



## (12)发明专利申请

(10)申请公布号 CN 111404751 A

(43)申请公布日 2020.07.10

(21)申请号 202010201945.5

(22)申请日 2020.03.20

(71)申请人 南京大学

地址 210046 江苏省南京市栖霞区仙林大道163号

申请人 成都云格致力科技有限公司

(72)发明人 张旭 张欣宇 薛雨 马展

(74)专利代理机构 江苏法德东恒律师事务所  
32305

代理人 李媛媛

(51)Int.Cl.

H04L 12/24(2006.01)

H04L 12/26(2006.01)

G06N 3/04(2006.01)

G06N 3/08(2006.01)

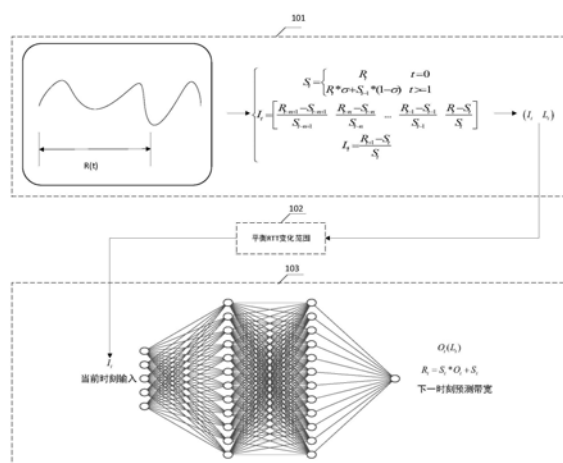
权利要求书1页 说明书4页 附图2页

## (54)发明名称

一种基于RNN神经网络的RTT预测方法

## (57)摘要

本发明公开了一种基于RNN神经网络的RTT预测方法。该方法的步骤为：首先测量得到RTT数据并按照时间先后排序；然后提取RTT数据的特征并转换为神经网络的输入数据与标签数据，对数据按照RTT数据特征进行分类，并平衡各分类中数据的个数；建立RNN神经网络模型，该模型利用历史时刻的RTT数据特征作为输入，其输出是未来时刻的RTT数据；接着对神经网络进行训练；最后利用训练后的神经网络将实时测量的RTT数据的特征转换得到预测的RTT数据。本发明提出了一种量化反映未来RTT变化的方法，在高测量间隔的条件下，弥补单纯测量RTT不足以反映未来RTT变化的缺陷，能有效指导互联网应用应对复杂网络变化。



1. 一种基于RNN神经网络的RTT预测方法,其特征在于,该方法的具体步骤为:

步骤1,测量得到RTT数据,并按照时间先后排序;

步骤2,提取RTT数据的特征并转换为神经网络的输入数据与标签数据,形成数据集;对数据集中的数据按照RTT数据特征进行分类,并平衡各分类中数据的个数;

步骤3,建立RNN神经网络模型用于训练与预测,所述神经网络模型利用历史时刻的RTT数据特征作为输入,其输出是未来时刻的RTT数据;

步骤4,从步骤2平衡后的数据集中分出一批数据作为训练数据,另一批数据作为交叉验证数据;从训练数据中随机选择N个数据作为神经网络的输入数据;重复迭代至误差满足需求或者迭代达到一定次数;

步骤5,使用交叉验证数据对神经网络进行验证,若效果满足需求则训练完成;

步骤6,训练完成之后,提取实时测量的RTT数据的特征,输入神经网络计算,经过转换得到预测的RTT数据。

2. 根据权利要求1所述的一种基于RNN神经网络的RTT预测方法,其特征在于,所述步骤2中,提取RTT数据的特征并转换为神经网络的输入数据与标签数据具体步骤如下:

步骤21:利用公式 $S_t | \sigma = \begin{cases} R_t & t = 0 \\ R_t * \sigma + S_{t-1} * (1 - \sigma) & t \geq 1 \end{cases}$ 平滑原始RTT数据,其中 $R_t$ 为原始RTT数据, $S_t$ 为平滑之后的RTT数据, $\sigma$ 为常数;

步骤22:取原始RTT数据减去对应时刻的平滑之后的RTT数据,再除以平滑之后的RTT数据,即 $\frac{R_t - S_t}{S_t}$ ,得到时序RTT特征;

步骤23:取多个连续时刻的时序RTT特征作为神经网络的输入,即 $I_t = [\frac{R_{t-n+1} - S_{t-n+1}}{S_{t-n+1}}, \frac{R_{t-n} - S_{t-n}}{S_{t-n}}, \dots, \frac{R_{t-1} - S_{t-1}}{S_{t-1}}, \frac{R_t - S_t}{S_t}]$ ,其中n为输入时间维度长度;取下时刻的原始RTT数据减去此刻的平滑之后的RTT数据,再除以此刻平滑之后的RTT数据,即 $\frac{R_{t+1} - S_t}{S_t}$ ,作为神经网络的训练标签 $L_t$ 。

3. 根据权利要求2所述的一种基于RNN神经网络的RTT预测方法,其特征在于,常数 $\sigma$ 取0.1。

4. 根据权利要求1至3之一所述的一种基于RNN神经网络的RTT预测方法,其特征在于,所述步骤2中,平衡各分类中数据的个数具体为:丢弃部分数据以保证每个分类中数据量在同一数量级。

5. 根据权利要求1所述的一种基于RNN神经网络的RTT预测方法,其特征在于,RNN神经网络模型至少包括一层RNN层或其变体;采用卷积层、全连接层或者池化层增加/减少特征维度,并采用BN层加速训练,以及采用Dropout层防止网络过拟合。

## 一种基于RNN神经网络的RTT预测方法

### 技术领域

[0001] 本发明涉及网络通信技术领域,特别涉及网络状况预测方法,具体涉及一种基于RNN神经网络的RTT预测方法。

### 背景技术

[0002] 随着互联网的逐渐普及,网络通信技术不断提高,用户对于优质网络的需求日益提高,许多日常互联网应用依赖于RTT (Round-Trip Time) 做出相应的调整策略。为了提高用户的实际体验,需要获知用户实时的RTT。例如,用户玩网络游戏时,服务器可以有多个选择,开发者根据用户与服务器之间的RTT选择中继结点以及转发路径,尽量减少用户玩游戏时的延时,增加操作连贯性,提高用户体验。

[0003] 目前主流的RTT测量的方法是服务器或者客户端发送一个带时间戳的RTT探测包,另一方收到RTT之后回复RTT探测包,最终服务器或客户端将当前时间减去RTT探测包内时间得到RTT。

[0004] RTT在用户网络状况良好时都是稳定的,随时间推移不会有太大改变。当用户网络状况较差时,RTT随时间变化大。同时,客户端密度较高时,由于数据包的优先度问题,低探测间隔会带来RTT探测误差,因此需要提高RTT的探测间隔也就意味着探测得到的RTT是当前时刻或之前一段时间内的RTT,无法体现未来一段时间内的RTT变化。因此,现有RTT探测方法的主要缺陷是在高测量间隔的条件下,单纯测量RTT以及采用滤波处理的方式,这不足以反映未来RTT变化,将使得基于RTT变化所做的应用策略受到影响甚至失效。

[0005] 基于神经网络的RTT预测方法对于训练方法有一定的要求,具体体现在训练集的选取以及神经网络模型结构两个方面。在训练集的选取方面,RTT数据在大部分网络空闲的情况下都是稳定的,变化幅度较小,若直接采用这部分数据进行训练,很容易导致神经网络过拟合,在网络状态良好时预测效果较好,在网络状态较差时,预测效果不是很理想。在神经网络模型方面,RNN网络可以提取序列中随时间变化的动力特征,因此其在学习和测试序列具有不同长度时也可以有稳定的表现,在整体上要优于非RNN网络。

### 发明内容

[0006] 针对上述现有技术存在的缺陷,本发明的目的是提供一种基于RNN神经网络的RTT预测方法,可以预测未来一段时间内RTT的变化。

[0007] 为了实现上述发明目的,本发明采用的技术方案如下:

[0008] 一种基于RNN神经网络的RTT预测方法,具体步骤为:

[0009] 步骤1,测量得到RTT数据,并按照时间先后排序;

[0010] 步骤2,提取RTT数据的特征并转换为神经网络的输入数据与标签数据,形成数据集;对数据集中的数据按照RTT数据特征进行分类,并平衡各分类中数据的个数;

[0011] 步骤3,建立RNN神经网络模型用于训练与预测,所述神经网络模型利用历史时刻的RTT数据特征作为输入,其输出是未来时刻的RTT数据;RNN神经网络模型至少包括一层

RNN层或其变体,采用卷积层、全连接层或者池化层增加/减少特征维度,并采用BN层加速训练,以及采用Dropout层防止网络过拟合;

[0012] 步骤4,从步骤2平衡后的数据集中分出一批数据作为训练数据,另一批数据作为交叉验证数据;从训练数据中随机选择N个数据作为神经网络的输入数据;重复迭代至误差满足需求或者迭代达到一定次数;

[0013] 步骤5,使用交叉验证数据对神经网络进行验证,若效果满足需求则训练完成;

[0014] 步骤6,训练完成之后,提取实时测量的RTT数据的特征,输入神经网络计算,经过转换得到预测的RTT数据。

[0015] 进一步地,所述步骤2中,提取RTT数据的特征并转换为神经网络的输入数据与标签数据具体步骤如下:

[0016] 步骤21:利用公式 $S_t|\sigma = \begin{cases} R_t t = 0 \\ R_t * \sigma + S_{t-1} * (1 - \sigma) t \geq 1 \end{cases}$ 平滑原始RTT数据,其中 $R_t$ 为原始RTT数据, $S_t$ 为平滑之后的RTT数据, $\sigma$ 为常数;

[0017] 步骤22:取原始RTT数据减去对应时刻的平滑之后的RTT数据,再除以平滑之后的RTT数据,即 $\frac{R_t - S_t}{S_t}$ ,得到时序RTT特征;

[0018] 步骤23:取多个连续时刻的时序RTT特征作为神经网络的输入,即

$I_t = [\frac{R_{t-n+1} - S_{t-n+1}}{S_{t-n+1}}, \frac{R_{t-n} - S_{t-n}}{S_{t-n}}, \dots, \frac{R_{t-1} - S_{t-1}}{S_{t-1}}, \frac{R_t - S_t}{S_t}]$ ,其中n为输入时间维度长度;取下时刻的

原始RTT数据减去此刻的平滑之后的RTT数据,再除以此刻平滑之后的RTT数据,即 $\frac{R_{t+1} - S_t}{S_t}$ ,作为神经网络的训练标签 $L_t$ 。

[0019] 本发明提出了一种量化反映未来RTT变化的方法,在实时测得用户RTT的基础上,将RTT数据的变化特征作为输入,通过RNN神经网络计算,预测出未来一段时间内的RTT。在高测量间隔的条件下,本发明可以弥补单纯测量RTT不足以反映未来RTT变化的缺陷,能有效指导互联网应用应对复杂网络变化。

## 附图说明

[0020] 图1为本发明方法的流程图;

[0021] 图2为本发明实施例中四层神经网络结构图。

## 具体实施方式

[0022] 下面结合实例对本发明作进一步的详细说明,有必要指出的是,以下的实例只用于对本发明做进一步的说明,不能理解为对本发明保护范围的限制,所属领域技术熟悉人员根据上述发明内容,对本发明做出一些非本质的改进和调整进行具体实施,应仍属于本发明的保护范围。

[0023] 图1展示了本发明基于RNN神经网络的RTT预测方法的流程图。过程101用于提取RTT的时序特征,并将测量RTT记录转换为神经网络的输入数据与标签数据,放入数据集。过程102根据需求划分数据集,保证每个分类的数据量在一个数量级。过程103使用神经网络进行训练,神经网络训练完成后,输入RTT数据经转换得到的时序特征,经过神经网络计算

与数学换算之后得到下一时刻的RTT预测值。

[0024] 本实施例中,过程101包括的具体步骤如下:

[0025] 步骤1:将测量得到的RTT记录整合、采样、清洗之后按照时间先后排序。取RTT数据,按照公式 $S_t|\sigma = \begin{cases} R_t & t = 0 \\ R_t * \sigma + S_{t-1} * (1 - \sigma) & t \geq 1 \end{cases}$ 平滑原始RTT数据,其中 $R_t$ 为原始RTT数据, $S_t$ 为平滑之后的RTT数据。 $\sigma$ 取0.1。对数据平滑之后,能够得到相对平稳的数据,降低数据噪声,减少噪声对模型拟合的影响。

[0026] 步骤2:取原始RTT数据减去对应时刻 $t$ 的平滑之后的RTT数据,再除以平滑之后的RTT数据,即 $\frac{R_t - S_t}{S_t}$ ,得到时序RTT特征。

[0027] 步骤3:神经网络的输入数据为: $I_t = [\frac{R_{t-n+1} - S_{t-n+1}}{S_{t-n+1}}, \frac{R_{t-n} - S_{t-n}}{S_{t-n}}, \dots, \frac{R_{t-1} - S_{t-1}}{S_{t-1}}, \frac{R_t - S_t}{S_t}]$ ,其中 $n$ 为输入时间维度长度,本实施例取10个连续时刻的时序RTT特征作为神经网络的输入;取下时刻的原始RTT数据减去此刻的平滑之后的RTT数据,再除以此刻平滑之后的RTT数据,即 $\frac{R_{t+1} - S_t}{S_t}$ ,作为神经网络的训练标签 $L_t$ 。

[0028] 本实施例中,过程102的具体步骤如下:

[0029] 步骤1:得到过程101输出的数据集。

[0030] 步骤2:样本的RTT数据变化范围不是均匀的,采用深度学习的方式必须要保证各个标签的样本数量不能相差太大,本实施例采用的是根据RTT的变化率平衡样本数量,同时将样本数量较多的标签中的数据丢掉一些。这里,根据 $\frac{R_{t+1} - S_t}{S_t}$ ,即下一时刻的RTT相对于此刻RTT的平滑值的变化率,将相对变化率的绝对值按照区间 $[0\%, 2\%)$ ,  $[2\%, 8\%)$ ,  $[8\%, 15\%)$ ,  $[15\%, \infty)$ 进行划分,记录五个区间内数据数量的最小值,记为最小样本数。将五个区间内数据分别随机采样,保证每个区间内数据数量为最小样本数。

[0031] 步骤3:将采样之后的数据集作为输出。

[0032] 通过神经网络计算预测RTT的具体步骤如下:

[0033] 步骤1:将过程102得到的数据集输入神经网络。

[0034] 步骤2:将神经网络的输出转换为下一时刻RTT的预测值,即 $R_{t+1} = S_t * 0_t + S_t$ ,其中 $R_{t+1}$ 为下一时刻的RTT预测值, $S_t$ 为这一时刻的RTT经过平滑之后的数值, $0_t$ 为神经网络的输出。

[0035] 时序排列的数据往往历史的输入与未来的输入具有一定的相关性,对于同样是时序序列的RTT数据,以单一时刻提取的RTT特征或者单一时刻提取的历史数据变化特征,不足以反映RTT整体的变化趋势。简单的BP神经网络只能找到某一特定时刻下,RTT特征与下一时刻的预测RTT的关联。RNN神经网络能够很好地处理时序数据,建立起历史记录之间的关系,更好地预测RTT数据。因此,本实施例的神经网络采用RNN神经网络,其具体结构如图2所示,网络具有四层结构,由LSTM(Long-Short-Memory)层、全连接层、Dropout层构成,具体结构如下:

[0036] 神经网络的输入为时间维度10、特征维度为1的RTT序列。

[0037] 第一层为LSTM层,隐层结点数为32,用于提取输入的时间序列的特征,并存储一部

分长期和短期的特征进网络里。

[0038] 第二层为全连接层,隐层结点数为64,用于将LSTM层提取的低维时间序列特征映射到高维空间。

[0039] 第三层为Dropout层,Drop概率为0.5,用于随机丢弃全连接层隐层结点,避免网络过拟合,提高神经网络的泛化能力。

[0040] 第四层为隐层结点数为1的全连接层,作为输出层,用于将高阶特征转换为预测的未来网络RTT,其输出为下一时刻的RTT特征。

[0041] 值得一提的是,提取的RTT特征可以替换为任何形式的特征,例如RTT变化的频率,RTT的变化率等。神经网络可以为包括任何形式的RNN神经网络的组合,此处的RNN神经网络应包括诸如LSTM (Long-Short-Memory),GRU (Gate-Recurrent-Gate),BRNN (Bidirectional-Recurrent-Neural-Networks)等一系列RNN网络变体,且网络模型并不局限于本实施例中给出的四层神经网络。神经网络输出到预测RTT的转换公式也不局限于本实施例给出的转换公式。

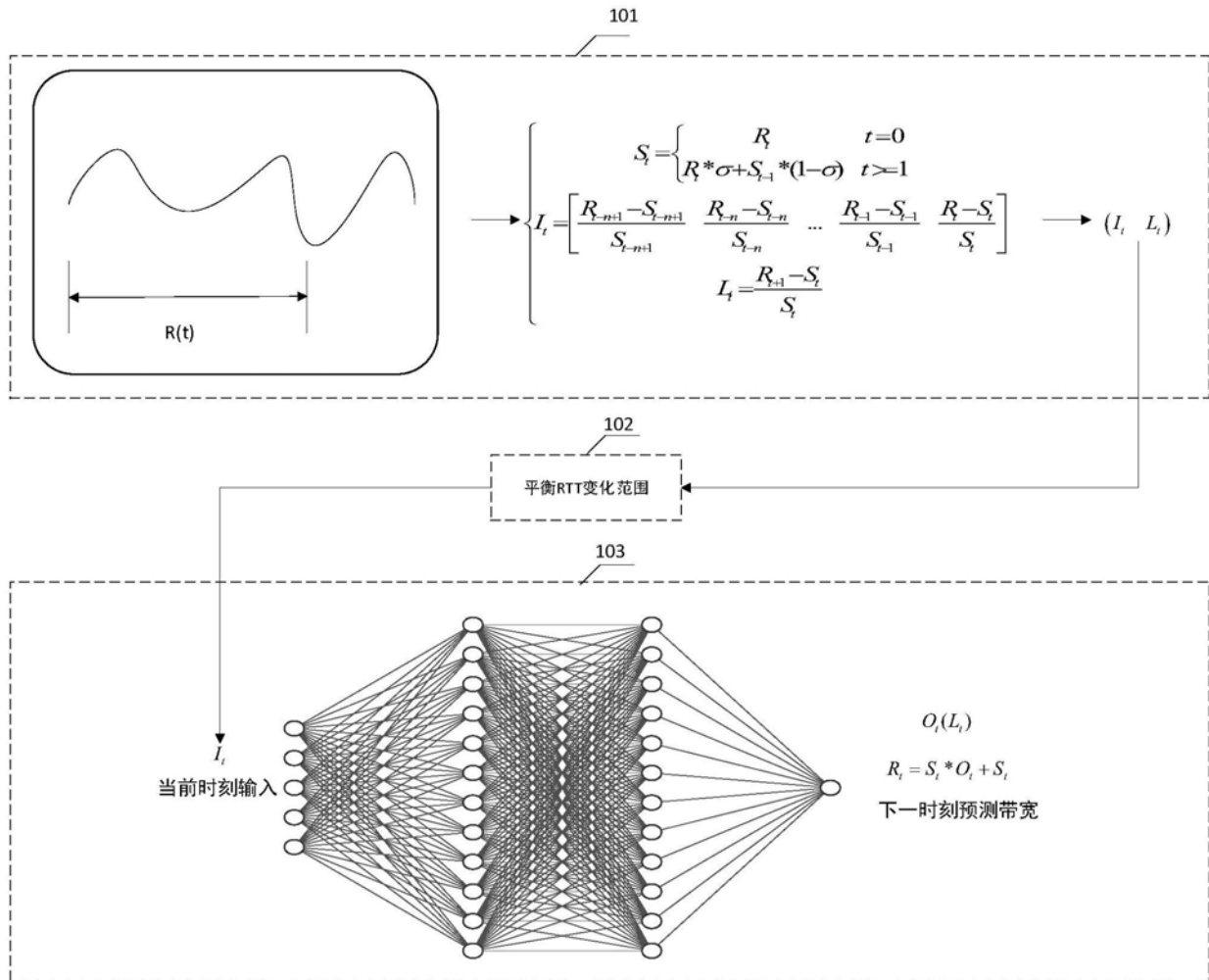


图1

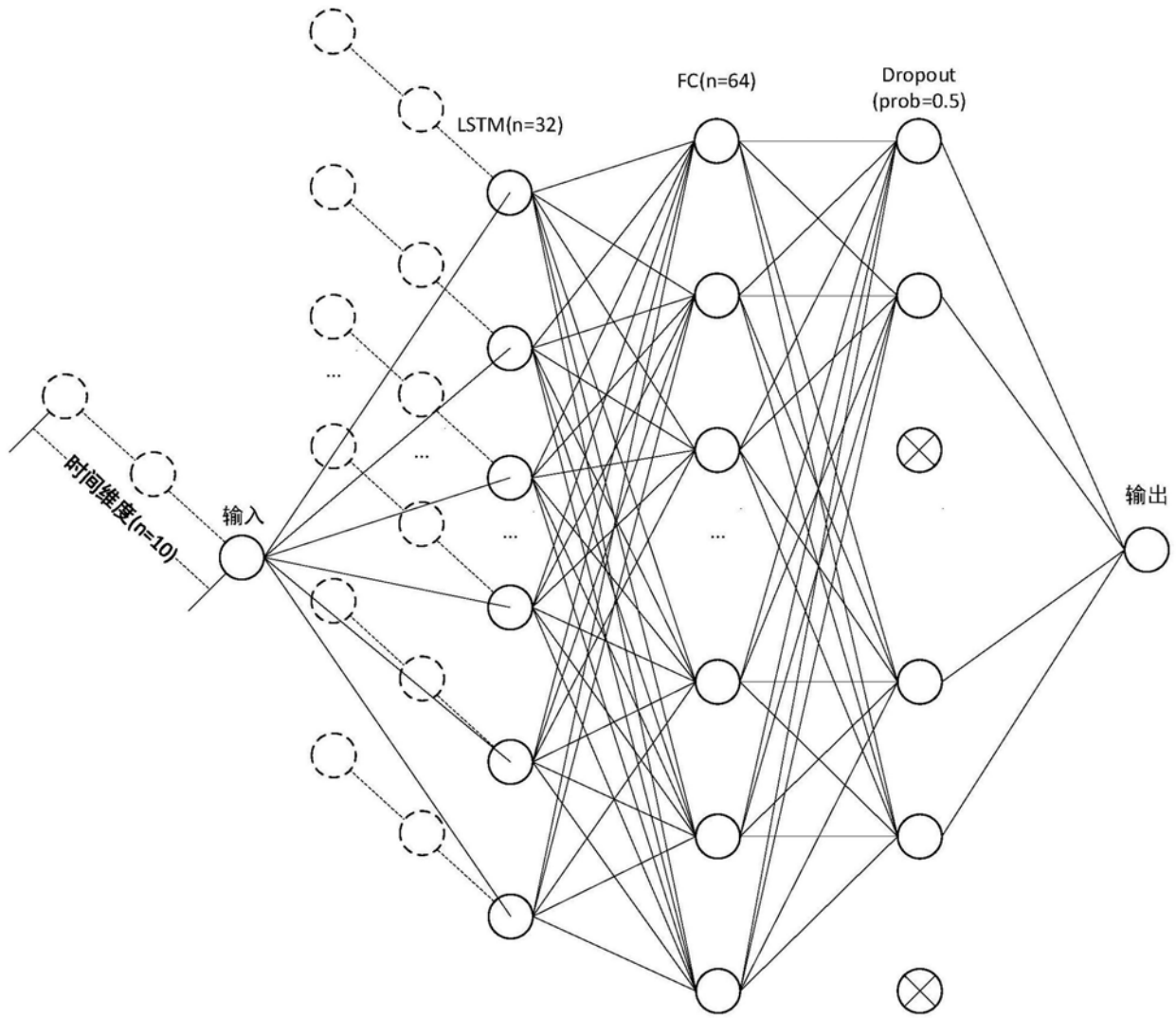


图2