基于李群的网络系统行为风险计算方法

肖禹名, 赵小林, 刘振岩, 宋　策, 常　悦

北京理工大学计算机学院 北京 中国 100081

摘要　随着互联网在生活中的广泛应用, 越来越多的软件开始收集更多用户信息以改进用户体验, 而这些存放在服务器上的用户隐私, 同时也存在着极大的泄露风险。对网络风险进行实时评估, 不仅可以关注到网络状态的变化, 也有利于随时调整网络防御手段, 及时防御网络攻击, 减小攻击损失。传统上, 网络风险评估往往采用统计计算的方法来进行, 本文通过采用数学中的李群模型, 对网络系统进行数学建模, 从而提出了一种实时计算网络风险的新算法。本文用李群运动学描述网络中的攻击行为。通过将网络系统中由指标和拓扑组成的矩阵映射到李群, 给出攻击行为路径以及网络攻防的数值定义, 使用测地线计算李群中元素的距离, 作为网络风险指标, 并提出了相应的网络风险损害评估方法, 从而将网络风险量化, 实现对网络安全状态的实时评价。为了检验这种网络安全风险评估的有效性, 本文使用现有数据集, 并编写代码进行了相关实验, 以评估这种方法的适用性与效率。实验结果证实, 该基于李群的网络系统行为风险计算方法对于网络攻防风险值的客观量化计算是有效的, 可以实现对网络风险的定量评估。在与其他的机器学习算法相比时, 各指标上均没有明显差距, 而其具有的一些特点则具备被深度开发的潜质。

关键词　网络安全; 李群; 数据安全; 风险计算; 行为风险

中图法分类号　TP309.2　DOI 号 10.19363/J.cnki.cn10-1380/tn.2025.07.03

A Method for Calculating Behavioral Risk in Network Systems Based on Lie Group

XIAO Yuming, ZHAO Xiaolin, LIU Zhenyan, SONG Ce, CHANG Yue

School of Computer Science &Technology, Beijing Institute of Technology, Beijing 100081, China

Abstract　With the widespread use of the Internet in our lives, more and more software is beginning to collect more user information to improve the user experience. However, there is also a significant risk of leakage of user privacy stored on servers. Real-time assessment of network risks can not only help to monitor changes in network status, but also facilitate the adjustment of network defense measures at any time, and timely defense against network attacks, reducing attack losses. Traditionally, network risk assessment often uses statistical methods to calculate. This article proposes a new algorithm for real-time calculation of network risks by using the Lie group model in mathematics to mathematically model network systems. This article uses Lie group kinematics to describe attack behaviors in the network. By mapping the matrix composed of indicators and topology in the network system to the Lie group, it gives the numerical definition of attack behavior paths and network attack and defense. Using geodesic to calculate the distance between elements in the Lie group, it serves as a network risk indicator, and proposes a corresponding network risk damage assessment method, thus quantifying network risks and achieving real-time evaluation of network security status. In order to test the effectiveness of this network security risk assessment method, this paper uses existing datasets and writes code to conduct relevant experiments to evaluate the applicability and efficiency of this method. The experimental results confirm that the Lie group-based network system behavior risk calculation method is effective for the objective quantitative calculation of network attack and defense risk values, and can achieve quantitative assessment of network risks. Compared with other machine learning algorithms, there is no significant difference in various indicators, and some of its characteristics have the potential for further development.

Key words　network security; Lie Group; data security; risk calculation; behavioral risk

1　绪论

近年来, 互联网发展取得长足进步的同时, 网络犯罪也隐藏着巨大的安全隐患。进入新世纪以来, 随着网络应用逐渐普及, 越来越多的应用程序开始在使用中持续收集用户数据, 用作算法更新和数据

通讯作者: 赵小林, 博士, 责任教授, Email: zhaoxl@bit.edu.cn。

本课题得到 国家重点研发课题《健身知识和线上指导数据安全与隐私保护技术研究》(No. 2022YFC3600404)资助。

收稿日期: 2023-11-30; 修改日期: 2024-02-01; 定稿日期: 2025-06-11

统计。目前, 世界各国普遍严禁本国用户数据离开境内服务器, 例如苹果公司在我国选择与云上贵州合作, 储存 iCloud 数据; 美国 Tik Tok 用户的数据则由当地的甲骨文公司保管。因为这些数量庞大的用户隐私数据, 一旦被不正确使用, 可能给用户、社会, 甚至国家带来严重的安全威胁, 维护网络空间安全, 就是在维护国家安全[1-2]。

现今大数据计算环境下, 数据处理流程主要存在三类隐私泄露问题: 数据输入阶段的原始数据被攻击者截获, 计算过程中的隐私数据被攻击者窃取, 以及不可信的数据消费者在结果输出阶段试图推断出数据隐私。与数据的上传与计算阶段相比, 数据在本地端的输入与存储环节往往保护强度更低, 更容易遭受网络攻击, 从而导致数据被窃取或保护机制失效, 所以对这些环节的网络行为需要做出安全评估。通过评估这种攻防力量的大小和网络连接的状态, 可以实现对攻击威胁的预测, 并防止损害进一步扩散。因此, 找到对此进行具体描述与量化的方法就显得尤为重要。

本文的主要贡献在于, 完善了将李群模型应用到网络空间安全状态评估的算法, 提出了一种有别于现有的网络安全评估方法的, 建立于数学模型基础上的新算法, 为网络安全风险计算提供了新思路。

2 相关研究

目前学术界最常见的隐私保护方案主要是差分隐私、联邦学习、安全多方计算等应用于数据计算阶段的隐私计算协议, 其对于计算阶段的数据泄露与推断攻击通常具有良好的防御效果[3-4], 但是并未涉及对于本地设备的网络攻击, 而这恰恰是危害最大的隐私泄露方式[5-6]。在对当前网络的风险评估方面, 一些研究提出了基于攻防对抗游戏的方法[7-9], 但存在状态爆炸问题。此外还有利用节点间存在漏洞对网络风险进行评估的复杂网络攻击建模的方法, 以及蔡等人在相关专利中[10]提出的基于层次分析法的贝叶斯网络评价方法等。Ramos 等人[11]对网络安全度量进行了深入的调查, 并讨论了它们的优缺点。本文在此基础上, 应用了一种新的方法。

微分流形用于描述微小的变化[12], 在图像处理[13-14], 信号处理[15]和机器学习中起着重要作用。Zlotnik 和 Forbes 提出了一种基于李群切线空间的不变代价函数梯度的状态估计[16]。高等人[17]提出了李群核学习的概念, 可以在高维空间中对数据集进行分类。若将网络的安全状态, 划分为不同类型的低维和高维空间, 就可以使用李群核学习的方法进行网络安全措施的应用。而对于网络的攻防计算, 同样也可以使用微分流形来进行建模, 在文献[18]中, 已经给出了与之相关的概念定义, 这为在网络安全评估中应用微分流形奠定了良好的基础。

在网络攻防及其效用度量上, 目前主要有基于 Stackelberg 博弈的攻防建模[19]、基于 Markov 博弈的攻防建模[20]、基于 Fliplt 博弈的攻防建模[21], 以及基于微分博弈的攻防建模[22]等。由于双方的决策具有顺序性, Stackelberg 博弈常被用于建模物理层安全中具备贯序决策特征的攻防过程; 多阶段的攻防过程通常基于具有离散多阶段特征的 Markov 博弈来分析; Fliplt 博弈在 2013 年被提出并广泛应用于 APT 攻击建模, 其通常表现为持久且隐蔽的计算; 而微分博弈则主要用于描述连续时间、多状态的动态安全攻防过程。

与其他评价方法相比, 李群法是一种更为客观的方法。本文将网络映射为李群结构, 从而通过网络状态的变化来描述网络攻击的状态。并通过研究流图中相邻节点的效用和距离对攻防效能进行评估。对于其在网络安全中的具体应用, 本文参考了一项专利, 即文献[23]。

3 网络风险的李群表示

3.1 符号表示

除非另有说明, 本文中使用以下符号, 如表 1 所示。

3.2 网络系统概念

网络系统

网络是一个支持各种服务和任务的虚拟空间, 网络对象由各种软件、硬件设备和信息进程组成。

网络行为

在网络系统中, 对于信息的收集、传输、处理、存储以及共享等, 共同构成了网络行为[18]。网络行为是网络系统中的一系列步骤, 具有一定的功能和意义。网络行为是网络不可分割的一部分, 也是网络实现具体能力的关键。

攻防能力

网络攻击的攻击力度用于衡量网络攻击突破网络防御的能力, 用 P 表示。攻击损害用于衡量网络攻击对网络系统造成破坏的能力, 用 L 表示。

网络系统防御能力用 F 表示, 衡量网络对于外界攻击的防御能力。网络系统恢复能力用 R 表示, 衡量网络系统在收到外界攻击后的修复能力。

表 1 符号说明

Table 1 Symbol description

符号	意义
x_i	第 i 个变量值
r_i	第 i 个实际观测值
std_i	第 i 个标准值
A	由常数组成的变换矩阵
y	李群映射中单个设备解析值
$d(x_1, x_2)$	李群中 x_1, x_2 元素间的距离
P	攻击力度
F	系统防御能力
E	指标观测矩阵
J	权重图
k_i	第 i 次运算中的随机数
X_{0i}	第 i 个设备的初始值矩阵
D	E 与 J 的内积
y^a_{ij}	李群中 a 点的第 i 行, 第 j 列元素
n	维数
Y	某点对应的李群表示
V	由指标构成的向量
U	攻防效用
L	攻击损害
R	网络系统恢复能力
Z	网络资产价值
b	攻击时间系数
T	代表攻击行为的攻击时间的常数
t	实际攻击时间
C	系统防御成本

3.3 李群和网络的相似性

3.3.1 李群的定义

李群是一类连续变换群, 在数学上, 一个李群可以表示成如下形式, 如公式（1）所示:

$$y = f_i\left(x_1, x_2, x_3, \ldots, a_1, a_2, \ldots, a_n\right), i = 1, 2, \ldots, n \quad (1)$$

对于 x_i 和 a_i, f_i 都是解析的。其中 a_i 是参数, 而 x_i 则是变量。所有的$(x_1, x_2, x_3, ..., x_n)$可以表示为 n 维空间中的一个点。

3.3.2 网络系统的李群构造

李群不仅是一个流形, 同时还具有群结构, 群中的加法和逆元运算构成了流形中的解析映射。定义一个李群为一个集合 G, 那么其满足:

1. G 是个群;

2. G 是个微分流形;

3. G 的群结构和微分结构相容。

李群保持了群运算的光滑结构, 同时具有微分流形的性质, 因此其常被用于描述数学和物理中的连续变换。而网络空间中状态转换恰好满足连续性条件, 此外网络和李群也有许多相似之处, 如表 2。

表 2 网络系统与李群的相似性

Table 2 The similarity between network systems and Lie groups

网络系统	李群
网络拓扑结构	拓扑图 G
网络行为	路径 $y=Ax$
指标	参数 x_i
指标的微小变化	导数 $x'(t)$
网络节点	李群中的点

作为微分流形, 李群的几何性质可以描述网络中的各项数据, 而群的代数性质则可以提供特定的解。对一个李群 G, 存在被称为 G 的李代数的向量空间, 包括距离计算在内的许多运算, 都可以通过李代数空间完成。

为了使用李群模型来描述网络行为, 需要把网络中各项数据和拓扑组成的矩阵映射到李群。关于网络可以表示为李群的证明, 文献[24]中已经进行了较为详细的阐述。

3.4 网络指标的连续性

3.4.1 指标的连续性构造

假设一个由诸多具有防御功能的设备组成的网络, 安装有防火墙和防病毒软件。相应地, 黑客也拥有许多攻击工具, 并利用网络中的漏洞发起攻击。攻击期间, 网络中的风险会随着状态不断变化, 反应在数学上, 就是对应的指标不断变化。

在本文所涉及的尺度和理论中, 时间的连续性毋庸置疑, 相应的, 随时间变化的各项指标也都是连续的。进一步, 对于离散数据, 认为其在各值之间是平滑过渡, 相应的导数也是光滑的。这样就实现了网络系统中指标的连续性构造。

3.4.2 指标的连续性构造

网络系统中的设备都可看成是李群空间中的点, 设备指标个数即对应该点坐标元素数量。指标的变化值即是对应流形的值。每个时刻设备的所有指标可以构成一个向量$(x_1, x_2, x_3, ..., x_n)$, 通过公式 1 可知, 可以使用该值来计算当前网络的状态, 这通过将f_i解析成 a_i 和 x_i 来实现。在代数表达式中, 公式 1 也可以被写成公式 2

$$y = Ax \quad (2)$$

其中 A 表示由常数组成的变换矩阵, x 表示变量矩阵。

3.4.3 李群的时空连续性

前面已经证明, 任意一个网络指标对时间是连续的。因此, 公式 2 可以对时间求导, 如公式 3 所示:

$$\frac{dy}{dt}=A\frac{dx}{dt} \tag{3}$$

基于相同的原因, 也可以进行空间连续性构造, 因此, 公式 2 对任意指标 x 的偏导也存在, 为公式 4:

$$\frac{dy}{dx_i}=A\frac{dx}{dx_i} \tag{4}$$

3.5 基于李群的网络风险计算

3.5.1 网络指标的高维映射

由于攻击, 网络设备之间的状态会发生变化。在李群中, 其体现为相邻元素间的距离变化。该变化可以用矩阵表示, 状态变化的方向即空间中沿切线向量的方向。因此, 变化的状态间的距离可以衡量网络风险。

一般来说, n 维流形 M_n 不是向量空间, 流形上的度量也不在欧式空间中。因此, 无法直接在李群中计算距离, 它需要通过测地距离来测量。局部欧氏变换是李群变换的特征, 因此可以先在欧式空间中收集指标, 其与初始值的差即为欧式空间中的距离。然后定义平滑的映射函数, 即可将其映射到更高维度的李群空间中。该变换过程是一个等距变换, 也是一种李群变换。经过变换后, 指标矩阵的维数大大增加, 从而可以进行李群中的距离计算。

本文使用了如下映射函数[24]使其可以保持求导不变性, 如公式 5 所示, 相比文献[23]中使用的公式, 公式 5 具有更强的对称性:

$$y_{ij}=\begin{cases} e^{-x_i*x_j} & i=j \\ e^{-\left(x_i-x_j\right)^2} & i\neq j \end{cases} \tag{5}$$

3.5.2 李群中的距离计算

对于李群空间, 任意两个元素之间的距离由公式 6 给出[24]:

$$d\left(a,b\right)=\|\log\left(a^{-1}b\right)\| \tag{6}$$

对任意的矩阵 X, 有公式 7:

$$\|X\|=\sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}\left|x_{ij}\right|^2} \tag{7}$$

因此有公式 8:

$$d(a,b)=\|\log(b)-\log(a)\| \tag{8}$$

文献[24]中, 使用公式 9 对公式 8 进行近似:

$$d(a,b)=\sum\left(-\ln y_{ij}^{a}+\ln y_{ij}^{b}\right) \tag{9}$$

其中, $y^a{}_{ij}$ 和 $y^b{}_{ij}$ 分别代表在李群中, a 点和 b 点所对应的矩阵里第 i 行、第 j 列的元素的实际取值。若在网络攻击中不相邻, 则认为二者间距离无限远。

综上, 任意两元素间的距离公式为公式 10:

$$d(a,b)=\begin{cases}\sum\left(\ln y_{ij}^{b}-\ln y_{ij}^{a}\right) \\ \infty\end{cases} \tag{10}$$

上下两种情况分别代表距离存在与不存在。

由于李群中的距离反映风险, 因此通过公式 10 可计算攻击过程中任意两设备间的风险。相应地, 根据网络攻击序列, 可以计算整个攻击路径的风险。

4 网络安全风险计算模型

在网络攻防效用计算上, 李群模型的数学基础是微分流形[18], 适用于描述网络的动态变化, 其针对攻防效用的综合计算通常采用攻击与防御能量相减的方式进行[24], 该方法简单有效, 且同时考虑了网络风险的内部和外部因素。不过由于实际情况较为复杂, 该方法仍然具有局限性。一方面, 该方法仅考虑了网络风险, 没有考虑攻击本身的差异, 由于网络攻击类型复杂多样, 每类攻击差异度很大, 因此对于不同的攻击, 往往需要采取不同的建模方式; 另一方面, 效用函数与网络风险相乘, 当某一个值较高时, 较小的值反而会对运算结果产生较大影响。

针对上述问题, 本文基于现有的微分博弈和李群模型[25], 提出了一种新的计算方法。

4.1 网络攻防效用计算

4.1.1 攻防成功的判定

使用网络攻击力度 P 衡量网络攻击的强度, 防御能力 F 衡量网络系统对对应攻击的防御能力。由于网络攻击类型的多样性, 因此需要对不同类别的攻击分别建模。对于暴力破解、恶意代码等网络攻击, 只有当 $P>F$ 时, 网络攻击才会取得成功, 攻击有效, 并对网络系统带来损害; 否则判定为攻击被网络成功防御, 几乎不对网络系统造成损害。

对于通过占用大量资源致使服务器瘫痪的 DoS 等攻击, 即便攻击未成功, 仍然会因为对资源的占用等方式对网络设备产生影响[26], 此时, 网络攻击力度 P 同时可以反映网络攻击对资源的占用情况, P 越大, 资源占用越明显。若 $P>F$, 则说明资源占用超出系统负载, 会导致服务器完全瘫痪, 此时攻击取得成功。

一般而言, 该模型适用于常见的主动攻击判定, 攻击的流量特征越明显, 评估结果越好。特别的, 对于被动攻击等不会使网络流量产生显著变化的攻击, 由于其流量特征不明显, 因此不适用于本文的评估方法。

4.1.2 攻击损害的判定

在文献[24]中, 对于网络攻防效用的计算直接使

用了攻击效用与防御效用相减来实现, 而实际情况更为复杂。根据文献[25], 在考虑了网络系统的恢复能力, 资产价值与对网络攻击的动态跟踪能力后, 攻击成功时的效用可以写成公式 11:

$$U=(L-R)*Z-\frac{bT}{t}+C \tag{11}$$

在攻击失败时, 效用可以写成公式 12:

$$U=-\frac{bT}{t}+C \tag{12}$$

其中第一项代表直接攻击损害, L 表示攻击损害, R 表示系统恢复能力, Z 表示系统资产价值; 第二项代表网络系统的跟踪能力, 即网络系统发现攻击行为的能力, 其是与攻击时间相关的函数, T 是代表攻击行为的攻击时间的常数, t 是实际时间, b 是系数, 随着 t 的减小, 跟踪能力增加, 攻击行为的成本也将相应增加。第三项 C 表示系统防御成本[25]。

在一般的主动攻击模型中, 网络攻防效用的指标可以分成攻击影响、攻击成本、防御效果、防御成本和资产五部分[25], 具体到效用计算中, 各指标所涉及的详细定义与计算方法可以参考文献[25]与国家重点研发计划《网络系统安全度量方法与指标体系》(2016YFB0800700)。例如防御成本包括网络架构、资源控制、安全事件策略等指标, 系统恢复能力包括恢复计划、数据备份计划、漏洞修复平均时间等指标, 攻击损害则取决于具体的攻击方式与攻击强度。

4.1.3 随机因素的影响

网络风险评估仅能作为对实际状况的参考, 只能估计网络攻击的趋势, 不能完全代表网络攻防能力的实际情况。因此, 若网络系统拥有较高的风险值, 意味着其多数情况下对于网络攻击的防御能力较差, 但即便攻击力度大于网络系统的评估防御能力, 由于多种因素的影响, 攻击未必 100%成功。为了模拟各种偶发因素的影响, 本文增加了随机变量的因素。

网络攻击成功与否受到多种因素的影响, 可能会使攻击更容易得手, 也可能会让防御更加有效。假设所有的因素都是随机的, 优先级等价, 所有的偶发因素都是有助于或不利于攻击的对立事件。显然, 所有对立事件的共同分布结果服从二项分布。当考虑的偶发因素足够多时, 根据中心极限定理, 其分布会趋近于正态分布。因此, 对于每次攻击, 实际的攻防情况会是以理论值为期望的正态分布。为了简化模型, 认为 F 固定, 将其转换为攻击力度 P 的正态分布。由于其中偏离中心太远的部分对于实际情况而言没有意义, 因此需要限定波动幅度与范围。

4.1.4 相关参数计算

网络风险值越大, 攻击成功的概率就越高。在不考虑其他因素的情况下, 网络中的风险完全是由网络攻击所导致的, 因此本文使用计算出的网络系统的总风险值来作为该情况下网络攻击的力度。显然, 网络攻击力度越大, 对网络带来的影响越明显, 网络风险值越高。

网络系统的防御能力 F 是网络系统面对不同攻击时的防御与耐受能力的综合体现。由于网络攻击方式不同, 同一网络系统在面对不同攻击时, F 的值也会有变化。例如某系统可能在面对 DDoS 攻击时具有较高的 F 值, 但由于其密码安全设置不合理, 使其可能在面临暴力破解攻击时极其脆弱。因此, 针对不同的攻击方式, 需要使用网络中的不同指标, 对 F 值分别计算。网络系统受到攻击时的承受能力取决于网络的规模、连接方式与安全措施。对于特定的攻击和网络系统, 其具体数值被唯一确定。其具体计算可以使用在国家重点研发计划《网络系统安全度量方法与指标体系》(2016YFB0800700)中, 对于网络防御能力的计算提出的指标。例如对于暴力破解类型的攻击, 网络系统防御能力的影响因素有口令强度、网络入侵防范、身份鉴别、访问控制等。而对于资源占用类型的攻击, 则需要考虑软件容错、设备资源控制、并发进程资源控制、网络架构等指标。其具体数值计算可以在指标采集后, 使用相关的数学或机器学习算法实现。

相应地, 攻击力度 P 的波动情况也与网络状态和攻击方式有关。系统恢复能力, 系统资产价值, 网络防御成本等指标则由具体的网络系统设备所确定。

4.1.5 攻击损害计算公式

综上, 网络系统的攻防效用计算公式为 13:

$$U=\sum_{i=0}^{n}(L_i-R_i)\times Z_i\times\left[1+\frac{k_iP_i-F_i}{k_{\max}P_i+F_i}\right]-\frac{b_iT_i}{t_i}+C_i \tag{13}$$

其中 k_i 为在每次攻击中相互独立的随机数, []为取整函数。不考虑每次攻击之间的相互影响, 则对于第 i 次网络攻击, 当且仅当结果大于等于 0 时, 说明该次攻击成功。公式 13 可评估在一段时间内, 网络系统整体的攻防效用, U 小于 0 时, 说明在该时间段内网络防御占优势, U 大于 0 时, 说明该时间段内网络攻击占优势, 网络面临较大风险, U 等于 0 时, 说明网络攻防实现了平衡。

4.2 网络攻防效用计算

本文应用李群模型, 建立了如下一套风险计算算法。

如图 1, 该算法的总体步骤如下:

1、获取设备指标并标准化;

2、构建向量 $V=(x_1, x_2, x_3, ..., x_n)$, 并根据初始数据计算变化量;

3、将变化映射到李群空间中;

4、构建网络权重图 J;

5、计算指标观测矩阵 E。

6、计算 J 和 E 的内积, 得到攻击力度 P。

7、收集与评估网络攻击危害、网络系统的防御和攻击耐受能力等各项指标;

8、计算攻防效用, 对网络风险进行评估。

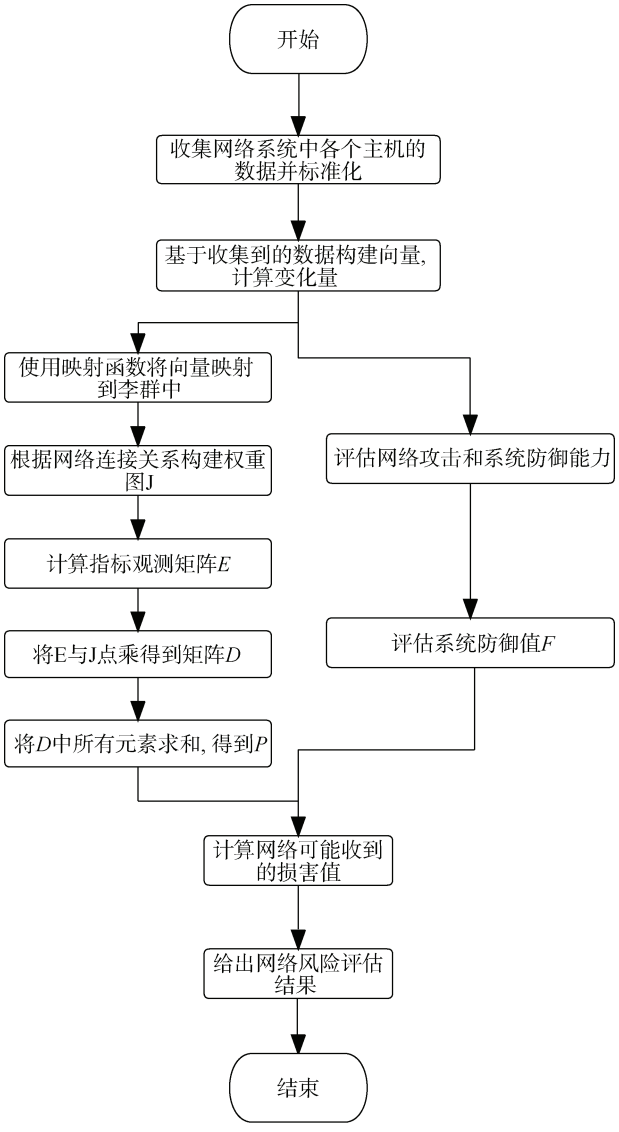

图 1 算法流程图

Figure 1 Algorithm flow chart

下文对各个步骤进行详细说明。

4.3 使用网络指标构造李群

4.3.1 收集网络指标

为了全面评估网络属性, 收集的指标不仅要涵盖网络的可用性、连接性和可靠性, 还要能衡量总体风险。

表 3 即给出了一个例子, 在实际应用中, 相当多的与网络相关的指标都是可用的, 指标数量越多, 对网络状态的评估越全面。当网络攻击带来相关指标的变化时, 不同设备在不同时刻的差距就会在李群空间中反映出来。

表 3 风险测量指标举例

Table 3 Example of risk measurement indicators

网络系统	矩阵
可用性	CPU 使用值
	内存使用值
	磁盘占用值
	系统负载值
连接性	开放端口值
	峰值流量
	宽带使用值
	节点连接值
	平均流量
可靠性	系统漏洞值
	程序漏洞值

对于每个设备, 某一时刻所有的观测值可以构成一个向量 $V=(x_1, x_2, x_3, ..., x_n)$。其中 n 是测量的指标数。

4.3.2 网络指标标准化

为了避免不同数据采用的量纲与数量级不同所带来的影响, 需要对数据标准化。每项指标都采用其在一段时间内收集的正常数据平均值作为标准, 将指标与标准值的比值作为实际数值, 这样使得所有数据具有相同的起点(即标准值为 1), 也使所有指标均无量纲。如果某个值在标准化后数值过大, 则说明其在该时间内变化幅度非常剧烈, 相应地, 此时网络中出现异常流量的可能性较大, 网络风险值较高。而网络攻击往往是在原有的网络正常使用的基础上进行的, 往往不会导致某个值在短时间内相比正常情况下剧烈减少。因此, 为了避免部分剧烈变动的数值在归一化后影响权重降低, 同时为了便于在后续处理过程中应用主成分分析法进行近似, 这里不对标准化后的数据进行归一化处理。相应的, 默认数据集中的所有的数据全部正确, 不会出现因数据收集错误导致的离散点。

数据标准化公式如下:

$$x_i = \frac{r_i}{std_i} \tag{14}$$

其中, x_i, r_i, std_i 分别代表向量中的第 i 个值, 及其对应

的观测值和标准值。

4.3.3 把网络系统映射为李群

对于选定的时刻, 使用指标的测量值减去初始值, 即得到了指标变化量 $X'=(x_1', x_2', x_3', ..., x_n')$。其中 $x_i'=x_i-x_{i0}$,。然后使用公式 5, 将低维(n 维)欧氏空间转换为高维(n^2-n 维)黎曼流形。对于有 n 个指标的设备, 在每个时刻对应的状态都会产生一个 n 维方阵。

4.4 在李群中进行网络风险计算

4.4.1 构建邻接矩阵

攻防效用集中体现在攻击过程中。本文使用邻接矩阵表示攻击路径。为了突出不同设备的重要性, 可以为链接添加不同权重。

4.4.2 构建指标观测矩阵

将某一时刻所有网络设备映射到李群后, 即可使用公式 10 计算任意设备间的风险。

n 个设备两两之间的距离构成了一个 n^2 维的指标观测矩阵 E。其中第 i 行, 第 j 列的元素代表第 i 到第 j 个元素的距离, 即从第 i 向第 j 个设备传输信息的通路所面临的网络风险。由公式 10 可以看出, 该矩阵为对称矩阵。

4.4.3 计算网络系统总风险

在得到指标观测矩阵 E 与邻接矩阵 J 后, 通过计算内积即可得到在该轮攻击中, 在 J 所表示的攻击序列上的总风险值。

4.5 损害计算

将计算得到的风险值作为此轮攻击的攻击力度 P, 在对网络系统与攻击的各项指标进行评估后, 即可通过公式 13, 对可能遭受的损失情况进行评估。

对于已经确定的 P, F 而言, 由于网络系统确定, 同时公式 13 中 k 分布已知, 因此最终结果分布也被唯一确定。相应地, 对于某种特定的结果(如网络攻击占据优势, 或网络攻防达到平衡), 都对应一个 k 的取值范围, 可以通过 k 的分布计算出结果出现的概率。

5 实验验证

5.1 小型网络实验过程

5.1.1 假设与基本数据

如图 2, 假设存在一个由防火墙、路由器、交换机、服务器 4 台设备组成的网络系统, 网络攻击也按照此顺序进行。

这里使用了系统负载值、网络漏洞值和瞬时流量三个指标进行度量, 假设收集到的数据如表 4, 表 5 所示。

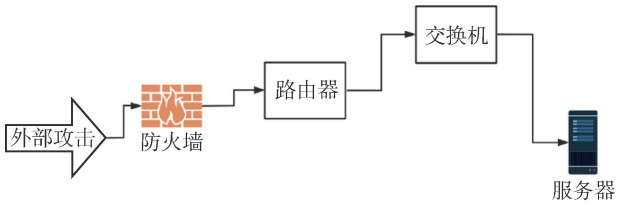

图 2 网络攻击路径

Figure 2 Network attack path

表 4 测量值

Table 4 Measurement value

id	系统负载值(%)	网络漏洞值(评估得分)	瞬时流量(kB/s)
1	63	52	704
2	50	59	100
3	53	85	430
4	78	76	907

表 5 初始值

Table 5 Initial value

id	系统负载值(%)	网络漏洞值(评估得分)	瞬时流量(kB/s)
0(标准数据)	40	5	703
1	28	2	217
2	58	8	140
3	34	50	341
4	40	15	103

5.1.2 数据处理与运算

对任意设备(例如 id=4 的服务器), 其观测指标构成了一个向量 $V_4=(load, loophole, flow)$。没有攻击时三个参数的值分别是 40%, 15, 103kB/s, 而标准数值为 40%, 5, 703kB/s, 标准化后得到公式 15:

$$V_4=(1.0,3.0,0.1465) \tag{15}$$

对于观测的时刻, 其对应值分别为 1.95, 15.2, 1.29, 变化量如公式 16 所示:

$$\begin{aligned}\Delta V_4 &= V_4 - V_0 \\ &= (1.95,15.2,1.2902)-(1.0,3.0,0.1465) \\ &= (0.95,12.2,1.1437)\end{aligned} \tag{16}$$

其他各点需要采用公式 5 中的函数将坐标映射到高维空间。例如对于 ΔV_4, 其在经过映射后, 数据变为一个 3 行 3 列的矩阵, 如公式 17:

$$Y_4=\begin{bmatrix} 0.40 & 1.08e^{-55} & 0.96 \\ 1.08e^{-55} & 2.29e^{-65} & 8.14e^{-54} \\ 0.96 & 8.14e^{-54} & 0.27 \end{bmatrix} \tag{17}$$

然后可以计算出任意连接的指标, 如公式 18 所示:

$$E_{12}=d\left(Y_1,Y_2\right)=68.79600701702122 \tag{18}$$

对角线上的元素对应设备自身的风险, 通过对设备的初始值和观测值之间的变化量计算得到, 例如, 对 id 为 4 的设备, 其初始值和观测值分别可映射成一个 3 行 3 列的矩阵, 如公式 19 与公式 20:

$$X_{04}=\begin{bmatrix}0.368 & 0.018 & 0.483\\ 0.018 & 0.0001 & 0.0003\\ 0.483 & 0.0003 & 0.979\end{bmatrix} \tag{19}$$

$$d_4=\begin{bmatrix}0.022 & 5.67\mathrm{e}^{-77} & 0.647\\ 5.67\mathrm{e}^{-77} & 4.58\mathrm{e}^{-101} & 9.36\mathrm{e}^{-85}\\ 0.647 & 9.36\mathrm{e}^{-85} & 0.189\end{bmatrix} \tag{20}$$

于是, 就有了公式 21:

$$E_4=d\left(X_{04},d_4\right)=939.705 \tag{21}$$

相应可以获得其他三个点的坐标,

于是获得指标观测矩阵, 如公式 22 所示:

$$E=\begin{bmatrix}443.48 & 68.80 & 211.97 & 207.64\\ 68.80 & 625.89 & 280.77 & 138.85\\ 211.97 & 280.77 & 870.98 & 419.61\\ 207.64 & 138.85 & 419.61 & 939.71\end{bmatrix} \tag{22}$$

构造的邻接矩阵如公式 23 所示, 由于所有设备之间都是单向连接, 因此所有设备之间连接的权重均为 1。对角线元素为设备本身的权重, 由设备在网络系统中发挥的作用与重要性所决定, 各设备在网络系统中的权重计算可参考文献[27]中提出的方法。由于本实验为计算方法示例, 因此采用主观方法对设备权重进行了简化, 考虑到防火墙(id=1)相对其他设备, 在面对网络攻击时的作用更加显著, 为了区分不同设备, 设置防火墙的权重为 2, 其他设备的权重为 1:

$$J=\begin{bmatrix}2 & 1 & 0 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 0 & 1\end{bmatrix} \tag{23}$$

随后得到公式 24:

$$D=J\cdot E \tag{24}$$

在公式 24 中, D 是权重矩阵和指标风险矩阵的点积, 因此有公式 25:

$$\begin{aligned}P&=\sum D\\ &=4092.7\end{aligned} \tag{25}$$

5.1.3 网络风险评估结果

假设该网络系统对此类攻击的防御值为 4000, 由于本次攻击的攻击力度为 4092.7, 则只要相应的随机数值大于 0.9773, 攻击就会成功。

为了说明随机因素如何影响攻击成功的概率, 假设 P 值会在平均值上下 25%的范围内波动。由于此区间内概率分布不明显, 因此将其分布映射到标准正态分布[–1,1]的区间中。则攻击成功的概率即标准正态分布在–1 到 1 上的截断正态分布中取值超过–0.09 的概率, 计算得出其值为 0.553, 因此, 本轮攻击有 55.3%的概率成功, 网络面临较大风险。

接下来计算攻防效用, 假设网络系统的防御成本为 1000, 跟踪能力为 1500, L=240, R=100, Z=70, 则根据公式 13, 可以算出在本次攻击中的攻防效用为 4919.4, 网络具有较高的安全隐患。

5.2 数据集验证实验

5.2.1 数据集说明

为了更好地验证模型的真实性和有效性, 本文采用了加拿大网络安全入侵检测系统研究所的 2017 版数据集(Canadian Institute for Cybersecurity-Intrusion Detection Systems-2017, CIC-IDS2017)进行验证[28]。由于数据集中都是节点间的连接信息, 而非节点本身的信息, 且数据量庞大, 因此本文采用了一些近似。

5.2.2 指标收集与标准化

本文从数据集中选用了 58 个指标, 并选用周一的正常数据作为初始值与标准值。对于数据异常(比如出现了 inf)的情况随机选取了一个时间的数据进行代替。

5.2.3 获取拓扑关系构造邻接矩阵

数据集提供了设备间的连接信息。实验采用了数据集中出现次数最多的 14 台主机, 包括了防火墙与 13 个其他主机, 主机信息在表 6 中列出:

表 6 选取的主机信息

Table 6 Selected host information

id	ip 地址	类型
1	172.16.0.1	Firewall
2	192.168.10.3	DNS+DC Server
3	192.168.10.50	Web server 16
4	192.168.10.51	Ubuntu Server 12
5	192.168.10.19	Ubuntu 14.4
6	192.168.10.17	Ubuntu 14.4
7	192.168.10.16	Ubuntu 16.4
8	192.168.10.12	Ubuntu 16.4
9	192.168.10.9	Win 7 Pro
10	192.168.10.5	Win 8.1
11	192.168.10.8	Win Vista
12	192.168.10.14	Win 10
13	192.168.10.15	Win 10
14	192.168.10.25	MAC

设备间的拓扑关系如图 3 所示:

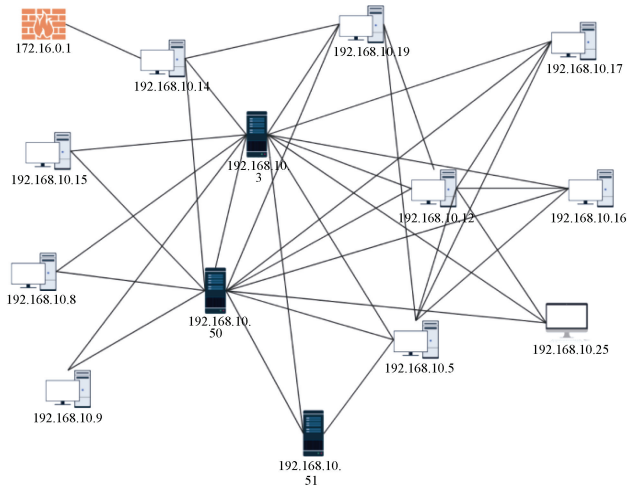

图 3　14 台设备之间的拓扑结构

Figure 3　Topology structure between 14 devices

由于数据集信息量不足, 因此不再考虑设备本身的风险, 即邻接矩阵 J 中所有的对角线元素均为 0。

为了区分不同设备, 本文对不同设备之间的链接设置了不同的权重。由于不考虑设备本身风险, 因此权重仅依赖于设备之间的链接, 不依赖于设备本身。考虑到设备与其他设备之间的链接越多, 网络流量越大, 越容易出现波动, 为了平衡这种影响, 本文采用设备与其他设备之间链接数的倒数作为权重, 填入到该邻接矩阵的对应列中, 从而使得与其他设备之间链接更少的设备具有更大的权重, 其网络指标发生的变化也更能反映出流量的异常。最终统计的各设备链接数如公式 26 所示, 对矩阵 J'中的每个元素取倒数, 即可获得邻接矩阵 J, 由于取倒数后, 计算数值会显著变小, 为了放大计算结果之间的差距, 下文中的数值均为将最终数值放大 50 倍后四舍五入的结果:

$$J'=\begin{bmatrix}
0&0&1&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&6&6&6&6&6&0&6&0&0&0\\
0&0&0&0&0&0&3&0&3&0&3&0&0&0\\
0&0&0&0&0&0&0&0&2&0&2&0&0&0\\
0&3&0&0&0&0&0&0&3&0&3&0&0&0\\
0&0&0&0&0&0&0&0&2&0&2&0&0&0\\
0&0&0&0&0&0&0&0&2&0&2&0&0&0\\
0&3&0&0&0&0&0&0&3&0&3&0&0&0\\
0&12&12&12&12&12&12&12&0&12&12&12&12&12\\
0&6&0&0&6&0&6&0&6&0&6&6&0&0\\
0&12&12&12&12&12&12&12&12&12&0&12&12&12\\
0&0&0&0&0&0&0&0&2&0&2&0&0&0\\
0&0&0&0&0&0&0&0&2&0&2&0&0&0\\
0&0&0&0&0&0&0&0&2&0&2&0&0&0
\end{bmatrix} \tag{26}$$

5.2.4　计算指标观测矩阵

数据集中的数据本身可被认为是欧式空间中的距离。这里借鉴文献[24]中的方法, 将流量数据作为变化向量映射到李群的切空间中, 并使用其中变化最大的量来估算切线方向, 用沿着切线的李群距离作为整个时空变化上的距离。

例如, 对于 IP 地址为 192.168.10.16 的主机, 向 IP 地址为 192.168.10.3 发送数据时, SYG Flag 的初始值是 1, 而在周三上午 10:00-10:05 这个时间段内, 该值平均为 0.003, 在将数据标准化后, 其变化量超过了 125, 远超其他指标, 因此采用该值来进行代表。假设 192.168.10.16 的下标为 i, 192.168.10.3 的下标为 j。则在矩阵 E 中, 有公式 27:

$$\begin{aligned}E_{ij}&=\ln\left(\frac{1}{0.007921942}\right)-\ln\left(\frac{0.00306278713629402}{0.007921942}\right)\\&=5.788\end{aligned} \tag{27}$$

相应可以计算出在该时间段内, 存在流量传输的所有点所对应的 E_{ij} 的值, 最终得到公式 28:

$$E=\begin{bmatrix}
0&0&0&0&0&0&0&0&0&0&6.6&0&0&0\\
0&0&0&0&0&0&0&0&1.6&0&4.8&0&0&0\\
0&0&0&0&0&0&0&0&0.5&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0.6&0&0.7&0&0&0\\
0&0&0&0&0&0&0&0&5.8&0&4.8&0&0&0\\
0&0&0&0&0&0&0&0&1.7&0&0.2&0&0&0\\
0&0&0&0&0&0&0&0&10.2&0&12.1&0&0&0\\
0&0&0&0&0&0&0&0&3.0&0&2.2&0&0&0\\
0&0&0&0&1.3&0&0&0&0&0&0.6&0&0&0\\
0&0&0&0&0&0&0&0&1.1&0&1.4&0&0&0\\
0&4.8&0&0.1&3.5&1.7&2.3&5.0&1.9&0.2&0&0&0.3&0\\
0&0&0&0&0&0&0&0&0.4&0&0.7&0&0&0\\
0&0&0&0&0&0&0&0&2.3&0&0.7&0&0&0\\
0&0&0&0&0&0&0&0&3.5&0&0&0&0&0
\end{bmatrix} \tag{28}$$

5.2.5　网络系统风险评估

由于数据集中并未给出更多的网络系统本身的信息, 因此, 不再具体评估计算网络系统对攻击的防御能力, 仅计算当前状态网络面临的风险值。根据公式 24, 对网络风险的计算结果为 1264。

作为对比, 周一的正常数据计算结果为 420, 说明该网络系统存在较高风险。

5.3　实验结果分析

为了提高准确性, 本文对所有网络攻击均以 5 分钟为时间段进行了检测, 并选取其中一个作为代表。结果如表 7 所示。

显然, 网络系统得分越高, 对应的网络风险越大。计算机在受到网络攻击时, 计算出的风险值与正常时刻相比会有显著变化。

对于一部分网络攻击, 例如XSS, 其评分显著低于其他攻击, 可能是由于攻击行为特征不明显, 因

表 7 网络风险计算值

Table 7 Network risk calculation value

攻击时间	选取的测量时间	攻击类型	风险值
周一	14:19-14:23	无(正常用户操作)	420
周二	10:05-10:09	FTP patator	2947
周二	14:28-14:32	SSH patator	2416
周三	9:48-9:52	DoS slowloris	3076
周三	10:31-10:35	DoS Slowhttptest	1935
周三	10:46-10:50	DoS Hulk	2208
周三	11:15-11:19	DoS GoldenEye	2168
周四	9:41-9:45	brute force	1992
周四	10:28-10:32	XSS	1052
周四	10:40-10:42	SQL injection	1962
周五	16:56-17:00	DDoS LOIT	2071
周五	16:03-16:07	Bot	1626
周五	13:52-13:56	Portscan	2240

此未检测到显著风险。而 SQL 注入攻击的高得分可能是由于攻击时间短, 因此部分流量指标发生了显著变化。数据集仅包含网络主机之间的流量特征, 不包括主机风险特征, 例如应用程序安全指标、web 服务安全、系统安全和其他指标等, 导致部分网络攻击相对来说难以被检测。

5.4 相关工作对比

对于 CIC-IDS2017 数据集, 常见的检测模型为根据网络流量判断攻击是否发生的二分类模型[29-31]。为了验证李群算法的有效性, 本文通过设置网络风险临界值的方法, 通过判断计算出的网络风险值是否超过临界值来判断网络是否受到攻击, 从而将李群转化为二分类模型, 对一周中的所有时间进行网络攻击判定。表 8 中列举了李群算法与部分其他算法在 CIC-IDS2017 数据集上的表现。从表 8 可以看出, 与其他机器学习算法相比, 李群度量方法在许多指标上都处于中等水平[24]。作为一种新的方法, 李群模型仍然有相当多的细节需要优化。

5.5 复杂度分析

对于具有 n 个设备的网络系统, 每个设备提取 m 项指标, 在空间开销上, 由于该方法建立的基础为数学运算, 因此在设备从欧式空间到李群空间的映射中, 计算过程只需要存储 n 个设备在初始时刻的 m 个指标值与用于标准化的 m 个平均值, 存储数量级为 mn。距离计算过程需要存储每个设备对应点的坐标, 为 n 个 m 行 m 列矩阵, 数据数量级为 m^2n。后续过程仅涉及对 n 行 n 列的观察指标矩阵的点乘与求和运算, 仅需存储两矩阵, 数据数量级为 n^2, 由于在指标采集时往往有 $m^2>n$, 因此总的空间复杂度为 $O(m^2n)$。

表 8 网络风险测量方法评价

Table 8 Evaluation of network risk measurement methods

算法	Pr	Rc	Fl
KNN	0.96	0.96	0.96
RF	0.98	0.97	0.97
CART	0.99	0.99	0.99
ID3	0.98	0.98	0.98
C4.5	0.93	0.94	0.94
SVM	0.97	0.97	0.97
Adaboost	0.77	0.84	0.77
MLP	0.77	0.83	0.76
Naive-Bayes	0.88	0.04	0.04
QDA	0.97	0.88	0.92
Lie Group	0.83	0.83	0.83

在时间开销上, 数据标准化与变化量计算需计算的数据量均为 mn, 时间复杂度为 $O(mn)$。在设备从欧式空间到李群空间的映射中, 对于每个设备均需计算数个 m 行 m 列矩阵, 时间复杂度为 $O(m^2n)$。整个算法中, 时间复杂度最高的部分为观察指标矩阵的运算, 该矩阵有 n 行 n 列, 由于其中每个元素的计算都需要将两个 m 行 m 列矩阵中的所有元素取对数并求和, 因此该部分的时间复杂度为 $O(m^2n^2)$。随后的矩阵点乘与求和操作的时间复杂度均不超过 $O(n^2)$。综上所述, 该算法总的时间复杂度是 $O(m^2n^2)$。由于每次结果互不相关, 因此李群算法适用于并行计算。同时因为算法建立的基础是数学模型, 因此与其他方法相比, 计算更加客观, 且可复现。

6 结论

针对隐私计算情境下, 系统面临的网络攻击的风险, 本文提出了一个基于李群的风险度量模型以实现对网络安全状态的实时动态评估, 并给出了相应的风险计算与危害评估方法, 同时进行了实验验证。本文中的内容总结如下:

1. 提出了基于李群的网络安全风险计算模型, 并进行了相关数学证明;

2. 给出了基于李群模型的网络风险计算方法;

3. 提出了网络安全风险的评估模型。

本文的创新之处在于, 完善了网络的李群模型, 以距离计算为基础, 应用李群模型和数学统计规律建立了一套完整的网络安全风险的动态评估方法。在应用文献[23]中提出的, 使用李群中的距离计算网

络风险的方法基础上, 增加了数学证明, 进一步改进了计算公式, 形成了一套完善的风险评估方法; 对不同的网络攻击方式进行了数学建模, 从而可以对不同攻击造成的网络风险进行针对性评估; 拓展了指标采集的方式, 增强了算法的适用性; 在网络服务调用的基础上增加了网络权重, 并给出了相应的计算方法。

鉴于目前的算法还不够成熟, 未来仍然有相当多的工作可以开展:

1. 对于在网络攻防效用计算中出现的系统防御值、攻击损害等指标, 以及随机数变量的选取, 由于暂未找到合适算法, 因此本文在举例时并未进行详细计算, 而是采用了主观值进行替代;

2. 本文对 CIC-IDS2017 的实验数据进行了理想化和近似处理, 且对于其中的错误数据缺乏解决方案。

致　谢　本课题得到了国家重点研发计划课题《健身知识和线上指导数据安全与隐私保护技术研究》(2022YFC3600404)的资助, 部分成果来源于国家重点研发计划《网络系统安全度量方法与指标体系》(2016YFB0800700)。感谢加拿大网络安全研究机构提供的 CIC-IDS2017 公开数据集。

参考文献

[1] Li S Q. Countermeasures and Enlightenment of Network Security Risks of SMEs in EU[J]. *Network Security Technology & Application*, 2022(4): 132-134.
(李舒沁. 欧盟中小企业网络安全风险应对与启示[J]. *网络安全技术与应用*, 2022(4): 132-134.)

[2] Mo W W. Domestic Network Security Risk Assessment and General Situation in 2021[J]. *China Information Security*, 2021(12): 39-43.
(磨惟伟. 2021 年国内网络安全风险评估与总体态势[J]. *中国信息安全*, 2021(12): 39-43.)

[3] Yan S, Lyu A L. Overview of the Development of Privacy Preserving Computing[J]. *Information and Communications Technology and Policy*, 2021(6): 1-11.
(闫树, 吕艾临. 隐私计算发展综述[J]. *信息通信技术与政策*, 2021(6): 1-11.)

[4] Zhao Z D, Chang X L, Wang Y X. A Survey of Privacy Preserving in Machine Learning[J]. *Journal of Cyber Security*, 2019, 4(5): 1-13.
(赵镇东, 常晓林, 王逸翔. 机器学习中的隐私保护综述[J]. *信息安全学报*, 2019, 4(5): 1-13.)

[5] Alanazi M, Aljuhani A. Anomaly Detection for Internet of Things Cyberattacks[J]. *Computers*, *Materials & Continua*, 2022, 72(1): 261-279.

[6] Bai B, Feng Y, Liu B X, et al. Research on Network Behavior-Based Cyberattack Grouping Method[J]. *Journal of Cyber Security*, 2023, 8(2): 66-80.
(白波, 冯云, 刘宝旭, 等. 基于网络行为的攻击同源分析方法研究[J]. *信息安全学报*, 2023, 8(2): 66-80.)

[7] Alhomidi M, Reed M. Risk Assessment and Analysis through Population-Based Attack Graph Modelling[C]. *World Congress on Internet Security*, 2013: 19-24.

[8] Lee J, Lee H, In H P. Scalable Attack Graph for Risk Assessment[C]. *2009 International Conference on Information Networking*, 2009: 1-5.

[9] Dai F F, Hu Y, Zheng K F, et al. Exploring Risk Flow Attack Graph for Security Risk Assessment[J]. *IET Information Security*, 2015, 9(6): 344-353.

[10] Cai Z Q, Zhao J B, Li Y, et al. Information Security Evaluation of System Based on Bayesian Network[C]. *2015 IEEE International Conference on Industrial Engineering and Engineering Management*, 2015: 315-319.

[11] Ramos A, Lazar M, Filho R H, et al. Model-Based Quantitative Network Security Metrics: A Survey[J]. *IEEE Communications Surveys & Tutorials*, 2017, 19(4): 2704-2734.

[12] Lee J M. Smooth Manifolds[M]. Introduction to Smooth Manifolds. New York, NY: Springer New York, 2013: 1-31.

[13] Peyré G. Manifold Models for Signals and Images[J]. *Computer Vision and Image Understanding*, 2009, 113(2): 249-260.

[14] Feng H L. Research on applying manifold learning algorithms to face recognition[D]. Chongqing: Chongqing University, 2008.
冯海亮. 流形学习算法在人脸识别中的应用研究[D]. 重庆: 重庆大学, 2008.

[15] Theodorakopoulos I, Economou G, Fotopoulos S, et al. Local Manifold Distance Based on Neighborhood Graph Reordering[J]. *Pattern Recognition*, 2016, 53: 195-211.

[16] Zlotnik D E, Forbes J R. Higher Order Nonlinear Complementary Filtering on Lie Groups[J]. *IEEE Transactions on Automatic Control*, 2019, 64(5): 1772-1783.

[17] Gao C, LI F, Shen C. Research on Lie Group Kernel Learning Algorithm[J]. *Journal of Frontiers of Computer Science & Technology*, 2012, 6(11): 1026-1038.
(高聪, 李凡长, 沈程. 李群核学习算法研究[J]. *计算机科学与探索*, 2012, 6(11): 1026-1038.)

[18] Hu C Z. Calculation of the Behavior Utility of a Network System: Conception and Principle[J]. *Engineering*, 2018, 4(1): 171-185.

[19] Wang Z, Duan C J, Wu T, et al. Research on Optimizing Security Control Mechanism of Networked System Based on Stackelberg Defender-Attacker Game[J]. *Journal of Cyber Security*, 2019, 4(1): 101-115.
(王震, 段晨健, 吴铤, 等. 基于 Stackelberg 攻防博弈的网络系统安全控制机制优化研究[J]. *信息安全学报*, 2019, 4(1): 101-115.)

[20] Huang K X, Zhou C J, Qin Y Q, et al. A Game-Theoretic Approach to Cross-Layer Security Decision-Making in Industrial Cyber-Physical Systems[J]. *IEEE Transactions on Industrial Electronics*, 2020, 67(3): 2371-2379.

[21] Zhang M, Zheng Z Z, Shroff N B. Defending Against Stealthy Attacks on Multiple Nodes with Limited Resources: A Game- Theo-

retic Analysis[J]. *IEEE Transactions on Control of Network Systems*, 2020, 7(4): 1665-1677.

[22] Gao Q Y. Research on network attack and defense competition based on differential game[D]. Beijing: Beijing University of Posts and Telecommunications, 2022.
高秋悦. 基于微分博弈的网络攻防竞争研究[D]. 北京: 北京邮电大学, 2022.

[23] 赵小林, 陈全保, 薛静锋, 等. 一种基于李群的网络系统风险度量方法: CN108777641A[P]. 2018-11-09.

[24] Zhao X L, Chen Q B, Xue J F, et al. A Method for Calculating Network System Security Risk Based on a Lie Group[J]. *IEEE Access*, 2019, 7: 70610-70623.

[25] Xie W Q. Research on network attack-defense utility based on game theory[D]. Xi'an: Xidian University, 2020.
谢卫强. 基于博弈论的网络攻防效用研究[D]. 西安: 西安电子科技大学, 2020.

[26] Wang F. Research on detection and reaction of distributed denial of service attacks[D]. Changsha: National University of Defense Technology, 2013.
王飞. 分布式拒绝服务攻击检测与响应技术研究[D]. 长沙: 国防科学技术大学, 2013.

[27] Nan Y, Chen L. Performance Evaluation Method Based on Objective Weight Determination for Data Center Network[J]. *Journal of Computer Applications*, 2015, 35(11): 3055-3058, 3091.
(南洋, 陈琳. 基于客观权重确定的数据中心网络性能评估方法[J]. *计算机应用*, 2015, 35(11): 3055-3058, 3091.)

[28] Sharafaldin I, Lashkari A H, Ghorbani A A. Toward generating a new intrusion detection dataset and intrusion traffic characterization[J]. *ICISSp*, 2018, 1: 108-116.

[29] Zhao Z B. Research on network abnormal traffic detection system based on machine learning[D]. Beijing: Chinese People's Public Security University, 2023.
赵忠斌. 基于机器学习的网络异常流量检测系统研究[D]. 北京: 中国人民公安大学, 2023.

[30] Zhou C, Yang D, Wei S J. Integrating GRU and CNN for Light-Weighted Model in Network Intrusion Detection[J]. *Computer Systems and Applications*, 2023, 32(8): 162-170.
(周璨, 杨栋, 魏松杰. 融合 GRU 和 CNN 的轻量级网络入侵检测模型[J]. *计算机系统应用*, 2023, 32(8): 162-170.)

[31] Zhao W B, Ma Z T, Yang Z. Model of the Malicious Traffic Classification Based on Hypergraph Neural Network[J]. *Chinese Journal of Network and Information Security*, 2023, 9(5): 166-177.
(赵文博, 马紫彤, 杨哲. 基于超图神经网络的恶意流量分类模型[J]. *网络与信息安全学报*, 2023, 9(5): 166-177.)

赵小林 现任北京理工大学软件工程专业责任教授。研究领域为网络空间安全、大数据、数据库技术。研究兴趣包括: 隐私计算、微分流形。Email: zhaoxl@bit.edu.cn

肖禹名 于 2022 年在北京航空航天大学计算机科学与技术专业获得学士学位。现在北京理工大学计算机科学与技术专业攻读硕士学位, 为中国计算机学会 CCF 学生会员。研究领域为网络安全、隐私计算。研究兴趣包括: 网络安全。Email: xiaoyuming@bit.edu.cn

常悦 于 2021 年在华北电力大学软件工程专业获得工学学士学位。现在北京理工大学计算机技术专业攻读硕士学位。研究领域为隐私计算、网络安全。Email: wymc6839@163.com

宋策 于 2020 年在华东理工大学电气工程及其自动化专业获得学士学位。现在北京理工大学计算机技术专业攻读硕士学位。研究领域为隐私计算。研究兴趣包括: 网络安全、数据安全。Email: 3220211062@bit.edu.cn

刘振岩 现任北京理工大学计算机学院网络攻防对抗技术研究所讲师。研究领域为人工智能、自然语言处理、网络空间安全等。研究兴趣包括文本分类, 恶意域名检测、隐私计算等。Email: zhenyanliu@bit.edu.cn