arXiv:2601.00384v1 [cs.CR] 1 Jan 2026

增材制造中的工程攻击向量与异常检测*

MdMahbub Hasan[1 [0009−0002−2226−6494]]、Marcus Sternhagen[1 [0009−0001−5605−6631]]和 Krishna Chandra Roy[1 [0000−0001−9388−8042]]

新墨西哥矿业与技术学院{mdmahbub.hasan,marcus.sternhagen}@student.nmt.edu, krishna.roy@nmt.edu

摘要。 增材制造（AM）正迅速应用于航空航天、汽车和医疗保健等关键领域。然而，这种信息物理融合带来了新的攻击面，尤其是在计算机辅助设计（CAD）与机器执行层之间的接口处。本文研究了针对两种广泛使用的熔融沉积成型（FDM）系统发起的定向网络攻击：Creality 旗舰机型 K1 Max 和 Ender 3。我们的威胁模型是一种多层中间人攻击（MitM）入侵，攻击者在用户界面向打印机固件上传G代码文件的过程中将其截获并篡改。该中间人攻击入侵链可实现多种隐蔽的破坏场景。这些攻击无法被传统切片软件或运行时接口检测到，导致打印部件在结构上存在缺陷，外观却看似正常。为应对这些隐蔽威胁，我们提出了一种无监督入侵检测系统（IDS），用于分析实时打印过程中生成的结构化机器日志。我们的防御机制采用冻结的基于Transformer的编码器（一种BERT变体）提取系统行为的语义表示，随后使用经对比训练的投影头学习对异常敏感的嵌入表示。之后，采用基于聚类的方法和自注意力自动编码器进行分类。实验结果表明，我们的方法能够有效区分良性执行与受感染执行。

关键词： 信息物理系统 · 入侵检测系统（IDS）· 3D打印机攻击 · 增材制造 · 异常检测。

1 引言

增材制造（AM）通常称为 3D 打印，它通过实现复杂部件的快速原型制作与生产，同时最大限度地减少材料浪费，彻底改变了现代制造业 [1, 2]。增材制造融入航空航天、医疗保健、汽车和国防等安全关键领域后，其底层基础设施已成为主要攻击目标

* 由美国国家科学基金会（NSF）资助

，面临信息物理威胁 [3,4]。随着联网3D打印机 [5],日益普及，攻击者可以利用贯穿数字线程的攻击向量，从 CAD设计、STL/G代码转换、网络通信直至固件执行，进而可能破坏打印部件 [6] 的机械完整性、尺寸精度或知识产权（IP地址） [7,8]。近期研究已证明，在增材制造流程的各个阶段发起攻击是可行的，并会造成严重后果。这些攻击包括篡改 CAD/STL文件 [9], 、恶意固件修改 [10,11], 、G代码级破坏 [12], ，以及通过功率、声学和磁辐射实施侧信道知识产权泄漏 [6,8,13,14]。与此同时，研究人员提出了采用过程监控 [15,16],、统计建模 [7], 和模拟辐射分析 [4]的攻击检测框架 [17] 。尽管已有这些努力，大多数既有方法仍依赖预先定义的攻击者模型或静态分析，或要求获取黄金参考模型（STL）。因此，在检测直接作用于 G代码（即机器可读指令）的实时隐蔽攻击方面，仍存在显著不足。G代码由切片软件根据 CAD模型生成，并在制造过程中执行。G代码上传后，人们对 G代码流程的信任假设便形成了一个严重漏洞。与 STL篡改不同，G代码级攻击能够悄然改变刀具路径、挤出量或打印时序，既不会触发目视检查，也不会违反基本几何约束。

在本研究中，我们探讨了能够绕过传统基于STL的验证、且不造成明显干扰便损害最终打印成品的隐蔽G代码篡改策略。我们提出了三种利用联网3D打印环境中现实威胁模型的策略：**延迟打印利用**、**访问阻塞型G代码替换**和**执行阶段篡改**。我们在现实威胁模型下实现了这些方法，包括欠挤出、过度挤出、噪声G代码注入、尺寸变化和内部空腔插入。我们假设攻击者能够访问打印机的文件系统或控制界面（例如遭到入侵的打印服务器、远程访问工具或内部人员威胁）。我们在两种广泛使用的FDM平台上开展了实验：Creality K1 Max和Creality Ender 3，它们分别采用不同的系统架构（Klipper和OctoPrint）。实验结果表明，这些威胁切实可行，且可推广至不同平台。

一项核心挑战在于，当没有真实标签STL或参考模型时，如何检测此类攻击；尤其是在偏差很细微的情况下（例如，轻微欠挤出或尺寸漂移）。为此，我们提出了一种基于日志的入侵检测系统（IDS），以无监督方式运行，且仅使用良性打印机日志进行训练。系统从打印机中提取遥测数据和执行日志。随后，我们使用冻结的“MiniLM Transformer”编码器进行自监督对比学习，生成具有判别性的潜在嵌入。这种方法无需标注攻击数据，同时保留语义结构。首先，使用主成分分析/UMAP 和 K-Means 聚类对嵌入进行可视化和分段，以确认良性模式与异常模式之间的区分。为了进行细粒度异常检测，我们进一步使用良性嵌入训练基于自注意力的自编码器，并在推

理过程中测量重构误差。重构损失较高的样本会被标记为异常。我们的主要贡献如下：

逼真的攻击场景： 我们基于一个自行开发的后门接口构造了多个逼真的对抗场景，该接口仿照合法的打印机控制面板。这些攻击包括欠挤出、过度挤出、噪声G代码注入、尺寸变化和内部空腔插入。

以日志为中心的入侵检测系统： 我们的方法利用打印机固件原生生成的结构化遥测日志，从而能够以非侵入方式实时进行异常检测。

表征学习： 我们采用预训练且冻结的 MiniLM 语言模型，将内容编码为具有语义意义的嵌入。我们仅使用良性日志训练对比投影头，使模型能够学习具有判别性的潜在空间，在无需标注攻击数据或标准 STL/G代码参考的情况下突出显示偏差。

混合评估： 我们通过无监督聚类和基于重构的检测来评估异常分离效果。我们使用良性嵌入训练一个基于自注意力的自编码器，以重构正常模式，并通过较高的重构损失标记异常。

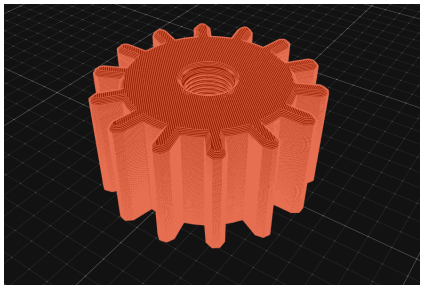
我们的方法通过将关注点从静态设计层面的验证转向动态运行时异常检测，补充了现有文献。

2 背景与预备知识

2.1 3D打印工作流

3D 打印是一种逐层制造工艺，可将数字模型转化为实体部件。在各种增材制造技术中，熔融沉积成型（FDM） [18] 因其价格实惠、易于使用且操作简单而最为普及。在 FDM 中，热塑性长丝 [19] 经熔化后通过加热喷嘴挤出，喷嘴沿 X、Y 和 Z 轴移动，将材料沉积到构建平台上。该过程逐层重复，直至整个物体成形 [20]。3D 打印工作流始于使用计算机辅助设计（CAD）软件创建数字 3D 模型。该模型以基于网格的格式导出，最常用的是 STL（立体光刻）文件格式 [21], ，该格式对物体的表面几何形状进行编1a码，如图所示。随后，STL文件由切片软件处理。切片软件会将 3D 模型切分为离散的水平层，并将其转换为机器可读指令。这些指令以 G代码 [22], 的形式编码，其中定义了打印机运动、挤出、温度控制和时序的精确命令。G代码文件通常会上传到打印机，并在制造过程中逐行执行，如图1b所示。

在我们的研究中，我们调查了两款广泛使用的 FDM 打印机：Creality K1 Max 和 Creality Ender 3。两者分别采用不同的系统架构和控制接口。



(a)

```
G1 X149.646 Y168.368 F30000
G1 Z.4
G1 Z.2
G1 E.6 F2400
;TYPE:Inner wall
;WIDTH:0.5
G1 F3000
G1 X131.279 Y150 E.9675
G1 X150 Y131.279 E.98612
```

(b)

图 1：(a) 立体光刻渲染，(b) G代码结构。

2.2 K1 Max 与 Ender 3：架构

K1 Max 运行基于 Klipper 的固件，并通过 Mainsail 和 Fluidd 等 Web界面连接到一台定制的、基于 Linux 的主机。Klipper 在底层将 G代码解析和运动规划任务卸载到主机，由主机通过串行接口与打印机的微控制器通信。这种架构显著提升了性能和灵活性，但也使主机与微控制器之间的通信通道暴露为潜在的攻击面。为了支持实时控制和 API访问，Klipper 配合 Moonraker 使用。Moonraker 是一项 Web服务，可提供 JSON-RPC接口以便与打印机状态交互。用户通过 Fluidd 或 Mainsail 等 Web前端上传 G代码文件；这些前端会将文件发送至 Moonraker API，将其存储在 Klipper 可访问的目录中，并通过 API调用启动打印。用户界面、API层与固件之间的这种分离，为能够访问本地网络或文件系统的攻击者提供了多个拦截点。

相比之下，我们的 Creality Ender 3 打印机采用基于树莓派的 OctoPrint 配置。在这种配置中，OctoPrint 软件运行于树莓派上，并通过 USB 直接与打印机连接。用户可访问由 OctoPrint 托管的 Web 仪表板，上传 G代码文件、监控温度和状态，并启动打印。OctoPrint 既充当控制器，也充当中继器，将命令直接转发给打印机固件。尽管 OctoPrint 的架构比 K1 Max 更简单，但它仍然引入了一个中心节点，G代码文件在执行前可能会在此被截获或篡改。

2.3 攻击面

在我们的威胁模型中，攻击者位于用户与固件之间，可通过入侵本地网络、获取代理接口（例如 Fluidd、Moonraker 或 OctoPrint）的访问权限，或直接操纵文件系统来实施攻击。我们实现了多种隐蔽的中间人攻击，这些攻击作用于 G代码层，如表 1所示。

表1：已实施攻击场景中使用的G代码类别

策略	**相关G代码**
延迟打印利用	M25、M0、G4、G28、G92
访问阻塞型G代码替换	G92、M28、M23、M24、M25
执行阶段篡改	G4、M25、M24、G92、G1、G0
攻击类型	**相关G代码**
欠挤出攻击	G0, G1, M83, M221, M404
过度挤出攻击	G0, G1, M82, M221, M404
噪声注入	G0、G1、G92、M104、M109
空腔插入	G0、G1、G92、M28
尺寸操纵	G92、M206、G1、G92.X、M579
知识产权盗窃	M154、M928、M111

3 相关工作

早期研究已展示了如何操纵 STL 等 3D 设计文件，从而引入结构弱点。Logan 等人 [9] 指出，对 STL 文件进行细微修改，既能逃过人工检查，也可能导致机械故障。同样，Sofia 等人 [3] 提出了一种端到端破坏攻击：通过操纵设计文件破坏无人机螺旋桨，使其在飞行途中失效。这些研究表明，未被发现的设计修改可能损害打印部件的完整性。作为一种机器可读的中间格式，G代码已成为攻击者关注的焦点。Caleb 等人 [7]研究了在无法获取黄金模型的情况下对 G代码进行细微恶意修改。他们的红队/蓝队设置量化了检测精细 G代码篡改的难度。“SOK” [11] 在此基础上，分析了 278 个潜在有害的 G-code，并提出了新的攻击者模型，这些模型即使只有有限的访问权限也能发起攻击。为防御此类威胁，研究人员提出对 G代码进行逆向工程，以验证打印几何形状。Tsoutsos 等人 [23] 开发了一种刀具路径重建方法，可根据 G代码重新生成近似的 3D 模型，并通过有限元分析（FEA）评估其结构可靠性。

研究表明，固件级入侵可提供持久且隐蔽的攻击向量。Mahmood 等人 [17] 证明，恶意固件能够拦截或替换打印任务，并篡改挤出设置。Muhammad 等人 [10]系统梳理了3D 打印多个阶段中的固件攻击，并提出“攻击可行性指数”（AFI），用于评估潜在威胁的等级。除固件攻击外，利用声学、电磁或功率发射的侧信道攻击也已得到探索。Song 等人 [8] 展示了如何通过分析声学和磁性数据，利用智能手机传感器重建G代码。“加密徒劳”一文 [14] 进一步证明，即使通信经过加密，功率侧信道泄漏仍可能实现近乎完美的模型重建。知识产权泄漏风险推动了保护增材制造流程的相关工作。Yampolskiy 等人 [24]提出了适用于涉及知识产权的制造的安全外包模型。相比之下，Chhetri 等人 [6] 指出，修改编译器可能会加剧侧信道泄漏，进而泄露G代码内容。在网络方面，McCormack

等人 [25, 26] 推出了 C3PO，这是一种用于评估联网的3D打印部署的安全态势和攻击路径的工具。实时过程监控是另一种防御手段。Gao 等人 [15] 开发了一种基于传感器的框架，用于监控速度、填充和风扇转速等物理参数，以检测异常打印行为。“KCAD”论文 [4] 介绍了一种动态网络攻击检测系统，利用模拟信号发射与G代码模式之间的统计映射。Mahmood 等人 [17] 和 Meleshko 等人 [27] 近期的研究侧重于在实验室设置中对网络攻击进行建模，并通过基于仿真的评估研究恶意篡改及其检测。

与以往聚焦于单一攻击向量的研究不同，例如篡改G代码、通过侧信道窃取信息或植入固件后门，我们的方法提出了一个全面的威胁模型，涵盖3D打印工作流中的多个入侵点。我们设计了一种切合实际的攻击者模型：攻击者利用恶意网页界面拦截用户活动，在打印前和打印过程中篡改上传的G代码文件，并劫持打印机控制。此外，我们提出了一种无监督异常检测系统，利用结构化的打印机主机日志以及结合对比学习的基于Transformer的日志嵌入，区分良性打印与遭到篡改的打印。

4 攻击策略

4.1 假设

我们的威胁模型假设，攻击者通过攻破身份验证凭据，获得了对3D打印机主机操作系统的root级访问权限。这一假设符合实际情况，并得到以往对嵌入式打印系统中身份验证实践不佳的观察以及我们的实验验证的支持。

在受控实验室环境中，我们通过暴力破解攻击和基于字典的密码攻击，验证了凭据泄露的可行性。这些实验在 Creality K1 Max 和 Ender 3 配置上进行，其中禁用了 SSH速率限制，并使用弱密码来模拟现实世界中的配置错误。暴力破解攻击通过并行枚举常见字符模式，成功破解了嵌入式系统中经常使用的短密码。基于字典的方法则进一步证明了其有效性：它利用精心整理的领域特定关键词列表，其中包括打印机型号（例如“Ender”“Creality”）、材料（例如“聚乳酸”“丙烯腈-丁二烯-苯乙烯”）和软件术语（例如“OctoPrint”“切片软件”“Klipper”）。这些词表模拟了攻击者基于对增材制造环境的了解而采用的现实策略。此类攻击的成功与更广泛的安全调查结果相符：许多 3D打印机部署时使用出厂默认凭据、由打印机名称衍生的密码，或其他薄弱的身份验证方案。此外，许多 3D打印控制器具有嵌入式、无人值守的特性，导致补丁更新不一致，安全卫生状况不佳。固件级身份验证机制要么缺失，要么

不足；而面向网络的控制界面（例如 Moonraker、OctoPrint 或 Fluidd）则扩大了攻击面。这些服务通常可通过局域网访问，有时也会为了方便远程打印而暴露在公共互联网上，从而进一步为远程攻击者提供可乘之机。

将凭据失陷确立为可行且可复现的步骤，使我们能够探索获得根权限后可能出现的后续攻击场景。

4.2 RogueInterface 与攻击

本研究提出的一种核心入侵技术，是创建一个伪造的控制接口，模仿用于控制运行打印机固件的 3D打印机的合法网页界面。这个恶意仿制界面托管在攻击者自己的机器上，或托管在与目标打印机位于同一网络中的已失陷设备上。通过凭据失陷获得根权限后，攻击者安装这个仿制界面，使其充当合法用户与底层打印机系统之间的透明代理。

它直接与打印服务器 API 通信，访问打印机的文件系统和配置目录。因此，攻击者能够实时查看所有用户操作和机器运行情况，包括 G代码上传、打印进度、系统温度、日志活动以及配置更改。通过将这个仿制界面部署为隐蔽的访问层，攻击者可以有效绕过传统的网络层监控工具。这使攻击者能够实施一系列后续 G代码篡改攻击，同时将被检测的风险降至最低，为下文所述的隐蔽破坏场景奠定基础。

威胁模型： 我们的威胁模型假设攻击者有一个进入增材制造流程的主要攻击入口，即对打印机文件系统和控制接口的系统级后门访问。图 2 展示了完整的入侵架构。

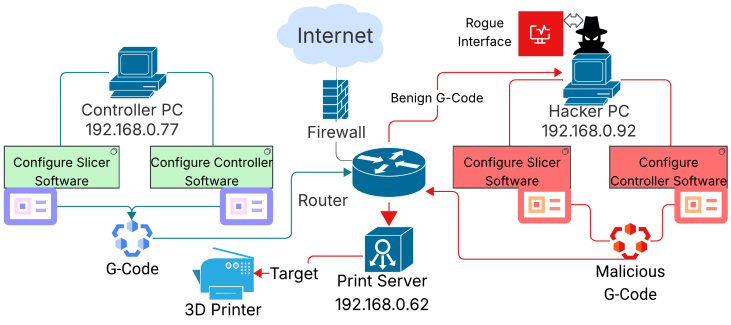


图 2：攻击策略流程

入侵点：通过 *G-code* 篡改实施中间人攻击。 攻击者通过 SSH 和伪造的网页界面，利用未经授权的访问权限，访问打印机的内部文件系统并控制 G代码 流程。我们实施了以下技术：

延迟打印利用： 攻击者监控打印机的 G代码上传目录，并等待合法用户通过标准网页界面（例如 Fluidd、Mainsail 或 OctoPrint）上传文件。一旦检测到新的 .gcode 文件，攻击者便等待短暂的延迟时间，以确保系统没有正在使用该文件。短暂等待后，攻击者下载已上传的文件，在本地注入隐蔽的扰动进行修改，然后将修改后的文件重新上传至原始路径。这种上传后的拦截方式很难被用户发现，因为大多数 3D 打印机用户界面既不会验证上传文件的完整性，也不会保留密码哈希 [28]。整个操作发生在文件上传之后、用户开始打印之前。这一策略针对的是从上传到执行之间的脆弱窗口。

访问阻塞型G代码替换：这种攻击改进了延迟替换策略：一旦检测到上传的 G 代码文件，便立即将其锁定。新 .gcode 文件一出现，攻击者就会暂时阻止用户访问该文件 5 到 10 秒。这可确保文件在修改过程中不会被过早执行。锁定后，攻击者下载该文件，注入针对性修改，再上传篡改后的版本。随后，文件将恢复为原始名称，从而完成替换。这种方法能够快速替换原始文件，并在操作过程中阻止用户访问。即使在用户触发打印速度很快的环境中，它也能提高破坏成功的可能性。

执行阶段篡改：攻击者的目标是正在进行的活动打印任务。通过打印服务器 API 查询打印机状态，攻击者可以确定当前正在执行的文件名以及打印状态。经过一段可配置的延迟（通常为 2 至 5 分钟）后，攻击者通过向后端控制接口直接发送命令，拦截正在进行的打印。随后，攻击者将当前正在打印的 G代码 文件替换为修改后的版本，并应用于剩余层。此方法使攻击者只需针对打印的后半部分进行篡改，从而让异常在早期检查时不易察觉，并可能绕过表面质量控制。此举也能避免引起用户怀疑，因为用户可能会将最终层的缺陷归因于切片问题或机械因素。

按入侵点划分的攻击：攻击者对 G代码控制的访问，使其能够实施范围广泛的恶意篡改，从隐蔽破坏到公然窃取知识产权。

知识产权（*IP*）盗窃。 攻击者访问3D打印机的文件系统后，可以窃取专有设计资产，例如上传的G代码文件或STL 衍生文件。这些文件可在任何修改之前或之后被截获，并秘密传输到攻击者控制的服务器。获取文件后，攻击者可以对其进行逆向工程，提取有价值的制造信息，例如路径规划逻辑、填充策略或经过精细调整的工艺参数。

此外，窃取的文件还可能被复制到远程3D打印机上，用于制造原始设计的仿制品。此类泄露在航空航天、生物医学设备或国防制造等高风险领域尤其具有破坏性，因为专有知识产权的泄露可能造成重大的战略和经济后果。

欠挤出： 通过降低E值，喷嘴输送的材料不足，导致层间结合薄弱、层间附着力差，最终降低零件的抗拉强度。这类缺陷可能存在于内部，无法通过打印后的目视检查发现。

过度挤出： 相反，增大E值会导致耗材沉积过量，造成尺寸不准确、表面粗糙，或打印头拖拽。这可能损坏周围层，或导致执行器过热。

噪声*G*代码注入： 另一种隐蔽的破坏手段是注入带噪声的运动指令，例如高频或随机的G0（非挤出移动）和G1（挤出移动）指令。这些指令可能通过引发高速振荡加剧步进电机的磨损，通过引入非均匀路径或锯齿状轨迹造成表面畸变。它们还可能导致微小延迟，使热控制失去同步，进而造成材料熔化不均匀。

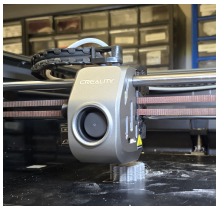
尺寸变化： 对坐标缩放进行细微操纵，例如将选定 G代码片段中的 X、Y 或 Z 数值相乘，可能会造成难以察觉却有害的几何畸变。沿 X/Y 轴缩放会影响占地尺寸和装配公差。沿 Z 轴缩放会影响层高，进而影响堆叠特性和垂直强度。局部变化（例如仅在打印件的中间三分之一处进行变化）可以避开仅测量包围盒的尺寸验证工具。此类攻击会损害对精度要求极高的部件（例如互锁部件或空气动力学敏感设计），同时保持整体外观相似。

内部空腔插入： 攻击者可能会在特定层范围内引入空洞或空腔，以破坏物体的结构核心。通过在预先设定的 Z高度区间内移除或跳过 E 挤出值来形成空腔。这些修改会在打印过程中途进行，以避开打印前预览或切片验证的检测。尽管外部几何形状保持完整，但这些区域缺少填充或轮廓线，会显著降低承载能力，导致物体在受力时悄然失效。这种攻击对安全关键零件尤为危险，因为内部缺陷不易被发现，却会直接影响可靠性。

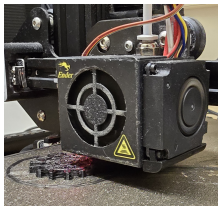
5 国防方法

5.1 数据收集

我们利用打印机日志作为高保真数据源，以捕获 3D 打印机的运行状态和运行时行为。日志是在两种主要条件下收集的：(1) 涉及良性切片和打印工作流程的良性运行，以及 (2) 恶意场景。



(a) 噪声注入对象：K1-Max。



(b) 空腔插入对象：Ender-3。

图 3：攻击期间 K1-Max 和 Ender-3 打印对象。

5.2 特征选择

为确保模型能够捕捉有意义的语义模式，同时尽量减少冗余，我们采用了结合领域驱动选择与统计筛选的混合方法。首先，我们根据特征与打印状态、热稳定性、通信健康状况和系统工作负载的相关性，从日志中确定了一组广泛的特征，如表 2 所示。为进一步优化选择，我们应用了两步统计筛选流程：

方差阈值筛选：良性数据集中方差接近零（阈值为 < 0.01）的特征被剔除，因为它们的判别能力有限。

相关性剪枝：从剩余特征中，我们计算了成对的皮尔逊相关矩阵，并剔除了具有较高线性相关性的特征（> 0.95），以减少冗余。

5.3 模型架构

我们提出了一种无监督异常检测框架。完整流程如图所示4。该架构由两个核心阶段组成：(i) 使用冻结的 Transformer 编码器进行对比表示学习；(ii) 使用自注意力增强型自编码器进行异常检测。

结构化日志表示为矩阵 $\mathcal{L} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\}$，其中每一行 $\mathbf{x}_i \in \mathbb{R}^d$ 表示一条具有 d 个数值特征的日志。经过低方差和高相关性筛选器处理后，得到的精简特征集表示为 $\mathcal{L}' = \{\mathbf{x}'_1, \ldots, \mathbf{x}'_N\}$，其中 $\mathbf{x}'_i \in \mathbb{R}^{d'}$、 $d' < d$。每个经过筛选的行 $\mathbf{x}'_i$ 都会被序列化为类似句子的结构。此转换保留了上下文，并支持基于词元的编码： $s_i =$ "$\text{key}_1 = v_1 \,|\, \text{key}_2 = v_2 \,|\, \ldots \,|\, \text{key}_d = v_{d'}$"

对比表示学习：我们采用 all-MiniLM-L6-v2 Transformer 模型作为冻结的编码器 $f_\theta : \mathbb{S} \to \mathbb{R}^{384}$，其中 θ 表示固定的预训练参数。对于给定的句子 s_i，冻结的编码器会生成一个 384 维的语义嵌入：$\mathbf{h}_i = f_\theta(s_i)$， $\mathbf{h}_i \in \mathbb{R}^{384}$

为使嵌入适用于异常检测，我们引入一个可训练的投影头 $g_\phi : \mathbb{R}^{384} \to \mathbb{R}^{128}$，其参数由 ϕ 定义。该投影头采用双层多层感知机实现： $g_\phi(\mathbf{h}) = \mathbf{W}_2 \cdot \text{ReLU}(\mathbf{W}_1 \cdot \mathbf{h} + \mathbf{b}_1) + \mathbf{b}_2$ ，

表2：日志中的特征

类别	特征（原始日志）	原因
打印进度	打印时间, SD卡位置, 缓冲时间，打印停滞, G代码输入	捕获打印任务已完成的进度并有助于检测检测异常进展或中断。中断。
热数据	微控制器温度, 腔室温度, 加热床温度, 挤出机温度, 加热床脉宽调制, 挤出机脉宽调制	反映热稳定性；有助于识别物理异常，例如欠挤出或过热。
系统资源：系统负载、CPU 时间、可用内存，表示 CPU 和内存压力。		
微控制器统计数据	写入字节数, 读取字节数, 字节重传, 无效字节数, 发送序列号, 接收序列号, 重传序列号	表示通信可靠性主机与微控制器之间；异常重传或丢包可能表明存在攻击。
微控制器时序	微控制器任务平均值, 微控制器任务标准差, 微控制器唤醒, rto 平滑往返时间, 往返时间方差,	时序延迟和任务调度方差可揭示异常处理行为或拥塞。
喷嘴/调平 MCU	喷嘴微控制器.*, 调平微控制器.*	比较时序和通信在不同的微控制器之间，可以检测局部问题或攻击痕迹。
树莓派 Pi 统计数据	rpi.微控制器任务平均值, rpi.字节重传, etc.	显示主机-控制器交互运行状况；此处的异常可能表明本地过载或网络中断。

权重矩阵为 $\mathbf{W}_1 \in \mathbb{R}^{256\times384}$、$\mathbf{W}_2 \in \mathbb{R}^{128\times256}$，偏置项为 $\mathbf{b}_1 \in \mathbb{R}^{256}$、$\mathbf{b}_2 \in \mathbb{R}^{128}$。完整模型的前向传播定义为：$\mathbf{z}_i = g_\phi(f_\theta(s_i))$，其中$\mathbf{z}_i \in \mathbb{R}^{128}$. 为以无监督方式训练模型，我们采用对比学习。对于每条良性日志行 s_i，我们通过领域感知扰动（例如，将“挤出机”替换为“tool- head”，或修改温度标签）生成一个合成扰动版本s_i'。目标是最小化良性嵌入与扰动嵌入之间的对比损失：$\mathcal{L}_{\text{contrastive}} = 1 - \cos(\mathbf{z}_i, \mathbf{z}_i') = 1 - \frac{\mathbf{z}_i \cdot \mathbf{z}_i'}{\|\mathbf{z}_i\| \cdot \|\mathbf{z}_i'\|}$. 训练期间仅更新投影头参数 g_ϕ，基础编码器 f_θ 则保持冻结。

推理与异常检测：在推理时，未见过的打印机日志会被序列化，并通过冻结的编码器 f_θ 和训练好的投影头 g_ϕ，生成紧凑的嵌入：$\mathbf{z}_i = g_\phi(f_\theta(s_i))$。这些嵌入作为两种互补异常检测策略的输入：

I. 基于聚类的检测：我们对良性嵌入应用 K-Means 聚类，以学习正常行为的质心表示。对于测试嵌入$\mathbf{z}_i$，异常分数计算为其到

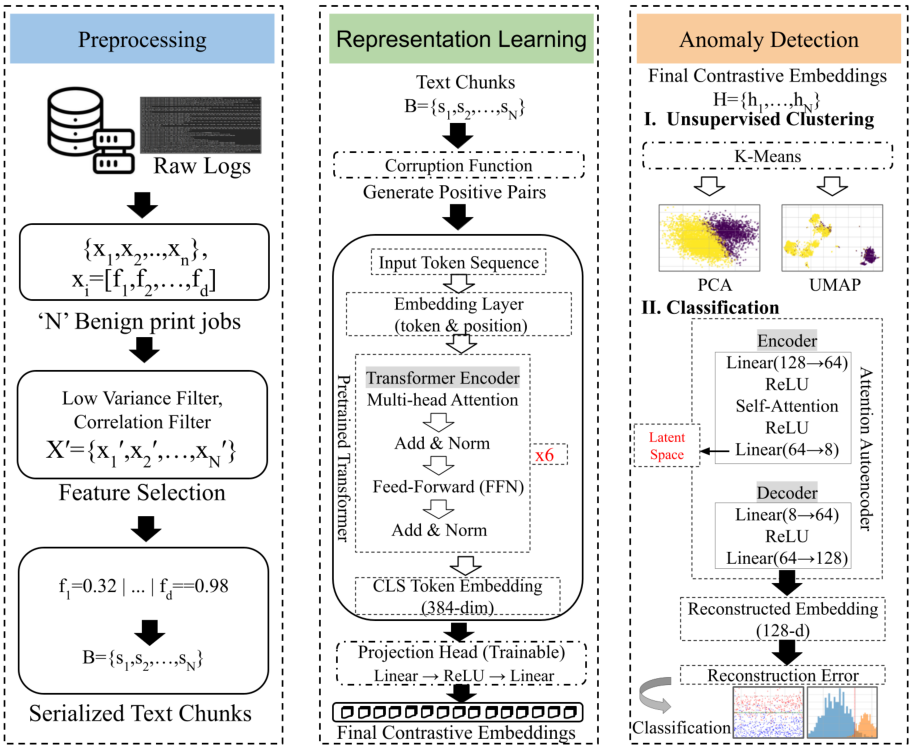


图4：模型架构：预处理、表征学习与异常检测

最近簇质心的最小欧氏距离： $\mathcal{A}_i^{(\text{cluster})} = \min_{c \in C} \|\mathbf{z}_i - c\|_2$. 簇距离较大的嵌入被视为异常。我们还应用主成分分析和 UMAP 等降维技术，对聚类边界进行可视化验证。

II. 基于重构的检测：为补充聚类结果，我们使用自注意力增强型自编码器重构由良性打印机行为生成的嵌入。较高的重构误差表明可能存在异常。$\mathbf{z}_i \in \mathbb{R}^{128}$ 表示序列化日志句子的对比嵌入。自编码器架构 $\mathcal{A}_\psi$ 由编码器 E 和解码器 D 组成，使得：$\hat{}\mathbf{z}_i = D(E(\mathbf{z}_i))$，对于 $i = 1, \ldots, N$.

编码器结构： 编码器将输入投影到维度为 8 的压缩潜在空间中。这通过以下方式实现：

$$\mathbf{h}_i^{(1)} = \mathrm{ReLU}(W_1 \cdot \mathbf{z}_i + \mathbf{b}_1), \quad W_1 \in \mathbb{R}^{64\times 128} \tag{1}$$

$$\mathbf{h}_i^{(attn)} = \mathrm{SelfAttention}(\mathbf{h}_i^{(1)}) \in \mathbb{R}^{64} \tag{2}$$

$$\mathbf{z}_i^{(enc)} = W_2 \cdot \mathrm{ReLU}(\mathbf{h}_i^{(attn)}) + \mathbf{b}_2, \quad W_2 \in \mathbb{R}^{8\times 64} \tag{3}$$

自注意力机制：我们实现了一个单头自注意力模块，用于捕获每个嵌入向量中特征之间的潜在交互。给定输入 $\mathbf{h} \in \mathbb{R}^{64}$，查询、键和值向量计算如下 [29]: $Q = W_Q\mathbf{h}$ $K = W_K\mathbf{h}$ $V = W_V\mathbf{h}$，其中 W_Q，W_K，$W_V \in \mathbb{R}^{64\times 64}$ 是可学习权重矩阵。注意力分数计算如下：$\alpha = \frac{Q\cdot K^T}{\sqrt{d_k}} \in \mathbb{R}$，$d_k = 64$，注意力输出 $=$ $\mathrm{softmax}(\alpha) \cdot V$。此操作使编码器能够在形成潜在表示时，动态强调信息量更丰富的特征分量。

解码器结构： 解码器根据潜在编码重构原始的 128 维嵌入：

$$\mathbf{h}_i^{(3)} = \mathrm{ReLU}(W_3 \cdot \mathbf{z}_i^{(enc)} + \mathbf{b}_3), \quad W_3 \in \mathbb{R}^{64\times 8} \tag{4}$$

$$\hat{\mathbf{z}}_i = W_4 \cdot \mathbf{h}_i^{(3)} + \mathbf{b}_4, \quad W_4 \in \mathbb{R}^{128\times 64} \tag{5}$$

异常评分：重构误差用于分配异常分数：$\frac{1}{128}\| \quad _i - \quad _i\|_2^2$ τ 95 $\mathcal{E}_i =$ ˆzz阈值根据良性重构误差的第 th 百分位确定。如果 $\mathcal{E}_i > \tau$，则将测试日志分类为异常。

6 实验分析

6.1 实验设置：

所使用的主要硬件平台是 Creality K1 Max 3D打印机，该打印机运行 Klipper固件，通过 Moonraker 进行接口连接，并通过 Fluidd 和 Mainsail Web前端访问。攻击自动化使用 Python 3.11 实现，采用基于SSH的文件访问、用于与 Moonraker API 交互的 requests，以及用于实时文件系统监控和 G代码注入的观察器。此外，我们还使用配备运行 OctoPrint 的树莓派4的 Ender 3，以验证我们的攻击在不同固件和控制栈之间的可移植性。两台打印机均连接到与攻击者主机相同的局域网（LAN），以便进行无缝监控和命令执行。所有实验均在一台高性能主机上完成，该主机配备了主频为 3.20 GHz 的 Intel Core i9-14900KF CPU 和 32 GB DDR5内存。我们使用了配备 24 GB GDDR6X显存、384 位内存总线和 512 个 Tensor核心的 NVIDIA RTX 4090 GPU，该 GPU 针对 AI 工作负载进行了优化。

6.2 数据集描述

表3展示了每个样本的规模。所有攻击类别均被特意排除在训练之外，以模拟现实中的零日威胁检测场景。类别比例经过选择，以确保多样性，并避免偏向任何特定攻击。

表 3：训练与验证的分布

类别类型	攻击类别	样本	用途
良性	–	98,720	训练
良性	–	57,373	验证
攻击	欠挤出	10120	验证
攻击	过度挤出	9950	验证
攻击	噪声注入	15456	验证
攻击	尺寸变化	9324	验证
攻击	空腔插入	10521	验证

6.3 训练细节

我们从日志中提取了共计 98,720 个良性样本，并将其整理为固定长度的时间窗口。冻结的 MiniLM Transformer 模型会为每个日志序列生成 384 维稠密嵌入。为适配这一维度，我们添加了一个可学习的多层感知机（MLP）投影头，将表示从 384 维映射到 128 维。

我们使用 Adam 优化器，学习率为 1×10^{-4}，批量大小为 64。模型仅使用良性数据进行 25 个训练轮次，不使用任何攻击标签。

7 定性评估

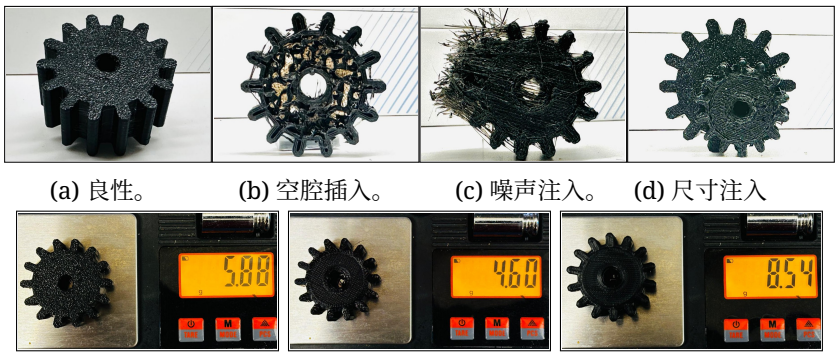
图5 和6分别展示了使用 Creality K1 Max 和 Ender-3 打印机打印的物理对象在各种攻击场景下的视觉影响。每幅图都并列对比了良性输出与遭篡改的输出，突出展示了几何变形和重量异常。在两种设置中，良性注入（图5a、6a）呈现出平滑的表面光洁度、均匀的齿轮齿以及结构完整的打印件，可作为完整性的视觉基线。在空腔插入攻击（b）下，两台打印机打印出的对象均出现内部材料缺失和填充不完整的情况，严重削弱了机械强度。噪声注入（c）会导致拉丝、不规则的外围路径和杂乱的表面纹理，揭示出由 G 代码扰动造成的挤出路径紊乱。尺寸操纵（d）在视觉上模仿了良性几何形状。



（a）良性。 （b）空腔插入。 （c）噪声注入。 （d）尺寸注入。

（e）正常挤出 (f) 欠挤出 (g) 过度挤出

图 5：Creality K1-Max 在不同攻击下打印的物体。

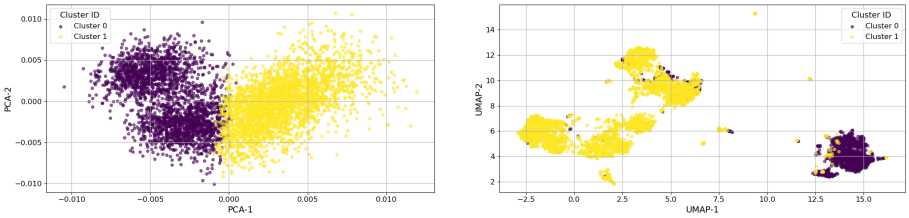


(a) 良性。 (b) 空腔插入。 (c) 噪声注入。 (d) 尺寸注入

(e) 正常挤出 (f) 欠挤出 (g) 过度挤出

图6：Ender-3 在不同攻击下打印的物体。

不过，它会微妙地改变整体尺寸（如果缩放为原始尺寸的 98%，肉眼无法察觉），可能会影响机械组件中的装配适配性和功能。下排重点展示了通过精确重量测量捕获的挤出级异常。对于 Creality K1 Max，良性挤出样本（e）的重量为 6.03g，而欠挤出（f）降至 4.32g，表明耗材沉积稀疏且零件强度较低。过度挤出（g）使质量增加至 9.00g，导致材料过多和表面结块。Ender-3 的结果呈现出类似趋势：良性挤出（e）为 5.80g，欠挤出（f）为 4.60g，过度挤出（g）为 8.54g，证实挤出攻击会在不同打印机架构中以一致的方式表现出来。



(a) 基于主成分分析的 K-Means 聚类

(b) 基于 UMAP 的 K-Means 聚类。

图 7：潜在空间中的聚类可视化

8 定量评估

8.1 聚类结果

图 7a展示了叠加在学习到的嵌入二维主成分分析投影上的 K-Means 聚类结果。以黄色和紫色标示的两个簇，沿主成分分析第一轴呈现出明显的线性分离。这种清晰的划分表明，对比学习框架已成功将良性行为和异常行为嵌入潜在空间中的不同区域。

为捕捉学习到的嵌入中的潜在非线性结构，我们使用 UMAP 将潜在空间降至二维。如图 7b所示，与主成分分析相比，K-Means 簇被投影到非线性流形后，分离效果更加明显。一个簇在右侧形成了紧密集中的群组，而另一个簇则分布在多个区域，这表明模型能够区分良性行为与多种攻击模式。

8.2 分类-重构误差：

为区分良性输入与恶意输入，我们采用了阈值机制。具体而言，我们使用良性验证数据上计算出的重构误差第95百分位数，将阈值设为 0.0019。这样既能确保仅标记偏差最异常的情况，又能最大限度地减少假阳性，这是在实际部署中的一项关键要求。

通过混淆矩阵、ROC 曲线和表格中的分类指标，对模型性能进行了定量评估。4自编码器的整体分类准确率高达 95.37%，精确率为 90.55%，召回率为 96.14%，F1 分数为 93.26%。值得注意的是，ROC 曲线下面积（曲线下面积）为 0.98，表明该模型对良性输入和异常输入具有很强的判别能力。

表 5 中的混淆矩阵进一步显示，94.99% 的良性样本被正确分类为良性，只有 5.01% 被误分类为攻击。对于攻击样本，96.14% 被正确识别为恶意样本，而 3.86% 被错误标记为良性。

表4：异常检测性能

指标	良性		攻击	
	数值	(%)	数值	(%)
精确率	0.9801	98.01%	0.9055	90.55%
召回率	0.9499	94.99%	0.9614	96.14%
F1 分数	0.9648	96.48%	0.9326	93.26%
总体准确率		0.9537 (95.37%)		
宏平均 F1 分数		0.9487 (94.87%)		
加权平均 F1 分数		0.9541 (95.41%)		
ROC 曲线下面积		0.9870 (98.70%)		
阈值（均方误差）		0.0019		

表5：自编码器输出

真实值 / 预测值	良性	攻击
良性	94.99%	5.01%
攻击	3.86%	96.14%

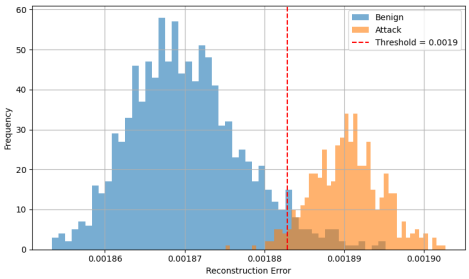
我们进一步可视化了良性样本和攻击样本的重构误差。图 8a 展示了各类别重构误差的分布。良性样本（蓝色）的误差分布集中紧密，均值较低，表明重构保真度较高。相比之下，攻击样本（橙色）的分布向右偏移，且离散程度更大。

图 8b进一步印证了这些发现，展示了测试样本序列中的重构误差。良性样本（蓝线）的重构误差始终低于设定的阈值。相反，攻击样本（红线）则频繁越过这一界限，表明其与学到的良性模式存在显著偏差。

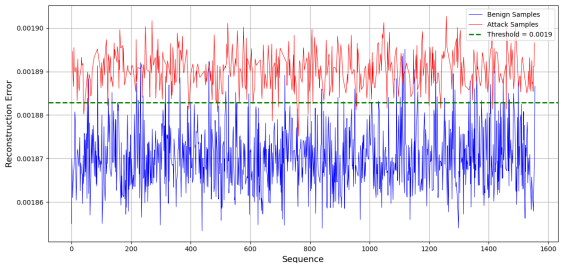
9 结论

本研究展示了多种攻击向量，包括劫持 G代码、空腔插入、欠挤出或过度挤出、尺寸操纵以及知识产权盗窃，这些攻击可通过后门访问或打印过程中篡改来发起。我们还开发并评估了一种无监督入侵检测框架，该框架采用对比表示学习和结构化日志分析。我们的方法学习良性打印机行为的语义嵌入，并采用基于聚类和基于重构的策略来检测偏差。我们在 Creality K1 Max 和 Ender 3 打印机上开展的实验验证了这些攻击的可行性以及国防模型的有效性。

尽管效果显著，当前系统仍存在一些局限。首先，该方法假设能够访问结构化且时间戳对齐的日志，但在专有或采用闭源固件的环境中，这类日志可能无法获取。其次，对于不会显著改变遥测模式的特定隐蔽攻击，



(a) 重建误差直方图



(b) 各序列的重构误差。

图 8：重构误差（均方误差）的可视化

（例如未引发挤出异常的隐蔽知识产权盗窃），系统可能无法检测到。未来的研究方向包括采用基于图的时序嵌入进行时序建模，以捕捉长程依赖。最后，将模型推广至多种打印机架构，并扩展为面向分布式打印机群的联邦异常检测，将进一步扩大其适用范围。

10 致谢

本材料基于美国国家科学基金会在奖项编号 2417062 项下支持的工作。本文所表达的任何观点、发现、结论或建议均属于作者本人，并不一定反映美国国家科学基金会的观点。

参考文献

1. Terry Wohlers、Tim Gornet、Noah Mostow、Ian Campbell、Olaf Diegel、Joseph Kowen、Ray Huff、Brent Stucker、Alex Doukas、Brian Drab、Drstvensek Igor、Nicholas Eitsert、David Espalin、Thomas Feldhausen、KhalidAbdel Ghany、Morgan Gillett-Crooks、Dan Guo、Arno Held 和 Joris Peels。增材制造史。 *SSRN Electronic Journal*，2017年1月。2. Osama Abdulhameed、Abdulrahman Al-Ahmari、Wadea Ameen 和 Syed Ham- mad Mian。增材制造：挑战、趋势与应用。 *Advances in Mechanical Engineering* 11 2 1687814018822880, 2019 ，():.

3. Sofia Belikovetsky、Mark Yampolskiy、Jinghui Toh、Jacob Gatlin 和 Yuval Elovici。dr

0 *11th USENIX* wned—— *Workshop on Offensive Technologies (WOOT 17)* 2017 增材制造中的网络物理攻击。发表于，温哥华（BC），8月。USENIX协会。4. Sujit Rokka Chhetri、Arquimedes Canedo 和 Mohammad Abdullah Al Faruque。Kcad：面向网络物理增材制造系统的动力学网络攻击检测方法。发表于 *2016 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)* 1 8, 2016 ，页码–。

5. Jens M¨uller、Vladislav Mladenov、Juraj Somorovsky 和 J¨org Schwenk。SOK：利用网络打印机。发表于 *2017 IEEE Symposium on Security and Privacy (SP)*，页码213–230，2017年。6. Sujit Rokka Chhetri、Anomadarshi Barua、Sina Faezi、Francesco Regazzoni、Ar- quimedes Canedo 和 Mohammad Abdullah Al Faruque。间谍工具：通过修改 3 *IEEE Transactions on Dependable and Secure Computing* 18 2 667 678, 2021 d打印机编译器泄露你的IP地址。，（）：–。7. Caleb Beckwith、Harsh Sankar Naicker、Svara Mehta、Viba R. Udupa、Nghia Tri Nim、Varun Gadre、Hammond Pearce、Gary Mac 和 Nikhil Gupta。大海捞针：检测增材制造G代码文件中的细微恶意修改。*CoRR*, abs/2111.12746，2021年。8. Chen Song、Feng Lin、Zhongjie Ba、Kui Ren、Chi Z hou 和 Wenyao Xu。我的智能手机知道你打印了什么：探索针对3D打印机的智能手机侧信道攻击。发表于 *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*，CCS '16，第895–907页，纽约（NY），美国，2016年。计算机协会。9. Logan D. Sturm、Christopher B. Williams、Jamie A. Ca melio、Jules White 和 Robert Parker。增材制造系统中的网络物理漏洞：一项针对.stl文件且涉及人类受试者的案例攻击。 *Journal of Manufacturing Systems* 44 154 164, 2017 ，：–。

10. Muhammad Haris Rais、Muhammad Ahsan 和 Irfan Ahmed。SOK：针对熔融丝材制造的3D打印机固件攻击。发表于 *18th USENIX WOOT Conference on Offensive Technologies (WOOT 24)*，第263–282页，费城（PA），2024年8月。USENIX协会。11. Samuel Moore、William Glisson 和 Mark Yampolskiy。恶意3D打印机固件的影响。2017年1月。12. Jost Rossel、Vladislav Mladenov、Nico W¨ordenweber 和 Juraj Somorovsky。3D 打印中恶意 G 代码的安全影响。13. Mohammad Abdullah Al Faruque、Sujit Rokka Chhetri、Arquimedes Canedo 和 Jiang Wan。针对增材制造系统的声学侧信道攻击。发表于 *2016 ACM/IEEE 7th International Conference on Cyber-Physical Systems (ICCPS)* 1 10, 2016 ，页码–。

14. Jacob Gatlin、Sofia Belikovetsky、Yuval Elovici、Anthony Skjellum、Joshua Lubell、Paul Witherell 和 Mark Yampolskiy。加密无济于事：利用功耗侧信道重建3D打印模型。发表于 *Proceedings of the 24th Inter-national Symposium on Research in Attacks, Intrusions and Defenses*，RAID '21，第135–147页，纽约（NY），美国，2021年。计算机协会。15. Yang Gao、Borui Li、Wei Wang、Wenyao Xu、Chi Zhou 和 Zhanpeng Jin。监测并保护你的3D打印机：针对网络物理攻击的在线过程监控。

Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.，2(3)，2018年9月。

16. Muhammad Ahsan、Muhammad Haris Rais 和 Irfan Ahmed。SOK：用于增材制造的侧信道监控——连接网络安全与质量保证领域。在 *2025 IEEE 8th European Symposium on Security and Privacy (EuroS&P)* 1160 1178, 2023 ，第 – 页。

17. Muhammad Arif Mahmood、Asad Waqar Malik 和 Frank Liou。用于识别增材制造中网络物理攻击的新型框架。*Progress in Additive Manufacturing*，10(2):1575–1586，2025。18. I. John Solomon、P. Sevvel 和 J. Gunasekaran。FDM 中各种工艺参数综述。*Materials Today: Proceedings*，37:509–514，2021。机械工程材料科学新趋势与创新国际会议。19. Vishal Mishra、Sushant Negi 和 Simanchal Kar。基于 FDM 的再生热塑性塑料及相关复合材料增材制造。*Journal of Material Cycles and Waste Management* 25 2 758 784, 2023 ，( ): –。

20. Kaufui V. Wong 和 Aldo Hernandez。增材制造综述。*International Scholarly Research Notices* 2012 1 208760, 2012 ，( ): 。

21. Ratnadeep Paul 和 Sam Anand。用于增材制造过程的基于 Steiner 补片的新型文件格式。*Computer-Aided Design*，63:86–100，2015。22. K. Rajaguru、T. Karthikeyan 和 V. Vijayan。增材制造——技术现状。*Materials Today: Proceedings*，21:628–633，2020。能源、环境与工程应用纳米材料最新趋势国际会议。23. Nektarios Georgios Tsoutsos、Homer Gamil 和 Michail Maniatakos。安全3D打印：利用刀具路径逆向工程重建并验证实体几何形状。在 *Proceedings of the 3rd ACM Workshop on Cyber-Physical System Security* 17, 15 20, 2017 ，CPSS ’

页面 – 纽约，NY，美国，。计算机协会。24. Mark Yampolskiy 、 Todd R. Andel 、 J. Todd McDonald、William B. Glisson 和 Alec Yasinsac。增材层制造中的知识产权保护：安全外包的要求。在 *Proceedings of the 4th Program Protection and Reverse Engineering Workshop* 4, 2014 ，

PPREW- 纽约，NY，美国，。计算机协会。25. Matthew McCormack、Sanjay Chandrasekaran、Tianlong Yu、Sandra Wolf 和 Vyas Sekar。C3po：联网3D打印机的安全分析工具（cmu-cylab-19-002）。26. Matthew McCormack、Sanjay Chandrasekaran、Guyue Liu、Tianlong Yu、Sandra DeVincent Wolf 和 Vyas Sekar。联网3D打印机的安全分析。在 *2020 IEEE Security and Privacy Workshops (SPW)*，第118–125页，2020。27. Aleksei Meleshko、Vasily Desnitsky 和 Igor Kotenko。用于攻击分析与检测的智能增材制造过程建模。在 *2025 33rd Euromicro International Conference on Parallel Distributed and Network-Based Processing ,,,* 第 – *(PDP)* 584 591, 2025
页。28. Md. Nazmul Islam、Yazhou Tu、Md. Imran Hossen、Shengmin Guo 和 Xiali Hei。关于增材制造的局限性、安全与隐私问题的调查。*CoRR*，abs/2103.06400，2021。29. Ashish Vaswani、Noam Shazeer、Niki Parmar、Jakob Uszkoreit、Llion Jones、Aidan N Gomez、 Ł ukasz Kaiser 和 Illia Polosukhin。注意力机制即所需的一切。载于 I. Guyon、U. Von Luxburg、S. Bengio、H. Wallach、R. Fergus、S. Vishwanathan 和 R. Garnett 编， *Advances in Neural Information Processing Systems*，第30卷。Curran Associates, Inc.，2017。