



基于知识图谱的泛在电力物联网安全可视化技术

余洋, 朱少敏, 卞超轶

(北京启明星辰信息安全技术有限公司, 北京 100193)

摘要: 随着泛在电力物联网的建设, 电力网络向互联互通转变、数据向交互共享转变、业务向横向贯通转变, 对原有网络安全防护体系提出了新的挑战。网络安全可视化技术通过对图模式的展示和分析, 帮助网络安全人员快速识别潜在攻击, 定位异常事件, 发现新的攻击类型, 快速捕获全局网络安全态势。基于知识图谱进行泛在电力物联网安全分析, 通过对内外部威胁情报进行建模, 构建概念实体间关系, 将威胁情报知识图谱化, 把复杂的警报信息转换成直观的、便于理解的图形图像信息, 为泛在电力物联网的网络安全防护决策提供精准支撑。

关键词: 泛在电力物联网; 安全可视化; 威胁情报; 知识图谱

中图分类号: TN929

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2019280

Knowledge graph based ubiquitous power IoT security visualization technology

YU Yang, ZHU Shaomin, BIAN Chaoyi

Beijing Venus Information Security Technology Co., Ltd., Beijing 100193, China

Abstract: With the construction of the ubiquitous power IoT, the power network transform to interconnection, the data become more interactive and shared, and the business transform to lateral link-up, which pose new challenges to the original network security protection system. Network security visualization technology helps network security personnel quickly identify potential attacks, locate abnormal events, discover new types of attacks, and quickly capture global network security situation by displaying and analyzing the graph patterns. Based on the knowledge graph, the ubiquitous power IoT security analysis was carried out. The internal and external threat intelligence was modeled firstly. Then the relationship between the conceptual entities was constructed. The threat intelligence as knowledge graph was modeled, and then the abstract and complex alarm information which was invisible was converted into a more intuitive and convenient style, which provided accurate support for ubiquitous network security protection decisions in the power IoT.

Key words: ubiquitous power IoT, security visualization, threat intelligence, knowledge graph

1 引言

泛在电力物联网是能源革命和信息通信技术

相互融合的产物, 是产业互联网在电力行业的具体体现和应用落地。泛在电力物联网逐步向着“数据采集+集中控制+区域自治”的模式转变, 终端

由单一功能型转向综合服务型,以开放型业务中台和数据中台方式提供以客户为中心的快速业务服务模式,面临着更为复杂的安全威胁和风险。随着综合能源服务、互联网金融等新业务的蓬勃发展,在安全交互等方面对现有防护体系提出了更高的要求。泛在电力物联网“人—机—物”安全互动和海量信息的高效快速处理需求,需要开展基于可视化的全场景网络安全分析和防护体系建设,有效应对新业务发展带来的安全风险。同时泛在互联对精准智能防护提出更高的要求,各类物联网设备的可视性、“云雾中心”的监测感知精准性有待进一步提升,而安全分析是达成上述愿景的前提。安全分析本质上是从现有安全数据中(包括安全日志、流量、恶意样本等)提取威胁情报,并综合分析的过程^[1-3]。然而随着泛在电力物联网网络规模增大,安全数据的量及种类激增,导致安全分析的过程越来越复杂,难度越来越大,尤其体现在异构数据的情报聚合问题上。虽然当前安全数据的种类增多了,然而不同种类数据的描述对象往往是同一类,甚至同一个,如何将这描述统一起来,共同描绘特定对象的行为和属性,进而得到更全面、更精确的安全分析结果,是一个值得研究的问题。网络安全可视化技术可以把不可见的、看似毫无关联的告警信息转换成直观的、便于理解的内外威胁情报,方便网络安全管理人员对海量告警事件精准分析,从而快速有效地感知网络态势,判别网络异常和攻击,是泛在电力物联网全场景安全防护的重要新技术手段。

知识图谱最早由美国谷歌提出,被用来提高搜索引擎的搜索效率,其本质是一个用来描述知识的可视化分析工具,能够简单直观地描述实体和实体之间的关系,符合人的思维习惯^[4-6];同时,知识图谱中描述的实体通常具有唯一标识,这一特性对于融合情报数据具有很大的优势。因此,本文基于知识图谱进行安全分析。首先,给出基于知识图谱进

行安全分析的总体架构;然后,分别给出部分典型场景下,利用知识图谱进行安全分析的实例;最后总结基于知识图谱进行安全分析的过程。

2 基于知识图谱的安全分析架构

图1是安全分析架构,整个安全分析架构大致分为3层,包括数据采集层、数据预处理层和安全分析层。

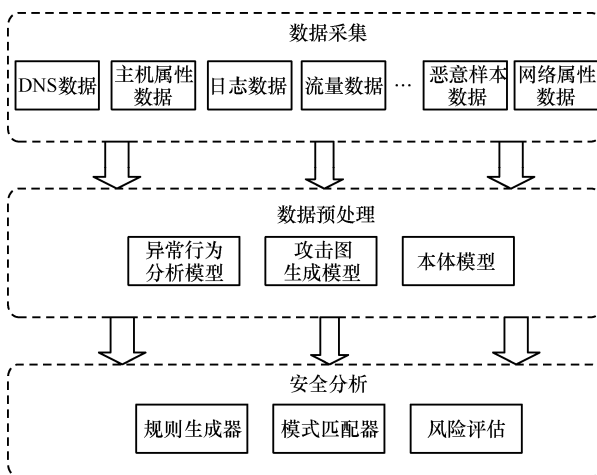


图1 安全分析架构

数据采集层的作用是通过IDS/IPS、防火墙以及漏洞扫描等设备,采集构建知识图谱所需的原始安全数据,主要包括主机属性数据、网络属性数据、恶意样本数据、日志数据、流量数据以及DNS数据等。通过采集层采集的数据大多是结构化的文本数据,为使安全分析更加有效,大多需要对这些数据进行进一步加工处理,从而获得更深层次的知识,例如基于主机属性信息和网络属性信息可以构建攻击图,进而将攻击图这种深层知识加入到知识图谱中,使得可以基于知识图谱查询特定特征的攻击路径。数据预处理层负责安全数据的进一步加工。安全分析层基于预处理的数据,构建安全分析知识图谱,安全分析层可以制定查询规则,用来查询异常行为、异常实体,同时可以基于已有的风险评估模型进行资产的风险评估。接下来本文着重阐述安全分析知识图谱



的构建,给出3个场景下知识图谱的构建过程及其在安全分析中的应用。

3 安全分析知识图谱

3.1 样本同源性知识图谱

安全数据可以从多角度体现样本的同源性,例如不同样本具有相同的回连服务器,就可以认为样本之间具有同源关系。根据恶意样本的分析结果,给出样本数据的形式化描述:

$$\begin{aligned} \text{SampleData} &::= (S, \xi_s) \\ S &::= (\text{sample}, \text{callback}), \\ \text{callback} &::= \{\text{tr}, \text{ur}, \text{hr}, \text{urr}, \text{cs}, \text{dr}\}, \\ \xi_s &::= \{\xi^{\text{TR}}, \xi^{\text{UR}}, \xi^{\text{HR}}, \xi^{\text{URR}}, \xi^{\text{CS}}, \xi^{\text{DR}}\}, \\ \xi^{\text{TR}} &: \text{sample} \rightarrow \text{tr}, \xi^{\text{UR}} : \text{sample} \rightarrow \text{ur}, \\ \xi^{\text{HR}} &: \text{sample} \rightarrow \text{hr}, \xi^{\text{URR}} : \text{sample} \rightarrow \text{urr}, \\ \xi^{\text{CS}} &: \text{sample} \rightarrow \text{cs}, \xi^{\text{DR}} : \text{sample} \rightarrow \text{dr} \end{aligned} \quad (1)$$

其中, SampleData 是样本数据,样本数据是一个二元组,包括实体集 S 和回连之间的关系 ξ_s , S 包括样本 sample 实体和回连 callback 实体, callback 包括6类,对应的关系也包括6类: tcp 请求 tr 以及 tcp 请求关系 ξ^{TR} 、udp 请求 ur 以及 udp 请求关系 ξ^{UR} 、http 请求 hr 以及 http 请求关系 ξ^{HR} 、url 请求 urr 以及 url 请求关系 ξ^{URR} 、连接

服务器 cs 以及有连接服务器关系 ξ^{CS} 、dns 请求 dr 以及 dns 请求关系 ξ^{DR} 。

根据上述数据描述给出恶意样本本体,如图2所示。

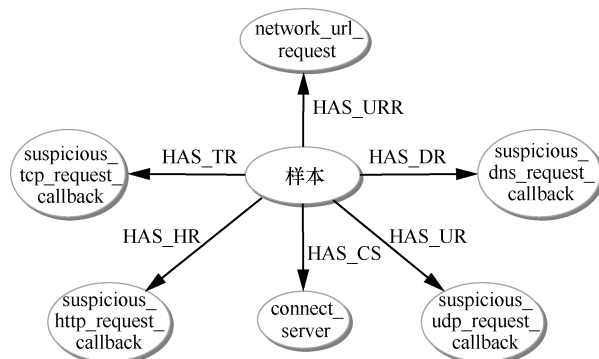


图2 样本同源性分析本体

该本体中共包括7种概念实体以及相应的6种关系。根据上述本体模型,将安全数据知识图谱化,可得到对应的知识图谱。在此基础上,通过执行如下的 cypher 查询语句:

```
MATCH p=()-[r:HAS_CONNECT_SERVER]
->() RETURN p LIMIT 25
```

可以得到如图3所示的结果。

图3给出了两个“簇”,簇中心表示的是回连服务器,而与簇中心连接的节点即为具有不同编号的样本,这种“簇结构”可以清晰地表示样

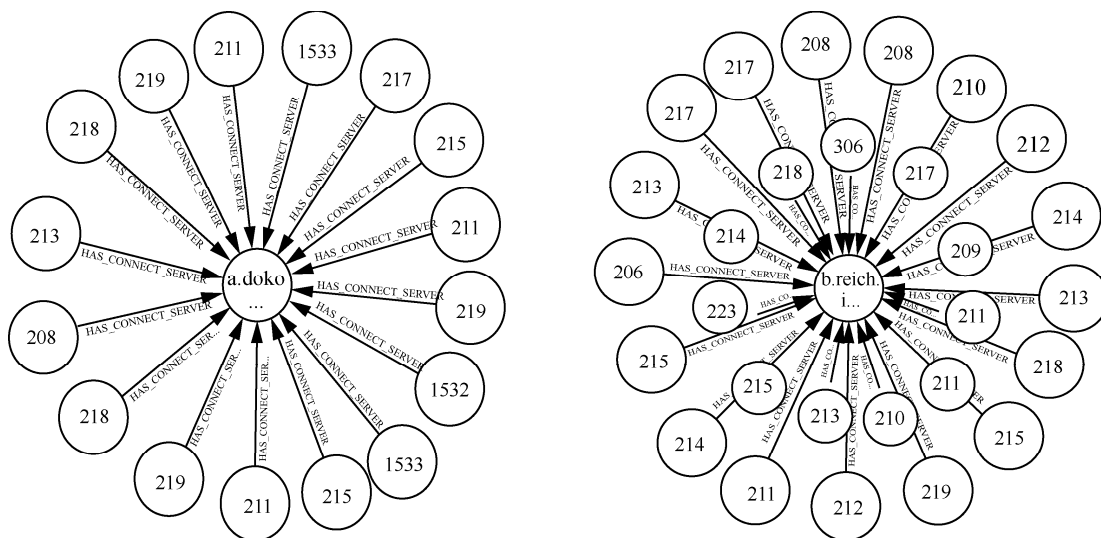


图3 同源性分析结果

本的同源性关系。

3.2 威胁狩猎知识图谱

威胁狩猎是威胁的搜索和关联过程，然而，仅发现孤立的异常行为已不能满足当前的需求。为了阻止威胁造成的影响，深入地分析威胁，需要分析威胁的扩散路径。而知识图谱能够很好地表达这种路径结构。在威胁狩猎场景下，使用了部分日志数据和 DNS 数据作为分析基础，下面给出数据的形式化描述。

$$\begin{aligned} \text{HuntingData} &::= (A, \xi_h), \\ A &::= \{\text{ip}, \text{host}, \text{mail}, \text{dns}\}, \\ \xi_h &::= \{\xi^{\text{RESL}}, \xi^{\text{MAL}}, \xi^{\text{QUERY}}, \xi^{\text{C2}}, \xi^{\text{SCAN}}\}, \\ \xi^{\text{RESL}} &: \text{host} \rightarrow \text{ip}, \xi^{\text{MAL}}: \text{mail} \rightarrow \text{ip}, \xi^{\text{QUERY}}: \text{ip} \rightarrow \text{dns}, \\ \xi^{\text{C2}} &: \text{ip} \rightarrow \text{ip}, \xi^{\text{SCAN}}: \text{ip} \rightarrow \text{ip} \end{aligned} \quad (2)$$

其中，HuntingData 是狩猎数据，狩猎数据是一个二元组，包括实体集 A 和回连之间的关系 ξ_h ， A 包括样本 ip 地址、主机地址 host、用户邮箱地址 mail 和 DNS 地址 dns，对应的关系包括 5 种：主机解析关系 ξ^{RESL} 、邮箱解析关系 ξ^{MAL} 、DNS 请求关系 ξ^{QUERY} 、命令控制请求关系 ξ^{C2} 以及扫描关系 ξ^{SCAN} 。

根据上述数据描述给出威胁狩猎本体，如图 4 所示。

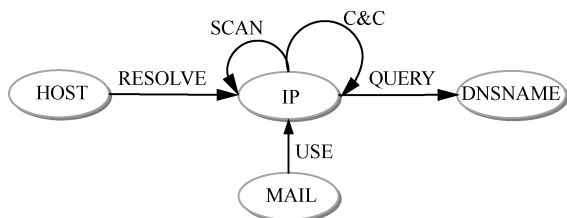


图 4 安全狩猎本体

图 5 标注了各个节点的种类，同时也对部分节点进行编号。其中，标注编号为 1、2 和 3 的节点为内网节点，从图 5 中可以看出内网节点 1 有 8 个 C&C 服务器，其中，编号 4~7 的 C&C 服务器被解析为同一个 URL 地址，说明该服务器上做了 IP

变换处理；同时，1 号主机同时还具有 8~11 号 C&C 服务器。根据上述分析，可以做相应的防御调整，例如针对性加强 1 号主机上的防御策略，或者将 1 号主机从内网隔离。

3.3 攻击图知识图谱

攻击图是能够反映主机间脆弱性关系的图结构，常用来分析攻击意图，也是分析攻击路径和进行风险评估的有效手段^[7-10]。将攻击图知识图谱化有利于更方便地展示和分析攻击路径。攻击图的生成过程需要多种数据，本文采用参考文献[11]中给出的方法生成攻击图，该过程需要的数据包括主机属性（主机上服务和脆弱性）和网络属性（主机间连接关系），基于定义的攻击模板，通过推理生成攻击图。攻击图生成过程使用的数据的形式化描述如下：

$$\begin{aligned} \text{AGData} &::= (\text{Data}, \xi), \\ \text{Data} &::= (\text{host}, \text{service}, \text{vulnerability}, \text{privilege}, \text{link}), \\ \xi &::= (\xi^{\text{hc}}, \xi^{\text{hs}}, \xi^{\text{hv}}, \xi^{\text{hp}}), \\ \xi^{\text{hc}} &: \text{host} \times \text{host} \rightarrow \text{link}, \xi^{\text{hs}}: \text{host} \rightarrow \text{service}, \\ \xi^{\text{hv}} &: \text{host} \rightarrow \text{vulnerability}, \xi^{\text{hp}}: \text{host} \rightarrow \text{privilege} \end{aligned} \quad (3)$$

攻击图生成数据 AGData 是由数据 Data 和数据间关系 ξ 构成，其中数据主要包括主机描述 host、服务描述 service、脆弱性描述 vulnerability、权限描述 privilege 和连接关系描述 link；数据关系主要有主机与主机间关系 ξ^{hc} 、主机与服务间关系 ξ^{hs} 、主机与脆弱性间关系 ξ^{hv} 、主机权限间关系 ξ^{hp} 。

攻击模板的形式化描述具体可见参考文献[11]，下面给出一条具体的远程提权类攻击模板，该模板使用 datalog 语言描述：

```
hostPrivilege (Attacker, DestName, 2):-
host (SourceName, IPaddress, Mask),
host (DestName, IPaddress, Mask),
hostService (DestName, ServiceName, Service-
Version, Protocol, Port, Privilege),
```

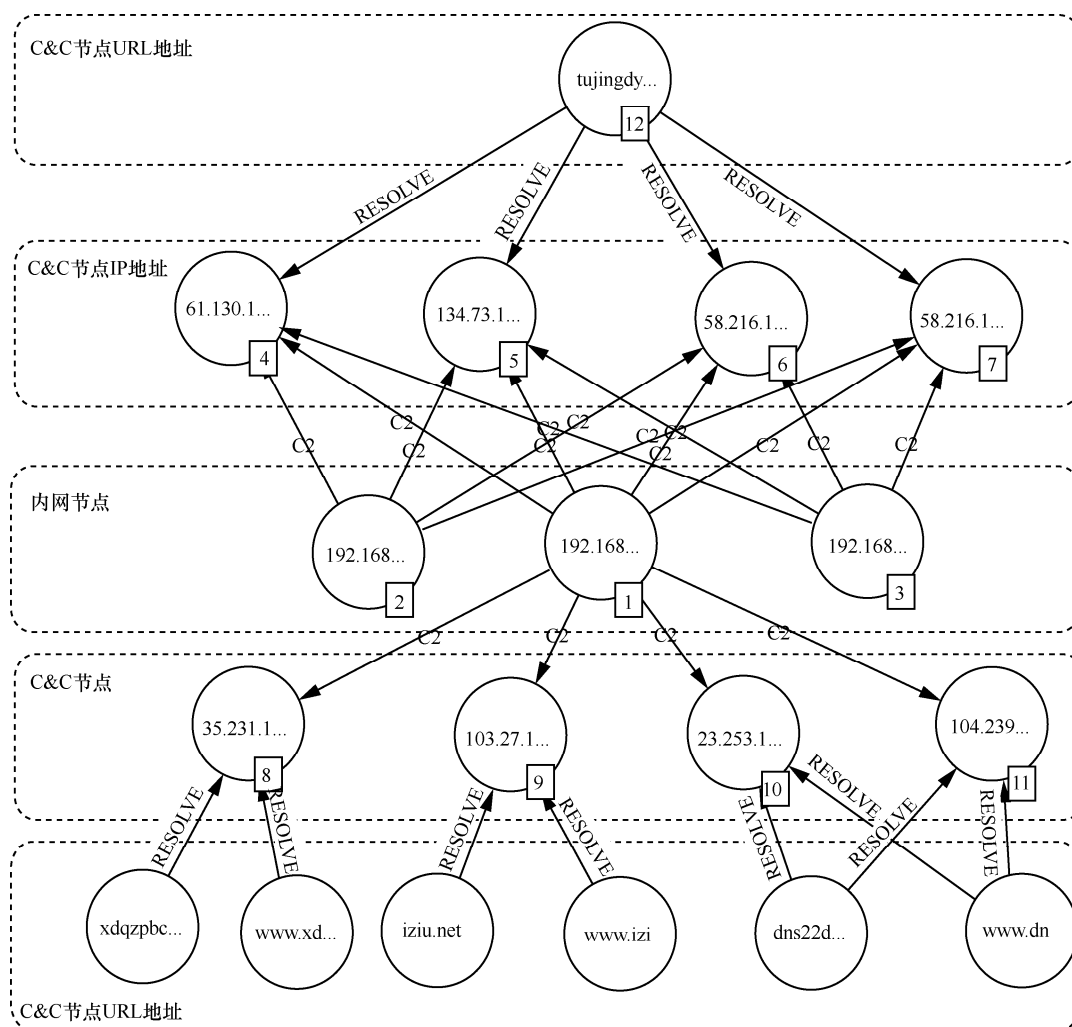


图5 威胁狩猎结果

Privilege ≥ 0 ,
 hostVulnerability (DestName, VulID, Service-
 Name, krpe),
 hostConnection (SourceName, DestName, Pro-
 tocol, Port),
 hostPrivilege (Attacker, SourceName, AttaPri),
 AttaPri ≥ 1 ,
 Exploit (Attacker, SourceName, DestName, VulID)

该攻击模板的意思是, 要完成源主机 SourceName 到目的主机 DestName 的提权操作, 必须具备 4 个前提条件: 目的主机 DestName 上存在特定服务 ServiceName; 目的主机 DestName 上

存在特定漏洞 VulID; 源主机 SourceName 和目的主机 DestName 可以通过协议 Protocol 及端口 Port 连接; 攻击者在源主机 SourceName 上具有特定权限 AttaPri。之后才能发生利用动作 exploit。

基于上述数据和攻击规则模板, 通过自动推理可得到攻击图, 如图 6 所示。

攻击图是数据加工过后的产物, 虽然攻击图中的元素也能表示攻击路径, 但是其表达能力没有经过重构过后的知识图谱强, 也没有知识图谱表达清晰; 同时, 知识图谱能够借助从其他数据中获取的知识集成到攻击图知识图谱中, 使得安全分析更加容易。图 7 是攻击图本体。

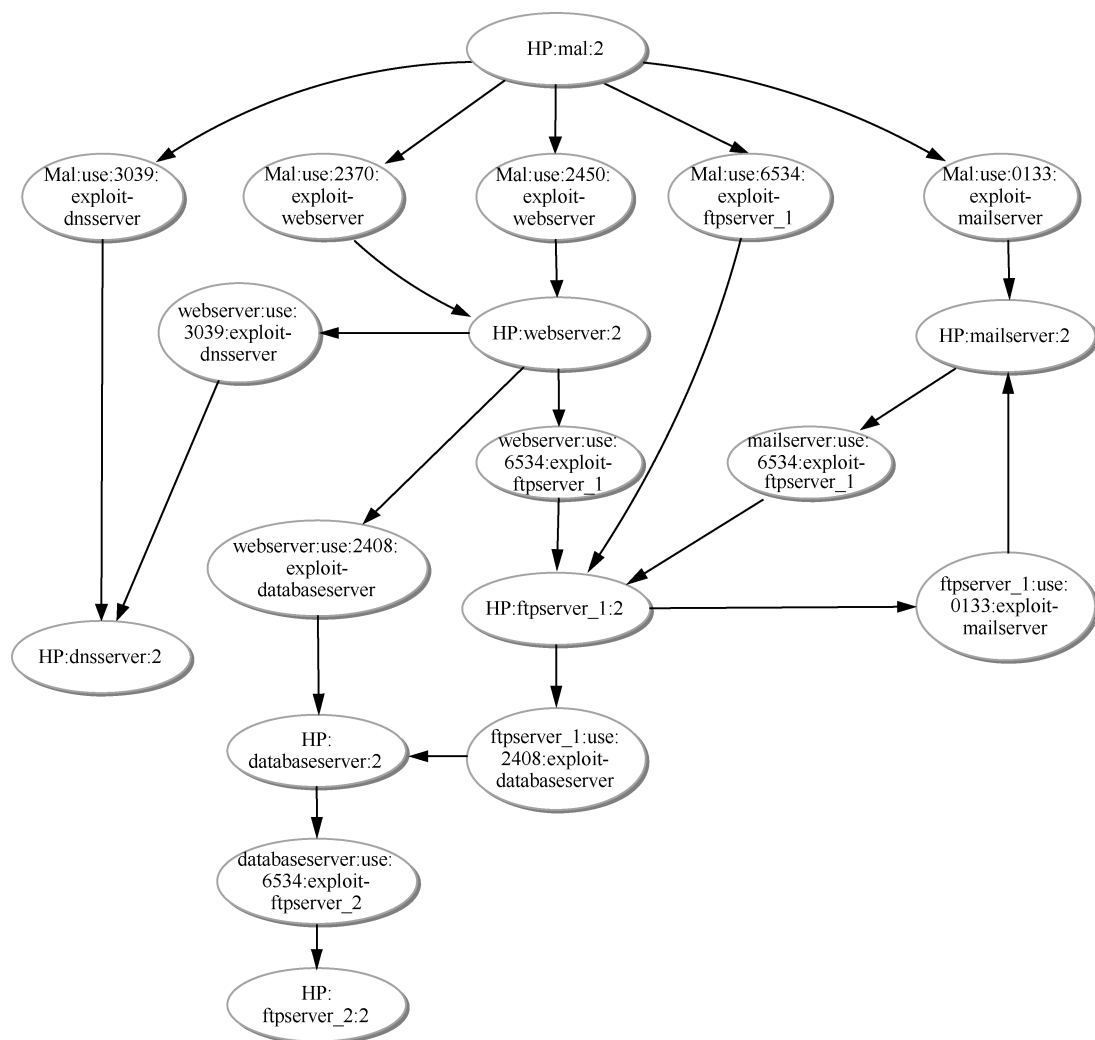


图 6 攻击图

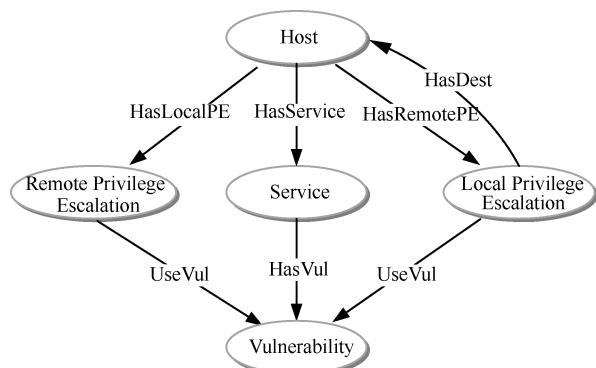


图 7 攻击图本体

基于上述攻击图本体可以构造攻击图知识图谱。可以通过查询查找任意两台主机之间的攻击路径，例如查找攻击者所在主机节点到 FTPserver_1

主机节点的所有攻击路径，如图 8 所示，共有 4 条从攻击者所在主机到达 FTPserver_1 主机的攻击路径。

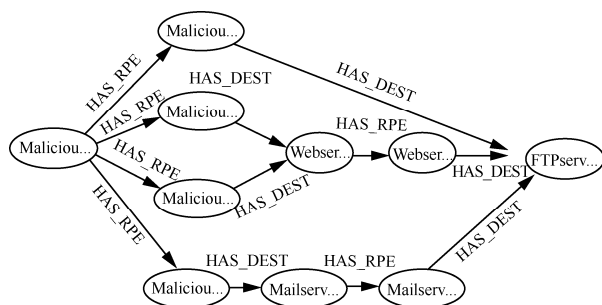


图 8 攻击路径

最短攻击路径如图 9 所示，查找攻击者所在主机节点到 FTPserver_1 主机节点之间的最短路径。



图9 最短攻击路径

图4和图7中描述的本体都有Host,且本质上是同一类概念,可将两套本体中同一概念的实体进行合并,将不同类数据隐含的知识进行融合,可获得更详细的关于该实体的信息,有助于安全分析。

4 结束语

泛在电力物联网覆盖了智能电网的信息感知、数据传输及智能处置等各个环节,在电网建设、生产管理、运行维护、用户服务等方面发挥巨大作用,可以全面提高智能电网各个环节的信息感知深度和广度^[12-13]。本文针对3个网络安全攻防场景,将知识图谱应用于内外部威胁情报整理和分析,不仅能有效地整理数据中隐含的知识,还能有效地挖掘知识之间的关联,使安全分析不仅能够分析局部数据,还能够总览全局,为泛在电力物联网可感知的网络安全提供了新的方法。

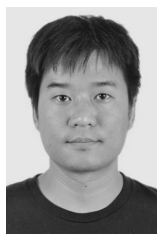
参考文献:

- [1] 修健. 网络安全分析中的大数据技术应用分析[J]. 网络安全技术与应用, 2018(3): 67-115.
XIU J. The analysis of big data technology application in network security[J]. Network Security Technology & Application, 2018(3):67-115.
- [2] 司德睿, 华程, 杨红光, 等. 一种基于机器学习的安全威胁分析系统[J]. 信息技术与网络安全, 2019, 38(4): 37-41.
SI D R, HUA C, YANG H G, et al. An intelligent analysis system for security threats based on machine learning[J]. Information Technology and Network Security, 2019, 38(4): 37-41.
- [3] 汪渊, 蒋凡, 陈国良. 基于图论的网络安全分析方法研究与应用[J]. 小型微型计算机系统, 2003(10): 1865-1869.
WANG Y, JIANG F, CHEN G L. A network security analysis method research and application based on graph theory[J]. Journal of Chinese Computer Systems, 2003(10): 1865-1869.
- [4] 侯梦薇, 卫荣, 陆亮, 等. 知识图谱研究综述及其在医疗领域的应用[J]. 计算机研究与发展, 2018, 55(12): 2587-2599.
HOU M W, WEI L, LU L, et al. Research review of knowledge

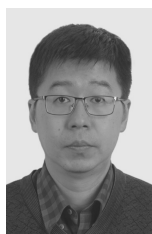
- graph and its application in medical domain[J]. Journal of Computer Research and Development, 2018, 55(12): 2587-2599.
- [5] 刘峤, 李杨, 段宏, 等. 知识图谱构建技术综述[J]. 计算机研究与发展, 2016, 53(3): 582-600.
LIU Q, LI Y, DUAN H, et al. Knowledge graph construction techniques[J]. Journal of Computer Research and Development, 2016, 53(3): 582-600.
- [6] 汤楠, 申德荣, 寇月, 等. 一种面向多源知识图谱的样例查询方法[J]. 计算机研究与发展, 2015, 52(S1): 1-8.
TANG N, SHEN D R, KOU Y, et al. A query-by-example method for multiple knowledge graphs[J]. Journal of Computer Research and Development, 2015, 52(S1): 1-8.
- [7] 王永杰, 鲜明, 刘进, 等. 基于攻击图模型的网络安全评估研究[J]. 通信学报, 2007(3): 29-34.
WANG Y J, XIAN M, LIU J, et al. Study of network security evaluation based on attack graph model [J]. Journal on Communications, 2007(3): 29-34.
- [8] 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.
YE Y, XU X S, JIA Y, et al. An attack graph-based probabilistic computing approach of network security[J]. Chinese Journal of Computers, 2010, 33(10): 1987-1996.
- [9] 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.
WU D, LIAN Y F, CHEN K, et al. A security threats identification and analysis method based on attack graph [J]. Chinese Journal of Computers, 2012, 35(9): 1938-1950.
- [10] 胡浩, 刘玉岭, 张玉臣, 等. 基于攻击图的网络安全度量研究综述[J]. 网络与信息安全学报, 2018, 4(9): 1-16.
HU H, LIU Y L, ZHANG Y C, et al. Survey of attack graph based network security metric[J]. Chinese Journal of Network and Information Security, 2018, 4(9): 1-16.
- [11] 余洋, 夏春和, 胡潇云. 采用混和路径攻击图的防御方案生成方法[J]. 浙江大学学报(工学版), 2017, 51(9): 1745-1759.
YU Y, XIA C H, HU X Y. Defense scheme generation method using mixed path attack graph[J]. Journal of Zhejiang University (Engineering Science), 2017, 51(9): 1745-1759.
- [12] 殷树刚, 许勇刚, 李祉岐, 等. 基于泛在电力物联网的全场景网络安全防护体系研究[J]. 供用电, 2019, 36(6): 83-89.
YIN S G, XU Y G, LI Z Q, et al. Research on full scene network security protection system based on ubiquitous power internet of things[J]. Distribution & Utilization, 2019, 36(6): 83-89.
- [13] 王静, 高昆仑, 张波. 基于网络隔离与安全数据交换的发电

集团双网体系研究与设计[J]. 电信科学, 2017, 33(2): 163-172.
WANG J ,GAO K L, ZHANG B. Research and design in dual network scheme of power corporation based on network isolation and secure data exchange[J]. Telecommunications Science, 2017, 33(2): 163-172.

[作者简介]



余洋（1985—），男，博士，北京启明星辰信息安全技术有限公司高级研究员，主要研究方向为大数据安全分析、人工智能安全应用、知识图谱。



朱少敏（1983—），男，博士，北京启明星辰信息安全技术有限公司高级工程师，主要研究方向为电力二次系统安全防护、多媒体信息安全、云计算安全。



卞超轶（1987—），男，博士，北京启明星辰信息安全技术有限公司高级研究员，主要研究方向为大数据安全分析、人工智能安全应用等。