01 | 安全的本质：数据被窃取后，你能意识到问题来源吗？

2019-12-09 何为舟

讲述：何为舟
时长 14:24 大小 13.20M

你好，我是何为舟。

今天是我们安全课程的第一讲，我们不会讲具体的细节内容。我打算先和你聊聊安全本身，以帮你建立整体的大局观。我确信，只要理解了安全的本质，在后续的课程中，你就更容易理解安全的概念和知识，也就能够建立解决安全问题的思维体系。

安全是什么？

首先，我们来看，安全是什么？

当你所在的企业内网被入侵，数据被窃取之后，你也许能知道，是某个业务漏洞导致黑客能够进入内网，但你是否意识到，数据安全保护机制上同样产生了问题？类似这种的问题有很多。当我们遇到某一个特定的攻击或者安全问题时，**往往看到的都是表象的影响，而能否找到根本原因并进行修复，才是安全投入的关键。**

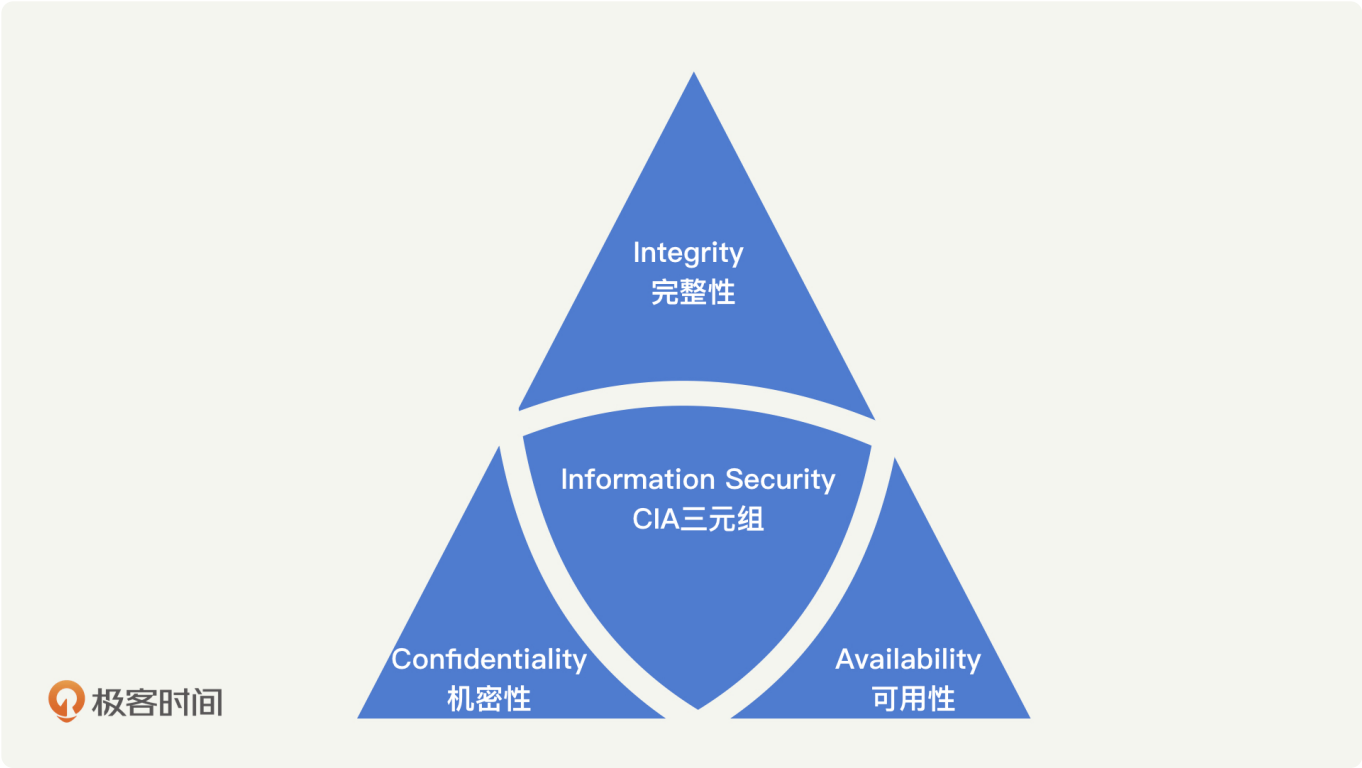
任何应用最本质的东西其实都是数据。用户使用产品的过程，就是在和企业进行数据交换的过程。比如，用户在使用微博时，或是将数据写入到微博（发博、评论、点赞等）中，或是从微博中获取数据（刷 feed、热门流）；用户在使用支付宝进行交易时，则是将资产以数据的形式进行转移。

因此，从另一个层面来说，安全的本质就是保护数据被合法地使用。怎么才叫“被合法地使用”呢？我们可以从机密性、完整性、可用性这 3 个方面具体来看。这也是在安全领域内最为基础的 3 个安全原则。

安全原则

机密性（Confidentiality）、完整性（Integrity）、可用性（Availability），我们可以简称为 CIA 三元组，是安全的基本原则。理论上来说，一个完整的安全保障体系，应该充分考虑到所有的 CIA 原则。当然，实际情况中，我们会根据企业需求，对安全在这三个方向上的投入做取舍。我们平时在评判一个企业的安全水平时，也会分别从这三个方向进行考量。

可以说，CIA 三元组原则，是安全领域内最基础也最重要的原则。你现在估计还没有感性认识，没关系，先有个整体印象，下面，我来给你详细讲解这三个原则的具体含义。



1. 机密性

我们先来看机密性。**机密性用一句话来说就是，确保数据只被授权的主体访问，不被任何未授权的主体访问。**简单用一个词总结就是“**不可见**”。

如何理解这个定义呢？举个例子，你不会允许陌生人查看你的个人隐私信息，但你可能会允许父母、朋友查看部分信息。同样的，对于应用中的数据，比如微信的朋友圈，你可以允许好友查看三天内的数据，但不允许好友查看三天前的数据。这些都是机密性在日常生活中的表现。

当然，首先你需要注意，机密性的一个前提是明确授权规则，也就是明确每一项数据可以被什么样的主体访问。在这个问题上，最安全的方法一定是，当每一次主体访问某一项数据时，都由相关负责人对该次行为进行审批。但是，这样显然是无法落地的，因为随着互联网的发展，每天都有万亿次的数据访问行为在发生。

因此，在安全领域我们提出了很多访问控制机制和安全模型，对数据和访问主体打上标签或者进行分类，并制定相应的访问控制规则去自动进行授权。关于访问控制机制， 在后续的内容中我们会再详细介绍，这里暂时不展开。另外，数据的存储、传输和处理过程也需要受到应有的保护。这些保护技术包括：加密、隔离、混淆、隐藏等等。

那么，针对机密性的攻击，都有哪些形式呢？

有的会直接针对保护技术进行破解。比如，去破解加解密算法、去逆向混淆代码等等。经过长期的发展，这些保护技术普遍都趋于成熟，安全性也在不断地提高。有了前人的积累，在保护技术上，我们其实不需要做太多投入，只需要采用最新的技术即可。

更多的时候，我们面临的机密性攻击，其实是人为原因导致的疏忽，也就是错误使用访问控制机制或数据保护技术。比如，因为权限滥用，导致开发人员拥有敏感数据的无限制访问权限；因为弱密钥，导致加密被破解；甚至显示器上的数据被别有用心的人窥探。所以说，**当前机密性保护的要点是引导人去做正确的事情，避免这类看似低级、实则普遍的漏洞发生。**

可以说，机密性是我们最容易理解的一个安全原则，也是企业在建立安全时最先想到的点。总的来说，机密性保护的技术都已经十分成熟了，但是在实施和落地的时候，往往会出现误用安全技术的情况。人的懒惰性是不可避免的，因此，机密性的安全保护往往都无法达到最佳状态，而是处于一个可用性和安全性的动态平衡点上。

机密性强调的是数据的“不可见”，但这并不代表数据是正确的。比如，将一个“True”存成了“False”，这就不是机密性要考虑的事了，而这种错误的存储，则是完整性需要考虑的事情。

2. 完整性

完整性就是确保数据只被授权的主体进行授权的修改，简单来说，就是“**不可改**”。

所谓“授权的修改”，就是对主体可进行的操作进行进一步的限制。比如，只能追加数据的主体无法执行删除的操作。以个人隐私信息为例，法律允许学校或者公司在个人档案内追加信息，但不能做任何修改。又或者说，你自己发的朋友圈，不希望被其他人进行修改。这些都是完整性的典型表现。

在授权方面，机密性中提到的访问控制机制同样适用。除此之外，完整性会更加强调对修改行为的日志记录，并有合适的监督机制进行审计。在保护技术方面，主要是利用加密、签名等技术，使得数据的完整性变得可验证。

你应该发现了，完整性和机密性是紧密相连的。因此，大部分的机制和技术都同时对完整性和机密性提供保护。

针对完整性的攻击也和机密性一样，更多的是由于人为原因导致的疏忽。除了黑客本身对数据的恶意篡改，已授权的主体也可能对数据完整性产生破坏，比如员工意外地误删除数据、程序 bug 导致错误数据被写入、正常用户的一些无效输入等。

相比于机密性，完整性往往容易被忽视。但是很多时候，机密性和完整性是共同出现的，做好了机密性的保护，基本也意味着做好了完整性的保护。因此，当我们在探讨安全问题、建设安全体系时，要将这两者结合起来，放在一起来研究。

机密性和完整性是为了保障数据是安全的，而数据的最终目的是要能够被看到或者使用。所以，对于数据来说，可用性也是很重要的一个方面。

3. 可用性

可用性应该是你最熟悉的原则。因为它不仅仅是安全方向上的问题，也是工程上面临的主要挑战。用一句话来说就是，**可用性就是确保数据能够被授权的主体访问到，** 简单来说，就是“**可读**”。

但事实上，可用性往往没有被划分到安全的责任中去，因为对于大部分企业来说，开发是最受到重视的，而开发会比安全首先去考虑可用性的问题。

举个典型的例子，面对高峰期的集中用户访问，如何保障用户能够正常地获取数据（“双 11”购物或者 DDoS 攻击等），你可以看到大量的研发人员对这个问题进行探讨和分享，但这其实都属于安全在可用性上的考量范围。

在安全机制上，我们要确保授权机制能够正确运行，使得拥有访问数据的主体能够及时地被授权，这是可用性的基本。那具体来说，可用性会面临哪些挑战呢？

在运维层面上，有很多技术在为可用性提供支撑，比如，在基础建设上的机房建设（如何在断电、高温、火灾等情况下保护设备）、多地冗余，以及在服务中的备份、资源冗余等。

在研发层面上，如何降低响应延迟、如何处理海量数据、如何在峰值进行扩容等，这些问题其实都是在可用性上的挑战。

在攻击的角度上，黑客也会对可用性发起攻击，也就是我们常说的 DoS（Denial of Service，拒绝服务）攻击。比如，通过发送大量的流量来占满带宽资源。

可用性一旦受到损害，其对企业的影响显而易见，也最容易受到关注。长久以来，无数研发和运维人员都投入了大量精力来进行完善。很多时候，可用性的投入，并不会非常精确地被划分到安全的责任中去。这正是我们最需要关注和去做的事情。

总结

好了，这一节的内容差不多了，我们来总结一下，你需要掌握的重点内容。

在所有的安全计划中，都会涉及对 CIA 三元组的取舍。不同的企业，在不同的发展阶段，CIA 都会有不同的优先级。什么是 CIA，你一定要牢记在脑海中，它将会贯穿我们整个专栏的学习。

通常来说，在互联网企业发展初期，可用性的优先级较高。如果涉及金钱相关的业务，则完整性的优先级更高；而涉及个人隐私相关的业务，则保密性的优先级更高。对于大部分企业而言，可用性在初期受到的挑战更多，则越发展越稳定，后期在可用性上的投入会逐渐降低。而完整性和机密性，会随着业务的发展，重要性越来越高，在企业的安全投入中，占比会越来越大。

因此，根据不同的发展阶段，列好 CIA 的优先级，是我们理解安全问题、定义安全需求、建设安全体系首先要做的事情。

课后思考

假设，你正在参加一个面试，面试官问：“你能否从 CIA 三元组的三个特性出发，结合你们公司的业务系统情况，和我分享下你理解的安全是什么？”你会怎么回答呢？

欢迎留言和我分享你的思考和疑惑，也欢迎你把文章分享给你的朋友。我们下一讲再见！

上一篇　开篇词 | 别说你没被安全困扰过

下一篇　02 | 安全原则：我们应该如何上手解决安全问题?

精选留言 (22)

写留言

攻城狮Ra
2019-12-10

公司是做ERP二次开发的，以可用性为主，在确保软件正常业务的基础上进行完整性和机密性的考量；理解：安全指得是自身合法利益的保障，自身相关的数据理应属于合法权益的一种，行使自身权益为可用性，争取自身权益为完整性，保障自身权益为机密性

展开

作者回复: 没太理解你描述的自身权益指的是什么?

1　2

张诚
2019-12-10

CIA，保密性，完整性，可用性。

主要做的是项目的支付模块，相应的支付借口的调用，敏感参数都进行了RSA非对称加密处理，并对总的做了签名校验。
对于订单支付操作做了完整的日志记录，以确保数据的完整性。
对于可用性只是采用项目的分布式部署以保证高可用。...
展开

作者回复: 挺好的，至少在各个方面都有考虑到，基础的安全能力都是有了的。后续无非考虑更深入的东西，比如rsa保证了接口调用的数据机密性，数据在服务器上是否有保护？日志的记录，有没有被篡改的可能?

2

Geek_98dc22
2019-12-14

看着老师开篇的内容，从安全方案的视角讲解CIA ，根据系统运行阶段划分，体现安全的实际应用，识别与认证，授权以及审计和问责。
以一个面对工业用户的产品为例，谈谈自己的安全理解。
1，识别与认证是非常重要的环节，因为一旦身份识破，对工业的损害会非常深远，因此是一个安全占比稍多的环节；多重识别手段+专用的认证工具。...
展开

作者回复: 有的，资深的黑客也一定会这么干。所以，日志一般会在不同的地方进行存储，就是为了增加黑客的难度，比如攻破了服务器A，但可能没办法删除服务器B上的日志。

1

小晏子
2019-12-09

从初创小型互联网电商的角度回答课后问题，考虑到CIA三要素，首先是A，如果可用性都不能保证，那么其他都毫无意义，其次是C，保护用户数据，订单数据是客户信任公司的关键，最后是I，之所以放在最后，是因为可能由于bug的存在导致用户操作权限混乱，但是这个至少不会导致公司垮掉，可以事后修补，所以放在最后。
展开

作者回复: 涉及金钱的对完整性还是有一定诉求的。比如用户的交易记录如果出现错误或者丢失了，可能最后对账都对不上，财务都过不去。不过在投入可以取舍，比如简单的多存几份日志，也能够提升完整性的保障。

1

鸵鸟
2019-12-09

嵌入式系统中bootloader开发主要会考虑机密性，对核心固件进行加密，同时将解密秘钥存储在安全存储中;完整性，对升级包镜像内容做SHA256，加载前做hash检验;同时还有最重要的一点——合法性，对镜像进行公私钥签名保证镜像来自合法的厂商。

展开

> 作者回复: 各领域都会有自己的特殊性，可以使用习惯的术语，但大体是相通的。即，你说的合法性，和完整性，其实做法和目的都是一致的。

1

Geek_70f787
2019-12-18

公司主要是做装修方便的，以可用性为主。其次是机密性，平台主要有多方用户——商家，厂家，装修用户。要确保各方用户不能看到别人的信息，特别是商家订单、装修用户。完整性的话，因为主要装修的话，是线下交易，所以在这方面会差一些

展开

一命赌快乐
2019-12-17

老师，这个面试题的讨论挺好的，希望多来一些。

展开

一命赌快乐
2019-12-17

安全就是保护对公司有价值数据的机密性，完整性，可用性。

二马
2019-12-15

机密性、完整性和可用性就像项目管理中的时间、范围和成本，三者在不同的场景中优先级是不一样的。安全防护的目标是进不来，看不见，改不了，拿不走，可追溯。
安全服务的对象是资产，数据是最核心的资产。不同的数据安全防护的侧重点不同，用户密码得看不见（加密存储），改不了（改动需要验证），拿不走（防止拖库）；用户资金余额侧重点则是改不了。...

展开

旺旺
2019-12-14

DDos攻击保护的什么数据呢？这应该是对系统的整体可用性产生了影响。

作者回复: 没错，DDoS主要破坏的是可用性。

进财
2019-12-13

当前情况下，我们是优先考虑可用性，但cia的三元组都比较重要，都应该考虑，但在具体资源投入，每一组要素投入的比重上怎么权衡，有没有什么标准去衡量。

展开

作者回复: 通过产出价值去衡量。评估一下，比如数据可用性受到影响，比如，数据不可用一小时，能产生多大损失。其他也一样，比如数据有10%被篡改了，会有什么损失。而这些损失，其实就是安全带来的价值。

Geek_114b64
2019-12-11

我们是做在线教育的，用户是上帝，我们做产品最终目的是要为用户服务，这样我们的产品才能有价值，所以可用我感觉是任何一个公司做研发首要考虑的，只有公司发展到一定阶段而且是发展的比较好，公司才会去或者说才有能力去考虑完整、机密两者，从我们公司来讲，线上直播教育课更倾向于可用，机密次之，完整性则最后考虑

展开

作者回复: 可以适当倾斜，但不要低估机密性和完整性带来的影响。比如用户在平台上填写的个人信息被泄漏，或者教育资源的内容被篡改，变成某些反动言论。这些安全事件的产生，是可以直接干跨一个公司的。

巨子联盟
2019-12-11

那么问题来了？数据泄露后，CIA怎么定位问题的来源呢？

作者回复: CIA不定位问题的来源，CIA只是告诉你要从哪些方面去考虑数据的安全性。比如数据泄漏了，说明机密性出了问题。至于具体是啥问题，可以根据后续的课程内容去具体分析。

丽莎
2019-12-11

作为一个出售动态防御的WAF公司，虽然不是传统业务，但是也可以使用CIA三元素进行分析。

对于我们最重要的也是可用性，由于我们的动态技术会对客户的HTML与HTTP通信进行修改，那么使得原有业务能在还原后被执行成为了第一重要性，我们通过一些自研的令牌...
展开

作者回复: 你好，感谢你的留言。总体分析的还是比较全面的。做防御产品的，可以从两个角度去分析CIA。一方面防御产品为客户的数据提供了哪些CIA保护，另一方面，防御产品本身自己又做了哪些CIA保护。WAF中的核心数据，我觉得有两种，一种是WAF自身的代码和策略，另一种是流经WAF的客户数据。如果后一种数据的CIA被攻破，导致客户数据在WAF中泄漏或者篡改了，作为一个安全产品，就很尴尬了。

Value
2019-12-10

计算密集型的应用的本质就不是数据...任何应用改为"数据密集型应用"或许更为恰当?

作者回复: 有一定道理。不过我理解所谓的计算密集型，其实就是处理某个数据的复杂任务，最终的输入和产出，其实还是某个数据吧。

张诚
2019-12-10

课程小记:
安全的基本原则：机密性，完整性，可用性。简称CIA。
机密性强调的是不可见性，数据只能被授权的主体访问。
完整性强调的是不可改，数据只能最追加操作，对数据的修改过程进行日志记录。
可用性强调的是可读，数据的可达性。

展开

米小亮
2019-12-10

信息安全的本质：保护对组织机构有价值的信息资产的CIA三要素不被破坏。公司的核心价值是什么，针对核心价值，确定CIA的优先级。

展开

作者回复: 可以尝试基于自己公司情况，总结分析一下。

我行我素
2019-12-10

目前考虑的是完整性，和机密性，之前刚开始的时候只要保证服务可用就行了（这个阶段除了问题都是靠人在后台处理数据）在目前趋于稳定后就渐渐的将重点放在完整性中了，确保在线上的问题，能通过补偿自动完成或撤销，还有大概1/4的人在处理机密性的问题，确保各个服务间的调用不可被外部其他人员获取；所以我的理解就是看目前公司处于什么阶段，那么所关注的重点也就不一样了

展开

作者回复: 赞

Cy23
2019-12-10

往往看到的都是表象的影响，而能否找到根本原因并进行修复，才是安全投入的关键。

星亦辰
2019-12-09

防盗刷和反爬属于 哪个范围呢?

感觉处于 机密性和可用性之间?

展开

作者回复: 盗刷更偏完整性，因为最终结果是用户的钱少了。当然，也会有诸如用户密码和个人丢失这种间接的机密性损失。爬虫，我认为属于机密性，因为爬虫相当于是通过你不允许的方式来获取的数据。