文章编号：2096-1472(2020)-08-52-04
DOI:10.19644/j.cnki.issn2096-1472.2020.08.015

基于语义本体的高校网络信息安全管理辅助决策系统研究及应用

邹 伟[1]，陈 恳[1]，欧阳昭相[2]

(1.云南师范大学，云南 昆明 650500;
2.德宏师范高等专科学校，云南 芒市 678400)
✉4236@ynnu.edu.cn; 3169@ynnu.edu.cn; 3451165293@qq.com

摘 要：面向海量异构多源的高校网络信息安全数据管理和辅助决策系统研究及应用是目前高校信息安全体系建设的重点和难点课题。本文首先提出了基于语义本体技术的多源异构高校网络信息安全知识整合策略，在此基础上构建了基于可计算和推理信息资源的网络信息安全辅助决策知识库，设计并开发了面向高校网络信息安全管理的辅助决策支持服务系统平台。通过常用网络攻击知识问答库数据对系统功能测试结果表明，该系统能够较好地实现对高校网络信息安全事件的关联查询、知识推理和辅助决策等支持服务，为高校网络信息安全领域多源异构数据整合、知识建模及智能管理提供了有效地解决方案。

关键词：语义本体；网络信息安全；知识库
中图分类号：TP301 **文献标识码**：A

Research and Application of Decision Support System for Network Information Security Management of Universities based on Semantic Ontology

ZOU Wei[1], CHEN Ken[1], OUYANG Zhaoxiang[2]

(1.*Yunnan Normal University*, *Kunming* 650500, *China*;
2.*Dehong Teachers' College*, *Mangshi* 678400, *China*)
✉4236@ynnu.edu.cn; 3169@ynnu.edu.cn; 3451165293@qq.com

Abstract: The research and application on data management and decision support system dealing with massive heterogeneous multi-source for network information security is the key and difficult subject in informationization process of universities. This paper firstly proposes a knowledge integration strategy for multi-source and heterogeneous university network information security knowledge based on semantic ontology technology. And then, a knowledge base for network information security decision support is built utilizing computable and inferential information resources, and a decision support service system is developed as well. The results of system function test on a constructed common network attack knowledge quiz library show that the proposed system can well implement the associated query, knowledge inference and decision-making support services for network information security incidents, which provides an effective solution to multi-source heterogeneous data integration, knowledge modeling and intelligent management in network information security domain.

Keywords: semantic ontology; network information security; knowledge base

1 引言(Introduction)

对海量异构多源的高校网络信息安全数据进行整合，结合专家知识库语义知识，为网络信息安全管理提供决策支持，是目前摆在高校网络信息安全管理者面前的一个难题。已有的研究主要集中于基于多源异构安全数据聚合、网络安全态势知识库模型构建、网络入侵知识库模型构建及网络威胁推演及信息安全区划分等方面[1−3]，这些研究在利用本体库进行网络入侵知识库构建、异构网络信息安全整合及威胁过程推演等方面，都取得了一定的进展。但是，缺乏以普通安全管理员为对象的决策支持知识库系统构建方面的研究。

利用语义本体集成海量异构多源的网络信息安全数据，将利用文档、表格、图片等形式存储的网络信息安全知识和专家知识转化为计算机可以直接计算和推理的资源[4,5]，为安全管理运维人员形成有价值的决策支持信息，帮助开展网络

基金项目：云南省应用基础研究计划重点项目(2016FA024)；民族教育信息化教育部重点实验室委托基金项目(2019002)；云南省教育厅科学研究项目(2019J0078).

信息安全防护。本文提出以语义技术对信息安全知识和专家知识进行整合，形成可计算和推理的信息资源，并在此基础上，整合网络及安全设备实时数据，开发了网络信息安全决策支持服务系统，实现了对网络信息安全事件的知识推理和决策辅助支持服务。

2　基于语义本体的网络信息安全辅助决策框架模型(Framework model of network information security assistant decision based on semantic ontology)

2.1　基于本体的网络信息安全辅助决策框架

通过对现有网络信息安全管理专家知识、安全防控规则及行业规范数据进行本体描述和语义集成，构建一套辅助决策框架，利用语义查询和语义推理技术，得到信息安全过程决策支持服务，协助网络信息安全运维人员进行辅助决策[6,7]。如图1所示。

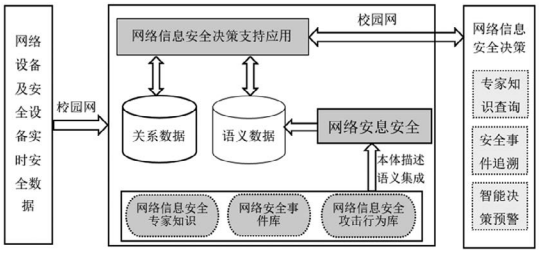

图1 基于语义本体的高校网络信息安全辅助决策框架

Fig.1 Framework model of network information security assistant decision based on semantic ontology

图1是基于语义本体的高校网络信息安全辅助决策框架。该框架的基础数据核心是网络信息安全设备实时数据及网络信息安全专家知识、安全事件库及网络信息安全攻击行为库等语义知识数据，它由局域网实时安全数据采集、网络信息安全专家知识及攻击行为知识库等知识的语义集成和决策辅助服务，核心是网络信息安全决策支持应用程序，通过RDF查询语言进行语义查询和语义推理，生成决策支持信息，通过不同类型终端，为安全管理运维人员提供辅助决策。

(1)局域网实时安全数据采集。通过部署在校园网络的多台日志采集服务器，将网络交换设备、防火墙设备、上网行为审计设备、应用防火墙、统一身份认证设备及漏洞扫描设备的相关日志数据进行采集，进行清洗、过滤和归并操作，存入关系数据库。

(2)知识数据语义集成。利用语义技术，将现有的网络信息安全中的专家知识、安全事件库及网络信息安全攻击行为等知识数据进行语义集成，转化为RDF三元组，构建出网络信息安全本体，实现对网络信息安全知识的一致性描述，为建立推理规则、进行自动推理打下基础。通过构建的网络信息安全过程本体库，生成语义数据库。

(3)辅助决策服务。形成了关系数据库和语义数据库的基础上，由网络信息安全支持应用程序进行调用，通过RDF查询语言进行语义查询和语义推理，生成决策支持信息，通过WEB终端，为网络信息安全管理者提供辅助决策支持。

2.2　基于本体的网络信息安全辅助决策三层结构模型

基于本体高校网络信息安全辅助决策系统，采用三层B/S架构体系，其结构体系分为三个层次，即数据资源层、决策层和用户接口层，如图2所示。数据资源层主要是对支持信息安全决策所需的数据资源进行获取和存储，决策层主要实现把接口层输入的用户任务，进行计算和推理，并映射成数据资源层中相应数据组合；用户接口层负责用户决策需求输入和决策结果显示。

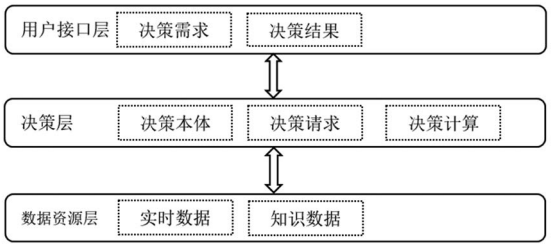

图2 基于语义本体的网络信息安全辅助决策三层结构模型

Fig.2 Three-tier structure model of network information security decision support based on semantic ontology

(1)数据资源层。数据资源层主要包含通过校园网络安全设备采集的实时日志数据及网络信息安全决策的各类知识数据。实时日志数据主要包含网络防火墙、上网行为管理系统、反病毒软件、远程漏洞检测系统、反垃圾邮件网关、应用防火墙、数据库审计系统、堡垒主机等日志数据等，进行标准化处理后存放在关系数据库。知识数据主要包含网络信息安全事件规则特征、防护规则，以及日常处理知识数据，进行本体描述和语义集成后，存放在语义数据库。

(2)决策层。决策层主要是通过RDF查询语言进行语义查询和语义推理，分为决策本体、决策请求和决策计算三部分。决策本体是对网络信息安全领域知识的形式化表达与描述。因此在决策层中，用户通过接口层输入决策请求，决策层对用户的决策请求进行解析，将以自然语言描述的决策请求转换为决策本体实例，决策计算模块进行计算和推理，并映射成数据资源层中相应数据组合。完成决策任务的计算，并向接口层中返回计算结果。

(3)用户接口层。用户接口层主要由信息安全决策需求输入与决策结果展示这两个功能模块组成。用户通过自然语言输入决策需求，随后系统开始对需求进行解析和计算，得到决策结果后，通过表格和图表方式进行决策结果展示。系统不仅可以在安全管理运维人员发出请求时提供决策支持信息，并且可以提前设置推送策略，将一些决策结论主动推送给相关人员。

3　网络信息安全本体库构建(Construction of network information security ontology library)

构建网络信息安全本体库，首先明确本体涉及的范围，在此基础上抽象出网络信息安全领域的关键概念作为本体的类，并通过定义类之间的关系来描述概念之间的联系、定义类与实例之间的属性来描述概念与个体之间的联系[8-10]。网络信息安全本体库构建流程图如图3所示。

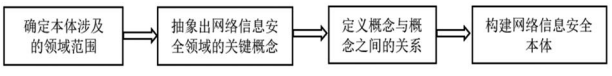

图3 网络信息安全本体库构建流程

Fig.3 Construction process of network information security ontology library

3.1　网络信息安全要素的分类提取

网络信息安全要素，是指在多源异构的数据源中能够引起网络信息安全状态发生变化的一些列基本元素[9]。根据这些

要素数据来源、实时性、可靠性及冗余度的不同，可以分为物理安全、网络安全和数据安全：

(1)物理安全：物理安全是包含计算机网络设备设备、网络拓扑配置、机房空调、供配电、监控系统等设施的日志记录。

(2)网络安全：主机安全、反病毒、系统安全检测、审计分析网络运行安全、WEB应用防火墙、网络安全检测、入侵检测、第二代防火墙、实名认证系统、上网行为管理系统等网络信息安全系统日志信息。

(3)数据安全：主要涉及信息传输的安全、信息存储的安全以及对网络传输信息内容的审计三个方面[11]，具体包括数据加密、数据完整性鉴别、防抵赖、信息存储安全、数据库安全、终端安全、信息的防泄密、信息内容审计、用户授权等数据库操作日志[12,13]。

3.2 从领域知识到RDF三元组

德国学者Studer等人首次提出了较完整的本体定义[14]。本体是对世界上客观存在事物的系统的描述，是领域知识的规范表示，它定义了概念以及概念之间的关系，是表达、组织、理解、展示和预测知识的常用方式之一[15-17]。目前本体被广泛应用于知识工程、智能信息集成、数据挖掘、海量信息的组织和处理等领域中[18,19]。目前普遍认为本体是对某一领域内的概念及其关系的一种概念化描述[18]。采用如下方式来定义三元组：

NetSecurity=(S，V，O)

其中，S为主语、V为谓语、O为宾语。一个三元组由主语、谓语、宾语构成，那么网络信息安全本体可以看成是由多个SVO三元组组成的集合。网络信息安全知识主要由文本型、表格型、图片型和网页型知识经预处理提取而来。文本型知识需要由领域专家进行分析，提炼成问答形式，然后由软件开发人员根据系统建设需求，建立相应的类、实例和属性，最终转化为RDF三元组。表格型知识可以直接转换为RDF三元组，表格的各个行的第一个单元格内容一般转换为三元组的主语，各个列的第一个元素的内容转换为三元组的谓语，各个行与各个列相交的单元格内容则转换为三元组的宾语。图片型知识通过属性rdfs:seealso把图片型知识连接进去，图片型知识直接转换为宾语。网页型知识把整个网页、网页的一部分、或者网页的集合的资源转换为主语，把描述某个资源的特征、性质或关系等网页型知识的属性转换为谓语，把描述网页型知识的资源的属性值转换为宾语。

3.3 网络信息安全本体SPARQL查询

当网络信息安全本体构建好以后，就要对所构建的本体创建一些测试实例，通过测试实例来检测所构建的本体模型是否达到预期目的，所构建的网络信息安全本体模型是否存在错误。语义本体可以通过SPARQL语言来处理，它是RDF数据的标准查询语言。采用SPARQL查询作为测试方法，检测计算机理解的网络信息安全本体三元组和人所理解的网络信息安全本体三元组的一致性。

SPARQL的SELECT查询语句查询变量和包含查询变量的三元组两个部分。比如，在前面对ARP攻击问题进行描述时，当关于ARP攻击及其类型的本体构建好以后，就可以运用SPARQL来做查询，可以把查询的问题表述为：“特洛伊木马攻击的特征及解决办法是什么？”对应的SPARQL查询语句为：

```
SELECT ?x?y?z
WHERE{
?x rdfs:label “特洛伊木马”,
?x:PorosityMin ?y,
?x:PorosityMax ?z
}
```

4 系统实现及功能验证(System implementation and functional verification)

4.1 系统实现

系统前端采集的网络信息安全日志，存放于ORACLE数据库。网络信息安全本体和推理规则存放于语义本体数据库。软件采用Protégé对网络信息安全本体库进行建模，TopBraid Composer进行开发，内置Pellet进行语义查询推理，构建了网络信息安全知识库检索及辅助决策功能。

4.2 功能验证

为了验证该模块的正确性，我们根据采集的“常用网络攻击知识问答库”进行测试。如表1是从网络信息安全知识检索库中提取的九条验证样本。

表1 网络信息安全知识查询测试样本

Tab.1 Test sample of network information security knowledge query

输入	输出	
攻击行为	攻击特征	解决办法
死亡之ping	攻击包加载的尺寸超过64K上限的ICMP包，系统就会出现内存分配错误，导致TCP/IP堆栈崩溃，出现宕机	对防火墙进行配置，阻断ICMP和任何未知协议
UDP洪水	各种假冒攻击利用简单的TCP/IP服务，如Chargen和Echo来传送毫无用处的占满带宽的数据	关掉不必要的TCP/IP服务。防火墙配置阻断UDP请求
Smurf攻击	通过使用将回复地址设置成受害网络的广播地址的ICMP应答请求数据包，导致该网络的所有主机都对此ICMP应答请求作出答复，最终网络堵塞	关闭外部路由器或防火墙的广播地址特性。为防止被攻击，在防火墙上设置规则，丢弃掉ICMP包
电子邮件炸弹	通过设置一台机器不断的大量的向同一地址发送电子邮件，攻击者能够耗尽接受者网络的带宽	对邮件地址进行配置，自动删除来自同一主机的过量或重复的消息
特洛伊木马	特洛伊木马是通过秘密手段将木马安装到目标系统。一旦安装成功并取得管理员权限，黑客就可以直接远程控制目标	避免下载可疑程序并拒绝执行，监视内部主机上的监听TCP服务
端口扫描	使用一些软件，向大范围的主机连接一系列的TCP端口，扫描软件报告它成功的建立了连接的主机所开的端口[20]	通过防火墙屏蔽端口扫描
DNS域转换	DNS协议不对转换或信息性的更新进行身份认证，黑客只需实施一次域转换操作就能得到你所有主机的名称，以及内部IP地址	在防火墙处过滤掉域转换请求
LDAP服务	黑客使用LDAP协议窥探网络内部的系统和它们的用户的信息	对于刺探内部网络的LDAP进行阻断并记录，对于公共LDAP服务，将其放入DMZ
DdoS攻击	攻击代理主机向目标主机发送大量的服务请求数据包，数据包所请求的服务往往要消耗大量的系统资源，造成目标主机无法为用户提供正常服务。甚至导致系统崩溃	在服务器前端加CDN中转，域名解析使用CDN的IP，隐藏服务器真实IP地址[21]

查询用户搜索相关内容，系统对内容进行解析，通过本体推理，将推理结果返回给用户。在系统验证过程中，输入为攻击行为，数据输出为攻击特征及解决办法。

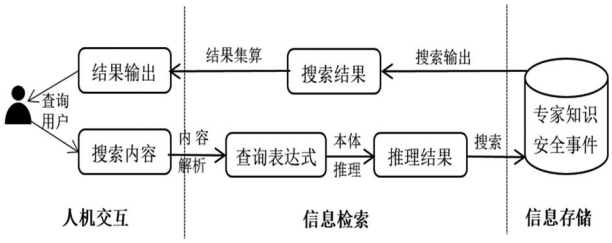

图4 基于语义本体的网络信息安全知识检索流程

Fig.4 The process of network information security knowledge retrieval based on semantic ontology

验证结论：将测试数据输入网络信息安全知识库查询系统，得到的运行结果如图5所示。将系统运行结果与测试数据输出结论相比较完全一致。由此可以证明，辅助决策知识库系统能够根据安全运维管理人员的输入问题，得到正确的描述及解决办法，为网络信息安全辅助决策提供支持。

图5 网络信息安全辅助决策知识库系统检索界面

Fig.5 Interface for retrieval of network information security assisted decision knowledge base system

5 结论(Conclusion)

本文提出了基于本体的网络信息安全辅助决策框架模型，并在此基础上通过网络信息安全本体库构建，实现了高校网络信息安全辅助决策知识库原型系统，经测试表明系统具有一定的辅助决策能力，本体技术作为网络信息安全决策支持系统的支撑技术是可行的。由于网络信息安全决策分析过程复杂，下一步需要继续对语义本体进行扩展，以保障系统在复杂决策分析过程中更加智能。

参考文献(References)

[1] 王世伟.论信息安全、网络安全、网络空间安全[J].中国图书馆学报,2015,41(02):72-84.

[2] 马民虎,张敏.信息安全与网络社会法律治理:空间、战略、权利、能力——第五届中国信息安全法律大会会议综述[J].西安交通大学学报(社会科学版),2015,35(02):92-97.

[3] 华辉有,陈启买.基于本体的网络安全态势知识库模型[J].计算机应用,2014,34(S2):95-98;107.

[4] 朱颖.基于语义技术的柑橘园土壤环境判定决策支持系统[D].西南大学,2014.

[5] 朱丽娜.本体论在网络安全态势感知中的应用[J].数字技术与应用,2018,36(05):188-189.

[6] 李彦旭,巴大志,成立.网络信息安全技术综述[J].半导体技术,2002(10):9-12;30.

[7] 韦勇,连一峰,冯登国.基于信息融合的网络安全态势评估模型[J].计算机研究与发展,2009,46(03):353-362.

[8] 司成,张红旗,汪永伟,等.基于本体的网络安全态势要素知识库模型研究[J].计算机科学,2015,42(05):173-177.

[9] 华辉有,陈启买.基于本体的网络安全态势知识库模型[J].计算机应用,2014,34(S2):95-98;107.

[10] 张殊.基于语义Web服务的组合方案研究及其在电子商务中的应用[D].南京理工大学,2009.

[11] 洪大翔.基于本体的信息系统安全域划分的研究与应用[D].重庆大学,2014.

[12] 张德祥,刘勋,扈炳良.校园信息网络安全体系构建研究[J].情报科学,2009,27(10):1542-1544.

[13] 邵雪航,周洪玉.企业网络安全体系构建[J].机械工程师,2007(09):104-105.

[14] Studer B, Benjamins V R, Fensel D. Knowledge engineering: Principles and methods[J]. Data and Know Ledge Engineering, 1998, 25(2): 161-197.

[15] Rodriguez G D, Velasco D. Ontologies for network secu-rity and future challenges[C]. 12th International Conferenceon Cyber Warfare and Security, 2017.

[16] Simmonds A, Sandilands P, Ekert L. An ontology for net-work security attacks[C]. Second Asian Applied ComputingConference on Applied Computing, 2004.

[17] Staab S,Studer R.Handbook on ontologies[M].International handbooks on information systems.Berlin:Springer, 2004: 227-255.

[18] 华辉有,陈启买.基于本体的网络安全态势知识库模型[J].计算机应用,2014,34(S2):95-98;107.

[19] 陶晓玲,韦毅,孔德艳,等.基于本体的网络流量分类方法[J].计算机工程与设计,2016,37(01):31-36;54.

[20] 张淑权.黑客攻击电脑的几种常见手法以及防御技巧[J].计算机与网络,2014,40(17):59.

[21] 周启惠,邓祖强,邹萍,等.基于区块链的防护物联网设备DDoS攻击方法[J].应用科学学报,2019,37(02):213-223.

作者简介：

邹　伟(1985-)，男，硕士，讲师.研究领域：教育信息化，高校网络信息安全和智慧党建.

陈　恳(1978-)，男，硕士，讲师. 研究领域：教育信息化.

欧阳昭相(1978-)，男，本科，讲师. 研究领域：本体技术及信息化.