

DOI:10.20033/j.1003-7241.(2026)10-0144-05

# 一种基于知识图谱的电力终端攻击行为快速溯源方法

王 晔, 郭 靛, 韦小刚, 刘 行, 金倩倩

(南京南瑞信息通信科技有限公司, 南京 210003)

**摘 要:**为了准确定位电网攻击事件的源头,研究基于知识图谱的海量终端接入电力网络攻击行为快速溯源技术。以电网攻击事件溯源本体模型为基础,塑造电力网络攻击行为知识图谱,利用知识图谱构建电网攻击溯源框架;以获取的海量终端接入电力网络后电网安全威胁情报和电网攻击事件线索两类数据为基础,采用基于图嵌入的电网攻击事件自动特征提取算法,提取电网攻击事件的关联特征向量,将其作为支持向量机的输入,通过挖掘电网攻击事件与攻击者间的分类关系,分类出电网攻击者,实现网络攻击行为快速溯源。实验结果表明:该技术能够对终端接入电网产生的攻击行为快速定位,找出被攻击数据的威胁源头并发出攻击报警,通过对电网攻击者的挖掘分析实现电网攻击行为快速溯源。

**关键词:**知识图谱;电力网络攻击;支持向量机;图嵌入;特征提取;网络攻击溯源

**中图分类号:** TP309; TM73

**文献标志码:** A

**文章编号:** 1003-7241(2026)10-0144-05

## A rapid source tracing method for power terminal attack behaviors based on knowledge graphs

Wang Ye, Guo Liang, Wei Xiaogang, Liu Xing, Jin Qianqian

(Nanjing NARI Information &amp; Communication Technology Co., Ltd., Nanjing 210003, China)

**Abstract:** In order to accurately locate the source of power grid attacks, the rapid tracing technology of massive terminals access to power network attacks based on knowledge graph is studied. Based on the ontology model of power grid attack event traceability, a knowledge graph of power grid cyber attacks is built and the framework of power grid attack traceability is constructed by using the knowledge graph. Based on the two types of data obtained from the power grid security threat intelligence and power grid attack event clues after the massive terminals are connected to the power grid, the automatic feature extraction algorithm of power grid attack events based on graph embedding is used to extract the correlation feature vector of power grid attack events, which is used as the input of support vector machine. By mining the classification relationship between power grid attack events and attackers, the grid attackers are classified, and then the network attack behavior can be traced quickly. The experimental results show that this technology can quickly locate the attack behavior generated by the terminal accessing to the power grid, find out the source of the threat of the attacked data and send out the attack alarm, and realize the rapid tracing of the power grid attack behavior through the mining and analysis of the power grid attacker.

**Keywords:** knowledge graph; power network attack; support vector machine; graph embedding; feature extraction; traceability of network attacks

我国电力行业智能化技术一直在持续提升,电网规模也越来越大,被人们广泛利用在各个电力领域,在电网建设过程中存在着海量的终端信息,系统无法确定其是否存在安全风险<sup>[1]</sup>,网络攻击者利用电网建设漏洞入侵电网系统并进行攻击,给电力网络带来了极大的安全威胁。网络攻击溯源技术是网络防御体系的关键技术,因此寻找有效的快速溯源技术实现对电网攻击事件的攻击者判断成了当下研究的重点。

罗予东等<sup>[2]</sup>提出一种遗传算法与神经网络相结合的

混合网络攻击检测溯源算法,利用 MLRM 与多目标遗传算法 NSGA-II 相结合生成封装特征选择算法实现对电网攻击事件的特征子集分类溯源,但该算法在检测分类过程中易出现检测疏漏,降低了电网故障诊断的准确率。李腾等<sup>[3]</sup>提出一种攻击溯源算法基于隐私保护的图卷积神经网络,实现对电网攻击行为特定类型的分类,为管理者提供具体的攻击类型,但该算法不能阻止未知的电网故障攻击进入,识别不出其安全性。

知识图谱是一种将非结构化的知识和数据转化为结

收稿日期:2025-08-28;录用日期:2025-10-14

基金项目:国网电力科学研究院有限公司南京南瑞信息通信科技有限公司科技项目资助(5246DR230021)

作者简介:王 晔(1984—),男,硕士,高级工程师,研究方向为电力信息化、网络与信息安全等。

通信作者:韦小刚(1983—),男,硕士,正高级工程师,研究方向为网络与信息安全。

引用本文:王晔,郭靛,韦小刚,等.一种基于知识图谱的电力终端攻击行为快速溯源方法[J].自动化技术与应用,2026,45(10):144-148.(Wang Ye, Guo Liang, Wei Xiaogang, et al. A rapid source tracing method for power terminal attack behaviors based on knowledge graphs[J]. Techniques of Automation and Applications, 2026, 45(10): 144-148.)

构化知识的数据库,将其应用于电力网络知识领域,能够高效实现电力网络信息的检索和呈现。为此,本文提出一种基于知识图谱的海量终端接入电力网络攻击行为快速溯源技术,通过攻击事件的关联特征向量分类出攻击者并实现威胁溯源,提高电力网络的安全性。

## 1 电力网络攻击行为溯源方法

### 1.1 电网攻击事件溯源本体模型

本体抽象化地描述现实世界某个领域的概念和概念之间关系,它由概念、实例和关系等基本元素组成<sup>[4]</sup>。本体中包含一种深层次的分类法,可以从不同的角度对事件进行一种更全面的分类描述。电网攻击事件溯源本体框架由电网攻击领域概念间的关联关系组成,利用本体构建电网攻击事件溯源本体模型,利用电网攻击事件中的分类层次判断溯源本体中概念间的层次关系<sup>[5]</sup>。

电网攻击溯源本体模型以层次模型作为基础,把攻击事件分为攻击领域本体和攻击原子本体,如下所示:

1)攻击领域本体。攻击领域本体是一种对攻击领域实体概念与互相领域活动等攻击领域的知识描述<sup>[6]</sup>。其中的攻击目标、过程、结果以及攻击所利用的漏洞用于表示攻击事件中概念和关系集;攻击所利用的漏洞、方法的选择、目标的确定和攻击所导致的结果这四种关系构成攻击类概念关系<sup>[7-8]</sup>。电网攻击者使用各种攻击方法对攻击事件目标的缺陷发起攻击,达到破坏性的效果,由图 1 所示。

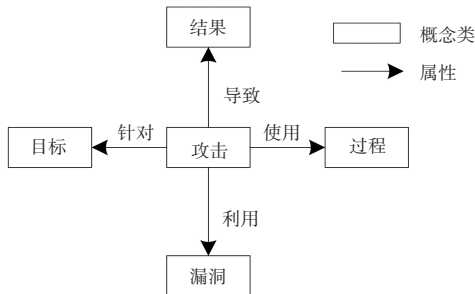


图 1 攻击领域本体

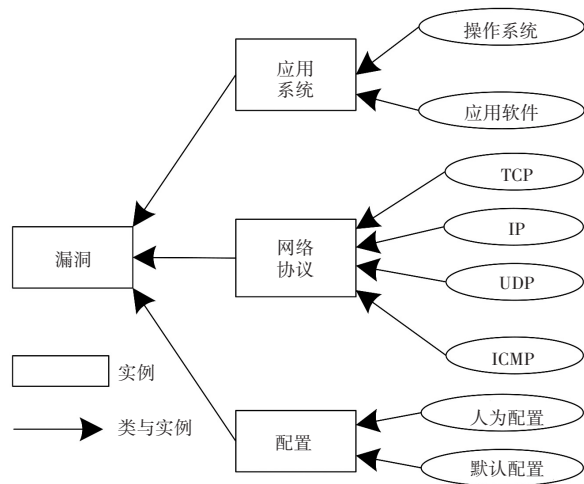
Fig. 1 Attack domain ontology

2)攻击原子本体。攻击原子本体是在攻击利用漏洞应用本体、攻击过程应用本体和攻击目标应用本体的基础上,对相关实体概念进行具体描述,细分为攻击利用漏洞原子本体、攻击过程原子本体、攻击目标原子本体<sup>[9-10]</sup>,如图 2 所示。

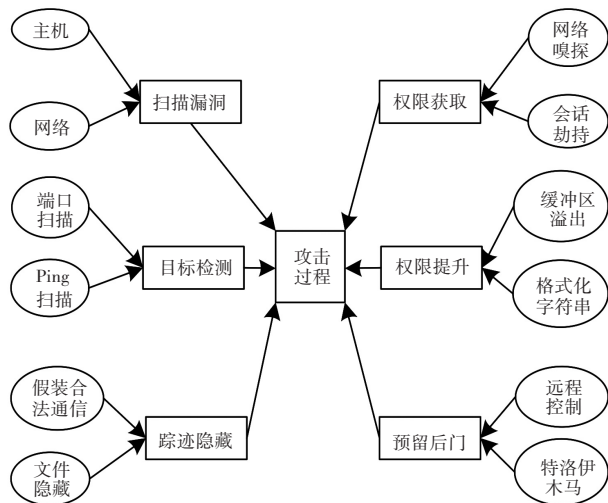
图 2(a)为攻击利用漏洞原子本体框架,包含软件系统漏洞、配置漏洞和电网协议漏洞;图 2(b)为攻击过程原子本体框架,它是一种单一的本体,能够独立完成功能和操作,与别的攻击行为无关;图 2(c)为攻击目标原子本体框架,它的分类取决于攻击目标的类型。

电网攻击领域的概念分层使攻击可以在较高的泛化抽象层中解决电网攻击事件,使攻击的层次观察更清晰<sup>[11]</sup>。替换高层概念集到底层概念集,使攻击本体的理

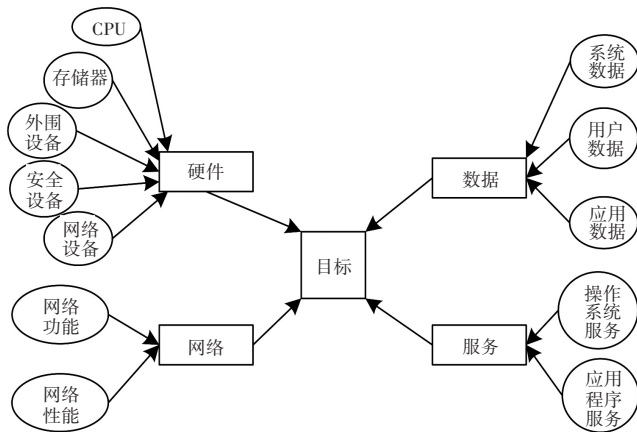
解性更强,实现不同攻击在不同细节层次上的描述,对整个攻击过程都进行了详细的分析,使攻击知识更容易被理解。



(a) 攻击利用漏洞原子本体



(b) 攻击过程原子本体



(c) 攻击目标原子本体

图 2 攻击原子本体

Fig. 2 Attack atomic ontology

1.2 塑造知识图谱

知识图谱通过不同的介质获取数据的属性参数后,通过实体间的联系获取检索内容、实体和属性间的关系,是一种有效的信息检索工具。知识图谱通过构建结构化数据库,用不同的符号描述数据间的关系以及相关描述<sup>[12-13]</sup>。实体-关系-实体这种形式是知识图谱描述数据的方式。

实体、属性以及关系构成知识图谱,通过这样的结构向逻辑关系中融入实体间的关系,塑造电力网络攻击行为知识图谱的过程如图3描述。

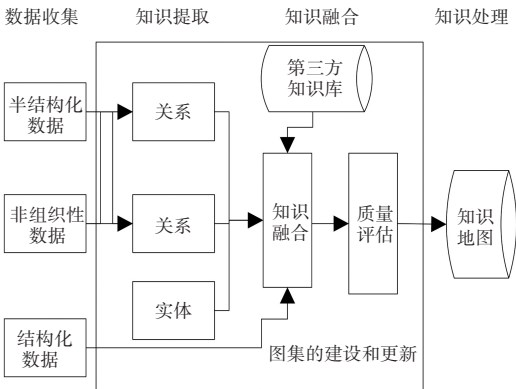


图3 知识图谱构建

Fig. 3 Knowledge graph construction

1.3 基于知识图谱的攻击溯源框架

知识图谱将实体、实体间关系和属性等知识图形成,是一种行业领域知识图谱,专业性较高、噪声较弱,通过面向电网终端信息的知识图谱以及电网攻击事件溯源本体模型,构建面对海量终端接入电网攻击行为的溯源框架,具体通过以下3个阶段构建。

1) 源数据获取。采集威胁情报源中电网攻击者的相关信息情报以及历史电网攻击事件有关的安全威胁情报,如已知攻击者、攻击者目标、攻击手段以及攻击模式等;电网攻击目标中存储大量数据源,如主机日志、流量日志、报警数据等,通过恶意代码分析、网络取证、数字取证调查等手段对电网攻击目标进行搜寻,获取电网攻击者的指纹、网络地址、网络域名等攻击线索数据。

2) 生成电网攻击事件溯源知识图谱。对获取到的电网攻击事件线索数据和电网安全威胁情报进行标准化处理,融合基于词向量和词典两种本体映射方法,实现实体数据的标准化对齐和多源异构数据的融合,利用如图3所示的知识图谱构建过程生成电网攻击事件溯源知识图谱。

3) 攻击者挖掘。利用基于随机游走的图嵌入算法生成电网攻击事件溯源实体,提取其中的电网攻击事件关联特征向量实现对SVM分类器的训练,进而完成对电网攻击者的判定与挖掘,最终构建出基于知识图谱的电网攻击事件溯源框架,如图4所示。

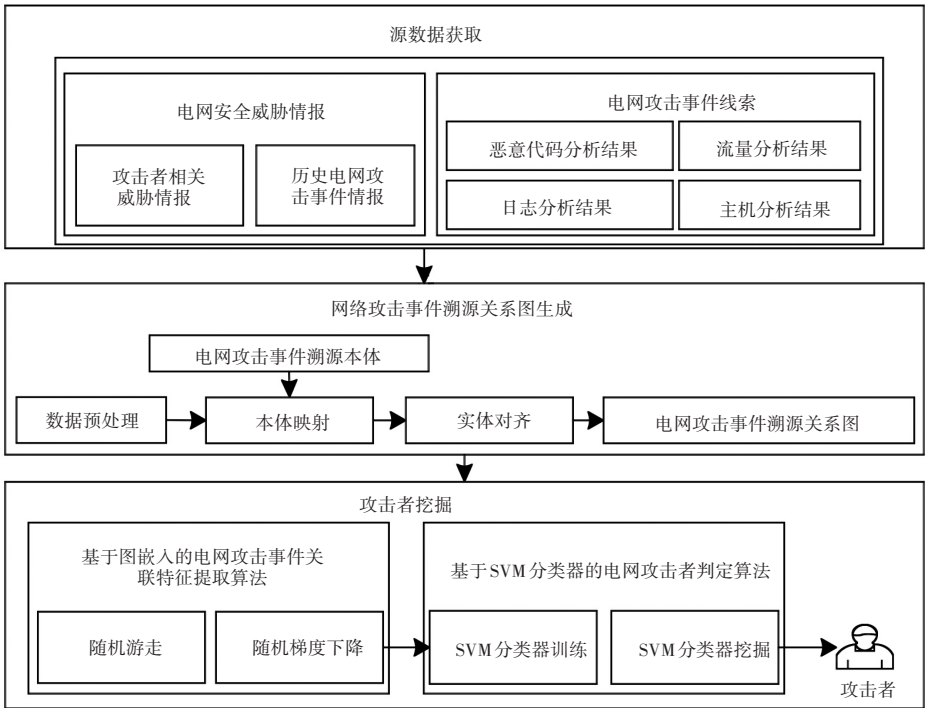


图4 基于知识图谱的攻击溯源框架

Fig. 4 Attack attribution framework based on knowledge graph

1.3.1 电网攻击事件特征提取算法

基于电网攻击溯源知识图谱,采用基于图嵌入的电网攻击事件特征自动提取算法,提取电网攻击事件的关联特征向量。主要有几个步骤:采用多源异构数据融合形成电

网攻击事件溯源知识图谱;利用基于随机游走的图嵌入算法生成电网攻击事件溯源实体序列后,计算电网攻击事件溯源实体的特征向量,具体算法如下。

1) 设实体集  $C_v^i$ , 其阶数为1, 且  $(u, v) \in C_v^i$ , 其中  $u$

表示电网攻击事件溯源实体  $v$  的直接关联集合。

2) 设实体  $v$  的间接关联实体集为  $C_v^k$ , 其阶数为  $k$ , 且  $\forall C_v^{k-1} \cap C_v^{k-1} \neq \emptyset$ 。

3) 设  $R$  是攻击事件的间接关联特征。

4) 设  $V_v$  为电网攻击事件溯源实体的特征向量, 表示 1 阶关联实体集  $C_v^1$  或  $k$  阶关联实体集  $C_v^k$  信息的数学向量, 即  $V_v = [x_1, x_2, x_3, \dots, x_d]$ , 其中  $x_d$  为特征向量。  $0 < d < k$ , 寻找函数  $f: v \rightarrow V_v$  的过程即是提取电网攻击溯源实体特征向量的过程。

为了确保后续溯源过程的进行, 需要依靠专家经验具体分析电网攻击事件, 并从电网攻击事件线索数据中挖掘出高质量的显性特征和隐性特征<sup>[14]</sup>; 当  $v$  发生时, 两种特征可能会同时存在, 因此可以同时进行  $v$  的特征提取和向量化两个过程。

该算法可以提取出电网攻击事件中隐含关系的关联特征向量, 为后续判定并挖掘电网攻击事件中的电网攻击者做基础<sup>[15]</sup>。

### 1.3.2 基于 SVM 分类器的攻击者判定

分类算法-支持向量机 (support vector machine, SVM) 是判断电网攻击者源头的核心算法, 具有结构风险最小化的优势。

SVM 分类器的目的在电网攻击行为快速溯源过程中准确地找到攻击事件的源头, 分类出电网攻击者, 可以采取基于图嵌入的电网攻击事件特征自动提取算法提取溯源电网攻击事件及历史电网攻击事件的特征向量并学习 SVM 分类器, 寻找出一个最优超平面, 使类间间隔最大, 确保攻击者分类效果最好。

设关联特征类别  $(x, y)$  的数据集  $P$ , 有

$$P = \{x_1, y_1\}, \{x_2, y_2\}, \dots, \{x_b, y_b\},$$

$$y_a \in \{+1, -1\} \quad (1)$$

式中,  $y_a$  表示第  $a$  个关联特征向量,  $b$  表示特征容量。再设一个分类超平面公式, 即

$$y = (x + v) \cdot \omega^T \quad (2)$$

式中,  $\omega$  表示法向量, 由  $\omega$  确定超平面方向; 位移项  $v$  确定原点和超平面之间的距离, 有

求超平面和两种异类支持向量的距离:

$$\gamma = 2 \times \|\omega\|^{-1} \quad (3)$$

将式(3)转为凸二次规划优化问题, 最小化  $\|\omega\|^2$ , 实现间隔最大化, 约束条件为

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 \quad (4)$$

$$\text{s. t. } y_i(\omega^T x_i + b) \geq 1 \quad (5)$$

式中,  $i = 1, 2, \dots, n$ 。

引入惩罚项到式(4), 目标函数更改为

$$\min_{\omega, b} \frac{1}{2} \|\omega\|^2 + Z \sum_{i=1}^n \xi_i \quad (6)$$

式中,  $\xi_i$ 、 $Z$  分别表示随机的松弛变量和惩罚参数。当  $Z$  是无穷大时, 式(6)中所有特征都需满足约束条件; 当  $Z$

是有限值时, 部分特征可以不满足约束条件, 为此提高模型的泛化能力。

转换式(6)为求解对偶问题, 引入核函数和 Lagrange 乘子, 解决非线性分类, 即

$$\max \tau = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j \kappa(x_i, x_j) \quad (7)$$

式中,  $\kappa(x_i, x_j)$ 、 $\alpha_i$  分别表示核函数和 Lagrange 乘子, 通过计算推算出相应的分类模型:

$$f(x) = \text{sgn}\left(\sum_{i=1}^n \alpha_i \kappa(x_i, x_j) + Z\right) \quad (8)$$

式中,  $\text{sgn}$  是标注特征类别的判别函数。

核函数的选择决定分类模型的计算效率, 高斯核 (RBF) 通过调整核函数半径  $\sigma$  同时实现线性和非线性两种分类, 解决了 SVM 和基于图嵌入的电网攻击事件特征自动提取的融合算法过于复杂的问题。因此本文利用 RBF 核完成解决电网攻击特征向量和攻击行为类型之间的分类问题:

$$\kappa(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (9)$$

综上采用 SVM 分类器获取最终的电力网络攻击行为溯源结果。

## 2 实验分析

为验证本文所述技术, 选取海量终端、防病毒软件和两台数据服务器, 搭建试验环境。每个终端中存储 Word、Excel、PPT、图片等多种数据文件信息, 服务器中存储邮件系统和点对点服务软件, 利用服务器对终端数据进行交互。采用较高传播性、感染性和控制性的病毒 (如蠕虫病毒、宏病毒等) 攻击电网数据交互过程中的数据, 实验环境拓扑如图 5 所示。

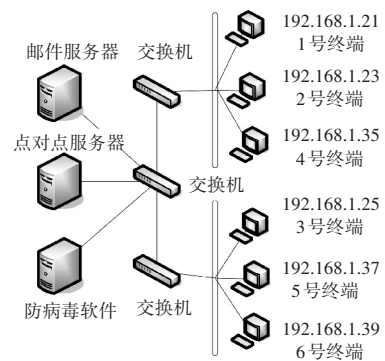


图 5 实验环境拓扑

Fig. 5 Experimental environment topology

设定终端数量为 5、10、15 台, 分别采用本文技术、基于人工神经网络和遗传算法的组合攻击检测技术、基于图卷积神经网络的攻击溯源技术进行对比实验。使用上述的网络病毒对电网中的数据进行攻击, 选择攻击时产生的不同病毒样本数据进行 3 组溯源实验, 取得最终的平均值, 实验结果如表 1 所示。

表 1 实验结果统计  
Tab. 1 Statistics of experimental results

| 指标      | 5            |                |       | 10           |                |       | 15           |                |       |
|---------|--------------|----------------|-------|--------------|----------------|-------|--------------|----------------|-------|
|         | 平均样本<br>数据/个 | 溯源实验平均<br>结果/个 | 准确率/% | 平均样本<br>数据/个 | 溯源实验平均<br>结果/个 | 准确率/% | 平均样本<br>数据/个 | 溯源实验平均<br>结果/个 | 准确率/% |
| 本文技术    | 5.3          | 4.9            | 92.5  | 9.4          | 9.1            | 96.8  | 20.8         | 20.2           | 97.1  |
| 文献[2]技术 | 5.3          | 2.1            | 39.6  | 9.4          | 5.1            | 54.2  | 20.8         | 10.7           | 51.4  |
| 文献[3]技术 | 5.3          | 1.2            | 22.6  | 9.4          | 3.8            | 40.4  | 20.8         | 10.9           | 52.4  |

由实验结果看出,本文技术能准确判断出被攻击的数据,并快速定位被攻击数据的源头,并且随着终端数量的增加,准确率不断提高,相比文献[2]和文献[3]技术,本文技术具有对攻击溯源较高的准确性。

为了快速追溯海量终端接入到电网中产生安全事件的源头,实验采用本文技术对攻击安全事件进行溯源,结果如图 6 所示。

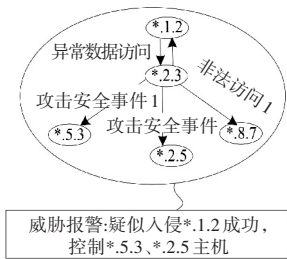


图 6 攻击溯源结果

Fig. 6 Attack traceability results

由图 6 可得,采用本文技术能够快速追溯到安全事件威胁源头为\*.1.2处,并发出疑似攻击事件的报警,生成攻击图并确定威胁源头,实现对大量终端进入电力电网后的安全事件威胁溯源,证明了本文技术的有效性。

设置实验电网受到病毒攻击的安全漏洞样本数量分别为 20、40、60、80 以及 100 个,漏洞攻击路径数量依次是 3、5、7、9、11 个。实验通过式(12)运算本文溯源技术在不同的攻击路径影响下的溯源均方误差,结果用表 2 描述。

$$MSE = \frac{1}{G} \sum_{s=1}^G (\Gamma_1 - \Gamma_2)^2$$

(10)

式中,  $\Gamma_1$ 、 $\Gamma_2$  分别用于描述本文技术的溯源结果和实际结果,  $G$  表示样本数量,  $s$  用于描述计算次数。

表 2 溯源结果

Tab. 2 Traceability results

| 攻击路径数量/个 | 安全漏洞样本数量/个 |         |         |         |         |
|----------|------------|---------|---------|---------|---------|
|          | 20         | 40      | 60      | 80      | 100     |
| 3        | 0.096 8    | 0.096 9 | 0.099 5 | 0.098 1 | 0.097 8 |
| 5        | 0.098 0    | 0.096 9 | 0.099 4 | 0.098 2 | 0.097 8 |
| 7        | 0.098 1    | 0.097 6 | 0.099 4 | 0.098 1 | 0.098 3 |
| 9        | 0.098 1    | 0.098 2 | 0.099 3 | 0.098 2 | 0.098 3 |
| 11       | 0.098 2    | 0.098 3 | 0.099 5 | 0.098 1 | 0.098 5 |

分析表 2 中数据能够看出,在不同的路径攻击数量前提下,本文技术的溯源均方误差都低于 0.1,说明本文溯源技术具有较高的精度,具有较强的对抗攻击路径数量影响能力。

3 结论

为避免海量终端接入电网产生的攻击事件行为,本文结合知识图谱与 SVM 共同实现电网攻击者的挖掘溯源。其中知识图谱的应用更加有利于追踪产生攻击事件行为的攻击者,助力攻击行为快速溯源,通过实验对本文采用技术的可行性及有效性进行了明确的验证,攻击溯源准确率高且可有效应对多种攻击并发下的攻击溯源。

参考文献

[1]张思龙,王兰成,娄国哲.基于知识图谱的网络舆情研判系统研究[J].现代情报,2021,41(4):10-16.

[2]罗予东,陆璐.基于人工神经网络和遗传算法的网络攻击检测[J].计算机工程与设计,2021,42(9):2446-2454.

[3]李腾,乔伟,张嘉伟,等.隐私保护的基于图卷积神经网络的攻击溯源方法[J].计算机研究与发展,2021,58(5):1006-1020.

[4]郭榕,杨群,刘绍翰,等.电网故障处置知识图谱构建研究与应用[J].电网技术,2021,45(6):2092-2100.

[5]李晓露,左璇,刘日亮,等.基于形状约束语言的电网模型知识图谱验证方法[J].中国电力,2022,55(1):119-125,228.

[6]梁若舟,高跃,赵曦滨.基于序列特征提取的溯源图上APT攻击检测方法[J].中国科学:信息科学,2022,52(8):1463-1480.

[7]孙澄,胡浩,杨英杰,等.基于网络防御知识图谱的Oday攻击路径预测方法[J].网络与信息安全学报,2022,8(1):151-166.

[8]冷涛,蔡利君,于爱民,等.基于系统溯源图的威胁发现与取证分析综述[J].通信学报,2022,43(7):172-188.

[9]陈志华,黄志宏.基于知识图谱的激光通信网络入侵攻击源定位方法[J].应用激光,2022,42(7):118-124.

[10]刘雪花,丁丽萍,郑涛,等.面向网络取证的电网攻击追踪溯源技术分析[J].软件学报,2021,32(1):194-217.

[11]徐伟.一种基于知识图谱的电力设备信息自适应检索方法[J].自动化技术与应用,2025,44(10):110-114.

[12]仇铭阳,赛煜,王刚,等.基于时间-概率攻击图的网络安全评估方法[J].火力与指挥控制,2022,47(1):145-149,155.

[13]朱红勤,周航,赵扬,等.基于知识图谱的电力调度控制系统运行状态识别模型[J].自动化技术与应用,2024,43(9):180-184.

[14]范敏,李昌茂,陈飞宇,等.基于EBLOF算法的攻击者IP分析系统及应用[J].工程科学与技术,2022,54(3):91-97.

[15]陈柏翰,王俊峰,唐宾徽,等.面向组织溯源的威胁行为技术关联研究[J].四川大学学报(自然科学版),2022,59(1):75-83.