第 46 卷 第 1 期 计 算 机 工 程 2020 年 1 月
Vol. 46 No. 1 Computer Engineering January 2020

·网络空间安全· 文章编号：1000-3428(2020)01-0136-08 文献标志码：A 中图分类号：TP309.2

面向攻击识别的威胁情报画像分析

杨沛安[1,2]，刘宝旭[1,3]，杜翔宇[1,3]

（1. 中国科学院大学，北京 100049；2. 中国科学院高能物理研究所，北京 100049；3. 中国科学院信息工程研究所，北京 100093）

摘 要： 新型网络攻击向高隐蔽性、高持久性和高扩散性的方向发展，导致攻击识别与检测难度骤增。为提高网络攻击识别的效率与准确性，提出一种面向攻击识别的威胁情报画像分析方法。建立攻击画像数据表达规范，基于 Killchain 模型和攻击原理，构建威胁属性状态转移关系的挖掘模型，提取属性状态转移序列。在此基础上，利用有色 Petri 网攻击图在因果关系处理和表达上的优势进行基于威胁属性的关联，并将相关要素与属性转换为要素原子图。通过要素融合算法对要素原子图进行融合，实现威胁情报画像分析。实际攻击事件分析过程中的应用结果表明，该方法能提高网络攻击识别准确度，并缩短攻击识别响应周期。

关键词： 攻击识别；威胁情报；情报分析；攻击图；关联分析

开放科学（资源服务）标志码（OSID）：

中文引用格式： 杨沛安，刘宝旭，杜翔宇. 面向攻击识别的威胁情报画像分析[J]. 计算机工程，2020，46(1)：136-143.

英文引用格式： YANG Peian，LIU Baoxu，DU Xiangyu. Portrait analysis of threat intelligence for attack recognition[J]. Computer Engineering，2020，46(1)：136-143.

Portrait Analysis of Threat Intelligence for Attack Recognition

YANG Peian[1,2]，LIU Baoxu[1,3]，DU Xiangyu[1,3]

（1. University of Chinese Academy of Sciences，Beijing 100049，China；2. Institute of High Energy Physics，Chinese Academy of Sciences，Beijing 100049，China；3. Institute of Information Engineering，Chinese Academy of Sciences，Beijing 100093，China）

【Abstract】 New network attacks are getting more covert and persistent with a high proliferation，resulting in a sudden increase in the difficulty of attack recognition and detection. To improve the efficiency and accuracy of network attack recognition，this paper proposes a portrait analysis method of threat intelligence for attack recognition. Based on the Killchain model and the principles of attack process，this method builds data representation standards for attack graph，so as to build a mining model of transition relationships between threat attribute states. Then the attribute state transition sequence is extracted. On this basis，the this method takes advantages of the Colored Petri Net（CPN）attack graph in causality processing and expression to associate threat attributes，and converts related elements and attributes to an Element Atomic Graph（EAG）. The EAG is fused using the element fusion algorithm to implement portrait analysis of threat intelligence. Application results in actual attack analysis demonstrate that the proposed method can improve accuracy of network attack recognition，and shorten the response period of attack recognition.

【Key words】 attack recognition；threat intelligence；intelligence analysis；attack graph；association analysis

DOI： 10.19678/j.issn.1000-3428.0051157

0 概述

随着威胁情报的发展，安全分析领域开始使用威胁情报数据对各类网络攻击进行画像分析，包括行业画像、用户画像、资产画像、威胁画像以及黑客画像等。目前该领域的研究和应用还处于起步阶段，研究内容分散，缺乏系统性。一方面，因为用于网络安全分析的威胁情报数据在数据结构、处理方法上与传统互联网用户数据有所不同，现有数据分析方法不能完全适用。另一方面，业界对威胁情报数据在攻击画像分析中的应用场景、分析需求还不明确，没有统一的威胁情报数据表达规范，不能为画

基金项目： 北京市科委基金（Z161100002616032）。

作者简介： 杨沛安（1988—），男，博士研究生，主研方向为网络信息安全；刘宝旭，研究员、博士生导师；杜翔宇，博士研究生。

收稿日期： 2018-11-10 **修回日期：** 2018-12-15 **E-mail：** 478649459@qq.com

像分析提供数据基础。

关联关系丰富是威胁情报最重要的特征,威胁情报的分析人员基于在网络安全分析领域多年积累的分析经验,通过多角度网络安全分析需要的知识进行关联融合实现了威胁情报的重要价值。所以关联分析在威胁情报分析领域具有十分重要的地位。目前,网络安全分析人员主要针对特定的安全分析需求,使用关联分析对低级安全数据进行融合,产生威胁情报。但是随着网络安全态势的急剧恶化,针对越来越多的攻击难以进行准确高效识别的问题,攻击分析人员对质量更高、适用性更强的威胁情报及更准确、全面、容易使用的攻击识别建模方法的需求越来越大。所以研究合适的关联分析方法对不同类别的威胁情报进行关联,为攻击识别提供包含更全面准确攻击特征的威胁情报具有重要的研究意义。

本文利用威胁情报在数据准确性和关联性上的优势,将威胁情报自动化分析得到的威胁要素作为画像轮廓,以威胁属性为画像内容进行画像分析,构建攻击识别模型,并从威胁画像数据表达规范、属性转移关系挖掘模型、属性关联分析和要素关联融合方面对面向攻击识别的威胁情报画像分析方法进行描述。

1　相关研究

1.1　画像分析技术

攻击画像分析是基于攻击画像模型和攻击画像数据表达规范,利用攻击要素数据进行攻击图绘制,形成基于攻击要素的攻击图攻击画像,提高识别准确率。画像分析技术早期出现于刑侦行业,用于对人物进行描述和分析,包括文献[1]利用多文档摘要技术实现人物传记的提取,文献[2]利用分类思想,实现多文档人物传记摘要系统,文献[3]提出“元事件”的概念,并将其应用到人物信息抽取领域。这一阶段的研究重点集中在人物实体属性的识别中,缺少对实体关系挖掘的研究。

随着信息技术的发展,开始出现面向抽象人物和人物群体的画像分析研究,如面向网络罪犯分析的计算机取证技术,用于将罪犯各项特征进行挖掘和关联,形成罪犯画像。当大数据技术逐渐成熟之后,开始出现基于大数据的用户画像分析,并通过这项技术对用户本身的各类社会属性以及用户上网偏好、购买力等信息进行分析,再结合用户分析框架将这些信息进行有机融合,实现对用户特征和轮廓的勾画。文献[4]面向移动互联网中的用户,对其行为和偏好进行研究和分类,并在用户行为分析方面展开重点讨论。文献[5]提出一种面向客户的安卓商品推荐软件,有效降低了客户进入商场后从海量产品中了解产品信息完成商品挑选的时间成本。文献[6]给出一种基于特征值的用户行为分析方法,并结合 SVM 模型对搜索与排序算法进行优化。文献[7]提出一种对互联网用户的需求和使用偏好等进行挖掘与分析的算法,并对包括个性推荐、定点营销与广告投放等内容进行深入介绍。该阶段的研究主要是基于用户的网络行为和属性,对其关系进行挖掘,分析用户行为模式特征,但还局限在单一的“人物行为-模式”分析中,缺少针对分析对象的更全面、深层次的综合分析。

1.2　基于图的关联分析方法

图论被广泛应用于计算机领域,计算和分析各类数据、状态和模式等。早期攻击图在网络安全分析领域的应用主要是利用攻击图的网络安全性和脆弱性进行分析与评估,后来出现应用攻击图进行漏洞分析和告警关联分析,接着出现基于攻击图的异常行为发现和网络攻击检测的相关研究。目前基于攻击图的关联分析在网络安全分析中的应用主要集中在风险评估与脆弱性分析、漏洞分析与告警关联、异常发现与攻击检测方面。

1.2.1　风险评估与脆弱性分析

该方面的研究从早期单一基于攻击图的脆弱性发现和风险评估,向脆弱性与安全评估相结合的方法发展,以实现对网络安全状态和相关安全威胁更准确有效的分析与发现。文献[8]针对网络弱点关联分析提出渗透图的概念,并基于渗透图设计一种网络风险评估模型。文献[9]提出一种基于攻击图的网络安全分析方法,该方法解决了由路径循环导致的攻击路径不可达及威胁概率计算错误问题,提高了攻击路径最大可达概率算法在复杂网络脆弱性分析中的适用性。文献[10]基于对大量脆弱性利用行为的研究,设计一种警报关联图用于网络脆弱性分析,并依此从属性分析角度给出一种网络脆弱性分析方法。文献[11]基于网络风险评估中,难以对全面安全性与局部脆弱性进行有效综合评估的问题,将层次化分析方法引入风险评估中。通过分析原子攻击和攻击证据的关联性得到攻击因果关系,以此因果关系构建贝叶斯攻击图,再通过对攻击图中脆弱点设计威胁度划分标准,实现对脆弱点严重程度和系统整体安全性的统一分析。文献[12]针对漏洞分析中对脆弱性进行有效量化困难的问题,给出一种基于攻击图的安全脆弱性量化评估方法。先通过脆弱性分析构建贝叶斯网络攻击图,然后对图中节点进行基于概率值的可利用性分析,最后结合漏洞评分系统和贝叶斯网络攻击图实现对网络系统脆弱性的量化评估。文献[13]针对当前系统安全管理中无法对系统整体态势进行评估的同时,对具体脆弱性和隐患进行识别和与发现这一问题,利用颜色 Petri 网(Colored Petri Net,CPN)构建系统脆弱性攻击图进行脆弱性分析和发现,取得了较好的效果。

1.2.2　漏洞分析与告警关联

该方面的研究在早期利用图的相似性分析方法

进行告警关联融合的基础上,逐渐向利用对漏洞、报警等局部显性威胁信息的分析结果,对网络整体安全性进行分析评估和对隐形威胁要素进行识别的方向发展。文献[14]针对 IDS 系统无法对具有复合模式的攻击进行有效识别和预测的问题,给出了一种告警预测图。基于网络脆弱性分析结果构建攻击图,并利用告警信息对攻击图进行优化。设计告警关联预测算法,利用该攻击图将脆弱性和告警信息进行融合,对可能的攻击位置进行预测。文献[15]通过多漏洞组合利用对攻击者可能采取的攻击路径进行分析。通过漏洞检测器确定本地漏洞信息,然后在对这些信息进行分析的基础上,基于权限提升的攻击/漏洞关联分析方法,对漏洞利用路径进行挖掘,并通过自动化的方法生成漏洞利用攻击图。文献[16]针对以主机为中心的漏洞分析方法对网络链路本身不确定性分析欠缺考虑的问题,提出一种基于不确定图的漏洞分析方法,采用不确定度准确地描述网络状态,并以此得到漏洞的最佳利用链路,实现对不确定网络中漏洞利用方法的有效分析。

1.2.3 异常发现与攻击检测

该方面的研究主要是从攻击性、脆弱性、漏洞利用率等角度,通过最优路径、最大脆弱性节点的发现对系统安全性进行研究,实现对系统安全状态的分析。文献[17]通过从告警数据中挖掘告警属性间的规律,构建基于扩展有向图的复合攻击模型,从而对攻击行为之间的逻辑关系进行表达。基于该模型通过向后匹配和缺项匹配的方式对新告警与已知告警进行关联,确定新告警属于已知攻击及其处于已知攻击的何种阶段。文献[18]面向入侵意图检测和漏洞发现困难的问题,通过对告警进行分析设计一种三层攻击图结构,并结合入侵意图的概率分析确定攻击意图概率图。然后通过分析图中可能的关键点实现对脆弱性较高主机的发现,进而提高分析人员的分析和检测效率。文献[19]定义了一种 SAGML 语言,基于该语言对攻击状态、行为和关系进行描述。然后对攻击图的状态和行为链结构进行深入研究,提出基于 XML 的攻击图绘制与分析方法。最后建立适合攻击图的检索和匹配方法,提高攻击图的利用率。文献[20]面对 APT 攻击不易识别、持续时间长等特点,将基于图的评估方法引入 APT 检测过程中。通过对 APT 攻击在行为、过程方面的特征进行分析,构建网络风险属性攻击图。然后对系统中各节点的行为和联通脆弱性进行评估,再结合两方面评估方法实现系统脆弱性的有效分析。文献[21]针对攻击过程中的攻击识别与发现问题,建立基于 Petri 网的攻击成本图,然后对其可承受的最大攻击力度进行分析和预测,并依此选择最优攻击路径。

2 威胁情报画像分析

为得到更准确全面的攻击识别模型,本文提出一种基于威胁情报的网络攻击画像分析方法。以威胁情报中的威胁要素作为画像骨架,以威胁属性作为画像要素。根据从威胁情报中提取的威胁属性与威胁情报库中的相关属性进行关联,实现“属性-属性”的关联,并利用 CPN 网在因果关系表达与分析上的优势,将要素与属性转换为要素原子图。而后通过融合要素原子图,实现“要素-要素”的关联,完成画像分析,形成更全面准确的威胁要素和属性的攻击识别模型,即网络攻击威胁情报画像。画像分析流程如图 1 所示。

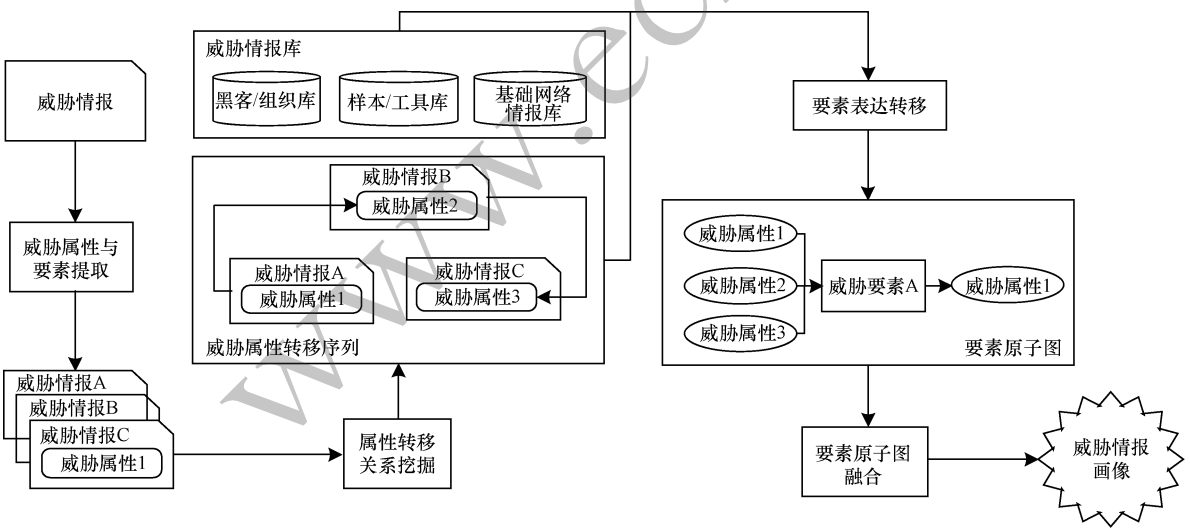

图 1 威胁情报画像分析流程

Fig. 1 Portrait analysis process of threat intelligence

2.1 威胁画像数据表达规范

威胁情报在安全分析中的优势在于其高度结构化的数据框架和具有丰富关联关系的数据内容。但这些特性也使得现有数据表达规范和通信传输协议在威胁情报数据表达的有效性与完整性、传输的准确性与安全性等方面存在问题,无法有效进行威胁

情报的表达和传输。本文旨在设计一种满足攻击画像分析的情报数据表达规范。

目前攻击画像分析通常只以攻击的某一指定要素为重点对象展开画像分析,如仅针对工具函数序列、网络行为特征等。这直接导致当面对多步的复杂攻击时,无法进行有效的检测和识别。根据相关研究可知,利用威胁情报作为攻击识别分析的数据进行攻击画像分析,目的是对攻击主体和主体行为进行挖掘和关联。通过从多要素、富关联的数据中提取攻击者、攻击目标、攻击工具等关键特征及相关属性,再对特征和属性进行关联关系挖掘,形成“特征-属性-关系”结构的攻击识别知识。

为满足以上对画像分析数据结构和内容上的要求,本文参考 STIX 情报表达规范的 12 个要素和 Cybox 规范的 88 个属性,对攻击画像数据表达模型进行设计。模型中包含威胁要素 9 类,威胁属性 56 种,具体如下:

1)攻击工具 Tool,使用实施攻击过程的合法工具,如远控工具和网络扫描工具等。威胁属性:id 表示代码,description 表示描述,kill_chain_phase 表示位于杀伤链阶段,tool_version 表示工具版本,behavior 表示工具功能。

2)恶意软件 Malware,攻击过程中植入到目标系统的恶意代码或者恶意软件。威胁属性:id 表示代码,description 表示描述,kill_chain_phase 表示位于杀伤链阶段,behavior 表示软件效果,release 表示释放内容,root 表示权限,reachable 表示目标可达性,exploit 表示利用脆弱性。

3)攻击模式 Attack_Pattern,用于表达攻击者尝试攻击的方法,用于攻击分类、生成攻击的固定模式及对攻击实施过程的详细描述。威胁属性:id 表示代码,description 表示描述,kill_chain_phase 表示位于杀伤链阶段。

4)观测线索 Observation,从系统或者网络中可观测到的数据,如日志信息或网络流量。威胁属性:first_observed 表示首次观测时间,last_observed 表示最后观测时间,number_observed 表示观测数量。

5)攻击指标 Indicator,用于表达可疑或者恶意的网络空间安全行为的指标。威胁属性:id 表示代码,description 表示描述,pattern 表示模式,pattern_lang 表示模式定义语言,pattern_lang_version 表示模式定义版本,valid_from 表示有效起始时间,valid_from_Reliability 表示有效起始可信度,valid_until 表示有效截止时间,valid_until_percison 表示有效截止精度,kill_chain_phase 表示位于杀伤链阶段。

6)攻击者 Threat_Actor,实施有恶意意图攻击的个体、团体和组织。威胁属性:id 表示代码,description 表示描述,aka 表示代号,name 表示姓名,roles 表示角色,goals 表示目标偏好,sophistication 表示复杂度,resource_level 表示资源级别,primary_motivation 表示初始动机,secondary_motivations 表示第二动机,private 表示偏好。

7)识别信息 Identity,用于代表并区分的个体、组织或者团体。威胁属性:id 表示代码,description 表示描述,identity_class 表示识别等级,group 表示小组,activitiyRegion 表示活动区域,country 表示国籍,contact_Info 表示联系信息。

8)入侵集合 Intrusion_Set,由一个组织精心组织的一系列行为和资源集合。威胁属性:id 表示代码,description 表示描述,aliases 表示身份类别,first_noticed 表示首次发现时间,first_noticed_Reliability 表示可信度,goals 表示目标偏好,resources_level 表示资源级别,primary_motivation 表示初始动机,region 表示地区,country 表示国籍。

9)组织战役 Campaign,描述一系列的针对特殊目标集合的恶意行动或攻击的敌对行为,经常被定义成目标或入侵集合的一部分。威胁属性:id 表示代码,description 表示描述,aliases 表示别名,first_seen 表示首次发现时间,first_seen_precision 表示首次发现精确度,objective 表示目的。

2.2　属性转移关系挖掘模型

根据 Killchain 模型中对 7 个阶段的描述,网络攻击行为由一系列分阶段的依据因果顺序发生的子攻击行为共同组成。而这一由多个依序发生子攻击组成网络攻击的特点反映其威胁情报中,多个威胁要素依据因果关系顺序出现。所以本文可以参考该思路设计威胁情报中威胁属性关系挖掘模型。假设该模型是一个属性转移图,该图为有向图,图中的顶点是威胁属性,边则表示属性的转移关系(主要是因果关系)。图中各点间的路径表示各属性间的因果关系,包括直接相关和间接相关,直接相关的属性间存在唯一路径,间接相关属性间存在多条路径。

定义 1　威胁属性集合 A,其中 p、q 表示具有关联关系的两个威胁属性。

定义 2　属性转移关系集合 T,表示威胁情报中所有属性转移关联的集合。

定义 3　属性转移关系 t,表示两点间路径,即两属性间的转移关系,由一个五元组表示 $<f,t,c,p,E>$,其中,f 表示该转移关系起点属性,t 表示该转移关系终点属性,c 表示该转移关系的置信度,p 表示关联转移关系属性,若该转移是直接转移,则用 0 表示,否则用 1 表示,E 表示该属性所属的威胁要素,即该属性包含在哪个要素中。

2.3　威胁属性关联

从威胁情报中得到的威胁属性可能只是所属威胁要素包含的众多威胁属性之一,需要通过以该属性为关联特征,与威胁情报库中的威胁要素进行关联。本文中借助 CPN 在因果关系表达与处理上的

优势,结合威胁属性状态转移序列对威胁属性进行关联,并实现表达转换,构建要素原子图。威胁属性关联过程如图 2 所示。

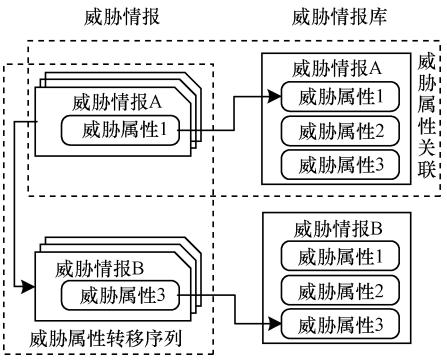

图 2 威胁属性关联过程

Fig. 2 Association process of threat attributes

在 CPN 中每一个原子攻击包括 3 类要素:原子攻击发生的条件,原子攻击本身和原子攻击产生的影响。攻击条件和攻击本身是从属关系,攻击本身和攻击影响是因果关系,以此来实现攻击条件向攻击影响的因果关系转移,如图 3 所示。

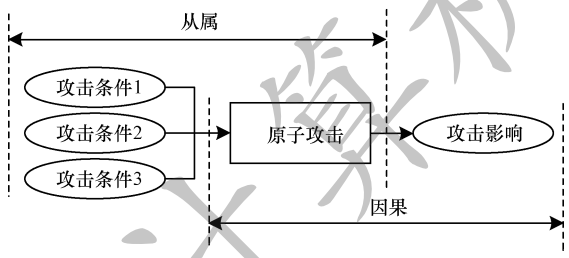

图 3 原子攻击中各要素的关系

Fig. 3 Relationship of elements in an atomic attack

在威胁情报中,威胁要素与威胁属性属于从属关系,而威胁要素与另一要素的属性为因果关系。例如攻击者和攻击工具分别是威胁要素,名字和代码分别是它们的属性之一,两者为从属关系。而攻击者因为要发动攻击,所以使用攻击工具,此为因果关系,得到的攻击者属性和攻击工具的关系如图 4 所示。

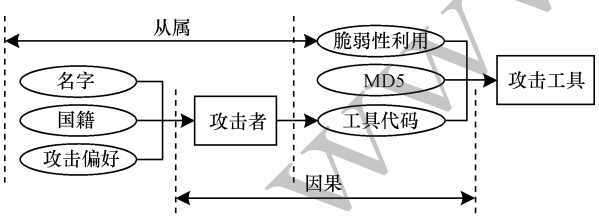

图 4 威胁情报要素与属性的关系

Fig. 4 Relationship between elements and attributes of threat intelligence

在图 4 中,椭圆形节点用于表示威胁要素从属属性和因果属性,等同于 CPN 结构中的库所;矩形节点用于表示威胁要素,等同于 CPN 结构中的变迁,左边的椭圆节点为要素原子图的输入库所 E_0,右边的椭圆节点要素原子图的输出库所 E_d。E_0-t 由威胁情报库中各类要素情报提供,t-E_d 由属性转移序列得到。

下文基于 CPN 给出要素原子图的形式化定义。要素原子图(Element Atom Grapic, EAG)是一个 CPN 结构,记为 $E_{EAG}=\langle E_{Ao},t,E_{Ad}\rangle$,其中,$E_{EAG}$ 为要素原子图的输入属性集合,t 为变迁,表示一个威胁要素,E_{Ao} 为威胁要素原子图的输出属性集合,E_{Ad} 表示与该威胁要素存在因果关系的另一个威胁要素的要素属性(称为目标属性)。需要说明的是,在通常应用 CPN 进行网络攻击关联分析时,确定攻击的输入库最困难,即“攻击发生的条件”,而在威胁情报数据中,由于威胁情报本来就经过分析加工后得到,因此直接给出了明确的攻击条件,可以大幅提高分析效率,这是威胁情报的优势所在。

2.4 基于图关联的威胁要素融合

通过“属性-属性”的关联,实现了对威胁属性的扩展和表达,但是攻击识别模型还需包含各个要素与属性间的关联关系,这就需要通过要素原子图的关联融合实现。下文对要素融合图进行定义。

威胁要素融合图(Element Atom Grapic, EFG),记为 $E_{EFG}=\langle P_{Eo}\cup P_{Ed},T_{Eo}\cup T_{Ed},E\rangle$,其中,$P_{Eo}$ 代表原始库所集合,包含所有要素原子图中的输入库所,即从属属性的集合,P_{Ed} 代表目标库所集合,包含所有要素原子图中输出库所,即目标属性的集合,T_{Eo} 代表唯一变迁集合,表示所有不包含因果关系的威胁要素的集合,这些威胁要素的属性都包含在 P_{Ed} 中,T_{Ed} 代表关联变迁集合,代表所有处于因果关系中的威胁要素的集合,这些威胁要素至少有一个属性包含在 P_{Ed} 中。因此 T_{Ed} 中各威胁要素的出现需要依赖于 T_{Eo} 中各从属关系的威胁要素,E 为 EFG 中所有要素、属性之间有向边的集合。

该要素融合图具有以下约束条件:

1)要素融合图中的有向连接只能用于威胁要素与属性相连,即 $E\subset((P_{Eo}\cup P_{Ed})\times(T_{Eo}\cup T_{Ed}))\cup((T_{Eo}\cup T_{Ed})\times(P_{Eo}\cup P_{Ed}))$。

2)初始要素集合 T_{Ed} 中变迁 t_o,$p_{pre}(t_o)$ 表示该要素包含的所有属性的集合,$p_{post}(t_o)$ 表示与要素有因果关系的所有属性的集合,即 $(p_{pre}(t)\subseteq P_{Eo})\wedge(p_{post}(t)\subseteq P_{Ed})$。

3)最终要素集合 T_{Ed} 中变迁 t_o,$p_{pre}(t_d)$ 表示该要素包含的所有属性的集合,$p_{post}(t_d)$ 表示与要素有因果关系的所有属性的集合,即 $(\exists p\in p_{pre}(t):p\in P_{Ed})\wedge(p_{post}(t)\subseteq P_{Ed})$。

在得到要素融合图的定义后,通过要素融合算法对要素原子图进行融合,融合过程如下:

1)将 EFG 初始设置为空,调用 Build 过程创建一个表示存在 AT 类要素中 name 类属性的库所。

2)以新创建的库所创建初始 EFG。

3)循环添加 EAG 到 EFG 中,迭代 EFG:

(1)对要素转移序列 ElementTransfer 中的每个

EAG 进行分解,确定其包含的属性 t_l。

(2)根据要素的键值对(r_k,r_n),在威胁情报库 TIDB 中定位到该要素的位置,然后返回该要素所包含的所有属性内容 T_{Iret}。

(3)根据 t_l和 T_{Iret}作为输入,构造相应的 EAG。

(4)将新构造的 EAG 追加到 EFG 中。

(5)当追加完成 ElementTransfer 中所有要素后,结束循环。

4)返回最终状态的 EFG。

3　实例分析

下文结合实际案例对基于属性的威胁要素融合方法进行介绍。从一份“海莲花”APT 分析报告中得到了包括“OceanLotus”“KVDropper”“SKI”“HUAWEI”“鱼叉攻击”“CVE-2010-20318”“网络通讯中断”“杀毒软件”“通过水坑攻击……后进行鱼叉攻击……”等在内的 8 类威胁要素和 10 个威胁属性。这些威胁属性构成了原始威胁情报,如图 5 所示。原始威胁情报知识对各攻击特征(威胁属性)进行简单罗列,不包含任何关系(从属或因果)。

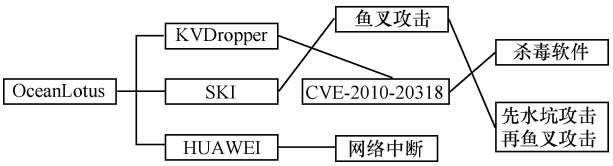

图 5　原始威胁情报

Fig. 5　Raw threat intelligence

通过对威胁属性的转移关系进行挖掘,得到威胁属性到威胁要素的转移关系序列,如“攻击源(OceanLotus)-->攻击工具(KVDropper)”“攻击工具(SKI)-->恶意代码(CVE-2010-20318)”。属性状态转移序列如图 6 所示。本文通过威胁属性关联方法,得到包含威胁要素与威胁属性及其关联关系的要素原子图如图 7 所示。

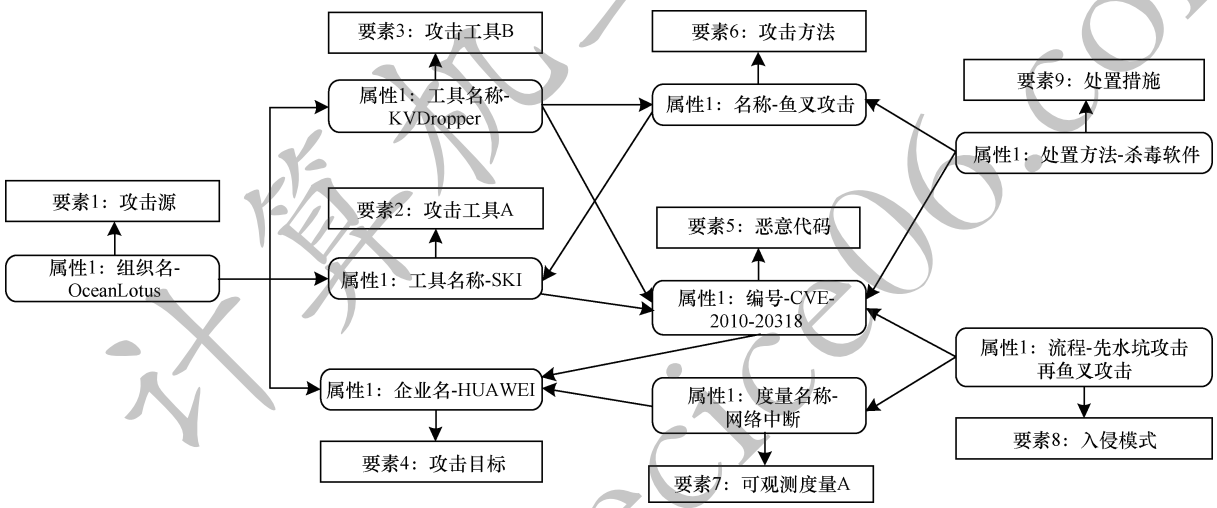

图 6　威胁要素关联关系挖掘后的威胁情报

Fig. 6　Threat intelligence with mined association relationships between threat elements

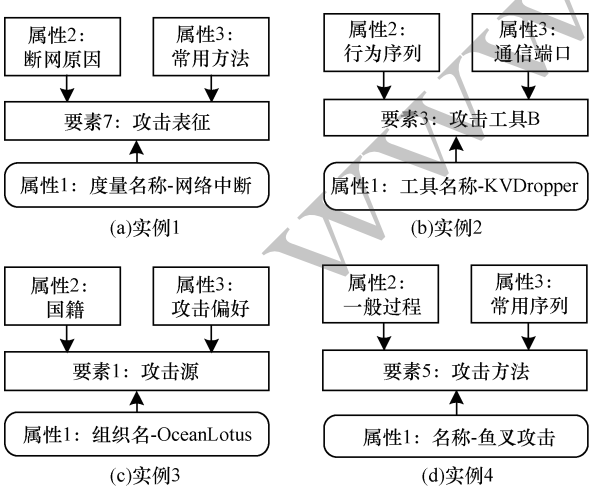

图 7　要素原子图

Fig. 7　Element atomic graph

以识别出的威胁要素与属性为画像轮廓,通过“属性-属性”关联和“要素-要素”融合,形成针对本次攻击的威胁情报画像如图 8 所示。可以发现,通过使用该画像分析方法对原始情报进行分析和处理,不仅实现了对原始情报中属性和要素的准确和全面表达,而且对属性隐藏关系进行挖掘,实现了对原始情报的丰富和补充,形成了完整、准确的攻击识别模型。相比原始情报,威胁情报画像分析方法得到的攻击画像对攻击的描述更加完整准确。以威胁情报画像作为相关攻击事件识别和跟踪的分析模型,可以有效提高分析准确率和效率。在后续针对 APT 攻击的追踪分析中,该画像将帮助分析人员定位多个同源攻击事件。

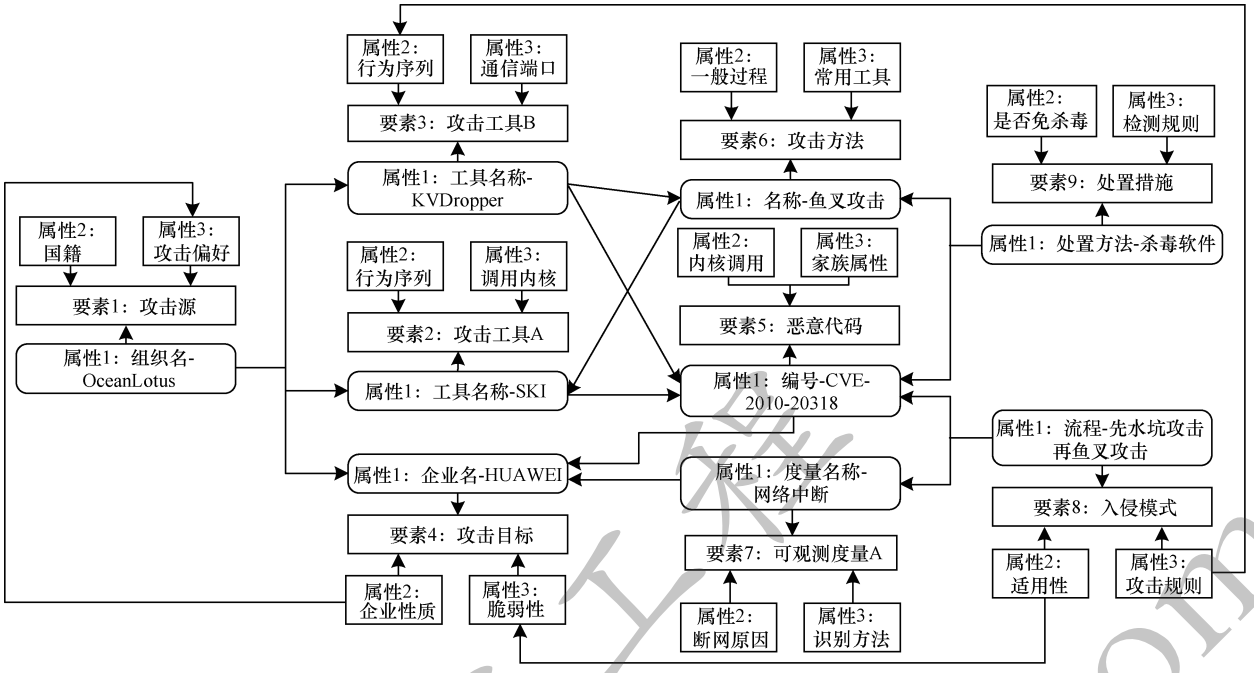

图 8 要素融合关联后的威胁情报画像

Fig. 8 Threat intelligence portrait after elements are fused and associated

4 结束语

为得到更准确全面的攻击识别模型，本文提出威胁情报画像分析方法。以威胁情报中提取的威胁属性转移序列为画像骨架，将威胁情报库中存储的相关要素及属性关联，实现基于属性的威胁情报融合，形成更丰富和完善的攻击特征，从而完成威胁情报画像的绘制。通过实际分析案例验证了本文画像分析方法的正确性和有效性。由于基于专家知识与分析经验得到的威胁要素与属性因果关系，可能会对画像分析结构产生影响，因此后续将在关联属性挖掘阶段引入置信度机制，并优化威胁属性关系转移序列的挖掘方法，进一步提高网络攻击识别准确度。

参考文献

[1] SCHIFFMAN B, MANI I, CONCEPCION K. Producing biographical summaries: combining linguistic knowledge with corpus statistics[C]//Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics. Philadelphia, USA: Association for Computational Linguistics, 2002: 450-457.

[2] ZHOU L, TICREA M, HOVY E. Multi-document biography summarization [EB/OL]. [2018-10-15]. https://www.researchgate.net/publication/1863946_Multi-document_Biography_Summarization.

[3] FILATOVA E, PRAGER J. Tell me what you do and I'll tell you what you are: learning occupation-related activities for biographies [EB/OL]. [2018-10-15]. http://www.doc88.com/p-9039789072425.html.

[4] ZHAO Zhijian. Building mobile user behavior analysis model[J]. China Science and Technology Information, 2014(1): 100-101. (in Chinese)
赵之健. 构建移动用户行为分析模型[J]. 中国科技信息, 2014(1): 100-101.

[5] ANIZA N, AMRAN N, ZAINI N, et al. User profile based product recommendation on Android platform [C]// Proceedings of the 5th International Conference on Intelligent and Advanced Systems. Washington D. C., USA: IEEE Press, 2014: 1-7.

[6] JIANG Zongli, ZHANG Ting. Optimizing of local search ranking algorithm based on user behaviors analysis[J]. Computer Technology and Development, 2014, 24(2): 15-18, 24. (in Chinese)
蒋宗礼, 张婷. 基于用户行为分析的本地搜索排序算法优化[J]. 计算机技术与发展, 2014, 24(2): 15-18, 24.

[7] HU Zhonggang, JIANG Minjuan. Information value mining based on user behavior analysis[J]. Jiangsu Communication, 2014, 30(1): 66-68. (in Chinese)
胡仲刚, 蒋敏娟. 基于客户移动互联网行为的信息价值挖掘应用[J]. 江苏通信, 2014, 30(1): 66-68.

[8] SU Jibin, XIAO Zongshui, XIAO Yingjie. Analysis and implementation of risk assessment based on penetration chart[C]//Proceedings of CACIS 2008. Hefei: China Instrument and Meter Association, 2008: 1160-1166. (in Chinese)
苏继斌, 肖宗水, 肖迎杰. 一种基于渗透图的风险评估分析与实现[C]//全国第 19 届计算机技术与应用学术会议论文集. 合肥: 中国仪器仪表学会, 2008: 1160-1166.

[9] YE Yun, XU Xishan, JIA Yan, et al. An attack graph-based probabilistic computing approach of network security [J]. Chinese Journal of Computers, 2010, 33(10):1987-1996. (in Chinese)
叶云,徐锡山,贾焰,等. 基于攻击图的网络安全概率计算方法[J]. 计算机学报,2010,33(10):1987-1996.

[10] ZHAO Bao. The study of network vulnerability based on attack graphs [D]. Changsha: National University of Defense Technology, 2009. (in Chinese)
赵豹. 基于攻击图的网络脆弱性分析技术研究[D]. 长沙:国防科学技术大学,2009.

[11] WANG Hui, WANG Yunfeng, YAN Xixi. HTV analysis method for hierarchical threats degree using Bayesian attack graph[J]. Computer Applications and Software, 2016,33(7):287-293. (in Chinese)
王辉,王云峰,闫玺玺. 基于贝叶斯攻击图的层次化威胁度 HTV 分析方法[J]. 计算机应用与软件,2016,33(7):287-293.

[12] LIU Shengwa, GAO Xiang, WANG Min. Application of attack graph method based on Bayesian network in network security assessment [J]. Modern Electronics Technique, 2013, 36(9):84-87. (in Chinese)
刘胜娃,高翔,王敏. 基于贝叶斯网络的攻击图方法在网络安全评估中的应用[J]. 现代电子技术,2013,36(9):84-87.

[13] WU Di, LIAN Yifeng, CHEN Kai, et al. A security threats identification and analysis method based on attack graph[J]. Chinese Journal of Computers, 2012, 35(9): 1938-1950. (in Chinese)
吴迪,连一峰,陈恺,等. 一种基于攻击图的安全威胁识别和分析方法[J]. 计算机学报,2012,35(9):1938-1950.

[14] SU Huaan. The study of IDS alert correlation and prediction based on attack gragh [D]. Changsha: National University of Defense Technology, 2010. (in Chinese)
苏华安. 基于攻击图的 IDS 警报关联预测技术研究[D]. 长沙:国防科学技术大学,2010.

[15] ZHAO Fangfang. Network security vulnerabilities detecting and attack gragh constructing [D]. Shanghai: Shanghai Jiao Tong University, 2008. (in Chinese)
赵芳芳. 计算机网络安全漏洞检测与攻击图构建的研究[D]. 上海:上海交通大学,2008.

[16] LIU Qiang, YIN Jianping, CAI Zhiping, et al. Uncertain-graph based method for network vulnerability analysis[J]. Journal of Software, 2011, 22(6):1398-1412. (in Chinese)
刘强,殷建平,蔡志平,等. 基于不确定图的网络漏洞分析方法[J]. 软件学报,2011,22(6):1398-1412.

[17] ZHANG Aifang. Network multistage attack model and detection approach based on extended directed graph[D]. Wuhan: Huazhong University of Science and Technology, 2008. (in Chinese)
张爱芳. 基于扩展有向图的复合攻击模型及检测方法研究[D]. 武汉:华中科技大学,2008.

[18] LUO Zhiyong, YOU Bo, XU Jiazhong, et al. Automatic recognition model of intrusive intention based on three layers attack graph [J]. Journal of Jilin University (Engineering and Technology Edition), 2014, 44(5): 1392-1397. (in Chinese)
罗智勇,尤波,许家忠,等. 基于三层攻击图的入侵意图自动识别模型[J]. 吉林大学学报(工学版),2014,44(5):1392-1397.

[19] QIAN Quan, ZHU Wei, LAI Yanyan, et al. Host-based attack graph for attack recognition [J]. Journal of Shanghai University (Natural Science Edition), 2013, 19(3):271-279. (in Chinese)
钱权,朱伟,赖岩岩,等. 基于主机攻击图的攻击识别[J]. 上海大学学报(自然科学版),2013,19(3):271-279.

[20] HUANG Yonghong, WU Yifan, YANG Haopu, et al. Graph-based vulnerability assessment for APT attack[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2017, 29(4):535-541. (in Chinese)
黄永洪,吴一凡,杨豪璞,等. 基于攻击图的 APT 脆弱节点评估方法[J]. 重庆邮电大学学报(自然科学版),2017,29(4):535-541.

[21] HUANG Guangqiu, CHENG Kaige. Expanded Petri net attack model based on attack graph [J]. Computer Engineering, 2011, 37(10):131-133. (in Chinese)
黄光球,程凯歌. 基于攻击图的扩充 Petri 网攻击模型[J]. 计算机工程,2011,37(10):131-133.

编辑　陆燕菲