

# 基于随机神经网络的多步网络时延预测模型

胡治国<sup>1,2</sup> 张大陆<sup>1</sup> 侯翠平<sup>1</sup> 沈斌<sup>1</sup> 朱安奇<sup>1</sup>

(同济大学计算机科学与技术系 上海 201804)<sup>1</sup>; (中国人民解放军 65583 部队 辽阳 111000)<sup>2</sup>

**摘要** 网络时延的动态变化反映了网络路径的负载特征,对时延的精确预测是实施网络拥塞控制、路由选择的重要依据,建立了基于随机神经网络的多步网络时延预测模型,该模型克服了传统时间序列预测方法受随机干扰因素影响大、模型结构辨识过程繁琐,以及传统神经网络预测方法易于陷入局部极值、偏离全局最优的缺点。仿真实验表明,在提前单步和多步的预测中该模型比 AR 模型、RBF 神经网络预测算法的准确度更高。

**关键词** 网络时延, RNN 神经网络, 预测

中图分类号 TP393 文献标识码 A

## Multi-step Network Delay Prediction Model Based on RNN

HU Zhi-guo<sup>1,2</sup> ZHANG Da-lu<sup>1</sup> HOU Cui-ping<sup>1</sup> SHEN Bin<sup>1</sup> ZHU An-qi<sup>1</sup>

(Department of Computer Science and Engineering, Tongji University, Shanghai 201804, China)<sup>1</sup>

(The Chinese People's Liberation Army, 65583 Troops, Liaoyang 111000, China)<sup>2</sup>

**Abstract** Network delay is an important performance metric of IP network which reflects network path's workload characteristics. The precise prediction for network delay is an important basis on congestion control and route selection. A new multi-step prediction method was proposed for network delay prediction based on the random neural networks, this method overcomes the disadvantage traditional time-series method and neural network method. Compared with traditional RBF network and AR model, the experimental result indicated that the proposed model has better accuracy for single steps and multi steps prediction.

**Keywords** Network delay, Random neural network, Prediction

## 1 引言

近年来随着 VoIP、网络游戏等其他时延敏感型业务在网络应用中的快速发展,需要网络提供服务质量保证,因而网络路径的时延特性越来越受到人们的重视。网络时延受网络拓扑、转发节点、路由算法以及背景流量等多重因素的影响,反映了路径的当前性能和对端主机提供的服务水平,但它们都随时间而随机变化<sup>[1,2]</sup>。跟踪和预测网络时延的动态变化可以获得网络资源的使用状况和性能趋势,为均衡网络负载和优化网络性能提供参考<sup>[3]</sup>。

文献[4]采用传统的分析理论——排队论,采用泊松到达的假定,认为数据的到达间隔具有不相关性,但是根据排队论预测的结果与网络实测数据有较大的差异<sup>[4]</sup>。文献[5-7]分别采用基于线性预测模型用时间序列的 AR (Auto-Regressive), ARMA (Auto-Regressive Moving Average), ARIMA (Auto-Regressive Integrated Moving Average) 预测模型,得到的结论是基于网络时延的时间序列可以通过低阶的 AR 或 ARIMA 模型精确建模,提供精确度较高的预测值。但模型参数的求解过程是通过离线求解方程实现,不是通过在线递

推估计实现,并且在预测的过程中模型的参数是固定的,不能很好地适应不确定性强的网络时延动态预测的要求,同时该方法在应用中受随机干扰因素影响大,模型结构辨识和模型检验的过程也比较繁琐。Parlos 在文献[8]、D. Srikar 在文献[9]用神经网络中的多层感知器(MLP)的不同算法对网络时延变化进行预测分析,但传统神经网络存在“过训练”和“训练不足”的问题,易陷入局部极值,外推能力有限。

随机神经网络(RNN)模型是 Gelenbe 在文献[10,11]中提出的,它第一次使用独立的数学模型描述了生物神经网络接收信号流激活而传导刺激的生理机制,克服了传统神经网络的诸多不足,在网络结构、学习算法、状态更新规则等方面都具有自身独特特点。作为仿生神经元数学模型,随机神经网络在联想记忆、组合优化问题上显示出了较强的优势,并得到广泛的应用<sup>[12-14]</sup>。本文工作主要在于设计和实现了基于随机神经网络(RNN)的网络时延预测模型,并与 AR 线性预测及径向基(RBF)神经网络模型进行了分析对比。实验结果表明,基于随机神经网络(RNN)的网络时延预测模型能够可靠、高精度地对网络时延的变化情况进行预测。

第2节分析网络时延的组成;第3节给出随机神经网络

到稿日期:2008-10-14 返修日期:2008-12-15 本课题受国家自然科学基金资助项目(90204010)资助。

胡治国(1978—),男,博士研究生,主要研究方向为网络性能评估, E-mail: huliang393@yahoo.com.cn; 张大陆(1949—),男,教授,博士生导师,主要研究方向为计算机网络; 侯翠平(1985—),女,硕士研究生,主要研究方向为网络管理; 沈斌(1983—),男,硕士研究生,主要研究方向为网络测量; 朱安奇(1984—),男,硕士研究生,主要研究方向为网络测量。

(RNN)预测模型及其它预测算法;第4节给出仿真实验结果并与传统的自回归和径向神经网络的预测结果进行对比;最后给出相应结论和下一步的研究内容。

## 2 网络时延组成分析

端到端时延由节点时延和链路时延组成,具体分为4部分:分组在链路上传播时延 $t_g$ (propagation delay)、分组的传输时延 $t_t$ (transmission delay)、分组在节点的处理时延 $t_p$ (processing delay)以及在缓冲区中的排队时延 $t_q$ (queue delay)。对于一个 $H$ 跳的链路,网络时延为

$$delay_i = t_{g,0} + t_{t,0} + \sum_{h=1}^H (t_{t,h} + t_{q,h} + t_{p,h} + t_{g,h}) \quad (1)$$

$t_{g,h}$ 为分组在第 $h$ 跳链路上的传播时延, $t_{t,h}$ 为分组在第 $h$ 个节点的传输时延, $t_{q,h}$ 为分组在第 $h$ 个节点的排队时延, $t_{p,h}$ 为分组在第 $h$ 个节点的处理时延。

严格讲,以上各部分时延都是变化的且都相互关联。但由于网络负载的变化发生在中间路由器上,因此说网络时延及其不确定性主要决定于数据所经过的跳数以及在每一跳上所花费的时间。网络传输时延看似随机,但其背后却有一定的规律。正如文献[15]分析的那样,在网络正常运行的情况下,因特网是采用动态路由机制来传输数据业务的。但这种动态并不等于随机路由,只有当互连网的情况发生严重变化时,例如路由失效、严重拥塞等,转发数据报的路由路径才会改变。对因特网路由行为的研究表明,通信双方或多方在会话期间路由由很少改变,大部分的数据流沿同一条物理路径传输。因此,各数据报的时延总是分布在一定的范围内的,看似随机的数据时延背后有着一定的规律可循,这就使预测网络时延成为可能。

## 3 RNN 神经网络预测模型

RNN是一种具有 $n$ 个神经元的开放随机网络,神经元 $i$ ( $i=1, \dots, n$ )的状态由 $t$ 时刻它的兴奋水平 $k_i(t) \in Z^+$ 表示,它是一个非负整数,称之为势。RNN中正信号(+1)表示兴奋,负信号(-1)表示抑制。正信号到达第 $i$ 个节点,该神经元的势加1,负信号到达使之减1(到零时不再减)。同时,当神经元势为正值,它将不断地释放信号,释放信号的时间间隔服从参数为 $r(i) > 0$ 的指数分布。若神经元 $i$ 释放一个信号,它作为正、负信号被传递到神经元 $j$ 的概率分别为 $p_{ij}^+$ 和 $p_{ij}^-$ ,这个信号也可能离开网络,此概率为 $d(i)$ 。假设 $p_{ii}^+ = p_{ii}^- = 0, 1 \leq i \leq n$ ,表示神经元不向自身传递信号。信号在传递过程中有损耗,即 $d(i) > 0$ 。显然

$$d_i + \sum_{j=1}^n [p_{ij}^+ + p_{ij}^-] = 1 \quad (2)$$

令 $w_{ij}^+ = r_i p_{ij}^+, w_{ij}^- = r_i p_{ij}^-$ ,用来表示节点发射正负信号的速率,它们与通常神经网络模型中的连接权值相似,并且是非负的。网络的神经元节点同样接受来自外界的信号,正、负信号分别服从参数为 $\Lambda_i, \lambda_i$ 的Poisson分布。

随机神经网络学习的主要目的是为了得到一个适当的权值矩阵,使得输入为一对兴奋和抑制的信号流速率的矢量时,网络输出为期望值,或者其与期望值的二次方差最小。即对于 $Q$ 组输入输出对 $(P, Y)$ , $(P^k = (\Lambda_k, \lambda_k), k=1, 2, \dots, Q; Y$ 为 $n$ 对兴奋和抑制信号流速率构成的矢量; $y^k$ 为 $n$ 个网络节点的期望输出构成的矢量),使得性能函数<sup>[11]</sup>

$$E^{(k)} = \frac{1}{2} \sum_{i=1}^n a_i (q_i^{(k)} - y_i^{(k)})^2 \quad (3)$$

为最小。

本文设计 RNN 预测算法如下。

Step1 样本的预处理

为了加快网络训练和收敛,需在网络训练之前对样本进行预处理。RNN 采用的归一化方法为

$$x_i^q' = \frac{x_i^q - x_{i \min}}{x_{i \max} - x_{i \min}} \quad (4)$$

$x_{i \max} = \max\{x_i^1, x_i^2, \dots, x_i^p\}, x_{i \min} = \min\{x_i^1, x_i^2, \dots, x_i^p\}$ ,  $p$ 为输入样本的个数, $x_i^q$ 为第 $q$ 个样本的第 $i$ 个变量。

Step2 神经网络结构选择

主要是依据对样本数据的初步实验、预测值与真实值的误差大小,对神经网络输入和输出节点个数、隐含节点数作出估计。

Step3 设神经元节点的正负信号输入为

$$x^{(k)} = (\Lambda^{(k)}, \lambda^{(k)}) \quad (5)$$

Step4 求 RNN 网络节点被激活的稳定概率状态 $q_i$

网络节点被激活的稳定概率状态分布可以简单表述为一个比值:所有兴奋信号的到达速度之和和所有抑制信号到达速率以及节点的发射频率之和<sup>[11]</sup>。

$$q_i = \lim_{t \rightarrow \infty} \Pr[k_i(t) > 0] = \lambda_i^+ / (r_i + \lambda_i^+) \quad (6)$$

$$\lambda_i^+ = \Lambda_i + \sum_{j=1}^n q_j^+ w_{ji}^+, \lambda_i^- = \Lambda_i + \sum_{j=1}^n q_j^- w_{ji}^- \quad (7)$$

$$r_i = \sum_{j=1}^n q_j^+ w_{ji}^+ + q_j^- w_{ji}^- \quad (8)$$

Step5 求得神经元节点的正负信号输入的权值矩阵 $W_k^+, W_k^-, W_k^+ = \{w_k^+(i, j)\}$ 和 $W_k^- = \{w_k^-(i, j)\}$ 这两个 $n \times n$ 矩阵,权值矩阵更新规则为<sup>[11]</sup>

$$w_{\omega}^{(k)} = w_{\omega}^{(k-1)} - \eta \sum_{i=1}^n a_i (q_i^{(k)} - y_i^{(k)}) [\partial q_i / \partial w_{\omega}^{(k)}] \quad (9)$$

$$\partial q_i / \partial w_{\omega}^{(k)} = r^+(u, v) q_u [I - W]^{-1} \quad (10)$$

$$\partial q_i / \partial w_{\omega}^{(k)} = r^-(u, v) q_u [I - W]^{-1} \quad (11)$$

$\eta > 0$ ,而且为常数。

Step6 神经网络训练与预测实现

在训练过程中,用已知 $n$ 个序列值为输入,随后 $l$ 个序列值为输出,并按顺序推移形成训练样本,这实质是一个 $R^n \rightarrow R^l$ 的映射问题。在预测过程中,可根据已知时间序列点来映射未来时间序列点。最后对 RNN 预测输出结果进行反归一化,验证实验结果精度。如不满足要求,返回 step2 去进一步调整神经网络结构。

## 4 仿真及结果分析

### 4.1 仿真实验

为了验证本文方法的有效性,使用网络模拟器 ns2 进行了仿真。为了和其它方法进行对比分析,仿真采用了与文献[9, 16]相同的拓扑结构,如图1所示。

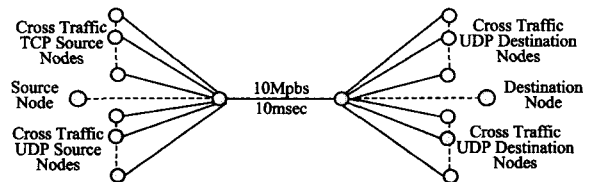


图1 仿真拓扑

为了模拟 Internet 中的背景流量,按照文献[17]中对 Internet 流量的统计数据设置 TCP 和 UDP 节点数,其中 TCP 节点占总的节点数量的 90%,仿真设定探测包发送间隔为 5ms,探测包大小 512bytes,紧链路为 10Mbps,延迟为 10ms,

各节点与紧链路连接带宽为 1Mbps,传播时延为 10ms,每个发生源 ON 或 OFF 的时长独立地符合重尾分布,测量时间为 5s,共测得 705 包的时延数据。

为了对比分析基于 RNN 神经网络预测算法的性能,本文同时实现了基于 AR 模型的预测算法和基于 RBF 神经网络的预测模型。选择将 1~505 个点用于 AR 模型的参数估计及神经网络训练,然后预测出 506~705 点的时延值。依据 AR 模型算法,确定了 AR 模型阶数  $p=5$ ,模型系数为  $\phi = [-0.9022 \ -0.4592 \ 0.1343 \ 0.2463 \ 0.0028]$ ,同时使用 RBF 算法,对神经网络的最佳结构进行了初步推算,在输入节点为 5、隐藏节点为 8 的情况下,预测值与真实值的均方误差最小。因此,在实验中把 RBF 网络和 RNN 网络的输入节点设为 5、隐藏节点设为 8,分别进行提前多步实验,其中提前 1,5,10 步实验结果如图 2 和图 3(提前 1 步预测)、图 4 和图 5(提前 5 步预测)、图 6 和图 7(提前 10 步预测)所示。

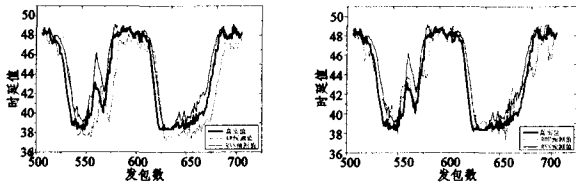


图 2 AR 与 RNN 时延预测比较 图 3 RBF 与 RNN 时延预测比较

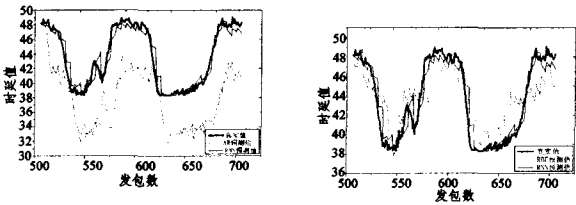


图 4 AR 与 RNN 时延预测比较 图 5 RBF 与 RNN 时延预测比较

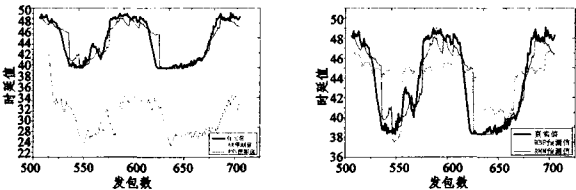


图 6 AR 与 RNN 时延预测比较 图 7 RBF 与 RNN 时延预测比较

#### 4.2 实验数值结果分析

为了检验本文提出的基于 RNN 网络预测的性能,分别与 AR 模型、RBF 神经网络预测算法从相关系数、预测值与真实值之间的均方误差、最大相对误差、预测时间 4 个方面进行了比较分析。

##### (1) 相关系数

又称 Pearson 相关系数,说明两个现象之间相关关系密切程度的统计分析指标。相关系数用希腊字母  $\gamma$  表示, $\gamma$  值的范围在 -1 和 +1 之间。 $\gamma > 0$  为正相关, $\gamma < 0$  为负相关, $\gamma = 0$  表示不相关。 $\gamma$  的绝对值越大,相关程度越高。

$$\gamma =$$

$$\frac{\sum x(k)\hat{x}(k|k-p) - \sum x(k) \sum \hat{x}(k|k-p)}{\sqrt{[\sum x(k)^2 - \frac{(\sum x(k))^2}{n}][\sum \hat{x}(k|k-p)^2 - \frac{(\sum \hat{x}(k|k-p))^2}{n}]}} \quad (12)$$

其中  $x(k)$  为真实值, $\hat{x}(k|k-p)$  为提前  $p$  步的预测值。

(2) 预测值与真实值之间均方差

$$MSE = \frac{\sum_{k=1}^N (x(k) - \hat{x}(k|k-p))^2}{\sum_{k=1}^N x(k)^2} \quad (13)$$

(3) 最大相对误差

$$MRE = \max_{1 \leq k \leq N} \frac{|\hat{x}(k|k-p) - x(k)|}{x(k)} \quad (14)$$

从实验结果可以看出,基于 RNN 神经网络的预测算法的预测性能明显优于其它两个方法。但随着预测步长的加大,预测性能开始有所下降,基于 AR 模型的预测算法性能下降更为剧烈,如表 1 所示。这是因为 AR 模型是一种线性预测模型,不适合用于网络时延这种数值变化较大、易受随机因素干扰的数据进行多步的预测。

表 1 预测方法性能比较

预测步长	分析方法	AR 模型	RBF 模型	RNN 模型
提前一步预测	$\gamma$	0.8871	0.9400	0.9718
	MSE	0.0023	0.0010	9.6492e-004
提前五步预测	$\gamma$	0.7694	0.8345	0.9635
	MSE	0.0214	0.0028	6.5014e-004
提前十步预测	$\gamma$	0.5807	0.7048	0.9086
	MSE	0.1146	0.0046	0.0015
	MAE	47.61%	16.53%	15.21%

##### (4) 算法运行时间

实验中使用 IBM R60i 笔记本,用 matlab7.0 实现相关算法,采用多次测量求均值的方法对基于 AR 模型、RBF 网络、RNN 网络的网络时延预测算法时间进行了比较。在预测输入为 200 组的情况下,AR 模型、RBF 网络、RNN 网络的预测算法运行时间分别是 16ms,78ms,47ms;在预测输入为 100 组的情况下,AR 模型、RBF 网络、RNN 网络的预测算法运行时间分别是 16ms,31ms,16ms。从实验情况看,AR 模型优于 RBF 网络、RNN 网络优于 RBF 网络预测算法。

同时,在实验过程中还进一步验证基于 RNN 算法的优越性。RBF 网络虽也克服易陷入局部极值的问题,但实验过程中最佳目标误差、径向基函数的分布密度的选择还是相当繁琐的。如选取不当,预测结果就很不理想。图 8 是本文在提前 2 步预测时遇到的情况。

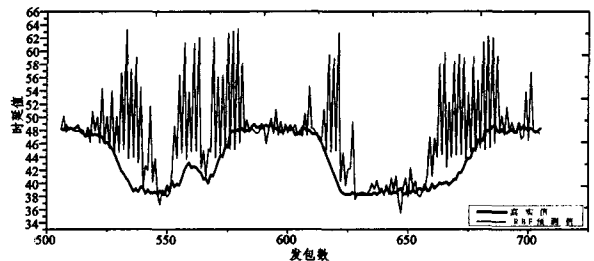


图 8 RBF 时延预测

**结束语** 本文提出了一种基于随机神经网络的预测模型,其建模过程比较简单,避免了传统的时间序列分析的模型结构辨识和模型检验的繁琐过程,并克服了传统神经网络易陷入局部极值及过训练的缺点。通过与 AR 模型及 RBF 神经网络的比较实验,显示了本模型对多步的网络时延预测有着优良的预测性能。如何运用该模型去预测网络时延超越阈

(下转第 112 页)

子商务语义服务资源在不同的机器上进行实验验证。下面是搜索语义服务测试示例:在搜索输入框输入“query the airline price from Wuhan to Beijing and query the start time and arrive time”,点击搜索后得到实验结果。

通过本文的方法建立的电子商务旅游语义服务搜索平台,在搜索的文本输入框输入动宾结构的陈述句,在部署好的以 RDF4S 为描述格式的语义服务中能够得到比较准确的实验结果。图 5 和图 6 是本文的一个实验过程。

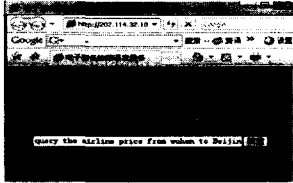


图 5 输入语义服务搜索条件

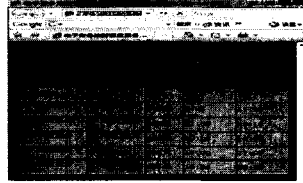


图 6 语义服务搜索结果

**结束语** 传统的搜索引擎都是按照关键字搜索网页,且大部分只是静态网页,因此对于语义服务,只能通过编写程序到 UDDI 中去提取已经注册了的网络服务。这种需要通过编写程序去搜索所需要的服务已经越来越难以满足普通人尤其是非程序员对网络服务的要求。为了解决这些面临的困难,本文试图设计一种面向非程序员的语义服务搜索方法。基于此,本文详细分析了一种基于类自然语言驱动、以 RDF4S 格式发布的语义服务的搜索方式。

基于类自然语言驱动的语义服务搜索在目前仍然是一个新的领域,还有很多的关键问题需要去解决。尤其这种面向非程序员的语义服务搜索方法,它面临一个“人”和“软件(这

里是指语义服务)”进行交互的一个巨大鸿沟。为了逐步解决人机交互的每一步,哪怕是取得一点点进展都需要进行大量的研究,这里不仅涉及到语义计算的研究、自然语言理解、自然语言处理,还涉及到形式化验证以及各种优化问题的研究。为了不断地完善这种基于类自然语言处理的语义服务搜索方法,今后的研究主要从以下几个方面进行:①分布式环境下的语义服务搜索优化;②并行环境下语义服务搜索优化;③问答式的语义搜索服务方式实现等等。

## 参 考 文 献

- [1] Verma K, Sivashanmugam K, Sheth A, et al. METEOR-S WSDI: A Scalable P2P Infrastructure of Registries for Semantic Publication and Discovery of Web Services[J]. Journal of Information Technology and Management, 2005, 6(1): 17-39
- [2] Paolucci M, Kawamura T, Payne T R, et al. Semantic matching of Web services capabilities[C]// Horrocks, ed. Proc. of the Int'l Semantic Web Conf. Sardinia; Springer-Verlag, 2002; 333-347
- [3] 马应龙, 金蓓弘, 冯玉琳. 基于进化分布式本体的语义 Web 服务动态发现[J]. 计算机学报, 2005, 28(4): 603-615
- [4] Lin Lin, Arpinar I B. Discovery of Semantic Relations between Web Services[C]// Web Services, 2006. ICWS '06. International Conference, on Sept. 2006; 357-364
- [5] Paolucci M, Mawamura T, Payne T R, et al. Semantic Matching of Web Service Capabilities[C]// First International Semantic Web Conference. 2002; 333-347
- [6] 吴健, 吴朝晖, 李莹, 等. 基于本体论和词汇语义相似度的 Web 服务发现[J]. 计算机学报, 2005, 28(4): 595-600

(上接第 87 页)

值的发生及其发生的时间,是我们下一步工作的目标。

## 参 考 文 献

- [1] Paxson V. End-to-end Internet packet dynamics[J]. IEEE ACM, Trans Networking, 1999, 7(6): 277-292
- [2] Allman M, Paxson V. On estimating end-to-end network path properties [C]// Proc. ACM SIGCOMM. 1999; 263-274
- [3] Johari R, Tan D K H. End-to-end congestion control for the Internet; Delays and stability[J]. IEEE/ACM Trans on Networking, 2000, 9(6): 818-832
- [4] Wong J W. Queuing network modeling of computer communication networks[J]. Computing Surveys, 1978, 10(3): 343-351
- [5] Jiao Liangbao, Zhang De. Differential AR algorithm for packet delay prediction[J]. Progress in Natural Science, 2006, 16(4): 437-440
- [6] Li Q, Millis D L. Jitter-based delay-boundary prediction of wide area networks[J]. IEEE Trans. on Networking, 2001, 9(5): 578-590,
- [7] Yang M, Li X R. Predicting end-to-end delay of the internet using times series analysis[R]. University of New Orleans, 2003
- [8] Parlos A G. Identification of the Internet end-to-end delay dynamics using multi-step neuro-predictors[C]// Proc. of the 2002 Int. Joint Conf. on Neural Networks, 2002
- [9] Srikar D. Empirical Modeling of End-to-End Delay Dynamics in Best-effort Networks[D]. Texas A&M Univ, 2004
- [10] Gelenbe E. Stability of the random neural network model[J]. Neural Computation, 1990, 2(2): 239-247
- [11] Gelenbe E. Learning in the recurrent random neural network [J]. Neural Computation, 1993, 5(1): 154-164
- [12] Mohamed S, Rubino G. A study of real-time packet video quality using random neural networks[J]. IEEE Trans on Circuits and Systems for Video Technology, 2002(12): 1071-1083
- [13] Gelenbe E, Liu P, Laine J. Genetic Algorithms for Route Discovery[J]. IEEE Trans on Systems, Man, and Cybernetics, 2006, 36(6): 1247-1254
- [14] Kocak T, Seeber J, Terzioglu H. Design and implementation of a random neural network routing engine[J]. IEEE Trans. on Neural Networks, 2003, 14(5): 1128-1143
- [15] Fujimoto K, Ata S. Statistical analysis of packet delays in the Internet and its application to playout control for streaming applications[J]. IEICE Transactions on Communications, 2001, E84-B(6): 1504-1512
- [16] Mallik K. Predictor development for controlling real-time applications over the Internet[D]. Texas A&M Univ, 2005
- [17] www. caida. org