第二部分：访问控制

03. 隐通道、推理控制和Captcha

厚德健行

内容

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- 隐通道的概念
- 隐通道的分类
- 推理控制
- Captcha

隐通道的概念

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- Lampson于1973年提出
- 比较通泛的定义是“能让一个进程以违反系统安全策略的方式传递消息的信息通道”，如：不同安全层次的主体共享的资源可以用来传送信息
- 隐通道：“一个并不是系统设计者所预期存在的通信路径”

隐通道的分类

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- 存储隐通道与时间隐通道
- 噪声隐通道与无噪声隐通道
- 聚集隐通道与非聚集隐通道

存储隐通道与时间隐通道

- 存储隐通道：一个隐通道涉及对一些系统资源或资源属性的操作(如是否使用了一个文件)，接收方通过观察该资源及其属性的变化来接收信息所形成的隐通道。
- 时间隐通道：接收方通过感知时间变化来接收信息所形成的隐通道。
- 通常，一些存储资源与一些临时行为(如查找时间)相关，所以涉及对它们的操纵同时表现出了存储隐通道和时间隐通道的特性

例1 目录结构隐通道

- 假定Alice有绝密许可而Bob仅有秘密级许可
- 如果文件空间被所有用户共享
- 若Alice要发送一个1给Bob，她将建立一个文件名为FileXYzW，如果Alice要发送0，她不会建立这样的文件
- 每分钟BOb更新文件列表
  - 若FileXYzW文件不存在，则Alice发送0
  - 若FileXYzW文件存在，则Alice发送1
- Alice能向Bob泄露绝密信息!

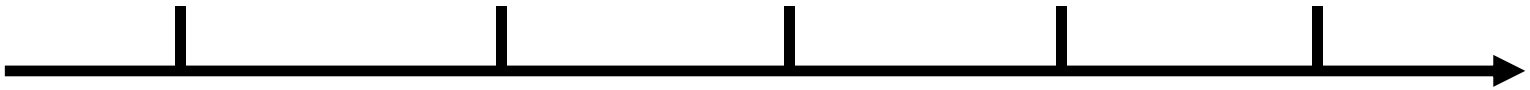
例1 目录结构隐通道

Alice: 建立文件　删除文件　建立文件　建立文件　删除文件

Bob: 检查文件　检查文件　检查文件　检查文件　检查文件

数据: 1　0　1　1　0

时间:



例2 CPU共享隐通道

- 假设有A和B两个进程，B的安全级高于A，B企图将信息传递给A
- 它们约定一系列间隔均匀的时间点t1，t1，t1，…（间隔时间至少允许两次CPU调度），A在每个时间点都请求使用CPU，而B在每个时间点，若要发送0，则不请求使用CPU；若要发送1，则请求使用CPU（假设B的优先级高于A）
- 在每个时间点
  - A若能立即获得CPU的使用权，则确认收到0
  - 若要等待，则确认收到1

例3 打印机联接隐通道

- 假设S和R分别是可以连接到打印机的设备，S想传信息给R。
- 每当S要传送“0”时，检查自己是否已联接到打印机，若联接打印机，则释放联接。每当S要传送“1”时，需要联接打印机。
- 在每个时间点，R每当试图联接打印机
  - R若他能联接，则确认收到“0”
  - 否则，确认收到“1”
- S和R要经过适当的同步处理（如R接收到“1”后，要立即释放打印机），便可以在S和R之间传送连续的比特流。

例4 进程号隐通道

- 进程号（PID）是系统中标志进程的唯一符号，在许多操作系统中采用连续递增的方法来管理进程号，即新建的进程的进程号在上一个进程的进程号的基础上加1，发送方想给接收方传递信息。
- 接收方建立一个子进程并立即结束它，记录下它的进程号；
- 发送方若发“0”，则什么也不做，若发“1”则建一个子进程并立即结束它；
- 接收方再建一个子进程并立即结束它，记录下它的进程号
  - 如果新的进程号与上一次得到的进程号相差1，则确认接收到“0”
  - 如果新的进程号与上一次得的进程号相差2，则确认接收到“1”；

- 其他例子
  - ACK消息
  - 网络通信等.
- 隐蔽通道存在的条件：
  1. 发送者和接收者必须能访问共享的资源
  2. 发送者必须能够修改共享资源的模型特性，且这些修改时接收者能够观察到的
  3. 发送者和接收者必须能够通信同步

噪声隐通道与无噪声隐通道

- 无噪音隐通道：在一个隐通道中，如果信息发送者发送的信息能够被接收方完全正确接收。
- 噪音隐通道：接收者所接收的信息要少于信息发送者所发送的信息。
- 可以通过使用纠错码将噪音隐通道转变为无噪音隐通道，但是这种转变在降低传输出错率的同时限制了原有的通道容量

聚集隐通道与非聚集隐通道

- 聚集隐通道：在一个隐通道中，为实现数据通信，多个数据变量(作为一个组)作为同步变量或信息。
- 非聚集隐通道：也称为单一隐通道，它仅影响单独的数据变量。
- 根据通信双方进程设置、读取和重置数据变量的方式，可以采用序列、并行或混合方式形成聚集信息传输通道以获得最大带宽。

结论

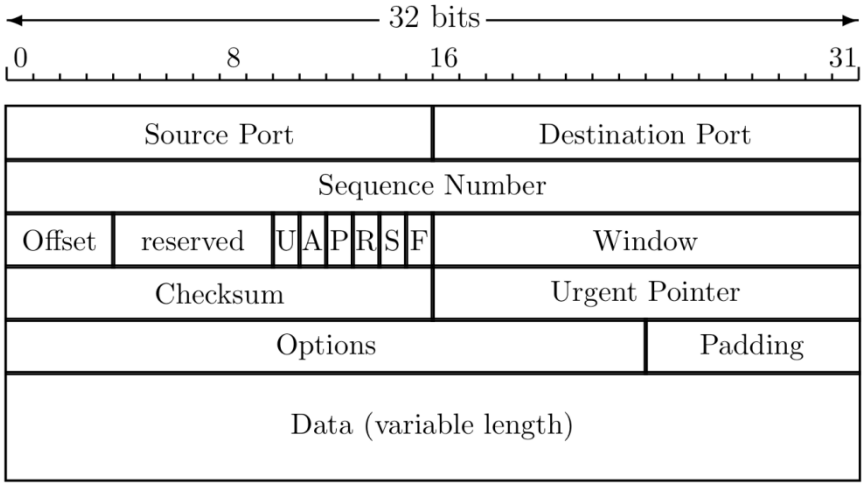
- 隐蔽通道是普遍存在的
- 简单消除隐蔽通道的方法...
  - 取消所有的共享的资源和所有通信
- 在任何有用的系统中消除所有隐蔽通道是不可能的
  - 美国国防部的指导方针: 将隐蔽通道的影响减少到 1b/s
  - 言下之意是其放弃了试图消除隐蔽通道

隐蔽通道

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- 考虑一个大小为100MB的绝密文件
  - 假定文件的明文版本存放在绝密文件系统中
  - 而加密版本(用AES和一个256位的密钥)存放在不保密的位置
- 假定能够将系统的隐蔽通道容量降低到1b/s
- 那么通过隐蔽通道将会需要25年以上才能泄露出整个100M文件
- 但是通过隐蔽通道花费不到5分钟就可以泄露出256位AES密钥!

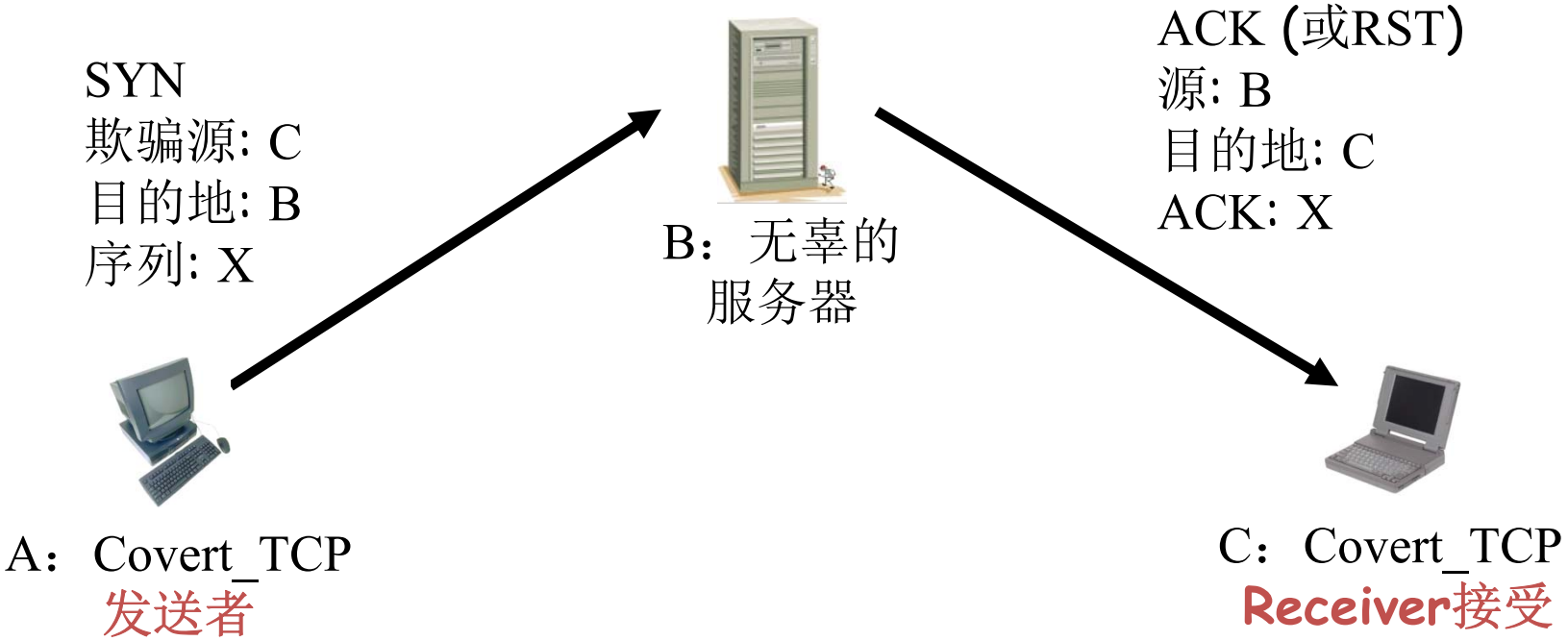
现实的隐藏通道



- 将数据信息隐藏在TCP头部的保留字段中
- 或使用covert_TCP工具将数据隐藏在：
  - 序列号
  - ACK字段

现实的隐藏通道

- 在TCP序列号或ACK字段中隐藏信息
- 工具: covert_TCP
- 发送方在序列号X中隐藏信息



浙江工业大学
ZHE JIANG UNIVERSITY OF TECHNOLOGY

推理控制

示例

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- 假定查询数据库
  - 查询: 在SanJose州立大学中女性计算机科学教授的平均工资是多少?
  - 答案: 95,000美元
  - 查询:在SanJose州立大学中女性计算机科学教授的人数为多少?
  - 答案: 1个
- 特定的信息从普通问题的回答中泄露出来了!

推理控制的研究方面

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- 如医疗信息。它属于私密信息，但又很有研究价值
- 什么办法既允许访问统计的重要数据，又能够保护隐私呢?
- 如何能让重要数据可被访问但同时又不泄露特定信息?

朴素推理控制

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- 在医疗记录上隐藏所有的名字和地址?
- 但仍然可以通过简单的方法从这些“匿名”的数据中得到特定的信息
- 隐藏姓名不是一种好的办法
- 还有其他方法可以提供更强大的推理控制吗?

- 控制查询集大小
  - 当结果集太小就不返回结果
- N回答, k%控制规则
  - 若一查询结果的k%或更多结果是由N个或更少的主体所提供的，则不允许查询结果的发布
  - 例如: 查询比尔盖茨所在街区的人均薪水
  - 这个技术被美国人口普查办公署应用在信息收集中
- 随机数据扰乱
  - 将一部分随机噪声加到数据中去
- 其他方法—但都不太理想

- 我们把表1中前4条记录当做数据集D1,前5条当做数据集D2。
- 计算D1和D2
- 隐私算法A=count（i）+噪音，

其中i=1,2,3.....。

- count(4)=2
- count(5)=3
- count(4)+噪音与count（5）+噪音其结果均以几乎完全相同的概率输出{2,2,3,4}

姓名	有糖尿病
Ross	是1
Monica	是1
Joey	否0
Phoebe	否0
Chandler	是1

推理控制的讨论

- 强的推理控制似乎实现起来是不可能的
- 有弱的推理控制是否比没有要好呢?
  - 是:肯定能降低信息泄露的数据
- 有弱的隐蔽通道保护是否比没有要好呢?
  - 是:肯定能降低信息泄露的数据
- 弱的密码加密手段是否比没有加密要好呢?
  - 不一定: 加密会指示出哪些是重要数据
  - 可能较容易地将一些数据从大量数据中过滤出来

浙江工业大学
ZHE JIANG UNIVERSITY OF TECHNOLOGY

CAPTCHA

全自动区分计算机和人类的图灵测试

图灵测试

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- 在1950年，由Alan Turing提出
- 测试过程是一个人向另一个人和一台计算机提问，但发问者看不到任何回答者，必须通过答案确定哪个回答者是人，哪个是计算机。
- 如果访问者不能解决这个问题，则该计算机将通过图灵测试
- 人工智能中的“黄金标准”
- 目前，还没有计算机能达到接近通过图灵测试的水准
  - 但一些声称通过了

- **CAPTCHA** — 全自动区分计算机和人类的图灵测试(**C**ompletely **A**utomated **P**ublic **T**uring test to tell **C**omputers and **H**umans **A**part)
- 自动 — 假定测试是由计算机程序生成
- 公开的 — 程序和数据都是公开的
- 图灵测试指出… — 人可以通过但计算机却不能通过的测试
  - **HIP** == **H**uman **I**nteractive **P**roof
- 反转的图灵测试

CAPTCHA悖论

- “…CAPTCHA是可以生成并评分的测试程序，但是它自己不能通过测试 …”
- 自相矛盾 — 计算机能够建立并评分的测试而自己不能通过!
- CAPTCHA 用于限制人类访问的资源
- CAPTCHA 可以看成是某种形式的访问控制

CAPTCHA的应用

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- 最初动机: 自动投票来选举计算机科学毕业生中设计做得最好的
  - 州立圣何塞大学 VS 斯坦福大学
- 免费邮箱服务——用于阻止从大量自动注册的邮箱账号发送的垃圾邮件
  - CAPTCHA被免费的e-mail服务所使用

CAPTCHA的规则

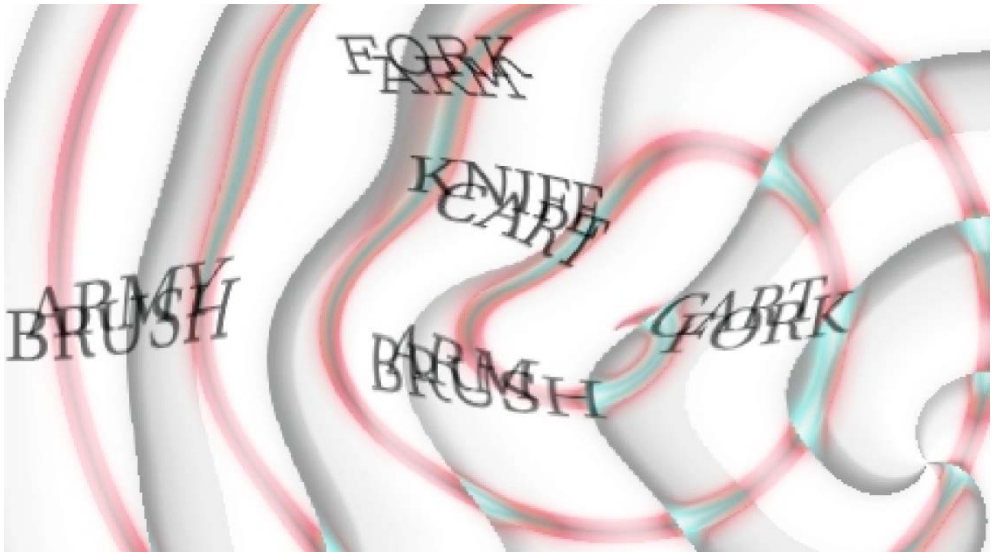
浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- CAPTCHA的要求是：
  - 必须对大多数的人来说通过非常容易
  - 对于计算机来说通过困难或不可能
    - 即使这些计算机能够访问CAPTCHA的软件
- 对某些人不能通过特定类型的情况下，希望有不同的CAPTCHA类型存在
  - 盲人不能通过视觉的CAPTCHA

CAPTCHA真的存在吗

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- 例: 从图像中找出三个单词



- 对大多数人来说很简单
- 但OCR问题对计算机来说却很难
  - OCR=光学字符识别

CAPTCHA的类型

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

- 当前的几种类型
  - 视觉的
    - 如上图的例子
  - 音频的
    - 对声音的歪曲处理
- 尚没有基于文本的CAPTCHA

练习

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

1.假如Alice是12306的后台管理员。

a.春运期间，Alice如何保证12306系统的可用性和安全性?

b.假如Alice用了图片验证，请问，Alice如何增强系统的安全性?

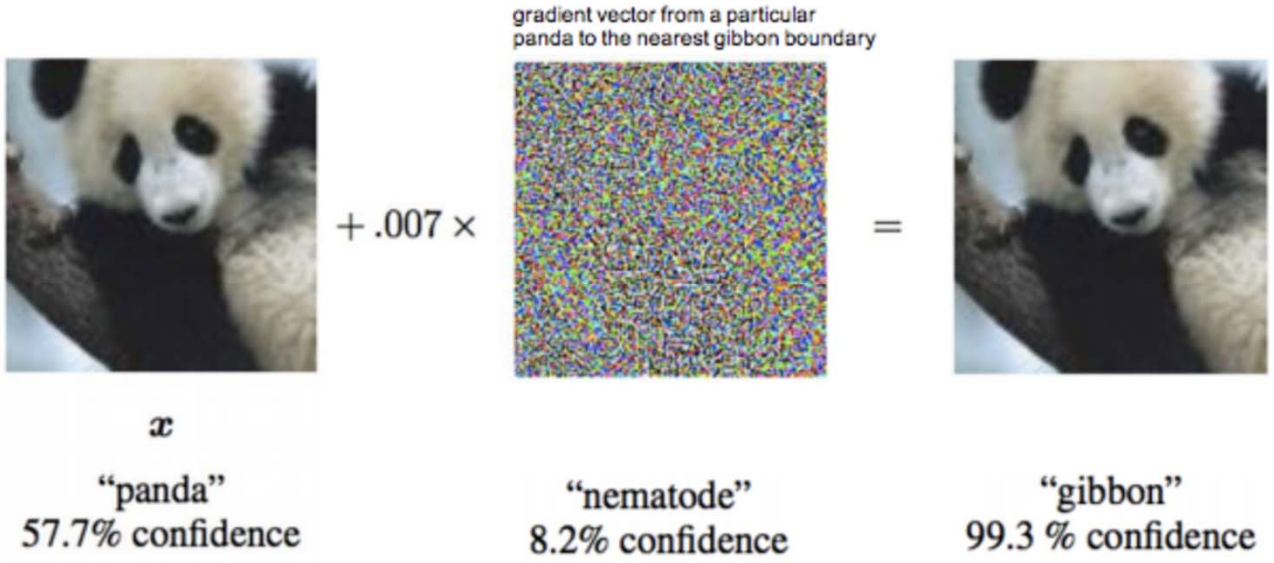
CAPTCHA和人工智能

浙江工业大学 ZHE JIANG UNIVERSITY OF TECHNOLOGY

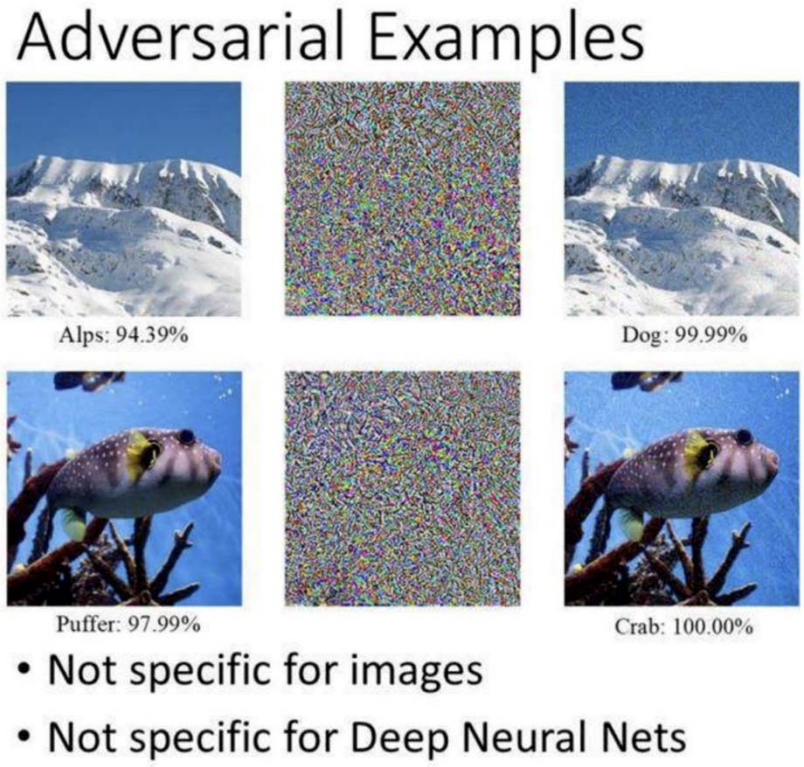
- OCR 是AI中具有挑战性的问题
  - 困难的部分是分割问题
  - 但人能很好的解决这个问题
- 歪曲的声音可能成为很好的验证码
  - 人还是能很好的解决这个问题
- 如果攻击者能够破坏CAPTCHA，那么他们就能解决了AI的疑难问题
  - 攻击者的这些努力将会有好的用途
- 可能有其他方法可以破坏CAPTCHAs...

- 深度学习的脆弱性

- 2015年，“生成对抗神经网络GAN之父”Ian Goodfellow在ICLR会议上展示了攻击神经网络欺骗成功的案例



- 2017NIPS对抗样本攻防竞赛案例



- 在2018年，Ian Goodfellow提出了首个可以欺骗人类的对抗样本

- **攻击方：**

通过生成更具迷惑性的对抗样本，使现有的深度学习模型识别出错。

- **防御方：**

训练更具鲁棒性的模型，使模型练就一双“火眼金睛”，正确识别对抗样本。

- **对抗训练：**

在训练模型的时候就加上对抗样本，相当于让深度学习模型在做一份考试真题，等真正上战场的时候，碰到对抗样本也无所畏惧。

关注我，下节内容更精彩：
04：防火墙、入侵检测