新闻热线：010—58884061
E-mail:kanaxi@126.com
■责编 陈 莹
科技日报
2017年10月9日 星期一

AI已经学会猜密码了 你还敢用“6个0”吗

本报记者 **张佳星**

近日，有消息说美国史蒂文斯理工学院开发出的人工智能可猜对25%的密码。

AI已经学会猜密码了？如果真是这样，你还敢用“6个0”“6个8”这样简单的密码吗？网络还有安全可言吗？

一些事之所以耸人听闻，是因为不明所以。为此，科技日报记者带着消息请教了中国科学院软件研究所可信计算与信息保障实验室主任张振峰，请他详细讲讲。

破解的是口令不是密码

讲正题前，插个花絮。单口相声大王刘宝瑞存世的经典作品里有个《黄半仙》，这位“半仙”捻动须髯，算得出村里张妈的戒指丢在哪，算得出艳阳天的后晌准下雨……却不是因为他真“仙”，而是他汇总、分析小细节，就能得出正确结论。这和AI的深度学习如出一辙，也就是说，美国这位“艾（AI）半仙”有25%的算中率，仰仗的是大量的基础数据，分析学习后，才能进行猜测，而不是因为它“天赋异禀”。

那么，问题来了，用于AI深度学习的基础数据库从哪儿来？这个技术会不会造成网络安全危机呢？

“先纠正一个概念，”张振峰说，“这则消息里说的密码，更准确地说是口令（password），而不是学术上密码学（cryptology）中研究的密码。”

它们最直观的区别是“字符串长度”，口令一般十几个字符，字符组成的所有可能可以被“穷举”，而一代RSA密码算法就有1024位，“穷举”在计算上是不可行的。打个不太恰当的比喻，攻破口令要踹开的是一扇门，而攻破密码是要征服一座迷宫。

“现在还没有研究表明AI能破解密码算法。”张振峰说，“密码”被业界认为是互联网的基础设施，一个国际标准的商用密码是非常复杂的，里面包含复杂的密码算法，凝结了研究人员多年的智慧，很难通过学习基础数据倒推其中规律，进而破解。张振峰所从事的正是密码学领域的研究。

AI领军人物、深圳先进技术研究院副院长汤晓鸥表示，AI很长时间以内是无法超越人类智慧的，也就是说它无法像人类一样进行创造性的脑力劳动。尽管这样的研究也在进行中，例如“谷歌大脑”项目，正是要开发出一款模拟人脑的软件。

猜对25%是怎样的成绩

这次不少报道标题都提到了“25%”的准确率，例如“准确率居然达到25%”“准确率逆天的25%”，潜台词是“25%”是个高不可攀的准确率，那事实是不是这样呢？

张振峰说：“AI独立猜测成功的比例不到12%，破解软件HashCat可以做到将近23%，这个25%是AI和HashCat两种方法相结合得到的数字。”

此外，单从准确率是“25%”判断它是否有效果是片面的。科学研究的做法是，以准确率为纵轴、实验攻击次数为横轴，得出不同攻击次数所对应准确率的曲线，“从曲线走向来看，准确率攀升幅度越大，那么口令猜测的成绩越好。”

可见，25%既不说明AI尝试4次就能猜对一次口令，也不说明它破解了1/4人群的口令，这个数字是AI创造出的新口令与它未知的另一部分旧口令比对之后，匹配的概率。

“25%说明AI在一定程度上提高了传统工具猜测密码的能力，对于口令强度测试具有积极作用。”张振峰说。

360网络攻防实验室负责人林伟告也持类似观点，他认为该研究可以加速破解口令的人工编程进程，或者用来测试口令的强度。

可供AI学习的数据猛增引担忧

那么AI破解口令，是怎么做到的呢？

原消息中提到，团队让一个人工智能程序利用数千万个泄露的密码来学习如何生成新密码。

数据显示，2016年，全球已知的用户数据泄露有40亿之多。2017年，这个数据可能更多。“猛增有可能是以前的存量，因为很多服务器的数据泄露，自己并不知情，”张振峰说，“或者即便知道，自己也不愿意主动公布。”

“也有可能是攻击手段越来越多导致的。”张振峰说。名噪一时的“永恒之蓝”背后，是网络攻击武器库的泄露。有消息称，这些攻击武器是在明码标价销售的。黑客获得更多的尖端攻击手段，可能是数据泄露猛增的另一个原因。

“暗网上有很多卖数据的，但并不一定是刚刚被泄露的，有可能有时间差，”张振峰说，“比如，有可能是5年前拿到的数据，因为偷盗的数据也不会是我们认为的‘明码’，而是服务器变换处理过的，还需要解密。”

无论何种原因，值得担忧的是，“艾半仙”能用来学习的基础数据越来越多。

“半仙”是如何修炼的

破解口令，目前大致有暴力攻击、启发式攻击、概率猜测等方式。

张振峰一一解释：暴力攻击是最原始的方法，把所有的可能都试一遍，计算机的计算能力越强大，破解越快；启发式攻击，也叫字典攻击，是根据泄露的口令进行分析，把规律“编写”成“字典”，并结合矫正规则进行猜测，用于攻击的“字典”不同，攻击的方式就不同，同等硬件条件下，字典越好，越快破解；概率猜测基于人们设置密码时，有着和自然语言类似的分布特征，通过数据集计算其概率分布，有些字符组合用的频率高，猜测就准。诸如国内网民最常用的25组密码，密码管理公司Keeper Security公布的2016年最常用的25个密码等就是这一类猜测依据的“冰山一角”。

AI破解口令是深度学习的一种应用，“它属于一种启发式方法，基于数据集来猜测口令，”张振峰说，“看起来还没有得到实战验证，只要用户在数据泄露之后及时修改自己的口令。”

那么，AI是如何进行口令猜测的呢？

有一个形象的比喻能说明这个过程。AI神经网络由大量“感知机”相互连接构成。感知机类似于生物神经系统中的神经元。它并非生来就具备强大的功能，而且需要训练才能掌握技能。例如希望神经网络通过西瓜的外形判断瓜的甜度，一开始AI并不懂如何去判断，这时就需要分别把西瓜的外形和对应的甜度输入神经网络，训练它学习两者的对应关系。训练过程实际上是通过学习数据来调整每一个感知机参数的过程。

神经网络读取数据样本后，感知机们会先根据现有模型参数进行计算，然后把输出的值与真实值进行比较，再将两者的差距反馈回去，以调整参数。经过反复多次“计算—比对—反馈—调整”的循环后，AI就能判断个八九不离十了。

但是，实际中，很多时候训练数据的真实结果信息难以获得——比如不能把每个瓜切开尝尝。这就用到了消息中美国史蒂文斯理工学院团队利用的“生成对抗网络”，巧妙避开“无法实时核实密码”这个问题。简单地说，研究团队设计出两个对抗的系统互相修炼，把获得数据一分为二，一部分用于生产，另一部分用于核验。通过训练，一个系统就像做赝品的画院学生练成了画家，另一个系统用“核验”数据充当“鉴定师”。

“但这些的基础都是源自已有的数据，这些数据是离线的，该消息中所用数据来自于2010年泄露的数据集合，其口令是明码存储在服务器上，而且长度不超过10个字符。”张振峰说。

无人驾驶技术革新：不联网也能实时处理

第二看台

本报记者 **李 伟**

是的，你没看错。10月份，无人驾驶公交车将正式登陆深圳。脑补一下现场画面：驾驶座上空无一人，方向盘却在自己转动，车辆按照既定的路线前行。人们会再一次感受到逆天的科技创新带来的惊悚观感。这要不是白天，你敢坐吗？

在5G网络覆盖、万物互联的时代，手机将失去现在独霸市场的优势地位。业内人士认为，汽车，作为万物互联的一个关键维度，将像今天的智能手机一样，成为业界未来争夺的重要阵地。“车辆的无人驾驶和高度人工智能化正成为新的趋势，目前整个行业正加速进入技术爆发期，其中先进驾驶辅助技术将首先形成千亿级市场。”腾讯副总裁陈菊红说。

要实现“无人驾驶”，需要利用车载传感器来感知车辆周围环境，并根据感知所获得的道路、车辆位置和障碍物信息，控制车辆的转向和速度，从而使车辆能够安全、可靠地在道路上行驶。而“先进驾驶辅助系统”（ADAS）是完成这一系列指挥方案的“大脑”中最为重要的技术之一。

10月份深圳市将推出两条无人驾驶公交线

“先进驾驶辅助系统（ADAS）的核心技术就是‘物体检测’，即检测行驶车辆周边环境。”前百度研究院副院长、深度学习实验室主任，如今的高科技公司地平线创始人兼CEO余凯在接受科技日报记者采访时透露，相对于物体分类，物体检测要困难得多。物体检测除了需要判断物体是什么，还需要给出物体的精确位置，“比如对于视觉检测来说，需要在10^6的搜索空间中精确地找出物体，考验的不仅仅是物体识别的准确度，还有计算的时间复杂度。地平线基于深度学习的高性能嵌入式视觉检测技术，就是为了满足这样的使用场景而研发的，是相对更高效的检测方法。”

在日前落幕的第三届军民融合装备展现场，记者看到了地平线展示的样机。本届装备展上，来自全国354家企业的422项技术成果参与了展出。其中，自主可控、人工智能、先进感知等信息技术领域的一大批“高、精、尖”的军民融合高技术装备占据了整整一层展区。余凯向记者介绍了地平线“嵌入式人工智能”在自动驾驶领域的一系列规划。

何为“嵌入式人工智能”？这实际上是地平线在成立之初就看准的发展方向。在人工智能领域深耕多年，余凯有自己的看法。“嵌入式人工智能是相对于当前联网下的人工智能而言，目前业界主要通过联网和数据中心的大规模计算来实现人工智能。”余凯说，“而‘嵌入式人工智能’要做的，是在本地进行实时环境感知、人机交互与决策控制。”

举个例子，当有小孩横穿马路时，如果依靠自动驾驶系统感知，然后把信号传送到云端再做决策，会相比本地运算产生更长的时间延迟。“假如遇到当时网络不稳定，结果不可想象。”余凯说。

对于10月份即将在深圳亮相的无人驾驶公交车，余凯认为，过去十年深度神经网络发展，让人工智能“感知”能力突飞猛进。但是所有的人工智能技术，包括自动驾驶，其实都是为了最终的“决策”服务。

从地平线展示的样机上可以看到，目前的研发水平能够满足“在高速公路和市区道路场景下，同时对行人、车辆、车道线及可行驶区域的实时检测和识别”。也就是说，地平线的ADAS系统已经能对交通环境中包含行人、车辆、隔离带等在内的多种可能影响行车安全的障碍物进行检测和识别，并筛选出可行驶的无障碍安全区域，“相较于在封闭道路行驶，其难度更大、重要性更高。”余凯说。

在余凯看来，自动驾驶乃至无人驾驶是人类出行方式的又一次颠覆式革新。要想在这一领域取得突破性进展，必须与产业上下游携手合作。“如果对自动驾驶的场景进行划分，可能会划分出成千上万种场景，并且每个场景都要做充分的测试，但如果上下游企业共同参与的话，整个产业会发展得更快。”另一方面，余凯认为，由于行业存在不确定性，通用的技术标准尚未被制定出来，资源又很有限，因此需要上下游企业联合，“抱团”推进技术研发和落地应用，在发挥各自优势的同时提高效率、降低风险。

新鲜事

迪拜将推“空中的士” 无人驾驶原型机试飞成功

据新浪科技报道，近日，阿联酋迪拜市进行了一次飞行测试，官方宣称其将会是世界上第一个提供无人机的士服务的城市。在飞行测试中，这架飞行器悬停在距离地面200米的高度，并沿着海边的沙滩飞越了迪拜的墨西哥湾，整个过程历时大约5分钟。

据了解，试飞原型机是由德国无人机公司Volocopter开发的一架小型两座直升机，机顶采用宽18英尺的螺旋桨，无需驾驶员即可搭载两位乘客，最高时速可达100km/h，目前最新的2X机型在75km/h的时速下，最大续航里程为27公里，且能够垂直起降。用户可以使用智能手机的应用程序召唤无人机到距离最近的站点，乘坐其到达目的地。另外，该机型的充电时间仅有2小时，快速充电时间则少于40分钟，还配有安全通信网络和紧急降落伞。

据悉，这款无人机的士从2010年起即投入研发，按照计划，Volocopter公司会在2020年将这种飞行器带到更多城市，5年之内，将“空中的士”服务项目落到实处，以缓解城市交通拥堵问题。

京东推出无人驾驶货车 已在指定路段测试

据环球网报道，近日，京东推出一款无人驾驶轻型货车，这是国内电商及物流领域首次推出无人货车产品，并且在交管部门指定的固定路段内开始路试。

这款货车是京东与上汽大通合作的EV80新能源无人轻型货车，是一款通过结合“无人驾驶功能”+“感知系统”+“新能源”等前沿技术武装的全新智能产品，以解决未来应用广泛的城市内物流运输需求。车辆通过搭载的雷达、传感器、高精地图及定位系统，在行进过程中，即使是150米外的障碍物也可以被提前探知，并且有足够的时间重新进行道路规划与障碍规避；当遇到信号灯时，前视摄像头可以准确感知，保障无人货车安全有序地平稳前行。

京东方面强调，虽然无人货车能够实现自动驾驶，但在进行路试时，依然在车上配备了驾驶员，应对可能产生的突发事件，确保路试安全的进行。他们还同时与东风汽车公司技术中心以及智行者开展了一系列技术合作，并联手推出东风电动无人轻型厢式货车。

MIT新机器人能自动“换衣” 可执行多种复杂任务

据新浪科技报道，近日麻省理工学院计算机科学与人工智能实验室（CSAIL）开发了一种具备极高适应力的机器人，能使用不同的“装备”让机器人拥有行走、涉水和飞行等执行不同任务的能力。

这种微型机器人尺寸仅有几厘米，外表是一个正方形，CSAIL称之为Primer。它能自己走进外包的“骨骼”中央，使用热量将其包裹起来，之后便可以完成不同的任务。研究人员称，Primer的移动受到内置磁铁的控制，它能够脱离外骨骼，找到新的外骨骼之后再自己“穿上”，像给自己穿上不同的“外衣”。它还能把不同的外骨骼拼接在一起，从而实现复合能力。在扩大体积并提供复杂的配置后，这些机器人可以具备更高的灵活性。例如，可以利用类似的装置来探索外星球的表面，或者地球的偏远地区，甚至利用不同的外骨骼在搜寻与救援任务之间切换。

一旦完成了特定任务，这种小型磁力机器人可以通过将自己浸入水中来摆脱外部的覆盖物。也就是说，机器人身上的外骨骼材料可以在水中溶解。

（本版图片除标注外来源于网络）