

AI疯狂网络攻击 防守如何降“魔法”

借助AI模型发现全球网站存在严重安全漏洞 黑客利用AI技术发起数百万次攻击试探

网络安全公司监测到，全球相关网络攻击尝试已达到数百万次。引发这场风波的，是全球大量网站共同使用的一项基础技术——PHP脚本语言，日前被发现存在严重安全漏洞。

各企业正赶在黑客利用AI（人工智能）技术发起攻击之前，争分夺秒地升级防护。网络安全专家估计，针对PHP漏洞的防护将于未来几周密集展开，但要彻底清除遗患可能需要数月甚至数年。

一个看似不起眼的软件漏洞，为何能牵动全球互联网？人工智能的介入，对网络安全又意味着什么？

回放

网络安全公司借助AI模型发现PHP漏洞

据路透社报道，美国卡利夫网络安全公司和隶属于“字母表”公司的网络安全企业维兹公司先前分别发布报告说，借助AI模型发现了PHP漏洞。此次安全问题涉及两个漏洞，攻击者可能通过触发内存耗尽状态，造成内存损坏，并在特定条件下利用漏洞远程执行代码。10月22日，PHP开发团队发布紧急安全更新，漏洞已被修复。

“这既是坏消息也是好消息。”英国皇家三军研究所网络技术部门负责人詹姆斯·沙利文说，黑客能借助AI发现新的漏洞并发起攻击，但另一方面AI也可用于发现并修补这些漏洞。

AI网络攻击有多“疯狂”？

如果我们把互联网想象成一座赛博城市，那么城里有将近70%的建筑地基，都是用一种叫“PHP”的底层代码砖头堆起来的。

过去的十几年，这一地基始终是稳固的。但前段时间，研究人员借助AI，在底层的砖头上发现了致命的小裂缝。而顺着这些小缝隙，攻击者可以绕过安保直通楼内。

这个消息一公开，整个互联网都倒吸了一口凉气。为什么？

因为攻击者速度太快了。安全网络监测显示，在PHP存在安全漏洞的消息传出后，短短45分钟内，网络上涌现了数万次针对这个漏洞的扫描和试探，两小时内增至数十万次，数小时后增至数百万次。

在金融圈，情况更像是一部科幻抢劫片。

在拉美，黑客组织通过精心设计的“套话”和伪装，骗过大模型的安全限制，让AI参与编排针对几百家金融机构的攻击试探。

在华尔街，银行高管发现很多用户把自己的账户授权给了第三方的“AI理财助手”。这些AI代理开始在系统里频繁越权操作。银行的防火墙分不清这是客户本人在操作，还是失控的AI在操作。

如果你觉得赛博空间的危机离自己还很远，那物理世界的失控可能就在你身边。

电影《豚鼠特工队》中，全世界的微波炉、咖啡机被植入了特制芯片，家用电器瞬间“黑化”。最近，这一幕在现实中上演。

黑客抓住老旧的物联网设备中的漏洞，入侵了法国几百家超市的冷链系统，远程关停了制冷压缩机，并篡改了温度报警器。这直接导致堆积如山的生鲜食品在室温下腐烂报废。

AI防守如何接招？

防守方的绝招是：用AI的魔法打败AI的魔法。让大模型化身“超级架构师”深入到最底层的汇编代码里，发起一场“数学级”的防御战。

过去寻找代码中bug，用的方法是“模糊测试”。这就像扛着一堆沙袋往墙上猛砸，墙没倒就以为是安全的。

现在用的方法是“形式化验证”。它没有去砸墙，而是写了1.8亿行严密的数学证明代码。这种可被机器检查的数学证明，让那些边边角角的隐秘漏洞无处遁形。

复盘这几起事件，我们会发现，攻防的速度已经从人类的以“天级”变成了机器的“毫秒级”。

普通人如何应对？

复旦大学计算与智能创新学院执行院长杨珉表示，个人层面，网络安全意识必须同步跟上技术发展的节奏。面对涉及个人隐私、资金财产等高度敏感网络操作，则需更加审慎。不随意给智能体开放自主交易权限，主动设置支付额度上限，从源头压缩风险空间；同时持续提升辨别深度伪造、AI生成欺诈信息的能力，守住个人安全第一道关口。

政府层面，要加快完善适配AI时代的网络安全监管制度体系。

企业层面，作为数字基础设施、AI产品与服务的运营方，企业更要压实主体责任，形成监管、管理、技术协同的立体化防护格局。

技术的狂飙从来都是双刃剑。AI可以帮助攻击者更快寻找裂缝，也可以帮助防守者加固墙体。真正决定它成为利刃还是坚盾的，不只是模型能力，更是人类如何授权、验证、审计，以及如何为每一次行动划定边界。

■综合新华社、央视新闻

连线

二维码别随便晒！打码的兑换券被AI修复

不知您有没有在网络上晒过快递单和兑换券二维码？近日，有网友在微信群里看到别人发的一张兑换券截图，发图人已经给这个二维码打了马赛克，本意是盖住二维码、不让别人扫码。

可网友随手把这张带马赛克的二维码截图丢给AI图片工具，试着修复，经AI复原后的二维码，却直接可以扫码，能读出里面内容。网友将相关内容发在了网上，帖子很快在网络上被热议，引发了大家对AI修复二维码能力的普遍担忧。

AI能复原被破坏的二维码吗？

那么，被遮挡了的二维码到底能不能复原，记者亲自测试了一下。

记者手上一张打印图是一个网站的完整二维码，扫它就能进入网站，另一张图将二维码放大，去掉了两边的部分，然而扫这张不完整的二维码一样可以进入网站。

记者将另一张二维码进行重度涂鸦，完全看不到涂鸦下二维码的内容。然后将这张图上传给AI让它复原。AI复原后再扫描它，居然还可以进入网站。

记者再将二维码涂鸦的范围扩大，但是二维码的部分纹路隐约可见，把它上传给AI，AI恢复了这张二维码，这张能不能扫描进入网站？结果是不能进入，AI未能完整复原可使用的二维码。

打马赛克尽可能覆盖90%区域

二维码由定位标记、功能模块、核心数据模块三部分组成。专家表示，如果承载内容的关键数据区被完全遮挡，原始编码彻底丢失，AI无法复原；如果只遮挡了四周冗余纠错区域，AI可以依靠未被遮挡的部分补齐图案，生成可正常扫码的二维码。简单来说，AI所谓的恢复，前提是遮挡区域没有盖住核心数据；一旦核心数据被彻底覆盖，AI是无法修复的。

中国计算机学会计算机安全专委会委员王媛媛指出，理论上AI是可以还原这些数据的。首先，二维码主要数据集中在中间位置，如果马赛克打在四周，AI可以通过还原图像的方式把二维码还原出来。

王媛媛表示，第二种可能性是，如果马赛克打得很浅或是覆盖区域不够，AI通过扩图能力，也可以把这些数据还原出来。

专家提醒，收款码、快递单等二维码都属于个人隐私，尽量不要直接发到网上。如果必须发布图片，要做好彻底遮挡。

王媛媛说，简单的防范方式是，如果我们需要对这些二维码进行有效处理时，我们尽可能把整张二维码90%的区域都打上马赛克，这样就可以避免AI通过还原能力把二维码的信息进行还原。

■据央视新闻

发行方式：邮政发行 发行单位：湖南省报刊发行局 发行热线：11185 邮发代号：41-118 定价：360元/年 印刷：湖南日报社印刷厂（地址：长沙市望城区黄金西路989号）