

网络安全态势多粒度分析的云方法

刘岱坪¹,董小华¹,张明威¹,陈佳²

(1. 武汉大学 计算机学院, 武汉 430079; 2. 武汉大学 软件工程国家重点实验室, 武汉 430079)

(liudaiping_123@126.com)

摘要:利用云理论的不确定性推理方法解决网络安全态势感知问题。在现有模型基础上,设计了一种基于知识库的多粒度网络安全态势感知模型。该模型具有动态更新知识库、多粒度态势展现以及态势可视化三大优势。另外,该模型实现了资产安全态势的扩展云推理算法,该算法能够处理连续值和离散值等多类参数,适用于多条件多规则推理,并且可以得到定性和定量两种安全态势结果。该模型具有较高的扩展性和实用性。最后,作者在武汉大学校园网中实验验证了模型的可行性及准确性。

关键词:网络安全;态势感知;云模型;粒度计算

中图分类号: TP393 **文献标志码:** A

Cloud method of multi-granularity network security situation analysis

LIU Dai-ping¹, DONG Xiao-hua¹, ZHANG Ming-wei¹, CHEN Jia²

(1. School of Computer Science and Technology, Wuhan University, Wuhan Hubei 430079 China;

2. State Key Laboratory of Software Engineering, Wuhan University, Wuhan Hubei 430079 China)

Abstract: This paper employs the uncertainty reasoning method in the Cloud Theory to solve the problem of network security perception. On the basis of the analysis of the existing models for security evaluation and their key technologies, the author proposed a situation awareness model using the knowledge base as well as the multi-granularity method. This model has three advantages over the conventional evaluation mechanisms: the dynamic updated knowledge base, multi-granularity method for security representation and visualization of the whole situation offers higher rate in accuracy and lower information redundancy. In addition, an extended cloud reasoning algorithm was designed and implemented to get both the qualitative and quantitative results of the monitored network and thus offers a more perceptible and scalable security situation. At last, experiments are done on the campus network of Wuhan University, and the results indicate that the feasibility and efficiency of our design.

Key words: network security; situation awareness; cloud model; granular computing

0 引言

为满足网络应用在安全方面更高的要求,起源于信息战战场态势感知技术的网络安全态势感知技术,为有效保护网络信息资产提供了一条崭新的思路。近年来,该方向已成为网络安全领域一个重要研究热点。

对于大型网络系统,网络管理人员一般不仅想知道一种粒度的态势情况。在不同层次、不同角度分析问题就是多粒度感知。面对海量安全态势数据,如果能够从一堆“杂乱”的数值得到网络系统空间和时间维度上多个粒度的安全态势,将更有利于网络安全状况的管理。

为得到多粒度网络安全态势,本文研究了网络安全态势感知领域的国内外经典模型,并结合本文的研究内容和作者的理解,给出一种基于知识库的多粒度网络安全态势感知模型。

1 网络安全态势感知

1.1 网络安全态势感知的概念

文献[1]首次提出网络态势感知的概念,即网络安全态势感知(Network Security Situation Awareness, NSSA)。但到目前,对网络态势感知还未能给出统一、全面的定义^[2]。作者

对于网络安全态势感知持以下观点:

1) 网络安全态势感知有一定时间和空间范围,是多粒度的。可以是实时的、最近 5 分钟的,或是当天的态势;可以是武汉大学,也可以是武汉市的网络安全态势。时间和空间的不同组合就得到了不同粒度的网络安全态势。

2) 网络安全态势是逻辑上分层次的。网络安全态势可以通过该网络范围内资产的安全态势体现,而资产的安全态势主要通过其各方面的运行情况体现。就好比人体是否健康,可以通过检查血液、测量体温等“运行状况”来判断。

3) 网络安全态势感知系统的最终目标是得到综合态势结果。系统不去处理一些细节问题,比如详细分析网络中到底正在发生着什么。

4) 网络安全态势感知系统可以综合态势预测功能,但不是必须的。

本文给出一个描述性定义:网络安全态势由一定时间和网络空间范围内的网络资产安全态势反映,网络安全态势感知是通过采集多种影响或衡量网络资产安全态势的数据源,对其加以分析得到网络资产安全态势,并通过粒度提升得到网络安全态势及变化趋势的技术。

收稿日期:2008-06-18;修回日期:2008-08-15。 基金项目:国家大学生创新训练计划(071048643)。

作者简介:刘岱坪(1987-),男,江苏扬州人,主要研究方向:网络安全、人工智能;董小华(1982-),女,湖北人,硕士,主要研究方向:网络安全;张明威(1985-),男,湖北黄石人,硕士研究生,主要研究方向:网络安全、信息安全、人工智能;陈佳(1982-),女,湖北武汉人,博士研究生,主要研究方向:数据挖掘。

1.2 NSSA 模型研究

网络安全态势感知是网络安全领域的一种新技术,其研究正处于起步阶段,因此系统方面多见于框架模型。其中有三种模型对该领域影响较大。文献[3]提出了 Endsley 态势感知层次模型,该模型强调与态势感知理解和量度有关因素的融合分析处理,将态势感知的实现分为态势要素提取、态势理解和态势预测三个层次,旨在更好地实现战场的态势感知。美国国防部联合指挥实验室的 JDL(Joint Director's of Laboratories)模型^[4]是一个数据融合模型,在军事领域有着广泛的应用,采用该结构能有效地辅助提高指挥员对战场态势的感知能力。在网络安全领域,针对入侵检测所构建的融合结构很多,但比较典型且被普遍接受的是 Tim Bass 所提出的融合结构^[5],该结构可用于网络安全态势感知和下一代入侵检测。

深入研究上述模型后,本文借鉴三种模型的思想,建立一种新的网络安全态势感知模型。

2 基于知识库多粒度网络安全态势感知模型

目前,网络体系结构日益复杂,新增资产带来新的安全隐患,安全边界不断变化,黑客攻击手段时刻更新等都需要 NSSA 系统不断制定新的安全规则及态势评估规则,从而得到更准确的态势结果。另外,如果能够提供更多层次、多角度的态势,则对于网络管理人员和决策者更有裨益。因此,本文建立了基于可更新知识库的多粒度网络安全态势感知模型,该模型的工作流程为:数据采集→数据融合→态势分析→态势预测/响应告警→态势可视化。

2.1 基于知识库的多粒度网络安全态势感知模型

本文借鉴 Endsley 的分层模式,引入 JDL 和 Tim Bass 的融合思想,设计了基于知识库的多粒度网络安全态势感知模型(Knowledge-based Multi-granularity Network Security Situation Awareness Model, KMG-NSSAM),如图1所示。

网络资产包括网络空间内的主机(包括主机上的软件系统)、网络设备(路由器、交换机等)以及网络流量等。在网络空间中提供资产安全状态数据的软件系统和硬件设备称为态势传感器。该模型应用多级信息融合与决策技术,通过多源

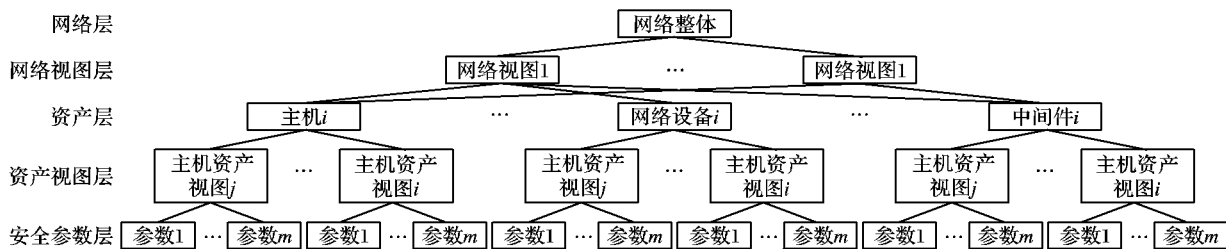


图1 基于知识库的多粒度网络安全态势感知模型(KMG-NSSAM)

2.2 网络资产粒度划分和态势参数选取

KMG-NSSAM 中,网络资产被抽取成一棵网络资产树,根节点为最高层,叶子节点为最低层。资产树与经典树概念不同的是,子节点可有多个父节点。把网络主机、软件系统(比如 DBMS、中间件等)和网络设备看作资产实体,构成资产层。每类资产的安全状态由选取的典型安全参数表示,这些参数构成安全参数层,即所有的叶子节点。每类资产的安全状态又由该资产的多个方面构成,各类资产每个方面的安全参数集在资产树中表示为资产视图层。另外,由多种资产的安全态势可以得到用户关心范围内的网络安全态势,即网络视图层。最后,资产树的根节点代表整个被监测网络。资产树如图2所示。

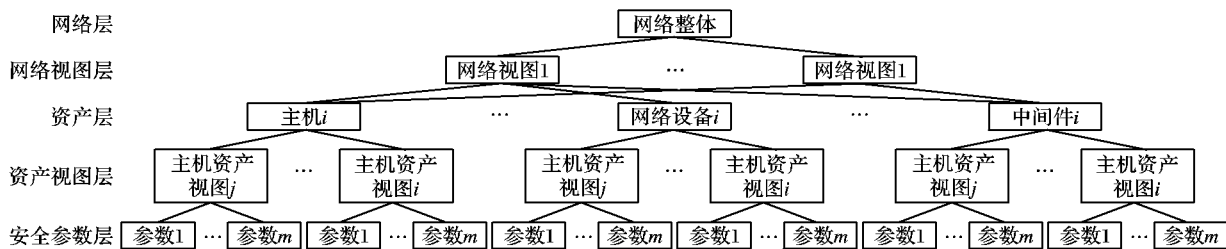


图2 网络资产树

网络资产树体现的不仅是一种层次关系,还体现了分析问题的粒度提升。上层概念(安全态势)比下层有所提升,表达的是一种更粗、更泛的概念。从较细的粒度世界泛化到较粗的粒度世界,是对信息或知识的抽象,可以使问题简化。用粗粒度观察和分析信息,可得到宏观的把握。这一点正是决策者所需要的。相反,如果细粒度观察和分析信息,则可发现纷繁复杂的表象,可得到更准确的差别。这一点可以使网络管理员更好地分析问题及危险所在。

作为资产树的叶子节点,安全参数是系统描述网络状态的原始数据,因此参数的选取直接决定了系统刻画网络安全态势的准确性。为达到高可信性的目标,KMG-NSSAM 选取参数时,综合参考了其他模型以及网络监管产品已成熟的指标体系。

根据参数代表资产态势方面的不同,可分为静态和动态参数;根据参数值性质的不同,可分为连续参数和离散参数。静态参数指资产的配置信息,如资产名称、主机磁盘容量等,这类参数大多在态势分析中作为参考信息。动态参数指资产的运行时状态信息,这类参数用于态势分析。静态参数大多采集频率较低,而动态参数大多采集频率很高。连续参数是指其值的变化是时间的连续函数。对于连续参数,若仅以某一时刻的值评估资产安全态势缺乏说服力,因为某时刻的参数值存在较大随机性。因此,需要将某个时间范围内的连续多个采样值进行粒度提升。可以通过逆向云发生器得到多个采样点构成的云图,并用该云图的期望值作为代表值。对于离散参数,若参数值发生改变,则用变化匹配相关态势规则;如果没有发生改变,则用采样值匹配态势规则。

3 网络安全态势的多粒度分析—云方法

推理规则的设定决定了推理的能力。由安全专家的知识汇聚成的规则库是一个好的方案。但安全专家的知识一般是自然语言表达,例如:如果在一段时间同一个端口的连接很多,那么服务拒绝攻击的可能性很大;如果一段时间内,登陆失败的次数很多,那么入侵的可能性较大。这些表达多是定性的,如“很多”、“较大”等。如何将自然语言的定性规则转化为计算机能够处理的定量规则呢?一个好的解决方法是借助云的不确定性推理。

3.1 由安全参数得到资产视图安全态势

KMG-NSSAM 使用云规则发生器进行不确定性推理。规则发生器可分为前件云和后件云两个部分。IF 部分是规则的条件,在这里用前件云实现,THEN 部分是规则的结果,用后件云实现。前件云的输入是待检值,输出是采样值激活某个规则前件的隶属度,该隶属度同时作为后件云的输入,输出是规则的结论。

一般来说,7 ± 2 个概念所对应的概念粒度比较适合人类认知的心理特点^[6],因此本文在不确定性推理中采用五规则生成器。不同的安全参数,表达的概念不一样,如 CPU 的占用率可以用“高”、“低”等概念来描述,数据传输速率却用“快”、“慢”等概念来描述。因此,本文不统一划分的概念名称,只统一划分的概念个数。KMG-NSSAM 借助专家知识规则库给定的各概念云的数字特征,得到规则发生器的前件云和后件云划分。后件云的概念划分可以统一为“很危险”、“危险”、“较危险”、“较正常”、“正常”五个概念。

现实中的推理问题,基本上属于多条件多规则推理问题。因此,本文借鉴云的多条件多规则推理方法。用于评价资产安全态势的参数分为连续参数和离散参数。对于连续参数,可以用云规则发生器来推理;而离散参数因其是确定值,所以云规则发生器的推理过程对其不适用。在此稍加改进。前件云发生器用来确定连续参数所属概念的隶属度,离散参数不经过前件云发生器,但仍作为规则的一个条件进入后件云推理。

基于上述思想,给出考虑连续和离散两种参数的扩展云推理算法。

算法 1 资产视图安全态势的扩展云推理算法

输入:待检参数向量 (X_1, X_2, \dots, X_n) 、规则前件(前件云)、规则后件(后件云)

输出:参数对应的资产视图安全态势(定性态势 C 、定量态势 Ex 、定量态势的隶属度)

步骤:

1) 将各连续参数代入相应的五概念前件云,得到其所属的一个或两个概念,如果有两个概念,则取激活隶属度较大者(极大判定法^[6]),最终得到隶属概念 C_p ;

2) 用 C_p 、离散参数值匹配规则库中的规则,激活的规则集为 $Rset$ (连续参数根据 3En 规则判断);

3) $Rset$ 中的每条规则 R_i :

① 将激活 R_i 的多个参数对应的激活隶属度求“软与”或“软或”,得到一个综合激活隶属度 SUB_i ;

② SUB_i 输入 R_i 对应结论的后件云,得到一个或两个云滴 $Drop_i$,加入云滴集 $Drop$ 。

4) 根据 3) 得到的 $Drop$,选取落在后件云最靠外侧的两点构造虚拟云(几何云)。得到几何云的期望 Ex ,则 Ex 在后件云中所属的概念即资产安全态势的定性概念 C , Ex 即定量态势, Ex 在后件云中对应的隶属度即该资产定量安全态势为 Ex 的概率。

算法说明:

1) 规则库中的规则一定要具有完备性。即任何输入条件都有相应的规则能够匹配。

2) “软与”或“软或”的处理可根据专家知识进行调节。

3) 算法时间复杂度 $O(N)$ 。本算法时间复杂度由前 3 步决定,第 1、2 步是由参数得到一个规则集 $Rset$,设有 N 个参数,其中 aN 个连续参数, $a \in [0, 1]$,则共需 $(aN + N)$ 次计算,输出 N 个规则的规则集;第 3 步,对规则中的每条规则运算,共需 N 次运算;因此,算法运行时间为 $(a + 2)N$,时间复杂度为 $O(N)$ 。

4) 在大量实验和实际运行基础上,我们发现对于大规模网络,事件报告延迟 5 ~ 10 min 仍能及时采取防护措施,且管理员能够容忍。这个延迟对本算法很足够了。

资产整体可以看作是一个特殊的资产视图,该视图包含所有该类资产的安全参数,因此得到资产安全态势与得到资产视图安全态势的算法是一样的。

3.2 由资产安全态势得到网络视图安全态势

一个网络视图由多个网络资产组成,且各资产的重要程度或对态势的影响程度不尽相同。因此,通过各资产的定量态势进行加权计算得到网络视图的定量态势。

设网络视图 S 由资产 A_i , $i = 1, 2, 3, \dots, n$ 组成,各资产对应的定性态势为 E_i ,权重为 W_i ,则该网络系统 S 的定量态势值可通过以下公式计算:

由式(1)知,网络视图安全态势值和资产安全态势值表达的是相当的概念,值越大,则态势越危险。

$$E = \frac{(\sum_{i=1}^n E_i \times W_i)}{\sum_{i=1}^n W_i} \quad (1)$$

仿照资产安全态势后件云的构建方法,通过态势知识库中网络视图安全态势前件云的三个数字特征,可以得到衡量网络视图安全态势的多概念云前件云。将网络视图的定量态势值输入网络视图安全态势前件云发生器,可以得到该值对于一个或两个概念的隶属度。重复输入多次,得到多个隶属度,根据极大判定法,取隶属度的最大值作为最终的隶属度,该最大隶属度对应的态势概念即为所求网络视图的定性态势。

网络整体可以看作一个特殊的网络视图,该视图包含所有资产对象,因此得到网络整体安全态势与得到网络视图安全态势的算法是一样的。

3.3 时间维度上的网络安全态势粒度提升

通过网络资产的运行状态分析网络安全态势,时间维度上的最小粒度取决于采集频率。比如 5 分钟采集一次,则可以得到 5 分钟内的近似网络安全态势。但是,决策人员关心的可能只是“今天”的态势。那么,就需要得到时间维度上更高粒度的态势。

根据数学中的研究方法,保持态势的另外一维“空间维度”不变,研究时间维度。以得到一定空间范围内资产“今天”的网络安全态势为例:

1) 对于某个资产。分两种情况求安全态势。第一种情况,“今天”内采集了多份该资产安全参数的样本;对样本点中的每个参数通过“无确定度的逆向云生成器”得到多个逆向云;将这些逆向云的期望输入 3.1 节的云规则发生器,得到该资产“今天”的安全态势。另一种情况,如果有比“今天”粒度低的该资产的安全态势,则可以用这些低粒度的态势通过逆向云或虚拟云中的几何云方法求得“今天”的安全态势。这样可以利用已有资源,而不用每次处理庞大的采样数据。

2) 对于某个网络视图。采用 3.2 节的方法,只是维度由“空间”变成了“时间”。

有了时间维度上的态势粒度提升,就可以画出以时间和定

量安全态势为横纵坐标的安全态势演变图,使安全态势演变可视化。这将更有利于网络管理人员对网络进行有效管理。

4 实验结果

为验证 KMG-NSSAM 的可行性及准确性,对其进行实验研究。实验内容分为两部分:网络安全态势感知多粒度分析模型的程序设计及实现,证明本文提出方法的可行性;采集武汉大学校园网网络状态数据进行分析,证明通过云方法得到的安全态势的准确性。

首先,通过专家知识配置模块配置云规则发生器前件云和后件云的三个数字特征,并得到相应云图,保存配置作为后续推理之用。以路由器接口丢包率为例,五概念划分如图4。

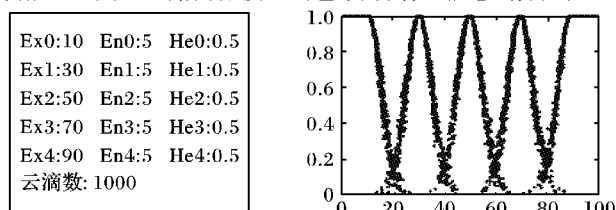


图4 路由器接口丢包率五概念划分

下面,可根据算法1实现资产视图安全态势的计算。以特殊资产视图路由器资产为例,输出云图和结论如图5所示。两条虚线是定位定性态势在五概念云中所属的概念。

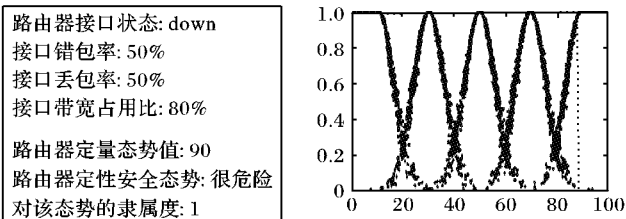


图5 路由器态势推理

接下来对网络视图安全态势的计算进行实验。其中各资产的权重由用户自定义。以武汉大学文理学院网络安全态势为例,输出云图和结论如图6所示。

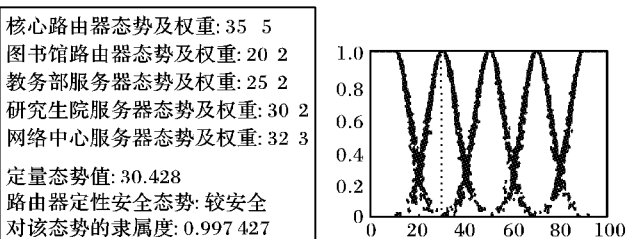


图6 武汉大学文理学院网络安全态势推理界面

最后,对时间维度安全态势的计算进行实验。要在时间维度上进行概念粒度的提升,需要借助逆向云发生器。用户输入云滴值(23.3 45.3 33.4 23.3 25.3 23.4 27.3 40.3 23.4 28.3 42.3 29.4 23.3 25.3 29.4),使用无确定度的逆向云发生器算法得到逆向云的云图及三个数字特征,如图7所示,其中逆向云期望:29.5,熵:7.213,超熵:1.8544。由一段时间内的态势可以得到安全态势演变图,如图8所示。有了安全态势演变图,网络管理员及领导就能更好地掌握网络资产的安全状况。

为验证本系统评估网络态势的准确性,笔者以网络主机进行实验,对比其染毒前后的态势计算结果。对系统监控网络中的某台运行情况较正常的主机进行参数采集。采集频率20s/次,采集时间:600s。将30次采样值送入逆向云发生器得到表1所示的结果。

用采样结果的期望作为送检值,推理结果:“主机安全态

势:较安全;定量态势值:30.28;隶属度:0.9991”。

对感染蠕虫病毒 Hiberium 后的该主机进行采样,采样频率和时间相同。送入逆向云发生器得到 CPU 使用率的逆向云期望是 42.7%,内存使用率的逆向云期望是 87.1%,将 42.7和 87.1 作为送检值,推理结果:“主机安全态势:危险;定量态势值:69.84;隶属度:0.9997”。

对比两次实验结果可知,系统得到的安全态势具有较高的准确性。

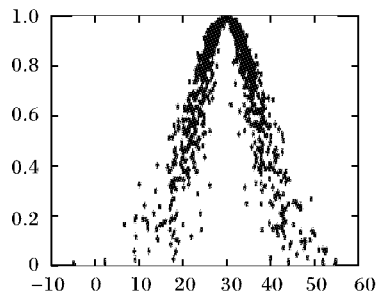


图7 逆向云发生器

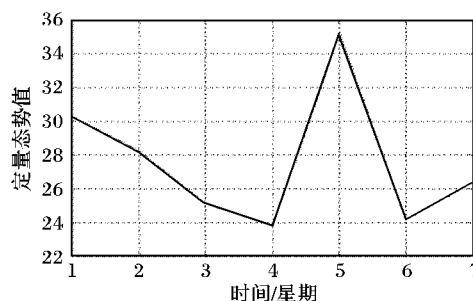


图8 教务部服务器安全态势演变图

表1 较正常主机采样数据结果

逆向云发生器	CPU 使用率/%	内存使用率/%
Ex	12.8	75.4
En	15.2	8.0
He	8.3	22.9

5 结语

本文提出一种基于知识库的多粒度网络安全态势感知模型。该模型提出将云理论用于网络安全态势感知的研究思路,给出了通过云方法得到多粒度网络安全态势的详细方法,从空间和时间两个维度出发,使得态势感知粒度由细到粗,步步提升;该模型能够得到定性和定量两种态势结果,并验证了该模型的可行性和准确性。

参考文献:

- [1] BASS T, GRUBER D. A glimpse into the future of ID. Special Issue Intrusion Detection, The USENIX Association Magazine [EB/OL]. (2005-09) [2008-04-23]. <http://www.usenix.org/publications/login/1999-9/features/future.html>.
- [2] 王慧强, 赖积保, 朱亮, 等. 网络态势感知系统研究综述[J]. 计算机科学, 2006, 33(10): 5-10.
- [3] ENDSLEY M R. Toward a theory of situation awareness in dynamic systems [J]. Human Factors Journal, 1995, 37(1): 32-64.
- [4] STEINBURG A N, BOWMAN C L, WHITE F E. Revisions to the JDL data fusion model [C]// Proceeding of Third NATO/IRIS Conference. Quebec, Canada: [s. n.], 1998: 430-441.
- [5] BASS T. Intrusion systems and multisensor data fusion: Creating cyberspace situational awareness [J]. Communications of the ACM, 2000, 43(4): 99-105.
- [6] 李德毅, 杜鹞. 不确定性人工智能[M]. 北京: 国防工业出版社, 2004: 90-248.