基于快速独立成分分析的 RoQ 攻击检测方法

荣 宏* 王会梅 鲜 明 施江勇
(国防科技大学电子信息系统复杂电磁环境效应国家重点实验室 长沙 410073)

摘 要：降质服务(Reduction of Quality, RoQ)攻击比传统的拒绝服务攻击(Denial of Service, DoS)攻击更具有隐秘性和多变性，这使得检测该攻击十分困难。为提高检测准确率并及时定位攻击源，该文将攻击流量提取建模为一个盲源分离过程，提出了基于快速 ICA (Independent Component Analysis)的攻击流特征提取算法，从若干观测网络和终端设备中分离出 RoQ 攻击流，然后提取表征攻击流的特征参数。接着设计了一种基于支持向量机的协同检测系统和检测算法，通过用已标记的有攻击和无攻击的样本训练 SVM 分类器，最终实现 RoQ 攻击的检测。仿真结果表明该方法能够有效检测并定位伪造 IP 地址的 RoQ 攻击，检测率达到 90%以上，而选取合适的 ICA 参数会提高检测效果。

关键词：网络安全；降质服务攻击；盲源分离；快速独立成分分析；特征提取
中图分类号：TN393.08 **文献标识码**：A **文章编号**：1009-5896(2013)10-2307-07
DOI: 10.3724/SP.J.1146.2013.00114

A Novel Method for Detecting Reduction of Quality (RoQ) Attack Based on Fast Independent Component Analysis

Rong Hong Wang Hui-mei Xian Ming Shi Jiang-yong
(*State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System, National University of Defense Technology, Changsha* 410073, *China*)

Abstract: RoQ (Reduction of Quality) attack is more stealthy and changeable than traditional DoS (Denial of Service) attack, which makes detection of RoQ extremely difficult. In order to improve detection accuracy and locate attack sources in time, this paper turns modeling attack flow extraction into a process of blind sources separation. A method is proposed based on fast ICA (Independent Component Analysis) to detach RoQ flow from several observation network devices and terminals. Then, some features' parameters that represent attack flow are extracted. After that, a system of collaborative detection system is designed on the basis of SVM (Support Vector Machine), using marked attack and no-attack samples to train the SVM classifier in order to detect RoQ attack finally. Simulation results illustrate that this method can detect IP spoofed RoQ attack as well as locate the attacker, accuracy of which reaches up to 90%. Moreover, choosing appropriate ICA parameters will improve results to some extent.

Key words: Network security; Reduction of Quality (RoQ) attack; Blind sources separation; Fast Independent Component Analysis (ICA); Feature extraction

1 引言

2004 年的 SIGCOMM 会议上，波士顿大学的 Guirguis 等人[1]受低速率 DoS 攻击[2]思想的启发，提出了一类新攻击方法——降质服务攻击(Reduction of Quality, RoQ)，与传统的 DoS 攻击相比，RoQ 攻击隐蔽性较强，不需要维持高速率攻击流，利用了网络中常见的适应性控制机制所存在的安全漏洞[3]，通过周期性地在短暂间隔内发送大量数据包来降低受害者提供服务的质量。由于只在短时间间隔内发送数据包，其它时间静默，这使得攻击流的平均速率较低，很难检测出来[4]。

目前，对 RoQ 攻击防范主要有两类方法：一类是修改有漏洞的协议或参数，如针对 LDoS 攻击的随机化 minRTO 方法[5]，修改 RED 算法的 $Q_{\min}$ 和 $Q_{\max}$ 参数[6]，但由于 Internet 协议已经被用户广泛使用，大规模修改协议不但会损害用户的利益而且也不现实；另一类是从攻击数据流特征角度，通过分析 RoQ 攻击特征来检测攻击的存在，主要有时域、频域和时频双域 3 种方法。

对于时域检测方法：Shevtekar 等人[7]提出通过

2013-01-22 收到，2013-06-05 改回
*通信作者：荣宏 ronghong01@gmail.com

统计相邻数据包到达的时间间隔估计脉冲持续时间和周期，并和 RTT 和 RTO 做比较来进行匹配，该方法只针对 TCP 超时重传引起的 RoQ 攻击。Luo 和 Chang 等人[8]提出采用离散小波变换分离出流入和流出过程，并通过非参数化 CUSUM 算法进行变点检测，该方法可以检测出周期变化的 RoQ 攻击，但误检率较高。Sun 等人[9]使用自相关分析法提取攻击数据流的周期性特征并采用基于动态时间环绕的匹配方法进行检测。该方法计算和存储的开销较大，对变周期检测效果欠佳。

对于频域检测方法：Chen 等人[10]通过傅里叶变换分析网络流量在频域内的能量分布，RoQ 攻击流的能量与正常流量相比主要集中在低频段，并提出针对低频特征的协同检测方法。但是该方法需要做离散傅里叶变换(DFT)，在传统的路由器操作系统上难以实现。

对于时频双域检测方法：文献[11,12]提出了一种基于小波分析的低速率 DoS 攻击单点检测方法和基于支持向量机的 RoQ 综合检测方法，这两种方法都利用了时频双域的 RoQ 攻击特征，对固定周期、变周期的 RoQ 攻击和传统 DoS 攻击有较高的检测率，但定位攻击源较困难。

我们认为，RoQ 的攻击检测应当是建立在区分正常和异常流量的参数模型基础上，对捕获的数据流分析，根据其统计特征判断是否存在攻击行为，系统能及时响应检测到的攻击，并采取相应的防御策略减小攻击带来的影响。即使异常流量只是整体流量的一小部分，也应能较快地检测出异常[13]。

本文首先对典型 RoQ 攻击进行了描述并且建立了盲源分离的数学模型，论证了可分离性；接着，提出了一种基于快速独立成分分析(Independent Component Analysis, ICA)的 RoQ 攻击流分离方法，并提出 5 种可提取的特征参数；然后设计了基于 SVM 的 RoQ 攻击协同检测系统以及具体的检测算法；最后的仿真实验检验了此方法的各项性能。

2 问题描述与分离模型

典型的 RoQ 攻击利用了 TCP 协议拥塞控制机制的缺陷(即检测到网络拥塞后迅速减小拥塞窗口，避免给网络带来更严重的拥塞)，以周期性的时间间隔发送持续时间很短的高强度数据流，造成拥塞窗口始终处于非稳定状态，进而使目标服务性能下降。攻击数据流的形成一般是从攻击者主机出发，然后淹没在正常网络流量中，经过若干路由器之后到达攻击目标。由于数据包的 IP 地址可以动态伪造，无法根据 IP 直接从混合数据流中提取出攻击数据流。

如果能协同地从若干节点的混合流量中提取出含有 RoQ 攻击特征的流量，就能够快速地检测出攻击并实现攻击源定位。我们将攻击流的提取建模为一个盲源分离过程，即从若干观测到的混合流量中提取、恢复出无法直接观测的各个原始流量的过程。对 RoQ 攻击流分离做一般的描述如下：

盲源分离是一个多输入多输出(MIMO)系统，其输入来自若干路由器、网关和终端的数据流量。假设在 m 个网络设备和终端上测得数据流量，$\boldsymbol{x}(k)=[\boldsymbol{x}_1(k),\boldsymbol{x}_2(k),\cdots,\boldsymbol{x}_m(k)]^{\mathrm{T}},i=1,2,\cdots,m$。其中，$k$ 表示时间，$\boldsymbol{x}_i(k)$ 表示数据流，即在时间段 $(t,t+\Delta t]$ 内到达路由器或终端系统的数据包数量，Δt 为采样时间间隔[9]，如图 1 所示。

寻找一个逆系统，由观察数据重构源数据流 $\boldsymbol{s}(k)=[\boldsymbol{s}_1(k),\boldsymbol{s}_2(k),\cdots,\boldsymbol{s}_m(k)]^{\mathrm{T}},i=1,2,\cdots,m$。源数据流 $\boldsymbol{s}_i(k)$ 之间统计独立，源数据流中就包含了攻击者的数据流，它们如何混合也未知。系统输出为

$$\boldsymbol{y}(k)=\boldsymbol{W}\boldsymbol{x}(k)=\boldsymbol{W}\boldsymbol{A}\boldsymbol{s}(k) \tag{1}$$

式中 $\boldsymbol{W}$ 是 $n\times m$ 的解混矩阵，$\boldsymbol{A}$ 是混合矩阵，$\boldsymbol{y}(k)$ 是真实源数据流的一种近似和估计。

根据中心极限定理，若随机变量 $\boldsymbol{X}$ 是由许多相互独立的随机变量 $\boldsymbol{S}_i(i=1,2,\cdots,N)$ 之和组成，只要这些随机变量有有限的均值和方差，则不论其为何种分布，随机变量 $\boldsymbol{X}$ 较 $\boldsymbol{S}_i$ 更接近高斯分布。因此，在分离过程中，可通过对分离结果的非高斯性度量来表示分离结果之间的相互独立性，当非高斯性度量达到最大时，表明已完成对各独立分量的分离。对于正常的网络通信环境，各个节点发送数据的时间和大小是随机不相关的，RoQ 攻击较正常的流量有着周期性、高强度的特点，可以通过非高斯性度量分离出 RoQ 数据流和正常的背景流量。

3 基于快速 ICA 的 RoQ 攻击流分离

根据信息处理的方式不同，盲源分离方法可以分为两大类：一类是基于非高斯最大化原则的批处理算法，如本文采用的快速 ICA 算法；另一类是基于各种代价函数的最小化的自适应算法[14]。快速 ICA，也称固定点算法是一种快速寻优算法，采用了批处理的方式，每一步有大量样本参与运算。该

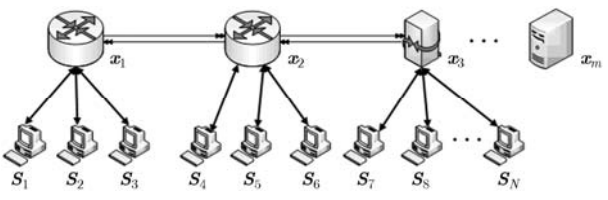

图 1 数据流混合过程示意图

方法收敛速度快、稳定性较强，已经成为提取数据信息特征的有利工具[15,16]。

3.1 基于快速 ICA 的盲分离算法

由信息论的相关理论可知：在所有方差相等的随机变量中，高斯变量的熵最大，因此我们利用熵来度量非高斯性。本文用熵的修正形式，即负熵作为目标函数[17]。

定义负熵如下：

$$N_g(\boldsymbol{Y}) = H(\boldsymbol{Y}_{\text{Guass}}) - H(\boldsymbol{Y}) \tag{2}$$

式中 $\boldsymbol{Y}_{\text{Guass}}$ 是一与 $\boldsymbol{Y}$ 具有相同方差的高斯随机变量，$H(\bullet)$ 是随机变量的微分熵。当 $\boldsymbol{Y}$ 具有高斯分布时，$N_g(\boldsymbol{Y}) = 0$；$\boldsymbol{Y}$ 的非高斯性越强，其微分熵越小，$N_g(\boldsymbol{Y})$ 值越大，所以 $N_g(\boldsymbol{Y})$ 可以作为随机变量 $\boldsymbol{Y}$ 非高斯性的测度。由于 $\boldsymbol{Y}$ 的概率密度分布函数未知，我们采用式(3)所示的近似公式：

$$N_g(\boldsymbol{Y}) = \{E[G(\boldsymbol{Y})] - E[G(\boldsymbol{Y}_{\text{Guass}})]\}^2 \tag{3}$$

快速 ICA 的学习原理是找一个方向使得 $\boldsymbol{W}^{\mathrm{T}}\cdot\boldsymbol{X}(\boldsymbol{Y} = \boldsymbol{W}^{\mathrm{T}}\boldsymbol{X})$ 具有最大的非高斯性。算法推导如下：$\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}$ 的负熵的最大近似值需通过对 $E\{G(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X})\}$ 进行优化来获得。根据 Kuhn-Tucker 条件，在 $E\{(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X})^2\} = \|\boldsymbol{W}\|^2 = 1$ 的约束下，$E\{G(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X})\}$ 的最优值在满足式(4)的点上取得。

$$E\left\{\boldsymbol{X}g\left(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X}\right)\right\} + \beta\boldsymbol{W} = 0 \tag{4}$$

式中 $g(\bullet)$ 是 $G(\bullet)$ 的导数，可以取下面几种函数，$g(u) = u^3$，$g(u) = u\cdot\exp(-a_2\cdot u^2/2)$，或 $g(u) = u^2$。以上 3 种函数分别称为 pow3, gauss 和 skew 函数。β 是一个定值，$\beta = E\{\boldsymbol{W}_0^{\mathrm{T}}\boldsymbol{X}g(\boldsymbol{W}_0^{\mathrm{T}}\boldsymbol{X})\}$，$\boldsymbol{W}_0$ 是 $\boldsymbol{W}$ 的初始值。利用牛顿迭代法解式(4)，可以得到式(5)，式(6)的近似公式：

$$\boldsymbol{W}_{k+1} = \boldsymbol{W}_k - \left[E\left\{\boldsymbol{X}g\left(\boldsymbol{W}_k^{\mathrm{T}}\boldsymbol{X}\right)\right\} - \beta\boldsymbol{W}_k\right] \Big/ \left[E\left\{g'(\boldsymbol{W}_k^{\mathrm{T}}\boldsymbol{X})\right\} - \beta\right] \tag{5}$$

$$\boldsymbol{W}_{k+1} = \boldsymbol{W}_{k+1} \Big/ \left\|\boldsymbol{W}_{k+1}\right\| \tag{6}$$

$\boldsymbol{W}_{k+1}$ 是经 $\boldsymbol{W}_k$ 计算更新后的值，$\beta = E\{\boldsymbol{W}^{\mathrm{T}}\cdot\boldsymbol{X}g(\boldsymbol{W}^{\mathrm{T}}\boldsymbol{X})\}$，简化后得到快速 ICA 算法的迭代公式：

$$\boldsymbol{W}_{k+1} = E\left\{\boldsymbol{X}g\left(\boldsymbol{W}_k^{\mathrm{T}}\boldsymbol{X}\right)\right\} - E\left\{g'\left(\boldsymbol{W}_k^{\mathrm{T}}\boldsymbol{X}\right)\right\}\boldsymbol{W}_k \tag{7}$$

$$\boldsymbol{W}_{k+1} = \boldsymbol{W}_{k+1} \Big/ \left\|\boldsymbol{W}_{k+1}\right\| \tag{8}$$

在执行快速 ICA 算法之前，通常要对数据进行中心化和白化的预处理，去除数据之间的相关性。快速 ICA 采用的牛顿迭代法有至少为 2 次方的收敛速度，可以有效地降低问题的复杂度[18]。算法流程如表 1 所示。

表 1 基于负熵的快速 ICA 的 RoQ 攻击流分离算法

输入：由网络核心路由器观测到的观测数据构成向量 $\boldsymbol{X} = (\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n)^{\mathrm{T}}$，$n$ 是观测点的数目

输出：经快速 ICA 算法分离出的攻击数据流 $\boldsymbol{y}_1$ 和正常数据流 $\boldsymbol{y}_2$

(1)对观测数据 $\boldsymbol{X}$ 进行中心化、白化处理，$\boldsymbol{X} \to \boldsymbol{Z}$；

(2)选择要估计的分量数量总数 num，分离个数 count 设为 1；

(3)选择一个随机的初始权向量 $\boldsymbol{W}_p$；

(4)选择合适的非线性函数 $g(\cdot)$，计算

$$\boldsymbol{W}_p = E\{\boldsymbol{Z}g(\boldsymbol{W}_p^{\mathrm{T}}\boldsymbol{Z})\} - E\{g'(\boldsymbol{W}_p^{\mathrm{T}}\boldsymbol{Z})\}\boldsymbol{W}_p$$

(5)更新 $\boldsymbol{W}_p$：

$$\boldsymbol{W}_p = \boldsymbol{W}_p - \sum_{j=1}^{p-1}(\boldsymbol{W}_p^{\mathrm{T}}\boldsymbol{W}_j)\boldsymbol{W}_j, \boldsymbol{W}_p = \boldsymbol{W}_p / \|\boldsymbol{W}_p\|$$

(6)如果 $\boldsymbol{W}_p$ 不收敛则返回第(4)步；

(7) count = count + 1，如果 count ≤ num，返回第(3)步；

(8)计算得到分离矩阵 $\boldsymbol{W}$，求得分离数据 $\boldsymbol{y}(i) = \boldsymbol{W}_i^{\mathrm{T}}\boldsymbol{Z}$，$i$=1, 2。

3.2 RoQ 攻击特征提取

在快速 ICA 算法下成功分离出 RoQ 攻击流后，应当对攻击流的特征进行提取分析，建立完备的特征库。由于攻击特征的隐蔽性，仅用一方面的特征来描述 RoQ 攻击是片面的，容易造成较高的虚警率。因此，为实现精确检测，本文提出 5 种特征信息来描述 RoQ 攻击，力争多维度全面描述 RoQ 攻击行为。

RoQ 攻击数据流量可以提取的特征参数有：平均网络流量强度(mean)、波动程度(variance)、攻击周期(period)、攻击持续时间(duration)、网络流量脉冲强度(intensity)。

(1)平均网络流量强度。RoQ 攻击的间歇式特性使得即使带有攻击的网络流量，其总体平均速率与正常流量也并无较大差别[11]。但是对于分离出来的流量，其攻击流的平均强度仍强于正常状态。这里用输出 $y_{\text{a}}(n)$ 的采样平均值对其评价，n 表示采样时间。

(2)波动程度。RoQ 的脉冲式特性导致流量强度变化剧烈，$|y_{\text{a}}(n)|$ 波动越大，方差越大，是反映攻击的重要特征之一。

(3)攻击周期。当攻击者发出典型的 RoQ 攻击脉冲后，TCP 发送方(受害者)进入超时重传状态，攻击者在 $\text{RTO}_{\min}$ 的基础上滞后一个 T_{lag} 时间，使得发送方能够维持 RTO 值为 $\text{RTO}_{\min}$，从而下次的攻击时间保持不变，攻击周期为 $\text{RTO}_{\min}+T_{\text{lag}}$，$T_{\text{lag}}$ 一般为 2~3RTT，而 $\text{RTO}_{\min}$ 一般为 1 s, RTT<RTO，

则周期应在 0.5~4 s 之间。

(4)攻击持续时间。攻击持续时间反映了攻击的强度，越短则说明强度越大，效果越好，更加隐蔽，是衡量攻击的一个重要特征。

(5)网络流量脉冲强度。RoQ 攻击发生时，由于其脉冲式特性，反映网络流量脉冲突发强度的脉冲因子将急剧增大。正常网络流量中也有突发数据流，因此我们对所有的极大值点做一个平均。

4 RoQ 攻击检测系统

本文设计了基于快速 ICA 的协同检测系统，主要由 3 部分构成：Agent 探针、数据分析模块和综合诊断模块。Agent 探针的功能是从部署了检测系统的路由器、网关和终端上收集实时流量信息。数据分析模块采用了盲源分离的快速 ICA 算法，从若干收集到的数据中分离出具有 RoQ 攻击特征的流量。

综合诊断模块由一个支持向量机(Support Vector Machine, SVM)二值分类器和攻击源定位模块构成。系统首先用已有的攻击特征样本数据对支持向量机进行训练，然后使用训练好的分类器对待诊断数据机进行分类，判别是否有攻击行为，待诊断数据由实际采集到的特征参数构成。攻击源定位模块根据探针的数据中是否含有攻击流的特征信息进逆向追溯，由于特征信息在攻击汇聚点和终端最为明显，可以定位到距攻击源最近的路由或网关。

4.1 基于 SVM 的检测流程算法

RoQ 攻击检测有两个因素要考虑，一是网络上有很多应用采用了周期性的数据推送或者心跳机制，有些心跳数据的周期和 RoQ 攻击周期类似，增加了检测的难度；二是由于网络设备性能、负载的差异，用传统的基于判断门限的方法难以将攻击分类，而且会造成较高的误检率。

综合上述考虑，本文采用具有学习功能的 SVM 做分类器，将检测归结为一个二值分类问题，即分成有 RoQ 攻击和无 RoQ 攻击两类。SVM 是对小样本统计估计和预测学习的最佳理论，由于二值分类过程采用了较多的特征值，我们引入核函数将其变换到高维空间以得到较好的分类结果，计算复杂度只取决于支持向量的个数。训练特征数据加上相应类标后交给 SVM 用于训练分类器。

在实际应用时，从网络上收集并经快速 ICA 提取的流量特征数据属于待诊断数据，首先根据待诊断数据中的周期是否在 0.5~4 s 内做一个粗分类，然后再用 SVM 分类器做判断，如果待诊断数据被判断为攻击类，系统就结合攻击流的特征信息通过攻击源定位模块追溯攻击源。由于观测数据是通过若干个节点采集的，可以结合攻击特征的强弱判断攻击流量产生的位置。检测系统最终将生成警报并把判断的信息提交给网络管理员。基于 SVM 的检测流程算法如表 2 所示。

表 2 基于 SVM 的检测流程算法

输入：数据流特征参数

输出：判别结果

(1)RoQ 攻击检测系统启动，接收各个核心路由、网关和终端流量统计实时数据；

(2)快速 ICA 算法分离攻击流量并提取特征参数；

(3)if (训练数据)

{加上 RoQ 类标交给 SVM}

else{直接交给 SVM，转(5);}

(4)生成综合诊断模块，转(1)；

(5)if (0.5 s≤周期参数≤4 s)

{基于 SVM 诊断模块进行 RoQ 攻击检测}

else{判断为非 RoQ 攻击，转(1); }

(6)if (检测到攻击)

{通知管理员，生成 RoQ 特征信息报告；

If (某个观测点(除被攻击终端外)的攻击特征最接近 RoQ 特征)

{定位攻击源到这个节点连接的网络，发出 RoQ 攻击预警，采取控制措施；}}

else {转(1); }

(7)算法结束

4.2 检测系统示意图

图 2 所示的是部署了 RoQ 攻击检测系统的示意图，RoQ 检测器是执行快速 ICA 算法提取攻击特征的检测预警中心，包括了数据分析和综合诊断模块。在核心路由器、边界网关和服务器上安装 Agent 探针，负责收集数据(各端口包速率、时间等)并执行 RoQ 检测器检测指令。整个系统是协同工作的：各个节点的 Agent 向 RoQ 检测器传实时流量统计数据，当 RoQ 检测器检测到异常的 RoQ 攻击流量时，就生成警报并将攻击特征信息和之前各 Agent 发来的数据进行比对，找出和攻击特征最匹配的节点就认为是离攻击源最近的节点。如图 2 所示：虚线的箭头表示 Agent 向 RoQ 检测器发送统计信息流，空心箭头表示网络管理员发出的控制命令流，带斜线的箭头表示 RoQ 攻击流，黑色实心箭头表示检测到 RoQ 攻击，追溯攻击源，直到找到最靠近攻击者的路由器或网关，管理员对攻击者的 IP 进行封锁或者流量控制和带宽重分配。

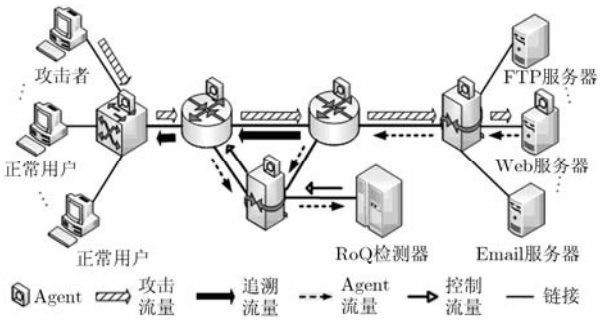

图 2 RoQ 攻击检测系统示意图

5　实验仿真

首先建立有 RoQ 攻击和无攻击的仿真场景，对检测过程中需要的流量数据进行统计，再采用基于快速 ICA 的算法提取攻击特征，加类标后训练 SVM 分类器，经大量样本测试后给出该检测系统的检测率、虚警率、漏检率等性能指标。

基于 OPNET 仿真平台搭建了一个典型的 RoQ 攻击场景。该场景中，一个有 100～150 个节点的 10 BaseT 子网，一个工作站设为攻击者，3 台服务器 (其中 Web 服务器是被攻击的受害者)，背景流量是访问 web, ftp 和 email 服务流量，部署了两台 Cisco 4000 路由器，它们之间链路为整个网络瓶颈，带宽为 10 Mbps，攻击者访问 Web 服务器需要经过这两个路由器。攻击脉冲已经构造好，攻击周期为 1.25 s，脉冲持续时间为 0.1～0.27 s，强度为 100 Mbps，仿真时间为 100 s，攻击开始时间为 70 s。实验场景配置如表 3 所示。

实验通过调整正常用户的数量、设定随机分布的访问服务开始时间和结束时间模拟突发的网络流量。攻击者所发出的流量数据图 3(a)所示，在 70 s 后周期性剧烈震荡。图 3(b), 3(c), 3(d)所展示的是在一次实验中(120 个正常用户)从路由器 1，路由器 2 和被攻击服务器捕获的流量，可以看出 RoQ 攻击发起后，网络流量的波动加剧，但是平均数据率仍然比较低，攻击特征被淹没在正常用户的流量中，无法直接判断是否遭受到攻击。

利用图 3(b), 3(c), 3(d)的数据执行快速 ICA 算法，分离出白化后的含有 RoQ 攻击特征流量和正常

表 3 实验场景配置表

设备名称	个数	服务	设置	开始时间(s)	结束时间(s)
攻击者	1	工作站	10 Mbps 攻击流	70	100
受害者	1	Web 服务	重量级负载	1	100
路由器	2	数据包转发	Cisco 4000	1	100
其它服务器	2	FTP 服务	中级负载	1	100
		EMAIL 服务	中级负载	1	100
合法用户	100~150	工作站	能访问所有服务器	泊松、高斯、均匀、指数分布	可重复直到仿真结束，重复间隔随机分布
路由间链路	1	连接路由器	10 BaseT	1	100
其余链路	5	连接路由和终端	100 BaseT	1	100

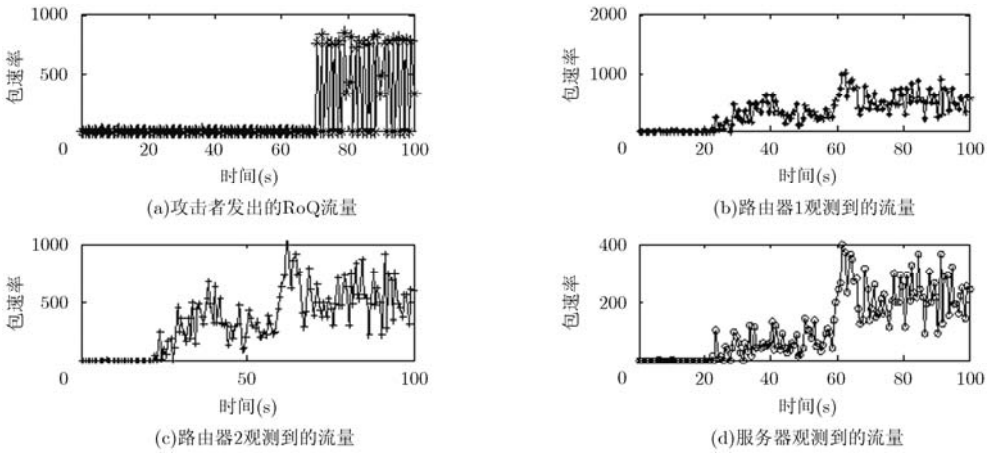

图 3　RoQ 攻击数据流和各观测点观测到的数据流量

用户的流量，其结果如图 4(a), 4(b)所示。图 4(a)表示 RoQ 攻击特征流量，在 70 s 后产生剧烈震荡，0～70 s 之间的小幅度波动是分离出的噪声，这与执行快速 ICA 算法时选取的非线性函数有关。图 4(b)标识正常用户的流量，其突发性与正常情况相符。

根据图 4 得到分离出的流量进行特征提取，分别提取网络流量平均强度、网络流量脉冲强度、波动程度、攻击周期和攻击持续时间 5 个参数。本实验中采用交叉验证方法，使用 150 组含有 RoQ 攻击的特征样本和 150 组正常数据流量的特征样本组成样本集。选取 RBF 核函数作为 SVM 分类器的核函数，用训练样本训练分类器，用测试样本对各项性能指标测试。

图 5 所展示的是执行快速 ICA 算法时选择不同的非线性函数拟合概率分布得出的检测率、虚警率和漏检率，以及它们和训练样本数量的关系。从图 5(a)可以看出检测率随训练样本的增长呈增长趋势，选择 pow3 和 gauss 函数的检测率要高于 skew 函数，但整体都在 85%以上；图 5(b)表示的是虚警率随训练样本比例的变化关系，pow3 和 gauss 函数均在 13%以下，只有 skew 函数的漏检率波动较大；图 5 (c)表示漏检率随着训练样本比例增加而漏检率降低，选择 pow3 和 gauss 函数的漏检率略低于 skew 函数。表 4 给出了选择各函数的性能指标的平均值，选择 pow3 和 gauss 函数的 ICA 算法精度较高，检测率都在 91%以上。这是因为这两种函数能较好拟合所要分离的流量的概率分布。因此，在选择快速 ICA 算法的非线性函数时要根据 RoQ 攻击这种周期性、高强度的特点选择合适的函数。

表 4 不同非线性函数下的平均检测率、漏检率和虚警率表(%)

函数	平均检测率	平均漏检率	平均虚警率
pow3	91.60	8.83	7.97
skew	88.87	13.38	8.88
gauss	91.62	9.12	7.65

实验验证，与 CUSUM 和 DFT 检测法相比，本文的方法在定位攻击源方面的时间开销较小，为其它方法的 1/10。这是因为该方法采用了多个节点的数据，省去了追溯的延时和多个节点 DFT 变换的时间。

6 结束语

现有的 RoQ 攻击检测方法难以兼顾准确率、实时性和对攻击源的定位。基于这种考虑，本文首先对 RoQ 攻击流量提取进行建模，提出了一种基于快速 ICA 的盲分离算法。该算法的优势是收敛速度较快，精度高，为进一步的检测提供了数据支持，并在此基础上设计了基于 SVM 的协同式检测流程算法。

实验证明，本文的检测方法兼顾了 RoQ 攻击检测的精度和效率，能够抵御伪造 IP 包头字段的 RoQ 攻击。后续工作将结合性能更优良的盲源分离算法，研究如何在复杂的 RoQ 攻击环境下(如分布式 RoQ 攻击)进一步提高系统检测率等性能指标。

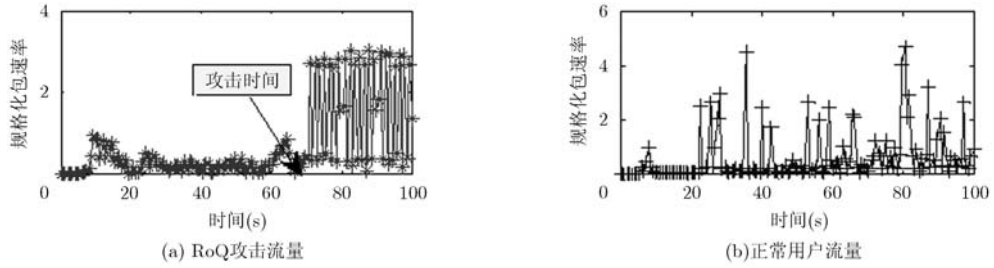

图 4 快速 ICA 算法分离出攻击流和正常的流量

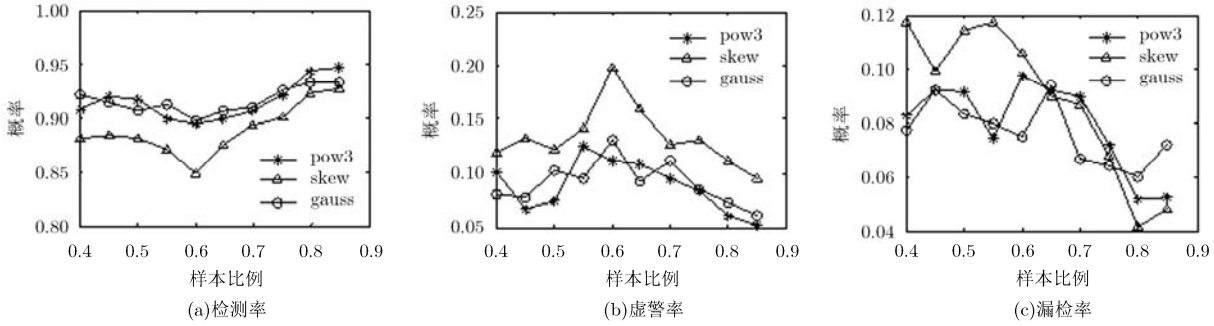

图 5 不同非线性函数下的检测率、漏检率和虚警率比较

参 考 文 献

[1] Guirguis M, Bestavros A, and Matta I. Exploiting the transients of adaptation for RoQ attacks on Internet resources[C]. Proceedings of the 12th IEEE International Conference on Network Protocols, Berlin, 2004: 184–195.

[2] Kuzmanovic A and Knightly E W. Low-rate TCP-target denial of service attacks: the shrew vs. the mice and elephants[C]. Proceedings of the 2003 Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications, New York, 2003: 75–86.

[3] Guirguis M, Bestavros A, Matta I, *et al.*. Reduction of Quality (RoQ) attacks on Internet end-systems[C]. Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Societies, Miami, 2005: 1362–1372.

[4] Qiao Zhu, Zhang Yi-zhi, and Xie Chui-yi. Research and survey of low-rate denial of service attacks[C]. IEEE 13th International Conference on Advanced Communication Technology, Seoul, 2011: 1195–1198.

[5] Kuzmanovic A and Knightly E W. Low-rate TCP-targeted denial of service attacks and counter strategies[J]. *ACM Transactions on Networking*, 2006, 14(8): 683–696.

[6] Zhang Jing, Hu Hua-ping, and Liu Bo. Robustness of RED in mitigating LDoS attack[J]. *KSII Transactions on Internet and Information Systems*, 2011, 5(5): 1085–1100.

[7] Shevtekar A and Ansari N. Do low rate DoS attacks affect QoS sensitive VoIP traffic?[C]. Proceedings of IEEE International Conference on Communications, Istanbul, 2006: 2153–2158.

[8] Luo Xia-pu and Chang R. On a new class of pulsing denial-of-service attacks and the defense[C]. Network and Distributed System Security Symposium, San Diego, 2005: 926–937.

[9] Sun Hai-bin, Lui C S, *et al.*. Defending against low-rate TCP attacks: dynamic detection and protection[C]. Proceedings of the 12th IEEE International Conference on Network Protocols, Berlin, 2004: 196–205.

[10] Chen Yu and Hwang Kai. Collaborative detection and filtering of shrew DDoS attacks using spectral analysis[J]. *Journal of Parallel and Distributed Computing-Special Issue: Security in Grid and Distributed Systems*, 2006, 66(9): 1137–1151.

[11] 何炎祥, 曹强, 刘陶, 等. 一种基于小波特征提取的低速率 DoS 检测方法[J]. 软件学报, 2009, 20(4): 930–941.
He Yan-xiang, Cao Qiang, Liu Tao, *et al.*. A low-rate DoS detection method based on feature extraction using wavelet transform[J]. *Journal of Software*, 2009, 20(4): 930–941.

[12] 何炎祥, 钟海, 刘陶, 等. 基于支持向量机的 RoQ 攻击综合检测方法[C]. 第三届信息安全漏洞分析与风险评估大会, 黄山, 2010: 167–178.
He Yan-xiang, Zhong Hai, Liu Tao, *et al.*. Support vector machine based integrated detection method for RoQ attacks[C]. The 3rd Conference on Vulnerability Analysis and Risk Assessment, Huangshan, 2010: 167–178.

[13] Sathya M John and Vincent Shweta. Survey on network anomaly detection[J]. *International Journal of Computer Science and Management Research*, 2012, 1(3): 600–603.

[14] 欧世峰, 高颖, 赵晓晖. 自适应组合型盲源分离算法及其优化方案[J]. 电子与信息学报, 2011, 33(5): 1243–1247.
Ou Shi-feng, Gao Ying, and Zhao Xiao-hui. Adaptive combination algorithm and its modified scheme for blind source separation[J]. *Journal of Electronics & Information Technology*, 2011, 33(5): 1243–1247.

[15] Peng Tian-liang, Liu Zeng-li, *et al.*. FastICA-EMD algorithm for analysis of the mixed signals in noise[C]. IEEE International Conference on Signal Processing, Communications and Computing, Xi'an, 2011: 1–4.

[16] Wan Xiang-kui and Chen Yi-yi. A fast ICA and its application in VEP feature extraction[C]. IEEE Sixth International Conference on Natural Computation, Yantai, 2010: 3673–3676.

[17] Gao Xiu-mei, Yuan Xiao-hua, *et al.*. One fast and automatic face recognition method[C]. IEEE Eighth International Conference on Fuzzy Systems and Knowledge Discovery, Shanghai, 2011: 1985–1988.

[18] 付卫红, 杨小牛, 刘乃安. 基于四阶累积量的稳健的通信信号盲分离算法[J]. 电子与信息学报, 2008, 30(8): 1853–1856.
Fu Wei-hong, Yang Xiao-niu, and Liu Nai-an. Robust algorithm for communication signal blind separation fourth-order-cumulant-based[J]. *Journal of Electronics & Information Technology*, 2008, 30(8): 1853–1856.

荣　宏：　男，1988 年生，硕士生，研究方向为网络安全.

王会梅：　女，1981 年生，讲师，研究方向为网络安全、电子信息系统建模仿真与评估.

鲜　明：　男，1970 年生，研究员，博士生导师，研究方向为网络安全、电子信息系统建模仿真与评估.

施江勇：　男，1990 年生，硕士生，研究方向为网络安全.