

(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
2020 年 11 月 5 日 (05.11.2020)



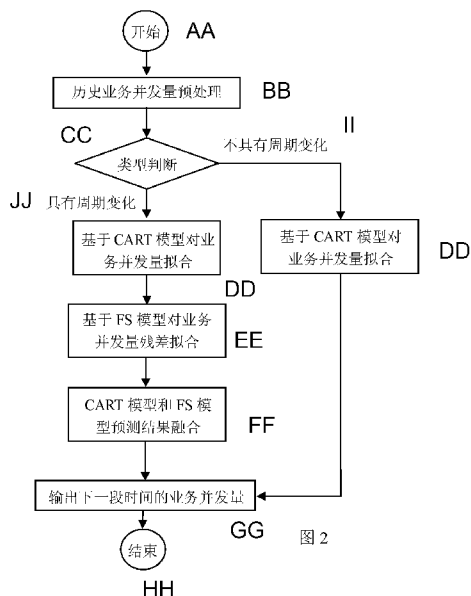
(10) 国际公布号
WO 2020/220438 A1

- (51) 国际专利分类号:
G06F 9/50 (2006.01) **H04L 12/24** (2006.01)
G06F 9/455 (2006.01)
- (21) 国际申请号: PCT/CN2019/090872
- (22) 国际申请日: 2019 年 6 月 12 日 (12.06.2019)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201910355147.5 2019 年 4 月 29 日 (29.04.2019) CN
- (71) 申请人: 东北大学 (NORTHEASTERN UNIVERSITY) [CN/CN]; 中国辽宁省沈阳市和平区文化路三巷 11 号, Liaoning 110819 (CN)。
- (72) 发明人: 郭军(GUO, Jun); 中国辽宁省沈阳市和平区文化路三巷 11 号, Liaoning 110819 (CN)。 王

馨悦(WANG, Xinyue); 中国辽宁省沈阳市和平区文化路三巷 11 号, Liaoning 110819 (CN)。 张斌(ZHANG, Bin); 中国辽宁省沈阳市和平区文化路三巷 11 号, Liaoning 110819 (CN)。 刘晨(LIU, Chen); 中国辽宁省沈阳市和平区文化路三巷 11 号, Liaoning 110819 (CN)。 侯帅(HOU, Shuai); 中国辽宁省沈阳市和平区文化路三巷 11 号, Liaoning 110819 (CN)。 侯凯(HOU, Kai); 中国辽宁省沈阳市和平区文化路三巷 11 号, Liaoning 110819 (CN)。 李薇(LI, Wei); 中国辽宁省沈阳市和平区文化路三巷 11 号, Liaoning 110819 (CN)。 柳波(LIU, Bo); 中国辽宁省沈阳市和平区文化路三巷 11 号, Liaoning 110819 (CN)。 王嘉怡(WANG, Jiayi); 中国辽宁省沈阳市和平区文化路三巷 11 号, Liaoning 110819 (CN)。 刘文凤(LIU, Wenfeng); 中国辽宁省沈阳市和平区文化路三巷 11 号, Liaoning 110819 (CN)。 张瀚铎

(54) Title: METHOD FOR PREDICTING CONCURRENT VOLUME OF SERVICES OF DIFFERENT TYPES FOR VIRTUAL MACHINE

(54) 发明名称: 一种针对虚拟机不同类型的业务并发量预测方法



- AA Start
BB Preprocessing of concurrent volume of historical services
CC Determination of type
DD Fit the concurrent volume of services on the basis of a CART model
EE Fit the residual of the concurrent volume of services on the basis of an FS model
FF Fuse the results of prediction from the CART model and the FS model
GG Output the concurrent volume of services in the next period of time
HH End
II Without a periodic change
JJ With a periodic change

(57) Abstract: Provided is a method for predicting the concurrent volume of services of different types for a virtual machine, wherein same relates to the technical field of cloud computing. A method for predicting the concurrent volume of services of different types for a virtual machine. The method comprises: first, collecting and preprocessing the concurrent volume of historical services of a virtual machine; then, determining the types of concurrent volume of the services of the virtual machine on the basis of an improved I-nearest neighbor-dynamic time warping (INN-DTW) method; and finally, fitting, by means of a classification regression tree, the concurrent volume of services without a periodic change, and fitting, by means of a Fourier series (FS) and a classification and regression tree (CART), the concurrent volume of services with a periodic change. The method for predicting the concurrent volume of services of

(ZHANG, Handuo); 中国辽宁省沈阳市和平区文化路三巷11号, Liaoning 110819 (CN)。 张娅杰 (ZHANG, Yajie); 中国辽宁省沈阳市和平区文化路三巷11号, Liaoning 110819 (CN)。

(74) 代理人: 沈阳东大知识产权代理有限公司 (SHENYANG DONGDA INTELLECTUAL PROPERTY AGENCY CO., LTD); 中国辽宁省沈阳市和平区文化路三巷11号, Liaoning 110819 (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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, 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)。

根据细则4.17的声明:

- 关于申请人有权申请并被授予专利(细则4.17(ii))

本国际公布:

- 包括国际检索报告(条约第21条(3))。

different types for a virtual machine provided in the present invention involves the prediction of the concurrent volume of each type of services of the virtual machine, thereby providing a basis for the increase or decrease in the virtual machine in the next step, and also facilitating the accurate estimation of the aging state of software of the virtual machine in order to achieve the aim of improving the performance and reliability of a working virtual machine.

(57) 摘要: 本发明提供一种针对虚拟机不同类型的业务并发量预测方法, 涉及云计算技术领域。一种针对虚拟机不同类型的业务并发量预测方法, 首先采集虚拟机的历史业务并发量, 并进行预处理, ; 然后基于改进的1最近邻-动态时间调整方法INN-DTW判断虚拟机业务并发量的类型; 最后采用分类回归树拟合不具有周期变化的业务并发量; 采用傅里叶级数FS和分类回归树CART拟合具有周期变化的业务并发量; 本发明提供的针对虚拟机不同类型的业务并发量预测方法, 对虚拟机各业务的并发量进行预测, 可以为下一步虚拟机的增加或者减少提供依据, 同时有助于准确估计虚拟机的软件老化状况, 以达到提高工作虚拟机性能和可靠性的目的。

一种针对虚拟机不同类型的业务并发量预测方法

技术领域

本发明涉及云计算技术领域，尤其涉及一种针对虚拟机不同类型的业务并发量预测方法。

背景技术

软件老化普遍存在于云服务系统中，在虚拟机处理业务并发请求的过程中，操作系统、应用软件等不断地积累错误，导致工作虚拟机的性能逐渐下降，进而影响到云服务系统的服务质量。云平台的高可伸缩、动态重构特性为确保不同并发条件下的云服务质量提供了技术基础，然而现有的虚拟资源动态调整方法仍存在很多缺陷。

一般来说，虚拟机上部署着各种各样的业务，而且不同时间各业务并发量的变化趋势不同，例如，有的业务并发量在白天某段时间持续增加，在晚上某段时间持续减少，有的业务并发量持续循环波动，而有的业务并发量一直保持平稳。通过对云平台各业务的并发量进行预测，可以为下一步虚拟机的增加或者减少提供依据，同时有助于准确估计虚拟机的软件老化状况，以达到提高工作虚拟机性能和可靠性的目的。

由于用户操作、虚拟机业务以及其他不确定性因素在时刻改变，所以业务的并发访问量不仅会随着时间平稳变化，往往还具有上升、下降以及循环波动等趋势，传统的负载模型比如指数平滑模型只能大致刻画出业务并发量的变化趋势，不能很好地捕获其中的非线性变化特征。

发明内容

本发明要解决的技术问题是针对上述现有技术的不足，提供一种针对虚拟机不同类型的业务并发量预测方法，实现对虚拟机中不同类型的业务并发量进行预测。

一种针对虚拟机不同类型的业务并发量预测方法，包括以下步骤：

步骤 1：采集虚拟机的历史业务并发量，并进行预处理，具体方法为：

步骤 1.1：扫描一段时间内虚拟机的业务并发量，发现业务并发量的缺失点；

步骤 1.2：对扫描到的业务并发量缺失点进行处理；

步骤 1.2.1：对于个别采样点缺失的情况，采用前一周期和后一周期业务并发量的平均值进行填补，虚拟机第 t 个时间段的业务并发量 $con(t)$ 缺失的计算如下公式所示：

$$con(t) = \frac{con(t+1) + con(t-1)}{2} \quad (1)$$

步骤 1.2.2：对于样本缺失达到百分九十以上的情况，舍弃全部样本并且将该段时间内业务并发量的值置为零；

步骤 1.3: 对于采集到的业务并发量中存在异常波动的极大极小样本进行异常值调整;

步骤 1.3.1: 结合四分位数计算 t 时间内虚拟机业务并发量正常取值的上限 H 和下限 L , 如下公式所示:

$$H = Q3 + k * (Q3 - Q1) \quad (2)$$

$$L = Q1 - k * (Q3 - Q1) \quad (3)$$

其中, $Q1$ 表示下四分位数, 即 t 时间内业务并发量升序数列的百分之二十五位点, $Q3$ 表示上四分位数, 即 t 时间内业务并发量升序数列的百分之七十五位点, k 用于描述不合理采样点的异常程度, 一般取 1.5 和 3, 分别代表中度和极度;

步骤 1.3.2: 通过图基检验方法判定各采样点数据是否正常, 并对异常值进行调整;

如果采样点数据值被判定为错误业务并发量样本, 则先将错误值丢弃, 再用均值填补法补充;

如果采样点数据值被判定为正常业务并发量样本, 则不做任何调整;

步骤 1.4: 对从日志数据库或者打点日志中采集到的业务并发量和 CPU 利用率数据进行数据间隔调整, 对采集的数据以秒、分钟或小时为单位进行合并;

步骤 1.5: 采用最大最小值归一法将步骤 1.4 处理后的数据进行归一化;

步骤 2: 基于改进的 1 最近邻-动态时间调整 (1-NearestNeighbor-Dynamic Time Warping, 即 1NN-DTW) 方法判断虚拟机业务并发量的类型, 具体方法为:

步骤 2.1: 对虚拟机的各业务并发量进行分类, 分为上升型、下降型、二次型、随机型、周期波动型、周期上升型和周期下降型;

步骤 2.2: 针对各种类型的业务并发量, 提前选取带标签的业务并发量数列作为已知样本;

步骤 2.3: 对每一个待分类的业务并发量数列, 依次扫描所有已知样本并通过临近算法计算出最相近的一条已知样本, 则该已知样本的类型即为待分类业务并发量的类型;

步骤 2.4: 将所有业务并发量归为两大类以简化 1 最近邻模型;

将随机型、上升型、下降型和二次型业务并发量归为不具有周期变化类;

将周期波动型、周期上升型和周期下降型业务并发量归为具有周期变化类;

步骤 2.5: 构造 $n \times m$ 矩阵, 使待分类的业务并发量数列 $\{x_1, x_2, \dots, x_n\}$ 和一条已知的业务并发量数列 $\{y_1, y_2, \dots, y_m\}$ 对齐, 其中, n 为待分类的业务并发量总数量, m 为已知的业务并发量总数量;

步骤 2.6: 将待分类的第 i 个业务并发量 x_i 和已知的第 j 个业务并发量 y_j 两点偏差作为矩阵中 (i, j) 位置的值 d_{ij} , 同时使用欧式距离和两点导数差的平方的方法, 计算待分类的业

业务并发量数列 $\{x_1, x_2, \dots, x_n\}$ 和已知的业务并发量数列 $\{y_1, y_2, \dots, y_m\}$ 对齐后各点的偏差 d_{ij} , 如下公式所示:

$$d_{i,j} = (x_i - y_j)^2 + (x'_i - y'_j)^2 \quad (4)$$

其中, x'_i 、 y'_j 分别为 x_i 、 y_j 的导数, 业务并发量 x_i 的导数 x'_i 的估计如下公式所示:

$$x'_i = \frac{(x_i - x_{i-1}) + ((x_{i+1} - x_{i-1})/2)}{2} \quad (5)$$

步骤 2.7: 在矩阵中从位置 (1,1) 开始, 根据除边界值外规定每个位置只能到达其上方、右方或者右上方的位置的约束条件迭代寻找出一条累积偏差最小的路径, 直到位置 (n,m) 结束;

步骤 3: 预测虚拟机不同变化类型的业务并发量, 具体方法为:

步骤 3.1: 采用分类回归树 (Classification And Regression Tree, 即 CART) 拟合不具有周期变化的业务并发量;

步骤 3.1.1: 遍历样本业务并发量数列的每个特征 F 的任意取值 f , 以 (F, f) 作为条件分割样本数据, 确定平方误差最小的分割位置, 从业务并发量数列中选择最好的切割点;

所述平方误差 $error$ 的计算公式如下:

$$error = \min \left[\sum_{x_i^F < f} (y_i - \hat{y}_i)^2 + \sum_{x_i^F > f} (y_i - \hat{y}_i)^2 \right] \quad (6)$$

其中, x_i^F 代表样本 x 中第 i ' 个业务并发量的特征, y_i 代表分割前的第 i ' 个序列样本, $\hat{y}_i$ 代表分割后的第 i ' 个子序列样本的拟合结果;

步骤 3.1.2: 保存作为切割点的业务并发量值, 并对业务并发量数列执行切分;

步骤 3.1.3: 依次构建特征 F 大于 f 的子树和小于 f 的子树, 进一步迭代对当前分割点左边和右边的业务并发量数列分割拟合, 直到无法再分记为叶子节点;

步骤 3.1.4: 从下而上重新遍历样本数据, 对所有业务并发量数列检查每个分割点, 判断分割之前与分割之后并发量数列的拟合误差,

若分割之后并发量数列的拟合误差降低, 则保留该分割点;

若分割之后并发量数列的拟合误差升高, 则取消该分割点并合并左右数列;

步骤 3.2: 采用傅里叶级数 FS 和分类回归树 CART 拟合具有周期变化的业务并发量;

步骤 3.2.1: 利用分类回归树 CART 拟合 $\{t_1, t_2, \dots, t_n\}$ 时刻的业务并发量得到拟合值 $\{y(0), \dots, y(n'-1), y(n')\}$, 刻画出业务并发量的上升或者下降趋势;

步骤 3.2.2: 把步骤 3.2.1 中所得到的业务并发量与真实业务并发量比较得到残差序列 $\{e(0), e(1), \dots, e(n)\}$;

步骤 3.2.3: 利用分类回归树 CART 预测 $\{t_{n+1}, t_{n+2}, \dots, t_{m'}\}$ 时刻的业务并发量为 $\{y(n+1), y(n+2), \dots, y(m')\}$;

步骤 3.2.4: 利用傅里叶级数 FS 拟合残差序列 $\{e(0), e(1), \dots, e(n)\}$, 刻画出业务并发量的周期趋势, 求得 $\{t_{n'+1}, t_{n'+2}, \dots, t_{m'}\}$ 时刻业务并发量的残差值 $\{e(n'+1), e(n'+2), \dots, e(m')\}$;

步骤 3.2.4.1: 使用函数 $w(t)$ 拟合残差序列 $e(0), e(1), \dots, e(n')$, 函数 $w(t)$ 如下公式所示:

$$w(t) = \frac{a_0}{2} + \sum_{j=1}^k (a_{j'} * \cos \frac{2\pi j' t}{P} + b_{j'} * \sin \frac{2\pi j' t}{P}) \quad (7)$$

其中, a_0 、 $a_{j'}$ 和 $b_{j'}$ 均为变量, $P = n'$, $k = \lfloor n'/2 \rfloor$, $\lfloor \cdot \rfloor$ 表示向下取整, $t = 1, 2, \dots, n'$;

步骤 3.2.4.2: 通过最小二乘法计算变量 $a_{j'}$ 和 $b_{j'}$ 的值, 如下公式所示:

$$(a_{j'}, b_{j'}) = (\frac{2}{n} * \sum_{j'=1}^n w_{j'} * \cos \frac{2\pi j' t}{P}, \frac{2}{n} * \sum_{j'=1}^n w_{j'} * \sin \frac{2\pi j' t}{P}) \quad (8)$$

其中, $w_{j'}$ 为第 j' 个用于拟合残差的函数;

步骤 3.2.5: 将 $\{t_{n'+1}, t_{n'+2}, \dots, t_{m'}\}$ 时刻的业务并发量与其对应的残差值相加, 得到 $\{t_{n'+1}, t_{n'+2}, \dots, t_{m'}\}$ 时刻业务并发量的预测值, 即 $\{y(n'+1) + e(n'+1), y(n'+2) + e(n'+2), \dots, y(m') + e(m')\}$ 。

采用上述技术方案所产生的有益效果在于: 本发明提供的一种针对虚拟机不同类型的业务并发量预测方法, 将业务并发访问量分为周期型、上升型、下降型、二次型和随机型, 不同类型的业务并发量所适用的预测方法不同, 在预测之前对各业务并发量进行分类, 不仅可以有针对性地训练业务并发量模型, 而且在对相同类型的业务并发量建模时还可以实现参数的共享。通过本发明方法对虚拟机各业务的并发量进行预测, 可以为下一步虚拟机的增加或者减少提供依据, 同时有助于准确估计虚拟机的软件老化状况, 以达到提高工作虚拟机性能和可靠性的目的。

附图说明

图 1 为本发明实施例提供的飞机票在线订购系统的实例拓扑图;

图 2 为本发明实施例提供的一种针对虚拟机不同类型的业务并发量预测方法的流程图;

图 3 为本发明实施例提供的二次型业务并发量预测结果的示意图;

图 4 为本发明实施例提供的周期上升型业务并发量预测结果的示意图。

图中, 1、客户端; 2、负载均衡 Nginx; 3、交换机; 4、服务端; 5、业务数据库 MySQL。

具体实施方式

下面结合附图和实施例, 对本发明的具体实施方式作进一步详细描述。以下实施例用于说明本发明, 但不用来限制本发明的范围。

本实施例以飞机票在线订购系统模拟 PC 端用户应用, 在曙光服务器上搭建该服务系统,

通过对飞机票在线订购系统加压模拟真实的业务并发场景，并采集不同的业务并发量数据为例，使用本发明的一种针对虚拟机不同类型的业务并发量预测方法进行业务并发量的预测。实例拓扑图如图 1 所示，客户端 1 使用 LoadRunner 软件产生业务并发访问，它可以模拟大量的用户同时点击飞机票订购系统页面，LoadRunner 发送页面请求后，由负载均衡 Nginx2 实现业务请求的接收和分配，最后服务端 4 安装 Tomcat 并部署飞机票在线预订系统，负责读写业务数据库 MySQL5，处理 LoadRunner 发送的请求。

一种针对虚拟机不同类型的业务并发量预测方法，如图 2 所示，包括以下步骤：

步骤 1：采集虚拟机的历史业务并发量，并进行预处理，具体方法为：

步骤 1.1：扫描一段时间内虚拟机的业务并发量，发现业务并发量的缺失点；

步骤 1.2：对扫描到的业务并发量缺失点进行处理；

步骤 1.2.1：对于个别采样点缺失的情况，采用前一周期和后一周期业务并发量的平均值进行填补，虚拟机第 t 个时间段的业务并发量 $con(t)$ 缺失的计算如下公式所示：

$$con(t) = \frac{con(t+1) + con(t-1)}{2} \quad (1)$$

步骤 1.2.2：对于样本缺失达到百分九十以上舍弃全部样本并且将该段时间内业务并发量的值置为零；例如在 20 个连续采样周期中，只有 2 个周期采集到业务并发量值，甚至全部数据为空，那么可以认为这段时间采集到的业务并发量都是不可信的，不能纳入历史数列进行预测；

步骤 1.3：对于采集到的业务并发量中存在异常波动的极大极小样本进行异常值调整；

步骤 1.3.1：结合四分位数计算 t 时间内虚拟机业务并发量正常取值的上限 H 和下限 L ，如下公式所示：

$$H = Q3 + k * (Q3 - Q1) \quad (2)$$

$$L = Q1 - k * (Q3 - Q1) \quad (3)$$

其中， $Q1$ 表示下四分位数，即 t 时间内业务并发量升序数列的百分之二十五位点， $Q3$ 表示上四分位数，即 t 时间内业务并发量升序数列的百分之七十五位点， k 用于描述不合理采样点的异常程度，一般取 1.5 和 3，分别代表中度和极度；

步骤 1.3.2：通过图基检验方法判定各采样点数据是否正常，并对异常值进行调整；

如果采样点数据值被判定为错误业务并发量样本，则先将错误值丢弃，再用均值填补法补充；

如果采样点数据值被判定为正常业务并发量样本，则不做任何调整；

步骤 1.4：对从日志数据库或者打点日志中采集到的业务并发量和 CPU 利用率数据进行

数据间隔调整，对采集的数据以秒、分钟或小时为单位进行合并；

在业务并发访问量建模时，以 1 秒为时间间隔采样的业务并发量波动频繁，趋势变化不明显，无法挖掘变化的特征，而且过密采样使得模型计算量加大，训练更加迟缓；因此，在本实施例中以 15 秒为间隔取平均值进行整理数据都是，虚拟机的其他数据也是以 15 秒为间隔；

步骤 1.5：采用最大最小值归一法将步骤 1.4 处理后的数据进行归一化；

步骤 2：基于改进的 1 最近邻-动态时间调整（1-NearestNeighbor-Dynamic Time Warping，即 1NN-DTW）方法判断虚拟机业务并发量的类型，具体方法为：

步骤 2.1：对虚拟机的各业务并发量进行分类，分为上升型、下降型、二次型、随机型、周期波动型、周期上升型和周期下降型；

步骤 2.2：针对各种类型的业务并发量，提前选取带标签的业务并发量数列作为已知样本；

步骤 2.3：对每一个待分类的业务并发量数列，依次扫描所有已知样本并通过临近算法计算出最相近的一条已知样本，则该已知样本的类型即为待分类业务并发量的类型；

步骤 2.4：将所有业务并发量归为两大类以简化 1 最近邻模型；

将随机型、上升型、下降型和二次型业务并发量归为不具有周期变化类；

将周期波动型、周期上升型和周期下降型业务并发量归为具有周期变化类；

步骤 2.5：构造 $n \times m$ 矩阵，使待分类的业务并发量数列 $\{x_1, x_2, \dots, x_n\}$ 和一条已知的业务并发量数列 $\{y_1, y_2, \dots, y_m\}$ 对齐，其中， n 为待分类的业务并发量总数量， m 为已知的业务并发量总数量；

步骤 2.6：将待分类的第 i 个业务并发量 x_i 和已知的第 j 个业务并发量 y_j 两点偏差作为矩阵中 (i, j) 位置的值 $d_{i,j}$ ，同时使用欧式距离和两点导数差的平方的方法，计算待分类的业务并发量数列 $\{x_1, x_2, \dots, x_n\}$ 和已知的业务并发量数列 $\{y_1, y_2, \dots, y_m\}$ 对齐后各点的偏差 $d_{i,j}$ ，如下公式所示：

$$d_{i,j} = (x_i - y_j)^2 + (x'_i - y'_j)^2 \quad (4)$$

其中， x'_i 、 y'_j 分别为 x_i 、 y_j 的导数，业务并发量 x_i 的导数 x'_i 的估计如下公式所示：

$$x'_i = \frac{(x_i - x_{i-1}) + ((x_{i+1} - x_{i-1})/2)}{2} \quad (5)$$

步骤 2.7：在矩阵中从位置 (1,1) 开始，根据除边界值外规定每个位置只能到达其上方、右方或者右上方的位置的约束条件迭代寻找出一条累积偏差最小的路径，直到位置 (n,m) 结束；

步骤 3：预测虚拟机不同变化类型的业务并发量，具体方法为：

步骤 3.1：采用分类回归树 (Classification And Regression Tree, 即 CART)拟合不具有周期变化的业务并发量；

步骤 3.1.1：遍历样本业务并发量数列的每个特征 F 的任意取值 f ，以 (F, f) 作为条件分割样本数据，确定平方误差最小的分割位置，从业务并发量数列中选择最好的切割点；

所述平方误差 error 的计算公式如下：

$$error = \min \left[\sum_{x_i^F < f} (y_{i'} - \hat{y}_{i'})^2 + \sum_{x_i^F > f} (y_{i'} - \hat{y}_{i'})^2 \right] \quad (6)$$

其中， x_i^F 代表样本 x 中第 i' 个业务并发量的特征， $y_{i'}$ 代表分割前的第 i' 个序列样本， $\hat{y}_{i'}$ 代表分割后的第 i' 个子序列样本的拟合结果；

步骤 3.1.2：保存作为切割点的业务并发量值，并对业务并发量数列执行切分；

步骤 3.1.3：依次构建特征 F 大于 f 的子树和小于 f 的子树，进一步迭代对当前分割点左边和右边的业务并发量数列分割拟合，直到无法再分记为叶子节点；

步骤 3.1.4：从下而上重新遍历样本数据，对所有业务并发量数列检查每个分割点，判断分割之前与分割之后并发量数列的拟合误差，

若分割之后并发量数列的拟合误差降低，则保留该分割点；

若分割之后并发量数列的拟合误差升高，则取消该分割点并合并左右数列；

步骤 3.2：采用傅里叶级数 FS 和分类回归树 CART 拟合具有周期变化的业务并发量；

步骤 3.2.1：利用分类回归树 CART 拟合 $\{t_1, t_2, \dots, t_n\}$ 时刻的业务并发量得到拟合值 $\{y(0), \dots, y(n'-1), y(n')\}$ ，刻画出业务并发量的上升或者下降趋势；

步骤 3.2.2：把步骤 3.2.1 中所得到的业务并发量与真实业务并发量比较得到残差序列 $\{e(0), e(1), \dots, e(n)\}$ ；

步骤 3.2.3：利用分类回归树 CART 预测 $\{t_{n+1}, t_{n+2}, \dots, t_m\}$ 时刻的业务并发量为 $\{y(n+1), y(n+2), \dots, y(m')\}$ ；

步骤 3.2.4：利用傅里叶级数 FS 拟合残差序列 $\{e(0), e(1), \dots, e(n)\}$ ，刻画出业务并发量的周期趋势，求得 $\{t_{n'+1}, t_{n'+2}, \dots, t_m\}$ 时刻业务并发量的残差值 $\{e(n'+1), e(n'+2), \dots, e(m')\}$ ；

步骤 3.2.4.1：使用函数 $w(t)$ 拟合残差序列 $e(0), e(1), \dots, e(n')$ ，函数 $w(t)$ 如下公式所示：

$$w(t) = \frac{a_0}{2} + \sum_{j=1}^k (a_{j'} * \cos \frac{2\pi j' t}{P} + b_{j'} * \sin \frac{2\pi j' t}{P}) \quad (7)$$

其中， a_0 、 $a_{j'}$ 和 $b_{j'}$ 均为变量， $P=n'$ ， $k=[n'/2]$ ， $[]$ 表示向下取整， $t=1, 2, \dots, n'$ ；

步骤 3.2.4.2：通过最小二乘法计算变量 $a_{j'}$ 和 $b_{j'}$ 的值，如下公式所示：

$$(a_{j'}, b_{j'}) = (\frac{2}{n} * \sum_{j'=1}^n w_{j'} * \cos \frac{2\pi j' t}{P}, \frac{2}{n} * \sum_{j'=1}^n w_{j'} * \sin \frac{2\pi j' t}{P}) \quad (8)$$

其中, $w_{j'}$ 为第 j' 个用于拟合残差的函数;

步骤 3.2.5: 将 $\{t_{n'+1}, t_{n'+2}, \dots, t_{m'}\}$ 时刻的业务并发量与其对应的残差值相加, 得到 $\{t_{n'+1}, t_{n'+2}, \dots, t_{m'}\}$ 时刻业务并发量的预测值, 即 $\{y(n'+1)+e(n'+1), y(n'+2)+e(n'+2), \dots, y(m')+e(m')\}$ 。

本实施例还提供了使用改进的 1NN-DTW 算法进行业务并发量的类型判断, 并与改进前算法进行对比, 验证改进后 1NN-DTW 的准确性, 具体为:

首先使用 LoadRunner 对服务端应用的浏览、查询、退票等各类业务的访问行为进行记录。然后对服务端虚拟机持续加压一小时并采集业务并发量, 按照预处理的方法处理缺失和异常的业务并发量值, 并以 15 秒为间隔调整并发量数据。

利用改进的 1NN-DTW 算法判断业务并发访问量类型, 并与 1NN-DTW、1NN-DDTW 对比, 采用正确率 Accuracy 和 F 值 F-measure 来衡量各算法的好坏。将第一步得到的并发量分别按每 80、120、160、200 个采样点截取为一个子序列, 并根据表 1 中列举的七种负载变化趋势打上类型标签作为一个样本序列, 最后得到 700 个样本序列, 选取其中 420 个作为类型判断的已知样本, 剩下的 280 个作为测试样本。

表 1 不同类型的业务并发访问量

并发量分类	具体趋势	并发量数据示例	标签
不具有周期变化	随机型	{102,105,103, 105,102,103,104,106,103,107}	1
	上升型	{102,310,309,420,510,518,616,530,721,856}	2
	下降型	{739,662,503,518,261,238,166,137,164,81}	3
	二次型	{727,687,524,451,256,318,366,523,718,825}	4
具有周期变化	周期波动型	{81,138,195,157,62,138,157,176,138,88}	5
	周期上升型	{84,142,169,125,268,237,184,271,286,241}	6
	周期下降型	{708,658,663,418,401,394,196,217,194,60}	7

采用本发明的改进的 1NN-DTW 和现有的 1NN-DTW、1NN-DDTW 这三种方法对业务并发量类型判断的对比结果如表 2 所示。从表 2 可以看出, 本发明方法的 Accuracy、F-measure 明显高于另外两种方法, 说明在判断业务并发量类型时, 从业务并发量的取值和变化趋势两方面考虑效果要优于只关注其中一个方面。另外, 虽然本发明方法同时计算相似点的欧式距离和导数差, 但是所用时间并未大幅度增加。

表 2 不同方法的业务并发量分类情况

方法	Accuracy	F-measure	Time(ms)
改进的 1NN-DTW	0.942	0.867	1120
1NN-DTW	0.873	0.751	984
1NN-DDTW	0.916	0.834	1097

本实施例还提供了使用本发明方法预测业务并发量，并与传统的 ARIMA 等方法进行对比，具体为：

首先使用 LoadRunner 对服务端应用的浏览、查询、退票等各类业务的访问行为进行记录。然后对服务端虚拟机持续加压一小时并采集业务并发量，按照预处理叙述的方法处理缺失和异常的业务并发量值，并以 15 秒为间隔调整并发量数据。

选取二次型和周期上升型两类相对复杂的并发量进行预测。通过分析过去 25 分钟的业务并发量值，估计未来 5 分钟的业务并发量，并选取均方误差 MSE、绝对误差 MAE、用时 Time 三项评价标准，借助 Python 工具包将本方法与 ARIMA、指数平滑 Holt-Winters 对比，验证本发明方法的准确性。

三种方法的业务并发量预测结果如表 3 所示，三种方法的业务并发量预测结果与真实并发量之间的对照结果如图 3 和图 4 所示。从图中来看，在设定的两种情况下本发明方法与 ARIMA、Holt-Winters 相比，对真实的业务并发量序列拟合更好，说明本发明方法在对各类并发量预测时较为有效。根据表 3 中结果进一步分析，本发明方法与 ARIMA、Holt-Winters 相比，在两种类型的并发量场景下 MSE 和 MAE 最低。在二次型并发量场景下三种方法的 MSE 和 MAE 较为接近，但是在周期上升型并发量场景下本方法明显更优，ARIMA、Holt-Winters 对这类复杂的并发量学习能力较差，这些表明在各种场景下本发明方法都具有可观的准确度。

表 3 不同方法的业务并发量预测结果

	二次型并发量			周期上升型并发量		
	MSE	MAE	Time(ms)	MSE	MAE	Time(ms)
本发明方法	1.2084	0.8639	2187	7.1271	2.1157	2439
ARIMA	3.3817	1.5537	1473	13.9613	3.0271	1652
Holt-Winters	4.1526	1.7906	1264	17.7694	3.6470	1138

最后应说明的是：以上实施例仅用以说明本发明的技术方案，而非对其限制；尽管参照前述实施例对本发明进行了详细的说明，本领域的普通技术人员应当理解：其依然可以对前述实施例所记载的技术方案进行修改，或者对其中部分或者全部技术特征进行等同替换；而

这些修改或者替换，并不使相应技术方案的本质脱离本发明权利要求所限定的范围。

权 利 要 求 书

1、一种针对虚拟机不同类型的业务并发量预测方法，其特征在于：包括以下步骤：

步骤 1：采集虚拟机的历史业务并发量，并进行预处理，具体方法为：

步骤 1.1：扫描一段时间内虚拟机的业务并发量，发现业务并发量的缺失点；

步骤 1.2：对扫描到的业务并发量缺失点进行处理；

步骤 1.3：对于采集到的业务并发量中存在异常波动的极大极小样本进行异常值调整；

步骤 1.4：对从日志数据库或者打点日志中采集到的业务并发量和 CPU 利用率数据进行数据间隔调整，对采集的数据以秒、分钟或小时为单位进行合并；

步骤 1.5：采用最大最小值归一法将步骤 1.4 处理后的数据进行归一化；

步骤 2：基于改进的 1 最近邻-动态时间调整方法 1NN-DTW 判断虚拟机业务并发量的类型，具体方法为：

步骤 2.1：对虚拟机的各业务并发量进行分类，分为上升型、下降型、二次型、随机型、周期波动型、周期上升型和周期下降型；

步骤 2.2：针对各种类型的业务并发量，提前选取带标签的业务并发量数列作为已知样本；

步骤 2.3：对每一个待分类的业务并发量数列，依次扫描所有已知样本并通过临近算法计算出最相近的一条已知样本，则该已知样本的类型即为待分类业务并发量的类型；

步骤 2.4：将所有业务并发量归为两大类以简化 1 最近邻模型；

将随机型、上升型、下降型和二次型业务并发量归为不具有周期变化类；

将周期波动型、周期上升型和周期下降型业务并发量归为具有周期变化类；

步骤 2.5：构造 $n \times m$ 矩阵，使待分类的业务并发量数列 $\{x_1, x_2, \dots, x_n\}$ 和一条已知的业务并发量数列 $\{y_1, y_2, \dots, y_m\}$ 对齐，其中， n 为待分类的业务并发量总数量， m 为已知的业务并发量总数量；

步骤 2.6：将待分类的第 i 个业务并发量 x_i 和已知的第 j 个业务并发量 y_j 两点偏差作为矩阵中 (i, j) 位置的值 d_{ij} ，同时使用欧式距离和两点导数差的平方的方法，计算待分类的业务并发量数列 $\{x_1, x_2, \dots, x_n\}$ 和已知的业务并发量数列 $\{y_1, y_2, \dots, y_m\}$ 对齐后各点的偏差 d_{ij} ，如下公式所示：

$$d_{i,j} = (x_i - y_j)^2 + (x'_i - y'_j)^2 \quad (1)$$

其中， x'_i 、 y'_j 分别为 x_i 、 y_j 的导数，业务并发量 x_i 的导数 x'_i 的估计如下公式所示：

$$x'_i = \frac{(x_i - x_{i-1}) + ((x_{i+1} - x_{i-1})/2)}{2} \quad (2)$$

步骤 2.7：在矩阵中从位置 (1,1) 开始，根据除边界值外规定每个位置只能到达其上方、

右方或者右上方的位置的约束条件迭代寻找出一条累积偏差最小的路径，直到位置 (n,m) 结束；

步骤 3：预测虚拟机不同变化类型的业务并发量，具体方法为：

步骤 3.1：采用分类回归树 CART 拟合不具有周期变化的业务并发量；

步骤 3.2：采用傅里叶级数 FS 和分类回归树 CART 拟合具有周期变化的业务并发量。

2、根据权利要求 1 所述的一种针对虚拟机不同类型的业务并发量预测方法，其特征在于：所述步骤 1.2 的具体方法为：

步骤 1.2.1：对于个别采样点缺失的情况，采用前一周期和后一周期业务并发量的平均值进行填补，虚拟机第 t 个时间段的业务并发量 $con(t)$ 缺失的计算如下公式所示：

$$con(t) = \frac{con(t+1) + con(t-1)}{2} \quad (3)$$

步骤 1.2.2：对于样本缺失达到百分九十以上的情况，舍弃全部样本并且将该段时间内业务并发量的值置为零。

3、根据权利要求 1 所述的一种针对虚拟机不同类型的业务并发量预测方法，其特征在于：所述步骤 1.3 的具体方法为：

步骤 1.3.1：结合四分位数计算 t 时间内虚拟机业务并发量正常取值的上限 H 和下限 L ，如下公式所示：

$$H = Q3 + k * (Q3 - Q1) \quad (4)$$

$$L = Q1 - k * (Q3 - Q1) \quad (5)$$

其中， $Q1$ 表示下四分位数，即 t 时间内业务并发量升序数列的百分之二十五位点， $Q3$ 表示上四分位数，即 t 时间内业务并发量升序数列的百分之七十五位点， k 用于描述不合理采样点的异常程度，一般取 1.5 和 3，分别代表中度和极度；

步骤 1.3.2：通过图基检验方法判定各采样点数据是否正常，并对异常值进行调整；

如果采样点数据值被判定为错误业务并发量样本，则先将错误值丢弃，再用均值填补法补充；

如果采样点数据值被判定为正常业务并发量样本，则不做任何调整。

4、根据权利要求 1 所述的一种针对虚拟机不同类型的业务并发量预测方法，其特征在于：所述步骤 3.1 的具体方法为：

步骤 3.1.1：遍历样本业务并发量数列的每个特征 F 的任意取值 f ，以 (F, f) 作为条件分割样本数据，确定平方误差最小的分割位置，从业务并发量数列中选择最好的切割点；

所述平方误差 error 的计算公式如下：

$$error = \min \left[\sum_{x_i^F < f} (y_{i'} - \hat{y}_{i'})^2 + \sum_{x_i^F > f} (y_{i'} - \hat{y}_{i'})^2 \right] \quad (6)$$

其中, x_i^F 代表样本 x 中第 i' 个业务并发量的特征, $y_{i'}$ 代表分割前的第 i' 个序列样本, $\hat{y}_{i'}$ 代表分割后的第 i' 个子序列样本的拟合结果;

步骤 3.1.2: 保存作为切割点的业务并发量值, 并对业务并发量数列执行切分;

步骤 3.1.3: 依次构建特征 F 大于 f 的子树和小于 f 的子树, 进一步迭代对当前分割点左边和右边的业务并发量数列分割拟合, 直到无法再分记为叶子节点;

步骤 3.1.4: 从下而上重新遍历样本数据, 对所有业务并发量数列检查每个分割点, 判断分割之前与分割之后并发量数列的拟合误差,

若分割之后并发量数列的拟合误差降低, 则保留该分割点;

若分割之后并发量数列的拟合误差升高, 则取消该分割点并合并左右数列。

5、根据权利要求 4 所述的一种针对虚拟机不同类型的业务并发量预测方法, 其特征在于: 所述步骤 3.2 的具体方法为:

步骤 3.2.1: 利用分类回归树 CART 拟合 $\{t_1, t_2, \dots, t_n\}$ 时刻的业务并发量得到拟合值 $\{y(0), \dots, y(n'-1), y(n')\}$, 刻画出业务并发量的上升或者下降趋势;

步骤 3.2.2: 把步骤 3.2.1 中所得到的业务并发量与真实业务并发量比较得到残差序列 $\{e(0), e(1), \dots, e(n)\}$;

步骤 3.2.3: 利用分类回归树 CART 预测 $\{t_{n+1}, t_{n+2}, \dots, t_m\}$ 时刻的业务并发量为 $\{y(n+1), y(n+2), \dots, y(m')\}$;

步骤 3.2.4: 利用傅里叶级数 FS 拟合残差序列 $\{e(0), e(1), \dots, e(n)\}$, 刻画出业务并发量的周期趋势, 求得 $\{t_{n'+1}, t_{n'+2}, \dots, t_{m'}\}$ 时刻业务并发量的残差值 $\{e(n'+1), e(n'+2), \dots, e(m')\}$;

步骤 3.2.5: 将 $\{t_{n'+1}, t_{n'+2}, \dots, t_{m'}\}$ 时刻的业务并发量与其对应的残差值相加, 得到 $\{t_{n'+1}, t_{n'+2}, \dots, t_{m'}\}$ 时刻业务并发量的预测值, 即 $\{y(n'+1) + e(n'+1), y(n'+2) + e(n'+2), \dots, y(m') + e(m')\}$ 。

6、根据权利要求 5 所述的一种针对虚拟机不同类型的业务并发量预测方法, 其特征在于: 所述步骤 3.2.4 的具体方法为:

步骤 3.2.4.1: 使用函数 $w(t)$ 拟合残差序列 $e(0), e(1), \dots, e(n')$, 函数 $w(t)$ 如下公式所示:

$$w(t) = \frac{a_0}{2} + \sum_{j=1}^k (a_{j'} * \cos \frac{2\pi j' t}{P} + b_{j'} * \sin \frac{2\pi j' t}{P}) \quad (7)$$

其中, a_0 、 $a_{j'}$ 和 $b_{j'}$ 均为变量, $P=n'$, $k=[n'/2]$, $[]$ 表示向下取整, $t=1, 2, \dots, n'$;

步骤 3.2.4.2: 通过最小二乘法计算变量 $a_{j'}$ 和 $b_{j'}$ 的值, 如下公式所示:

$$(a_{j'}, b_{j'}) = (\frac{2}{n} * \sum_{j'=1}^n w_{j'} * \cos \frac{2\pi j' t}{P}, \frac{2}{n} * \sum_{j'=1}^n w_{j'} * \sin \frac{2\pi j' t}{P}) \quad (8)$$

其中， $w_{j'}$ 为第 j' 个用于拟合残差的函数。

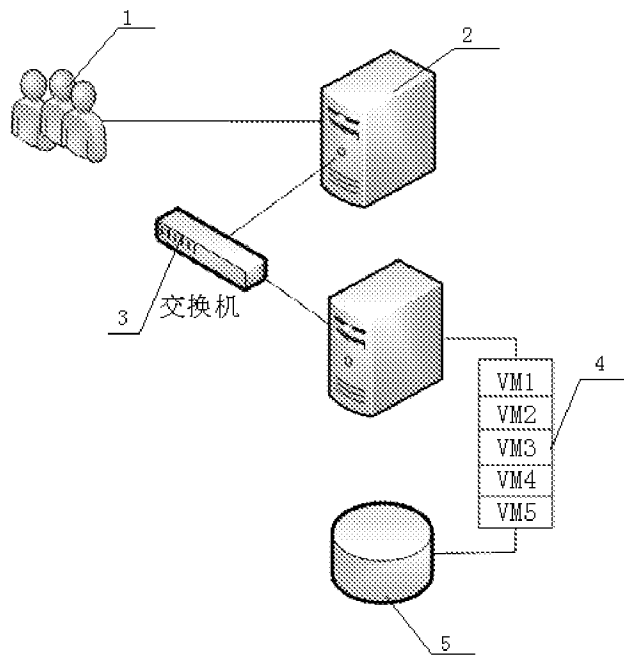


图 1

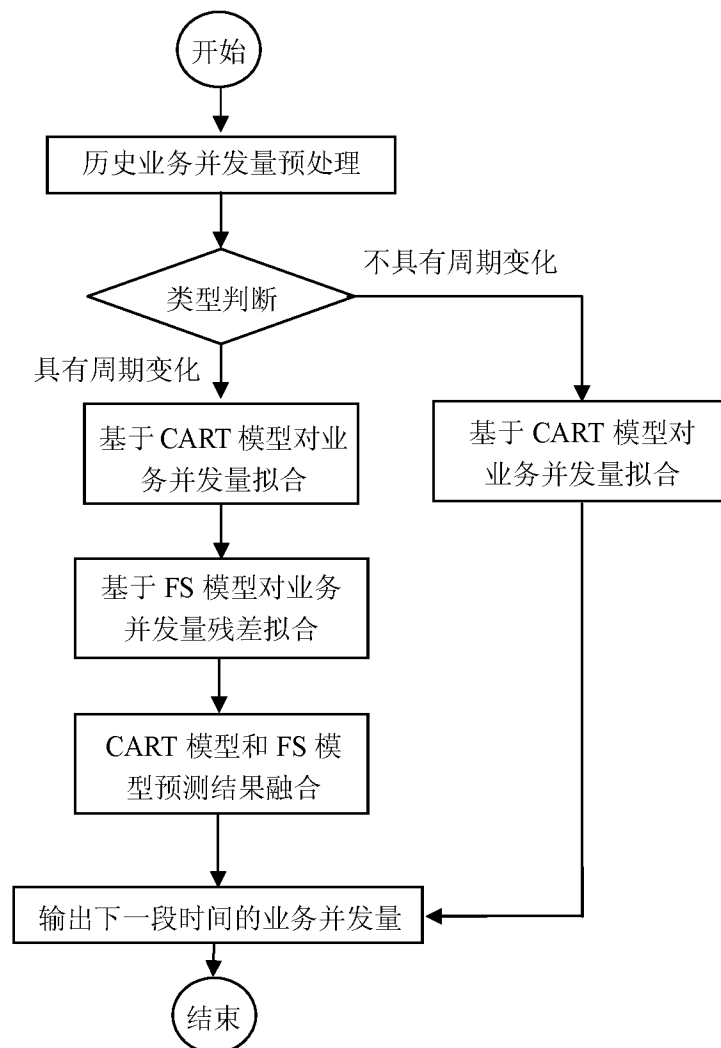


图 2

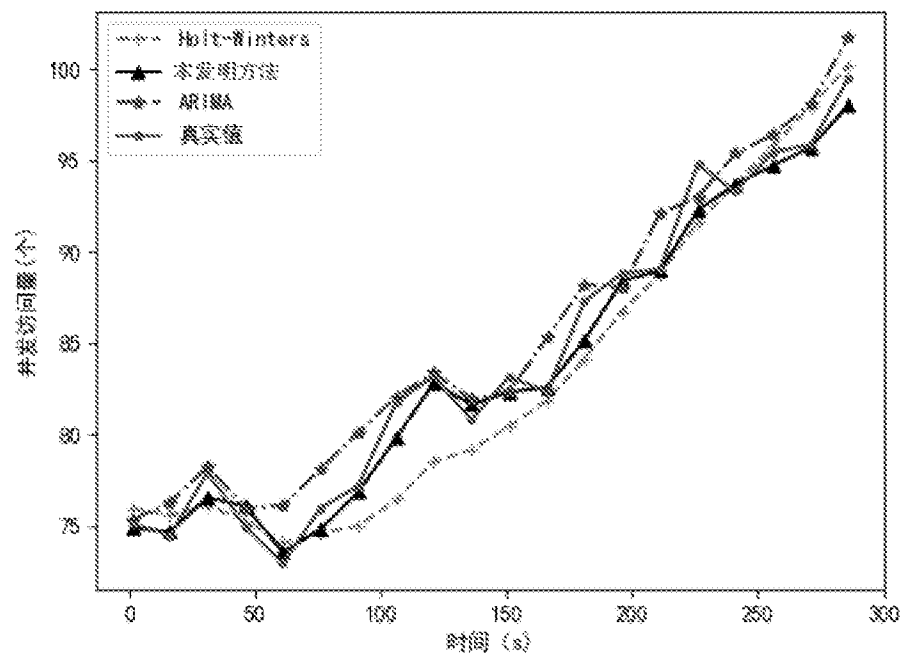


图 3

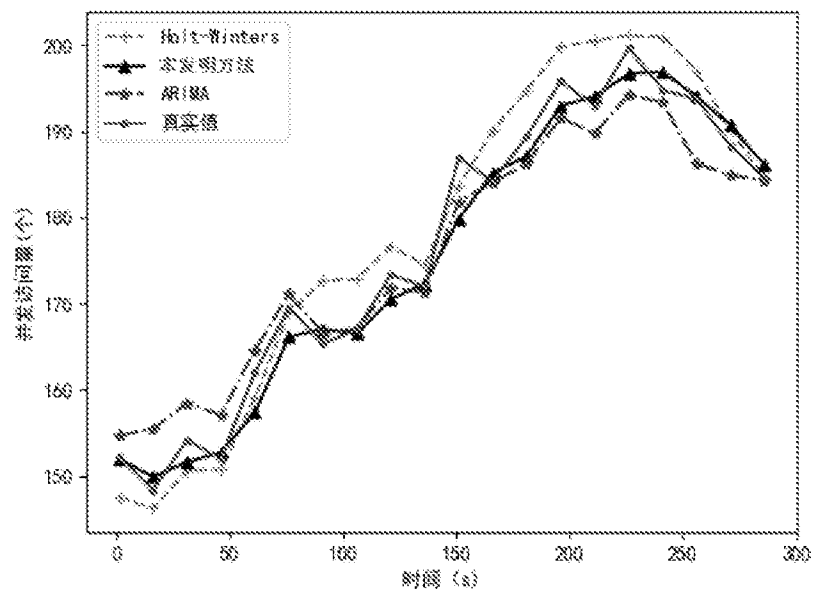


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/090872

A. CLASSIFICATION OF SUBJECT MATTER

G06F 9/50(2006.01)i; G06F 9/455(2006.01)i; H04L 12/24(2006.01)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,H04L

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)

CNABS; CNTXT; VEN; EPTXT; USTXT; WOTXT; CNKI; IEEE: 虚拟机, 业务, 负荷, 负载, 预测, 规律, 周期, 最近邻模型, 分类回归树, 矩阵, 训练序列, 训练数列, 样本, 偏差, 最小, 路径, 欧氏, 欧几里得, 距离, 动态时间规整, VM, predict, CART, 1NN, KNN, Dynamic time wrapping, DTW, 1 nearest neighbor, adaptive, scalar, rule, matrix, Euclidean distance

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	孟煜 等 (MENG, Yu et al.). ""云计算环境下云服务用户并发量的区间预测模型 (Prediction Interval Estimation Model of User Concurrent Requests for Cloud Service in Cloud Environment)"" <i>计算机学报 (Chinese Journal of Computers)</i> , Vol. 40, No. (2), 28 February 2017 (2017-02-28), ISSN: 0254-4164, pp. 378-396	1-6
A	原继东 等 (YUAN, Jidong et al.). ""时间序列的表示与分类算法综述 (Review of Time Series Representation and Classification Techniques)"" <i>计算机科学 (Computer Science)</i> , Vol. 42, No. (3), 31 March 2015 (2015-03-31), ISSN: 1002-137X, section 4.1	1-6
A	CN 109088747 A (ZHENGZHOU YUNHAI INFORMATION TECHNOLOGY CO., LTD.) 25 December 2018 (2018-12-25) description, paragraphs [0002] and [0062]-[0122]	1-6



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

03 January 2020

Date of mailing of the international search report

19 January 2020

Name and mailing address of the ISA/CN

China National Intellectual Property Administration
No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing
100088
China

Facsimile No. (86-10)62019451

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2019/090872

C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	CN 109409496 A (CHONGQING UNIVERSITY OF POSTS AND TELECOMMUNICATIONS) 01 March 2019 (2019-03-01) entire document	1-6
A	CN 109034179 A (HENAN POLYTECHNIC UNIVERSITY) 18 December 2018 (2018-12-18) entire document	1-6
A	CN 104407688 A (HARBIN INSTITUTE OF TECHNOLOGY SHENZHEN GRADUATE SCHOOL) 11 March 2015 (2015-03-11) entire document	1-6
A	CN 104915434 A (HARBIN INSTITUTE OF TECHNOLOGY) 16 September 2015 (2015-09-16) entire document	1-6
A	US 2018018533 A1 (UNIVERSITY OF CENTRAL FLORIDA RESEARCH FOUNDATION, INC.) 18 January 2018 (2018-01-18) entire document	1-6

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/CN2019/090872

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	109088747	A	25 December 2018	None			
CN	109409496	A	01 March 2019	None			
CN	109034179	A	18 December 2018	None			
CN	104407688	A	11 March 2015	None			
CN	104915434	A	16 September 2015	CN	104915434	B	27 March 2018
US	2018018533	A1	18 January 2018	US	10133949	B2	20 November 2018
				WO	2018014018	A1	18 January 2018

国际检索报告

国际申请号

PCT/CN2019/090872

A. 主题的分类 G06F 9/50(2006.01)i; G06F 9/455(2006.01)i; H04L 12/24(2006.01)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, H04L 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)) CNABS;CNTXT;VEN;EPTXT;USTXT;WOTXT;CNKI;IEEE: 虚拟机, 业务, 负荷, 负载, 预测, 规律, 周期, 最近邻模型, 分类回归树, 矩阵, 训练序列, 训练数列, 样本, 偏差, 最小, 路径, 欧氏, 欧几里得, 距离, 动态时间规整, VM, predict, CART, INN, KNN, Dynamic time wrapping, DTW, 1 nearest neighbor, adaptive, scalar, rule, matrix, Euclidean distance		
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	孟煜 等. “云计算环境下云服务用户并发量的区间预测模型” 计算机学报, 第40卷, 第2期, 2017年 2月 28日 (2017 - 02 - 28), ISSN: 0254-4164, 第378-396页	1-6
A	原继东 等. “时间序列的表示与分类算法综述” 计算机科学, 第42卷, 第3期, 2015年 3月 31日 (2015 - 03 - 31), ISSN: 1002-137X, 第4.1节	1-6
A	CN 109088747 A (郑州云海信息技术有限公司) 2018年 12月 25日 (2018 - 12 - 25) 说明书第[0002]、[0062]-[0122]段	1-6
A	CN 109409496 A (重庆邮电大学) 2019年 3月 1日 (2019 - 03 - 01) 全文	1-6
A	CN 109034179 A (河南理工大学) 2018年 12月 18日 (2018 - 12 - 18) 全文	1-6
A	CN 104407688 A (哈尔滨工业大学深圳研究生院) 2015年 3月 11日 (2015 - 03 - 11) 全文	1-6
☒ 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 1月 3日		2020年 1月 19日
ISA/CN的名称和邮寄地址		受权官员
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		匡仁炳 电话号码 86-(20)-28950886

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 104915434 A (哈尔滨工业大学) 2015年 9月 16日 (2015 - 09 - 16) 全文	1-6
A	US 2018018533 A1 (UNIV CENTRAL FLORIDA RES FOUND INC) 2018年 1月 18日 (2018 - 01 - 18) 全文	1-6

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2019/090872

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	109088747	A	2018年 12月 25日	无	
CN	109409496	A	2019年 3月 1日	无	
CN	109034179	A	2018年 12月 18日	无	
CN	104407688	A	2015年 3月 11日	无	
CN	104915434	A	2015年 9月 16日	CN	104915434 B 2018年 3月 27日
US	2018018533	A1	2018年 1月 18日	US	10133949 B2 2018年 11月 20日
				WO	2018014018 A1 2018年 1月 18日