Sensor Network Cache Attack Defense Based on Alarm Attribute Clustering*

LIU Haisong[1], NI Zhen[1], GUO Hongjian[2]*
(1.Informatization Construction and Management Office, Nanjing Xiaozhuang University, Nanjing Jiangsu 211171, China;
2.School of Computer Science, Nanjing Audit University, Nanjing Jiangsu 211815, China)

Abstract: Sensor network high-speed cache attackers can crack passwords or obtain sensitive data by simply monitoring the response time of shared resources. In order to improve the network's defense capabilities, a sensor network cache attack defense based on alarm attribute clustering is proposed. Firstly, the information entropy filtering method is used to remove false alarms of sensor network cache attacks and improve data quality. Secondly, the dynamic time threshold clustering method is used to identify alarms belonging to the same attack scenario, providing a basis for the formulation of defense strategies. Finally, based on the analysis of the relationship between attack patterns and content request locality in the clustering results of content request locality, a high-speed cache attack defense strategy for sensor networks is designed and implemented to effectively resist cache pollution attacks and ensure the normal operation of sensor networks. The simulation results show that the proposed method maintains an attack resistance ability of over 90%, and the damage ratio has increased from 10% to 28%, indicating high attack defense capability and efficiency.

Key words: sensor network; high-speed; cache attack defense; alarm attribute clustering; information entropy filtering; removal of false alarms; dynamic time threshold
EEACC: 7230 **doi**: 10.3969/j.issn.1004-1699.2026.09.025

基于告警属性聚类的传感网络高速缓存攻击防御*

刘海松[1], 倪 震[1], 郭红建[2]*
(1.南京晓庄学院信息化建设与管理办公室, 江苏 南京 211171; 2.南京审计大学计算机学院, 江苏 南京 211815)

摘 要: 传感网络高速缓存攻击者仅通过监控共享资源的反应时间等即可破解密码或获取敏感数据, 为了提高网络的防御能力, 提出基于告警属性聚类的传感网络高速缓存攻击防御。首先, 通过信息熵过滤方法去除传感网络高速缓存攻击误告警, 提高数据质量; 其次, 利用动态时间阈值聚类方法识别出属于同一攻击场景的告警, 为防御策略的制定提供基础; 最后, 基于内容请求局部性分析聚类结果中攻击模式与内容请求局部性之间的关系, 设计并实施传感网络高速缓存攻击防御策略, 有效抵御缓存污染攻击, 保障传感网络的正常运行。仿真结果表明, 所提方法攻击抵御能力保持在 90%以上, 损害比从 10%提升至 28%, 说明具有较高的攻击防御能力和效率。

关键词: 传感网络; 高速缓存攻击防御; 告警属性聚类; 信息熵过滤; 误告警去除; 动态时间阈值
中图分类号: TP393 **文献标识码**: A **文章编号**: 1004-1699(2026)09-2124-08

传感网络[1]作为物理空间与数字空间交互的关键枢纽, 广泛渗透于各类智能领域, 其深度应用极大地推动了社会信息化和智能化的发展。然而, 随着传感网络规模的日益扩大和复杂性的增加, 其面临的安全威胁也日益严峻, 其中高速缓存攻击作为一种隐蔽性强、影响范围广的恶意手段, 正逐渐成为传感网络安全防护中的一大挑战。高速缓存攻击的独特之处在于, 攻击者无需直接访问或修改受害系统的内部核心状态, 而是巧妙地利用传感网络中普遍存在的共享资源, 如高速缓存、内存访问模式等, 通过监控这些资源的使用情况及其产生的微妙反应时间差异, 间接地获取受害系统的敏感信息或破解加密数据。这种攻击方式不仅难以被传统安全机制直接察觉, 还因其非侵入性和低噪声特性, 使得防御工作变得尤为复杂和困难。与此同时, 传感网络中海量的传感器节点持续不断地生成着庞大的原始安全数据流。这些数据中蕴含着丰富的网络状态和异常行为信息, 但同时也夹杂着大量的冗余、无关甚至误导性的数据, 直接分析这些数据以发现潜在的安全威胁难度较高, 且效率低下容易遗漏关键信息。

项目来源: 江苏省产学研合作基金项目(BY20230619); 江苏省产学研合作基金项目(BY20230589)
收稿日期: 2024-08-28 **修改日期**: 2024-09-27

为此,针对相关的缓存攻击防御方法,引起了学者的广泛关注。王少清等[2]通过监测分析传感网络中微架构(缓存)的状态变化,通过智能算法识别潜在的异常缓存行为;动态调整缓存分配与管理策略,以阻止攻击者利用缓存污染、缓存耗尽等手段。但该算法复杂度较高,在识别潜在异常缓存行为时伴随误报和漏报的风险,可能干扰网络的正常运转并对网络性能造成损害,还可能对算法的精确度和稳定性产生影响。Choi[3]设计了安全感知分区策略。动态调整缓存分区,依据写入计数避免缓存刷新,消除驱逐缓存行;在分区变更时评估新状态写入量,优化缓存使用,减少刷新开销,提升缓存寿命与写持久性,增强传感网络对高速缓存攻击的防御能力。该方法基于写入计数来避免缓存刷新和消除驱逐缓存行,但该策略因复杂的动态分区调整和状态评估,带来较高计算成本,影响系统性能。同时,写入计数的准确性和及时性对分区调整效果至关重要,且其普适性有限,难以全面适应多样化的传感网络和缓存需求。Ma 等[4]构建智能主动防御机制,建模为随机博弈,运用 PPO 优化深度强化学习,设计自检组件应对缓存状态变化;通过模拟不同攻击策略训练,提升收敛速度与稳定性,使原型系统有效应对传感网络高速缓存攻击。但在资源受限的传感网络环境中,该方法通常需要大量的计算资源,影响系统的实时性和响应速度,导致系统在面对突发攻击时无法迅速做出有效应对。邢光林等[5]提出基于包标记的高速缓存攻击防御方法。在 Interest 包中嵌入边缘节点标识,检测攻击时识别恶意数据特征,确定恶意源前缀;向下游发送溯源包定位攻击源,并动态调整缓存策略隔离攻击。该方法主要依赖于包标记和溯源技术,但在网络设备上实施复杂的配置可能增加额外成本与复杂性,同时,面对伪装合法流量的攻击,其检测能力受限,且溯源过程耗时,可能影响攻击定位的准确性和时效性。

综上所述,虽然这些方法都在提升传感网络对高速缓存攻击的防御能力方面有贡献,但也都存在计算复杂度、实时性、准确性、普适性和额外成本等方面的问题。因此,本研究提出基于告警属性聚类的传感网络高速缓存攻击防御。通过信息熵过滤技术降低误告警的数量;然后,用动态时间阈值聚类算法将告警分类,提高告警信息的有效性;最后,分析告警信息类别,并设计针对性的防御策略,以提高对抗缓存污染攻击的准确度,确保传感网络稳定高效运行。

1　误告警过滤预处理

传感网络中传感器节点众多,它们持续不断地生成大量原始数据,其中包括了真实的安全事件数据和误告警信息。误告警过滤预处理能够有效剔除这些冗余和噪声数据,使得后续的数据分析更加聚焦于有价值的信息,从而提高数据分析的效率和准确性。因此,首先从特征提取和误告警去除两个步骤,减少误告警的数量,完善告警质量。

步骤 1　特征提取

为了应对安全威胁[6],首先提取传感网络关键事件特征:传感网络告警周期、告警密度、源 IP 至目的 IP 数量、攻击源威胁度。针对这些特征展开离散化处理,转换为离散型随机变量,以便分析。计算每个特征值的概率分布,并基于该分布计算每个特征的熵值。

①告警周期值的熵

将告警数据按源 IP 和告警类型分组,形成时间序列$\{x(t):t\in M\}$,对其分别展开自相关分析[7-8]后获取间隔 m 分钟告警发生的关系度量:

$$R(m)=\frac{1}{M-m+1}\sum_{n=0}^{M-m}[x(n)x(n+m)] \tag{1}$$

式中:$0\leqslant m\leqslant n$,M 为统计分钟数,n 为初始分钟数,对上式应用傅里叶变换得到各个频率的能量密度函数 $\mathrm{PSD}(g)=|\mathrm{DFTR}(m),g|$,$0<g\leqslant 0.5$,从中寻找最大频率点,以“周期=1/频率”的计算方式求得告警周期值 $A(A_i)=A_i^M/M$,其中 A_i^M 为周期值为 A_i 的告警数量。根据该公式得到告警 S_i 的告警周期 $A_i(A_i\in A)$ 相应的熵为:

$$H(A_i)=-A(A_i)\log_2 R(m) \tag{2}$$

②告警密度的熵

通过计算每个告警 S_i 在特定时间窗口 T 内的相关告警数,并除以时间窗口长度得到告警密度 F_i,公式为:

$$F_i=\frac{\mathrm{count}(S_i\times\mathrm{relevant})}{T_{i_{\max}}-T_{i_{\min}}},\quad |T_{i_{\max}}-T_{i_{\min}}|<T \tag{3}$$

式中:$\mathrm{count}(S_i\times\mathrm{relevant})$ 为告警数量;$T_{i_{\max}}$、$T_{i_{\min}}$ 分别为最大、最小分布时间差。告警 S_i 的告警密度 F_i $(F_i\in F)$ 所对应的熵表示为:

$$H(F_i)=-A(F_i)\log_2 A(F_i) \tag{4}$$

③源 IP 对应目的 IP 数的熵

设加入告警源 IP 对应的目的 IP 为 Q_i,数量为 q_i^N,源 IP 数量为 N,则其对应概率 $A(Q_i)=q_i^N/N$,由此得到告警的源 IP 对应目的 IP 数 $Q_i(Q_i\in Q)$ 相应的熵:

$$H(Q_i)=-A(Q_i)\log_2 A(Q_i) \tag{5}$$

④攻击源威胁度的熵

考虑告警优先级和告警类型多样性,计算每个源 IP 在给定时间内的总威胁度,并根据告警优先级加权求和,结合告警类型的多样性展开调整,假设攻击源威胁度为 T_i 的告警数量为 t_i^N,则威胁度 $T_i(T_i\in T)$ 的概率公式为 $A(T_i)=t_i^N/N$,对应的熵为:

$$H(T_i)=-A(T_i)\log_2 A(T_i) \tag{6}$$

步骤 2 误告警去除

①将上述待检测告警 S_i 的特征信息熵组合,形成序列 $G(i)=[H(F_i),H(A_i),H(Q_i),H(T_i)]$,对其展开归一化处理得到:

$$\begin{cases}G'(i)=[H'(F_i),H'(A_i),H'(Q_i),H'(T_i)]\\H'=H/H_{\text{sum}}\\H_{\text{sum}}=H(F_i)+H(A_i)+H(Q_i)+H(T_i)\end{cases} \tag{7}$$

②对误告警特征信息熵 $G(f)$ 实施归一化[9]得到 $G'(f)$,计算 $G'(i)$ 和 $G'(f)$ 的互雷尼信息熵:

$$I_{0.5}(G'(i),G'(f))=2\log_2\sum_d\sqrt{a(i)a(f)} \tag{8}$$

式中:d 为告警事件特征维数;$a(i)$、$a(f)$ 为 $G'(i)$、$G'(f)$ 的概率分布函数。

③将计算出的互雷尼信息熵 $I_{0.5}(G'(i),G'(f))$ 与基础阈值 λ 实施比较。若 $I_{0.5}(G'(i),G'(f))<\lambda$,则认为待检测告警 S_i 是误告警,作过滤处理。反之为真正的传感网络高速缓存攻击告警。根据以上步骤,生成并输出误告警过滤后的最终告警列表 $G''(i)$,以便后续处理。

2 传感网络高速缓存攻击防御

2.1 告警属性聚类

经过误告警过滤的筛选后,仍可能残留大量内容、时间或来源上高度相似的告警信息[10],这些告警虽被独立处理,却显著增加了后续分析与管理的复杂性。为此,引入了基于动态时间阈值的告警属性聚类策略,旨在将原本分散却内在关联的告警特征进行有效整合,形成超级告警。每一超级告警均作为其所涵盖告警类别的标志性代表,不仅简化了告警信息的呈现,还极大地促进了告警数据的聚合与价值提升。

步骤 1 数据准备

输入数据经过误告警过滤后的告警记录集合,记作日志 $\text{RA}=\{\text{ra}_1,\text{ra}_2,\cdots,\text{ra}_n\}$,$\{\text{ra}_i,i=0,1,2,\cdots,n\}$,$\text{RA}\in G''(i)$。其中每个 ra_i 为一个告警记录,n 为告警总数,明确超级告警的结构 SA = {srcIP, srcPort, dstIP, dstPort, attackTypeList, warning-Level, count, startTime, endTime},其中 count 为聚合告警数量; attackTypeList 为聚合后的攻击类型序列; startTime 为告警开始时间; endTime 为告警结束时间。

步骤 2 告警属性相似度计算

确定用于计算相似度的告警属性[time, srcIP, srcPort, dstIP, dstPort, attackType]并利用下式描述告警相似度:

$$D(\text{ra}_i,\text{ra}_j)=\sum_k w_k s(\text{ra}_{ik},\text{ra}_{jk})\Big/\sum_k w_k \tag{9}$$

式中:k 为告警属性索引;ra_{ik} 和 ra_{jk} 为告警第 k 属性值;w_k 为告警的第 k 属性权重。

步骤 3 IP 地址相似度计算

检查两条告警 ra_i、ra_j 的源 IP srcIP(sip)和目的 IP dstIP(dip)是否有交集,若有交集,则 IP 地址相似度较高,可以赋予较高的相似度分数,计算公式为:

$$s_{\text{IP}}=\begin{cases}1 & \text{sip}_{\text{ra}_i}=\text{sip}_{\text{ra}_j}\ \|\ \text{sip}_{\text{ra}_i}=\text{dip}_{\text{ra}_j}\\1 & \text{dip}_{\text{ra}_i}=\text{sip}_{\text{ra}_j}\ \|\ \text{dip}_{\text{ra}_i}=\text{dip}_{\text{ra}_j}\\0 & \text{other}\end{cases} \tag{10}$$

步骤 4 端口相似度计算

使用端口的概念层次树[11]来分类端口。利用层次树中两个端口到它们共同祖先节点的最长路径长度 $l(a_{\text{ra}_i},a_{\text{ra}_j})$ 作为衡量这两个端口相似程度的一个指标,其计算公式为:

$$s_{\text{port}}=\begin{cases}1 & a_{\text{ra}_i}=a_{\text{ra}_j}\\1-l(a_{\text{ra}_i},a_{\text{ra}_j})/3 & a_{\text{ra}_i}\neq a_{\text{ra}_j}\end{cases} \tag{11}$$

步骤 5 攻击类型相似度计算

直接检查两个告警的攻击类型(attackType)是否一致,若完全一致,则设定它们的相似度为最高值 1;若不同,则相似度被判定为最低值 0。

步骤 6 动态时间阈值计算

初始化一个动态时间阈值 T,用于判断告警是否属于同一时间窗口内。根据新接收的告警动态更新时间阈值 T,告警日志时间差 τ_i,使其更接近于实际攻击的持续时间。更新公式基于告警日志时间差的均值 τ_{avg} 和更新系数 $\sigma^*(\tau)$,即:

$$\begin{cases}T=\tau_{\text{avg}}+\tau_{\text{avg}}\times\sigma^*(\tau)\\\tau_{\text{avg}}=\sum_i^{n-1}\tau_i/(n-1),i=1,2,\cdots,n-1\\\sigma^*(\tau)=\sigma(\tau)/\tau_{\text{avg}}\\\sigma(\tau)=\sqrt{\sum_i^{n-1}\left[\sum_i^{n-1}\tau_i/(\tau_i-\tau_{\text{avg}})\right]/(n-1)}\end{cases} \tag{12}$$

步骤 7 告警属性聚类

创建一个空的超级告警集合 **SA**{},对于每个

告警 r_{a_i},执行以下步骤:

遍历 **SA** 中的每个超级告警 sa_i,计算 r_{a_i} 与 sa_i 的总体相似度,若 r_{a_i} 与某个 sa_i 的总体相似度超过某个阈值,则将 r_{a_i} 加入 sa_i(更新 attcakTypeList, count,startTime,endTime)。若没有找到相似度足够的超级告警,则创建一个新的超级告警 sa_i_new,将 r_{a_i} 作为其第一个成员。如果加入了新的告警到某个超级告警,根据新告警的时间信息更新该超级告警的动态时间阈值。遍历完所有告警后,**SA** 中将包含所有聚类后的超级告警 $\mathbf{SA}\{\mathrm{sa}_1,\mathrm{sa}_2,\cdots,\mathrm{sa}_m\}$,将其输出供后续攻击防御使用。

2.2　基于内容请求局部性的缓存污染攻击防御策略

缓存污染攻击以其隐蔽性著称,攻击者往往利用伪装或篡改告警信息的手法来掩盖其真实意图。即便经过聚类处理的告警信息,也可能因污染而误导为正常或关联告警,从而加剧了识别和防御的挑战性。为此,首先通过超级告警记录分析内容流行度与路径多样性,验证请求内容的实际流行度并评估请求来源的多样性。随后,结合以上内容分析,将告警记录融入 CPDP-CRL 缓存策略中,通过引入基于告警记录的调整因子动态调整各类内容的缓存概率,有效识别并防御缓存污染攻击,确保网络对合法内容的高效访问。

2.2.1　内容流行度与路径多样性分析

在传感网络高速缓存攻击防御的初步阶段,已得到了超级告警记录。记录中包含了请求的详细信息(请求次数和来源),通过将这些信息与全局或区域性的流行度数据库比对,可以验证被请求内容的流行度[12]是否与其实际流行度相符。通过计算每个请求经过的不同路径数目,可以评估内容的路径多样性,路径多样性低意味着请求来自少数固定源,增加了缓存污染的风险。若一个内容被频繁请求,但其请求路径却非常有限(如只来自少数 IP 或地理位置),那么这很可能是缓存污染攻击的迹象。

假设从一个用户到一个核心路由节点有 Q_{path} 条不同的路径,一个请求流行内容的兴趣包通过第 i 条路径到达核心路由节点的概率为 $p_i=\sum_{i=1}^{Q_{\mathrm{path}}}p_i=1$。在超级告警记录的支持下,估计这些概率值,并计算 y' 个兴趣包所经过的平均路径数目 $E[X']$:

$$\begin{cases}E[X']=Q_{\mathrm{path}}-\sum_{i=1}^{Q_{\mathrm{path}}}(1-p_i)\\ X'=\mathrm{SA}\{\mathrm{sa}_1,\mathrm{sa}_2,\cdots,\mathrm{sa}_m\}\sum_{i=1}^{Q_{\mathrm{path}}}x_i\end{cases}\tag{13}$$

式中:X' 为 y 个兴趣包经过的不同路径数目;x_i 为布尔变量,用于描述 y' 个兴趣包是否经过第 i 条路径,若经过 $x_i=1$,否则 $x_i=0$。内容流行度越高,请求该内容的兴趣包所经过的不同路径数目也越多。

2.2.2　缓存污染攻击防御

在传感网络环境中存在 W 个合法的内容提供源,它们共同发布了合法内容。这些内容被划分为 K 个类别,每个类别中均包含等量 $b=W/K$ 的内容文件。网络中的每个节点 x 被设计为能够存储 C 个内容文件,以支持内容的缓存与快速访问。所有节点接收到的内容请求,其兴趣包(即请求信号)的到达遵循泊松分布规律[13],且整个网络内所有节点 x 的请求总到达率被设定为 μ_x。这种设计确保了内容请求的高效管理与资源的合理分配。在当前时刻 t,对于节点 x 各接口 face 上第 k 类内容兴趣请求次数的变异系数 $\mathrm{CV}_x^t(k)$ 的计算公式如下:

$$\begin{cases}\mathrm{CV}_x^t(k)=\dfrac{\mathrm{SD}_x(k)}{\mathrm{mean}_x(k)}\\ \mathrm{SD}_x(k)=\sqrt{E[X']\sum_{\mathrm{face}\in\mathrm{Face}_x}^{|\mathrm{Face}_x|}(\mathrm{Num}_x^{\mathrm{face}}(k)-\mathrm{mean}_x(k))}\\ \mathrm{mean}_x(k)=\dfrac{1}{|\mathrm{Face}_x|}\sum_{\mathrm{face}\in\mathrm{Face}_x}^{|\mathrm{Face}_x|}\mathrm{Num}_x^{\mathrm{face}}(k)\end{cases}\tag{14}$$

式中:face 为任意接口号;$\mathrm{mean}_x(k)$ 和 $\mathrm{SD}_x(k)$ 分别为节点 x 各接口 face 上第 k 类内容兴趣请求次数的均值和标准差;$|\mathrm{Face}_x|$ 为节点 x 上的总接口数;$\mathrm{Num}_x^{\mathrm{face}}(k)$ 为内容请求数目。

为了更有效地防御缓存污染攻击,将超级告警记录中的信息融入 CPDP-CRL 策略中[14]。具体地,根据告警记录调整每类内容的缓存概率,特别是对那些被频繁请求但来源可疑或路径多样性低的内容。在原始 CPDP-CRL 策略的基础上,引入一个基于告警记录的调整因子 k^{α},该因子反映了第 k 类内容受缓存污染攻击的风险程度。缓存概率的计算公式修改为:

$$\begin{cases}P_{\mathrm{CPDP\text{-}CRL}}(k,t)=\dfrac{p_k}{1+\mathrm{e}^{\mathrm{CV}_x^t(k)/\rho}}\\ p_k=\dfrac{c}{k^{\alpha}},c=\left(\sum_{k=1}^{K}\dfrac{1}{k^{\alpha}}\right)^{-1}\end{cases}\tag{15}$$

式中:p_k 为第 k 类内容的请求概率;ρ 为概率调节参数。基于式(15),兴趣请求在接口上分布越均匀($\mathrm{CV}_x^t(k)$ 越小),反映兴趣广泛,缓存概率越大。反之,变异系数大则可能预示缓存污染,缓存概率降低。同时,内容流行度(请求概率 p_k)也影响缓存概率,低流行度内容缓存概率低,优化缓存使用。理想

状态下,兴趣请求完全均匀分布时,缓存概率接近$\frac{p_k}{2}$,有效抑制低流行度内容缓存,并结合 APA 策略优势,高效应对多种传感网络高速缓存污染攻击的防御。

至此,按照图 1 所示的流程,实现基于告警属性聚类的传感网络高速缓存攻击防御。

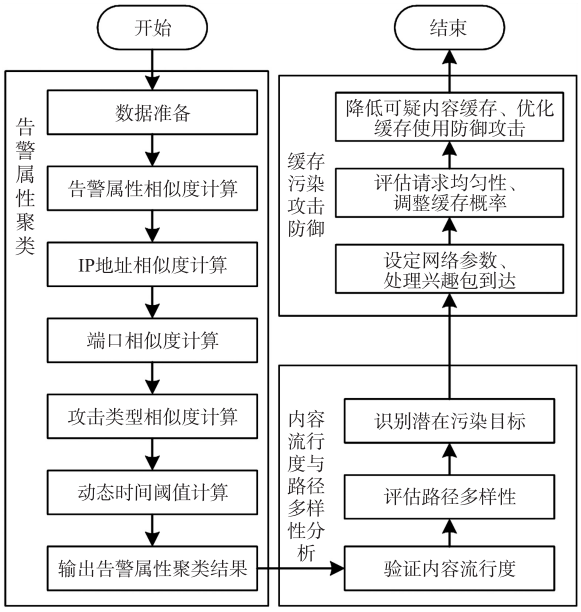

图 1 传感网络高速缓存攻击防御流程

所提方法通过信息熵过滤提升传感网络环境中的数据质量,去除误告警数据;利用动态时间阈值聚类识别传感网络中同一攻击场景的告警属性;基于内容请求局部性分析攻击模式与请求特征,设计并实施防御策略,有效抵御缓存污染,保障传感网络稳定运行。

3 仿真与分析

为了验证基于告警属性聚类的传感网络高速缓存攻击防御的整体有效性,需要对其展开仿真。

3.1 仿真环境与参数

仿真环境如图 2 所示。

图 2 所示环境中,攻击发起者坐在电脑前,其旁边摆放的显示器和笔记本电脑则暗示了这是一个日常的工作环境。从攻击发起者出发,虚线箭头直指右侧的多个计算机设备,这些设备构成了网络的核心部分,包括多台显示器和其对应的服务器。加粗实线箭头的指向不仅代表了数据传输的路径,更代表了潜在的攻击路径。这些箭头相互连接着右侧的计算机设备,展示了攻击在网络中的传播和扩散过程。

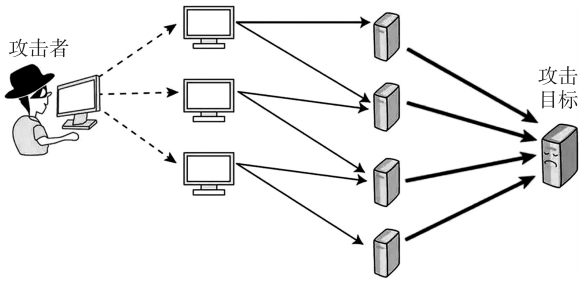

图 2 仿真环境图

基于污染内容数量和攻击强度,仿真将传感网络高速缓存污染攻击分为五类:轻量攻击(少量内容,低强度)、集中式攻击(少量内容,高强度)、分散式轻量攻击(大量内容,低强度)、分散式攻击(大量内容,高强度)和中量攻击(内容与强度均中等),并设定如下参数:①正常内容总数:10 000;②Zipf 分布参数(用于生成正常请求规律):1.2;③Interest Packet 发送频率:1 000 个/s;④节点缓存空间大小:500 个内容项;⑤仿真周期时长:60 s;⑥检测算法参数统计周期:10 个周期;⑦攻击开始周期:第 20 周期。

所提方法所需参数为:①信息熵过滤阈值:0.05;②最小聚类大小:5;③时间窗口:10 s;④相似度阈值:0.8;⑤局部性阈值:0.6;⑥分析周期:5 个周期。

3.2 所提方法的验证

为了验证所提方法的传感网络高速缓存攻击告警属性聚类效果,在仿真平台上实施了为期两周的高速缓存污染攻击模拟仿真,并采集了相应的告警数据。原始告警数据数量庞大,包含了许多周期性误告警,通过应用所提方法选用的误告警过滤机制后,有效减少了误告警数量,在此基础上采用所提方法选用的动态时间阈值告警属性聚类方法对误告警过滤后的告警属性记录展开聚类处理,聚合效果用告警聚合率评估,定义如下:

$$\eta=(Z-z)/z \tag{16}$$

式中:Z 为原始告警数量;z 为聚类后告警数量。聚合率是用于反映聚合算法去除重复和冗余告警的效率,聚合率越高,表明冗余告警的去除效果越明显,提供给下一步告警防御策略的告警数据质量也就越高。由此得到的结果如表 1 所示。

表 1 告警属性聚类前后对比结果

攻击类型	原始告警数量	误告警过滤后告警数量	聚类后告警数量	聚合率/%
轻量攻击	200	50	10	95.00
集中式攻击	150	30	6	96.00
分散式轻量攻击	300	70	14	95.30
分散式攻击	400	95	19	95.25
中量攻击	250	60	11	95.60
总计	1300	305	50	96.15

聚类结果显示,所提方法聚合率高达 95%,表明冗余告警的去除效果十分明显。这是因为所提方法通过信息熵过滤方法,有效去除了传感网络高速缓存攻击中的误告警。这一步骤极大地提高了后续分析的数据质量,减少了噪声和干扰,使得后续的攻击识别和防御策略能够基于更准确、更可靠的数据集进行。

考虑到所提方法使用的动态时间阈值聚类参数对聚类结果及传感网络高速缓存攻击防御效果具有一定影响,因此,分析时间弯曲窗口、相似度阈值和聚类最小规模三个参数不同组合(如表 2 所示)下,聚类结果的准确性和稳定性。

表 2　不同参数组合

参数组合编号	窗口大小	相似度阈值	聚类最小规模
1	5	0.7	3
2	5	0.8	5
3	5	0.9	7
4	10	0.7	3
5	10	0.8	5
6	10	0.9	7
7	15	0.7	3
8	15	0.8	5
9	15	0.9	7

图 3 所示结果为在表 2 不同参数组合下的轮廓系数、Davies-Bouldin Index、攻击识别率和误报率仿真结果。其中,轮廓系数是一种衡量聚类效果好坏的指标,其值范围在-1 到 1 之间。值越大表示聚类效果越好,即类内数据点越相似,类间数据点越不相似。Davies-Bouldin Index 用于评估聚类算法中各个聚类簇的分离程度。值越小,表示聚类簇之间的分离度越好,聚类效果越佳。攻击识别率是指在攻击场景下,能够正确识别出攻击行为的比例。高攻击识别率意味着所提方法能够更有效地检测出潜在的威胁。误报率是指在非攻击场景下,错误地将正常行为识别为攻击行为的比例。

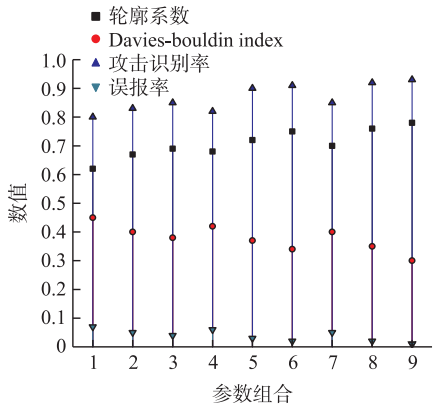

图 3　不同参数组合下的聚类性能

由图 3 可知,随着窗口大小、相似度阈值和聚类最小规模的增加,轮廓系数总体上升,Davies-Bouldin Index 下降,表明聚类质量提升,聚类内更紧密,聚类间更分离。同时,攻击识别率提升至 85%,误报率降至 5%左右,显示增加这些参数值能提高攻击识别能力并减少误判,从而增强方法的准确性和可靠性。综合考虑各个指标,参数组合 9(窗口大小为 15,相似度阈值为 0.9,聚类最小规模为 7)表现出较好的性能。在这个组合下,轮廓系数达到 0.78,Davies-Bouldin Index 为 0.30,攻击识别率为 93%,误报率为 1%。与其他参数组合相比,这个组合在聚类效果和攻击识别性能方面都表现得最为出色。因此,确定动态时间阈值方法最佳聚类参数为组合 9。

为验证所提方法的优越性,在仿真平台上进行了准确度仿真,并采集了相应的仿真数据,如图 4 所示。

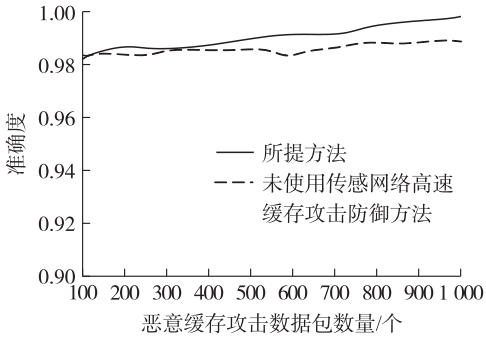

图 4　传感网络高速缓存攻击防御方法准确性结果

由图 4 可知,随着恶意缓存攻击数据包数量的增加,与未使用传感网络高速缓存攻击防御方法相比,所提方法的准确度较高,趋于 1,这表明所提方法在聚类结果的基础上进一步分析了攻击模式与内容请求局部性之间的关系,能够准确识别并拦截恶意缓存攻击数据包,为制定更加精准、有效的防御策略提供了重要依据。

为了验证所提方法在不同攻击场景下的可扩展性时,在 LDA 缓存污染攻击、FLA 缓存污染攻击和 DDoS 攻击场景下,以缓存污染检测率、缓存污染防御成功率和合法内容访问延迟为指标进行仿真。结果如表 3 所示。

表 3　所提方法的可扩展性分析

攻击场景	缓存污染检测率/%	缓存污染防御成功率/%	合法内容访问延迟/ms
LDA	95	88	50
FLA	92	90	55
DDoS	N/A	75	150

由表 3 可知,所提方法在 LDA 场景下,缓存污染检测率高达 95%,防御成功率为 88%,合法内容访问延迟保持在 50 ms 的较低水平。FLA 场景下,检测率为 92%,防御成功率提升至 90%,但访问延迟略增至 55 ms。而对于 DDoS 攻击,由于其主要目的非直接针对缓存污染,检测率不适用,防御成功率虽有 75%但受攻击复杂性和规模影响,合法内容访问延迟显著升高至 150 ms,显示出 DDoS 对整体网络性能的较大影响。综上可知,所提方法能够在多种攻击场景下工作,显示了其广泛的适用性。这表明方法具有一定的可扩展性,能够针对不同类型的攻击进行有效的检测和防御。

3.3 仿真结果分析

选取攻击抵御能力作为评估所提方法对高速缓存攻击的防御能力,并引入文献[3]安全感知分区策略和文献[5]包标记攻击防御方法作为对比方法,验证所提方法的优越性。攻击抵御能力值越高,进入网络的恶意缓存攻击数据包越少,对网络性能的影响越小,计算公式为:

$$\text{Attack_resisitance}=\frac{\text{InterestPacket}_{\text{dropped}}}{\text{InterestPacket}_{\text{total}}} \tag{17}$$

式中:$\text{InterestPacket}_{\text{dropped}}$为攻击者发出的包含恶意缓存攻击的数据包总数;$\text{InterestPacket}_{\text{total}}$为被丢弃未进入网络的恶意缓存攻击数据包数量。根据上述公式,在不同数量的恶意缓存攻击数据包下得到三种方法在该指标上的表现结果,结果如图 5 所示。

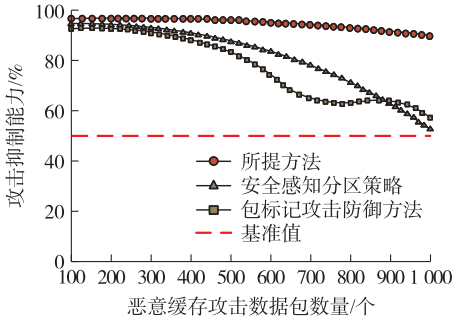

图 5 各方法攻击抵御能力

由图 5 可知,所提方法、安全感知分区策略和包标记攻击防御方法均在基准值(50%)以上,证实了三种方法具有较优的攻击抵御能力。其中,所提方法的攻击抵御能力一直保持在 90%以上,远超安全感知分区策略和包标记攻击防御方法。通过精准识别并有效拦截恶意缓存数据包,显著降低了网络受攻击风险,随着恶意缓存攻击数据包的增加,其攻击抵御能力也只出现略微降低。这是因为所提方法利用动态时间阈值聚类方法,获取了属于同一攻击场景的告警聚类结果。这一过程不仅提高了告警处理的效率,还使得防御策略能够针对具体的攻击场景进行定制化设计,增强了防御的针对性和有效性。相比之下,其他两种方法在应对复杂多变的缓存攻击时难以达到同等水平的防御效果,尤其是包标记攻击防御方法,还带有一定波动。因此,所提方法是提升传感网络安全性的更佳选择。

选用分散式攻击作为主要的传感网络高速缓存攻击类型,攻击持续时间为 50 s,记录时间变化下的响应时间,根据损害比=(攻击前响应时间-攻击时响应时间)/攻击前响应时间×100%的思路,在无防御措施的情况下的平均损害比为 50%。以此作为基准,分别应用所提方法、安全感知分区策略和包标记攻击防御方法展开攻击防御,记录每种方法时间变化下的损害比,结果如图 6 所示。

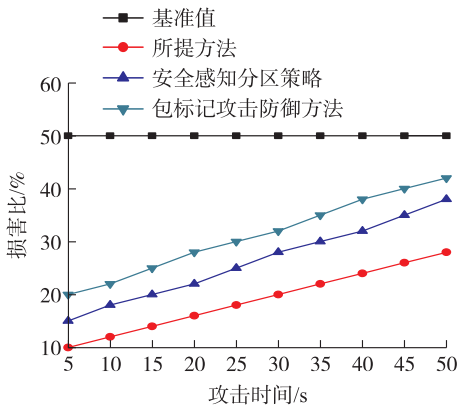

图 6 各方法损害比对比

由图 6 可以看出,随着时间从 5 s 增加至 50 s,所提方法损害比从 10%提升至 28%、安全感知分区策略从 15%提升至 38%,包标记攻击防御方法从 20%提升至 42%。所提方法在每个时间点上均表现最佳,10 s 时损害比仅为 12%,显著低于安全感知分区策略的 18%和包标记攻击防御的 22%,表明其在抵御损害上具有明显优势。安全感知分区策略虽不及所提方法,但防御效果优于包标记攻击防御,后者在各时间点上损害比均较高,如 30 s 时达 32%,显示出其防御能力相对较弱。由此可以看出所提方法防御传感网络高速缓存攻击方面的优势。

4 结论

所提方法通过信息熵过滤误告警,提升数据准确性;利用动态时间阈值聚类识别相同攻击场景告警;分析攻击模式与内容请求局部性关系,设计针对性防御策略,有效抵御缓存污染攻击,保证传感网络运行的稳定性。经验证,所提方法通过高效的信息

熵过滤和聚类分析,实现了 95%的聚合率,显著去除了冗余告警,提升了数据质量。在优化聚类参数组合后,显著提升了聚类质量,攻击识别率高达 93%,误报率低至 1%,验证了方法的高准确性和可靠性。在 LDA 和 FLA 攻击场景下,缓存污染检测率分别达到 95%和 92%,防御成功率分别为 88%和 90%,同时保持了较低的合法内容访问延迟。此外,所提方法在攻击抵御能力上远超其他方法,保持在 90%以上,且在长时间测试中损害比增长缓慢,表明其在抵御损害方面具有明显优势。综上所述,所提方法在不同攻击场景下均表现出色,具备高可扩展性和较优的攻击检测与防御能力。

参考文献:

[1] 周祖德,姚碧涛,谭跃刚,等. 光纤传感在制造领域应用的分析与思考[J]. 机械工程学报,2022,58(8):3-26.

[2] 王少清,赵有健,吕志远. exLCL:一种针对 spectre 攻击的防御方法[J]. 清华大学学报(自然科学版),2021,61(11):1221-1227.

[3] Choi J. A Method of Defense Against Cache Timing Attack in Non-Volatile Memory[J]. IEICE Electronics Express,2023,20(6):20220477.

[4] Ma T, Xu C, Yang S, et al. An Intelligent Proactive Defense Against the Client-Side DNS Cache Poisoning Attack via Self-Checking Deep Reinforcement Learning[J]. International Journal of Intelligent Systems, 2022, 37(10):8170-8197.

[5] 邢光林,陈璟,余俊乐,等. 命名数据网络中基于包标记的 Interest 泛洪攻击缓解研究[J]. 中南民族大学学报(自然科学版),2021,40(2):204-209.

[6] 李瑞. 制革企业集成自动化系统网络安全防护体系研究[J]. 中国皮革,2023,52(7):27-30.

[7] Zhang C,Lv W,Zhang P,et al. Multidimensional Spatial Autocorrelation Analysis and It's Application Based on Improved Moran's I[J]. Earth Science Informatics,2023,16(4):3355-3368.

[8] Parlak F,Pham D D,Spencer D A,et al. Sources of Residual Autocorrelation in Multiband Task fMRI and Strategies for Effective Mitigation[J]. Frontiers in Neuroscience,2023,16:1051424.

[9] 张立国,章玉鹏,金梅,等. 基于归一化注意力机制的特征自适应融合目标跟踪算法[J]. 计量学报,2023,44(9):1383-1389.

[10] 邓翠艳,姚旭清. 基于 DBSCAN 算法的告警数据聚类研究[J]. 太原理工大学学报,2021,52(1):111-116.

[11] 万润君,郭嗣琮,刘海涛,等. 适于高维数据的多标记学习层次树模型[J]. 辽宁工程技术大学学报(自然科学版),2022,41(1):73-78.

[12] 桂易琪,叶子鉴,陈永康. 一种基于区域划分和内容流行度的 NDN 缓存策略[J]. 江苏大学学报(自然科学版),2023,44(2):186-193.

[13] 段星德,伍震寰,张钟妮,等. 半参数双重 Tweedie 复合泊松回归模型的贝叶斯分析[J]. 应用数学,2024,37(1):272-279.

[14] 黎秀,潘虹,朱镕申. 移动终端软件高速缓存侧信道脆弱性识别仿真[J]. 计算机仿真,2022,39(2):496-500.

刘海松(1981—),男,汉族,江苏南通人,硕士,实验师,主要研究方向:模式识别、网络安全;

倪　震(1980—),男,汉族,江苏兴化人,博士,副教授,主要研究方向:智能信息处理、网络安全;

郭红建(1977—),男,汉族,江苏泰兴人,硕士,教授,硕士研究生导师,研究方向为信息安全,信息系统审计。