数据团队“隐形守护者”

从被动应对到资源输出，腾讯安全 20 年成长记

■作者 刘俊寰

根据最新发布的《IDC全球网络安全支出指南》预测，2020 年全球网络安全总投资将达到 1202.8 亿美元，较 2019 年同比增长 10.1%，而中国网络安全市场总体支出将达到 87.5 亿美元，较 2019 年同比增长 24.0%。与全球相比，中国网络安全市场近几年在国家政策法规、数字经济、威胁态势等多方需求驱动下，整体市场规模发展快速。

但同时，我们也应清醒地意识到，虽然中国网络安全投入在整体 IT 投入中的占比有所提升，但相对于全球平均水平还存在一定差距。国内的大量企业要么处于被动防御状态，要么缺乏自己的专业安全团队。

随着产业互联网的发展，尤其是随着年初疫情的暴发，加快了实体业务数字化上云的步伐。随之而来的，企业的安全防护特别是数据安全的防护也在面临新的挑战。

这也让企业的安全团队承担起“隐形守护者”的角色：这不是一个高调光鲜的工程，需要日积月累的持续建设，甚至你在日常工作中可能都感受不到它的存在；但是一旦有重大事故发生，就可能给公司造成巨大不可挽回的损失。在高速发展的产业互联网时代，大量新技术的出现让安全问题已经不是单点领域问题，而是一个系统工程，数据安全的应用范围早已超出了数据本身的安全，还涉及到整个安全体系。

作为科技领域的领军企业，腾讯安全团队在二十年的运营过程中，为云上安全积累了宝贵的经验。6 月初，我们跟腾讯安全副总裁黎巍聊了聊腾讯安全团队的建设问题：作为腾讯整个数据系统“保护者”，安全团队如何一步步发展起来，又是如何将这种能力输出到行业的？

要做好内部建设，也要走出去
安全团队建设的三个阶段

腾讯自身的安全建设，在二十年的历程中，经历了三个阶段。

第一个阶段是启蒙阶段，成立初期的腾讯和其他公司一样，安全建设以防御和对抗黑客入侵为主。

当然，要建立自己的安全团队不是那么简单的事，制定安全规范、构建安全体系，这些都是必要的。随着后期腾讯业务不断扩展，团队发现很多安全问题具有共性，如果只是一味被动应对，不仅会陷入“持久战”，团队也很容易进入疲态。

所以 10 年前，腾讯安全开始主动做一些安全的运营和建设，也正是这个时期安全团队的建设进入了第二个阶段，即把安全体系化和产品化，进行主动地运营。这个阶段，安全团队就总体目标达成一致——保证核心资产数据不会被窃取和丢失。这也是二十多年的发展中腾讯安全一直在践行的理念。

到第三个阶段，腾讯自身的安全生态已经做得比较系统了，但是放眼国内，还有不少企业在安全方面处于非常原始的阶段，这就触发了腾讯安全想要走出去，把 20 多年的安全经验和能力资源输出到整个产业中，帮助产业数字化转型。

黎巍坦言道，腾讯安全希望未来不只是为企业提供产品或解决方案，还能够为企业转型打造合适的安全战略观，更多维度帮助企业解决安全问题。

“超过 90% 的企业一定是在云上更安全”

但是也正如黎巍所说，“数据不是静态的，是动态的”，数据安全问题本身就是一个复杂的系统，这也是企业不愿意主动处理安全问题的原因之一。

疫情之下这个趋势更为明显，传统行业都在经历着数字化转型，原来的静态资产和业务逐渐数字化，进而用于服务客户，这是典型的数据流动问题。在这个高速流动的数据体系里，想要更好地实现防御安全，系统管理的视角必不可少。

去年，在腾讯的全球数字生态大会上，黎巍曾经表示，“90% 的企业一定是在云上更安全”。经过了疫情考验，黎巍告诉我们，从目前的实际情况来看，这个数字肯定要超过 90%。

受疫情的持续影响，以线下业务为主的中小企业收入明显下滑，线下模式受创后，众多企业纷纷转型线上。

但是面对迫在眉睫的生存问题，这些临时转型线上的企业往往会忽视重要的安全问题，或者因为资金和技术原因，暂时把安全事宜放在一边。这种情况下，从云上获取安全保护对企业来说是一个很好的选择。

由于云上已经构建好了完善的安全体系，企业只需在这个基础上投放业务即可，这与从零开始建设安全团队和安全防御产品相比要容易得多。

新技术触发新的安全危机：伴生型的安全问题
需要前瞻性的安全意识 与体系化建设

近年来，随着物联网、AI 大数据、5G 等一众新技术的出现，安全问题变得更为复杂。

这种技术的发展一定会带来新的变革，因为安全是伴生型的，一定会伴随着新的业态，演变出新的安全状态。因此安全团队的主要挑战在于，快速适应新兴的业态。

根据多年累积的经验，腾讯安全也总结出了一些不变的原则，对于此，黎巍分享道，首先是对安全问题的零容忍，只要出现安全问题，就当做头等大事应对，这是很重要的第一点。

另外，腾讯始终坚持践行最小安全权限原则，用现在比较流行的说法就是零信任。其实在 10 年前，腾讯安全就已经进行了初步贯彻，具体来说，就是对企业网络内外的任何人、设备和系统，都基于不信任原则进行系统建设。

最后就是安全能力的积累，腾讯有二十年黑灰产对抗的经验，这些经验的积累再加上威胁情报的分析，可以根据不同的业务、不同的业态进行快速地研究，构建产品和解决方案。

但是就国内整体环境而言，在安全方面的投入和意识都存在较大的缺陷。对此黎巍总结道，首先多数的企业缺乏组织安全团队的经验，其次，企业大多处于被动响应状态，不出问题就当成没问题，第三，企业的安全意识不够，与专业安全公司的互动也不强。

至于在安全方面国内企业能否实现“弯道超车”，黎巍坦言，与其期待超车，不如持续扎实地进行安全建设，毕竟底盘不稳的话，弯道翻车的情况更多。等底座扎实后，自然就有能力应对各种威胁了。

同时，国外的不少企业和团队安全体系相较而言都要系统得多，可以进行适当的参考学习。

对于如何将安全意识下沉到公司内部，黎巍表示，这一定是自上而下的过程。近几年从全球来看，安全方面的立法立规不断增强，比如欧洲 GDPR、美国 CCPA，中国也出台了一系列数据安全法规。这其实就是从顶层驱动告诉大家安全的重要性。

二十年前，腾讯就已经重视起了安全问题，如今腾讯内部有两大安全团队，安全平台部与企业 IT 部，这两个部门会根据国内外的法律法规的制度层面进行相应的内部建设。除此之外，运营和产品也都遵循着严格的安全体系和制度规范，比如在产品研发阶段，甚至会细化到某些高危函数的使用。

用 AI 预警新型威胁
人机协同仍是安全团队主要工作方式

去年，腾讯安全推出 AI 预警技术，进一步保障线上安全。

黎巍介绍道，其实在很多年前，AI 就已经在安全领域做出了不小贡献。就国内互联网行业传统的安全观念来看，会更偏向硬件，而近年来随着物联网、云的出现，安全问题就变得复杂起来。对企业而言，面对越来越多的新型未知的威胁，需要更加智能高效的系统来进行预测和判断。

就腾讯的所有安全产品而言，它都有一个基座，AI 就扮演了这样的一个角色，在这个基础上进行协同，深入到各种产品中，再进行更广泛地应用。

在腾讯安全内部 AI 的主要运用有两点，一个是对一些新型未知的威胁的预测和预判，还有就是协同的联动防御。

举例来说，腾讯安全有终端也有云，有时候在终端上发现了一些新型威胁，这就需要一个全链条的联动，把这种威胁贯通到云上。反过来也同样成立，这其实算是一种产品跨终端的联动，也是 AI 在基础驱动上的价值体现。

不管是对新型未知威胁的预测，还是协同联动防御，人机协同仍然是腾讯安全在提供服务时主要的工作方式，黎巍介绍道，人类和机器其实是一个相辅相成的过程。

腾讯安全每天的业务流量是非常庞大的，在这种海量的业务流量情况下，要找到新型未知的威胁，本身就是非常大的挑战，这其中我们就能通过一些算法、AI 技术，更好地进行分析和预警。

目前 AI 和大数据处理技术已经广泛应用于安全领域，不过黎巍认为 AI 仍然处于初级阶段，一些高级威胁分析还是离不开专业安全工程师的深度参与，但面对复杂的企业数字化业态以及云时代的海量数据安全挑战，他相信基于 AI 的安全技术演进将重塑安全产业，也将助力企业更加高效地应对数字化转型过程中伴生的各类安全威胁。E

监测污染排放发电厂，机器学习从太空怎么做

■来源 IEEE
■编译 杨威

矿物燃料发电厂是引起温室效应的最大气体排放源之一。

18000 座电厂温室气体排放量占全球总排放量的 30%，其中包括了每年大约 150 亿公吨的二氧化碳，燃烧矿物燃料产生的污染物也严重降低了空气质量和公共卫生，因此会导致心脏病和呼吸系统疾病以及肺癌，这将接近全世界十分之一的人口死亡数量。

为了避免空气污染和气候变化带来的严重影响，人们需要了解排放源。现有的技术可以检测大气中的 CO2 和其他气体含量，但其颗粒度还不够精确，无法确定谁排放了多少碳。

今年 7 月，一项名为 Climate TRACE 的新计划启动了，旨在准确跟踪的人为 CO2 排放源头，当然无论该源头在世界上的任何地方。由 9 个组织和前美国副总统阿尔·戈尔组成的联盟已经开始在七个部门（包括电力、交通运输和森林火灾）中追踪此类排放。

本文作者是机器学习研究员，与非营利组织 WattTime，Carbon Tracker 和世界资源研究所一起正在研究 Climate TRACE 的发电厂项目，他们主要使用了现有的卫星图像和人工智能技术，来估算世界上每一个化石燃料发电厂的排放量。

太空监测气体排放的瓶颈

美国是公开发布有关单个发电厂排放量的高分辨率数据的几个国家之一。美国的每个工厂都有现场排放监测设备，并向环境保护局报告数据。但是安装和维护这些设备的成本对许多其他国家而言简直高得离谱。因此其他国家报告的年度排放总量可能只是粗略估算，而不是实际测量值。这些估算值缺乏核实，可能会低估其排放量。

温室气体排放量之所以难以评估。一方面原因是，这些气体并非全部来源于人造，例如从海洋、火山分解以及土壤、植物和动物的呼吸中释放出的 CO2 和甲烷也将温室气体排放到大气中。另一方面，则是人类间接产生的，例如水泥生产和肥料。即使你知道排放源，由于排放量波动，估算量也可能很困难。而燃烧矿物燃料的发电厂会根据当地需求和电价等因素调整发电量。

在夏威夷的莫纳罗亚（Mauna Loa）天文台和美国国家航空航天局（NASA）的 OCO-2 等卫星，对 CO2 的浓度进行了局部测量。卫星不是直接测量浓度，而是根据从地球反射的阳光中有多少被空气中的二氧化碳分子吸收来估算浓度。欧洲航天局的 Sentinel-5P 使用类似的技术来测量其他温室气体。光谱测量非常适合创建大气 CO2 浓度的区域图。在大流行期间，这样的区域估计尤其显著，因为居家定单导致污染物减少。据报道，在城市周围，主要是由于交通运输量的减少。

但是这些测量的分辨率太低。例如，OCO-2 的每次测量都代表地面上 1.1 平方英里（2.9 平方公里）的区域，因此它无法揭示单个发电厂的排放量（更不用说自然发电厂的 CO2 了）。OCO-2 每天提供每个位置的观测结果，但由于云、风和其他大气变化而产生大量噪声。为了获得可靠的信号并抑制嘈杂的数据点，应在一个月内对同一站点的多次观测求平均值。

要估算源头的排放量，我们既需要足够高的空间分辨率以查看工厂运行情况，又需要不断观察这些测量值随时间的变化情况。

来自欧洲航天局哥白尼前哨卫星网络的图像显示了该发电厂运行时的烟雾和水蒸气流。

如何使用 AI 对电厂排放进行建模

我们很幸运，数十个卫星网络和数百个卫星正在实时捕获我们所需的那种高分辨率图像。大多数的地球观测卫星都在可见光谱中进行观测。此外，我们还使用热红外来检测热信号。

让分析人员查看来自多颗卫星的图像并将它们与其他数据进行交叉引用非常耗时，这种方式不仅非常昂贵且容易出错。随着我们合并来自其他卫星的数据，图像的数量将会增加。一些观测包含多个波长的信息，这意味着需要分析甚至更多的数据，并且需要经过微调，眼睛才能准确地识别。目前没有任何一个团队能在一定的时间内处理那么多数据。

借助 AI，游戏准则发生了变化。人类已经将深度学习应用于自动驾驶汽车的语音识别和避障，我们通过将相同的深度学习方法应用于气体检测，从而可以更快地预测排放，并增强从多种波长的卫星图像提取模式的能力，该算法的准确度取决于卫星的类型和电厂的技术。

我们首先将历史卫星图像与工厂报告的发电量进行匹配，以创建可以学习它们之间关系的机器学习模型。只要给定一个新的工厂图像，该模型就可以预测工厂的发电量和排放量。

我们在发电方面有足够的基础知识和数据来训练模型。美国和台湾是报告每小时排放量和发电量的少数几个国家中的两个。澳大利亚和欧洲国家仅报告发电量，而其他国家 / 地区则报告每日累计发电量。了解发电量和燃料类型后，我们可以估算未报告该数据的排放量。

在使用已知发电能力的电厂对算法模型进行训练后，我们便可以将模型应用于全球任何发电厂。我们的算法为各种类型的发电厂创建了预测模型，并且可以根据预测结果估算一段时间内的排放。

深度学习模型在卫星图像中寻找什么特征

在典型的矿物燃料发电厂中，温室气体通过烟囱排出，生成了模型可以发现的烟雾。效率更高的植物或采取二次收集措施以减少排放的植物可能会很难看到羽状流。在这种情况下，当发电厂的特性已知时，我们的模型会寻找其他视觉和热指标特征。

模型寻找的另一个特征是冷却。矿物燃料发电厂燃烧燃料使水沸腾，产生水蒸气，使涡轮旋转发电，然后必须将蒸汽冷却回水中，以便可以重复使用以产生更多的电能。根据冷却技术的类型，可能会从冷却塔中产生大量的水蒸气羽流，或将热量排放为附近的热水而释放出热量，而我们便使用可见光成像和热成像来量化这些特征。

将深度学习模型应用于监控全球电厂排放

到目前为止，我们已经使用来自美国和欧洲的发电数据创建并验证了一套燃煤电厂的初始模型。一支由科学家和工程师组成的跨学科团队，将继续收集和分析其他国家的真实数据。当我们开始在全球范围内测试我们的模型时，我们还将根据报告的年度国家总数和燃油消耗数据对它们进行验证。我们从 CO2 排放开始，但希望扩展到其他温室气体。

我们的目标是覆盖化石燃料发电厂的全球排放量，也就是说，对于任何国家的任何化石燃料发电厂，我们都能够准确预测其温室气体排放量。

接下来是什么？我们将公开排放数据。可再生能源开发商将能够使用它来确定新的风电场或太阳能发电场将产生最大影响的位置。监管机构将能够制定和执行新的环境政策。公民个人可以看到他们当地的发电厂对气候变化的贡献。它甚至可能有助于追踪《巴黎气候协定》的进展，该协定将于 2021 年重新谈判。E

BIG DATA DIGEST 大数据文摘

（本版由《大数据文摘》杂志授权转载）