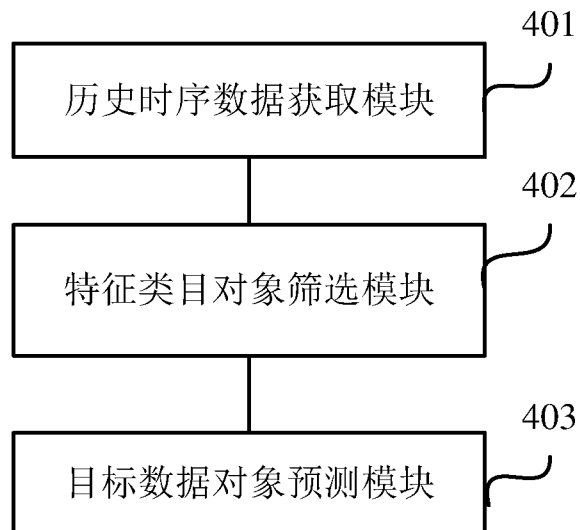


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2017/070356

A. CLASSIFICATION OF SUBJECT MATTER

G06Q 50/02 (2012.01) 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)

G06Q

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)

WPI, EPODOC, CNPAT, CNKI: category, time sequence, history, sort, catalog, commodity, goods, future, forecast

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 102346894 A (ALIBABA GROUP HOLDING LIMITED), 08 February 2012 (08.02.2012), description, paragraphs [0029]-[0090], and claims 1-16	1-18
A	CN 104517224 A (ZHEJIANG UNIVERSITY OF TECHNOLOGY), 15 April 2015 (15.04.2015), the whole document	1-18
A	CN 105184618 A (GUANGZHOU VIPSHOP INFORMATION TECHNOLOGY CO., LTD.), 23 December 2015 (23.12.2015), the whole document	1-18
A	US 2014122155 A1 (WAL-MART STORES, INC.), 01 May 2014 (01.05.2014), the whole document	1-18

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 23 March 2017 (23.03.2017)	Date of mailing of the international search report 11 April 2017 (11.04.2017)
Name and mailing address of the ISA/CN: State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No.: (86-10) 62019451	Authorized officer SU, Fei Telephone No.: (86-10) 62413632

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2017/070356

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102346894 A	08 February 2012	EP 2577591 A1	10 April 2013
		JP 2013534340 A	02 September 2013
		US 2012036037 A1	09 February 2012
		WO 2012018388 A1	09 February 2012
CN 104517224 A	15 April 2015	None	
CN 105184618 A	23 December 2015	None	
US 2014122155 A1	01 May 2014	None	

国际检索报告

国际申请号

PCT/CN2017/070356

A. 主题的分类

G06Q 50/02 (2012.01) 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)

G06Q

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))

WPI, EPODOC, CNPAT, CNKI: 时间序列, 历史, 类别, 类目, 分类, 商品, 未来, 将来, 预测, time sequence, history, sort, catalog, commodity, goods, future, forecast

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 102346894 A (阿里巴巴集团控股有限公司) 2012年 2月 8日 (2012 - 02 - 08) 说明书第[0029]-[0090]段、权利要求1-16	1-18
A	CN 104517224 A (浙江工业大学) 2015年 4月 15日 (2015 - 04 - 15) 全文	1-18
A	CN 105184618 A (广州唯品会信息科技有限公司) 2015年 12月 23日 (2015 - 12 - 23) 全文	1-18
A	US 2014122155 A1 (WAL-MART STORES, INC.) 2014年 5月 1日 (2014 - 05 - 01) 全文	1-18

☐ 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

2017年 3月 23日

国际检索报告邮寄日期

2017年 4月 11日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局 (ISA/CN)
中国北京市海淀区蓟门桥西土城路6号 100088

受权官员

苏菲

传真号 (86-10) 62019451

电话号码 (86-10) 62413632

国际检索报告
关于同族专利的信息

国际申请号
PCT/CN2017/070356

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	102346894	A	2012年 2月 8日	EP	2577591	A1	2013年 4月 10日
				JP	2013534340	A	2013年 9月 2日
				US	2012036037	A1	2012年 2月 9日
				WO	2012018388	A1	2012年 2月 9日
CN	104517224	A	2015年 4月 15日	无			
CN	105184618	A	2015年 12月 23日	无			
US	2014122155	A1	2014年 5月 1日	无			

表 PCT/ISA/210 (同族专利附件) (2009年7月)