计算机研究与发展
Journal of Computer Research and Development
DOI:10.7544/issn1000-1239.2014.20130619
51(10):2308-2317,2014

基于多部图的云用户行为认定模型

田俊峰　　曹　迅
(河北大学网络技术研究所　河北保定　071002)
(545481066@qq.com)

A Cloud User Behavior Authentication Model Based on Multi-Partite Graphs

Tian Junfeng and Cao Xun
(*Institute of Network Technology*, *Hebei University*, *Baoding*, *Hebei* 071002)

Abstract　Cloud computing is developing rapidly, and the trustiness of cloud platform is the key issue relating to its success or failure. The authentication of the trustiness of user behavior is an important part of ensuring the credibility of cloud platform. In order to solve the problem of trustiness of cloud users' behaviors, a cloud user behavior authentication model based on multi-partite graphs (BAM) is proposed. It includes the layer of user behavior evidence, the layer of building behavior multi-partite graphs and the layer of behavior authentication. The behavior evidence is the basis, the multi-partite graphs is the method and the behavior authentication is the purpose. In the layer of user behavior evidence, the model determines the type of evidence, collects behavior evidences and analyzes user behavior quantitatively; in the layer of building behavior multi-partite graphs, the model builds two multi-partite graphs based on the layer of behavior evidence and the knowledge of graph theory; in the layer of behavior authentication, the model builds the cloud user behavior authentication module to verify that users are trusted. Identity re-certification and risk game are introduced to enhance security and accuracy of the model. The analysis of small-scale cloud user behaviors in simulation experiments show that, the model is accurate and effective in measuring the normal behavior of cloud users and in distinguishing malicious user with the risk user, and it has higher detection ratio and lower false positive ratio.

Key words　cloud computing; behavior evidence; behavior multi-partite graphs; trustiness; behavior authentication

摘　要　云计算迅猛发展,云平台的可信性是关乎其成败的关键问题,而用户行为可信性认定是保证云平台可信的重要环节.提出一种基于多部图的云用户行为认定模型,通过行为证据层、行为多部图构建层和行为认定层3个层次来解决云服务中用户行为可信性问题;同时引入身份再认证和风险博弈来增强模型的安全性与准确性.仿真实验通过对小规模云子域用户行为的分析表明,该模型可以准确地描述云用户的正常行为,对恶意用户有较高检测率,同时能有效地区分恶意用户与风险型用户,降低误报率.

关键词　云计算;行为证据;行为多部图;可信性;行为认定

中图法分类号　TP393.08

收稿日期:2013-04-25;**修回日期**:2013-12-16

基金项目:国家自然科学基金项目(60873203,61170254,61163050);河北省自然科学基金项目(F2012201145);河北省高等学校科学技术研究重点项目(ZH2012029)

云计算改变了传统的IT方式,为人们带来诸多便利,同时也面临着更为严峻的信息安全挑战[1].在云计算环境中,云平台为用户提供了开放的访问接口,云用户可以直接使用和操作云服务提供商提供的软件、平台,甚至是网络基础设施,因此云用户的不良行为对云资源安全的影响是非常严重的.云用户的身份是否真实,用户的行为是否可信是保障云资源安全的关键内容.如今身份认证技术已经比较成熟,例如生物特征识别、数字证书、动态口令等,但身份认证无法阻止合法用户的恶意行为,因此对云终端用户行为的可信性进行有效分析和认定对保障云端安全具有十分重要的意义.

1 相关工作

国内外学者在网络用户行为的可信性方面已取得了部分成果:林闯等人[2]提出了用户行为信任的评估、预测与控制架构;田立勤等人[3]借鉴社会信任的特性和计算机对信任评估的要求,提出了一种基于行为证据的双滑动窗口的行为信任量化评估机制;蒋泽等人[4]采用更完善的多维决策属性来衡量用户行为可信性;陈亚睿等人[5]建立了云服务提供商和云终端用户之间的重复博弈模型,通过不完全信息多阶段博弈来分析终端用户的类型;Almenarez等人[6]提出的PTM(pervasive trust management model based on D-S theory)模型,定义了基于普适环境的域间动态信任模型,采用改进的证据理论(D-S theory)方法进行建模,信任度的评估采用概率加权平均的方法;Brosso等人[7]提出了基于用户行为分析的连续认证系统,在环境信息中提取出用户的行为证据,将用户划分为不同的信任等级,在模糊化过程中依据相关规则确定各个参数的权重,通过神经模糊逻辑不断更新用户行为数据库,保持用户行为的准确和可靠;Elaine等人[8]提出了一种基于用户行为的隐式认证用户身份的方法,该方法通过手机等移动设备来收集用户行为信息,模拟用户行为,从而隐式认证用户.Khazzar等人[9]借鉴心理学方法研究用户行为认证系统的可行性,通过用户在3D迷宫中的反应来获取用户行为信息,认证用户身份,实验表明其准确率达到了88.33%.

通过以上的分析发现,对用户行为的研究大部分只注重了用户行为分析值的获取和相关权重的计算方法,但没有深入讨论如何有效地利用用户行为分析值来进行用户认定这一关键问题.SaaS级云用户行为的表现形式具有多样性,由于云服务多是基于Web的,所以用户的Web浏览行为能够准确地反映云用户的行为特征,不同的用户点击网页的时间间隔不同、浏览内容不同、内嵌链接的选择不同并且网络环境也不同.本文基于图论理论,将云用户行为分析值与多部图结合,提出了基于多部图的云用户行为认定模型.模型根据云服务应用程序本身特征以及时间因素在整个云服务过程中插入若干个观察点,在每个观察点处利用软硬件工具收集用户行为证据,结合层次分析法(analytic hierarchy process, AHP)分析用户行为,构造云服务行为多部图和用户行为多部图,计算最佳行为路径,获得用户行为偏离度,认定用户行为是否可信.

2 相关概念与整体逻辑框架

2.1 相关概念

定义1. 云用户行为.指用户作为主体,通过身份认证后,在使用云服务的过程中进行的一系列动作或操作.

定义2. 云用户行为证据.指在云服务过程中可以直接由软硬件检测(或者经过简单计算)获得的用来定量分析用户行为的基础数值.

定义3. 观察点(observation point, OP).指云服务过程中获取用户行为信息的关键点,即对云服务安全有较大影响的点.本文选取云服务功能选择处、比较重要的Web页面(用户信息页面、支付页面等)处作为观察点.

定义4. 云服务行为多部图(cloud behavior multi-partite graphs, CG).用来描述所有云用户使用云服务的行为,用三元组 $T_{CG}=\langle V,E,N\rangle$ 来表示,其中 V 表示节点集,任意一个节点 v_i 表示用户使用云服务的一次行为,E 是边集,连接两个节点,N 表示该图包含 N 个部分(N 个观察点).“边”表示行为的先后顺序,行为图的任意一边和两个节点可以表示为 $v_k e_i v_j$,k,i,j 都是整数,$k=j-1$.

定义5. 云用户行为多部图(user behavior multi-partite graphs, UG).用于描述一个用户在一项云服务中能够进行的全部行为.用户行为多部图是云服务行为多部图的子图.

定义6. 行为路径.指用户使用一次云服务的过程.在多部图中表示为一组节点与边的序列,其端点为“开始”、“结束”2个虚拟节点.

定义7. 云服务最佳行为路径 B_BP.用来描述

云服务中所有用户行为中最可信的那条行为路径.本模型取节点值加权和最大的那条路径.

定义8. 云用户行为可信.指用户在云服务中的行为路径与最佳行为路径基本一致(偏离度在规定的阈值之内).

定义9. 风险型用户.指用户的身份是合法的,但在某些时刻或环境下可能进行一些异常行为.

定义10. 恶意用户.指用非法手段获得了合法身份,其行为在大部分时间内是异常的.

2.2 整体逻辑框架

传统网络用户由于物理资源的限制性,只能在较为固定的范围内使用网络;而云用户由于脱离了物理硬件资源的限制,可以在更自由的环境中使用云服务,因此云用户的行为有更强的动态性,行为证据也更加的复杂多样.在这样的云环境下,传统的用户行为认证方法无法有效地区分恶意用户和风险型用户.本模型基于行为多部图,细粒度地分析了云用户在整个云服务过程中的行为,并且引入了身份再认证技术和风险博弈技术,使云服务过程中恶意用户的识别更加准确与高效.

云用户整体的规模是巨大的,但是根据云服务内容、用户的自身属性等可以将云用户划分到不同的域当中,在同一个域中,又可以依据用户的偏好或云服务商的相关规定将域中用户划分到不同的子域中,这样云用户的规模就会变得相对较小,本文提出

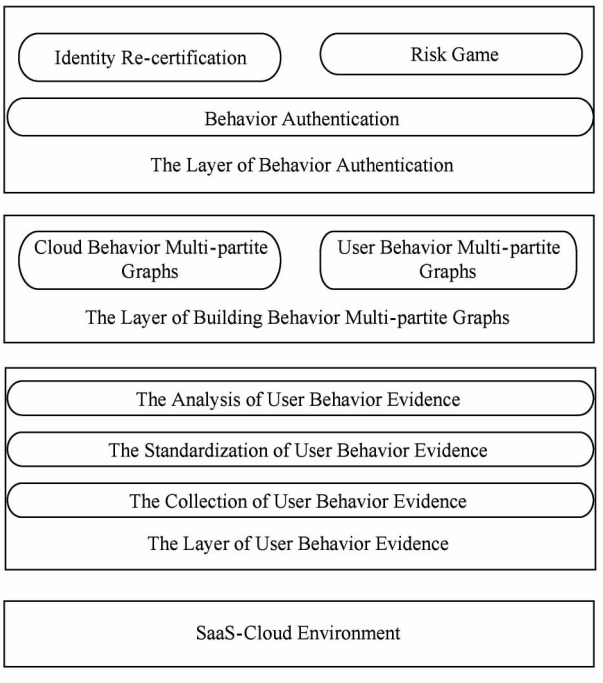

Fig. 1 Logical framework of the model.

图1 模型整体逻辑框架图

的基于多部图的云用户行为分析方法适用于这种小规模的云用户情况.基于多部图的云用户行为认定模型采用层次结构,分为3层:行为证据层、行为多部图构建层和云用户行为认定层.其中行为证据是基础,多部图是方法,用户行为可信认定是目的,模型整体逻辑框架如图1所示.

3 云用户行为证据

3.1 行为证据的获得

用户行为以具体的用户行为证据来表现,获得全面的、粒度适当的、可信的证据是云用户行为可信性分析的基础.本文综合分析其他学者提出的行为证据收集方法[10-11],采用了以下3种获取行为证据的方法:1)网络流量监测工具,如Bandwidthd,它可以在IP地址的基础上获取HTTP, TCP, UDP, ICMP, VPN和P2P的数据流;2)现有的入侵检测系统,如Tcpdump,它可以获得访问次数、操作失败次数、网络传输延时等证据;3)Web日志文件,从高层协议看,用户的浏览过程是通过一系列HTTP请求/响应构成的.由于一个页面通常包含多个内嵌链接,例如图片、广告条、背景音乐和框架页面等,因此,用户的每一次浏览行为(例如点击页面链接、前进、后退、刷新等)都会触发浏览器发出一系列的HTTP请求.这些HTTP请求到达服务器后,其属性(源地址、请求时间、请求对象等)会被记录在服务器的日志文件中.因此,通过Log文件可以分析出用户的浏览行为.

3.2 行为证据规范化

通过软硬件工具获得的行为证据的表现形式具有多样性,主要有:1)越大越优型,如数据传输速度、无障碍服务次数等;2)越小越优型,如云服务响应时间、越权访问次数等.为了方便云服务多部图中最佳路径的计算,需要对证据进行规范化.在此引入优属度的概念,即证据对于模糊概念"优"的隶属程度.证据优属度可由查德公式[12]导出.

1) 对于越大越优型证据:$g=\dfrac{e-\inf(e)}{\sup(e)-\inf(e)}$,

2) 对于越小越优型证据:$g=\dfrac{\sup(e)-e}{\sup(e)-\inf(e)}$,

其中,g为证据的优属度,即行为证据规范化后的值;e为直接获得或经过简单计算得到的证据数值;$\sup(e)$,$\inf(e)$分别为证据值的上界、下界.证据的优属度在$[0,1]$之间取值,其值越大越优.

3.3 基于 AHP 的云用户行为分析

为了科学地认定云用户行为的可信性,还必须对云用户行为进行定量分析.为此,本文将经典的 AHP 方法[13]运用到云用户行为分析中,将云用户行为作为最高层,行为证据作为最底层,通过对行为证据的层层分析,确定各行为证据对云行为的权重,最后,通过行为证据权重和行为证据值求出云用户的行为分析值.具体过程如下:

1) 建立递阶云用户行为层次结构

云用户行为层次结构包含 3 个层次,由上到下依次为:云用户行为层(目标层)、行为属性层(中间层)和行为证据层(措施层).

2) 构造判断矩阵并赋值

根据递阶层次结构构造判断矩阵.矩阵中的元素两两比较,依据重要性标度含义表[13]对重要性程度按 1~9 赋值.

3) 计算权向量并作一致性检验

计算判断矩阵的最大特征根及其所对应的特征向量,并进行一致性检验.若检验通过,特征向量(归一化后)即为权向量;若没通过,重新构造判断矩阵.具体的权向量计算如下:

$$\boldsymbol{W}_i=\frac{1}{n}\sum_{j=1}^{n}\frac{a_{ij}}{\sum_{k=1}^{n}a_{kl}},$$

其中,n 代表判断矩阵中元素总数,$1\leqslant i\leqslant n$,$1\leqslant l\leqslant n$.

一致性检验的步骤如下:

步骤 1. 计算一致性指标(consistency index, CI),$CI=\frac{\lambda_{\max}-n}{n-1}$($\lambda_{\max}$ 表示判断矩阵的最大特征根).

步骤 2. 查 RI 值表[12]确定相应的平均随机一致性指标(random index, RI).

步骤 3. 计算一致性比例(consistency ratio, CR)并进行判断 $CR=\frac{CI}{RI}$.

当 $CR<0.1$ 时,认为判断矩阵的一致性是可以接受的;$CR\geqslant 0.1$ 时,认为判断矩阵不符合一致性要求,需要对该判断矩阵进行重新修正.

4) 计算组合权向量并作组合一致性检验

计算每一个判断矩阵各元素针对目标层(最上层)的相对权重.这一权重的计算采用从上而下的方法逐层合成.同样,也需要对组合权向量进行一致性检验.

5) 计算云用户行为可信度

设用户行为的属性向量为 $\boldsymbol{A}=(a_1\cdots a_i\cdots a_n)^{\mathrm{T}}$,属性的权重向量为 $\boldsymbol{W}_A=(w_1\cdots w_i\cdots w_n)^{\mathrm{T}}$,则 $V_{\mathrm{cub}}=\boldsymbol{A}^{\mathrm{T}}\boldsymbol{W}_{\boldsymbol{A}}=(a_1\cdots a_i\cdots a_n)(w_1\cdots w_i\cdots w_n)^{\mathrm{T}}=\sum_{i=1}^{n}a_iw_i$ 称为用户行为分析值(value of user behavior analysis, VB),它表示用户行为对于"可信"这一模糊概念的隶属度,因此,也称作云用户行为可信度.

4 多部图构建与行为可信性认定

4.1 多部图构建

本模型所涉及的多部图包括云服务行为多部图和云用户行为多部图.云服务行为多部图是在不同云用户多次使用云服务行为的基础上构建的.具体的构建过程如下:

1) 将云服务的全部过程划分为 N 个部分,建立 N 个观察点(不包括开始节点和结束节点),在每个观察点处收集用户行为证据,每个观察点所收集的行为证据类型是不同的,因为其侧重点不同,有的偏重效率、有的偏重安全等,此外观察点也有权重,权重的确定方法依然采用第 3 节介绍的层次分析法的思想,各个观察点间相互对比,最终确定各自相对于云用户行为的权重.

2) 任取第 i 个观察点,它包含 S_i 个节点,即 S_i 种行为(S_i 小于等于用户使用云服务的次数),不同的观察点处 S_i 的取值是不同的.

3) 将第 i 个观察点处的每一个节点(代表 1 种行为)与第 $i+1$ 个观察点处的每一个节点相连接,得到一个完全多部图,这就是云服务行为多部图(如图 2 所示,OP 代表观察点,VB 代表用户行为分析值,S_i 代表用户在第 i 个观察点处的行为总数).通过这种方法构建的行为多部图是一个完全多部图,有些边在实际实验中不一定存在,即云服务行为多部图仅在理论上存在,实际中只是无限接近它.

云服务行为多部图中由开始节点到结束节点的任意一条路径,表示用户使用一次云服务的过程.由于节点代表着用户行为分析值,并且此值越大越好,所以将各观察点中值最大的节点连接起来就是云服务的最佳行为路径.云服务的最佳行为路径不一定实际存在,它仅是用户行为认定的一个标准,当用户使用云服务时相对于最佳行为路径的偏离度将是确定用户行为是否可信的重要依据.

云用户行为多部图是在一个用户多次使用某项云服务的基础上构建的,其构建过程与云服务行为多部图的构建过程类似,将整个云服务过程划分为 N 个部分(即建立 N 个观察点),观察点的数目和位

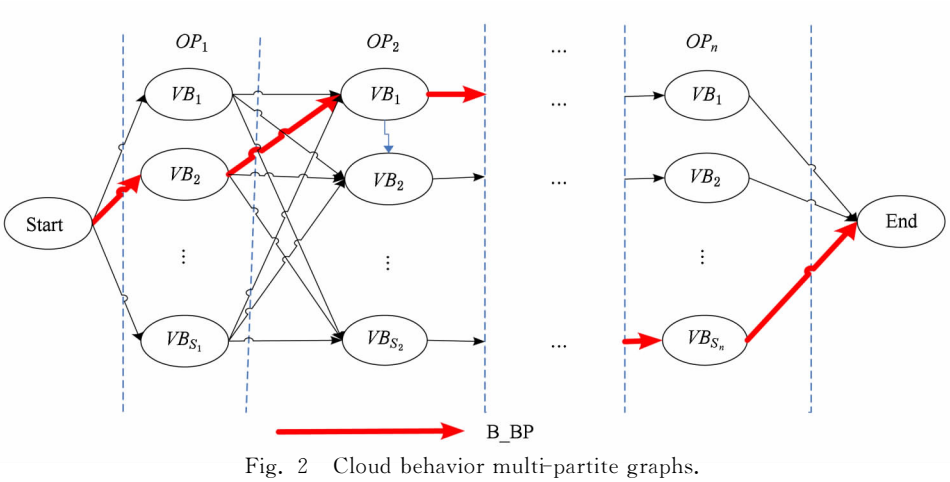

Fig. 2 Cloud behavior multi-partite graphs.

图 2 云服务行为多部图

置与云服务行为多部图是一致的,在每个观察点处只收集被观察用户的行为证据,最后将相邻观察点处的节点连接,构成云用户行为多部图. 云用户行为多部图的节点数和边数小于等于云服务行为多部图的节点数和边数.

云用户最佳行为路径的确立过程同样是将用户行为多部图各部节点中行为分析值最大的节点相连接,它同样是理论上存在(实验中也可能恰好存在). 通过比较云用户最佳行为路径和云服务最佳行为路径,得到用户行为偏离度,从而将用户分为保守型和激进型,同时这也是用户行为可信阈值判定的依据.

在实际运行维护过程中,随着用户与云服务间交互次数的增加,多部图的规模会不断的增大,存储资源和计算资源都是有限的,所以必须控制多部图的规模,综合考虑时间因素和节点频率定期去除多部图中的失效节点,实现多部图的动态性维护,这与实际用户行为的阶段性(不同的时间段内,行为的特征是不一样的)是相符合的.

4.2 云用户行为可信性认定

云服务行为多部图和云用户行为多部图是云用户行为认定的基础. 在此基础上构建了云用户行为认定模块,工作流程如图 3 所示:

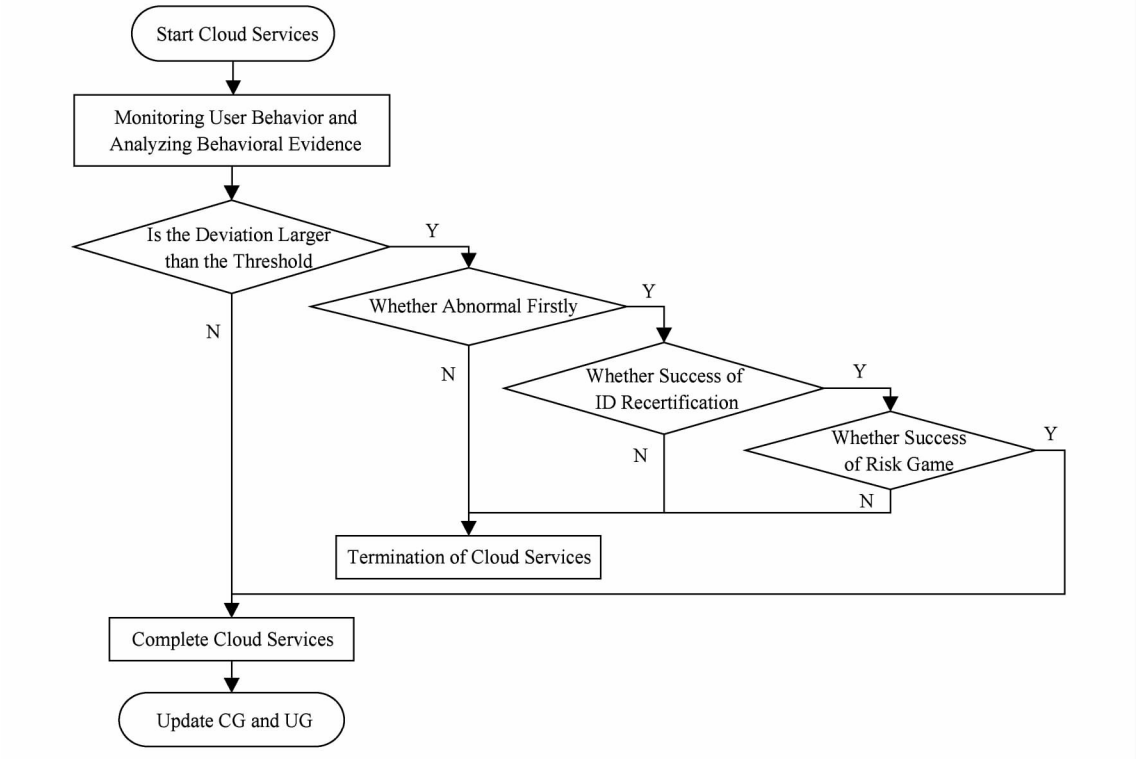

Fig. 3 The cloud user behavior authentication flowchart.

图 3 云用户行为认定流程图

云服务进行中的用户行为认定具体步骤如下:

步骤 1. 根据云服务的特征,在云服务过程中选取 N 个观察点,利用第 3 节介绍的云用户行为分析法的前 4 步计算出每个观察点的权重 $\omega_i(1\leqslant i\leqslant N)$.

步骤 2. 在每个观察点处收集云用户行为证据,计算证据优属度,依据云用户行为层次分析法计算云用户行为分析值 $V_i(1\leqslant i\leqslant N)$.

步骤 3. 在每个观察点处计算用户行为分析值与标准行为分析值(最佳行为路径中各观察点处节点代表的值)差的绝对值,差值乘以观察点的权重,所得的值作为该观察点处的用户行为偏离度 $D_i=|V_i-V_i|\times\omega_i(V_i'$ 为标准行为分析值,$1\leqslant i\leqslant N)$.

步骤 4. 累加已认定过的观察点处的用户行为偏离度,即 $D=\sum_{i=1}^{k}D_i$(k 为已认定过的观察点总数)当其值大于用户行为可信阈值 α 时,对用户进行身份再认证,如果失败则终止服务,如果成功则进行风险博弈[5],如果博弈失败则终止服务;如果博弈成功则继续服务,将偏离度归零,但是降低其行为可信阈值,并对其进行标记,若再次超出阈值则立即终止服务.

步骤 5. 成功完成云服务,更新云服务行为多部图和用户行为多部图.

用户行为认定的前提是已经通过大量实验建立了云服务行为多部图. 整个认定分为 2 个阶段:初始阶段和稳定阶段. 在初始阶段,由于用户初次使用云服务,用户行为信息不全面,用户行为多部图没有建立,用户类型也尚未明确,所以选用云服务行为多部图中的最佳行为路径作为认定用户行为的标准,该路径上的节点值就是标准行为分析值. 在稳定阶段,用户已经与云服务进行了多次交互,形成了自己的用户行为多部图,确立了自己的最佳行为路径,此时用户行为多部图中的最佳行为路径作为用户行为认定的标准.

用户行为可信阈值的初始值是云服务行为多部图中的最大偏离度,即 $\alpha_0=\max(D_1,D_2,D_3,\cdots,D_n)$,其中 n 代表不同用户使用云服务的总次数(或者由云服务提供商、网络安全专家和用户协商确定 α_0). 在稳定阶段,用户行为可信阈值 α_1 通过以下公式获得:$\alpha_1=\alpha_0+\sum_{i=1}^{n}|VU_i-VC_i|\times\omega_i$,其中 α_0 为可信阈值初始值;n 为观察点总数;VU_i 为用户行为多部图的最佳行为路径中第 i 个观察点处的用户行为分析值;VC_i 为云服务行为多部图的最佳行为路径中第 i 个观察点处的用户行为分析值;ω_i 为第 i 个观察点的权重. 当进行过风险博弈后行为可信阈值变为 $\alpha_2=\alpha_0-\sum_{i=1}^{n}|VU_i-VC_i|\times\omega_i$.

5 模型仿真及结果分析

5.1 仿真数据

实验数据来源于微软提供的非商务 Web 服务器 MyNonCS2KSrv 的日志文件[14]. 首先从数据集中随机、不重叠地选出两组用户,分别记为 DS1 和 DS2,其中 DS1 用于构造云服务行为多部图;DS2 用于模型测试. 由于该数据集中不包括恶意用户和风险用户,为了验证模型对 2 种用户的识别能力,我们根据相关资料[15],模拟了这 2 种类型用户的行为,具体方法如下:恶意用户模拟 DDoS 攻击,发送 GET 请求的时间间隔设为均值 20 ms 的随机数(HTTP 请求的发送速率约为每秒 50 个),攻击时间长度为云服务运行时间,GET 请求的内容有 2 种方法生成:1)截取一段正常用户的 HTTP 请求序列片段;2)随机生成每个 GET 请求的内容. 风险型用户在大多数情况下的行为是正常的,但在某些时刻可能会出现异常行为,实验通过在某时刻大量发送 HTTP 请求的方式模拟风险用户的行为,与恶意用户类似,但时间很短.

5.2 构造云服务行为多部图

在该服务中标记 5 个观察点,根据 DS1 中的数据分析每个观察点处的用户行为,获得用户的各项行为证据,行为证据主要包括:1)环境属性的行为证据,包含网络吞吐量、IP 包传输延时和 IP 包丢包率;2)操作属性的行为证据,包含点击次数、访问页面数、敏感页面访问次数和页面停留时间. 环境属性主要用于判断用户所处网络环境是否安全、正常,其中吞吐量表示在单位时间内通过某个网络(或信道、接口)的数据量,反映了当前网络负载情况;传输延时表示 IP 包在传输介质中传输所用的时间;丢包率是数据包丢失部分与所传数据包总数的比值,丢包的原因主要有物理线路故障、设备故障、病毒攻击、路由信息错误等. 操作属性主要用于判断用户行为是否符合其行为习惯,其中点击次数表示用户访问某特定资源(如页面上的图片、链接等)的次数;访问页面数反映用户在观察点规定时间内浏览页面的总数;敏感页面访问次数是用户在观察点处访问敏感信息(用户个人信息、支付信息等)的次数;页面停留

时间反映用户浏览某页面所用时间. 根据第 3 节介绍的分析方法计算观察点处的云用户行为分析值. 用户行为递阶层次结构如图 4 所示,各个行为证据的权重如图 5 所示.

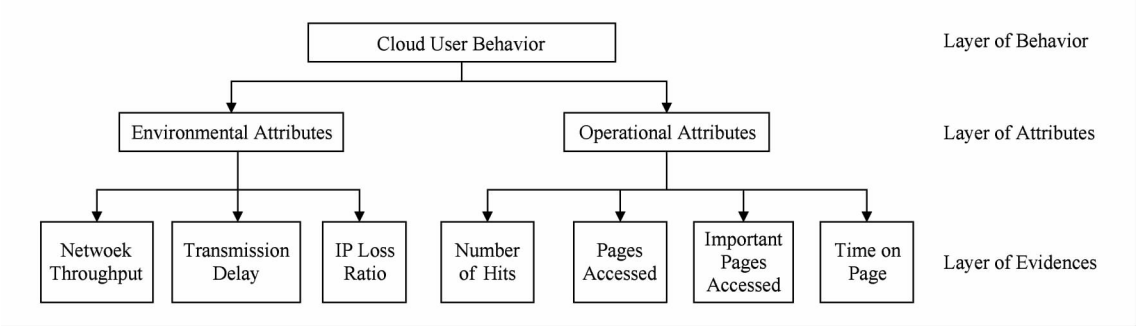

Fig. 4 User behavior hierarchical structure.

图 4 用户行为递阶层次结构图

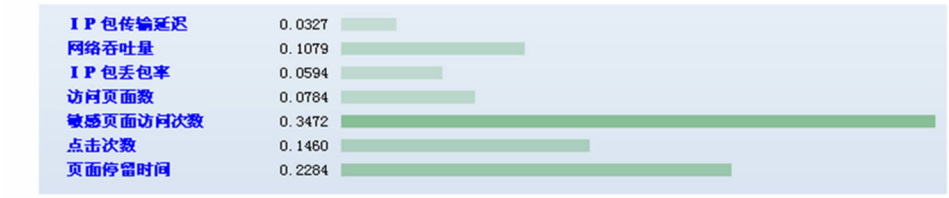

Fig. 5 The weights of cloud user behavior evidences.

图 5 云用户行为证据权重

通过对数据集 DS1 的分析,约 95%的用户行为集中在稳定范围内,观察点 1～5 的用户行为分析值的稳定范围依次是: 0.56～0.63, 0.69～0.80, 0.44～0.52, 0.49～0.59, 0.48～0.55. 得到正常用户的行为多部图,如图 6 所示,由于用户行为分析值分布密集,采用散点图表示,图 6 中的一条路径表示用户使用一次云服务,由于用户行为分析值越大表示用户行为越可信,因此连接各观察点处的最大值,构成最佳行为路径,如图 6 中的粗虚线所示. 图 6 中各种点表示云用户行为分析值,路径表示用户与云服务的交互过程,最上面的粗虚线代表云服务最佳行为路径.

为了计算用户行为可信阈值,首先计算各个观察点的权重,结果如图 7 所示.

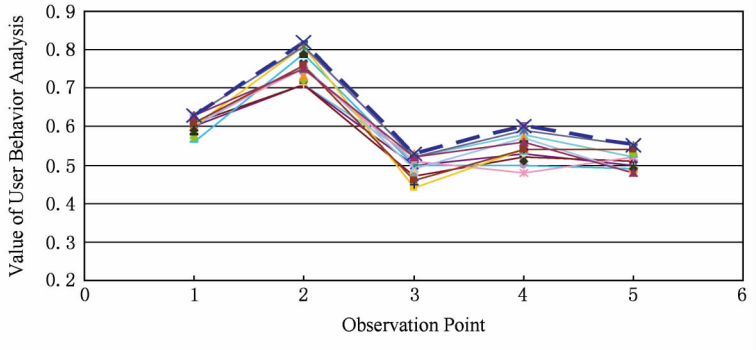

Fig. 6 Cloud behavior multi-partite graphs.

图 6 云服务行为多部图

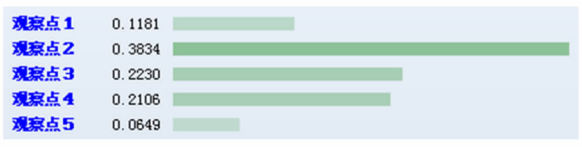

Fig. 7 Weight of each observation point.

图 7 观察点权重

统计各观察点处的最大偏离度分别为:0.07,0.11,0.08,0.10,0.07,加权求和,最终得到云用户行为可信阈值为 0.09.

5.3 模型验证与分析

本实验选用数据集 DS2 测试云用户行为认定模型,并将本模型与其他 2 个模型进行了对比. 在 DS2 中共有用户 100 个,其中恶意用户 2 个,风险用户 1 个,其余为正常用户. MU1(第 1 个恶意用户)采用第 1 种 DDoS 攻击手段,在云服务时间内持续发送大量的 Index.asp 页面请求,致使吞吐量、访问页面数、点击次数等行为证据表现异常;MU2 采用第 2 种攻击手段,在云服务时间内持续发送大量 HTTP 请求,请求内容为随机页面,致使吞吐量、访问页面数、点击次数、敏感页访问次数等行为证据表现异常. 风险用户(risk user, RU)的表现为在某个随机时刻发送大量随机内容的 HTTP 请求. 通过统计获得 DS2 中各用户行为分析值,如图 8 所示:

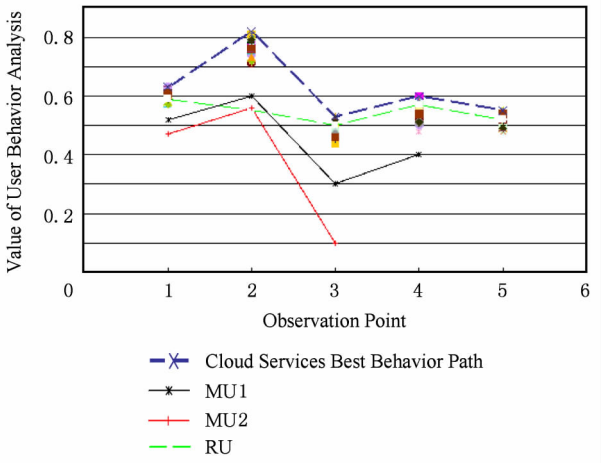

Fig. 8 The values of user behavior analysis in DS2.

图 8 DS2 云用户行为分析值

本实验中由于用户类型一致,所以云用户行为多部图与云服务行为多部图基本一致,认为云用户最佳行为路径与云服务最佳路径的偏离度为 0,即可信阈值为 0.09. 本实验假设异常用户都通过了身份再认证和风险博弈. 通过图 8 可以清晰地发现 3 条异常的路径,其中 1 条只在第 2 观察点处出现较大偏离,其余 4 个观察点处行为表现正常,能够完成服务,说明此用户为风险用户;其余 2 条路径在各观察点处均与最佳路径偏离较大,在第 2 观察点处累积偏离量超过可信阈值后,分别在第 3、第 4 观察点处再次超出可信阈值,服务终止,说明两者为恶意用户. 由此可见本文提出的云用户行为认定模型能够准确地检测出异常用户,并且能够正确地区分恶意用户与风险用户.

在异常检测中,检测率(detection ratio, DR)与误报率(false positive ratio, FPR)是 2 个重要的性能衡量指标. 本实验中检测率主要体现在恶意用户的检出量,而误报率主要体现在对风险用户的误报. 为此通过改变恶意用户所占的比例来测试模型的检测率和误报率,并且与经典的 PTM(pervasive trust management model based on D-S theory)模型[6]以及基于多维决策属性的用户行为评估模型(MDA)[4]进行了比较,结果如图 9、图 10 所示:

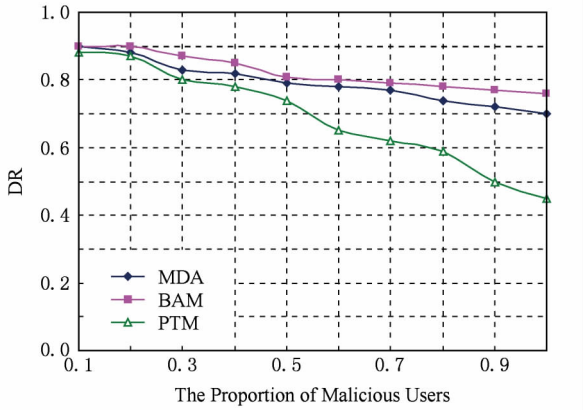

Fig. 9 The DR of the three models.

图 9 恶意用户检测率

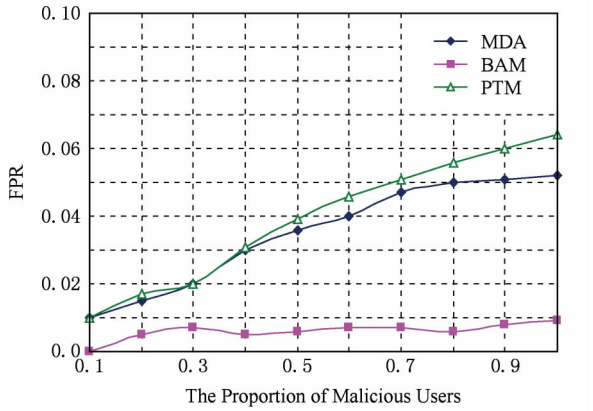

Fig. 10 The FPR of the three models.

图 10 风险用户误报率

由图 9 可知,在恶意用户所占比例较小时,3 个模型都具有较高的检测率,随着恶意用户所占比例的增加,PTM 模型的恶意行为检测率明显下降,波动较大,MDA 模型出现小幅度下降,本文模型 BAM 依然保持较高的检测率. 在图 10 中,在恶意用户所占比例较小时 3 种模型误报率都比较低,随着恶意用户所占比例的增加,PTM 模型和 MDA 模型的误报率明显上升,本文模型 BAM 依然维持在较

低的范围.本模型能维持较高检测率和较低误报率的原因有2点:1)多部图中不含恶意用户的行为信息,用户行为可信性的认定过程不受恶意用户干扰,模型对不同环境具有较好的适应性;2)模型将云服务过程划分为若干部分进行分布式认证,最后再进行综合认定,保持较高的稳定性.此外,2图中曲线的波动与行为证据或观察点的选择及其权重有关,当权重较大的行为证据或观察点(如本实验中的敏感页面访问次数、第2观察点处)出现异常情况,检测率和误报率会出现较大偏差.因此,我们需要合理准确地选择行为证据和观察点,更加科学地确定它们的权重.

6 总 结

本文应用AHP方法对云用户行为进行量化分析,并在云用户行为分析值的基础上结合图论中多部图的相关知识,提出了一种基于多部图的云用户行为认定模型,最后通过实际的网络数据及仿真实验验证了模型具有如下优点:1)能够正确地、细粒度地描述云环境下的用户行为;2)通过行为多部图中的最佳行为路径可以直观、高效地检测出恶意用户;3)有效地区分了恶意用户和风险型用户,具有较低的误报率.

同时研究中发现:1)行为证据是用户行为认定的基础,因此行为证据的选择及其权重的确定显得尤为重要,AHP是对定性问题进行定量分析的一种简便、灵活而又实用的多准则决策方法,但是该方法具有较强的主观性,对AHP的改进和证据类型的选择将是下一步工作的主要内容之一;2)行为之间是具有关联性的,这种关联性通过多部图中的"边"来体现,对行为多部图中"边"的深入挖掘将会进一步提高模型的准确性.

参 考 文 献

[1] Foster I, Zhao Y, Raicu I, et al. Cloud computing and grid computing 360-degree compared [C] //Proc of Grid Computing Environments Workshop (GCE'08). Piscataway, NJ: IEEE, 2008: 1-10

[2] Lin Chuang, Tian Liqin, Wang Yuanzhuo. Research on user behavior trust in trustworthy network [J]. Journal of Computer Research and Development, 2008, 45(12): 2033-2043 (in Chinese)
(林闯, 田立勤, 王元卓. 可信网络中用户行为可信的研究[J]. 计算机研究与发展, 2008, 45(12): 2033-2043)

[3] Tian Liqin, Lin Chuang. Evaluation mechanism for user behavior trust based on DSW [J]. Journal of Tsinghua University: Science and Technology, 2010, 50(5): 763-767 (in Chinese)
(田立勤, 林闯. 基于双滑动窗口的用户行为信任评估机制[J]. 清华大学学报: 自然科学版, 2010, 50(5): 763-767)

[4] Jang Ze, Li Shuangqing, Yin Chengguo. Evaluating network user behavior trust based on multiple decisions attributes [J]. Application Research of Computers, 2011, 28(6): 2289-2293 (in Chinese)
(蒋泽, 李双庆, 尹程果. 基于多维决策属性的网络用户行为可信度评估[J]. 计算机应用研究, 2011, 28(6): 2289-2293)

[5] Chen Yarui, Tian Liqin, Yang Yang. Model and analysis of user behavior based on dynamic game theory in cloud computing [J]. Acta Electronica Sinica, 2011, 39(8): 1818-1823 (in Chinese)
(陈亚睿, 田立勤, 杨扬. 云计算环境下基于动态博弈论的用户行为模型与分析[J]. 电子学报, 2011, 39(8): 1818-1823)

[6] Almenarez F, Marin A, Campo C, et al. PTM: A pervasive trust management model for dynamic open environments [C] //Proc of the 1st Workshop on Pervasive Security, Privacy and Trust. Los Alamitos, CA: IEEE Computer Society, 2004: 1-8

[7] Brosso I, Neve A, Bressan G, et al. A continuous authentication system based on user behavior analysis [C] // Proc of 2010 Int Conf on Availability, Reliability and Security. Piscataway, NJ: IEEE, 2010: 380-385

[8] Elaine S, Niu Yuan, Jakobsson M, et al. Implicit authentication through learning user behavior [G] //LNCS 6531: Proc of ISC 2010. Berlin: Springer, 2011: 99-113

[9] Khazzar A, Savage N. Graphical authentication based on user behaviour [C] //Proc of the 2010 Int Conf on Security and Cryptography. Piscataway, NJ: IEEE, 2010: 86-89

[10] Tian Liqin, Lin Chuang, Ji Tieguo. Quantitative analysis of trust evidence in internet [C] //Proc of the 10th Int Conf on Communication Technology. Piscataway, NJ: IEEE, 2006: 194-198

[11] Tian Liqin, Lin Chuang, Ni Yang. Evaluation of user behavior trust in cloud computing [C] //Proc of 2010 Int Conf on Computer Application and System Modeling (ICCASM 2010). Piscataway, NJ: IEEE, 2010: 567-572

[12] Wang Peizhuang. Fuzzy Set Theory and Its Applications [M]. Shanghai: Shanghai Scientific and Technical Publishers, 1983 (in Chinese)

(汪培庄. 模糊集合论及其应用[M]. 上海: 上海科学技术出版社, 1983)

[13] Baidu Encyclopedia. AHP (analytic hierarchy process) [EB/OL]. (2006-07-21)[2013-02-25]. http://baike.baidu.com/view/364279.htm

[14] Microsoft. How to analyze non-commerce Web server log files [EB/OL]. (2007-11-16)[2013-03-04]. http://support.microsoft.com/kb/293887/zh-cn

[15] Xie Yi, Yu Shunzheng. Anomaly detection based on Web users' browsing behaviors [J]. Journal of Software, 2007, 18(4): 967-977 (in Chinese)
(谢逸, 余顺争. 基于 Web 用户浏览行为的统计异常检测[J]. 软件学报, 2007, 18(4): 967-977)

Tian Junfeng, born in 1965. PhD, professor, and PhD supervisor. His current research interests include distributed computing, network technology, trusted computing, cloud computing, etc (tjf@hbu.edu.cn).

Cao Xun, born in 1987. Master. His current research interests include cloud computing, trusted computing, cloud user behavior, etc.