

IPv6 网络攻击事件溯源中的 攻击树节点特征定位

唐晓萌^{1†}, 任凯², 滕俐军³, 汪敦全³, 梁鼎铭³

(1. 福建华电可门发电有限公司, 福建 福州 350500; 2. 华电电力科学研究院有限公司, 浙江 杭州 310000;

3. 深圳融安网络科技有限公司, 广东 深圳 518000)

摘 要: 在网络普及的现代, 各类网络协议层出不穷, 而在当下最为普遍的便是 IPv6 网络协议, 其虽然优化解决了前网络协议中存在的部分问题, 但是依旧存在各类安全隐患, 为了能更好地预防并构建网络防御系统, 需要对过往攻击事件中的攻击树节点特征实施提取。为此, 提出了 IPv6 网络攻击事件溯源中的攻击树节点特征定位。该方法首先利用攻击类型实施扩展, 得到攻击树模型, 并以此展开计算, 之后结合深度学习网络对攻击树的节点特征实施提取, 最后利用 DV-Hop 方法对提取到的攻击树节点特征实施定位。所提方法在定位过程中不仅能耗低, 而且误差小, 不易受环境干扰, 定位稳定性较好, 定位效率较高, 可以更为有效地为改进网络防御系统提供数据支持。

关键词: IPv6 网络; 攻击树; 节点特征; 特征定位; 事件溯源

中图分类号: T393.08

文献标识码: A

Attack Tree Node Feature Location in IPv6 Network Attack Event Source Tracing

TANG Xiaomeng^{1†}, REN Kai², TENG Lijun³, WANG Dunquan³, LIANG Dingming³

(1. Fujian Huadian Kemen Power Generation Co., Ltd., Fuzhou, Fujian 350500 China;

2. Huadian Electric Power Research Institute Co., Ltd., Hangzhou, Zhejiang 310000 China;

3. Shenzhen Rongan Networks Technology Co., Ltd., Shenzhen, Guangdong 518000 China)

Abstract: In the modern era of network popularization, various network protocols emerge one after another. At present, the most common is the IPv6 network protocol. Although it has optimized and solved some problems in the previous network protocols, there are still various security risks. In order to prevent and build a network defense system better, it is necessary to extract the characteristics of attack tree nodes in past attacks. Therefore, the node feature location of attack tree in IPv6 network attack event traceability is proposed. This method first uses the attack type to expand and obtain the attack tree model, and then calculates based on this model. Then it combines the deep learning network to extract the node features of the attack tree, and finally uses the DV-Hop method to locate the extracted node features of the attack tree. The proposed method not only has low energy consumption, but also has small error and good positioning performance, which can provide more effective data support for improving the network defense system.

Key words: IPv6 network; attack tree; node features; feature location; event source tracing

随着当代社会科技的高速发展,IPv6 成为新一代网络协议,其相较于 IPv4 虽然增加了丰富的地址空间、简便的配置以及各类安全性和移动性的支持,但是 IPv6 并没有从根源上解决 IPv4 中存在的问题,依旧存在各类安全隐患。在 IPv6 网络受到攻击时,攻击树的节点定位是预防下次攻击并改进网络防御系统的重要环节,这也使得攻击树的节点定位成为当下研究工作的热点。

刘浩然等^[1]利用下层无迹卡尔曼滤波算法与侧节点状态,结合信标节点的测距延迟传播修正节点状态,之后利用上层滤波对其开展新一轮的预测与修正,该方法定位误差较小,但在不同环境下的单位时间内定位时波动性较大,稳定性较差。余修武等^[2]利用 Bounding-box 方法划分出节点存在区域,之后再在区域内实施初始化启发个体并结合群居蜘蛛优化算法与加权中心反向学习策略,以获取节点位置,该方法计算效率高,但得到的节点定位误差较大。郝佳星^[3]利用圆的平分线必过圆心的原理结合锚点的移动获取定时反馈的定标消息,以此获取准确的标定点和弦,在得到两条和弦后,根据上述原理计算得出节点位置,该方法具有较为准确的定位性能,但在定位过程中能耗较大。

为了解决上述问题,提出 IPv6 网络攻击事件溯源中的攻击树节点特征定位。

1 定位方法

建立 IPv6 网络攻击事件溯源中的攻击树节点特征定位整体框架图,如图 1 所示。

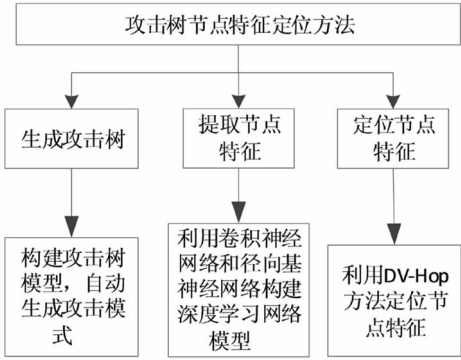


图 1 攻击树节点特征定位整体框架图

1.1 攻击树的生成

攻击树的构成:根为攻击的目标,而中间代表子目标,叶表示分解步骤;虽然已经有研究手工构建出攻击树并证明其有效性,但手工构造的攻击树存在效率低的弊端,为此,需要构建便于检测且可

以自动生成攻击树的算法,由此得到如图 2 所示的攻击树模型,其可以很好地实现攻击模式自动生成和重用。

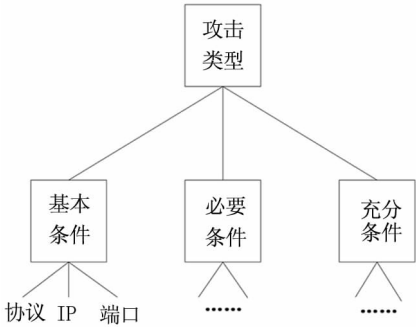


图 2 攻击树扩展定义图

算法目标:通过源数据构建可以准确表述攻击行为的攻击树^[4],攻击树的具体构建步骤如下:

输入:源数据 $E = \{B_1, B_2, \dots, B_n, B\}$ 。输出:攻击树。

步骤 1:对 $E = \{B_1, B_2, \dots, B_n, B\}$ 实行归约处理,获取数据集 $E' = \{B_{i_1}, B_{i_2}, \dots, B_{i_n}, B\} (\{j_1, j_2, \dots, j_k\} \subseteq \{1, 2, \dots, n\})$ 。

步骤 2:将 B 中的数据按属性代表各个攻击取值 $\{a_1, a_2, \dots, a_l\} (a_j \neq a_k, j, k \in \{1, 2, \dots, l\} \text{ 且 } j \neq k)$,将 E' 转变为子集集合 $\{C_1, C_2, \dots, C_l\}$ 并满足 $C_1 \cap C_2 \cap \dots \cap C_l = \varphi$ 且 $C_1 \cup C_2 \cup \dots \cup C_l = E'$ 。其中 C_j 中所有样本 B 属性的取值均为 b_j (假定 b_1 为正常行为)。

步骤 3: for $x = 2$ to l

①攻击树的根节点用 b_x 表示,记作 s_x 。

②从 $C_x = \{B_{j_1}^x, B_{j_2}^x, \dots, B_{j_k}^x\}$ 中获取基本条件信息,若其中的属性只有唯一取值,那么便将其看作叶结点添加至攻击树的基本分支^[5],若还有其他取值,则不添加。

③扫描 C_x , 获取 $\{B_{j_1}^x, B_{j_2}^x, \dots, B_{j_k}^x\}$ 各属性的取值与出现频率,将属性取值为 1,频率为 100% 的属性单独列出,记作 $N_e = \{B_{c_1}^x, B_{c_2}^x, \dots, B_{c_s}^x (\{c_1, c_2, \dots, c_s\} \subseteq \{j_1, j_2, \dots, j_k\})$ 并将其放置在攻击树的必要条件分支下,若 $N_e = \varphi$,则继续步骤 4。

④计算 $q_x = Q(b_j | B_{c_1}^x, B_{c_2}^x, \dots, B_{c_s}^x)$, 若 $100\% - X \leq q_x \leq 100\% + X$ 且 X 为误差范围内的值,则第 x 棵攻击树构建完成,否则继续步骤 5。

⑤对 $q_x = Q(b_j | B_{c_1}^x, B_{c_2}^x, \dots, B_{c_s}^x, B_{c_{s+1}}^x, B_{c_{s+2}}^x, \dots, B_{c_s}^x)$ 实施重新计算,若符合 $100\% - X \leq q_x \leq 100\% + X$ 且 X 为误差范围内的值的条件,那么第 x 棵攻击树构建完成,否则继续步骤 5;若该步骤执

行次数超出预定值,那么继续步骤 4。

步骤 4:结束。

步骤 5:在由上述步骤所获取的子集中,根据“逐步向前选择”的方法增添属性,得到 $T_e = \{B_{c_{s+1}}^r, B_{c_{s+2}}^2, \dots, B_{c_s}^n\}$, 将其放置在充分条件分支下,然后转为步骤 3 中的⑤。

在完成上述计算后,可以将攻击类型分为两种,一种为可以直接从源数据中获取的攻击;另一种为不能直接从源数据中获取的攻击。所以第一类攻击可以直接从源数据 E 中获取其产生的必要条件,即 $N_e \neq \varphi$;而第二类攻击则与第一类攻击相反,即 $N_e = \varphi$ 。因此,所用算法必须在有限步骤内实施收敛;在复杂度上,算法属于 $P(|E|)$,其中, $|E|$ 代表 E 中所含总样本数,综上所述,算法类似于多项式。

1.2 节点特征提取

为了更好地实现攻击树节点特征定位的目的,在构建出攻击树模型后,还需要对其实施节点特征提取。

实施节点特征提取所利用的深度学习网络是由卷积神经网络和径向基神经网络构建而成的^[6-7]。后者的优势主要体现在其对于输入输出产生的显著非线性映射,在前向网络内其映射效果最为突出,因此在利用深度学习网络提取节点特征时,可以得到其网络模型如图 3 所示。



图 3 深度学习网络结构

在图 3 所示的网络结构内,卷积神经网络由多个 D 层与 T 层构成^[8],位于平面内的神经元具有各自独立的特性。 D 为特征提取层,其中各个神经元输入均与上层连接,以此获取攻击树节点的关键

特征;在降采样层中包含多个映射平面,各平面分别对上层特征码的局部特征实施提取或对其特征实施二次提取。为了达到逐级获取到攻击树节点本质特征的目的,可以将数据作为图 3 中的最底层数据输入。

设 $Y_{j,k}$ 代表各卷积层特征,其中 j 为卷积层, k 为特征,由此可以得到 $Y_{j,k}$ 数学模型为:

$$Y_{j,k} = g\left(\sum_{j \in N_k} c_{j,k} + Y_{j-1,k} x_{u,j,k}\right) \quad (1)$$

其中, N_k 表示 $Y_{j-1,k}$ 的集合, $x_{u,j,k}$ 代表 $Y_{j,k}$ 的权值, $c_{j,k}$ 为卷积特征图的偏置。

上层输入 O 个特征在降采样层均有对应的映射结果,想要对 u 降采样层的 k 特征实施表述,可以利用下式实现:

$$Y_{u,k} = g[c_{u,k} + \gamma_{u,k} \text{down}(Y_{j-1,k})] \quad (2)$$

其中, $\text{down}(\cdot)$ 代表降采样函数, $\gamma_{u,k}$ 表示 $Y_{u,k}$ 的位移乘性偏置向量, $c_{u,k}$ 为 $Y_{u,k}$ 的加性偏置向量。

在深度学习网络训练过程中^[9], $x_{u,j,k}$ 、 $c_{u,k}$ 和 $\gamma_{u,k}$ 都会直接影响最终的攻击树节点特征提取结果,所以需要利用量子粒子群算法对深度学习网络实施优化。

量子粒子群算法中的粒子出去量子空间中,设 O 为在空间内寻优解的粒子个数,并用 $Y_j = (y_{j1}, y_{j1}, \dots, y_{jE})$ 代表 E 维空间内的 j 的位置,用 $qc_j = (qc_{j,1}, qc_{j,2}, \dots, qc_{j,E})$ 代表 j 每次迭代时的最优位置,用 $hc_j = (hc_{j,1}, hc_{j,2}, \dots, hc_{j,E})$ 代表整个粒子群的全局最优位置。可以将波函数 $\varphi(y)$ 作为依据,从而引申出空间粒子状态,利用 $|\varphi(s)|^2$ 表述在 s 处粒子出现的概率 $H(s)$,若 $H(s)$ 符合归一化标准,则可以得到:

$$\int_y f(x) dx |\varphi(s)|^2 = f_y f(x) dx |H(s)|^2 = 1 \quad (3)$$

在式(3)的基础上可以利用下式对粒子在 σ 势阱内的定态薛定谔方程做出表述^[10]:

$$\frac{2n}{i^2} \varphi[\eta \text{ 其不意 } \sigma(Y - q_j) + F] + \frac{e\varphi}{(Y - q_j)^2} = 0 \quad (4)$$

其中, n 为粒子质量, $Y - q_j$ 表示能量本征值, i 为普朗克常量, F 为粒子能量, e 为 j 所在维度, $e = 1, 2, \dots, E$ 。经上述计算,可以得到波函数的数学模型如下所示:

$$\begin{cases} \varphi(s) = \frac{1}{\sqrt{M}} f^{-\frac{1}{M}} \\ M = i^2 n \eta = \frac{1}{\varphi} \end{cases} \quad (5)$$

假定, $qc_{je}(u)$ 为 e 维 j 在 u 次迭代下的势阱, 那么 j 在 $u+1$ 次迭代时的 $\varphi(y)$ 为:

$$\varphi(y_{je}(u+1)) = \frac{1}{\sqrt{M_{je}(u)}} \exp \left[-\frac{|x_{je}(u+1) - qc_{je}(u)|}{M_{je}(u)} \right] \quad (6)$$

粒子的位姿实施随机采样, 可以得到在 $u+1$ 次迭代时 j 在 e 维的分量 $y_{je}(u+1)$ 。由此可以得到势阱特征长度 $M_{je}(u)$ 的数学模型为:

$$M_{je}(u) = 2\beta(u) |n_e(u) - y_{je}(u)| \quad (7)$$

其中, β 为收缩扩展系数, n 代表平均最优位置, 并且其可以用下式获取:

$$n(u) = \frac{1}{O} \sum_{j=1}^O qc_j(u) = \left[\frac{1}{O} \sum_{j=1}^O qc_{j1}(u), \frac{1}{O} \sum_{j=1}^O qc_{j2}(u), \dots, \frac{1}{O} \sum_{j=1}^O qc_{jE}(u) \right] \quad (8)$$

利用上述算法优化深度学习网络模型, 结合对比散度算法, 可以得到在上升迭代过程中参数 $x_{u,j,k}$ 、 $\varphi_{u,k}$ 和 $c_{u,k}$ 等的偏导数 Δx_{jk} 、 $\Delta \varphi_j$ 和 Δc_k , 可以用下式表达其更新过程:

$$\begin{cases} x_{jk}^{u+1} = x_{jk}^u + \theta_x \Delta x_{jk} \\ \varphi_j^{u+1} = \varphi_j^u + \theta_\varphi \Delta \varphi_j \\ c_k^{u+1} = c_k^u + \theta_c \Delta c_k \end{cases} \quad (9)$$

式中: θ_x 、 θ_φ 、 θ_c 代表学习率, u 表示迭代次数。

在利用量子粒子群算法获取粒子的适应度 G 时, 可以根据深度学习网络产生的误差来确定, 其数学模型如下所示:

$$G = \frac{1}{O} \sum_{j=1}^O (t_j - v_j)^2 \quad (10)$$

其中, t_j 代表攻击树的数据样本 j 的重构值, v_j 代表数据样本 j 的标签值。当深度学习网络产生的误差值越小, 得到的攻击树节点特征效果就越好, 其定位准确率就会越高。

1.3 节点特征定位

在得到攻击树节点特征之后, 便可以根据其特征实现攻击树节点定位。在此利用 DV-Hop 方法实施对于攻击树的节点特征定位。

DV-Hop 算法步骤: ①利用距离矢量交换协议实现节点间的信息交换, 以此获取节点与导标节点间的跳距。②在得到上述信息后, 求取网络平均每跳距离, 之后将其看作校正值应用于网络中。③实行三角测量定位法。

设导标节点 (Y_j, Z_j) 的校正量 d_j 表示如下:

$$\begin{cases} d_j = \frac{\sum \sqrt{(Z_j - Z_k)^2 + (Y_j - Y_k)^2}}{\sum i_j} \\ j \neq k, \forall k \end{cases} \quad (11)$$

其中, (Y_k, Z_k) 表示 (Y_j, Z_j) 的坐标, i_j 则当前节点至其附近节点的跳数。

如图 4 所示, 设 B 代表需定位的节点特征, 在 M_1 、 M_2 和 M_3 之间的位置距离以及跳数均为已知量的情况下^[11], 可以得到 M_2 的校正值为 $\frac{40+75}{2+5} = 16.43$ 。假定节点特征 B 的校正值从 M_2 中取得, 那么 B 与 M_1 、 M_2 和 M_3 之间的距离分别为 $M_1 - 3 \times 16.43$ 、 $M_2 - 2 \times 2 \times 16.43$ 、 $M_3 - 3 \times 16.43$, 之后在此基础上利用三角测量定位的方法对节点特征 B 实施定位。

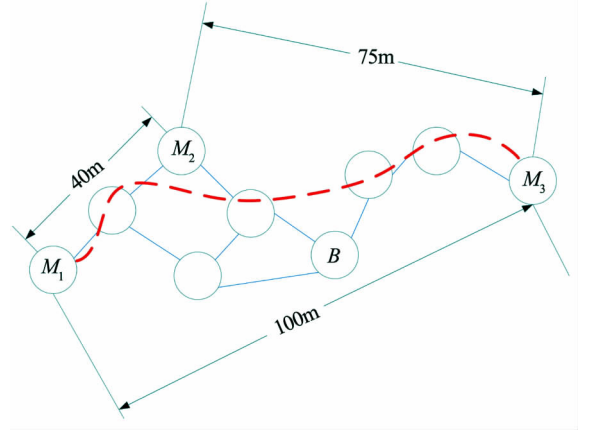


图 4 节点特征定位示例

2 实验与分析

为了验证 IPv6 网络攻击事件溯源中的攻击树节点特征定位的整体有效性, 需对其展开测试。为验证本文所提方法在 IPv6 网络攻击事件溯源中的攻击树节点特征定位中的应用效果, 采用 MATLAB 仿真平台搭建短距离 D2D 无线 IPv6 通信网络, 该网络为单基站, 网络中用户分布均匀, 网络的仿真详细参数见表 1。

表 1 网络仿真详细参数

参数名称	取值
网络覆盖半径/m	50
用户数量/人	45
发射功率/dBm	25
最小速率/b. s. Hz	1~15
路径损耗/dB	53
链路间距离/m	5~30
噪声功率/dBm	-110
信噪比/dB	6

对本文所提方法实行相关测试前,需先对网络中基站允许的干扰门限值 I_0 实行设定,该值的设定对于网络的通信效果存在直接影响。由于用户最大发射功率的逐渐增加,导致网络中信号接收用户的干扰功率结果增加,因此采用本文方法将接收用户干扰功率的最小值与 I_0 值结果对应,作为设定干扰门限结果,见图 5。

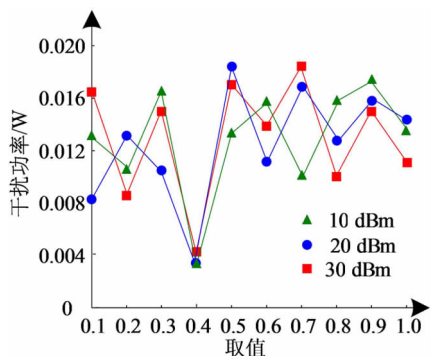


图 5 干扰门限值测试结果

依据图 5 结果得出:当用户的最大发生功率发生变化后,随着门限取值的逐渐增加,信号接收用户的干扰功率发生明显的波动变化。最大发射功率分别为 10 dBm、20 dBm、30 dBm 时,干扰功率的最小值分别为 0.004 W、0.0042 W、0.0038 W、0.0041 W,其对应的干扰门限值 I_0 的取值为 0.4,因此确定干扰门限值 I_0 的取值为 0.4,并用于后续实验中。

本次实验采用 MATLAB 2012 软件,在 CIC-IDS2017 数据集 (<https://www.unb.ca/cic/datasets/ids-2017.html>) 中随机选取 35 个节点数据,并利用所提方法,文献[1]、文献[2]和文献[3]方法对其攻击树节点特征实施定位,为更准确地对其定位误差实施描述,可利用下式实施计算:

$$\text{AverageError} = 100\% \times \frac{\sum_{j=1}^O \sqrt{(\bar{z}_j - z_j)^2 + (\bar{y}_j - y_j)^2}}{O \times S} \quad (12)$$

其中, $(\bar{y}_j, \bar{z}_j)$ 为估计坐标,而实际坐标为 (y_j, z_j) , O 为未知节点数, S 代表通信半径。

在上式的基础上得到所提方法,文献[1]、文献[2]和文献[3]方法对节点特征实施提取,得到的结果误差如图 6 所示。

由图 6 可以看出,文献[1]、文献[2]和文献[3]方法的节点定位误差较大,其中,文献[3]方法的最大定位误差超过 1 m。而所提方法的定位误差较

为稳定,均维持在 0.6 m 以下,而且低于对比方法的误差,因此所提方法得到的定位结果更接近于真实情况,可以更精确地实现攻击树的节点特征定位。

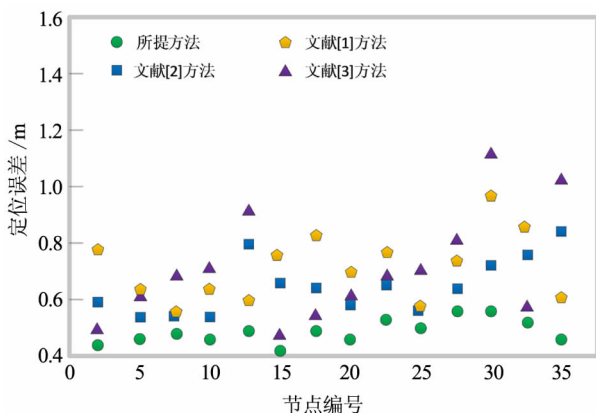


图 6 各方法定位误差对比

在得到各方法节点定位误差后,还需要对所提方法,文献[1]、文献[2]和文献[3]方法的在单位节点内定位效率展开对比,取七组不同环境下单位时间内各方法所定位的节点数,在单位时间内处理节点个数越多,表示该方法的处理效率越高。得到的结果如图 7 所示。

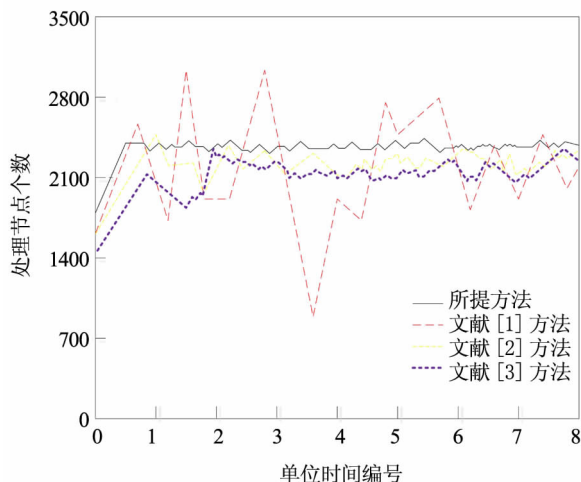


图 7 各方法节点定位效率

由图 7 可以看出,各方法在不同环境下的单位时间内所定位的攻击树节点数据都存在不同程度的波动,虽然文献[1]方法在其中几个环境节点下的定位效率高,但其存在较大的波动。文献[2]和文献[3]方法的处理节点个数分别在 2300 和 2100 左右,而所提方法处理的节点个数维持在 2400 个左右,高于文献[1]方法、文献[2]和文献[3]方法。因此从整体来看,所提方法的定位效率以及定位稳

定性更好,更适用于攻击树节点特征定位的实际应用。

在实施攻击树的节点特征定位时,不仅仅需要对其误差及性能要求较高,还需要对其在定位过程中的能耗展开对比,得到的结果如下所示:

表 2 各方法能耗性能对比/KJ

网络负载	0.2	0.4	0.6	0.8	1.0
所提方法	2.24	3.25	4.98	5.43	6.77
文献[1]方法	2.27	4.58	6.64	10.89	15.73
文献[2]方法	3.40	5.76	11.45	15.77	23.62
文献[3]方法	3.12	4.92	10.32	14.99	22.46

由表 2 可以看出,随着网络负载的增加,四种方法对攻击树节点特征开展定位所需的能耗也随之提高。文献[1]、文献[2]和文献[3]方法的最高能耗分别为 15.73 kJ、23.62 kJ 和 22.46 kJ,而所提方法的最高能耗为 6.77 kJ,低于对比方法。因为所提方法增加了节点特征提取的环节,大大降低了定位过程所需能耗,因此在定位过程中能耗变化较为稳定,消耗能量也更低,由此表明所提方法的能耗优于其他方法。

3 结 论

在现代网络发展的过程中,IPv6 网络协议逐渐成为主力,但是其仍然存在各类安全隐患。为及时发现并预防各类攻击的生成,需要对过往攻击事件中的攻击树节点特征实施定位。目前存在的各类节点特征定位方法中均存在能耗高、定位误差

大、在不同环境下单位时间内的定位稳定性较差且定位效率较低的问题,为此提出 IPv6 网络攻击事件溯源中的攻击树节点特征定位,该方法首先还原出事件的攻击树模型,之后对于其特征实施精确提取并展开定位,很好地解决了目前方法存在的问题,为网络协议的安全推广及日后更新提供保障。

参考文献

[1] 刘浩然,覃玉华,邓玉静,等. 基于双层修正无迹卡尔曼的水下节点定位算法[J]. 仪器仪表学报,2020,41(3):142—149.

[2] 余修武,张可,刘永,等. 基于反向学习的群居蜘蛛优化 WSN 节点定位算法[J]. 控制与决策,2021,36(10):2459—2466.

[3] 赫佳星. 基于几何原理的无需测距传感器节点定位[J]. 计算机工程与设计,2021,42(10):2718—2724.

[4] 潘刚,米士超,郭荣华,等. 基于攻击树和 CVSS 的网络攻击效果评估方法[J]. 电子技术应用,2022,48(4):76—80.

[5] 杜金莲,孙鹏飞,金雪云. 一种用于威胁检测的反目标攻击树模型[J]. 计算机科学,2021,48(S1):468—476.

[6] 章杜锡,谢宏,李力. 基于神经网络的变电站异常网络流量识别[J]. 信息技术,2020,44(5):140—144.

[7] 于怀征. 基于树分类器神经网络的雷暴预测方法[J]. 信息技术,2020,44(7):17—22.

[8] 王丽君,史二娜. 基于卷积神经网络的智能交通信号控制研究[J]. 信息技术,2020,44(10):56—60.

[9] 宋蓓蓓,马穗娜,何帆,等. Res2-Unet 深度学习教育网络的 RGB-高光谱图像重建[J]. 光学精密工程,2022,30(13):1606—1619.

[10] 雷一凡,宋刚,高成勇,等. 基于非线性薛定谔方程的时间序列去噪[J]. 计算机辅助设计与图形学学报,2021,33(8):1202—1209.

[11] 任克强,潘翠敏. 融合 RSSI 跳数量化与误差修正的 DV-Hop 改进算法[J]. 传感技术学报,2020,33(5):718—724.