研究与开发

基于FP-Growth算法的新能源配电网CPS网络攻击检测方法

李瑞[1]，刘珊[1]，闫磊[2]
（1. 国网山西省电力公司电力科学研究院，山西 太原 030001；
2. 国网山西省电力公司，山西 太原 030021）

摘　要：为有效分析识别有源配电网信息物理系统（cyber physical system，CPS）状态，提出基于FP-Growth算法的有源配电网信息物理系统网络攻击检测方法。首先分析考虑网络攻击的有源配电网控制模型及CPS网络攻击影响机理，通过实时仿真平台对有源配电网CPS信息侧和物理侧进行监测来获取原始数据；然后通过额定电压、电流值制订数据离散化规则，并根据规则对原始数据进行离散量化处理来生成事件序列。在此基础上，采用FP-Growth算法挖掘历史数据异常信号的频繁项集和强关联关系，通过已有频繁序列特征对新的攻击类别和故障点进行识别，实现对有源配电网CPS网络攻击的检测。最后，仿真实验验证了所提方法的可行性和有效性。

关键词：有源配电网；信息物理系统；网络攻击；FP-Growth算法；事件序列
中图分类号：TM73；TP181
文献标志码：A
doi: 10.11959/j.issn.1000−0801.2024229

Network attack detection method for CPS of active distribution network with renewable energy based on FP-Growth algorithm

LI Rui[1], LIU Shan[1], YAN Lei[2]
1. State Grid Shanxi Electric Power Research Institute, Taiyuan 030001, China
2. State Grid Shanxi Electric Power Company, Taiyuan 030021, China

Abstract: In order to effectively analyze and identify the cyber physical system (CPS) status of active distribution network, a network attack detection method based on FP-Growth algorithm for active distribution network cyber physical system was proposed. Firstly, the active distribution network control model considering network attack and the impact mechanism of CPS network attack were analyzed, and the raw data was obtained by monitoring the CPS information side and physical side of the active distribution network through the real-time simulation platform. Then, the data discretization rules were formulated through the rated voltage and current values, and the original data were discretized and quantized according to the rules to generate event sequences. On this basis, the FP-Growth algorithm was used to mine the frequent items and strong correlations of abnormal signals in historical data, and new attack catego-

收稿日期：2024−06−07；修回日期：2024−10−15
基金项目：国网山西省电力公司科技项目（No.520530230009）
Foundation Item: The State Grid Shanxi Electric Power Company Science and Technology Project (No.520530230009)

ries and fault points were identified through the existing frequent sequence features. The detection of CPS network attack on active distribution network was realized. Finally, the feasibility and effectiveness of the proposed method were verified by a simulation experiment.

Key words: active distribution network, cyber physical system, network attack, FP-Growth algorithm, event sequence

0 引言

随着配电网容量和规模的持续扩大，以及分布式新能源和多元化负荷的不断接入，配电网由“无源”变为“有源”，潮流由“单向”变为“双向”，从而保证了发电侧与用电侧之间的双向信息交互[1-3]。分布式新能源的接入使配电网的结构、潮流和故障电流都发生了很大变化，有源配电网中信息物理系统的高度融合，进一步导致针对信息系统的网络攻击直接对物理系统的安全稳定运行构成威胁，进而加剧了电网故障检测和定位的复杂性[4-6]。

目前针对入侵检测系统，通常基于网络通信行为提取关键特征以检测与识别网络攻击。文献[7-8]对软件定义网络（software defined network，SDN）架构进行分析，提出熵与PSO-BP神经网络相结合的分布式拒绝服务（distributed denial of service，DDoS）攻击检测方法，提取异常开关的流量特征检查是否发生DDoS攻击。文献[9-10]通过建立虚假数据注入（false data injection，FDI）攻击模型，基于同步相量测量数据，提出基于集员滤波的自动发电控制系统FDI攻击检测方法，以及抵御FDI攻击的有效策略。上述研究仅限于DDoS攻击或FDI攻击等单一的攻击类型，对其他攻击类型并不适用。针对不同类型的网络攻击，有学者利用攻击树模型评估不同攻击场景的成功率[11]，提取系统中周期性通信行为特征，构建白名单、黑名单，以实现有效的入侵检测[12-13]。然而，上述方法只能检测到已知的异常行为，随着新的攻击不断出现，无法判断新的网络故障出现的原因及故障点位置。随着信息物理的高度融合，现有识别方法可以通过信息端的事件检测物理侧的异常行为，文献[14-16]针对智能设备，采用基于主机的分布式入侵检测系统，同时引入面向安全的信息物理状态估计系统，但攻击检测方法尚未结合有源配电网信息侧与物理侧的综合特征，且效率和准确率均不高。

考虑有源配电网脆弱性数据属于布尔型的离散数据，因此，本文选择了关联规则数据挖掘算法进行分析。现有研究多使用Apriori算法对数据进行关联性分析[17-19]，但该算法寻找频繁项集时遍历次数多，导致分析效率低、匹配能力不足。相比于Apriori算法，FP-Growth算法生成的FP-Tree只需要对数据遍历两次，有效提升了数据匹配效率和准确性。然而，FP-Growth算法大多用在对二次设备缺陷的关联性分析研究中[20-22]，对于有源配电网CPS网络攻击检测，此类关联规则算法没有得到很好的应用。

针对上述问题，本文考虑有源配电网信息物理系统的综合特征，提出基于FP-Growth算法的有源配电网通信系统网络攻击检测方法。首先分析有源配电网控制模型和CPS网络攻击模型，采用信息物理联合仿真平台获取原始攻击数据，并根据量化规则对原始数据进行离散量化处理，生成事件序列；然后将预处理后的数据导入FP-Growth算法进行有效关联分析，得到的频繁事件序列作为特征序列，以此作为攻击类型和故障点位置的识别依据；最后利用迁移学习提高匹配数据的准确性。

1 有源配电网网络攻击分析模型

1.1 总体框架

本文提出一种有源配电网信息物理系统网络攻击检测方法，总体实现框架如图1所示。以断

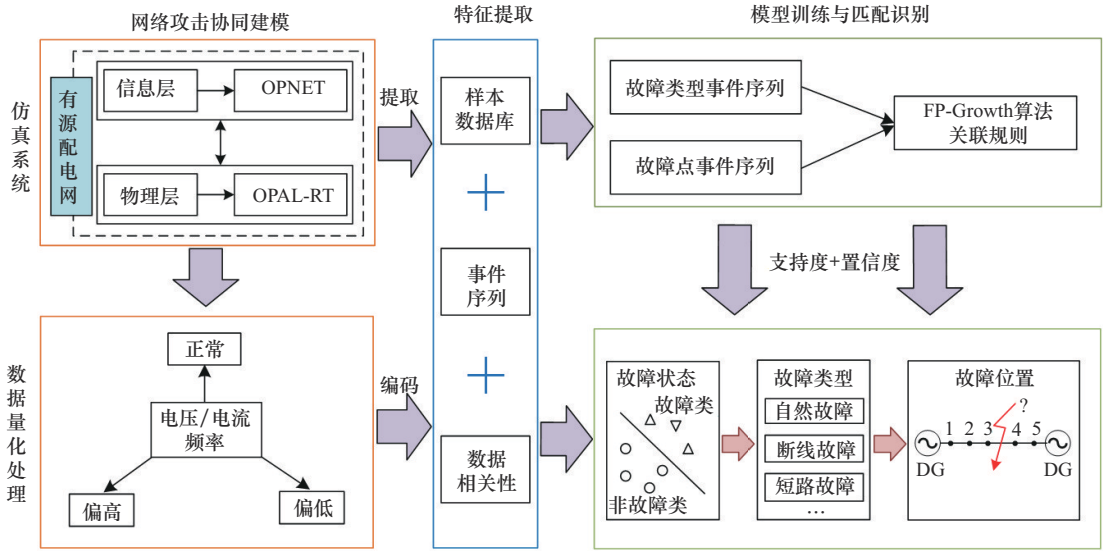

图1　总体实现框架

线攻击为例，首先，深入分析有源配电网网络攻击控制模型；其次，在仿真平台上设置特定场景，对网络攻击中的信息侧和物理侧数据进行精准采集；最后，根据事先设定的离散化规则对数据进行量化编码，生成涵盖不同故障类型和故障点的攻击事件序列。为了有效识别潜在的网络攻击，本文采用了 FP-Growth 算法对量化编码整合的数据进行样本训练，计算支持度和置信度。通过分析频繁事件序列，系统能够准确识别有源配电网的故障状态、故障类型和故障点位置。此外，本文通过匹配新的检测数据和生成频繁序列，使用迁移学习技术优化模型参数，提高匹配的精确性。

1.2　网络事件数据获取

随着分布式电源、分布式储能以及可控负荷的发展，配电网的结构发生了变化，电力系统各个环节中部署了具有通信和信息处理能力的智能电子设备，使得物理侧和信息侧高度耦合。为深入研究有源配电网在网络攻击场景下的表现，本文首先构建了有源配电网控制模型，有源配电网控制模型结构如图 2 所示。在该模型中，有源配电网的综合评价指标称为被控量，涵盖了经济指标、安全指标和供电质量指标等多个方面。被控量的具体数值受电网公司的保护动作、调度控制策略以及用户行为等多方面因素的综合影响，当考虑可能存在的恶意控制行为时，这些因素的综合作用导致有源配电网的源荷之间出现动态不平衡现象，这会使得被控量偏离其正常的参数要求范围，进而可能引发安全稳定事故。

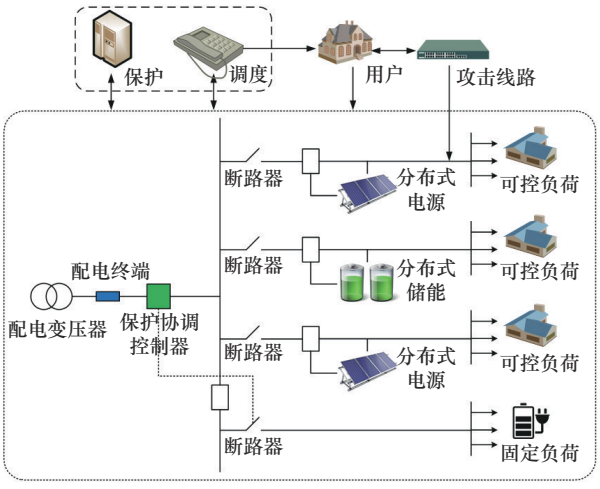

图2　有源配电网控制模型结构

基于有源配电网控制模型，本文所考虑的有源配电网 CPS 网络攻击模型如图 3 所示。电力网络实时上传系统信息（节点电压、电流等）和状

态信息（断路器开关状态等），信息网络控制中心下达控制指令，外部攻击行为作用于信息层信息侧，并对物理层物理侧产生影响，既包含信息元素，又包含物理元素，体现了网络攻击的跨空间性。

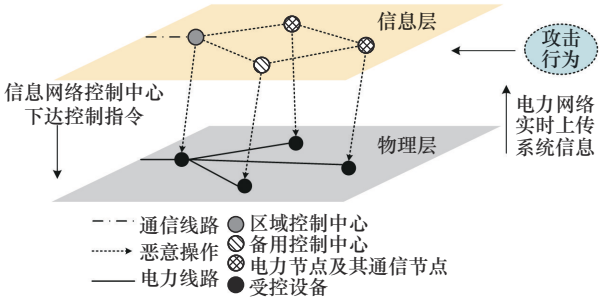

图3　有源配电网CPS网络攻击模型

1.3　仿真平台架构

本文建立包含电力系统仿真工具RT-LAB和信息通信系统仿真工具OPNET的信息物理联合仿真平台，对有源配电网信息物理系统网络攻击进行半实物仿真。有源配电网的实时监测管理由RT-LAB和变电站系统获取电力系统和可控负荷的数据来实现。电力仿真系统RT-LAB是一个分布式实时平台，具有成本低、采集信号响应速度快、操作性和可靠性强的特点，其量测单元模块通过确定采样频率和采集数据内容（如电压、电流、频率等），同时给数据内容加上时间戳，用于延时分析和序列产生等后续数据处理。包含TCP/IP和UDP的RT-LAB与外部通信使用协议是利用以太网实现的，本文主要采用TCP/IP方式来实现数据交互。通信仿真系统OPNET是一款离散事件网络仿真平台，采用软件在环（software in the loop，SITL）外部仿真接口方式。

本文通过在仿真平台上设置特定场景来获取网络攻击过程的原始实时数据，模拟事故过程（故障类/非故障类）发生的状态。信息物理系统联合仿真平台模型如图4所示。

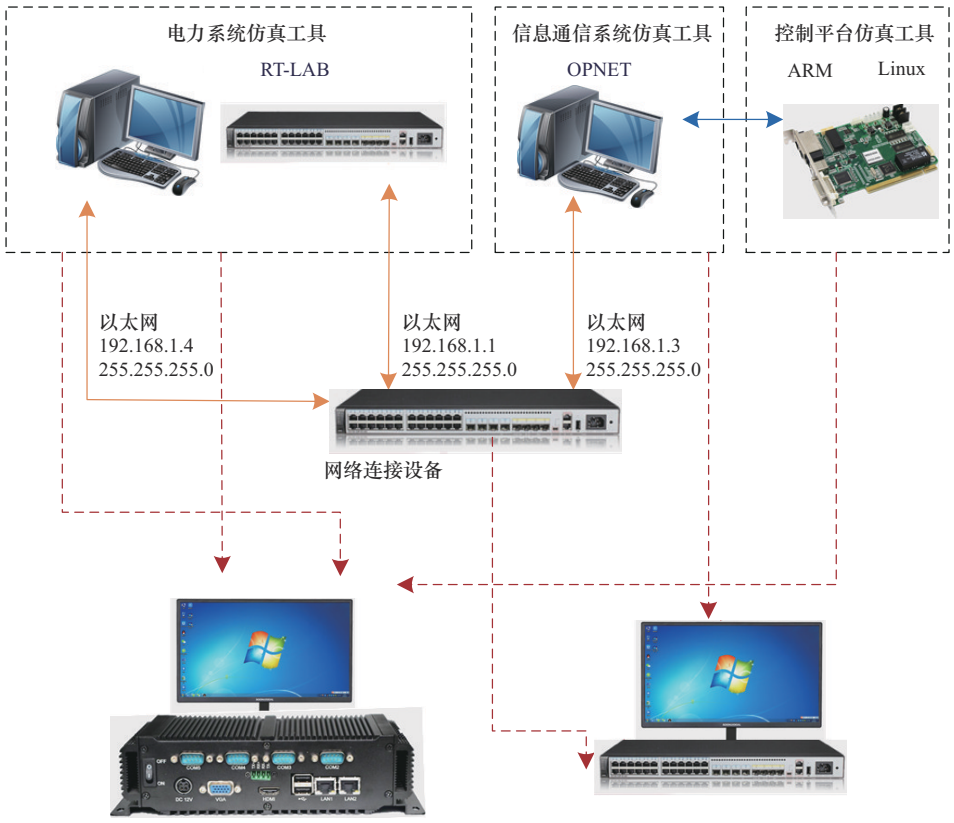

图4　信息物理系统联合仿真平台模型

2　数据处理与事件序列生成

通过电力系统仿真工具RT-LAB和信息通信系统仿真工具OPNET所获取的原始数据包含电压、电流、频率和通信状态等关键参数，这些关键参数具有周期性采样的特性，针对离散通信数据与连续物理数据之间的问题，本文采用了一种连续数据离散化的处理方法，解决通信和电力数据两者之间的时序和尺度差异。本文所使用的数据离散化方法如下。

2.1　数值离散化

（1）电压离散化

设备在额定电压（U_n）下正常运行，在传输线上，电压的波动范围为额定电压的±0.05倍之间，将得到的原始数据分为3个不同的量化级别，得到原始电压值与量化电压值的对应关系。电压离散量化对应关系见表1。

表1　电压离散量化对应关系

原始值	量化值	物理意义
$<0.95U_n$	2	偏低
$0.95U_n$～$1.05U_n$	1	正常
$>1.05U_n$	0	偏高

（2）电流离散化

由于配电网设备运行状态会影响电流值，支线的负荷水平在系统运行过程中呈现较大的变化，因此，为了评估系统的运行状况并检测是否存在故障异常，本文选择以初始稳态下的电流量测值作为量化的基准，将量化电流值划分为3个不同的等级。设定电流测量的初始稳态值为I_n，得到原始电流值与量化电流值的对应关系。电流离散化对应关系见表2。

表2　电流离散化对应关系

原始值	量化值	物理意义
$<0.85I_n$	2	偏低
$0.85I_n$～$1.15I_n$	1	正常
$>1.15I_n$	0	偏高

（3）频率离散化

我国规定电力系统的额定频率为50 Hz，允许频率在额定值的基础上有±0.2 Hz的波动。为了更有效地评估频率的运行状况，将原始值划分为3个不同的等级，得到原始值与量化值的对应关系。频率离散化对应关系见表3。

表3　频率离散化对应关系

原始值/Hz	量化值	物理意义
<49.8	2	偏低
49.8～50.2	1	正常
>50.2	0	偏高

2.2　时间离散化

为了将系统中连续的物理数据抽象为更有代表性的状态，以达到有效降低系统状态数量的目的，本文采用时间离散化的方法，将连续的时间过程分割为瞬时事件，从而可以更加直观地反映事件状态的改变。事件离散状态模型可表示为：

$$M=\left\{t,[\boldsymbol{U}_m],[\boldsymbol{I}_m],[\boldsymbol{f}_m]\right\} \tag{1}$$

其中，t表示一条事件记录的时间戳，$[\boldsymbol{U}_m]$、$[\boldsymbol{I}_m]$、$[\boldsymbol{f}_m]$分别表示电压、电流、频率量测值离散化后组成的向量。

2.3　通信数据处理

离散化处理信息侧的原始数据后，需要使其和物理侧的数据保持一致，得到通信状态量与量化值的关系。通信数据离散化对应关系见表4。

表4　通信数据离散化对应关系

通信状态量	量化值
断路器状态	0（打开）/1（闭合）
通信链路状态	0（正常）/1（不正常）

利用前文所述的物理侧和信息侧原始数据离散化方法，采用以下方式来表示协同攻击事件序列：

$$N=\left\{t,[\boldsymbol{U}_m],[\boldsymbol{I}_m],[\boldsymbol{f}_m],[\boldsymbol{X}_{\text{off}}],[\boldsymbol{X}_{\text{on}}]\right\} \quad (2)$$

其中，$[\boldsymbol{X}_{\text{off}}]$、$[\boldsymbol{X}_{\text{on}}]$表示不同通信状态的向量。

3 FP-Growth算法与网络攻击辨识方法

3.1 FP-Growth算法与关联规则

传统辨识恶意网络攻击的方法一般为黑名单或白名单方法，而对于已知的异常项集，传统方法常采用Apriori算法进行关联分析，但此类方法存在一些缺陷，如匹配率偏低、辨识度弱等，本文选择的FP-Growth算法则较好地解决了Apriori算法在处理过程中需要多次扫描数据库并生成大量候选项集，进而导致计算效率低的问题。

事务集合$S=\{S_1,S_2,\cdots,S_n\}$用来表示一系列项集中发生的事件，对应本文中的S事件序列候选项集。FP-Growth算法的核心思想在于构建FP-Tree，通过将原始数据集中的事务映射到该树上，从而发现频繁项集。构建FP-Tree的过程只需要对数据进行两次遍历，相对于Apriori算法而言，这有效提高了扫描效率。FP-Tree的建立和对其上的频繁项集进行挖掘是FP-Growth算法的主要步骤。

（1）构建事件S序列FP-Tree：依次遍历候选S序列项集，计算其支持度，并设置支持度和置信度的最小阈值S_{supmin}和C_{confmin}，并根据支持度将频繁项集S序列按降序排列；后续扫描数据集时，每个候选频繁项集都按次序处理，并依次创建分支，以此形成事件S序列的FP-Tree。

（2）挖掘频繁项集：为每个候选S序列项集找到其条件模式基，递归调用树结构，剪枝不满足最小阈值的项，并连接列举所有组合，直至形成单一路径的树结构。FP-Growth算法流程如图5所示。

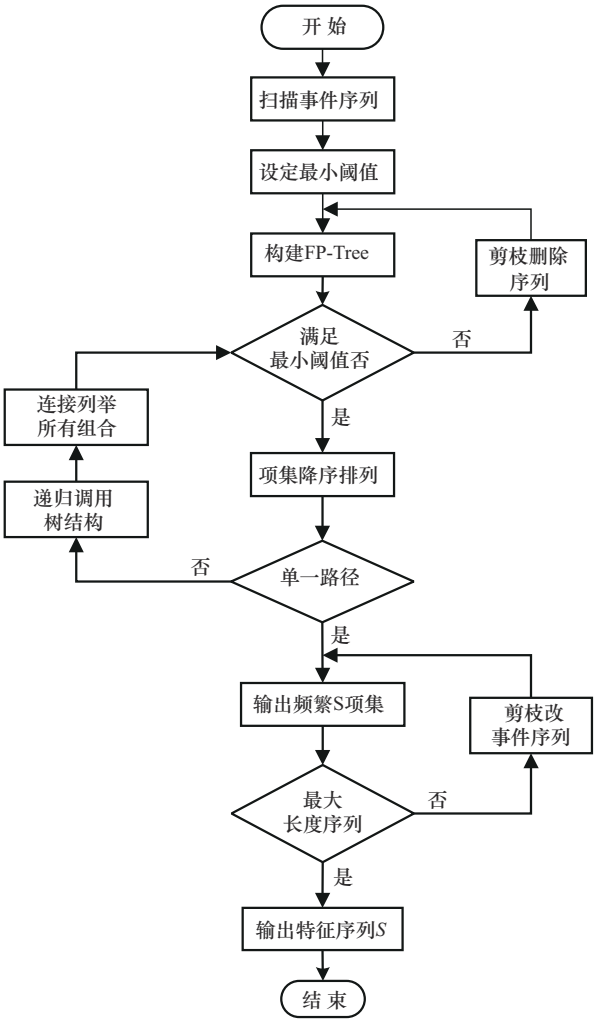

图5 FP-Growth算法流程

关联规则是一种在大型数据库中通过寻找频繁项集，进而发现变量之间的相互联系的数据挖掘方法，本文采用此方法能够迅速比对新的网络攻击事件序列和已经获得的特征序列，从而准确判断网络攻击的类别和故障点的位置。已得出的频繁特征序列用A表示，新的事件序列用B表示。已知特征序列A的支持度表示为：

$$\text{support}(A)=\frac{\text{count}(A)}{\text{count}(\text{Dateset})} \quad (3)$$

其中，count(A)表示特征序列A在事件序列数据库中出现的数量，count(Dateset)表示事件序列数据库的数量。关联规则$A\Rightarrow B$的置信度则表示为：

$$\text{confidence}(A\Rightarrow B)=\frac{\text{count}(A\cup B)}{\text{count}(A)} \quad (4)$$

其中，count($A \cup B$)表示关联规则$A \Rightarrow B$在事件序列数据库中出现的数量。

设置关联规则支持度和置信度的最小阈值，计算网络攻击事件序列的支持度，挖掘其频繁项集，并根据生成的事件特征序列，计算新的事件序列与特征序列的置信度，将两者对比匹配，从而判断攻击类型和故障点位置，以便采用相应的抑制措施。

3.2　攻击与故障事件序列生成

基于上文提出的事件序列S表达式，一个完整的事件过程可以通过状态转移过程来描述。事件状态转移矩阵如下。

$$\begin{bmatrix} & t & f & i & \cdots & im & v & \cdots & vm & \text{switch} & \cdots & \text{switch}m & \text{comm} & \cdots & \text{commm} \\ S1 & t1 & f1 & i1,1 & \cdots & i1,m & v1,1 & \cdots & v1,m & \text{switch}1,1 & \cdots & \text{switch}1,m & \text{comm}1,1 & \cdots & \text{comm}1,m \\ S2 & t2 & f2 & i2,1 & \cdots & i2,m & v2,1 & \cdots & v2,m & \text{switch}2,1 & \cdots & \text{switch}2,m & \text{comm}2,1 & \cdots & \text{comm}2,m \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ Sn & tn & fn & in,1 & \cdots & in,m & vn,1 & \cdots & vn,m & \text{switch}n,1 & \cdots & \text{switch}n,m & \text{comm}n,1 & \cdots & \text{comm}n,m \end{bmatrix} \tag{5}$$

其中，$S1, S2, \cdots, Sn$为一系列事件S，从$S1$到Sn，$t1, t2, \cdots, tn$为$S1, S2, \cdots, Sn$发生的时间，$f1, f2, \cdots, fn$为每个事件$S1, S2, \cdots, Sn$的频率，$i1, i2, \cdots, in$为每个事件$S1, S2, \cdots, Sn$的电流，$v1, v2, \cdots, vn$为每个事件$S1, S2, \cdots, Sn$的电压，$\text{switch}1,1, \cdots, \text{switch}1,m$为事件$S1$中第1个到第$m$个开关的状态，$\text{switch}n,1, \cdots, \text{switch}n,m$为事件$Sn$中第1个到第$m$个开关的状态，$\text{comm}1,1, \cdots, \text{comm}1,m$为事件$S1$中第1到第$m$个通信状态；$\text{comm}n,1 \cdots \text{comm}n,m$为事件$Sn$中第1个到第$m$个通信状态。

对于正常故障和网络攻击，完整的事件过程对应不同的表达式，这些表达式用于将不同故障类型和不同故障地点进行分类。若将基本状态序列作为理想序列，即$S1$-$S2$-$S3$-$S4$-$S5$-$S6$，那么可以基于此来定义其他非标准状态：$S1$-$S2$-$S4$-$S5$-$S6$表示数据丢失状态，$S1$-$S2$-$S3$-$S9$-$S4$-$S5$-$S6$表示额外状态，$S1$-$S2$-$S3$-$S4$-$S5$-$S6$表示延迟状态。

本文根据信息物理联合仿真平台得到原始攻击数据，并对数据进行量化离散处理，将预处理后的数据生成事件序列导入FP-Growth算法，提取频繁项集中支持度最高的项集，这些项集将同一类型事件按照一定的顺序排列形成的事件序列S视为该类型的事件特征序列，被用作后续对比匹配的依据，以识别新的事件序列是否符合该类型的特征。若不同的频繁候选序列都满足最小阈值，则取子序列最长的作为特征序列。所提出的基于FP-Growth算法的有源配电网信息物理系统网络攻击检测模型如图6所示。

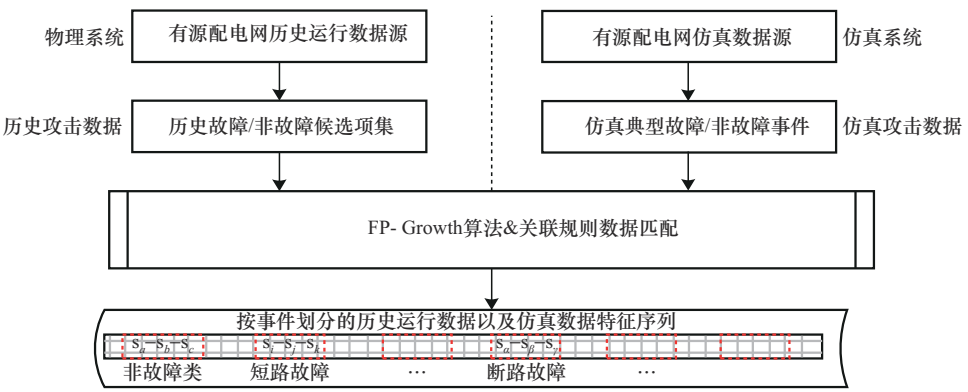

图6　基于FP-Growth算法的有源配电网信息物理系统网络攻击检测模型

4 算例仿真

4.1 攻击事件特征序列分析

为验证所提出的基于FP-Growth算法的有源配电网信息物理系统网络攻击检测方法的有效性和可靠性，本文以有源配电网断线攻击为例，网络攻击系统仿真模型如图7所示。采集信息物理联合仿真平台得到的原始数据，经过物理侧和信息侧的离散量化处理，通过事件状态转移形成事件序列，使用FP-Growth算法对预处理后的事件序列进行关联分析，生成事件特征序列。基于仿真中获得的特征事件序列中的未知事件可以分为正常故障和网络攻击两个类别。在匹配过程中，需要根据可能出现的情况适当调整设置的最小阈值，从而达到最佳的匹配效果。根据新生成的事件序列与得到的特征序列进行对比匹配，识别未知事件为正常故障或网络攻击。设置不同场景，得到不同类型的网络攻击及故障点位置特征序列，从而对不同的攻击类型和故障点进行识别。

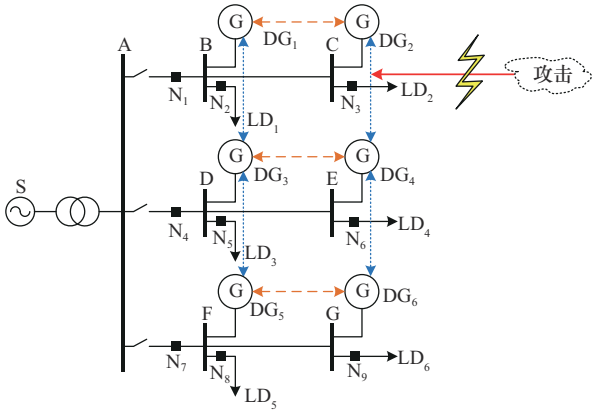

图7 网络攻击系统仿真模型

选取N_1、N_3、N_6、N_9不同时间节点的物理量测值，对其频率、电压、电流等原始数据进行分析，根据量化规则，将离散量化后的数据合并到同一序列中，以40%故障点为例，得到断线攻击事件序列量化处理结果，见表5。

表5 断线攻击事件序列量化处理结果

t	f	$v1$	$v2$	$v7$	$v8$	$i1$	$i2$	$i7$	$i8$
12.518 2	0	$2v_1$	1	2	2	0	1	1	2
12.519 2	0	2	1	1	2	0	1	1	2
12.520 2	1	2	1	1	1	0	1	1	2
12.521 2	1	2	1	1	1	0	1	1	1
…									
22.563 6	1	2	1	1	0	0	0	1	1
22.564 6	1	1	2	2	0	1	0	1	1
…							1		
32.769 5	1	1	1	1	2	0	1	1	1
32.770 5	1	1	1	2	1	0	2	1	0
…									
44.257 8	1	1	1	2	1	1	1	1	1
44.258 8	1	1	0	2	2	1	1	1	1

本文采用离散化处理，使用0、1、2的量化规则来处理连续数据，为保证不同类型的原始数据不会混合在同一事件序列中，设置不同故障点在联络线全长30%～80%的故障场景，将量化后的数据导入FP-Growth算法来构建FP-Tree，对数据库进行遍历，得到不同故障点位置的事件序列，并计算不同事件序列的支持度，得到最长且支持度最高的一条子序列作为特征序列，当新的故障事件序列生成时，将优先与该特征序列进行对比匹配。断线攻击故障事件序列及特征序列见表6。

表6 断线攻击故障事件序列及特征序列

故障点位置	事件序列	特征序列	支持度
30%	*S*10535-···- *S*11467	*S*10634-···- *S*11467	66.7%
40%	*S*9803-···-*S*9745	*S*9702-···- *S*9615	50.0%
50%	*S*12479-···-*S*11534	*S*11389-···- *S*12034	60.0%
60%	*S*12024-···-*S*9957	*S*10352-···- *S*9978	66.7%
70%	*S*9803-···-*S*10264	*S*9859-···- *S*10397	54.5%
80%	*S*9608-···- *S*9711	*S*9946-···- *S*9839	45.0%

通过设置不同的场景，在信息物理联合仿真平台仿真不同类型的网络攻击，以此来验证所提的基于FP-Growth算法的有源配电网信息物理系统网络攻击检测方法的可行性和有效性。本文根

据设计出的检测模型，在仿真平台模拟了多种DDoS攻击类型，包括Http Flood、ICMP Flood、UDP Flood、SYN Flood攻击。经过对原始数据的离散量化处理，生成DDoS攻击的事件序列，导入FP-Growth算法进行关联分析，得到不同类型的DDoS攻击故障事件序列及特征序列，见表7，以此作为新的故障事件的识别匹配依据。

表7　不同类型的DDoS攻击故障事件序列及特征序列

攻击类型	事件序列	特征序列	支持度
Http Flood	*S*18752-····- *S*8890	*S*8956-····- *S*9012	72.8%
ICMP Flood	*S*9578-····-*S*9573	*S*9368-····- *S*9593	66.7%
UDP Flood	*S*10157-····-*S*13167	*S*9047-····- *S*12064	80.0%
SYN Flood	*S*8476-····- *S*8078	*S*8904-····- *S*8484	82.6%

在信息物理联合仿真平台模拟信息泄露攻击故障事件，得到不同故障点位置的信息泄露攻击故障事件序列及特征序列，见表8。

表8　不同故障点位置的信息泄露攻击故障事件序列及特征序列

故障点位置	事件序列	特征序列	支持度
10%	*S*13534-····- *S*14793	*S*13498-····- *S*13268	90.3%
20%	*S*10258-····-*S*11267	*S*10376-····- *S*11209	85.6%
30%	*S*12486-····-*S*12826	*S*12406-····- *S*12467	82.8%
40%	*S*13826-····-*S*12793	*S*12379-····- *S*12972	78.5%
50%	*S*10756-····- *S*10267	*S*10286-····- *S*9996	73.7%

4.2　攻击类型和故障点匹配分析

基于上述得到的不同攻击类型和故障点位置的特征序列，将新生成的事件序列与特征序列进行对比匹配，计算其置信度，判断新的事件属于自然故障还是网络攻击。若属于网络攻击，继续检测识别其攻击类型及故障点位置。通过设置不同攻击场景，生成不同特征序列，利用迁移学习不断提高匹配的效率和准确性。不同攻击场景匹配率如图8所示。

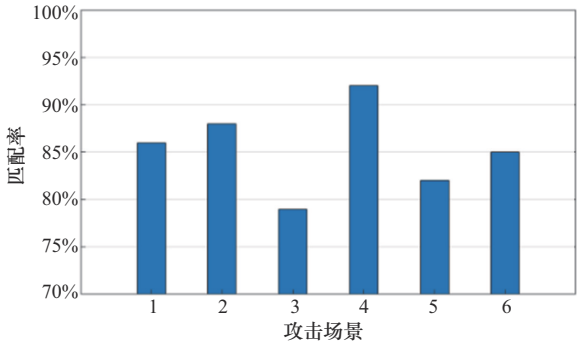

图8　不同攻击场景匹配率

4.3　不同算法计算效率对比

对信息物理联合仿真平台产生的原始攻击数据进行离散量化处理，并对生成的事件序列进行关联规则分析，传统关联规则的Apriori算法，需要多次扫描数据库，每次利用候选项集产生频繁项集，而本文采用的FP-Growth算法生成的FP-Tree只需要遍历数据库两次，有效提高了生成频繁项集的效率，两种算法的计算效率对比如图9所示。由图9可知，数据集在5 000左右时，两种算法计算效率差别不大，随着数据集的增加，从10 000增加到15 000时，FP-Growth算法相对于Apriori算法有了显著提高，且针对某些特定场景的网络攻击生成的事件序列，FP-Growth算法不仅计算效率高，识别匹配特征序列的准确性也有所提高。

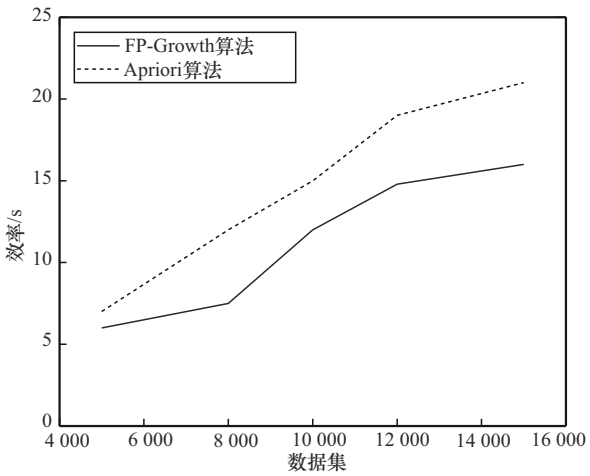

图9　两种算法的计算效率对比

5 结束语

本文提出了基于FP-Growth算法的有源配电网信息物理系统网络攻击检测方法，通过仿真算例验证本方法的卓越性，得出了以下结论。

（1）该方法能有效监测网络攻击实时数据，离散量化原始数据后，通过FP-Growth算法对事件序列进行关联分析，从而溯源攻击类型和故障点位置。

（2）该方法适用于各类型网络攻击场景，结合信息侧与物理侧双侧特点，使用FP-Growth算法有效提高了新事件与特征序列匹配的效率和准确性。

参考文献：

[1] 吴建旭, 于永进. 基于改进和声搜索算法的多目标配电网重构优化[J]. 电力系统保护与控制, 2021, 49(19): 78-86.
WU J X, YU Y J. Multi-objective distribution network reconfiguration optimization based on an improved harmony search algorithm[J]. Power System Protection and Control, 2021, 49(19): 78-86.

[2] 于洋, 汤伟, 叶远波, 等. 基于5G通信和改进DTW的配电网差动保护技术[J]. 哈尔滨理工大学学报, 2023, 28(4): 110-117.
YU Y, TANG W, YE Y B, et al. Distribution network differential protection technology based on 5G communication and improved DTW[J]. Journal of Harbin University of Science and Technology, 2023, 28(4): 110-117.

[3] 周成瀚, 邹贵彬, 杜肖功, 等. 基于正序电流故障分量的有源配电网纵联保护[J]. 中国电机工程学报, 2020, 40(7): 2102-2112, 2390.
ZHOU C H, ZOU G B, DU X G, et al. A pilot protection method based on positive sequence fault component current for active distribution networks[J]. Proceedings of the CSEE, 2020, 40(7): 2102-2112, 2390.

[4] 陈家琪, 王琦, 汤奕, 等. 考虑双侧特征的电力信息物理系统异常检测方法[J]. 电网技术, 2022, 46(6): 2339-2348.
CHEN J Q, WANG Q, TANG Y, et al. Anomaly detection method for cyber physical power system considering bilateral features[J]. Power System Technology, 2022, 46(6): 2339-2348.

[5] 汤奕, 李梦雅, 王琦, 等. 电力信息物理系统网络攻击与防御研究综述(二)检测与保护[J]. 电力系统自动化, 2019, 43(10): 1-9, 18.
TANG Y, LI M Y, WANG Q, et al. A review on research of cyber-attacks and defense in cyber physical power systems part two detection and protection[J]. Automation of Electric Power Systems, 2019, 43(10): 1-9, 18.

[6] 彭莎, 孙铭阳, 张镇勇, 等. 机器学习在电力信息物理系统网络安全中的应用[J]. 电力系统自动化, 2022, 46(9): 200-215.
PENG S, SUN M Y, ZHANG Z Y, et al. Application of machine learning in cyber security of cyber-physical power system[J]. Automation of Electric Power Systems, 2022, 46(9): 200-215.

[7] NAN D L, WANG W Q, ZHANG L, et al. Risk assessment of the operation state of power grid secondary equipment based on association rule mining and combination weighting-cloud model[J]. Power System Protection and Control, 2021, 49(10): 67-76.

[8] DINH P T, PARK M. BDF-SDN: a big data framework for DDoS attack detection in large-scale SDN-based cloud[C]//Proceedings of the 2021 IEEE Conference on Dependable and Secure Computing (DSC). Piscataway: IEEE Press, 2021: 1-8.

[9] 吴英俊, 汝英涛, 刘锦涛, 等. 基于集员滤波的自动发电控制系统虚假数据注入攻击检测[J]. 电力系统自动化, 2022, 46(1): 33-41.
WU Y J, RU Y T, LIU J T, et al. False data injection attack detection for automatic generation control system based on set-membership filtering[J]. Automation of Electric Power Systems, 2022, 46(1): 33-41.

[10] 何西, 涂春鸣, 于力. 配电网状态估计量测处理方法及抗虚假数据注入攻击策略[J]. 高电压技术, 2021, 47(7): 2342-2349.
HE X, TU C M, YU L. Measurements processing method of distribution network state estimation and anti false data injection attack strategy[J]. High Voltage Engineering, 2021, 47(7): 2342-2349.

[11] 丁明, 李晓静, 张晶晶. 面向SCADA的网络攻击对电力系统可靠性的影响[J]. 电力系统保护与控制, 2018, 46(11): 37-45.
DING M, LI X J, ZHANG J J. Effect of SCADA-oriented cyber attack on power system reliability[J]. Power System Protection and Control, 2018, 46(11): 37-45.

[12] 黄长慧, 胡光俊, 李海威. 基于URL智能白名单的Web应用未知威胁阻断技术研究[J]. 信息网络安全, 2021, 21(3): 1-6.
HUANG C H, HU G J, LI H W. Research on unknown threat blocking technology of Web application based on URL intelligent whitelist[J]. Netinfo Security, 2021, 21(3): 1-6.

[13] 彭勇, 向憧, 张淼, 等. 工业控制系统场景指纹及异常检测[J]. 清华大学学报(自然科学版), 2016, 56(1): 14-21.
PENG Y, XIANG C, ZHANG M, et al. Scenario fingerprint of an industrial control system and abnormally detection[J]. Jour-

nal of Tsinghua University (Science and Technology), 2016, 56(1): 14-21.

[14] MITCHELL R, CHEN I R. Behavior-rule based intrusion detection systems for safety critical smart grid applications[J]. IEEE Transactions on Smart Grid, 2013, 4(3): 1254-1263.

[15] 蒋正威, 张超, 孙伟乐, 等. 基于神经网络的电网假数据注入攻击检测方法研究[J]. 浙江电力, 2020, 39(11): 45-50.
JIANG Z W, ZHANG C, SUN W L, et al. Detection method of false data injection attack on power grid based on neural networks[J]. Zhejiang Electric Power, 2020, 39(11): 45-50.

[16] 杨剑锋, 秦钟, 庞小龙, 等. 基于深度学习网络的输电线路异物入侵监测和识别方法[J]. 电力系统保护与控制, 2021, 49(4): 37-44.
YANG J F, QIN Z, PANG X L, et al. Foreign body intrusion monitoring and recognition method based on Dense-YOLOv3 deep learning network[J]. Power System Protection and Control, 2021, 49(4): 37-44.

[17] 陈勇, 李胜男, 张丽, 等. 基于改进 Apriori 算法的智能变电站二次设备缺陷关联性分析[J]. 电力系统保护与控制, 2019, 47(20): 135-141.
CHEN Y, LI S N, ZHANG L, et al. Association analysis for defect data of secondary device in smart substations based on improved Apriori algorithm[J]. Power System Protection and Control, 2019, 47(20): 135-141.

[18] 张延旭, 胡春潮, 黄曙, 等. 基于 Apriori 算法的二次设备缺陷数据挖掘与分析方法[J]. 电力系统自动化, 2017, 41(19): 147-151, 163.
ZHANG Y X, HU C C, HUANG S, et al. Apriori algorithm based data mining and analysis method for secondary device defects[J]. Automation of Electric Power Systems, 2017, 41(19): 147-151, 163.

[19] CONG Y. Research on data association rules mining method based on improved apriori algorithm[C]//Proceedings of the 2020 International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE). Piscataway: IEEE Press, 2020: 373-376.

[20] 叶万余, 苏超, 罗敏辉, 等. 基于关联规则挖掘的输电线路缺陷状态预测[J]. 电力系统保护与控制, 2021, 49(20): 104-111.
YE W Y, SU C, LUO M H, et al. Transmission line defect state prediction based on association rule mining[J]. Power System Protection and Control, 2021, 49(20): 104-111.

[21] 刘思怡, 苏运, 张焰. 基于 FP-Growth 算法的 10kV 配电网分支线断线故障诊断与定位方法[J]. 电网技术, 2019, 43(12): 4575-4582.
LIU S Y, SU Y, ZHANG Y. Open-line fault diagnosis and positioning method for 10kV power distribution network branch line based on FP-Growth algorithm[J]. Power System Technology, 2019, 43(12): 4575-4582.

[22] 马月坤, 刘鹏飞, 张振友, 等. 改进 FP-growth 算法及其分布式并行实现[J]. 哈尔滨理工大学学报, 2016, 21(2): 20-27.
MA Y K, LIU P F, ZHANG Z Y, et al. Improved FP-growth algorithm and its distributed parallel implementation[J]. Journal of Harbin University of Science and Technology, 2016, 21(2): 20-27.

[作者简介]

李瑞（1981– ），男，博士，国网山西省电力公司电力科学研究院正高级工程师，主要研究方向为电力系统自动化及其智能数字化技术。

刘珊（1987– ），女，国网山西省电力公司电力科学研究院高级工程师，主要研究方向为电网数字化及网络攻防技术。

闫磊（1985– ），男，国网山西省电力公司高级工程师，主要研究方向为电网调度自动化及通信。