

网络空间安全态势感知研究

赵淦森¹, 吴杰超¹, 陈梓豪¹, 任雪琦¹, 谭昊翔¹, 马朝辉^{1,2}

(1. 华南师范大学计算机学院, 广东 广州 510631;

2. 广东外语外贸大学思科信息学院, 广东 广州 510006)

摘要: 网络空间是继海、陆、空、太空后的第五大主权空间, 其安全与国家安全紧密相关。如何在网络空间中及时发现安全威胁、防御攻击、溯源攻击等是当下面临的困境之一。对网络空间中多源、异构、割裂的资产, 在构建可信 IaaS 的基础上, 通过基于本体论的键联数据语义网关联技术, 构建网络空间安全态势感知模型, 在保证融合数据访问控制安全的同时, 对网络空间资产进行综合分析, 以感知网络空间的安全态势, 对网络空间治理具有一定的参考意义。

关键词: 网络空间安全; 态势感知; 本体论; 语义网

中图分类号: TP309

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2016.00090

Research on cyberspace security situation awareness

ZHAO Gan-sen¹, WU Jie-chao¹, CHEN Zi-hao¹, REN Xue-qi¹, TAN Hao-xiang¹, MA Zhao-hui^{1,2}

(1. School of Computer Science, South China Normal University, Guangzhou 510631, China;

2. Cisco School of Informatics, Guangdong University of Foreign Studies, Guangzhou 510006, China)

Abstract: Cyberspace is the fifth sovereignty space after the sea, land, sky and space. The security of cyberspace is closely relative to national security. How to discover threat in time, how to defense network attack, how to trace attackers were the problems to be solved. Facing the polyphyletic, isomeric, disconnected properties in the cyberspace, bases on a credible IaaS, a cyberspace security situation awareness was constructed by ontology, semantic Web and linked data theory. While guarantee the access security of integrating data, the model could make a comprehensive analysis on the properties in the cyberspace, which contributes to cyberspace governance.

Key words: cyber security, situation awareness, ontology, semantic Web

1 引言

在信息大爆炸的当下, 网络空间治理面临 3 个问题: 1) 超大规模的网络基础设施和资源, 如何实现有效的管理和利用? 2) 网络空间中巨量的数据, 如何将与业务相关的价值提取出来并且利用好? 3) 如何保障网络空间中的资产安全? 为应对网络空间安全治理挑战, 研究者们已取得一些成果。

在网络空间资产语义化方面, 1998 年,

Berners-Lee 首次提出语义网 (SW, semantic Web)^[1], 语义网给网络空间中的数据添加语义标签, 方便数据分类和分析挖掘。随着语义网的发展, 缺少关联关系的语义数据已经不能满足业务需求, Berners-Lee^[2]在语义网的基础上提出了键联数据, 键联数据使用 URI 来唯一标识资源, 基于语义网, 强调数据关联性。建立数据源链接的方法分为: 基于规则、基于图算法以及基于属性的算法等^[3]。基于语义网的理论体系——本体论, 为构建键联数据提供了方法论和模型^[4]。本

收稿日期: 2016-08-03; 修回日期: 2016-08-28。通信作者: 赵淦森, Gzhao@scnu.edu.cn

体论里互相连接的资产本体可以作为资产的标准表示形式^[5]。在本体论应用方面,文献[6]中对安全信息和事件管理(SIEM, security information and event management)中的信息和操作创建了本体模型。文献[7]创建了 SIEM 的漏洞本体数据模型以及混合架构本体库。文献[8]描述了 SIMS 的知识库本体。基于计算机的安全性和可靠性,可以设计安全本体^[9]。基于攻击的前件和后件,可以为攻击行为建立本体^[10]。利用本体模型可检测攻击行为的上下文^[11]。基于软件漏洞建立本体模型,可对漏洞进行静态分析^[12]。针对资源访问控制,也可以构建安全数据本体模型^[13]。

在网络空间安全管理方面,文献[14]中描述了安全管理平台的主要组成部分。针对事件进行分析是目前信息安全管理平台进行关联分析的主要分析方法^[15]。其中,对于分布式入侵行为,时间模式匹配主要针对时间维度进行分析^[16]。基于 SIEM 的数据模型,混合本体存储架构,可以利用描述逻辑对资源做形式化的描述^[17]。保障安全的重要方法之一是主动防御,利用多种安全工具对系统进行扫描和监控,通过科学的分析方法找到系统漏洞,提前采取行动提高系统的安全性^[18]。常见的安全工具^[19]分为扫描工具、监控工具(如 nmap、tcpdump 等^[20])、入侵检测工具(如 NFR、Snort 等^[21])和防火墙工具。安全管理平台统一管理和运营分离的安全体系和工具,通过集成各种安全工具综合管理系统^[22]。

在网络空间基础设施方面,云计算技术将大规模的 IT 物理资源虚拟化,实现大资源管理,大能力构建,用户以租户的角色通过互联网使用这些资源和服务。云计算面向用户的服务方式主要有 3 种:1) 基础设施即服务(IaaS),如国内的阿里云、盛大云,国外的亚马逊云;2) 平台即服务(PaaS),如国内的新浪云、国外的 GAE;3) 软件即服务(SaaS),如 Gmail。用户不必再购置机器、不必再购买软件光盘,取而代之的是通过互联网去租用云计算提供的这些服务。将海量的数据清洗、融合、分析挖掘,结合相关业务,可以让原本孤立的、多源的、异构的数据产生巨大的价值。

2 网络空间安全态势感知挑战

1) 采集的数据多源、异构、割裂

目前,网络空间的安全管理基于多种安全工具,如 nmap、awvs、openvas、ossec 等,数据采集方式的多元化和网络监测方式的多样化,为网络空间态势感知提供丰富的数据资源。然而,网络空间数据源很多,不同数据源产生的数据结构不一致,不同安全工具所获得的数据结构也不一致。由于融合和关联多层次数据方法的缺失,使对网络空间安全监控获得的多维度的数据变孤立、割裂。这些问题对于综合多方面的数据分析网络空间安全态势造成了一定的挑战。

2) 网络空间基础设施不可信

云计算时代,IaaS 将云计算中心的物理资源转化为由虚拟机组成的资源池,实施资源池化管理。这种资源池化管理方式将底层物理节点的管理细节对外隐藏。从安全的角度看,资源池化管理模式限制了用户鉴别部署服务的虚拟机是否由可信的物理节点承载的能力,使用户对从 IaaS 云服务商获取的虚拟机存在固有的不信任。网络空间安全态势感知平台应该搭建于可信的基础设施之上。

3) 用户身份难以鉴别

一方面,网络空间用户的匿名化造成即使通过攻击溯源找到了攻击者的 IP 地址、MAC 地址等信息,仍难以确认攻击者的身份。另一方面,攻击者可以通过僵尸机、肉鸡等来实施攻击行动,隐藏自己的身份,进一步加大了网络空间攻击的精确溯源难度。

4) 融合数据时访问控制冲突

面对在网络空间中采集到的多源、异构、割裂的数据,为了整合资源,解决信息孤岛问题,需要进行数据融合。但在数据融合过程中,会遇到如何保证数据机密性的问题。在数据融合过程中,需要保护数据不被非授权用户访问,避免数据泄露。此外,不同的数据源使用各自的访问控制策略,面对融合后的数据,不能直接使用原来某一数据源的访问控制策略,需要产生新的融合访问控制策略。其中,如何解决融合过程中的访问控制策略冲突问题是网络空间数据关联分析

的一个挑战。

3 网络空间安全态势感知模型

3.1 模型简介

网络空间安全态势感知模型如图 1 所示。融合云计算技术和大数据技术,在构建可信 IaaS 的基础上,通过基于本体论的键联数据语义网关联技术,构建网络空间安全态势感知模型,在保证融合数据访问控制安全和融合数据搜索安全的同时,对网络空间资产进行综合分析,以感知网络空间的安全态势。

3.2 模型关键技术

3.2.1 云计算池化信任管理

为了构建可信的 IaaS,支撑上层的网络空间安全态势感知平台。根据面向资源池化管理的信任管理模型,设计由具有匿名信任关系的可信的物理机组成的可信集群和能够提供可信虚拟机的

可信资源池。将 IaaS 云服务中心的物理机组织成为物理主机集群,以集群为单位承载用户的软件服务。集群与用户服务形成一对一的对应关系。以一种匿名的群签名密码系统为依托,在集群内部的物理主机之间构建起具有匿名的可验证的信任关系,将普通的物理主机集群变为物理主机之间具有信任关系的可信集群,实现可信集群内物理主机的信任关系的建立、验证和撤销。可信集群能够满足资源池化管理模式下对物理主机匿名的信任管理需求,允许在物理主机之间建立信任关系同时又实现对物理主机真实身份的隐藏。以可信集群为基础,构建能够提供可信虚拟机的可信资源池。可信资源池让用户参与到可信资源池的创建过程,将创建可信资源池的责任由用户和 IaaS 提供者共同承担。根据资源的实时安全属性和基准安全属性,用户在使用虚拟机之前,先通过虚拟机提供的信任验证接口获取虚拟机的实时

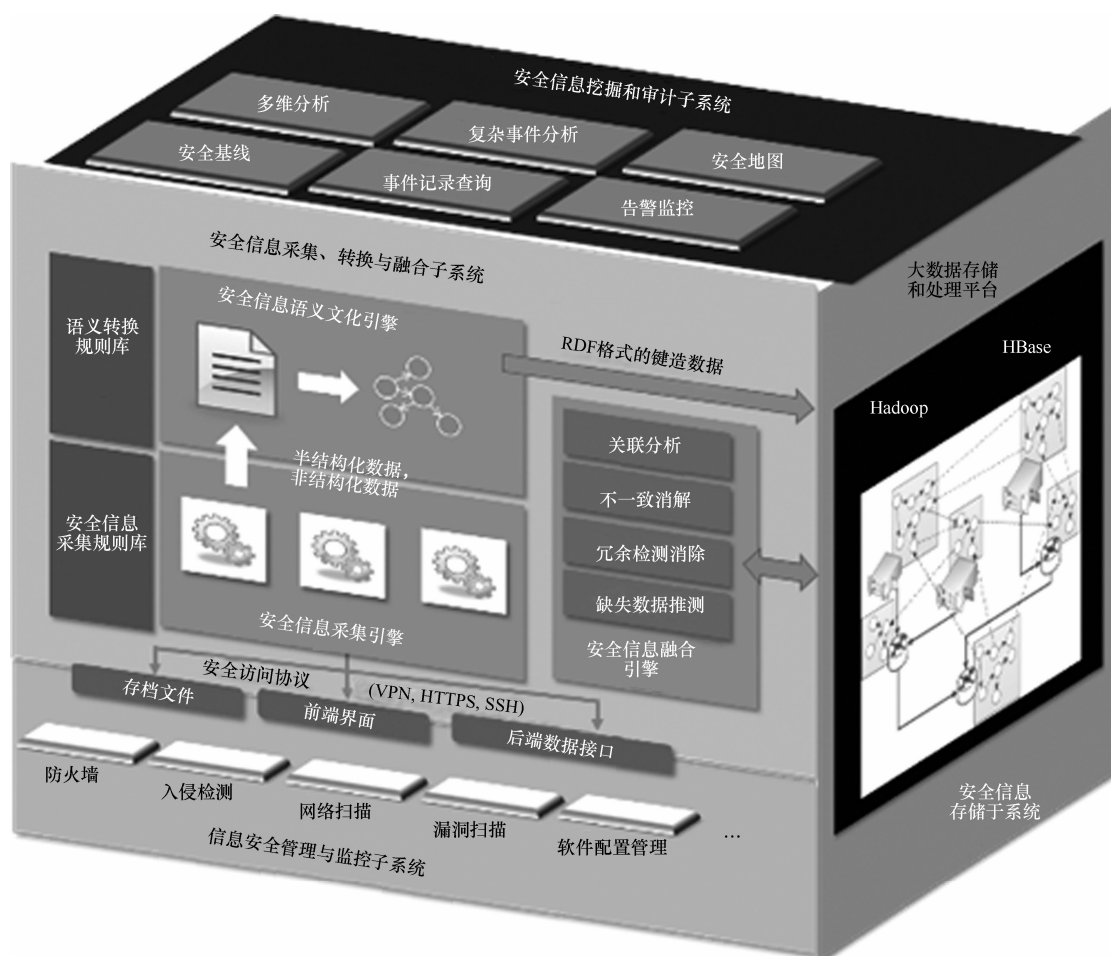


图 1 网络空间安全态势感知模型

安全属性, 然后与虚拟机的基准安全属性对比, 如果一致就表明虚拟机是可信的; 否则, 虚拟机就是不可信的。另外, 根据虚拟机迁移协议能够在迁移源节点和迁移目的节点建立可靠的传输通道, 能够抵抗中间人攻击。

3.2.2 融合数据搜索的多级安全模型

为了解决融合网络空间数据时出现的访问控制冲突问题, 设计面向多级安全访问控制的数据融合模型。假设所有数据源都基于 BLP 访问控制模型, 即访问策略由 Lattice、安全级别映射函数以及访问矩阵 3 个部分组成。将访问策略融合的 Lattice 合并转化为 Hasse 图的合并, 访问策略融合的 3 个步骤分别为: 初始处理、冲突处理和化简处理。在 Hasse 图中添加关联线段、删除冲突线段并进行化简。映射函数转换是根据原始 Hasse 图与融合后 Hasse 图上的节点进行映射; 而策略融合中的访问矩阵融合是对访问用户和数据集的扩充合并。对于融合数据的安全级别标识问题, 根据原始数据集之间的关系选择最小上界安全级别方法, 保证其在原有数据集都被访问的情况下新融合数据集才允许访问。通过数据机密性进行形式化定义和理论推导, 可以证明基于多级安全的融合数据搜索模型是正确的、安全的。结合访问策略融合中的 3 个处理步骤, 可以证明融合策略与原始的访问策略具有一致性。

3.2.3 基于键联数据语义关联分析的安全管理

为了对割裂的数据进行关联分析、攻击溯源以及管理, 基于本体论和键联数据的理论依据, 为安全数据融合和关联分析构建安全本体模型, 实现基于键联数据语义关联分析的安全管理。首先, 根据综合分析安全管理的需求以及安全数据的特征, 利用本体论和键联数据建立安全本体模型, 通过计算公共属性的语义距离的方式, 基于属性和规则构建安全数据之间的关联关系, 融合不同数据源的安全数据, 关联同一实体的不同实例数据, 实现数据的整合和关联; 其次, 针对安全本体模型实例化的数据, 从不同维度对资产的关联数据进行分析, 通过设定基于属性和语义距离的方式关联资产与其他实体, 关联不同数据源的同一资产实体, 实现去冗余的过程; 最后, 结合安全本体模型, 对键联数据进行语义关联分

析, 设计安全本体实例化算法和溯源分析算法, 基于资产的日志语义属性对日志进行密度聚类, 从日志的语义属性中分析资产状态, 根据资产日志从键联数据中找到资产异常日志相关数据, 从多角度的语义关联关系得到发生异常状态的关联数据。

3.2.4 基于语义网技术的安全态势评估

为了对网络空间安全态势进行客观全面的评估, 运用语义网技术, 将多源异构的安全数据进行有效的融合及语义化, 对大规模分布式基础设施系统进行形式化的描述, 将物理世界的分布式系统以更有利于机器理解的 RDF 图形式呈现到信息世界中。并在此基础上运用逻辑推理推导分布式系统中资产的安全态势情况, 进而得到整个分布式基础设施系统的安全态势情况。首先, 运用语义网技术将收集到的数据语义化, 从而对分布式系统进行形式化描述。构建针对分布式系统态势评估的安全本体, 根据基于语义网技术的原始数据实例化算法和关联跨数据集的资产实例的关联规则, 合并同一资产实体的资产实例。其次, 基于推理规则推导分布式基础设施系统中资产的安全态势情况, 根据推理 Host 的安全态势的推理规则模型, 汇总 Host 安全态势得出网络空间安全态势。

4 模型实现

1) 可信 IaaS 搭建

基于 OpenStack 构建可信集群和可信资源池。用户与 IaaS 资源之间的信任认证关系如图 2 所示。信任验证原理就是用户在使用虚拟机之前, 先获取虚拟机和虚拟机所在集群成员主机(物理机)的实时安全属性; 然后分别对比虚拟机的实时安全属性和虚拟机的基准属性、集群成员主机的实时安全属性和基准安全属性, 如果都一致, 说明虚拟机是可信的, 虚拟机所在物理机也是可信的。

2) 数据管理模块

数据管理模块基于 jena 框架及 TDB 存储模块设计与键联数据操作相关的接口, 包括本体模型的创建、提取, 键联数据的增加、删除、修改等操作, 以及键联数据规则推理接口。

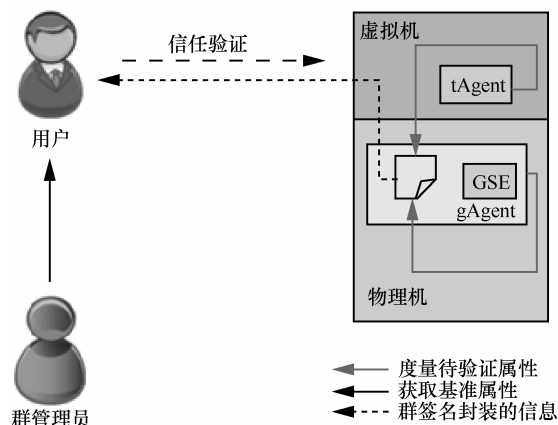


图 2 可信 IaaS 的资源信任验证

3) 数据采集模块

数据采集模块集成多种安全扫描工具，负责扫描和监控网络空间环境，采集相关数据（如资产分布情况、目前的开放情况、提供的服务信息、当前的运行状态、存在的漏洞等信息）。该模块程序根据一个预设的频率 f 定时执行，周期性地采集和更新网络空间中的资产状态（如图 3 所示）。

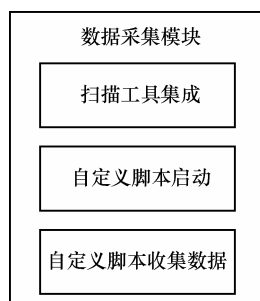


图 3 数据采集模块

4) 语义化模块

语义化模块对数据采集模块中不同的安全工具采集回来的数据进行融合，针对安全数据设计安全本体模型，用融合后的数据实例化安全概念，根据规则对实例数据设定关联关系，为数据添加语义标签，将关联后的数据以三元组的方式存储为键联数据（如图 4 所示）。



图 4 语义化模块

5) 分析模块

分析模块利用键联数据对融合后的安全日志数据进行分析，使用基于语义属性密度聚类的算法对日志内容进行分析，并在键联数据中回溯异常日志相关的数据，对其关联数据进行分析以得到异常的相关信息，最终发现网络空间中的异常事件（如图 5 所示）。

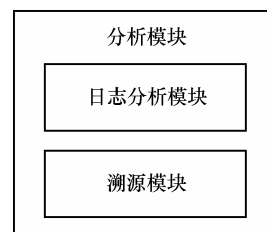


图 5 分析模块

6) 可视化模块

可视化模块通过 Web 网页为系统提供友好的可视化界面，用户可以在界面上设定扫描的工具、目标监控的网络范围、采集数据的频率等参数，同时为用户提供了该网络范围相关的键联数据关系情况，以及资产分析结果，方便用户查看相关的数据（如图 6 所示）。

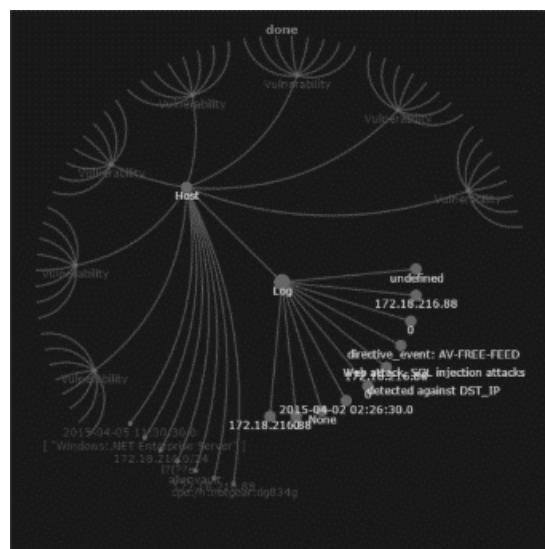


图 6 可视化模块

5 结束语

没有网络安全就没有国家安全，网络空间作为继海、陆、空、太空后的第五大主权空间，其安全与国家安全紧密相关。云计算技术能够

有效地管理和利用超级大规模的基础设施跟资源,大数据技术在巨量的数据中将与业务相关的价值给提取出来并且利用好。融合云计算和大数据技术,在网络空间中发现资产、对资产进行提取、分析挖掘,进而掌握整个网络空间的安全态势,有助于为网民提供一个安全的网络环境、有助于为企业保驾护航、有助于为国家发现和防御网络攻击保障国家网络空间安全提供支撑。但无论是哪一种技术,其本身都存在着一定的安全问题。本文提出的网络空间安全态势感知模型,融合云计算技术和大数据技术,在构建可信 IaaS 的基础上,通过基于本体论的键联数据语义网关联技术,构建网络空间安全态势感知模型,在保证融合数据访问控制安全和融合数据搜索安全的同时,对网络空间资产进行综合分析,以感知网络空间的安全态势,对网络空间安全治理有一定的参考意义。然而目前本文所能获得的网络空间资产有限,非实时发现与分析,数据处理规模较小,未来可以改进的地方还很多。

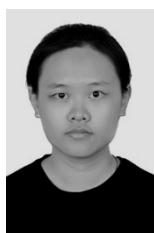
参考文献:

- [1] BERNERS-LEE T, HENDLER J, LASSILA O. The semantic Web[J]. Scientific American, 2001, 284(5): 28-37.
- [2] BERNERS-LEE T, CHEN Y, CHILTON L, et al. Tabulator: exploring and analyzing linked data on the semantic Web[C]//The 3rd International Semantic Web User Interaction Workshop. 2006.
- [3] POULOVASSILIS A, WOOD P T. Combining approximation and relaxation in semantic Web path queries[M]//The Semantic Web-ISWC 2010. Berlin Heidelberg: Springer, 2010: 631-646.
- [4] MCCRAE J, SPOHR D, CIMIANO P. Linking lexical resources and ontologies on the semantic Web with lemon[M]//The Semantic Web: Research and Applications. Berlin Heidelberg: Springer, 2011: 245-259.
- [5] BIRKHOLZ H, SIEVERDINGBECK I, SOHR K, et al. IO: an interconnected asset ontology in support of risk management processes[C]//The 7th International Conference on Availability, Reliability and Security (ARES). 2012: 534-541.
- [6] GRANADILLO G G, MUSTAPHA Y B, HACHEM N, et al. An ontology-based model for siem environments[M]//Global Security, Safety and Sustainability & e-Democracy. Berlin Heidelberg: Springer, 2012: 148-155.
- [7] KOTENKO I, POLUBELOVA O, SAENKO I. The ontological approach for siem data repository implementation[C]//2012 IEEE International Conference on Green Computing and Communications. 2012: 761-766.
- [8] DEVERGARA J E L, VILLAGRÁ V A, HOLGADO P, et al. A semantic Web approach to share alerts among security information management systems[M]//Web Application Security. Berlin Heidelberg: Springer, 2010: 27-38.
- [9] EKEHART A, FENZ S, KLEMEN M, et al. Security ontologies: improving quantitative risk analysis[C]//International Conference on System Sciences. 2007: 156a.
- [10] AHMAD S, MORTEZA A. Predicting network attacks using ontology-driven inference[J]. International Journal of Information and Communication Technology, 2013, 4(9):13.
- [11] RAZZAQ A, ANWAR Z, AHMAD H F, et al. Ontology for attack detection: an intelligent approach to Web application security[J]. computers & security, 2014, 45: 124-146.
- [12] YU L, WU S Z, GUO T, et al. Ontology model-based static analysis of security vulnerabilities[M]//Information and Communications Security. Berlin Heidelberg: Springer, 2011: 330-344.
- [13] MENG Y L, YIN G S, WANG H Q. Study on the ontology modeling method of Heterogeneous security data integration[J]. Computer Engineering and Applications, 2010, 46(19):138-140.
- [14] BIDOU R. Security operation center concepts & implementation[EB/OL]. <http://www.iv2-technologies.com>.
- [15] YUAN S, ZOU C. The security operations center based on correlation analysis[C]//International Conference on Communication Software and Networks (ICCSN). 2011: 334-337.
- [16] ZHANG D, ZHANG D. The analysis of event correlation in security operations center[C]//International Conference on Intelligent Computation Technology and Automation (ICICTA). 2011(2): 1214-1216.
- [17] KOTENKO I, POLUBELOVA O, SAENKO I. The ontological approach for siem data repository implementation[C]//Green Computing and Communications (GreenCom). 2012: 761-766.
- [18] KUMAR G, KUMAR K. Network security-an updated perspective[J]. Systems Science & Control Engineering: An Open Access Journal, 2014, 2(1): 325-334.
- [19] KAUR A, SALUJA M. Study of network security along with network security tools and network simulators[J]. International Journal of Computer Science & Information Technologies, 2014, 5(1): 88-92.
- [20] BEJTILICH R. The practice of network security monitoring: understanding incident detection and response[M]. No Starch Press, 2013:100-120.
- [21] ANAND V. Intrusion detection: tools, techniques and strategies[C]//The 2014 ACM Siguccs Annual Conference on User Services. 2014:69-73.
- [22] SUNDARAMURTHY S C, CASE J, TRUONG T, et al. A tale of three security operation centers[C]//The 2014 ACM Workshop on Security Information Workers. 2014: 43-50.

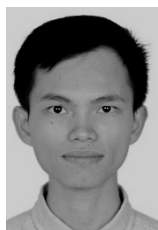
作者简介：



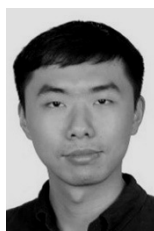
赵淦森（1977-），男，广东东莞人，博士，华南师范大学计算机学院副院长、教授、博士生导师，主要研究方向为云计算、大数据、信息安全。



任雪琦（1993-），女，广东惠州人，华南师范大学硕士生，主要研究方向为云计算与信息安全。



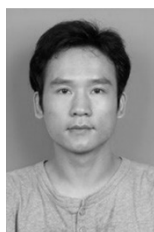
吴杰超（1993-），男，广东廉江人，华南师范大学硕士生，主要研究方向为云计算与信息安全。



谭昊翔（1990-），男，广东江门人，华南师范大学硕士生，主要研究方向为云计算与信息安全。



陈梓豪（1992-），男，广东广州人，华南师范大学硕士生，主要研究方向为云计算与信息安全。



马朝辉（1974-），男，湖南湘潭人，华南师范大学博士生、广东外语外贸大学讲师，主要研究方向为网络安全、云计算和大数据等。