基于 Hadoop 架构云平台的动态行为信任评估方法

章玉英
（电信工程系 四川管理职业学院，四川 成都 611732）

摘要：针对 Hadoop 架构云平台不能验证“合法”用户的行为是否可靠的问题，文章提出了基于 Bayes 均值信任模型和动态行为信任链，用于计算每个客户端的信任值，使得信任值满足阈值的用户获得相应服务资源。该方案能够防止合法客户端受到“被动”的拒绝服务攻击，提高了客户端获取服务资源的公平性，同时有效防止恶意节点的协同欺骗。

关键词：云平台；Hadoop；信任链路
中图分类号：TP393 **文献标识码：**A **文章编号：**1673-1131（2018）02-0100-03

Dynamic behavior trust evaluation method based on Hadoop architecture cloud platform

Zhang Yuying
(Department of Telecom Engineering, Sichuan Vocational Institute of Management Chengdu 611732)

Abstract:According to the problem of the cloud platform for Hadoop architecture not verifying the reliability of the behavior of its "legitimate" user, this paper proposes the Based on Bayes mean trust model and the trust link of dynamic behavior to compute the trust value of each client. The client with satisfying the threshold value will gain the corresponding the service resources. The scheme prevents the legitimate clients from being "passive" denial-of-service attacks and improves the fairness of the client access to service resources. It is also effective to prevent the cooperative deception of malicious nodes.

Key Words:Cloud platform; Hadoop; Trust link

0 引言

Hadoop 架构云平台是一款开源的分布式计算与存储平台[1]。由于 Hadoop 构建云平台特有的可扩展性、部署灵活性等特征，越来越多使用者对其关注有加。Hadoop 构建云平台作为最流行的开源云平台之一，该平台的安全性影响其进一步的发展。近年来，Hadoop 生态体系安全是众多学者关注的重要方面之一。陈玺等[2]分析了当前 Hadoop 架构云平台的安全性，然后从可信平台、加密演算法、混合加密算法、三重数据加密算法等进行的方法优劣对比，同时进一步分析了 Hadoop 架构云平台的安全策略的细粒度、模块化、可扩展性等研究情况。典型的 Hadoop 架构云平台的安全性是基于 Kerberos 认证协议的安全策略[3]，虽然该方法有效解决了接入云平台用户身份的合法性，但该策略无法防止“合法”客户端通过不正常行为获取更多的利益，即客户端引起的“被动”拒绝服务攻击。因此针对客户端行为的信任分析[4]是本文研究的重点。

针对 Hadoop 架构云平台不能验证“合法”用户的行为是否可靠的问题，本文提出了基于 Bayes 均值信任模型，该模型

3.3 构建新型社交网络框架

构建新型的社交网络框架，拓宽大数据的缓存空间，以数字公共服务平台（CCN）为节点或中继节点，进行移动社交用户分组，进行用户兴趣包升级，是时代发展的必然，也是移动社交网络可持续发展的要求。近期麦肯锡报告显示，谷歌、百度等技术巨头花费开源框架开发方面的资金，每年都多达 3000 万美元。因为在技术迭代迅速的年代，封闭而陈旧的基础框架根本没有任何商业价值，建立开源框架，吸引、融入人工智能社区，才是每个行业和商业主体该考虑的问题。因此，我们要重视和加强新型社交网络框架构建，研发社交框架互转项目，增强公共社交框架模型的转换模式，以了解用户基于多源信息融合的隐式表达方式的内涵，并通过用户的隐式表达了解他们的喜好。构建新型社交网络框架的重点是先建立模型，进行移动社交网络架构，因为不论是用户想要发送还是获取移动数据，它都要在移动社交网络条件下进行。然后，将移动数据缓存在数字领域公共服务平台上，使用兴趣数据包分析、传递移动数据的请求信息，使用代理节点和中继节点传送相关数据，实现移动社交用户之间的频繁连接，同时使上层开源接口库向上拓展，促使移动社交社区摆脱框架限制。

3.4 保护用户隐私和安全

各种终端智能设备和移动网络的出现，为用户进行开放、自由社交提供了便利。数据显示，2016 年，全球移动社交用户数量约为 27 亿。借助移动社交网络平台，用户可以随时进行信息分享和传递，不过，移动服务器的记录功能，也给社交网络用户的隐私和安全带来了挑战。如在移动社交网络中，许多用户为了表现自我、维护朋友圈的关系，经常会有意无意地将个人信息暴露出来，而移动服务器会将用户的个人信息、位置信息、人际交往关系和轨迹信息等都记录下来。借助新兴技术和防御技术，对公共社交领域常见的诈骗窃密等问题进行跟踪、分析和监督，同时提示用户培养良好的网络习惯和社交行为，尤其是在网络社交中，更要谨慎对待有关人身财产安全的问题，以确保移动社交平台和网络生态安全，维护自身的合法权益。

参考文献：

[1] 于喆. 移动社交网络中大数据聚类算法的研究与应用[D]. 南京邮电大学,2017.
[2] 师耀,王咏霖. 大数据在移动社交网络的应用研究[J]. 通讯世界,2017,(13):58-59.
[3] 齐晓娜,张宇敬,封二英. 移动社交网络用户隐私保护问题研究[J]. 产业与科技论坛,2017,16(16):35-36.

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

能够防止 Hadoop 架构云平台其他合法客户端受到“被动”的拒绝服务攻击，提高了客户端获取服务资源的公平性，同时构建动态行为信任链路能够更为准确地分析客户端节点在接入 Hadoop 架构云平台后的服务行为，可以有效防止恶意节点的协同欺骗。

1 构建 Hadoop 安全体系

一般来说，典型 Hadoop 架构云平台运行是假设运行在信任环境下，再利用现有的 Kerberos 认证协议，实现 Hadoop 架构云平台集群对客户端的安全认证和票据授权服务。对于具有合法身份的客户端用户是可以正常登录云平台服务器，获取数据存储和计算服务，然后这些合法用户可以在“不受监控”状态下恶意获得更多服务资源，从而导致 Hadoop 架构云平台其他合法客户端受到“被动”的拒绝服务攻击，影响了客户端获取服务资源的公平性。

典型的 Hadoop 架构的云平台包括 Hadoop 集群核心节点和辅助终端接口，其中集群核心节点包括：NameNode、备份 NameNode、DataNode 以及 JobTracker 等成员，辅助终端接口涉及网关、交换机、资源控制器等硬件设备，因此基于 Hadoop 架构云平台的体系架构分为三层 数据存储层、中间件控制层和应用节点层[1]，以加密、认证[5]、身份授权[6]、安全存储[7]实现 Hadoop 架构云平台的安全性，如图 1 所示。

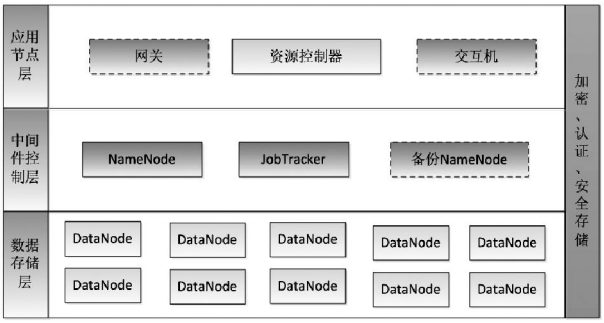

图 1 基于 Hadoop 架构的云平台体系

其中，DataNode 是 Hadoop 架构云平台的存储资源和计算资源，其主要作用为存储分布系统数据以及执行 MapReduce 的任务，Hadoop 架构云平台数据信息都存储在 DataNode 终端设备上。当该云平台接收到 MapReduce 作业任务时，NameNode 节点将根据当前需要完成的任务情况，通过资源调度策略为当前任务分配相应的 DataNode 节点的存储资源或计算资源。

NameNode 节点在 Hadoop 架构云平台中起中心控制作用，云平台中所有数据文件以索引方式存储到不同的 DataNode 节点上，然后将索引信息存储到 NameNode 节点，同时将索引信息与“数据片”相关联。即使是一个占用存储空间很小的文件存储到云平台上，它也可能按照 NameNode 对照索引分片存储到多个 DataNode 节点上。Hadoop 架构云平台中的 NameNode 起着资源管理的作用。为此 Hadoop 架构云平台设计了备份 NameNode 节点，实现对关键节点备份功能。

对于以分布式文件系统的 Hadoop 架构云平台，JobTracker 起着与 NameNode 节点相当的中心控制作用。它负责向运行 TaskTracker 每个 DataNode 节点上的服务提交用户任务，因此该节点同样起着资源调配的作用。

资源控制器用于监控网关和路由器的运行状态，实时监控每个硬件关键参数，并为硬件设备提供自适应的参数修改功能。网关和路由器是 Hadoop 架构云平台直接进行数据通信的中转站，搭起多个 Hadoop 架构云平台共享数据的通信链路。

2 基于 Bayes 均值信任模型

在 Hadoop 架构云环境下，当恶意客户端接入 Hadoop 架构云平台，它可能通过“非正常竞争”方式获取存储资源和计算资源，严重影响整个 Hadoop 架构云平台的服务资源调度策略。由于 Hadoop 架构云平台下客户端访问云端的信任关系具有动态性和不确定性，那么贝叶斯推演机制较好模拟客户端登录云端的服务信任行为评价，即多个云端对每个客户端登录 Hadoop 架构云平台的服务行为的可靠性评判。

当客户端 x 与云端 y 发生 n 次服务交互后，在客户端获取服务的过程中，它表现为正常行为的次数为 u，非正常行为次数为 v，则客户端 x 和云端 y 直接交互成功的后验概率服从 Beta 分布，其密度函数[3]为：

$$Beta(\theta|u,v)=\frac{\Gamma(u+v+2)}{\Gamma(u+1)\Gamma(v+1)}\theta^{u}(1-\theta)^{v} \tag{1}$$

直接信任度[3]为：

$$\theta_{xy}=E[Beta(\theta|u+1,v+1)]=\frac{u+1}{n+2} \tag{2}$$

每个客户端的信任度由 Hadoop 架构云平台的所有云端共同决定。假设基于 Hadoop 架构云平台中云端数为 m 个，且通常情况下每个云端评判客户端行为的重要性等级是相同的，则云平台中的 NameNode 节点对客户端 x 的信任值为：

$$\theta_{x}=\frac{\sum_{y\in\{1,\ldots,m\}}\theta_{xy}}{m} \tag{3}$$

3 构建动态行为信任链路

本文利用 DataNode、NameNode/备份 NameNode 和 JobTracker 共同参与对客户的信任行为的评价，再采用评价结果控制客户端获取的服务资源，从而防止恶意客户端非法竞争资源，进而降低 Hadoop 架构云安全平台的服务质量。

Hadoop 架构云安全平台信任链路由 DataNode、NameNode/备份 NameNode 和 JobTracker 实现对客户端节点的信任行为评分，Hadoop 架构云平台中的每个 DataNode 节点对客户端进行周期性的信任服务评价，然后将自己对客户端的评估结果传输给 NameNode/备份 NameNode。当 NameNode 节点正常工作时，备份 NameNode 处于“休眠”状态，否则 DataNode 节点返回的信任数据由备份 NameNode 节点接收，再进行信任计算处理。JobTracker 主要用于 Hadoop 架构云平台的监控功能，该节点输入信息是来自 NameNode/备份 NameNode，输出给每个 DataNode 节点用于修正 DataNode 节点对每个客户端节点行为的信任值，则每个客户端的信任链路的状态信息流如图 2 所示。

当客户端节点接入 Hadoop 架构云平台时，信任监控节点 JobTracker 与信任链路控制节点 NameNode/备份 NameNode 进行信任信息交互，在完成信任计算初始化之后，NameNode/备份 NameNode 将选择某些 DataNode 节点对客户端提供存储和计算数据服务，并驱动正在提供服务的 DataNode 节点按照部署的位置次序形成逐一对当前客户端节点的行为进行考核的信任数据链，即上一个 DataNode 节点将信任值通过 I/O 传输链路传递给下一个 DataNode 节点，最后末尾 DataNode

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

节点返回所有信任值给 NameNode/备份 NameNode。当信任链路控制节点 NameNode/备份 NameNode 收到信任评价值时，信任监控节点 JobTracker 与信任链路控制节点 NameNode/备份 NameNode 再次进行信任信息“交换”后，信任监控节点 JobTracker 确定客户端的信任值。

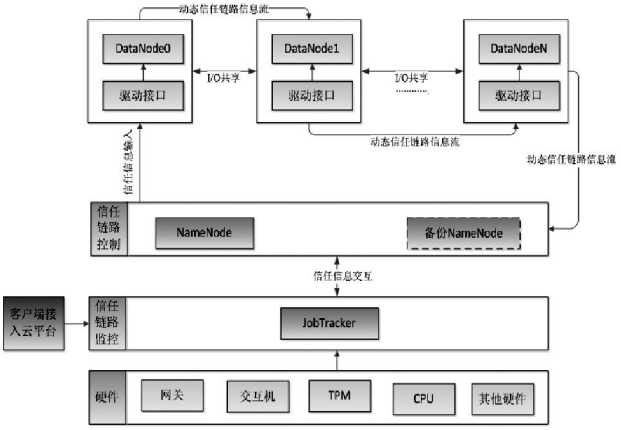

图 2 Hadoop 架构的云平台信任链路的状态信息流

4 动态行为信任计算方法

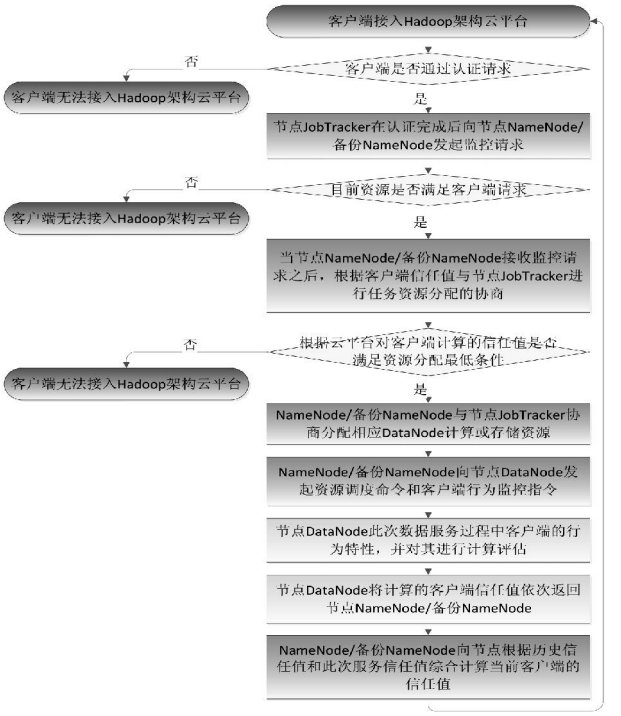

图 3 动态行为信任计算方法

根据 Hadoop 架构云平台信任信息流传输处理与计算结果，每个客户端节点行为信任值由三个参与者决定，它们实体包括：信任监控节点 JobTracker、信任链路控制节点 NameNode/备份 NameNode 和存储节点 DataNode。首先客户端节点在接入 Hadoop 架构云平台之前，信任监控节点 JobTracker 要利用 Kerberos 认证协议对其合法性进行验证，然后给与合法的接入云平台票据授权。客户端节点只有通过了云平台的身份识别，信任监控节点 JobTracker 才对其行为信任进行评估计算，客户端的行为信任值计算流程如图 3 所示。

对于每个接入 Hadoop 架构云平台的客户端，本文采用所有云端参与评价客户端的信任链逐个传递信任值得方式，最后将行为信任评估情况返回给信任监控节点 JobTracker，从而解决“合法”客户端的在接入 Hadoop 架构云平台的“非法”行为，有效提高 Hadoop 架构云平台的服务质量。

5 安全性分析

基于 Bayes 均值信任模型是在充分考虑 DataNode 节点提供服务的质量和效率为基础，分析了 Hadoop 架构云平台“合法”客户端可能受到“被动”的拒绝服务攻击，通过 Hadoop 架构云平台所有可信节点参与计算评估的策略，其主要安全性如下：①Kerberos 认证协议保证了客户端节点接入 Hadoop 架构云平台之后的身份合法性，实现对客户端节点的单点认证，有效防止非授权用户“非法”占用资源。②基于 Bayes 均值信任模型能够防止 Hadoop 架构云平台其他合法客户端受到“被动”的拒绝服务攻击，提高了客户端获取服务资源的公平性。③动态行为信任链路通过 JobTracker 节点、NameNode/备份 NameNode 节点和 DataNode 节点之间动态信任链路的信息传输，能够更为准确地分析客户端节点的接入 Hadoop 架构云平台受服务行为，可以有效防止恶意节点的协同欺骗。④客户端节点安全评分机制是由 Hadoop 架构云平台自动执行，具有自适应特性，可根据客户端执行任务情况实时动态给出每个客户端节点评分值。

6 结语

针对 Hadoop 架构云平台不能验证“合法”用户的行为是否可靠的问题，本文提出了基于 Bayes 均值信任模型，该模型能够防止 Hadoop 架构云平台其他合法客户端受到“被动”的拒绝服务攻击，提高了客户端获取服务资源的公平性，同时采用动态行为信任链路能更为准确地分析客户端节点的接入 Hadoop 架构云平台受服务行为，可以有效防止恶意节点的协同欺骗。然而该方案没有与基于 Kerberos 认证协议融合设计，且云平台通过信任计算之后应该为管理人员提供有效的预警机制，保证其高效的服务质量。

参考文献：

[1] Danil Zburivsky, Sudheesh Narayanan. Hadoop 集群与安全[M]. 机械工业出版社, 2014.
[2] 陈玺, 马修军, 吕欣. Hadoop 生态体系安全框架综述[J].信息安全研究, 2016, 2(3), 684-698.
[3] 林果园. 云计算环境下基于行为信任的访问控制安全技术研究[M].人民邮电出版社, 2016.
[4] 赵科军, 葛连升, 刘洋,等. 基于 Hadoop 和 Spark 构建扩展的网络安全分析平台[J].华中科技大学学报(自然科学版), 2016, 44(增刊 I), 25-28.
[5] 王志华, 庞海波, 李占波. 一种适用于 Hadoop 云平台的访问控制方案[J].清华大学学报(自然科学版), 2014, 54(1), 53-59.
[6] 涂云杰, 白杨. 基于 Hadoop 和双密钥的云计算数据安全存储策略设计, 2014, 22(8), 2629-2631.
[7] 李颖超. 基于 Hadoop 的云存储系统文件处理与安全研究[J].现代电子技术, 2016, 39(21)

基金项目：四川省教育厅 2017 科研项目(No.17ZB0271)。
作者简介：章玉英，讲师，研究方向为信息安全技术。

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net