

The Study on the Prediction of Dynamic Workload of the Network Equipment

Jing ZHANG

Campus Network Center, Panzihua University, Panzihua 617000, China

E-mail: zjpzh@tom.com

Abstract: The prediction of the workload of the network equipment can help to keep abreast of the operation of the equipment, ensure the stable operation of the network and improve the operation and the management level. A comprehensive prediction method is presented after the analysis and study of the dynamic workload of the network equipment. With the comprehensive utilization of the grey prediction and the BP neural network prediction model and the prediction accuracy improved after the change of the algorithm, a comprehensive prediction model is presented, and the specific method and its process of the prediction are given. By using computer simulation, the result of the prediction is analyzed, and it proves that this model can greatly improve the prediction accuracy. The model and the method work well in the quantitative prediction of the network equipment workload.

Keywords: algorithm improvement; comprehensive prediction; network equipment; prediction model; workload

网络设备工作动态负载预测研究

张 靖

攀枝花学院网络中心, 攀枝花, 中国, 617000

E-mail: zjpzh@tom.com

摘 要: 对网络设备工作负载进行预测可以全面掌握设备运行情况, 保障网络稳定运行, 提高网络运行和管理水平。经过分析和研究网络设备动态工作负载的指标因素, 对比研究相关预测模型, 提出了一种综合预测方法。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, 设计了一种综合预测模型, 给出了具体的预测方法以及过程, 采用计算机模拟和仿真, 对预测结果进行分析, 预测精度得到了较好的提高, 达到了预测目的。实际应用证明得出结论, 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, 有较好的参考和实用意义。

关键词: 算法改进; 综合预测; 网络设备; 预测模型; 工作负载

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。计算机网络建设之初, 往往需要进行用户数、系统容量等工作负载方面的设计, 满足在负载之内的正常运, 主要是参考相关历史的、静态的数据进行设计。但由于目前 IPV4 协议本身、网络资源竞争模式、开放性等存在制约因素, 同时各种计算机网络病毒影响, 时常导致超过设备本身负载能力, 使网络传输十分缓慢, 甚至中断, 因此, 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。负载可能有多种因素影响, 一些预测针对一种影响因素, 有的预测受到模型本身限制, 预测范围和精度也不高, 因此, 提出一种利用多种模型优势综合预测的方法来进行网络设备负载^[1], 为网络设计和网络运行管理提供重要参考。

2 网络设备工作负载

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也很多, 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、网络协议等等, 目前还没有统一的工作负载计算

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)

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]的参考依据。

3 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。

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。综合利用灰色理论与神经网络^[3]建立预测模型思路^[4,5]，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。

4 预测方法

4.1 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^[6]。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^[7]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。

4.2 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^[8]，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：

$$Y_j = f\left(\sum_{i=1}^n W_{ij} X_i\right), f(x) = 1 / (1 + e^{-ax})$$

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

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

学习过程： $W_{ij}(n+1) = W_{ij}(n) + \eta \cdot \delta_j X_i$ ，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 $W_{ij}(n+1) = W_{ij}(n) + 1.1\eta \cdot \delta_j X_i$ ，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 $W_{ij}(n+1) = W_{ij}(n) + 0.5\eta \cdot \delta_j X_i$ ，判断新的误差比原来大，继续减少学习率，直到新点的误差比我们保留的误差还小，则替换原保留的权重值继续网络训练。

5 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。

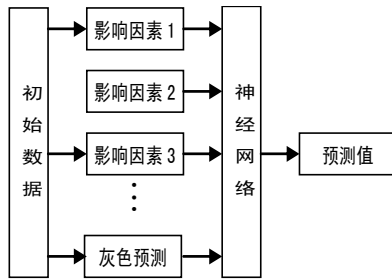


Figure 1. Comprehensive prediction model
图 1. 综合预测模型

5.1 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 个节点的参数向量为 $[w_{j1}, w_{j2}, \dots, w_{jm}, \theta_j]$ ，一个输出节点 y 。

5.2 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[9]，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

不妨设置数据更新窗口 $T=2$ ，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，将得预测数据为 $p = \{29294, 30093, 31552, 33090, 34710, 36418, 38218, 40115, 42117, 44228, 46455, 48806, 51287, 53906, 56673, 59595, 62682, 65944, 69392, 73038, 76893, 80971, 85286, 89851, 94086, 99083, 102750\}$ 。

5.3 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。

5.4 结果分析

5.4.1 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。

Table 1. Initial data
表 1. 初始数据

序号	内存使用百分比	接口利用百分比	连接数 (1×104)	流量 (100M) (1×104)
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

Table 2. Prediction result
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	1024950	90665	96115.43	99850.9
4	106613.5	94949	97762.20	102288.5
5	111659.5	99435	85020.17	103021.26

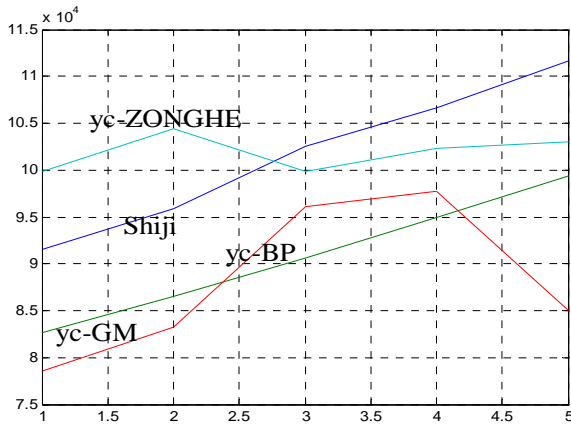


Figure 2. The result map of the prediction data
图 2. 预测数据结果图

5.4.2 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。

5.4.3 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。

6 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。根据本文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。但在具体实践应用过程中，还需要进一步完善和研究，如网络设备工作负载指标和影响因素的确立、预测模型的完善、算法的改进、病毒的影响等。

致谢

本文的研究中，网络设备负载指标测试和负载预测应用得到了周伟、蒋灿两位同事的大力帮助，并采纳了他们相关建议，感谢他们给予的帮助。

References (参考文献)

- [1] TAN Hua, XIE Chi, SUN Bo, et al. Combination Forecasting Model of Securities Market Based on Grey Model and Neural Network[J]. Journal of Hunan University (Natural Sciences). 2007,34(9):86-89 (Ch).
谭华, 谢赤, 孙柏等. 证券市场灰色神经网络组合预测模型应用研究[J]. 湖南大学学报(自然科学版). 2007,34(9):86-89.
- [2] QI Xue-dong; LI Xia; LIANG Hong, et al. Study Dynamic Load Balancing Strategy in Computing Grid[J]. Microelectronics & Computer. 2009, 26(5):21-24 (Ch).
齐雪冬, 李霞, 梁鸿等. 计算网格中动态负载均衡策略研究[J]. 微电子学与计算机, 2009, 26(5):21-24.
- [3] ZOU Xin-yao, YAO Ruo-he. Lifetime Prediction of Electronic Devices Based on Forecast System of Back Propagation Neural Network[J]. Microelectronics & Computer, 2009, 26(1):52-53 (Ch).
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.
- [4] CAO Jian-hua, LIU Yuan, DAI Yue. Network traffic prediction based on grey neural network integrated model. [J]. Computer Engineering and Applications, 2008, 44(5):155-157 (Ch).
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.
- [5] CAO Jian-hua. Network Traffic Prediction Based on Grey Neural Network Model[D]. Wuxi: Jiangnan University, 2009 (Ch).
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.
- [6] LI Fu-qin. Study on the Stability and the Modeling Precision of Grey Model[D]. Wuhan: Wuhan University of Technology, 2007 (Ch).
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.
- [7] HU Chun-guang, GAO Yan, LI Ying. An Extended Sliding Window Method[J]. Microelectronics & Computer, 2007,24(8):106-112 (Ch).
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.
- [8] WU Chang-you. The Research and Application on Neural Network[D]. Harbin: Northeast Agricultural University, 2007 (Ch).
吴昌友. 神经网络的研究及应用[D]. 哈尔滨: 东北农业大学, 2007.