

# 一种基于历史信任数据的 DDOS 防御模型

李金良, 王文国, 何裕友

(曲阜师范大学 计算机科学学院, 山东 日照 276826)

**摘要:** 分布式拒绝服务攻击给网络安全和网络服务质量带来了巨大的威胁。通过对分布式拒绝服务攻击原理及现有防御措施的分析, 为了更有效防御这类攻击的发生, 可以考虑在边界路由器上建立一种基于历史信任数据的源地址库的防御模型。该模型以历史信任数据库为依托, 通过对异常 IP 包使用核心无状态公平排队算法进行源地址检测并对其处理结果做出相应的处理, 可以有效、快速过滤掉异常的 IP 包数据, 提前防止网络受到分布式拒绝服务攻击的侵害。

**关键词:** 分布式拒绝服务攻击; 历史信任数据; 异常 IP 包; 源地址检测; 核心无状态公平排队算法

**中图分类号:** TP393.08

**文献标识码:** A

**文章编号:** 1673-629X(2007)07-0160-03

## A Model Based on Historical Trusted Data to Defense DDOS

LI Jin-liang, WANG Wen-guo, HE Yu-you

(Dept. of Computer Science, Qufu Normal University, Rizhao 276826, China)

**Abstract:** The distributed denial of service attacks brought an enormous threat to network security and network quality of service. Through analysing principle of distributed denial of service attack and the existing preventive measures, in order to more effectively defend against such attacks, the border routers can be taken to establish a defense model based on historical trusted data of the source address. The model is based on a historical trusted database, and to the abnormal IP used CSFQ source address detection and makes treatment with outcome of the algorithm. The model can be effective, rapid filter abnormal IP packet data, in advance to prevent networks against distributed denial of service attack.

**Key words:** DDOS; historical trusted data; abnormal IP; source address detected; CSFQ

## 0 引言

近年来, 蠕虫 (Worm)、分布式拒绝服务攻击 (DDOS) 和垃圾邮件 (Spam) 已经成为当今网络安全领域面临的三大威胁, 并在世界各国引起了高度重视。DDOS 以其容易实施、难以防范、难以追踪等特点成为当前最常见的攻击技术, 极大地影响了网络和业务主机系统的有效服务<sup>[1]</sup>。全球包括 Yahoo, CNN, eBay 在内的著名网络都曾遭受过 DDOS 攻击, 使得公司损失惨重。同时, 分布式拒绝服务攻击也必将成为未来信息战的重要手段之一。因此, 研究分布式拒绝服务攻击及其防御对策是极为重要的。

针对日益严重的 DDOS 攻击, 人们也提出了各种防御措施<sup>[2]</sup>, 比如边界过滤、速率限制、随机丢包、SYN

Cookie、SYN Cache、消极忽略等, 但是由于 DDOS 攻击中使用了源 IP 地址伪装技术, 在这些防御方法中大都无法区分正常和恶意 IP 包, 并且通常在它们所丢弃的 IP 包中也包含了大量的合法 IP 包, 因此在高强度的 DDOS 攻击中, 这些措施的防御效果就比较差。

## 1 DDOS 攻击的原理

分布式拒绝服务攻击 (DDOS, Distributed Denial of Services) 是指采用分布、协作的大规模的拒绝服务攻击 (DOS) 方式, 攻击者联合或控制网络上能够发动攻击的若干主机同时发动攻击, 制造数以百万计的数据流入欲攻击的目标, 消耗其网络带宽或者系统资源, 致使目标主机的服务请求极度拥塞, 无法提供正常的网络服务<sup>[3,4]</sup>。

DDOS 的攻击原理: 如图 1 所示, 攻击者首先将入侵主控端并安装拒绝服务攻击程序 (主控端可能是一些防护能力较差的服务器, 也可能是存在系统漏洞或配置上错误的主机), 然后攻击者通过主控端来入侵更多的傀儡机并安装拒绝服务攻击程序。分布式拒绝服

收稿日期: 2006-09-23

基金项目: 国家人事部高层次留学人员回国工作资助项目 (国人部发[2004]61 号)

作者简介: 李金良 (1982-), 男, 山东潍坊人, 硕士研究生, 研究方向为网络安全; 王文国, 博士, 教授, 硕士研究生导师, 研究方向为网络通信与信息安全。



务攻击的程序一般分为守护程序与服务端程序,这些程序可以协调,使分散在互联网各处的计算机共同完成对一台主机攻击的操作。攻击者控制多个主控端,每一个主控端又控制多个傀儡机。这样攻击者就可以在特定的时间对其所控制的主机发布命令,让它们来访问特定的目标网络,以达到DDOS的目的。另外,攻击者还可以通过设定访问指令的并发任务个数、重复次数、包长等来调节攻击的强度。这样即使处理能力最强、带宽最大的服务器资源也会被耗尽。

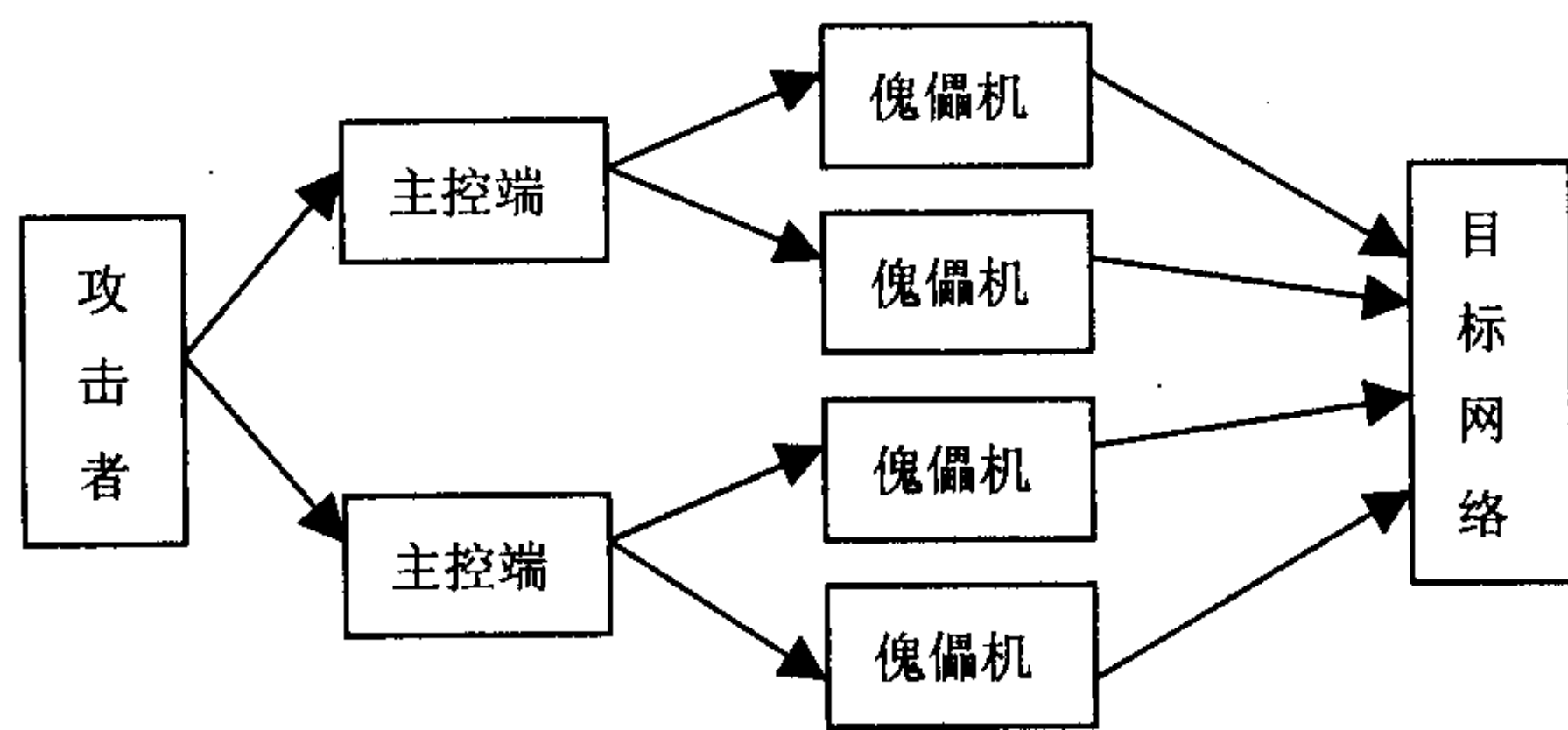


图1 DDOS攻击原理图

反弹分布式拒绝服务攻击(DRDoS)比分布式拒绝服务攻击多了一个反弹器,攻击者或其控制下的傀儡机并不是直接向受害者发送攻击数据包,而是向第三方的反弹器发送特定的数据包(如TCP,UDP和各种ICMP消息协议等具有自动消息生成的协议),反弹器根据消息自动生成协议规定,自动生成攻击者所希望的数据包,发送到目标网络,形成反弹分布式拒绝服务攻击。

僵尸网络(BOTNET)是指攻击者利用网络搭建的可以集中控制、可以相互通信的计算机机群。根据CipherTrust网站对Zombies的统计,每天出现在中国的Zombies大约有3~5万个,可见其规模之大、发展速度之快、破坏能力之强<sup>[5]</sup>。使用僵尸网络发起DDOS攻击也是当前僵尸网络的主要威胁之一。由于僵尸网络可以形成庞大规模,而且利用其进行DDOS攻击可以做到更好的同步,还可以完全使用正常的访问指令,等等,这都使得这种DDOS比原来的DDOS手段危害更大、防范更难。

## 2 基于历史信任数据的DDOS防御模型

### 2.1 模型的提出

为了有效地解决目前DDOS的攻击破坏行为,考虑在边界路由器与主机网络之间部署这样一个基于历史信任数据的模型。那么为什么要选择在边界路由器上部署这样的一个模型呢?主要是因为边界路由器相对于核心路由器来说,它有更多的CPU空闲时间。并且在DDOS攻击过程中,为了减轻路由器的负担,模型设置只有在进入路由器的IP包数据发生异常时才启

动源地址检测,这样可以节省CPU的开支。

模型大致思想如下:以通过边界路由器的历史IP包数据为参考,建立一个带有信任机制的源IP地址库,赋予每一IP一定的信任级别,并随时间、访问次数、访问频率等参数实时修改其信任级别,以提供良好的服务。当通过路由器的IP包数据发生异常时,启动模型的源地址检测功能,首先通过拥塞控制算法对数据进行处理,以选择有效数据进入下一步;然后对上一步的数据判断是否是信任数据库里的数据,并做出相应的处理;对于存在于信任数据库里的数据判断其是否为合法源地址,并做相应的处理。

模型要正确有效地建立起来,关键在于怎样建立信任数据库和怎样来进行源地址检测。

### 2.2 相关概念定义

下面给出文中涉及到的几个概念的抽象数据类型定义<sup>[6]</sup>。

定义1:IP包:表示经由边界路由器进入内网的IP包数据(可能是TCP包或UDP包),用一元组Packet(Sadd)来表示,Sadd表示该IP数据包的源IP地址。

定义2:源地址项:记录经由边界路由器进入内网的IP包所使用过的源地址,用三元组Sitem(Sadd,N,Timeout)来表示。其中,Sadd表示一个使用过的源地址;N表示对该源地址所设置的信任度,N的初值为1,最大值为10000;Timeout表示该地址项的时间戳,可以设为7天。

定义3:信任数据库(RD,Reputation Database):所有源地址项的集合。

定义4:合法源地址:在进行源地址检测时,如果源地址项的N达到了临界值 $N_0$ ,那么这个源地址项的IP地址为合法的源地址(临界值 $N_0$ 的选取为所有Sitem项中N的平均值的 $1/3$ 取整)。

### 2.3 信任数据库的建立

为了防止在发起攻击前,攻击者先发送少量的恶意IP包(其中也包括非法的源IP地址)让信任数据库为其建立非法源地址项的行为,在通过边界路由器的网络流量正常时,模型按一定的概率 $P$ 统计已成功建立访问连接的IP包的源地址(包括TCP,UDP,ICMP数据包等,但对三次握手所使用的TCP SYN包不做统计)。并且提取该数据包的源地址项Sadd存入信任数据库中。

对信任度及信任数据库的处理。在提取的IP包第一次访问时,记 $N=1$ ,并连同该IP包的源IP地址Sadd写入信任数据库,以后每次正常访问时 $N++$ 。在整个过程中,若 $N$ 小于1或大于10000,则自动将信任库中该IP包的数据项删除。对于超出规定时间



Timeout 的源 IP 地址也应该从信任库中删除。

### 2.4 源地址检测

当有异常 IP 包进入边界路由器时,启动源地址检测功能。这里把异常 IP 包的进入看作是即将发生拥塞,并且当作拥塞来处理。所以选择拥塞控制算法对异常的 IP 包进行控制,这里选择核心无状态公平排队算法。

核心无状态公平排队算法(CSFQ, Core- Stateless Fair Queuing)的主要特点:将路由器划分为边界路由器和核心路由器。在典型的部署中,边界路由器会处理上千个流,而核心路由器会处理 50k~100k 个流。CSFQ 利用这个差异,指定边界路由器来管理基于每个流的状态,每个流的状态以标签的形式表示;核心路由器依次使用标签来为每个输入流公平地分配带宽。标签初值是边界路由器根据每个流的情况设定的,标签值的更新是基于路由器的聚集信息的。边界路由器记异常 IP 包进入的速率是  $r(t)$ ,并按速率  $C$ (可选路由器的出口速率)往所设置的缓冲区送出 IP 包数据。异常 IP 包进入算法模块后,算法对每一异常 IP 包进行管理,首先读出 IP 包的速率  $r(t)$ :如果  $t_i^k, l_i^k$  分别为流  $i$  中第  $k$  个报文的到达时间和长度,则  $r_i(t)$  按照如下公式计算:

$$r_i^{new} = (1 - e^{-T_i^k/K}) l_i^k / T_i^k + e^{-T_i^k/K} r_i^{old}$$

其中,  $T_i^k = t_i^k - t_i^{k-1}$ , 并且  $K$  是常量。 $T_i^k$  是第  $k$  个报文到达时间差,  $t_i^k - t_i^{k-1}$ 。

这样给定一组流的到达速率,  $r_1(t), r_2(t), \dots, r_n(t)$ , 算法按照 max-min 确定公平共享带宽  $\alpha$ ,  $\alpha$  确定了算法处理报文的速率:

$$\sum_{i=1}^n \min(r_i, \alpha) = C$$

这样可以得出随机的报文丢弃概率:

$$P = \max(0, 1 - \alpha / r_i(t))$$

这样异常 IP 包通过 CSFQ 算法的处理后,将处理后的 IP 包送入到一个缓冲区内。当所送出的 IP 包大小达到所设置缓冲区的最大值时将 IP 包集合送回源地址检测模块。然后在信任库中查询 IP 包集合中 IP 包的 Sadd, 若存在,则判断该 IP 包是否为合法源地址,合法,将对应  $N++$  并转发该 IP 包,若不合法,则丢弃该 IP 包并置  $N--$ ; 若不存在,则丢弃该 IP 包。

基于历史信任数据的防御 DDOS 模型如图 2 所示。

### 3 算法形式化描述

综上所述的讨论,模型的形式化描述如下:

main()

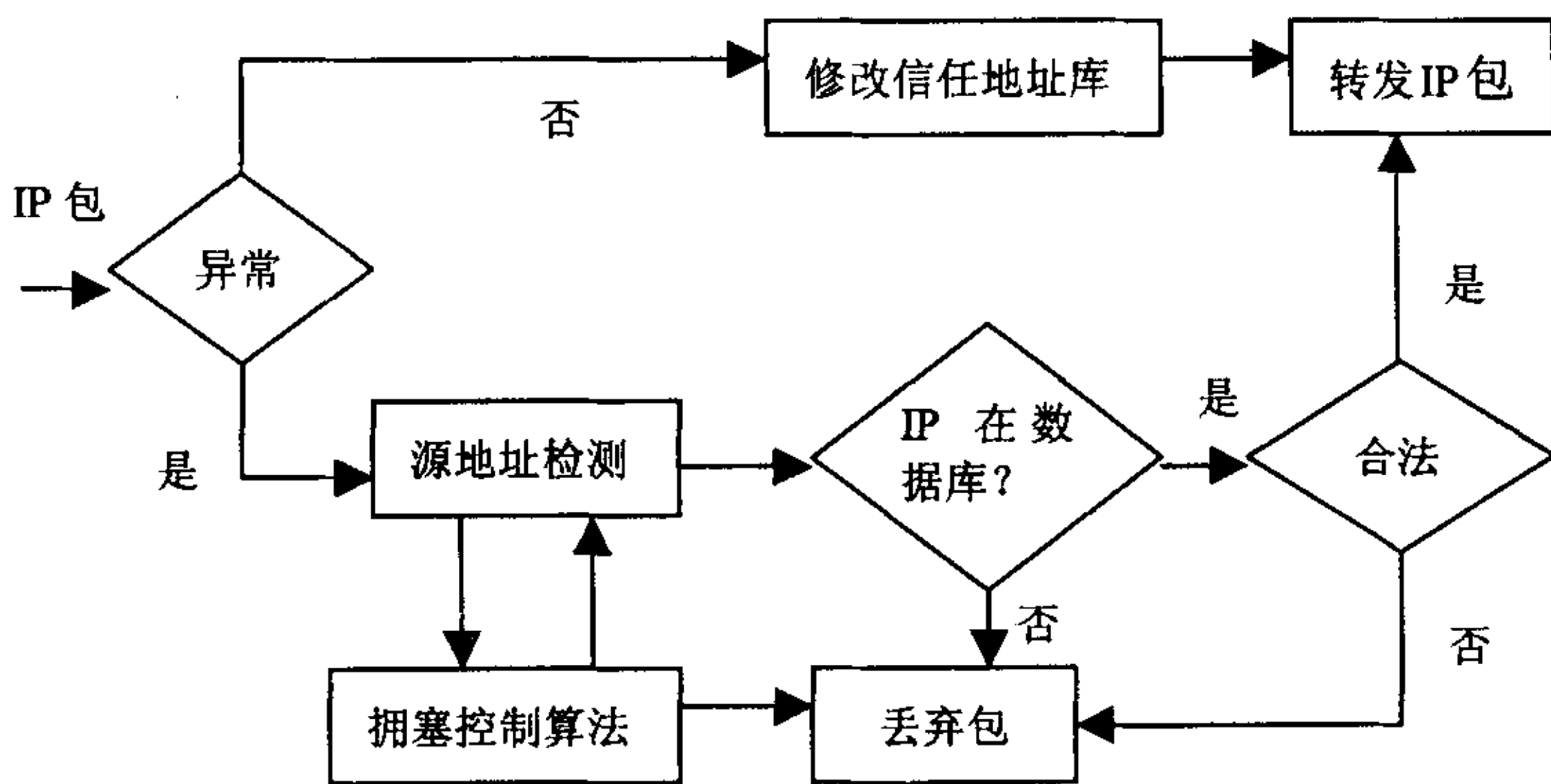


图 2 基于历史信任数据的防御 DDOS 模型

```

{如果网络流量正常
{边界路由器按概率 P 抓包;
  如果 IP 包在信任库中不存在
    {如果 IP 包不是 TCP SYN 包
      {取该 IP 包的 Sadd;
        N = 1;
        Timeout = 7;
        将该 IP 包三元组 Sitem(Sadd, N, Timeout)写入信任
        数据库;
      }
    }
  else //说明该 IP 包已经存在于信任库中
    { N++;
      修改该 IP 包所对应的三元组;
    }
  }
else //说明网络流量出现异常
  启动源地址检测功能;
  对于源地址检测后的 IP 包集合取出每一个 IP 包
  如果该 IP 包不在现在的信任数据库中
    丢弃该包;
  else //说明该包目前在信任数据库中
    {计算此时 N0 = 所有 N 值的 1/3 取整;
      取出该 IP 包所在三元组 Sitem;
      如果三元组 Sitem 的 N 值 >= N0 //该 IP 包合法
        修改 Sitem, N++;
      else //该 IP 包不合法
        {丢弃该包;
          N--, 修改 Sitem;
        }
    }
}

如果信任库中源地址项 Sitem 的 N 值小于 1 或大于 10000
  将该源地址项删除;
}
    
```

### 4 结束语

DDOS 攻击所造成的危害日益严重,各种传统的

(下转第 199 页)



始数据如表 3 所示。

表 3 用例缺陷记录表

符号	<i>i</i>	<i>U</i>	<i>N<sub>ui</sub></i>	<i>S<sub>ui</sub>M<sub>ui</sub>T<sub>ui</sub></i>	<i>W</i>
含义	阶段编号	用例编号	某用例阶段缺陷总数个数	分别表示严重、中等、轻度缺陷总数	缺陷权重

说明:  $W_s$  = 严重缺陷权重(缺省为 10);  $W_m$  = 中等缺陷权重(缺省为 3);  $W_l$  = 轻度缺陷权重(缺省为 1)。

在每个软件开发阶段,根据缺陷严重性和数量获取公式(6):

$$P_i = W_s \times (S_{ui}/N_{ui}) + W_m \times (M_{ui}/N_{ui}) + W_l \times (T_{ui}/N_{ui})$$

(6)

用例缺陷指数(DI) 通过把软件开发各阶段的  $P_i$  通过下面公式(7) 计算得到:

$$DI = \frac{\sum_{i=1}^n i \times P_i}{CU} = (P_1 + 2P_2 + 3P_3 + \cdots)/CU$$

(7)

软件生命周期各阶段也被标以一个权重,软件开发的进展越多,例如阶段 2 或阶段 3,那么赋予的权重越大(也就是 2 或 3),最后把结果按该用例的复杂度(CU)规格化。该指标可以将从以往的项目或用例得到的数据用作比较的基线数字,并将为下次迭代更新基线。可以确定的是,相关用例的缺陷密度和缺陷指数在用例的审查和评估中,能提供客观有效的参考数据。

4 总 结

对用例相关的度量,由于度量对象的特殊性导致该研究具有独特的意义:

(1)预防性:与以往的一些研究(比如针对源代码的度量)不同,这里是针对需求分析模型和设计模型进

行的度量。将问题发现并杜绝在软件开发的早期阶段,会大大减少由于错误或不合理而导致的花费。

(2)事前估计性:进行项目的估计和预算是软件开发较为关键的一环。所提出的度量指标能够对项目进行有效的估计,从而能够帮助开发部门合理调度资源,做好软件开发计划。

(3)评估性:所做的工作为比较开发过程提供了一定的量化依据。对于软件过程改进开始时基线的识别、改进进行到一定阶段后改进效果的衡量,提供切实可行的操作方法。

(4)实用性:研究是针对目前的主流开发过程 RUP——以用例驱动为中心,涉及的建模语言 UML 也是在软件工程界中占据着主导地位,因此对此的研究具有积极的意义。

可以预见到的是:随着用例技术的更普遍使用,基于用例的相关度量分析将更加深入而广泛。

参考文献:

[1] Rumbaugh J, Jacobson I, Booch G. The Unified Modeling Language Reference Manual [M]. [s. l.]: Addison Wesley Longman, Inc,1999.

[2] Jacobson I. Object - Oriented Software engineering - A Use Case Driven Approach [M]. Harlow, England: Addison - Wesley,1992.

[3] Schneider G. 用例分析技术[M]. 第 2 版. 北京:中信出版社,2002.

[4] 潘秋菱. 基于过程和度量的软件质量管理方法研究[D]. 合肥:合肥工业大学,2002.

[5] Schulmeyer G G, McManus J I. Handbook Software Quality Assurance[M]. 北京:机械工业出版社,2003.

[6] IEEE Standard for a Software Quality Metrics Methodology, IEEE Std[S]. 1989.

(上接第 162 页)

防御措施已难以应付。文中提出的基于历史信任数据的 DDOS 防御模型,引入了信任数据库作为辅助。在网络流量发生异常时,启动源地址检测,对进入边界路由器的异常包通过拥塞控制算法和一定的信任机制进行过滤,丢弃掉大部分具有威胁的包。但对于所丢弃的包是否可以导入缓冲区并结合源地址追踪技术,以进一步提高该防御模型防御 DDOS 攻击的能力还有待于进一步探讨。

参考文献:

[1] 黄志洪. 现代计算机信息安全技术[M]. 北京:冶金工业出

版社,2004:234 - 244.

[2] 朱良根,张玉清,雷振甲. DOS 攻击及其防范[J]. 计算机应用研究,2004(7):82 - 84.

[3] 薛立军. 分布式拒绝服务攻击检测与防护[D]. 西安:西安电子科技大学,2003.

[4] 罗光春. 入侵检测若干关键技术 with DDOS 攻击研究[D]. 西安:西安电子科技大学,2003.

[5] CipherTrust. CipherTrust's Zombie Stats[EB/OL]. 2006 - 07 - 15. <http://www.ciphertrust.com/resources/statistics/zombie.php>.

[6] 严蔚敏. 数据结构(C语言版)[M]. 北京:清华大学出版社,2002.