网络空间安全中的密码学 - 课程报告

安全领域的“密码”之我见 - 浅谈我自己与密码学

个人信息

姓名：宋林轲 学院：网络空间安全学院 学号：202228018670016

缘起

王老师您好，作为一个课程报告，鄙人起这个标题确实“大而无当往而不返”，不过我还是希望能斗胆谈谈自己至今的经历。

我现在是网络空间安全专业的学生。其实一开始选择学习这个专业确乎是我的一时兴起，2017年到2018年正是勒索病毒猖獗的年代，当时我抱着搞清楚勒索病毒这东西“应该是一件很酷的事情”这样很质朴的想法就成为了本科院校第一届做安全方向的学生，也全然不知道数学与密码学在这个领域有什么应用。

经过四至五年的学习，我在安全领域尝试了很多方向，也大体对网络安全的一些主流研究方向有了一定的认识。然而，对于本科的很多专业课，我大多保持着不求甚解的学习态度，只是享受掌握新知的片刻欢愉，而对网络空间安全的全局缺乏把握和联想。

现在回想起来，我个人感觉确实还是在密码学老师的讲授下对RSA、AES算法等产生粗浅的认识，在研究生师兄们的指导下在课余做CTF关于crypto的题目的那一段经历最让我回味：那段日子不管是上课还是业余实践都是很愉快的，看了victor shoup的密码学书籍并且乐在其中。然而，那时的我全然不知课本上的那些案例、CTF题目中的Demo之于现如今的系统安全，乃至网络空间安全究竟存在着多大的意义。我个人当时以为的密码的应用基本就局限于课本中所说的那些组件：

在这里我希望简单解答一下老师您上课时希望我们回去思考的一些问题：

1. 校验码与单向散列函数的区别：

> 虽然都在密码学中得到应用，但实际上这两个东西的区别很大，前者是简单的用于检验数据是否传输正确的方法，如循环冗余检验等等。在计算机组成原理中也有使用，如奇校验、偶校验方法等等，在信息论中也常常用于纠错。
>
> 而单向散列函数则是用于将数据压缩成固定程度的一类函数并散列开的函数，常用于数字签名等场景。

2. 为什么消息认证码不能向第三方证明？为什么不能防止否认？

> 消息认证码用于通信双方交互时检查消息是否受到篡改，即检查消息的完整性。但是，从第三方的视角来看，任何拥有秘钥的个人都能发送合法且完整的消息认证码，无从让第三方从消息认证码得知发送者的身份。发送者面对第三方也可以轻松抵赖，不承认是自己发送的消息认证码。
>
> 为避免这种情况，可以利用RSA密码这样的非对称加密技术设置公私钥对，以数字签名的形式防止抵赖。

3. 数字证书由谁签发？

> 是由CA来签发的，CA还是比较权威的第三方组织。

4. 单向散列函数与加密顺序问题？

> 一般的场景是先用单向散列函数进行散列化然后再进行秘钥加密，这样可以减轻一部分计算负担。

其时的我只是享受着做数学玩具的乐趣，也因此在大三夏令营保研的时候，在信工所老师们的询问下审视自己时，我陷入了犹豫。

“既然要确定接下来三年到五年的研究方向了，我需要看看我到底喜欢的是什么？我究竟是喜欢做数学还是真的很喜欢做密码？”

我最终得出的答案是我本科时期对密码学的热爱可能单纯只是搞清楚密码背后的数学原理，用代码算出crypto题目中的flag的喜悦感。我怀疑自己对密码学的爱源于自己对数学的喜欢：实际上直到今天我依然保留着业余时间翻翻数学书的习惯，哪怕自己很多时候也是不求甚解地享受片刻的欢愉。

于是，出于这种怀疑心理，我又换了一个研究方向，现在做的是系统安全。不过我很惊讶的是，方向的转变并不代表着与密码学的诀别。

系统安全方向？密码学？

我现在的导师在博士时期从事密码分析研究，现在从事系统安全相关研究。在保研的时候我联系他时，不曾想到现在有很多新颖的东西还是和密码学息息相关。当然，我的导师给了我很高的自由度，让我补充了很多本科前三年落下的知识，阅读了很多经典的书籍，也对网络空间安全和计算机科学这个整体有了一个粗浅的认识：相比于过往的盲人摸象，我感觉自己已经进步很多了。

大三的时候，信工所从事网络攻防的老师曾很鼓励我去认真学习密码学：“密码学是信息安全的根基。”从听说这个想法，到切身意识到这一点花费了我很长时间，在这过程中我对计算机、网络空间安全的认知也愈发呈现出一个整体。即便看上去是和密码学没有太大关系的系统方向：所谓挖掘系统存在的漏洞、或者构筑安全的系统这样的和计算机关系过大的研究姑且不论，毕竟他们设计出来就带着安全的目的，哪怕是出于性能优化、效率提高的一些系统设计，都带着很多密码学的影子。

密码学之于网络空间安全，乃至整个计算机科学的根基效应，在我看来有两个角度：

首先是密码学带来的反光效应。我们简单提两个例子：

1. 为了防止自身密码被轻易攻破而提出的“最小熵”概念，在victor shoup和dan boneh的书籍中提到它是为了把可能性尽可能地分摊，为此需要依赖一个强大的哈希函数把东西散列开，同时这个哈希函数还得防止被彩虹表撞库攻击到。而如果我们从它的反面来思考，就知道为何哈希函数不只是在应用于密码学之中了：为了防止多核系统、多个磁盘、多个网卡等中出现部分资源被大量使用，而其余资源却鲜有人问津造成系统寿命缩短问题，可以利用哈希函数实现负载均衡和资源调度。
2. 用于纠错码的安全设计，实际上在很多处理器设计分支预测器的时候使用上了。由2个bit构成的纠错码，能够有效地将10和01这样的错误码纠正成11和00，而传统分支预测器中同样采用了这样的结构来记录历史状态，实现分支预测和性能提高。

其次是密码学在新兴场景下的应用与强大生命力，计算机工程能够解决架构的问题，是为骨架，而面向安全的需求则需要密码学作为血肉与灵魂：

在部分云计算场景下，用户虚拟机需要依赖强大的秘钥实现内存的加密，以保护自身隐私。如果仅从计算机工程的角度实现隔离，则无法建立可信执行环境，因为缺乏厂商用于背书的信任根TPM和一系列信任链的传递，在远程认证的时候，依旧无从证明自身的可信。

密码学拥有强大的生命力，有很多安全的需求都可以通过密码学问题来解决。随着chatgpt等大模型语言模型的横空出世，计算机领域其实有很多方向受到了巨大的冲击。然而，密码学作为网络空间安全的根基，或许在较长一段时间内都将处于一个较为稳固的地位。