

网络设备工作负载综合预测研究

张 靖

(攀枝花学院 网络中心, 四川 攀枝花 617000)

摘 要: 充分利用现有成熟预测模型的优势, 取长补短, 合理选取预测模型, 结合负载影响因素, 串联或者并联各模型输入或者输出, 综合预测, 进行网络设备动态负载预测. 综合利用灰色和 BP 神经网络预测模型, 并对其算法加以改进来提高预测准确度, 分析并提出了一种综合预测模型, 给出了具体的预测方法以及过程, 并对预测结果进行分析, 提高了预测精度, 达到了预测目的. 实际应用证明, 该模型及方法在网络设备负载定量预测方面具有良好的效果, 有较好的参考和使用意义.

关键词: 算法改进; 综合预测; 网络设备; 预测模型; 工作负载

中图分类号: TP393.07

文献标识码: A

文章编号: 1000-7180(2010)11-0151-03

The Study on the Comprehensive Prediction of the Network Equipment Workload

ZHANG Jing

(Campus Network Center, Panzhihua University, Panzhihua 617000, China)

Abstract: By making full use of the advantages of the existing prediction model and selecting reasonably the prediction model, the dynamic workload of the network equipment will be predicted with the combination of the influential factors of the workload and the connection of the models in series or parallel to input or output the result. With the comprehensive utilization of the grey prediction and the BP neural network prediction model and the algorithm improvement of the prediction accuracy, this article analyzes and presents a comprehensive prediction model and gives the specific method and process of the prediction. The analysis of the result of the prediction has proved that this model has improved the prediction accuracy. This model and method work well in the quantitative prediction of the network equipment workload.

Key words: algorithm improvement; comprehensive prediction; network equipment; prediction model; workload

1 引言

计算机网络作为当今社会重要通信平台, 支撑其运行的网络设备成为十分关键和重要的部分, 网络设备的可靠性、负载等已成为重要的研究方向和内容. 开展网络设备动态负载预测研究, 提前定量预测负载变化情况, 防止设备过载, 保证运行正常, 具有重要的应用价值和意义.

网络设备工作负载是支撑网络应用的能力指标, 反映工作运行状态, 在设计负载范围内能够正常运行. 一旦超过就无法保证正常运行. 测试或者反映

设备工作负载的指标和因素很多, 如 CPU、内存、流量、时延、用户数、应用进程数、数据包大小、网络丢包率、网络协议等, 目前还没有统一的工作负载计算方法. 由于影响负载的因素较多, 且随时发生变化, 影响程度也不一样, 采用加权思想, 设备工作负载^[1] 指标因素 $\{a_1, a_2, \dots, a_n\}$, 权值 $\{q_1, q_2, \dots, q_n\}$, 则工作负载为 $G = \sum_{i=1}^n a_i q_i$. 网络流量经常作为网络设备负载动态情况的标志性特征指标, 以下根据设备工作相关因素来预测网络流量, 及时反映设备动态负载, 也可作为负载均衡处理策略^[2] 的参考

收稿日期: 2009-09-27; 修回日期: 2009-12-08

基金项目: 攀枝花学院科研重点项目(JJ0830); 攀枝花市科技项目(2009CY-R-13)

依据.

2 预测模型

国内外对预测开展了很多研究,如灰色预测、神经网络预测、回归分析、概率统计、Markov 模型等,但是还没有形成普遍适用的方法,各有特点,应用范围、前提条件以及预测精度不一样.设备工作负载动态预测有很多因素影响,复杂程度不同,也难用一个单独的预测模型达到较好的预测效果.因此,结合多种预测模型,取长补短,综合预测,研究网络设备动态工作负载.各模型输出形成综合预测结果,常采用最小二乘法来确定各单项模型的权系数^[1].

灰色系统建模方法简单,但没法并行计算,不能充分考虑序列之间的关联,基于累加生成和最小二乘法,原始数据列累加处理后呈现明显的指数规律,在弱化原始数据随机性的基础上,容易找出数据变化规律,使得神经网络中的非线性激励函数易于逼近,可修正灰色模型的结果^[2].综合利用灰色理论与神经网络建立预测模型的思路^[3-4],提出了一种根据灰色理论和人工神经网络预测模型,加以改进,形成综合预测模型以及方法.

3 预测方法

3.1 灰色预测的改进

灰色预测应用最广泛的是 GM(1,1)模型,在实际应用中,其使用范围与发展系数 a 相关, $|a| \leq 2$ 才有使用价值^[5].同时,它没有考虑随着时间推移相继进入系统的扰动或驱动因素的影响,预测精度较高的是未来 1~2 个数据,越是远离时间原点,预测意义就越弱.鉴于此,采用滑动窗口^[6]的机制进行数据更新,将老数据的信息意义逐步降低,及时不断补充新信息的数据,建立改进的预测模型.

预测方法如下:

(1) 设原始数据 $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ 滑动窗口大小 $T, T < n, T$ 的确定和预测精度有关系,可参考预测精度和经验设定.

(2) 用原始数据 $X^{(0)}$ 建立 GM(1,1) 模型,用该模型预测出第 $n+T$ 个数据 $Y^{(0)} = (y^{(0)}(n+1), y^{(0)}(n+2), \dots, y^{(0)}(n+T))$;

(3) 置入新信息 $Y^{(0)}$, 去掉老信息.用 $X^{(0)} = (x^{(0)}(T+1), x^{(0)}(T+2), \dots, x^{(0)}(n), x^{(0)}(n+1), \dots, x^{(0)}(n+T))$ 建立一新的 GM(1,1) 模型,预测出结果再加入预测数据 $x^{(0)}$, 并去掉相应原始数据;类似做法直到预测出所需全部数据.

3.2 神经网络的改进

BP 网络预测通用性强且应用广泛,但是 BP 网络也存在许多不足之处需要改进^[7],主要是采用的学习步长是一个定值,在网络的训练初期如果学习步长取得比较合适,它会使网络的误差不断减小,当网络在不断训练过程中,学习步长已经不能满足要求,导致网络在训练过程中出现振荡.需要对学习步长不断调整,改进神经网络,优化预测.

BP 神经网络进行改进:

X_i 为输入, Y_i 为输出, T_i 为期望输出, η 为学习率, $f(x)$ 为作用函数,选择 Sigmoid() 为转移函数, w_{ij} 是第 i 个单元到第 j 个单元连接的权系数,则神经网络正向传播:

$$r_j = f(\sum_{i=1}^n w_{ij} X_i), f(x) = 1/(1 + e^{-ax})$$

$$\text{导函数: } f'(x) = af(x)(1 - f(x))$$

$$\text{误差为: } E(n) = \sum_{i=1}^n (T_i - Y_i)^2$$

学习过程: $W_{ij}(n+1) = W_{ij}(n) + \eta \delta_j X_i$, 其中 δ_j 为函数 $f(x)$ 在 $X^{(j)}$ 的梯度,如果输出误差 $E(n+1) < E(n)$,保留误差及权值,增大学习率 η 为 1.1η ,网络权重更新为 $W_{ij}(n+1) = W_{ij}(n) + 1.1\eta \delta_j X_i$,判断新的误差比原来小,继续增大学习率,直到该梯度方向搜索的误差不再减少为止,依次保留误差及对应的权值.如果输出误差 $E(n+1) > E(n)$,保留误差及权值,减少学习率 η 为 0.5η ,网络权重更新为 $W_{ij}(n+1) = W_{ij}(n) + 0.5\eta \delta_j X_i$,判断新的误差比原来大,继续减少学习率,直到新点的误差比保留的误差还小,则替换原保留的权重值继续网络训练.

4 综合预测

综合预测就是充分利用多种成熟模型各自优势,有效组合,综合预测.模型之间主要是并联、串联的方式.并联是灰色模型、神经网络模型预测结果的线性加权值作为预测结果,却没有考虑到预测结果之间的非线性关系.串联是将多个灰色模型的预测结果作为神经网络模型的输入,然后神经网络预测,输入忽略了其它影响因素对预测结果的影响.因此,提出综合预测将灰色模型的预测结果和影响预测指标的主要因素同时作为神经网络的输入,来实现综合预测,如图 1 所示.以网络路由交换机为例进行流量综合预测,采用 MATLAB 软件进行仿真.

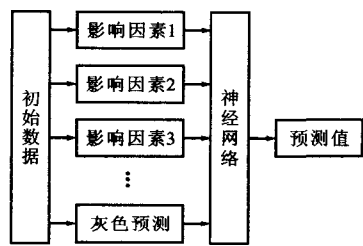


图1 综合预测模型

4.1 数据准备

n 个输入 $r_1, r_2, \dots, r_n$, 其中前 $n - 1$ 个输入为影响设备工作负载的主要因素, 第 n 个输入为灰色模型的预测结果, 隐含层共有 m 个节点, 第 j 个节点参数向量为 $[w_{j1}, w_{j2}, \dots, w_{jm}, \theta_j]$, 一个输出节点 y .

4.2 灰色模型预测

(1) 数据序列光滑性检验
初始数据序列 $x = (x(1), x(2), \dots, x(n))$, $\zeta_k = x(k - 1)/x(k)$, 称 ζ_k 为序列 x 的级比偏差^[4], 须满足 $\zeta_k \in (e^{(-2/(n+1))}, e^{(2/(n+1))})$, 否则进行数据变换来提高预测精度, 例如对数变换、指数变换、幂函数变换等.

(2) 改进的灰色预测
设置数据更新窗口 $T = 2$, 第一次预测 2 个新数据, 然后去掉初始的第 1 和 2 个数据, 再加入预测产生的 2 个数据开始预测, 产生新的数据后按照前述方法继续预测, 得预测数据.

4.3 神经网络结构设计

网络输入包括预测指标主要因素为 3 类, 每类 $R(i) = \{r_{i1}, r_{i2}, \dots, r_{in}\}$, 灰色预测结果为 $p = \{p_1, p_2, \dots, p_n\}$, 实际值为 $T = \{t_1, t_2, \dots, t_n\}$; 网络输出(即预测值) $Y = \{y_1, y_2, \dots, y_k\}$, 其中 $i \in (1, 3), n \in (1, 22), k \in (1, 5)$.

网络采用三层 BP 网络, 网络训练采用 Levenberg - Marquardt, 输入层采用 5 个神经元, 输出层采用 1 个神经元, 隐含层采用 12 个神经元.

4.4 结果分析

4.4.1 预测结果

初始数据如表 1 所示. 采用灰色模型、神经网络和综合模型预测, 结果见表 2 和图 2 所示.

4.4.2 精度检验

采用平均残差来衡量预测的精度. 平均残差分别为: 采用灰色预测为 0.1058, BP 神经网络为 0.1318, 综合预测为 0.0647.

表 1 初始数据

序号	内存使用 (%)	接口利用 (%)	连接数 (1×10^4)	流量($100M$) (1×10^4)
1	21.90	35.65	4.8495	2.9294
2	21.95	36.10	4.9062	3.01375
3	20.95	36.35	4.9699	2.97235
4	20.30	36.85	5.0431	3.10335
5	19.90	37.10	5.1165	3.302
6	19.35	37.65	5.1842	3.5452
7	19.15	37.90	5.2552	3.8341
8	19.30	37.90	5.3339	4.0425
9	19.00	38.10	5.4202	4.331
10	19.20	38.10	5.5081	4.64985
11	19.10	38.00	5.5932	4.8467
12	18.35	38.10	5.6759	4.93515
13	18.55	38.05	5.7539	5.18915
14	19.10	37.85	5.8248	5.4585
15	20.10	37.35	5.8922	5.79965
16	20.20	37.50	5.9592	6.13685
17	20.55	37.30	6.0243	7.5588
18	20.70	37.35	6.0877	6.9474
19	20.85	35.85	6.1504	6.90865
20	20.15	34.80	6.2097	6.80
21	20.00	34.55	6.2637	8.19155
22	20.15	33.90	6.3132	6.92765
23	19.90	33.35	6.3592	9.15995
24	19.70	33.15	6.4020	9.58985
25	20.25	34.20	6.4420	10.2495
26	20.40	34.00	6.4804	10.66135
27	21.00	34.45	6.5186	11.16595

表 2 预测结果数据

序号	实际值	灰色	神经网络	综合预测值
1	91599.5	82668	78521.88	99887.06
2	95898.5	86574	83194.22	104463.5
3	102495.0	90665	96115.43	99850.9
4	106613.5	94949	97762.20	102288.5
5	111659.5	99435	85020.17	103021.26

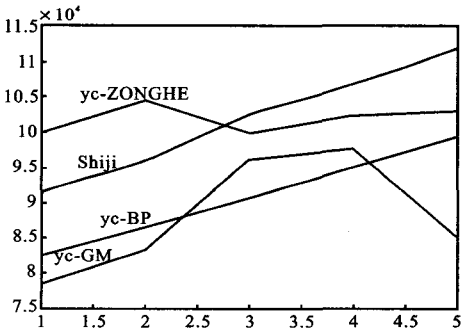


图2 预测数据结果图

(下转第 157 页)

- [2] Zhang Y, Skolnick J. Spicker: approach to clustering protein structures for near-native model selection[J]. Journal of Computational Chemistry, 2004(25):865-871.
- [3] 张昕, 彭宏, 郑启伦. 基于微粒群算法的聚类分析[J]. 微电子学与计算机, 2006, 23(9): 94-95.
- [4] 高尚, 汤可宗, 杨静宇. 一种新的基于混合蚁群算法的聚类方法[J]. 微电子学与计算机, 2006, 23(12):38-40.
- [5] Brendan J Frey, Delbert Dueck. Clustering by passing messages between data points[J]. Science, 2007, 315(5814): 972-976.
- [6] Yang Zhang. I-tasser server for protein 3d structure prediction[J]. BMC Bioinformatics, 2008(9):40-47.
- [7] Andriy Kryshchak, Maciej Milosian, Lukasz Szajkowski, et al. Casp6 data processing and automatic evaluation at the protein structure prediction center[J]. Proteins: Structure, Function, and Bioinformatics, 2005(7):19-23.
- [8] Sitao Wu, Jeffrey Skolnick, Yang Zhang. Ab initio model-

ing of small proteins by iterative tasser simulations[J]. BMC Bioinformatics, 2007(5):17-26.

- [9] Machael Tress, Iakes Ezkurdia, Osvaldo Grana, et al. Assessment of predictions submitted for the casp6 comparative modeling category[J]. Proteins: Structure, Function, and Bioinformatics, 2005(7):27-45.
- [10] Yang Zhang, Jeffrey Skolnick. Scoring function for automated assessment of protein structure template quality[J]. Proteins: Structure, Function, and Bioinformatics, 2004, 57(4):702-710.

作者简介:

黄旭 男,博士研究生.研究方向为智能信息处理新技术.

吕强 男,教授,博士生导师.研究方向为生物信息计算、元启发搜索、并行分布计算.

杨凌云 男,硕士研究生.研究方向为智能信息处理新技术.

钱培德 男,教授,博士生导师.研究方向为中文信息处理理论.

(上接第 153 页)

4.4.3 结果评估

通过表 2 和图 2 可见,采用设计的综合预测模型和算法,能够实现预测,预测数据精度得到提高,满足预测目的和要求,具有实际的应用价值.

5 结束语

网络设备工作负载由于工作环境、影响因素等复杂性,没有具体标准,特别是定量预测、计算和分析,不容易准确预测.根据文中提出的综合预测模型,采用给出的具体预测方法及步骤,实际应用预测效果良好,为网络设备动态工作负载定量预测提出了参考方法,在网络建设、管理和应用工作中有一定的实用价值.

参考文献:

- [1] 元雪冬,李霞,梁鸿,等. 计算网格中动态负载平衡策略

研究[J]. 微电子学与计算机, 2009, 26(5):21-24.

- [2] 邹心遥,姚若河. 基于 BP 神经网络的电子元器件寿命预测[J]. 微电子学与计算机, 2009, 26(1):52-53.
- [3] 曹建华,刘渊,戴悦. 一种基于灰色神经网络的网络流量预测模型[J]. 计算机工程与应用, 2008, 44(5): 155-157.
- [4] 曹建华. 基于灰色神经网络模型的网络流量预测算法研究[D]. 无锡:江南大学, 2009.
- [5] 李福琴. 灰色模型的稳定性和建模精度研究[D]. 武汉: 武汉理工大学, 2007.
- [6] 胡春光,高燕,李颖. 一种扩展滑动窗口算法[J]. 微电子学与计算机, 2007, 24(8):106-112.
- [7] 吴昌友. 神经网络的研究及应用[D]. 哈尔滨:东北农业大学, 2007.

作者简介:

张靖 男, (1972-), 硕士, 副教授. 研究方向为软件工程和计算机网络.