

利用统计特征的网络应用协议识别方法

徐莉¹, 赵曦², 赵群飞¹, 秦涛³,

(1. 上海交通大学图像处理与模式识别研究所, 200240, 上海; 2. 上海金融学院信息管理系, 201209, 上海;

3. 西安交通大学电子与信息工程学院, 710049, 西安)

摘要: 在网络流统计特征的基础上, 提出了一种应用协议识别算法. 根据网络流概念, 在网络层建立应用协议特征的描述方法, 并采用数据量、数据包、时间 3 种属性全面地描述网络协议的特征. 采用主成分分析方法来确定网络流特征属性的主要成分, 以减少环境因素的影响. 结合 BP 神经网络算法建立的网络协议识别模型, 其网络特征具有良好的持久性和稳定性, 模型分类结果也不易受网络环境的影响. 真实网络环境下的实验结果显示, 所提方法能够准确识别目前网络中常用的应用协议, 包括 HTTP、FTP、BitTorrent 及 TELNET, 识别准确率达到 97% 以上.

关键词: 神经网络模型; 网络流; 应用协议; 协议识别

中图分类号: TP393 **文献标志码:** A **文章编号:** 0253-987X(2009)02-0043-05

Network Application Protocol Identification Based on Statistical Methods

XU Li¹, ZHAO Xi², ZHAO Qunfei¹, QIN Tao³

(1. Institute of Image Processing and Pattern Recognition, Shanghai Jiaotong University, Shanghai 200240, China;

2. Department of Information Management, Shanghai Finance College, Shanghai 201209, China; 3. School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: An application protocol identification method is proposed based on the statistical characteristics of network flows. The flow features at the network level are extracted according to the concept of network flow, and three attributes: the number of packets, the number of bytes, and time, are used to capture the flow characteristics roundly. Then the principal component analysis algorithm is used to determine the main characteristics of the flow attributes to reduce the effect of environment. Finally, a BP neural network model is given to identify the application protocols. As the features used in the proposed method are more stable, the output results of the model are hence accurate with the change of the network environment. Experimental results in real network environments show that the proposed method can identify several major application protocols accurately, such as HTTP, BitTorrent, FTP and TELNET, and the identification precision is above 97%.

Keywords: neural network model; traffic flow; application protocol; protocol identification

随着信息化大潮的到来, 计算机网络应用日益推广, 应用协议种类逐渐增加, 加之网络代理和端口的动态变化策略的采用, 使网络管理愈加困难. 网络应用协议的准确识别可以实现网络运行状况的准确监控和网络资源的有效调配, 同时也可提高网络入

侵检测的效率和准确性, 改善网络服务质量, 确保网络安全、高效运行.

最简单的协议识别方法是通过分析网络数据包的端口信息来分析 HTTP 协议^[1], 而现在广泛应用的 P2P 协议所采用的通信端口都是非知名、动态变

化的,因此基于端口的网络协议识别技术逐渐失去效力.通过对数据包负载的分析也可实现网络应用协议的划分,这种方法虽然识别准确率较高,但计算复杂^[2].为了适应 Internet 流量数据庞大、应用属性动态变化的特点,根据数据包头统计特征进行应用协议识别已成为当前计算机网络领域一个新的研究热点.数据包头的统计特征可以用网络流描述,它是一组具有相同五元组(源 IP、目的 IP、源端口、目的端口、传输层协议)的数据包序列,如图 1 所示.

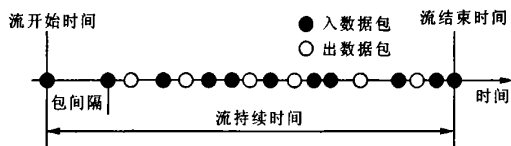


图1 网络流描述

网络流由一系列流量属性构成,包括数据包大小、数据包方向、数据包间隔和网络流持续时间等. Moore 等人通过直接测量和利用傅立叶变换技术给出了网络流的 249 种属性^[3],试图为协议分类研究提供一个完整的属性集合.

在网络流特征属性获取的基础上,国内外学者在协议识别方面也做了大量的研究工作,如引入基于概率模型的朴素贝叶斯方法^[3],用聚类方法处理协议识别问题^[4],将支持向量机方法应用于 P2P 流的识别领域^[5],采用高斯滤波方法实现网络协议识别^[6]等.上述方法虽然在一定程度上解决了基于统计特征的应用协议识别问题,但是所获取的网络流特征属性随着网络环境的变化而变化,缺乏持久性和稳定性,所以其广泛应用受到了限制.

针对以上问题,本文提出了一种网络协议识别方法.

1 网络流特征提取

根据流量统计特征实现协议识别的关键是网络流量属性集的获取. Moore 等人给出了双向 TCP 流的 249 种属性^[3],然而在实际的网络环境中,随着网络带宽的不断增加,待分类的网络流数目通常达到了数十万乃至数百万条,过多网络流导致计算负载过于沉重,而且在所提取的网络流属性中,由于众多网络流属性相互关联,使得分类模型的识别率较低,难以满足现有网络管理的需要.为此,以利于网络管理为基础,以属性易于获取和分析为出发点,本文利用 11 种属性来描述流量数据中双向 TCP 流的主要

特征,各属性的具体说明如表 1 所示.

表1 网络流特征属性及描述

属性	网络流特征	特征描述
x_1	存活时间	网络流持续的时间
x_2	流方向	流入或者流出监控端
x_3	入包个数	监控端接收的数据包数目
x_4	出包个数	监控端发出的数据包数目
x_5	入字节量	监控端接收到的数据量
x_6	出字节量	监控端发送的数据量
x_7	最大数据包	包含的最大数据包大小
x_8	包大小均值	传送数据包的平均大小
x_9	最大包间隔	最大的数据包时间间隔
x_{10}	包间隔均值	数据包间隔均值
x_{11}	出入字节比	接收和发送的数据量之比

2 应用协议识别算法

2.1 协议识别方法框架描述

本文的网络协议模型共分为网络流特征数据采集、网络流主要属性信息提取、BP 神经网络模型建立以及测试验证等 3 个步骤,其流程框架如图 2 所示,具体描述如下.

步骤 1:根据网络流的定义,将网络数据包汇聚为网络流,提取网络流的特征.

步骤 2:采用主成分分析(PCA)方法提取网络流属性的主要特征.

步骤 3:根据提取出的网络流主要特征,结合 BP 神经网络,构建网络应用协议识别模型.

2.2 网络流主要属性的提取

网络应用的每一次交互行为都可以用一个包含 11 种属性的特征向量来描述,将其分别表示为 x_1, x_2, \dots, x_{11} ,对应关系详见表 1. 这样,网络用户的行为集合就可以用一个 $n \times 11$ 的矩阵 X 来表示

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,11} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,11} \\ \vdots & \vdots & & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,11} \end{bmatrix} = [X_1, \dots, X_{11}] \quad (1)$$

式中: $x_{i,j}$ 表示第 j 个流量属性在第 i 条网络流样本上的观察值; n 为网络流的条数.然后,采用 PCA 方法实现网络流主要特征的提取^[7]. $\lambda_1, \lambda_2, \dots, \lambda_{11}$ 是协方差矩阵的特征值序列, u_1, u_2, \dots, u_{11} 是每个特征值分别对应的特征向量,其中特征值满足 $\lambda_1 > \lambda_2 > \dots$

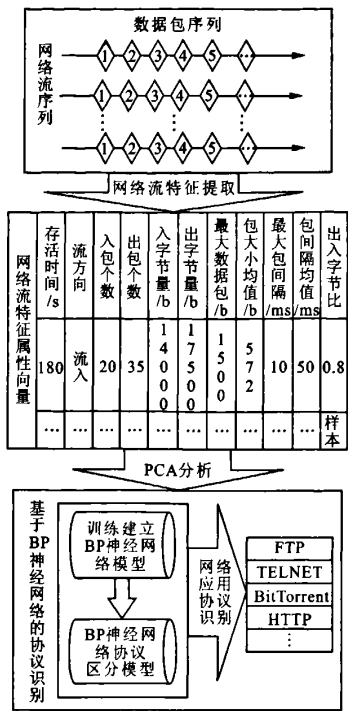


图 2 协议识别流程框架图

$>\lambda_{11}$. 通过方差贡献率 α 可衡量不同特征值和特征向量在重构原始空间时的重要性,通常 $\alpha \geq 80\%$. 满足式(2)的 m 值对应的特征向量构成一个特征子空间,它继承了原始数据集的主要特征,而余下的 $11-m$ 维特征属性主要是包含着噪声.

$$\alpha \leq \sum_{i=1}^m \lambda_i / \sum_{i=1}^{11} \lambda_i \tag{2}$$

2.3 网络应用协议识别模型

BP 神经网络是一种多层前馈型神经网络,它包含输入层、隐含层和输出层,可以逼近任意连续函数,广泛应用于模式分类并取得了很好的结果[8]. 根据协议识别需要,可建立 (I, J, K) 的协议识别网络模型,其中 I 为输入层神经元数, J 为隐含层神经元数, K 为输出层神经元数. 隐含层神经元数 J 采用经验式确定[9],即

$$J = (K + I)^{1/2} + H \tag{3}$$

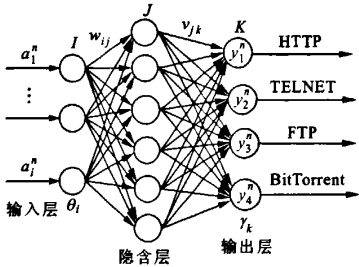
式中: H 为介于 $[1, 10]$ 之间的整数. 经过实验对比,选择隐含层节点数为 6,可以取得理想的实验结果. 将通过 PCA 算法提取后的特征属性 $P_n = (a_1^n, \dots, a_I^n)$ 作为神经网络的输入,其中的 n 表示训练集中的第 n 条网络流记录, $a_i^n (i=1, \dots, I)$ 为提取后的网络流特征属性. 本文根据实验结果选取 $I=3$, 输出层

神经元节点数为 4,期待输出值 $Y_n = (y_1^n, \dots, y_4^n)$,期待输出值与应用协议间的对应关系如表 2 所示.

表 2 神经网络期待输出与应用协议对应关系

应用协议	Y_n			
	y_1^n	y_2^n	y_3^n	y_4^n
HTTP	1	0	0	0
FTP	0	1	0	0
BitTorrent	0	0	1	0
TELNET	0	0	0	1

据此,网络协议识别的 BP 网络结构如图 3 所示.



w_{ij} : 输入层与隐含层的连接权重; v_{jk} : 隐含层与输出层的连接权重;
 θ_i, γ_k : 输入层、输出层神经元阈值

图 3 基于神经网络的协议识别模型

3 实验结果

3.1 网络数据收集

为了验证本文方法,在实验室环境下采用被动监听的方式采集网络流量. 图 4 所示是网络流量数据采集平台的拓扑结构图,其中电脑终端是监控对象,它们经过路由器连接至 Internet,在路由器处设置镜像服务器,用于采集流经路由器的网络流量.

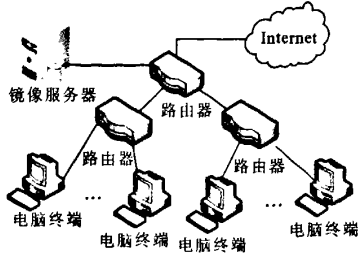


图 4 网络流量收集系统拓扑图

对收集到的网络流量,根据网络流定义将数据包汇聚为网络流并提取 11 种网络流特征. 采用特征

字段和网络端口相结合的方式标定应用层所采用的应用协议。在本文实验中,共收集了 HTTP、FTP、BitTorrent 及 TELNET 4 种网络应用数据,数据情况如表 3 所示。鉴于建立网络应用节拍模式的需要,实验中提取的网络流为完整的 TCP 交互流,即数据集不包含数据包个数少于 3 的网络流。随机抽取每种协议网络流的 50% 作为训练网络流集合,并建立网络应用协议识别模型,其余数据集将作为测试网络流集合,最后测试所建模型的准确性。

表 3 采集到的网络应用协议数据

应用协议	HTTP	FTP	BitTorrent	TELNET
流数	2 724	388	768	1 062
样本比	0.551	0.079	0.155	0.215
训练数	1 362	194	384	531

3.2 PCA 分析结果

实验中收集到的每一条网络流数据都是一个 11 维的向量, n 条流的特征数据就形成了 $n \times 11$ 的特征矩阵。采用 PCA 分析方法寻找网络流属性的主要特征,由式(3)计算所得的不同应用协议的方差贡献率,结果如图 5 所示。由图 5 可知,FTP 应用协议和 HTTP 应用协议的方差贡献率在属性为 3 维时就已经超过了 80%,BitTorrent 和 TELNET 应用协议的方差贡献率在属性为 3 维时分别接近和超过 70%,可见网络应用协议特征可以用 3 维数据有效描述。

经过 PCA 分析认为,可以用 3 维特征有效地描述原始网络流特征的主要属性。图 6 给出了提取出的主要属性在 3 维空间中的分布特征。由图 6 可以看出,经过特征提取处理后,网络应用协议特征数据在 3 维空间中呈现不同的特性,相同的应用协议在 3 维特征空间中聚集为相同的簇,表现出了很强的相似性,具有良好的可区分性。这为识别应用协议奠定了有效的基础。

3.3 BP 神经网络识别结果

经过 PCA 分析认为,可以将原有的 11 维网络流特征映射到 3 维空间。随机选取网络数据集的 50% 作为训练样本集,并建立 BP 神经网络模型,再利用其余数据集作为测试样本来验证本文算法的准确性。在本文中,选取神经网络系统误差为 0.01,协议识别模型各步的误差曲线如图 7 所示。由图 7 可以得出,经过近 4 次迭代,协议识别模型达到规定的系统误差,经过 10 次迭代后,识别模型趋于收敛。

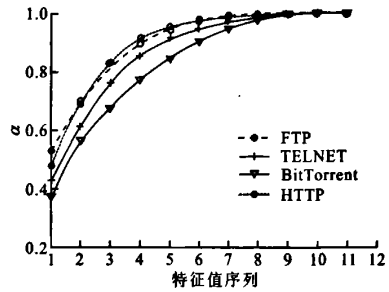


图 5 应用协议的低维方差贡献率

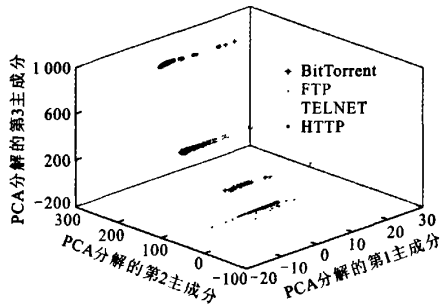


图 6 降维处理后不同应用协议特征的分布

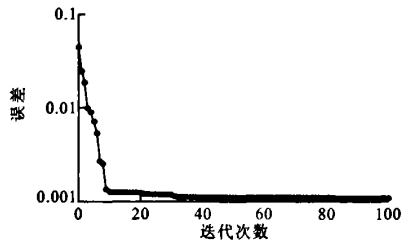


图 7 协议识别模型各步的系统误差

利用测试数据集得到的测试结果如表 4 所示。由表 4 结果可以看出,本文模型可以准确实现网络流协议类型的划分。根据原始网络流特征,仅采用 BP 神经网络识别算法的协议识别结果如表 5 所示。结合表 4、表 5 可以看出,单独使用 BP 神经网络算法的识别准确率要低于 PCA 与 BP 相结合的协议识别算法。这是因为,所提网络流的 11 类特征数据之间存在着一定的关联性,由此影响了识别准确率。

表 4 PCA 结合 BP 网络协议识别算法的测试结果

应用协议	HTTP	FTP	BitTorrent	TELNET
流数	1 362	194	384	531
识别数	1 341	190	375	523
识别准确率/%	98.5	97.9	97.6	98.5

表5 BP网络协议识别算法的测试结果

应用协议	HTTP	FTP	BitTorrent	TELNET
流数	1 362	194	384	531
识别数	1 251	180	365	503
识别准确率/%	91.8	92.7	95	94.7

经过PCA降维处理后,去除了原始数据之间的关联性,减少了网络流量属性的记录数量,这不但提高了BP识别模型的识别准确率,也降低了计算复杂度。

4 结束语

网络应用协议的有效识别是近几年网络安全领域的研究热点,也是提高网络管理效率的基础。本文提出了一种网络流特征刻画方法,即利用网络层参数描述网络应用协议的特征模式,采取PCA方法提取网络流特征的主要属性,最后结合BP神经网络建立不同网络应用协议识别模型。真实网络环境下的实验结果表明,本文方法能够有效识别网络应用协议的类型。我们的下一步工作是研究基于流量分类方法的实时应用系统。

参考文献:

- [1] MOORE D, KEYS K, KOGA R, et al. The Coral-Reef software suite as a tool for system and network administrators [C]//Proceedings of the 15th USENIX Conference on Systems Administration. Berkeley, CA, USA: USENIX Association, 2001: 133-144.
- [2] ROESCH M. SNORT: lightweight intrusion detection for networks [C]//Proceedings of the 13th USENIX Conference on Systems Administration. Berkeley, CA, USA: USENIX Association, 1999: 229-238.
- [3] ZUEV D, MOORE A W. Traffic classification using a statistical approach [C]//6th International Workshop on Passive and Active Network Measurement. Berlin, Germany: Springer-Verlag, 2005: 321-324.
- [4] ERMAN J, ARLITT M, MAHANTI A. Traffic classification using clustering algorithms [C]//Proceedings of ACM SIGCOMM Workshop on Mining Network Data. New York, USA: ACM, 2006: 281-286.
- [5] 王锐,王逸欣,樊爱华,等.一种跨层P2P流量检测方法[J].计算机应用,2006,26(12):30-32.
WANG Rui, WANG Yixin, FAN Aihua, et al. A P2P traffic detection method between levels [J]. Computer Applications, 2006, 26(12): 30-32.
- [6] CROTTI M, DUSI M, GRINGOLI F, et al. Traffic classification through simple statistical fingerprinting [C]//Proceedings of ACM SIGCOMM Computer Communication Review. New York, USA: ACM, 2007: 5-16.
- [7] DUDA R O, HART P E, STORK D G. Pattern classification [M]. 2nd ed. Beijing, China: China Machine Press, 2004: 576-580.
- [8] 程洪,郑南宁,高振海,等.基于主元神经网络和K-均值的道路识别算法[J].西安交通大学学报,2003,37(8): 812-815.
CHENG Hong, ZHENG Nanning, GAO Zhenhai, et al. Road recognition algorithm using principal component neural networks and K-means [J]. Journal of Xi'an Jiaotong University, 2003, 37(8): 812-815.
- [9] HECHT-NIELSON R. Neuron computing [M]. Reading, MA, USA: Addison Wesley, 1990: 124-125.

(编辑 苗凌)

西安交通大学两团队获2008年度国家级教学团队荣誉称号

日前,教育部公布了2008年度国家级教学团队名单,西安交通大学2个教学团队榜上有名,获得国家级教学团队荣誉称号。这2个团队分别是电气学院王兆安教授作为带头人的电力电子与新能源技术研究中心教学团队和理学院马知恩教授作为带头人的大学数学系列课程教学团队。

本次评比经过学校推荐、省级评审、专家评审等程序,全国共评选出300个国家级教学团队。

(来源:交大新闻网 <http://xjtunews.xjtu.edu.cn>)

利用统计特征的网络应用协议识别方法

作者: 徐莉, 赵曦, 赵群飞, 秦涛, XU Li, ZHAO Xi, ZHAO Qunfei, QIN Tao
作者单位: 徐莉, 赵群飞, XU Li, ZHAO Qunfei (上海交通大学图像处理与模式识别研究所, 200240, 上海), 赵曦, ZHAO Xi (上海金融学院信息管理系, 201209, 上海), 秦涛, QIN Tao (西安交通大学电子与信息工程学院, 710049, 西安)
刊名: 西安交通大学学报 ISTIC EI PKU
英文刊名: JOURNAL OF XI'AN JIAOTONG UNIVERSITY
年, 卷(期): 2009, 43(2)
被引用次数: 0次

参考文献(9条)

1. MOORE D, KEYS K, KOGA R [The Coral-Reef software suite as a tool for system and network administrators](#) 2001
2. ROESCH M [SNORT: lightweight intrusion detection for networks](#) 1999
3. ZUEV D, MOORE A W [Traffic classification using a statistical approach](#) 2005
4. ERMAN J, ARLITT M, MAHANTI A [Traffic classification using clustering algorithms](#) 2006
5. 王锐, 王逸欣, 樊爱华 [一种跨层P2P流量检测方法](#)[期刊论文]-[计算机应用](#) 2006(12)
6. CROTTI M, DUSI M, GRINGOLI F [Traffic classification through simple statistical fingerprinting](#) 2007
7. DUDARO, HART P E, STORK D G [Pattern classification](#) 2004
8. 程洪, 郑南宁, 高振海 [基于主元神经网络和K-均值的道路识别算法](#)[期刊论文]-[西安交通大学学报](#) 2003(08)
9. HECHT-NIELSON R [Neuron computing](#) 1990

相似文献(10条)

1. 学位论文 曹建华 [基于灰色神经网络模型的网络流量预测算法研究](#) 2008

随着Internet及其应用的迅速发展,网络规模日益增大,网络应用日益复杂,由于互联网是一个高度复杂的非线性系统,为了实现可靠的数据传输及合理的网络资源分配,就需要深刻了解网络的控制机制和复杂的行为特性。网络流量分析和建模对网络性能评价具有重要的意义,完全符合网络流量复杂统计特性的模型,能够帮助对网络流量进行精确的分析和仿真,非常有助于网络的设计和控制。

网络流量预测分析及建模一直是分析网络性能的重要研究课题,流量预测结果为网络管理中带宽分配、流量控制、选路控制、接纳控制和差错控制等提供主要参考依据。网络流量具有一定的动态性、实时性、相关性、随机性和含噪声性。预测精度的高低、所选或所建立模型的表达能力,对于分析和仿真,理解网络的动态行为,以及指导流量控制的设计工作均具有重要的指导意义。

本论文的研究目的是为了探索一种新的网络流量模型来更好的描述、预测网络流量的特性,文章首先分析了网络流量的一些主要特性,在真实环境中的网络流量呈现出相当明显的多尺度特性,例如自相似性、长相关性、单分形和多分形等特性;接着,分析、比较几种传统的网络流量分析模型的优缺点,如半马尔柯夫模型、Poisson模型、ARMA模型及ON/OFF模型。

论文提出了基于残差改进的灰色预测方法、灰色神经网络组合模型等多种综合预测分析方法,预测实际网络流量,确定适合的模型输入参数,详细分析了各个模型的精度和预测效果,为高精度的短期(以及中长期)网络流量预测软件系统的集成开发奠定了基础。

针对常用的动态全局链路状态法和基于流量的路由算法存在的不足,提出了一种新的基于流量预测的路由改进算法,可结合流量预测的优势对路由算法加以有效补充,理论分析预计该算法能获得更佳的路由效果。

2. 学位论文 谢纪军 [网络OD流估计的非线性方法研究](#) 2008

OD(Origin-Destination)流量矩阵在网络设计,路由配置及调试,网络流量监测等方面具有重要的意义。然而随着互联网的飞速发展,网络拓扑结构越来越复杂,通过网络测量的方法直接获得网络OD流量矩阵变得越来越困难或昂贵。而链路的负载流量相对OD负载流量更容易获得,因此通过收集链路的流量来估计OD流量矩阵成为网络层成像研究热点之一。

由于测量获得的链路流量数通常远小于OD流量数,因此OD流量矩阵估计是一个典型的欠定性问题,传统的方法往往假定每个OD流服从某种模型(如高斯分布等)或同一时刻的多个OD流服从某种模型(如重力模型等),以该模型作为约束求解获得OD流量矩阵。我们认为网络OD流量的变化是一个复杂的过程,难以使用上述预先给定模型获得精确的OD流量矩阵。但同一个网络在不同时刻的流量矩阵存在一定的关联性,我们可以从不同角度提取同一网络历史OD流量存在的一些规律,求解当前时刻的OD流量矩阵。为此本文在如下两个方面进行了一些有意义的探索:

(1)根据网络OD流量估计需要结合网络流量时空分布相关联这一观点,本文将递归神经网络模型这一非线性方法引入到OD流的估计领域,提出了基于时空关联的OD流的估计方法-基于递归神经网络的OD流估计。具体思路如下:利用已获得的链路数据和部分OD流矩阵信息训练估计模型,以获得链路数据到OD流矩阵的映射权值,利用已保存的映射权值来预测未来时刻的OD流矩阵信息。本文还就初值的约束以及部分估计信息的约束展开了讨论,仿真结果表明,该算法极大的提高OD流矩阵的估计的精度。(2)本文提出了基于模拟退火算法的OD流矩阵估计这一非线性的估计算法。我们利用当前时刻的OD流矩阵与前一时刻的OD流矩阵的相互关系,以通过增加状态转移矩阵方式,来约束OD流量矩阵估计,从而提高了OD流矩阵的估计精度。具体实验验证了我们提出算法的有效性。

3. 期刊论文 雷霆, 余镇危, LEI Ting, YU Zheng-wei [一种网络流量预测的小波神经网络模型](#) -[计算机应用](#)

2006, 26(3)

结合小波变换和人工神经网络的优势,建立一种网络流量预测的小波神经网络模型。首先对流量时间序列进行小波分解,得到小波变换尺度系数序列和小波系数序列,以系数序列和原来的流量时间序列分别作为模型的输入和输出,构造人工神经网络并且加以训练。用实际网络流量对该模型进行验证,结果表明,该模型具有较高的预测效果。

4. 学位论文 [吴文清 基于系统动力学的网络业务研究](#) 2004

本文应用混沌理论对网络流量的系统动力学特征进行了认真分析,并结合网络流量的具体数据,计算了网络流量的Hurst参数,关联维数, Lyapunov指数.在此基础上,运用小波理论对网络流量数据进行了去长相关处理,进一步对网络流量混沌成因进行了探讨,指出网络流量的混沌与网络流量的长相关存在着联系.在此基础上,运用BP神经网络理论,建立了相关模型,对实际网络流量进行了预测.研究表明,基于混沌理论建立的BP神经网络模型和FARIMA(p, d, q)模型都能较准确地对网络流量进行预测,而BP神经网络模型能够经过学习,获得较长的预测步长.

5. 期刊论文 [余健. 郭平. Yu Jian. Guo Ping 一种优化的RBF神经网络模型用于网络流量预测](#) -[计算机应用与软件](#)

2008, 25 (12)

采用径向基RBF神经网络对网络流量数据的时间序列进行建模与预测.采用传统的学习算法对RBF网络训练时,对网络流量数据容易出现过拟合现象,提出了自适应量子粒子群优化AQPSO算法,用于训练RBF神经网络的基函数中心和宽度,并结合最小二乘法计算网络权值,改善了RBF神经网络的泛化能力.实验结果表明,采用AQPSO算法获得的RBF神经网络模型具有泛化能力强、稳定性良好的特点,在网络流量预测中有一定的实用价值.

6. 学位论文 [刘杰 基于神经网络的非线性网络流量分析与预测](#) 2005

随着Internet的迅猛发展,网络规模日趋庞大,网络拓扑结构渐趋复杂,越来越需要一种新的网络规划和设计手段来提高网络设计的客观性和设计结果的可靠性.网络流量模型的建立是网络性能分析、网络协议设计、网络规划等的基础.

本文根据实际网络中测量的大量网络流量数据,建立一个时间相关的基于神经网络的流量模型,预测和分析网络流量状况.传统的流量分析建立在线性模型的基础上,但是由于复杂的拓扑结构和网络行为,网络流量表现为一个非线性的系统.本文基于ANN理论提出精度较高的非线性网络流量预测模型,该模型具有很高的自适应性和鲁棒性.

本文不仅阐述了神经网络模型的构造、预测模型的建立,而且介绍了网络流量的采集和预处理的一系列具体方法和实现.包括去除离群点,缺损值补足以及白噪声消除在内的数据处理方法被应用于网络流量数据的预处理和后处理中,对于预测效果有很大提升.

本文提出的模型,通过在实际网络环境上的应用,取得较好效果.本文还根据此模型给出一个实际的流量监测与预测系统,它由网络流量采集系统及网络流量预测系统,可更方便的监控预测网络状况,为网络性能分析等工作服务.

7. 期刊论文 [白燕. 马光思. BAI Yan. MA Guang-si 基于灰色径向基神经网络模型的流量预测与分析](#) -[计算机工程与科学](#)2008, 30 (10)

根据神经网络能有效修正灰色预测模型的思路[1],本文提出了基于灰色系统及径向基神经网络的组合预测模型.通过采集园区节点交换机的流量数据,在分析网络流量时间序列特性的基础上建立灰色GM(1, 1)模型,并采用径向基神经网络对预测模型残差进行修正.实验结果和仿真实验表明,组合模型效果及预测精度远优于单一灰色预测模型.

8. 学位论文 [田妮莉 基于小波和FIR神经网络的流量预测模型研究](#) 2007

随着网络规模的日益庞大和复杂,发生各种故障的可能性增大,网络性能更容易受到影响.网络流量能直接反映网络性能的好坏,如果网络所接受的流量超过它实际的运载能力,就会引起网络性能下降.所以网络流量的建模与预测对于大规模网络的规划设计、网络资源管理以及用户行为的调节等方面都具有积极意义.本论文是基于国家自然科学基金项目“基于共变正交和联合优化的多媒体网络性能预测模型”的研究需要而展开的.

本文介绍了网络流量的特征,在分析了小波理论和FIR神经网络特点的基础上提出了一种基于小波变换和FIR神经网络的网络流量预测模型.该模型采用小波分解把网络流量数据分解成小波系数和尺度系数,即高频系数和低频系数,将不同频率成分的系数单支重构为高频流量分量和低频流量分量,用FIR神经网络得到不同频率成分的预测分量,将它们的合成作为对原始网络流量的预测结果.本文采用研究人员公认的网络流量数据作为测量分析对象,对该模型进行了仿真实验:对局域网和广域网网络流量的单步预测、在不同小波基下的单步预测性能及分析、与现有的小波变换神经网络模型和FIR神经网络模型在不同结构下的性能比较及分析.这些实验的结果验证了本文提出的预测模型的有效性和优越性.最后对本文的主要工作进行了总结,并针对本文提出的模型有待改进的地方和下一步的工作提出了几点看法.

9. 期刊论文 [吴文清. 赵黎明. 刘嘉焜. Wu Wenqing. Zhao Liming. Liu Jiakun 基于混沌理论的网络流量BP神经网络](#)

[测](#) -[计算机工程与应用](#)2005, 41 (17)

应用相空间重构技术,论证了网络流量存在混沌现象,并计算了实际网络流量的关联维, Lyapunov指数;在此基础上对网络流量时间序列建立BP神经网络模型对网络流量进行预测,实例表明,与FARIMA(p, d, q)模型相比,基于混沌理论的BP神经网络模型有较好的预测能力,并具有较长的预测步长.

10. 学位论文 [汪志勇 结合分形神经网络理论的网络流量预测研究](#) 2009

网络流量特性的分析、流量建模以及流量预测对于新一代网络协议设计、网络管理和提高网络服务质量等都有重要的意义.本文也正是以网络流量特性分析为出发点,对网络流量预测模型算法及其应用进行了相关的研究.

论文首先对网络流量特性进行了详细分析,介绍并实现了六种网络流量分形维度的估算方法,通过仿真实验对比验证了各种估算方法的精度以及周期和噪声信号对估算方法的影响;同时论文探讨了网络流量自相似的成因和对网络性能的影响.

其次本文归纳总结了网络流量建模预测发展过程中的传统的短相关模型、自相似模型各自的优缺点,讨论了网络流量领域最新的分析方法和理论,分析得出:针对同时具有长相关和短相关的自相似网络业务流应当采用混合 N_0 模型才能更精细的刻画其特性.

在详细分析网络流量特性的基础上,提出基于分形滤波神经网络的混合流量预测模型,改进神经网络训练算法,应用该模型对真实网络流量的变化趋势进行拟合和预测,实验结果表明:分形滤波神经网络模型能同时描述自相似流量的长相关和短相关特性,并且其拟合真实流量效果比FARIMA模型、AR模型和BF神经网络模型更好;Hurst参数值越大的仿真流量,其预测精度越高;分形滤波去长相关后真实网络流量短期预测精度高于传统回归模型;分形滤波去短相关的预测模型,在多步预测中更能抓住业务流的未来变化趋势.

最后论文将分形预测模型与网络动态带宽分配相结合,仿真实验表明针对具有分形特性突发性较强的业务流,基于预测的动态网络带宽分配策略能有效的减少信元丢失率、传送时延及队列长度,大大提高了网络传输质量.

本文链接: http://d.wanfangdata.com.cn/Periodical_xajtdxxb200902010.aspx

授权使用: 山东大学(sddx), 授权号: b7f7a074-ad1e-48be-8442-9ec600f09258

下载时间: 2011年4月15日