基于全局网络 PCA 的 DDoS 攻击检测方法*

柳 祎[1], 付 枫[1,2], 孙 鑫[2]
(1. 解放军 63778 部队, 黑龙江 佳木斯 154002; 2. 吉林大学 计算机科学与技术学院, 长春 130012)

摘 要: 随着网络规模的不断扩充,对于 DDoS 攻击的集中式检测方法已经无法满足实时性和准确性等要求。针对大规模网络中的 DDoS 攻击行为,提出了一种基于全局 PCA 的分布式拒绝服务攻击检测方法(WPCAD)。该方法由传统的 OD 矩阵得出各节点的 OD_{in} 矩阵,各分布式处理单元通过 PCA 分析到达该节点的多路 OD 流之间的相关性,利用 DDoS 攻击流引起流量之间相关性突变的特性来完成检测。该方法采用分布式处理的方式,降低了检测数据所消耗的带宽,并满足了检测的实时性。实验结果表明该方法具有更好的检测效果。

关键词: 分布式拒绝服务攻击; 全局网络主成分分析; OD 矩阵; 分布式检测
中图分类号: TP393 **文献标志码**: A **文章编号**: 1001-3695(2012)06-2205-03
doi:10.3969/j.issn.1001-3695.2012.06.053

DDoS detection method based on network-wide PCA

LIU Yi[1], FU Feng[1,2], SUN Xin[2]
(1. *Unit* 63778 *of PLA*, *Jiamusi Heilongjiang* 154002, *China*; 2. *College of Computer Science & Technology*, *Jilin University*, *Changchun* 130012, *China*)

Abstract: With the extension of network scale, the centralized detection method against DDoS had failed to meet requirements such as real-time and accuracy. This paper presented a distributed method based on WPCAD to detect increasingly serious DDoS attacks. DDoS attack flows could cause correlation between the abnormal traffic generated by certain tools and originated from different nodes. By taking advantage of this feature, this method first got OD_{in} matrixes from the original OD matrixes, then each processing unit extracted the correlation between potential anomalous traffic by principle component analysis(PCA). This method cunsumed less network bandwidth and met the requirement of real-time with the distributed structure. The experimental results show that this method has better detection effect.

Key words: DDoS attacks; network-wide PCA; OD matrix; distributed detection

随着 Internet 的广泛普及,IPERF、TFN2K 等众多发起 DDoS 攻击的恶意软件可随意下载,导致了 DDoS 攻击的泛滥。如何及时有效地检测出 DDoS 攻击,对于当前的网络安全检测系统来说仍是一个艰巨的任务[1]。已有的一些 DDoS 攻击检测方法常常是通过监视某一条关键链路上的流量特性来进行检测,这些方法仅能掌握网络的局部状况,并不能对分布于整个网络中的攻击行为进行分析。近年来提出了一些全局的网络异常检测方法并逐渐成为研究的热点。Lakhina 等人[2]采用主成分分析(PCA),将高维流量矩阵数据降维后分为正常子空间和异常子空间,然后将数据投影到异常子空间中来检测突变的异常行为。Rubinstein 等人[3]则分析了 PCA 检测方法所存在的缺陷,例如将相关的异常流量部分划分到正常子空间等。Kanda 等人[4]将 PCA 方法与 Sketch 方法相结合,采用动态 Top-K 主成分策略对于全局异常达到了更好的检测效果。

DDoS 攻击是由大量受控制的僵尸主机同时向攻击目标发起恶意连接而形成的,这些攻击流量具有相同的特征,在攻击起始时刻、攻击强度、数据包类型等方面具有相似性。传统 PCA 检测方法只针对不相关的突发异常流量,却将相关的 DDoS 攻击流量划分到正常空间。本文将传统 PCA 方法中的全局 OD 流量矩阵变换为 OD_{in} 输入矩阵,采用 PCA 方法分析其中一个节点的 OD_{in} 输入矩阵就能得到所有到达该节点的各支路流量之间的相关性信息。一个节点的所有输入支路之间的相关性呈现一定的规律性,而 DDoS 攻击会使发起攻击的各支路之间的相关性急剧增加从而脱离正常模式。本文提出了一种基于全局 PCA 的分布式拒绝服务攻击检测方法 WPCAD(network-wide principle component analysis detection)。该方法的优势为:a)可以进行分布式处理,既减轻了节点间的通信负载,又提高了检测运算速度;b)在 OD_{in} 输入矩阵上进行分析,能够在多路攻击 OD 流尚未大规模汇聚前完成检测。

1 OD_{in} 输入矩阵

1.1 OD 矩阵

流量矩阵(TM)是研究整体网络状态的基本输入元素,它可以描述网络中各节点或者各链路之间的流量关系[5],表示为

$$Y = AX \tag{1}$$

其中:$Y=(y_1,\cdots,y_I)'$ 表示所有链路之间流量大小的向量,y_i 表示第 i 条链路上所通过的流量,I 为该网络中链路的数量;

收稿日期: 2011-11-15; **修回日期**: 2011-12-28 **基金项目**: 国家自然科学基金资助项目(60973136)

作者简介:柳祎(1977-),女,辽宁北票人,工程师,硕士,主要研究方向为网络安全、通信系统(mislyxx@126.com);付枫(1984-),男,硕士研究生,主要研究方向为网络安全、网络流量建模;孙鑫(1984-),男,博士研究生,主要研究方向为网络安全、复杂网络.

$\boldsymbol{X}=(x_1,\cdots,x_J)'$表示所有 OD 对之间流量大小的向量,$x_j$ 表示第 j 个 OD 对所通过的流量,J 是该网络中 OD 对的数量,当网络中共有 N 个节点时 $J=N^2$;$\boldsymbol{A}$ 是一个 $I\times J$ 的路由矩阵,其中的每一个元素均为 0 或者 1。当链路 i 属于 OD 对 j 时 $a_{ij}=1$,否则 $a_{ij}=0$。$\boldsymbol{Y}$ 可以通过 SNMP 数据获得,$\boldsymbol{A}$ 可由每个路由器的路由表得出,$\boldsymbol{X}$ 可以直接通过部署 NetFlow 软件获得,或者由已知的 $\boldsymbol{Y}$ 和 $\boldsymbol{A}$ 通过式(1)得出。

采用一个固定的时间间隔取得网络中所有 OD 流状态的 t 个采样时段,这些采样点形成 $t\times J$ 的 OD 流矩阵 $\boldsymbol{TX}$。$\boldsymbol{TX}$ 中的第 i 行表示第 i 个采样点网络中所有 J 个 OD 流的状态,第 j 列表示第 j 个 OD 流在所有 t 个采样时段下的状态。传统的 PCA 方法通过分析 $\boldsymbol{TX}$ 矩阵,将降维生成的多个主成分向量分别划分到正常子空间和异常子空间从而进行检测。本文则将 $\boldsymbol{TX}$ 矩阵变换为 $\mathrm{OD_{in}}$ 输入矩阵,从而能反映汇聚到同一个目的节点的所有 OD 流量之间的相关性,同时也更加有利于分布式处理。

1.2 $\mathrm{OD_{in}}$ 输入矩阵生成方法

取矩阵 $\boldsymbol{TX}$ 中的某一行 X_i,X_i 包含第 i 个采样时段网络中所有 J 个 OD 流的状态。网络共有 N 个节点,$J=N^2$,则可以将 X_i 转换为 $N\times N$ 的矩阵 $\boldsymbol{S}$,其中的每个元素 s_{ij} 表示起点 i 到终点 j 的流量大小。因为从起点 i 到终点 j 的流量和从起点 j 到终点 i 的流量不一定相等,所以 $\boldsymbol{S}$ 并不是一个对称矩阵。

考虑网络中的某一个节点 k,它在 $\boldsymbol{S}$ 的第 k 行 $s_{k*}=(s_{k1},\cdots,s_{kN})$ 表示在第 i 个采样时段节点 k 向网络中所有 N 个节点发出的流量;$\boldsymbol{S}$ 的第 k 列 $s_{*k}=(s_{1k},\cdots,s_{Nk})'$ 则表示在第 i 个采样点时网络中所有 N 个节点向节点 k 发出的流量。将所有 t 个采样时段形成的 $\boldsymbol{S}$ 中有关节点 k 的列向量组成 $t\times N$ 的矩阵 $\boldsymbol{TS}_k^{\mathrm{in}}$,$\boldsymbol{TS}_k^{\mathrm{in}}$ 中的每一行表示某一采样时段网络中所有节点向终止节点 k 所发出的流量。当然也可以将所有 t 个采样点形成的 $\boldsymbol{S}$ 中有关节点 k 的行向量组成 $t\times N$ 的矩阵 $\boldsymbol{TS}_k^{\mathrm{out}}$,则 $\boldsymbol{TS}_k^{\mathrm{out}}$ 中的每一行表示某一采样点时节点 k 向网络中所有节点所发出的流量。本文主要研究以节点 k 为终点的多路流量之间的相关性,所以主要采用 $\boldsymbol{TS}_k^{\mathrm{in}}$,其他节点的 $\mathrm{OD_{in}}$ 输入矩阵的计算方法与节点 k 相同。

2 WPCAD 方法原理及步骤

2.1 WPCAD 检测原理

主成分分析(PCA)是把原来多个变量划为少数几个综合指标的一种统计分析方法。从数学角度来看,这是一种降维处理技术[6]。PCA 将待处理数据映射到一组新的坐标轴上,这些新的坐标轴就叫做主成分。每个主成分之间是彼此正交的。对 $\mathrm{OD_{in}}$ 输入矩阵 $\boldsymbol{TS}^{\mathrm{in}}$ 应用 PCA 方法,求出该矩阵的主成分就相当于求 $\boldsymbol{TS}^{\mathrm{in}'}\boldsymbol{TS}^{\mathrm{in}}$ 的特征向量。主成分 $\boldsymbol{v}_i$ 是通过 $\boldsymbol{TS}^{\mathrm{in}'}\boldsymbol{TS}^{\mathrm{in}}$ 计算出的第 i 个特征向量,如式(2)所示:

$$\boldsymbol{TS}^{\mathrm{in}'}\boldsymbol{TS}^{\mathrm{in}}\boldsymbol{v}_i=\lambda_i\boldsymbol{v}_i \tag{2}$$

其中:$i=1,\cdots,N$,λ_i 为特征向量 $\boldsymbol{v}_i$ 所对应的特征值且 $\lambda_1\geqslant\lambda_2\geqslant\cdots\geqslant\lambda_N$。$\boldsymbol{v}_1$ 作为 $\boldsymbol{TS}^{\mathrm{in}}$ 的第一主成分向量,捕获了数据中关联性最强的部分,可以表示为

$$\boldsymbol{v}_1=\arg\max_{\|v\|=1}\frac{v'\boldsymbol{TS}^{\mathrm{in}'}\boldsymbol{TS}^{\mathrm{in}}v}{v'v} \tag{3}$$

其他的主成分 v_k 可以表示为

$$v_k=\arg\max_{\|v\|=1}\|(\boldsymbol{TS}^{\mathrm{in}}-\sum_{i=1}^{k-1}\boldsymbol{TS}^{\mathrm{in}}\boldsymbol{v}_i\boldsymbol{v}'_i)v\| \tag{4}$$

$\boldsymbol{TS}^{\mathrm{in}}$ 共会产生 N 个特征向量,但这 N 个特征向量的前 r 个就能很好地近似 $\boldsymbol{TS}^{\mathrm{in}}$,此时 r 明显小于 N。前 r 个特征向量为 $\boldsymbol{TS}^{\mathrm{in}}$ 的本质有效成分,其他部分只包含了很少的信息可以被省略掉。

网络中的可疑异常流量不都是由 DDoS 攻击引起的,例如瞬间聚集的大规模流量就会常常引起 DDoS 网络检测系统的误报。由于网络流量具有突发、重尾和长相关等特性,这种流量的突变是时常发生的。有效地检测 DDoS 攻击并尽可能地降低误报,就是要能区分 DDoS 攻击流量与瞬时突发流量之间的区别。DDoS 攻击常常是由位于不同地域的多个僵尸网络在同一种或同一类型攻击软件控制下同时向目标主机发起异常访问,多路攻击流量之间在攻击开始时间、间隔时间、攻击类型等多方面存在一定的相似性,导致了到达目标主机的多个 OD 流之间的相关性也同时增强,而突发流量的多路 OD 流之间却不具有这种明显的相关性。本文利用 DDoS 攻击爆发时会引起到达目标节点的多路 OD 流之间相关性明显上升的特性来进行 DDoS 攻击检测,这种相关性的测量通过对 $\mathrm{OD_{in}}$ 输入矩阵进行主成分分析来完成。

2.2 WPCAD 步骤

WPCAD 方法是由很多独立的处理单元组成,这些处理单元可以运行在一台主机上,也可以分别运行在不同的主机上。WPCAD 方法的具体步骤如图 1 所示。

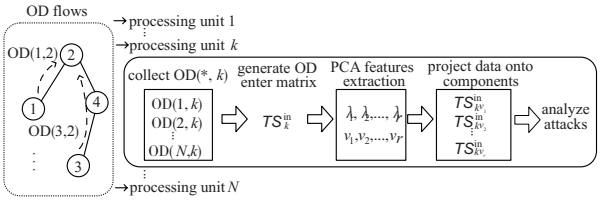

图1 WPCAD的步骤

WPCAD 方法首先要完成对网络中所有 OD 流状态的收集,将与某一处理单元相关的 OD 流状态数据传递给该处理单元所在的主机。这种 OD 流状态数据的分向传递并不需要通过一个统一的处理中心,而是网络中的每个节点自动完成的。传统集中式处理的 DDoS 检测方法常常需要网络中的全部节点把所有信息都传递给中心处理主机,而 WPCAD 方法使各节点只向每个处理单元提供该单元所需的信息即可。这样降低了处理单元所在链路的负荷,提高了检测运算速度。如图 1 所示,以处理单元 k 为例,处理单元 k 收集目的端为节点 k 的所有 OD 流信息,通过这些信息构造 k 节点的 OD 流输入矩阵 $\boldsymbol{TS}_k^{\mathrm{in}}$。对 $\boldsymbol{TS}_k^{\mathrm{in}}$ 进行主成分分析得出 N 个特征值及对应的特征向量,选取其中的前 r 项作为 $\boldsymbol{TS}_k^{\mathrm{in}}$ 的有效成分。分别计算 $\boldsymbol{TS}_k^{\mathrm{in}}$ 在 r 个特征向量下的投影,当计算结果大于门限 M 时,表示该时刻有指向节点 k 的攻击发生。预设门限值与网络的实际情况有关,可以通过对历史数据的分析得到,选取可接受误报率条件下的阈值作为门限值。

3 实验与讨论

由于很难获得具有准确攻击标记信息的网络流量数据,本

文将从真实网络环境中收集到的流量信息作为背景数据,通过注入强度可调的 DDoS 攻击流量数据来验证本文方法的有效性。本文采用的背景流量数据集来自于欧洲科研与教育网 GÉANT[7],这个数据集由 GÉANT 中以每 15 分钟为间隔所取得的网络 OD 流矩阵构成,共进行采样 4 个月。GÉANT 主要连接欧洲的大学和科研机构,在每一个欧洲国家都布置有一个 PoP 节点,共由 23 个边界路由器和 38 条主干链路组成。

WPCAD 方法首先将 GÉANT 数据集中以 XML 格式存储的 OD 流矩阵变换得到各个节点的 OD_{in} 输入矩阵,然后对其进行主成分分析。以该数据集中编号为 7 的节点一周的流量数据为例,对所形成的 OD_{in} 输入矩阵 TS_7^{in} 进行主成分分析,结果如图 2 所示。

对 TS_7^{in} 进行主成分分析所得到的前五项主成分的累积贡献率达到了 82.073%,所以只采用这五个主成分就能很好地刻画所有到达节点 7 的多路 OD 流之间的关系。计算 TS_7^{in} 在这五个主成分上的投影,就能有效地捕获到达该节点的多路 OD 流之间相关性的变化情况,分析其映射到第一主成分上的曲线,如图 3 所示。

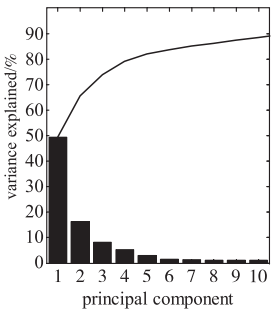

图2　各个主成分的贡献率

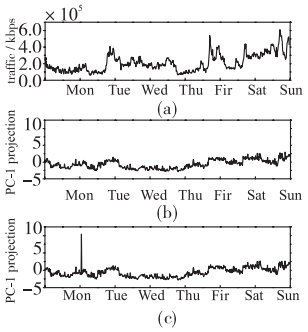

图3　TS_7^{in}在第一主成分上的投影曲线

选取了节点 7 一周的流量数据,图 3 (a)表示网络中的所有节点向节点 7 传输的数据包之和,也就是节点 7 的总输入流量;(b)列出了 TS_7^{in} 在第一主成分上的投影曲线。由图 3(a)和(b)的对比可知各支路流量之间的相关性与到达节点的总流量大小没有必然联系,(a)中出现的突发流量是由到达相同目的节点的各流量汇聚而成,但是各流量之间相关性不强,也就是说并不是同增或同减,所以没有引起第一主成分上的突变。关于其他主成分上的变化情况将在图 4 中进行分析。在图 3(a)所示的背景流量下,于周二 1∶15 ~ 1∶30 向其他各节点注入攻击强度为2 000 kbps,指向节点 7 的 DDoS 攻击流量,节点 7 在这时间段内原输入流量为133 470 kbps,加入攻击流量后输入流量为179 470 kbps,并没有造成流量上的突变;(c)显示了发动 DDoS 攻击时节点 7 的 OD_{in} 输入矩阵在第一主成分上的突变,未发生攻击时该点值为 -0.978,攻击发生时刻该点值跳变为 10.435。图 3 通过对比节点在正常时刻和 DDoS 攻击爆发时刻其 OD_{in} 输入矩阵在第一主成分上的投影曲线变化情况,验证了 WPCAD 方法的有效性。同时为了考察 DDoS 攻击对其他主成分的影响,采用在1 ~ 5 000 kbps的攻击强度下分别列出了前五项主成分的数值变化情况,如图 4 所示。

由图 4 可知在每一种攻击强度下第一主成分(PC-1)的变化都要明显于其他主成分,所以通过分析节点 OD_{in} 输入矩阵映射到第一主成分上值的变化,就能够达到检测 DDoS 攻击的目的。为了进一步验证 WPCAD 方法的有效性,将该方法与文献[3]中 PCA 与 Sketch 相结合的方法进行对比实验。由 GÉANT 数据集包含的 23 个节点,每个节点设置含有 100、200、500、1 000、2 000 kbps五种攻击强度的 115 个攻击样本,并且每个节点随机抽取 10 段背景流量作为正常样本,共组成含有 345 个样本的测试数据集。通过调节门限值 M 对比两种检测方法的 ROC 曲线如图 5 所示。

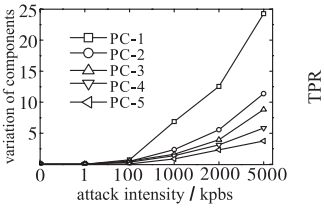

图4　各主成分值在不同强度DDoS攻击下的变化

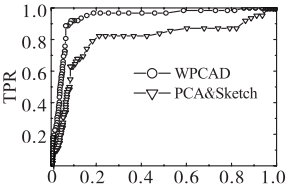

图5　ROC曲线

在图 5 中每一个相同的误报率下,WPCAD 方法相比文献[3]中 PCA 与 Sketch 相结合的方法具有更高的检测率,并且在 ROC 曲线下的面积大于文献[3]中的方法,说明了本文提出的 WPCAD 方法具有更好的检测性能。

4　结束语

本文研究了网络流量中多路 OD 流之间的相关特性,针对 DDoS 攻击提出了一种有效的分布式检测方法,实验结果表明该方法具有较高的检测率同时具有较低的误报率。本文的主要贡献:a)在原 OD 流矩阵基础上生成了新的 OD_{in} 输入矩阵,该矩阵包含了到达同一目的端所有 OD 流的信息;b)利用 PCA 方法从全局的角度通过分析多路 OD 流之间的相关性来进行 DDoS 检测,提出了一种 WPCAD 方法;c)该方法易于分布式部署,适于网络规模的扩充。

参考文献:

[1] LOUKAS G, OKE G. Protection against denial of service attacks: a survey[J]. Computer Journal, 2010, 53(7): 1020-1037.

[2] LAKHINA A, CROVELLA M, DIOT C. Diagnosing network wide traffic anomalies[C]// Proc of ACM SIGCOMM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications. New York: ACM Press, 2004: 219-230.

[3] RUBINSTEIN B, NELSON B, HUANG Ling, *et al*. Evading anomaly detection through variance injection attacks on PCA[C]// Proc of the 11th International Symposium on Recent Advances in Intrusion Detection. Berlin: Springer-Verlag, 2008: 394-395.

[4] KANDA Y, FUKUDA K, SUGAWARA T. Evaluation of anomaly detection based on sketch and PCA[C]// Proc of IEEE Global Telecommunications Conference. Miami: IEEE Press, 2010: 1-5.

[5] LIU Zi-qian, CHEN Chang-jia. Principal component analysis on estimated OD flows[C]//Proc of the 7th International Conference on Electronic Measurement & Instruments. Beijing: Academin Publishers LTD, 2005: 177-183.

[6] HUBERT M, ENGELEN S. Fast cross-validation of high-breakdown resampling methods for PCA[J]. Computational Statistics and Data Analysis, 2007, 51(3): 5013-5024.

[7] UHLIG S, QUOITIN B, LEPROPRE J, *et al*. Providing public intradomain traffic matrices to the research community [J]. Computer Communication Review, 2006, 36(1): 83-86.