基于机器学习的心电图分类中的侧信道辅助推理泄漏

Jialin Liu电气与计算机工程，天普大学jialinliu@temple.edu

Ning Miao电气与计算机工程，加利福尼亚大学戴维斯分校nmiao@ucdavis.edu

Chongzhou Fang电气与计算机工程，加利福尼亚大学戴维斯分校czfang@ucdavis.edu

Houman Homayoun电气与计算机工程，加利福尼亚大学戴维斯分校hhomayoun@ucdavis.edu

Han Wang电气与计算机工程，天普大学han.wang.hw@temple.edu

arXiv:2304.01990v1 [cs.CR] 4 Apr 2023

摘要—**心电图（ECG）测量心脏产生的电活动，用于检测异常心跳和心脏病发作。然而，异常情况并非规律出现，因此需要连续监测心跳。为减少监测所需的人力工作，研究人员利用机器学习技术实现任务自动化。近年来，许多公司推出了具有心电图监测和异常心跳提醒功能的产品。在各种分类算法中，基于时间序列的算法动态时间规整（DTW）被广泛用于执行心电图分类任务。尽管取得了进展，基于动态时间规整的心电图分类也带来了新的攻击途径，即泄露患者诊断结果。本文表明，输入心电图样本的标签可通过侧信道攻击 Flush+Reload 窃取。具体而言，我们首先识别出动态时间规整在心电图分类中的漏洞，即扭曲路径选择与预测结果之间的相关性。随后，我们实现了一种攻击，利用 Flush+Reload 监测已知心电图数据的扭曲路径选择，并据此构建预测器，以建立扭曲路径选择与输入心电图样本标签之间的关系。实验结果表明，基于 Flush+Reload 的推理泄漏能够以 84.0% 的攻击成功率识别动态时间规整中两个样本的标签。**

索引词—**隐私泄露、动态时间规整、心电图、侧信道攻击**

I. 引言

心电图（ECG）反映心脏的电活动，广泛用于检测和诊断心律失常。传统上，训练有素的心脏病专家会通过目视检查心电图波形来发现异常。然而，心律失常间歇性发作，妨碍了持续检测和诊断。为应对这一挑战，基于机器学习的分类算法已被广泛用于检测心电图轨迹中的异常心跳 [12]–[25]。随着机器学习的发展，许多公司推出了具备心电图监测和不规则心律检测功能的产品。这些功能显著帮助人们追踪自身健康状况，也为医生确认诊断提供了更多参考。然而，基于机器学习的心电图分类也带来了新的安全挑战。机器学习算法的计算依赖于心电图轨迹的标签，而这些标签可能被恶意利用来推断样本的标签。

在本研究中，我们选取一种最常见的机器学习算法，即时间序列分类算法动态时间规整（DTW） [1], 作为案例研究，以展示其造成的严重隐私泄露威胁。由于即使数据集经过缩放和移位，DTW 仍具有鲁棒性，因此长期以来一直被用于判断心电图轨迹 [41] 是否正常。DTW 首先通过允许将一个时间序列非线性映射到另一个时间序列，寻找两个时间序列的最优对齐，从而使两者之间的距离最小。然后，它根据两条轨迹中对齐的数据点计算欧几里得距离。以图 1 所示的两条心电图时间序列为例，这两条心电图序列分别表示为 A（a_1，a_2，...，a_n）和 B（b_1，b_2，...，b_m）。序列 Ai 和序列 B j 中的数据点有三种扭曲选择：i+1、j+1和i+1&j+1，称为**扭曲方向**。三种扭曲方向中的选择依据公式 1 确定，其中会选取使距离最小的扭曲方向。序列 A 和序列 B 在一轮 DTW 计算中的所有扭曲方向称为**扭曲路径** 或**对齐路径**。随后，根据扭曲路径计算 A 与 B 之间的总距离，并将其用于分类。

$$\gamma(i,j) = (a_i, b_j)^2 + \min \begin{cases} \gamma(i-1,j) \\ \gamma(i-1,j-1) \\ \gamma(i,j-1) \end{cases} \quad (1)$$

我们发现，输入到 DTW 模型、属于同一类别的时序心电图序列，其扭曲路径往往更为相似。我们发现，在式（1）中从 $\gamma(i-1,\ j)$、$\gamma(i-1,\ j-1)$ 和 $\gamma(i,\ j-1)$ 中进行选择的计算取决于心电图轨迹的标签。这是因为标签相同的序列很可能具有相似的扭曲选择。尽管 DTW 中的计算行为不会直接泄露任何私人信息，但计算行为（即扭曲

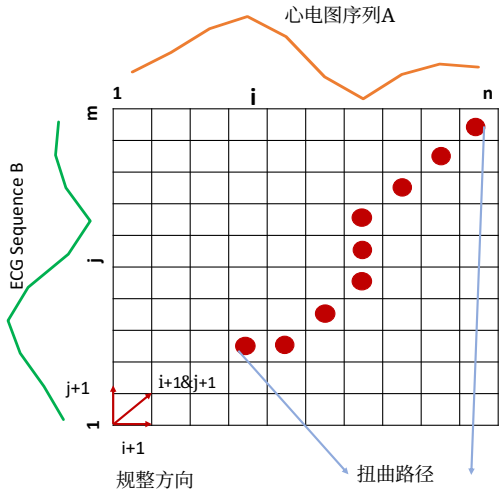


图 1：动态时间规整的规整方向和扭曲路径。

路径）与标签之间的相关性，暴露出一种新的漏洞：攻击者可以通过窥探扭曲对齐来泄露输入的标签。

基于扭曲路径与心电图样本标签之间的相关性，我们进一步利用侧信道攻击（SCA）Flush+Reload 监测规整方向，并推断发送至动态时间规整的输入标签。然而，大多数动态时间规整实现中的三个规整方向在内存位置上彼此相近，而 Flush+Reload 无法监测所有这些方向。为应对这一挑战，我们提出仅监测一个规整方向，并探索不同类型的机器学习分类器，以展示攻击者恢复标签的全部潜力。我们使用加州大学河滨分校时间序列仓库中的心电图数据集 [2]，发现基于 Flush+Reload 的动态时间规整心电图分类攻击能够成功推断输入标签，准确率最高可达 84.0%。

主　　要贡献总结如下　　如下。

- 据我们所知，这是首项探究基于动态时间规整的心电图分类中侧信道辅助推理泄漏，并揭示扭曲路径与两个输入序列标签之间相关性的研究。

- 本文进一步构建了一种基于 Flush+Reload 的推理攻击，可监测规整方向并推断输入的标签。

- 我们还评估了三种规整方向在区分标签方面的重要性，为未来的防御方法提供见解。

II. 背景与相关工作

A. 心电图（*ECG*）

心电图（ECG）是一种用于发现心律异常或心脏疾病的标准检查。传统上，医疗从业者通过目视检查患者的纸质记录来识别异常，这显著妨碍了对心脏病高风险人群进行长期监测，

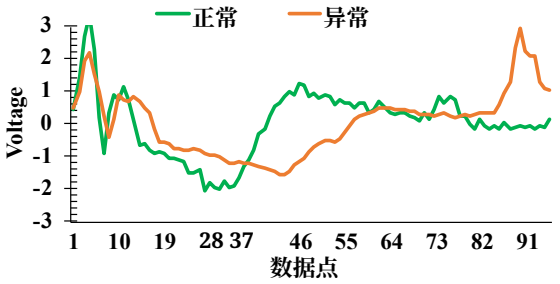


图 2：来自心电图200 [21]数据集的正常和异常心电图样本。

例如吸烟者。然而，心脏病的高死亡率要求持续进行心电图监测，以便及早发现并准确分类。业界和学术界都研究了基于机器学习的心电图分类，以解决这一问题。近年来，苹果、三星和华为均推出了心电图监测和分类产品。这为医生提供了重要指导，也使追踪更加容易。图2展示了来自 ECG 200 数据集 [21] 的一个正常样本和一个异常心电图样本。Y 轴表示心电图信号的电压强度，X 轴表示数据点。心电图样本的长度为 100。

B. 基于缓存的侧信道攻击

引入缓存层次结构，是为了弥合内存与 CPU 之间的延迟差距；这些缓存由多个应用程序和用户共享。共享缓存组件使攻击者能够通过制造争用并测量访问时间，观察受害应用程序的行为，从而推断受害者的数据访问模式。过去十年间，人们开发了多种基于缓存的侧信道攻击，包括 Flush+Reload[40], Flush+Flush [5], Prime+Probe [19], ，用于监视共享缓存活动并窃取密码、密钥等机密信息。本文利用 *Flush+Reload* 观察受害者 DTW 心电图分类模型的计算行为。*Flush+Reload* [40]通过监视共享页面中的内存访问行，利用页面去重技术的漏洞。该攻击针对 CPU 中的末级缓存，并持续刷新缓存中的受害应用程序数据。随后，攻击者尝试访问这些数据并测量访问时间（延迟）。访问时间的长短决定了受害应用程序是否访问过这些数据。

C. 相关工作

过去十年间，机器学习取得了显著进展，并已应用于疾病诊断 [7], [24],智能监控[14], [42], 金融决策 [13],等多个关键领域。安全与隐私问题日益受到关注。先前研究表明，利用面向硬件的攻击可以恢复架构、参数和输入 [10], [11], [16],[20], [23], [38], [39]。 [16], [28]。例如，Honget al. [9],

[10] 利用基于缓存的侧信道攻击 Flush+Reload，并监测推理阶段使用的层。该方法以 0% 的错误率成功恢复了 MalConv 和 ProxylessNAS-CPU。Hua 等人 [11]研究了硬件加速器上的内存侧信道泄漏，并发现内存访问模式可用于逆向分析卷积神经网络模型的结构和权重。除了机器学习设计之外，输入和标签也面临威胁。Xiang 等人[38]利用功耗侧信道，在基于FPGA的加速器上重建输入图像。该攻击在最具代表性的图像数据集之一 MNIST数据集 [15], 上进行评估，准确率达到 89%。Liu 等人 [16]指出，利用基于软件的侧信道遥测数据有助于推断输入图像的输出类别。Luo 等人 [20], 也开展了类似研究，他们利用统计学习模型建立了 Prime+Probe 获取的缓存访问模式与推理结果之间的相关性。推断出的推理结果还可结合自适应蒙特卡洛定位（AMCL）算法，进一步揭示车辆路线或位置。

上述研究展示了机器学习在面对侧信道攻击时存在的信息泄漏，但缺乏对受害模型的内部行为及由此产生的漏洞的分析。相比之下，我们的工作旨在通过 Flush+Reload 提供漏洞分析和一种概念验证攻击。

清单 1：包含输出指令的扭曲路径/*calculatetheEuclideandistanceofthethreewarpingdirections,choosetheonewithminimaldistancecost*/fp =fopen("warpingpath.txt","w");**double**i1j =DistanceCost(i+1,j);**double**i1j1 =DistanceCost(i+1,j+1);**double**ij1 =DistanceCost(i,j+1);

```
if (i1j < i1j1 && i1j < ij1){
        prev_cost = i1j;
        fprintf(fp, "i+1\n");}

否则如果(ij1 <i1j&&ij1 <i1j1){prevcost =ij1;fprintf(fp,"j+1\n");}

否则 {
        prev_cost = i1j1;
        fprintf(fp, "i+1&j+1\n");}
```

III. 基于动态时间规整的心电图分类漏洞

本节将介绍心电图轨迹的扭曲路径与标签之间的相关性。为此，我们考察加州大学河滨分校时间序列仓库中的心电图数据集[2], , 即 ECG200 [21]。接下来，我们将首先展示来自不同类别的心电图序列中，$i+1$、$j+1$、$i+1$和$j+1$的扭曲路径分布。

我们探究动态时间规整的扭曲路径是否与输入心电图样本的标签相关。为此，我们记录标签相同和标签不同的样本所对应的扭曲方向，即 $i+1$、$j+1$、$i+1$和$j+1$。以 ECG200 为例，我们将收集两条正常样本、两条异常样本，以及一条正常样本与一条异常样本之间进行动态时间规整时的规整路径方向。我们在三种扭曲方向代码中插入一条输出指令，以获得清晰的扭曲路径，如列表1所示。为简便起见，我们将三种动态时间规整输入称为：正常、异常和混合。正常表示动态时间规整所处理的两条心电图序列均为正常样本；异常表示两条心电图序列均为异常样本；混合表示一条心电图序列为正常样本，另一条为异常样本。我们统计每 1000 次扭曲选择中各扭曲方向出现的次数。由于心电图样本的长度为 100，两条心电图序列之间的扭曲选择总数约为 100×100 次。

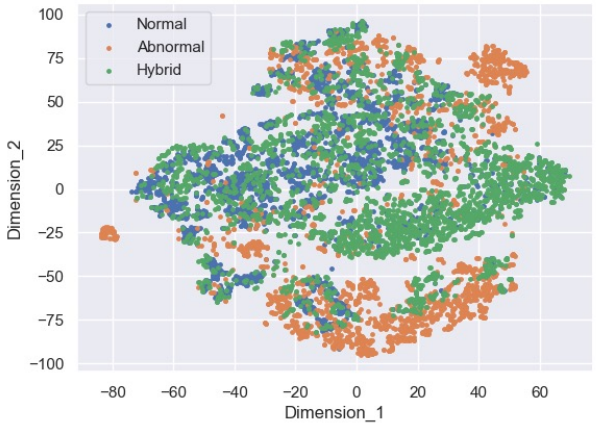


图 3：正常、异常和混合类型输入心电图样本的动态时间规整扭曲选择 t-SNE 图。

我们评估了发送至 DTW 的三类心电图输入样本所对应的三种路径扭曲方向的总数，即 $i+1$、$j+1$、$i+1$&$j+1$。我们利用 t-SNE [26]探究 DTW 的两个具有不同标签的心电图输入样本之间的扭曲方向是否具有可分性。t-SNE 可以将扭曲选择转换为二维图，如图 3 所示。我们发现，不同心电图样本之间的扭曲选择是可分的。这也符合这样的直觉：具有相同标签的样本很可能以相似的方式进行扭曲。扭曲选择的可分性使我们能够根据对扭曲方向的观测，推断发送至 DTW 的两个心电图样本的标签。三类输入之间规整路径分布差异的根本原因在于扭曲方向选择算法。如列表1所示，距离代价决定了 $i+1$、$j+1$、$i+1$&$j+1$之间的扭曲方向选择。因此，扭曲方向选择仍然包含距离代价的信息，而这也是基于动态时间规整的分类的基础。

IV. 攻击实现

基于第三节中所述的漏洞，我们实现了一种基于缓存的侧信道攻击 Flush+Reload，用于窃取规整路径方向，并推断发送至 DTW模型 的心电图样本标签。该实现包括两个主要部分：规整路径收集和基于机器学习的分类模型。

A. 威胁模型

受害目标是一个用于预测心电图样本的DTW分类模型。本文提出的攻击者旨在通过监测扭曲路径的选择，推断发送至DTW的心电图样本的标签。我们利用基于缓存的Flush+Reload来监视DTW的计算，并假设攻击者（没有sudo权限）与受害DTW模型位于同一台物理机器上。这种共置可通过利用近期工作 [4] 在云端实现，也可通过诱使用户下载恶意软件在边缘设备上实现。我们假设攻击者无法直接访问输入心电图样本、DTW分类模型和预测结果，即该心电图分类模型在软件层面是安全的。为了建立侧信道观测与心电图样本标签之间的关联，攻击者还需要能够查询受害者的基于DTW的心电图分类模型，并拥有自己的已标注心电图样本来查询该模型。

我们考虑的威胁模型在现实世界中同样具有可行性。当企业推出心电图监测和分类产品时，攻击者可以购买一款产品，植入Flush+Reload，并使用自己的已标注心电图数据对其进行查询，以收集扭曲路径，建立侧信道观测与心电图样本标签之间的关联。一旦攻击者诱使用户下载包含Flush+Reload的恶意代码，就可以利用这种关联推断其他良性用户心电图样本的标签。

表I：不同分割数和扭曲方向下随机森林的分类准确率。

分割数	$i+1$	$j+1$	$i+1$&$j+1$	完整扭曲路径
100	81.1	83.7	84.8	85.1
10	71.5	69.2	74.3	79.4
1	37.4	43.6	38.5	53.1

B. 通过刷新+重新加载获取路径选择

我们将使用攻击者的心电图样本，通过 Flush+Reload 收集观测到的规整路径，并利用该数据集训练分类模型以推断标签。首先，我们按照列表1所示实现 Flush+Reload，将目标设为规整方向选择代码。然而，一个挑战是，三个方向选择代码对应的内存位置过于接近。Flush+Reload 无法获取这三个扭曲方向的执行观测。因此，我们只能监测三种路径选择之一，并统计每个特定时间间隔内的出现次数。根据对心电图规整路径的分析，可以确定 Flush+Reload 跟踪记录中要监测的规整方向和时间间隔。

如表I所示，我们研究了选择不同扭曲方向和分割数量对标签分类准确率的影响。随后将根据分割数量把一条规整路径跟踪记录划分为多个部分，每部分的时间间隔相同。分割数量为1时，跟踪记录不进行划分。表I显示，分割数量从1增加到100时，分类准确率随之提高。另一个观察结果是，在所有三种分割设置下，完整的规整路径都能取得最高准确率。尽管如此，选择最优的监测规整方向可以缩小准确率差距。因此，本文选择100作为分割数量，并选择 $i+1$&$j+1$ 作为监测的规整方向，因为在区分心电图标签方面， $i+1$&$j+1$ 优于$i+1$ 和 $j+1$ 。

C. 基于机器学习的分类模型

我们将上一步收集的整个数据集划分为 70% 的训练集和 30% 的测试集。此外，我们在训练数据集上进行十折交叉验证，以构建一个准确率高且稳健的分类器，避免过拟合问题。此外，本研究考察了大量机器学习分类器，以选出准确率最高的最优分类器。我们选取了五种分类器：基于树的随机森林（RF）、基于支持向量机的序列最小优化、基于规则的 JRiP、多层感知机（MLP）和基于惰性学习的 IBK。我们使用 Weka [6] 实现这五种分类算法。

D. 在线攻击流程

分类模型训练完成后，攻击者即可进行在线部署。在线攻击流程包括以下步骤：

- 插入 Flush+将 Reload 恶意代码植入同样提供基于动态时间规整的心电图分类服务的目标设备。
- 预处理 Flush+Reload 观测结果，并将每条跟踪记录划分为 100 个时间间隔相等的片段。
- 将处理后的数据发送给攻击者拥有的远程分类器。
- 利用分类器推断 DTW 输入心电图样本的标签。泄露的标签可用于其他恶意活动。

V. 评估

A. 实验设置

1) 攻击平台：所有评估均在一台戴尔台式电脑上进行，该电脑配备 12 个英特尔 I5-10400 核心和三级缓存系统。其 L1 和 L2 缓存彼此独立，容量分别为 192 KB 和 1.5MB。L3 缓存存储器由所有核心共享，容量为 12 MB。L3 缓存的包含性使攻击者能够观察在同一台机器上运行的受害应用程序的缓存行为。在本文中，我们利用 Flush+Reload，它通过共享的 L3 缓存实施攻击。

表II：使用随机森林的混淆矩阵。

预测 / 实际	正常	异常	混合	假阳性	攻击成功率(%)
正常	2204	2	426	0.143	83.7
异常	6	643	48	0.001	92.3
混合	470	5	2166	0.142	82.0

2) 受害者数据集：如第 II 节所述，我们从加州大学河滨分校时间序列数据集库[2]中选取 ECG 200 数据集。ECG 200 包含 200 个样本，其中 133 个标记为正常，67 个标记为异常，每个样本的长度为 100。因此，ECG 200 为正常样本的 DTW 计算提供 $\binom{133*132}{2}$ 条路径，为异常样本的 DTW 计算提供 $\binom{67*66}{2}$ 条路径，并为混合样本之间的 DTW 计算提供 $133 * 67$ 条路径。我们使用攻击成功率（即分类准确率）来衡量所提出攻击的有效性。

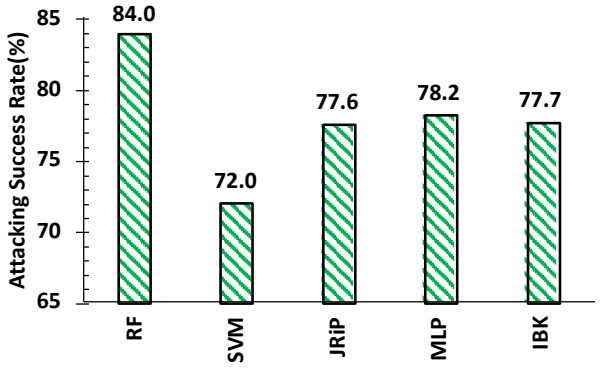


图 4：不同分类器的攻击成功率。

B. 不同分类器的攻击成功率

我们研究了不同的分类算法，以选出最优算法，从而根据 Flush+Reload 观测推断受害者模型的推理结果。图 4展示了来自不同类别的五种分类算法的攻击成功率。我们发现，所有五种分类器都能有效推断输入心电图样本的标签，攻击成功率均超过 70%。其中，随机森林优于其余四种算法，成功率达到 84.0%，这表明对扭曲路径的 Flush+Reload 观测与心电图样本的标签高度相关。根据这些结果，随机森林是攻击者推断标签最有效的分类器，因此将在后续评估中使用。

此外，对于 $i+1\&j+1$ 方向，利用侧信道观测得到的分类准确率仅略低于表I中使用实际扭曲路径得到的准确率，分别为 84.0% 和 84.8%。对于规整方向 $i+1\&j+1$，分类准确率略有差距，这是因为 Flush+Reload 观测仍会受到瞬态执行或其他正在运行的程序所产生的中断的噪声影响，尽管 Flush+Reload 旨在监视程序中的特定行为。

C. 攻击置信度

除了总体攻击成功率外，我们还通过表II中的随机森林混淆矩阵评估了三种动态时间规整输入的置信度。我们发现，大多数误分类来自正常类和混合类。异常类的推断准确率最高，达到 92.3%，且假阳性率最低，为 0.001。尽管随机森林的总体攻击成功率为 84.0%，但所呈现的攻击者在异常类上的准确率可达 92.3%，同时假阳性率较低，仅为 0.001。该混淆矩阵进一步揭示了通过侧信道攻击监控扭曲路径而导致的基于动态时间规整的心电图分类泄露威胁。

VI. 防御讨论

部署有效的防御方法，对于应对针对基于机器学习的医疗应用的侧信道辅助隐私攻击至关重要。

一种解决方案是部署有效的检测模块，以获知正在遭受侧信道攻击。先前研究 [8], [29]–[31], [33], [35] 提出利用硬件性能计数器构建侧信道攻击检测器，保护计算机系统免受隐私泄露。另一种应对隐私受损的方法是设计缓解措施并干扰侧信道观测，以保护医疗信息。近期已有多项研究提出修改缓存层次结构或缓存存储器架构，以缓解侧信道攻击。缓存分区技术 [3], [17], [22]通过静态或动态地为每个应用程序进程划分缓存存储器，旨在缓解基于缓存的侧信道攻击，从而使攻击者无法观测受害应用程序的“侧信道信息”。一些研究采用访问随机化 [18], [36], [37] ，主要通过随机化缓存干扰、重新映射缓存索引，或以随机缓存填充取代按需提取，来消除硬件架构中的安全漏洞。其余防御方法 [27], [32], [34] 则提出调整硬件和系统参数设置，包括频率和预取器，以增加侧信道观测中的噪声。

VII. 结论

心电图已广泛用于测量心跳并诊断心律不齐。基于机器学习的分类技术使自动异常检测成为可能，并显著减少了对医务人员的需求。然而，这些技术也带来了新的安全挑战。本研究以时间序列分类器动态时间规整为例，揭示了规整方向与心电图样本的

标签之间的相关性。我们利用 Flush+Reload 监测规整方向，设计了一种利用这种相关性的攻击，并采用机器学习分类器推断动态时间规整输入的标签。结果表明，所提出的攻击成功率最高可达 84.0%。进一步分析该攻击的混淆矩阵发现，对于异常输入，攻击成功率最高，达到 92.3%。本研究揭示了基于动态时间规整的心电图分类中的隐私泄露，并为未来的防御方法提供了参考。

参考文献

[1] D. J. Berndt 和 J. Clifford，《使用动态时间规整在时间序列中查找模式》。载于*KDD*研讨会，第10卷，第16期。美国华盛顿州西雅图，1994年，第359–370页。[2] H. A. Dau、E. Keogh、K. Kamgar、C.-C. M. Yeh、Y. Zhu、S. Gharghabi、C. A. Ratanamahatana、Yanping、B. Hu、N. Begum、A. Bagnall、A. Mueen、G. Batista 和 Hexagon-ML，《UCR时间序列分类档案》，2018年10月，https://www.cs.ucr.edu/~ –eamonn/time series data 2018/.[3] L. Domnitser、A. Jaleel、J. Loew、N. Abu-Ghazaleh 和 D. Ponomarev，《不可独占的缓存：以低复杂度缓解缓存侧信道攻击》，《*ACM*体系结构与代码优化汇刊》（*TACO*），第8卷，第4期，第35页，2012年。[4] C. Fang、H. Wang、N. Nazari、B. Omidi、A. Sasan、K. N. Khasawneh、S. Rafatirad 和 H. Homayoun，《Repttack：利用云调度器引导共置攻击》，*arXiv*预印本 *arXiv:2110.00846*，2021年。[5] D. Gruss 等人，《Flush+ flush：一种快速且隐蔽的缓存攻击》，载于入侵与恶意软件检测及漏洞评估国际会议，2016年。[6] M. Hall、E. Frank、G. Holmes、B. Pfahringer、P. Reutemann 和 I. H. Witten，《Weka数据挖掘软件：更新》，*ACM SIGKDD*探索通讯，第11卷，第1期，第10–18页，2009年。[7] S. Hassantabar、M. Ahmadi 和 A. Sharifi，《使用卷积神经网络方法，根据肺部X光图像诊断和检测COVID-19患者的感染组织》，《混沌、孤子与分形》，第140卷，第110170页，2020年。[8] C. U. M. P. D. C. U. W. H. D. C. U. R. S. S. J. C. U. Homayoun，Houman（圣何塞），《双层侧信道攻击检测方法及设备》，专利 20 230 092 190，2023年3月。 [在线]。可访问：https://www.freepatentsonline.com/y2023/0092190.html [9] S. Hong、M. Davinroy、Y. Kaya、D. Dachman-Soled 和 T. Dumitras.，《如何利用闲暇时间攻破0*06776*、2020*2002*wn NAS》，*arXiv*预印本 *arXiv:* 。[10] S. Hong、M. Davinroy、Y. Kaya、S. N. Locke、I. Rackow、K. Kulda、D. Dachman-Soled 和 T. Dumitras.，《存在缓存侧信道攻击时运行的深度神经网络安全分析》，*arXiv*预印本 *arXiv:1810.03487*，2018年。[11] W. Hua、Z. Zhang 和 G. E. Suh，《通过侧信道信息泄露对卷积神经网络进行逆向工程》，载于*2018*年第*55*届*ACM/ESDA/IEEE*设计自动化会议（*DAC*）。IEEE，2018年，第1–6页。[12] B. Huang 和 W. Kinsner，《使用动态时间规整进行心电图帧分类》，载于*IEEE CCECE2002*。加拿大电气与计算机工程会议。会议论文集（目录号 *02CH37373*），第2卷。IEEE，2002年，第1105–1110页。[13] G. Jeong 和 H. Y. Kim，《使用深度Q学习改进金融交易决策：预测股票数量、行动策略及迁移学习》，《专家系统与应用》，第117卷，第125–138页，2019年。[14] C.-B. Jin 等人，《面向静态视频监控摄像头，利用时间图像上的卷积神经网络进行实时人体动作识别》，载于泛太平洋多媒体会议。Springer，2015年。[15] Y. LeCun，《MNIST手写数字数据库》，*http://yann.lecun. com/exdb/mnist/*。[16] C. Liu、M. Kar、X. Wang、N. Chawla、N. Roggel、B. Yuce 和 J. M. Fung，《评估通过软件可访问的遥测数据发生的信息泄露的方法》，载于*2021*年*IEEE*硬件安全与可信国际研讨会（*HOST*）。IEEE，2021年，第259–269页。[17] F. Liu 等人，《Catalyst：在云计算中抵御末级缓存侧信道攻击》，载于*2016*年*HPCA*。

[18] F. Liu 和 R. B. Lee，《随机填充缓存架构》，载于第47届IEEE/ACM 国际微体系结构研讨会，2014年。[19] F. Liu、Y. Yarom、Q. Ge、G. Heiser 和 R. B. Lee，《末级缓存侧信道攻击在实践中可行》，载于2015年 IEEE S&P 研讨会，2015年。[20] M. Luo、A. C. Myers 和 G. E. Suh，《利用缓存侧信道隐蔽追踪自动驾驶车辆》，载于第29届 {USENIX} 安全研讨会（{USENIX} 安全 20），2020年，第859–876页。[21] R. T. Olszewski，《面向时间序列数据结构模式识别的广义特征提取》。卡内基梅隆大学，2001年。[22] D. Page，《分区缓存架构作为一种 侧信道防御机制》，2005年。[23] K. Patwari、S. M. Hafiz、H. Wang、H. Homayoun、Z. Shafiq 和 C.-N. Chuah，《针对 CPU-GPU 边缘设备的 DNN 模型架构指纹识别攻击》，载于2022年 IEEE 第7届欧洲安全与隐私研讨会（EuroS&P）。IEEE，2022年，第337–355页。[24] A. H. Ribeiro、M. H. Ribeiro、G. M. Paix˜ao、D. M. Oliveira、P. R. Gomes、J. A. Canazart、M. P. Ferreira、C. R. Andersson、P. W. Macfarlane、W. Meira Jr 等，《使用深度神经网络自动诊断12导联心电图》，Nature Communications，第11卷，第9期，2020年1，1期，第–页。[25] V. Tuzcu 和 S. Nas，《动态时间规整：一种用于识别心律失常中心电图变化的新工具》，载于2005年 IEEE 系统、人和控制论国际会议，第1卷。IEEE，2005年，第182–186页。[26] L. Van der Maaten 和 G. Hinton，《使用 t-SNE 可视化数据》。机器学习研究杂志，第9卷，第11期，2008年。[27] H. Wang，《用于分析和防御侧信道威胁的应用机器学习》。加利福尼亚大学戴维斯分校，2022年。[28] H. Wang、S. M. Hafiz、K. Patwari、C.-N. Chuah、Z. Shafiq 和 H. Homayoun，《通过基于缓存的侧信道攻击对 DNN 发起隐蔽推理攻击》，载于2022年欧洲设计、自动化与测试会议及展览（DATE）。IEEE，2022年，第1515–1520页。[29] H. Wang、S. Salehi、H. Sayadi、A. Sasan、T. Mohsenin、P. S. Manoj、S. Rafatirad 和 H. Homayoun，《评估基于机器学习的检测方法对自动驾驶车辆侧信道攻击的防御效果》，载于2021年 IEEE 第3届人工智能电路与系统国际会议（AICAS）。IEEE，2021年，第1–4页。[30] H. Wang、H. Sayadi、S. M. P. Dinakarrao、A. Sasan、S. Rafatirad 和 H. Homayoun，《在硬件层面实现微型人工智能以保障边缘设备安全》，IEEE电路与系统新兴及精选主题期刊，第11卷，第4期，第803–815页，2021年。[31] H. Wang、H. Sayadi、G. Kolhe、A. Sasan、S. Rafatirad 和 H. Homayoun，《Phased-Guard：用于检测和识别零日微体系结构侧信道攻击的多阶段机器学习框架》，载于2020年 IEEE 第38届计算机设计国际会议（ICCD）。IEEE，2020年，第648–655页。[32] H. Wang、H. Sayadi、T. Mohsenin、L. Zhao、A. Sasan、S. Rafatirad 和 H. Homayoun，《通过随机化缓解基于缓存的侧信道攻击：系统与架构级综合分析》，载于2020年欧洲设计、自动化与测试会议及展览（DATE）。IEEE，2020年，第1414–1419页。[33] H. Wang、H. Sayadi、S. Rafatirad、A. Sasan 和 H. Homayoun，《Scarf：利用底层硬件特征实时检测侧信道攻击》，载于2020年 IEEE 第26届在线测试与鲁棒系统设计国际研讨会（IOLTS）。IEEE，2020年，第1–6页。[34] H. Wang、H. Sayadi、A. Sasan、S. Rafatirad 和 H. Homayoun，《Hybrid-Shield：用于运行时检测和缓解基于缓存的侧信道攻击的精准高效跨层对策》，载于第39届计算机辅助设计国际会议论文集，2020年，第1–9页。[35] ——，《HybridG：结合动态时间规整与高斯分布模型检测新出现的零日微体系结构侧信道攻击》，载于2020年 IEEE 第19届机器学习与应用国际会议（ICMLA）。IEEE，2020年，第604–611页。[36] Z. Wang 和 R. B. Lee，《一种兼具更高性能与安全性的新型缓存架构》，载于第41届IEEE/ACM 国际微体系结构研讨会论文集。[37] ——，《用于阻止软件缓存侧信道攻击的新型缓存设计》，载于 ACM SIGARCH计算机体系结构新闻，2007年。[38] L. Wei、B. Luo、Y. Li、Y. Liu 和 Q. Xu，《我知道你看到了什么：针对卷积神经网络加速器的功耗侧信道攻击》，载于第34届年度计算机安全应用会议论文集，2018年，第393–406页。

[39] M. Yan 等，《缓存心灵感应：利用共享资源攻击学习 {DNN} 架构》，载于第*29*届安全研讨会（{*USENIX*} 安全 *20*），2020年，第2003–2020页。[40] Y. Yarom 和 K. Falkner，《Flush+ reload：一种高分辨率、低噪声的 L3 1，2014，22 25缓存侧信道攻击。》，载于*USENIX*安全研讨会，第卷，第–页。[41] G. Zhang、W. Kinsner 和 B. Huang，《基于动态时间规整匹配和帧分类的心电图数据挖掘》，《生物力学与生物医学工程中的计算机方法》，第12卷，第6期，第701–707页，2009年。[42] L. Zhou 等，《利用深度神经网络对智能电网中的网络攻击进行分类》，载于第二届计算机科学与应用工程国际会议论文集，2018年，第1–5页。