2018-2019学期秋季
信息隐藏课程
第1讲 绪言

中国科学院信息工程研究所
INSITUTE OF INFORMATION ENGNEERING CAS

SKLOIS
信息安全国家重点实验室

赵险峰

中国科学院信息工程研究所
信息安全国家重点实验室

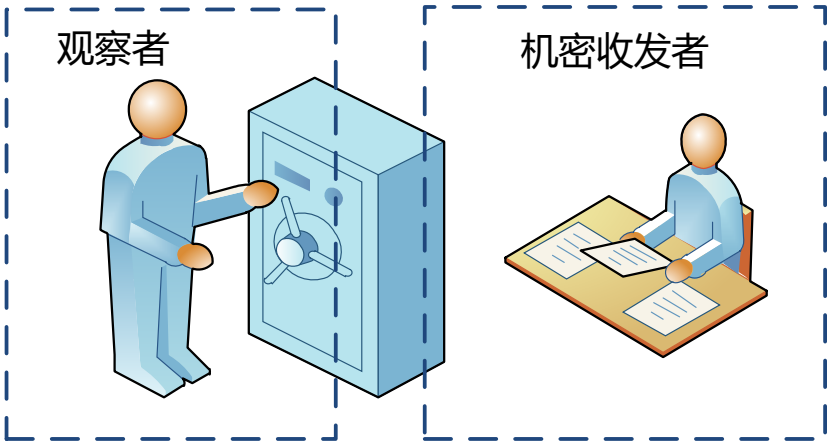
2018年9月

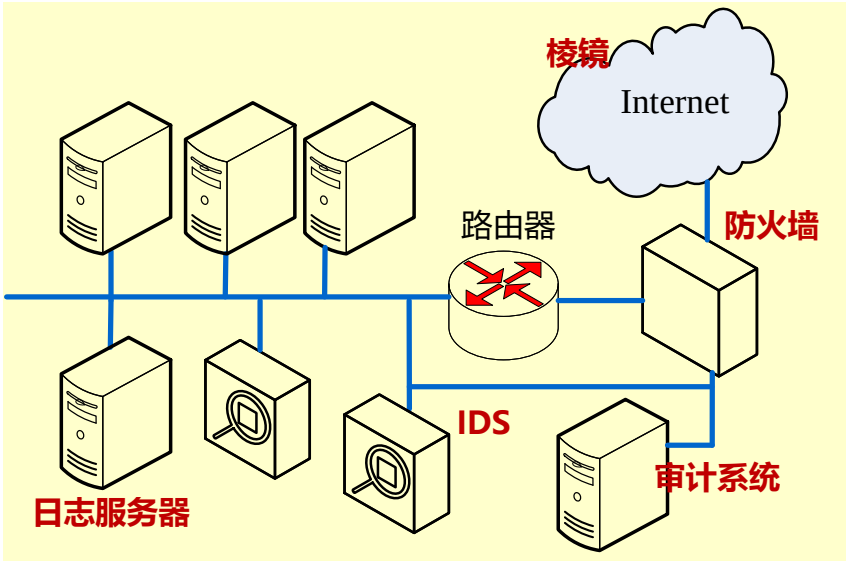
纲要

1. 从密码到信息隐藏与隐写
2. 隐写的发展与应用
3. 隐写安全指标
4. 文献阅读推荐

1-1 一个问题：保密通信 = 加密通信？

- 两种保密性：
  - 保护保密通信的内容——加密
  - 保护保密通信的事实——隐蔽通信（隐写），某些环境下更需要

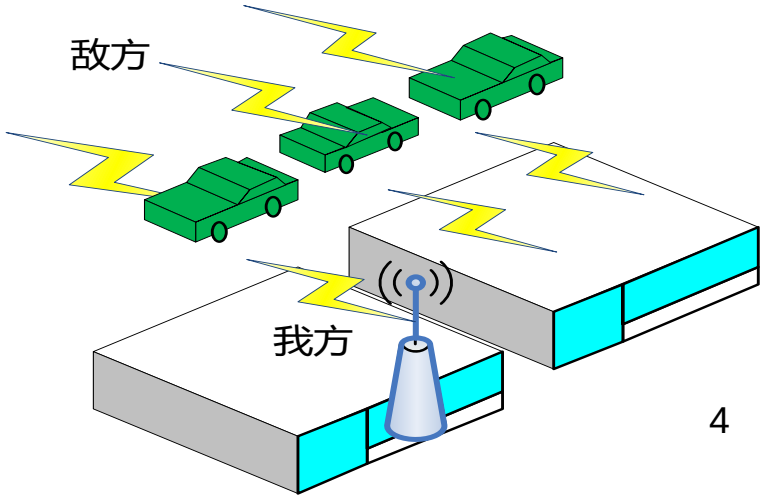




看守者-囚犯模型

1-2 未进行行为隐藏而暴露的保密通信例

- 电影/电视剧《永不消逝的电波》，延安情报人员李侠与日本鬼子的对抗:
  - 我方：隐姓埋名，假扮夫妻（日常行为有隐蔽性）；加密发报，仅仅信息内容有隐蔽性
  - 敌方：日本特高科在上海试用无线电探测车，可识别新信道和加密电波的存在，可定位
- 结论：没有保护保密通信的事实，会造成情报员的暴露

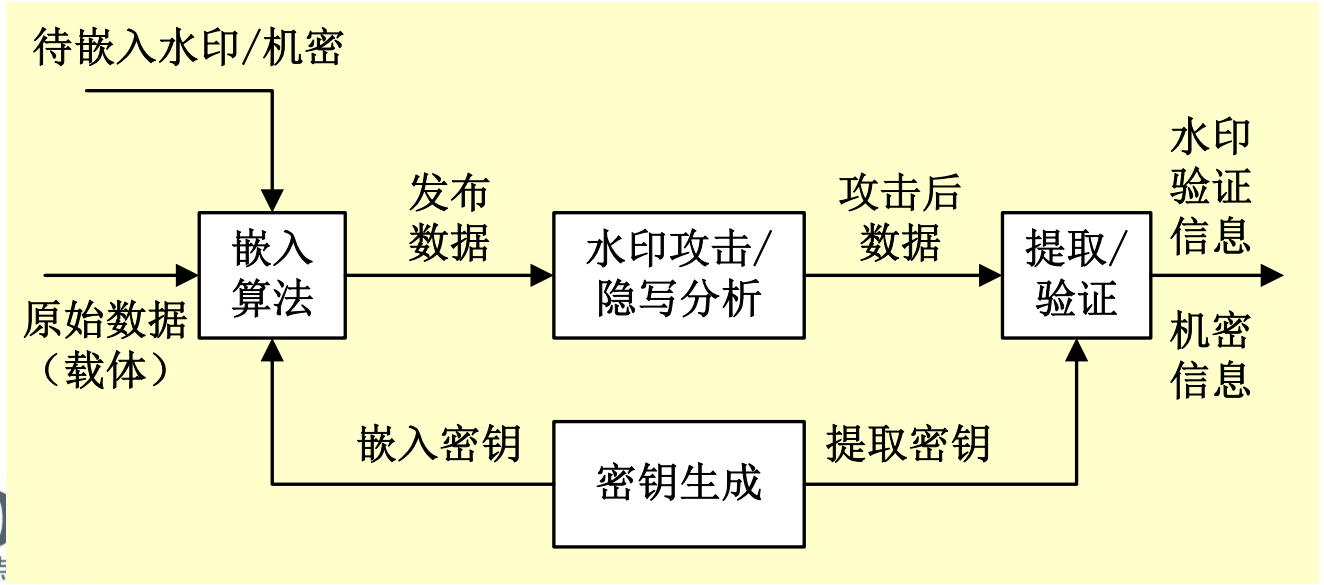


1-3 密码的局限

- 密码主要解决了消息保密传输、数据来源认证与完整性认证等信息安全问题。但是，密码方法并不解决以下两方面的问题：
  - **保密通信的行为隐蔽性问题**：伪随机性（Pseudo Randomness）是密码的标志性特征
  - **松散环境下的内容保护与内容认证问题**：密码保护数据，但内容≠数据？怎么保护多媒体内容？内容肆意散布怎么管？转码了怎么认证（密码过于敏感）？合法用户居然不可信！怎么约束他？内容怎么进行完整性保护和认证？松散环境下数字签名被删除了怎么办？篡改了能否定位？

1-4 信息隐藏基本概念

- 信息隐藏（Information Hiding）针对上述安全需求，建立了自己的理论基础与方法体系。
- 信息隐藏又名数据隐藏（Data Hiding），是指将特定用途的信息隐蔽地藏于在其他载体（Cover）中，使得它们难以被发现或者消除，通过可靠提取隐藏的信息，实现隐蔽通信、内容认证或内容保护功能
- 主要包括鲁棒水印（Robust Watermarking）、脆弱（Fragile）水印与隐写（Steganography）



1-5 鲁棒水印

- 20世纪末期，社会对数字内容的版权保护问题日益关注，人们认识到，仅仅靠法律保护版权是不够的，出现了数字产权管理（Digital Rights Management，DRM）技术
- 鲁棒水印是重要的DRM与安全标识技术之一，指将与数字媒体版权或者购买者有关的信息嵌入数字媒体中，**使攻击者难以在载体不遭到显著破坏情况下消除水印**，而授权者可以通过检测水印实现对版权所有者或内容购买者等标识信息的认定

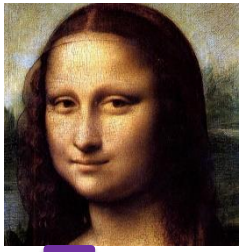
旋转			0.9857 (水印提取有效)
上下采样			0.9984 (水印提取有效)

1-6 鲁棒水印——性能需求

- 鲁棒性（Robustness）。指在主动攻击下，授权用户仍然能够提取水印信息
- 水印容量（Capacity）。指水印能够可靠传输的信息量
- 安全性（Security）。指水印攻击者难以从水印的算法、应用协议或者实现方法上获得有益于攻击的信息
- 盲性（Blindness）。指水印的检测不依赖与原始媒体的存在

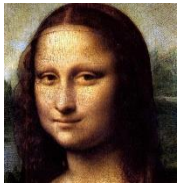
- 当前实现完全有效的鲁棒水印难度非常大。尤其是在尺寸缩放和裁剪等几何攻击（Geometric Attacks）下，水印检测或提取非常困难，设计抗几何攻击的水印已成为挑战
- 但是也可认为，由于水印攻击降低了媒体感知质量，水印在某种程度上是成功的，就像门锁一样，虽然防止不了破门而入，但是还是起到了防范作用
- 以上情况毕竟影响了鲁棒水印的应用

鲁棒数字水印面临解决的难题

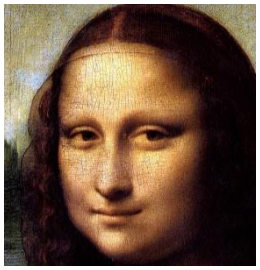


原始提取

水印在一些攻击下的提取效果



缩放



剪裁

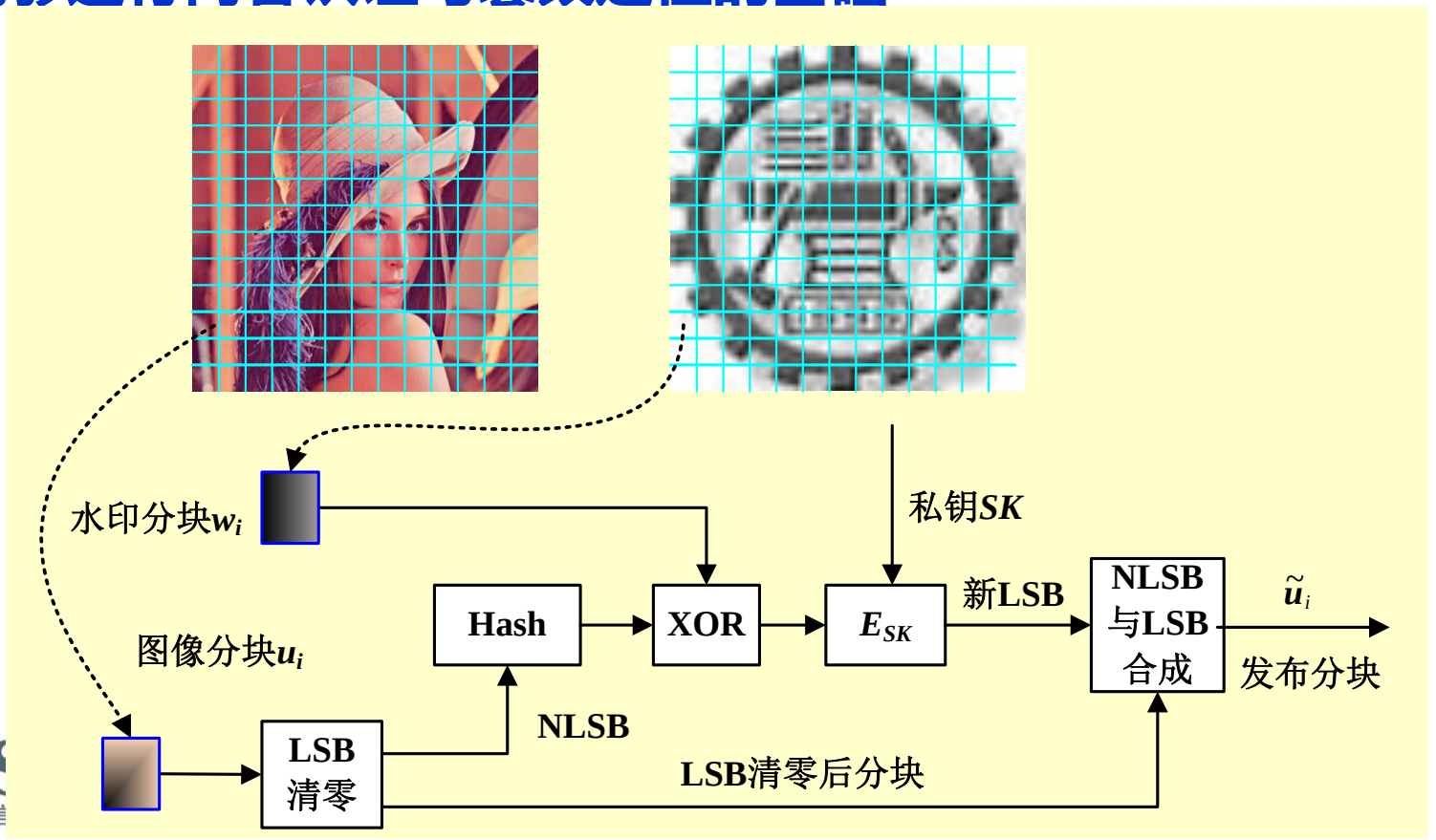
在保证感官质量的基础上嵌入水印（感知透明性）
在内容受到一定处理或干扰情况下仍可以提取水印（鲁棒性）

旋转

SKLOIS
信息安全国家重点实验室

1-7 脆弱水印

- 脆弱水印技术将防伪信息隐藏在内容本身中，以后通过水印检测发现篡改，并发现篡改的位置，方便地支持了被保护内容的安全流动，避免了由于数字签名是单独数据成分造成的协议开销
- 隐藏在被保护内容中的信息会随着内容的改动而变化，这就是它能够进行内容认证与篡改定位的基础



1-8 脆弱水印——性能需求

- 脆弱性（Fragileness）。嵌入被保护内容的水印应随着内容的改动而变化
- 定位精度。被嵌入水印随着内容的改动而变化需要反映内容被篡改的位置
- 可逆性（Reversibility）。指嵌入的水印可以被授权者完全消除，并且原始媒体能够得到还原
- 安全性（Security）。指水印攻击者难以从水印的算法、应用协议或者实现方法上获得有益于攻击的信息
- 盲性（Blindness）。指水印的检测不依赖与原始媒体的存在

- 多数脆弱水印方案距离对有效的内容（而不是对数据）保护的需求还有一定距离。有的脆弱水印被称为是半脆弱的（Semi-fragile），它们只对内容的变化敏感而允许内容可以接受编码等正常处理，显然这样的水印更接近于实现内容认证与保护的目标，但难度很大

1-9 插话：主动取证 vs 被动取证

- 主动取证：存在安全预处理环节，如水印技术；更准确
- 被动取证：不存在安全预处理环节，如内容盲取证；适用面更宽

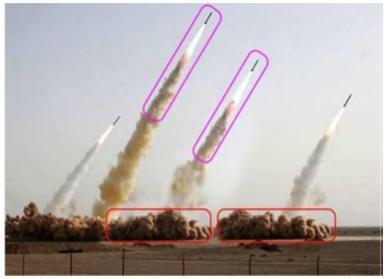
(a)

(b)

(c)

(d)

(e)



SKLOIS
信息安全国家重点实验室

1-10 隐写

- 隐写是基于信息隐藏的隐蔽通信或者隐蔽存储方法，它将机密信息难以感知地隐藏在内容可公开的载体中，在保护保密通信或者存储的内容同时，也保护了这种行为事实
- 称隐写后的载体为隐文（Stego-text）或者隐写媒体（Stego-media)
- 相比密码方法，隐写一直在信息的传输率上处于较大的劣势，因此有很长一段时间没有获得显著的发展。当前，随着网络与数字媒体的普及，这个情况正在迅速改变

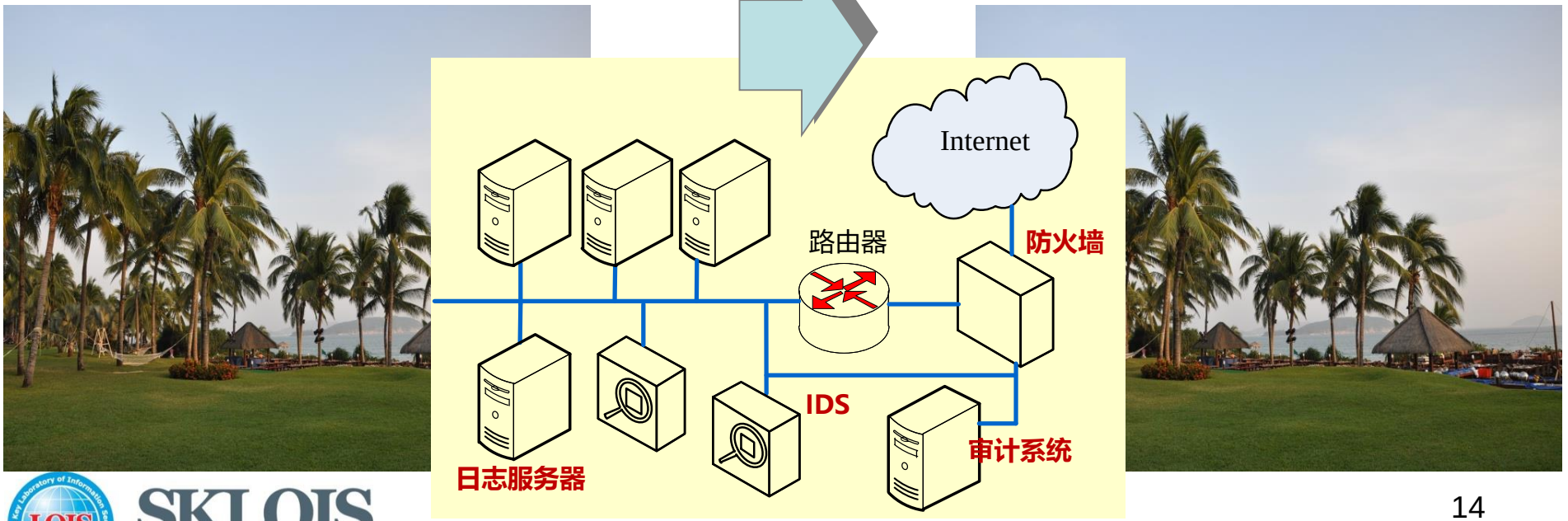
冗余性是隐写的安全基础，多媒体信息冗余大

8MB

3MB

1-11 隐写带来的挑战

- 多媒体的应用普及加大了当前隐写信道的容量：
  - 数字图像、视音频隐写（1千万像素可隐藏40万字节、20几秒视频能隐藏1个word文件）
  - 批隐写将大尺寸文件隐藏在多个媒体文件中
- 当前的安全防护措施仅仅审计“明信道”；盲隐写识别非常困难



1-12 隐写的性能需求（1）

- **安全性。**隐写的首要安全性是特征隐蔽性，因此，隐写的安全性一般就是指隐写后媒体的特征变化隐蔽性；这里分为知晓或者不知隐写算法的两种情况，后者情况下的安全也称为隐写在盲检测（Blind Detection）下的安全
- **隐写容量。**可用负载率（Payload）表示，它表示平均每一个嵌入位置所能承载的隐蔽信息量，令 m 表示传输的消息量，n 为嵌入样点的数量，则负载率 α 的计算公式为 $\alpha = \frac{m}{n}$；bpp（bits per pixel）、bpnac（bits per nonzero AC coefficient）
- **嵌入效率**（Embedding Efficiency）。嵌入效率 e 的意义是，平均每修改一个位置所能传输的隐蔽消息信息量，若信息量用比特表示，则计算公式为

$$e = \frac{\text{平均每载体样点承载的信息比特}}{\text{平均每载体样点被修改量}}\ bits/\text{次修改}$$

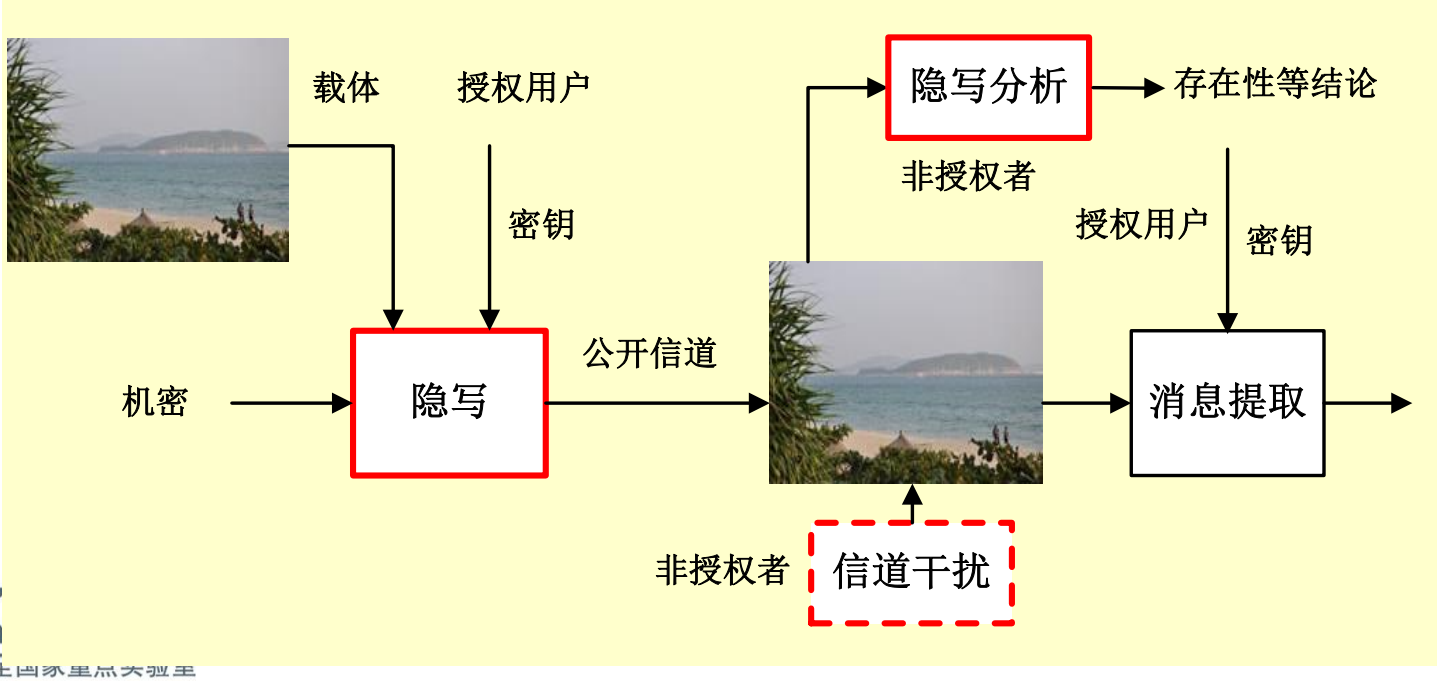
1-13 隐写的性能需求（2）

- **应用安全性**。指敌手难以从隐写应用协议与实现上发现有利于检测隐写媒体的方法；对应用协议，这里也分为敌手知晓或者不知协议设计的两种情况
- **计算效率**。指隐写的算法执行效率；实际任何方法都存在这个指标，但是，由于隐写**多用于不安全的物理环境**，这个指标显得更重要，是隐写方法的标志性指标之一
- **鲁棒性**。隐写信道存在有损和无损两种情况，在后者情况下，隐写载体面临有意或者无意的干扰（**如社交网络的转码**），隐写在这类条件下也需要有抗干扰的能力

- 本课程主要讲述理论、方法与应用相对成熟的隐写及其分析

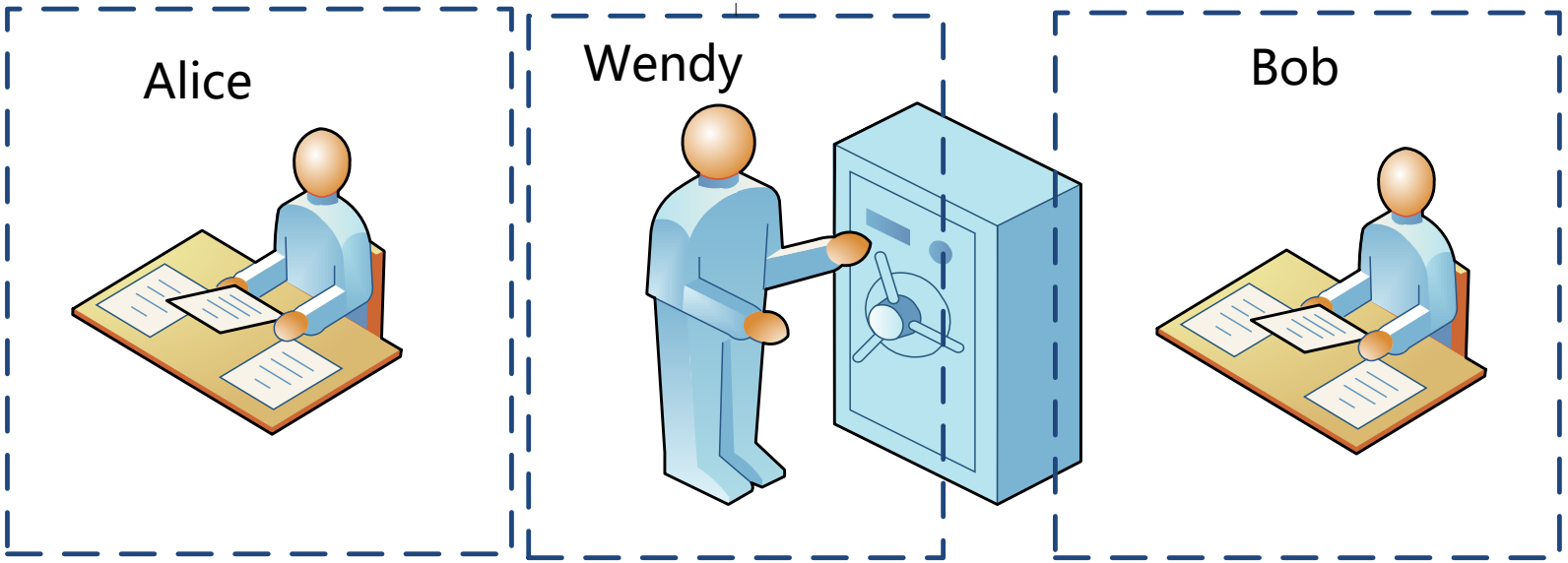
1-14 隐写与隐写分析对抗模型

- 隐写失败的标志是隐写事实的暴露
- 隐写分析（Steganalysis）泛指针对隐写的攻击，它主要通过检测隐写后载体特征的变化判定隐写的存在，也有少量隐写分析的技术目的还包括对隐写算法、参数的估计或者对隐藏信息的非授权提取等
- 隐写分析：被动攻击；信道干扰：主动攻击



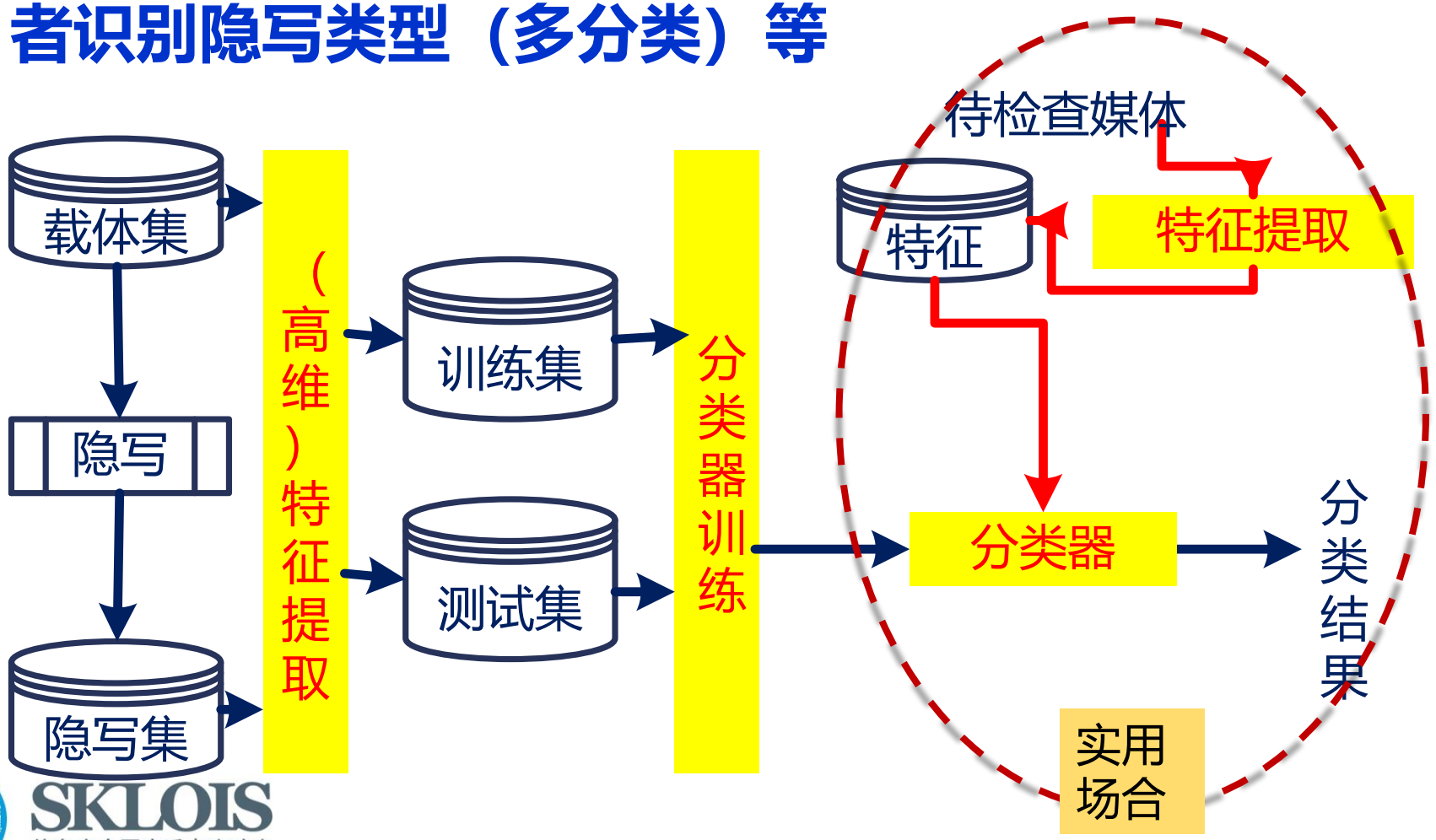
1-15 “囚犯问题”

- 1983年，Simmons提出了著名的“囚犯问题”
  - Alice与Bob是分别关押的犯人，Wendy是看押他们的狱监
  - Alice与Bob实际正在谋划共同越狱，他们之间难以见面，但却有权委托Wendy相互捎带消息，只能将机密消息隐藏在一般的文字描述中
  - Wendy不但可以仔细分析这些文字，也可能修改和伪造文字对犯人进行秘密考察。

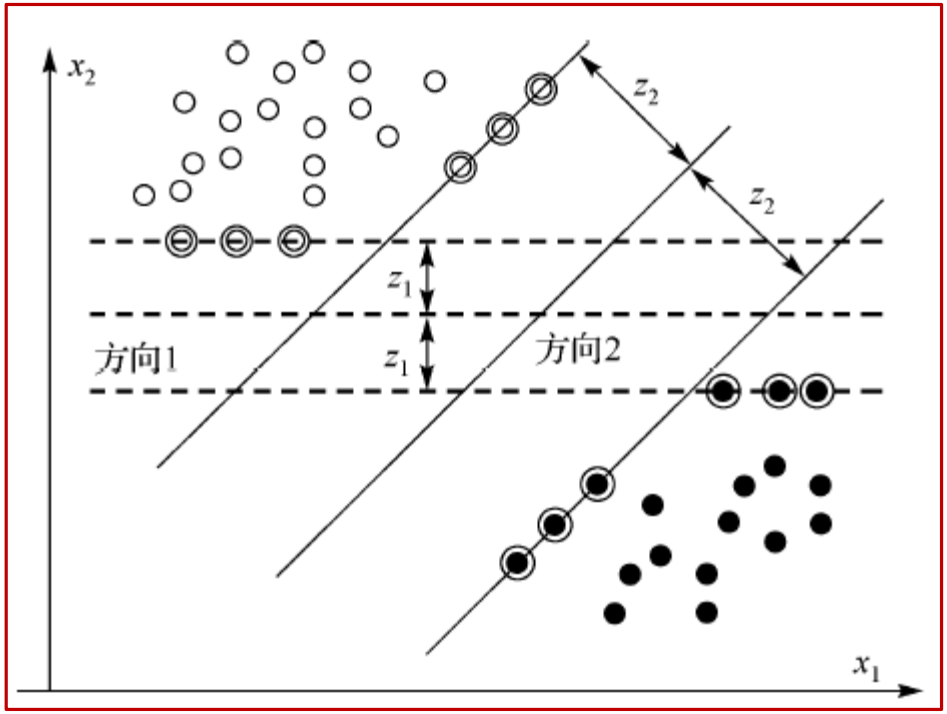


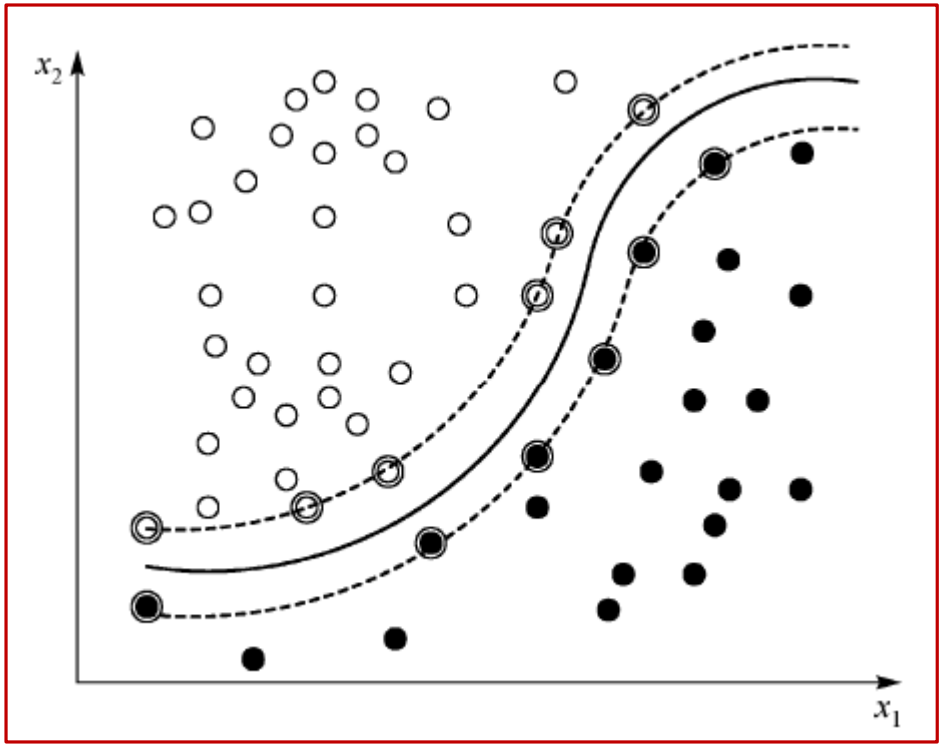
1-16 隐写分析一般过程

- 样本制作、训练，得到分类器
- 用分类器区分隐写媒体和自然媒体（二分类），或者识别隐写类型（多分类）等



特征空间中样本特征的分类示意





1-17 隐写分析的主要性能

- **漏检率（False-detection Rate）。将隐写媒体判断为自然媒体的比率；与之相对的概念是：**

$$\text{真阳性率或检测率} = 1 - \text{漏检率}$$

- **虚警率（False-alarm/positive Rate）。将自然媒体判断为隐写媒体的比率；与之相对的概念是：**

$$\text{真阴性率或不误报率} = 1 - \text{虚警率}$$

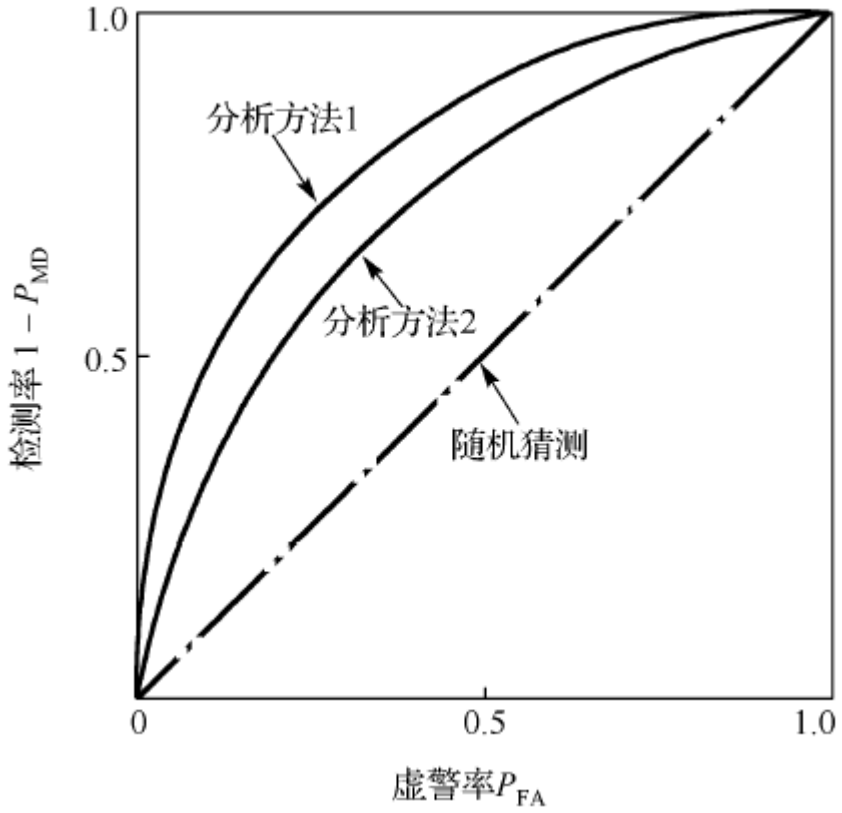
- **正确率（Accuracy Rate）。是隐写分析的主要技术指标，一般认为真阳性率与真阴性率同等重要，则可表示为**

$$\text{正确率} = 1 - \frac{\text{漏检率} + \text{虚警率}}{2} = \frac{\text{真阳性率} + \text{真阴性率}}{2}$$

$$\text{错误率} = \frac{\text{漏检率} + \text{虚警率}}{2} = 1 - \frac{\text{真阳性率} + \text{真阴性率}}{2}$$

检测性能的更确切表达：ROC曲线

- 接收操作特性（Receiver Operating Characteristic, ROC）曲线；一般通过调节分类器判决参数获得



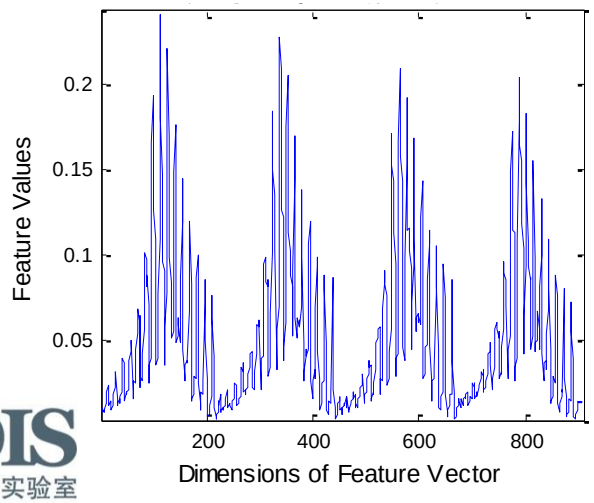
1-18 隐写对抗隐写分析的主要方法

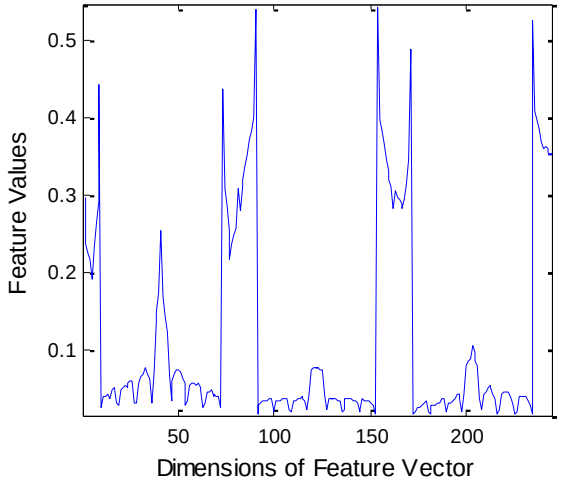
- 隐写对抗隐写分析的主要手段是，降低隐写对各类特征的扰动。当前，在保持一定负载率前提下，隐写主要可以通过以下方法不同程度地实现这一目标：
  - 特征保持。隐写的嵌入尽量保持载体的原有特征（非常困难的问题）
  - 降低修改次数。在一定负载率下，提高嵌入效率，减少对载体的修改次数
  - 降低修改扰动。降低对载体修改的信号幅度或者能量
  - 降低被检测代价。对风险进行定量描述，对载体的修改方式考虑了降低被隐写分析检测的风险
  - 提高应用方式的安全。在应用中考虑了隐写协议安全、抗关联分析等因素
  - 以上未考虑鲁棒性，鲁棒隐写当前研究很少（前期社交网少）

1-19 隐写分析对抗隐写的主要方法

- 隐写分析对抗隐写的主要手段是，发现与识别隐写对各类特征的扰动。当前，隐写分析主要可以通过以下方法不同程度地实现这一目标：
  - 有效提取隐写分析特征。发现与提取对隐写敏感的特征
  - 有效构造隐写特征识别系统。构造与训练能有效识别隐写分析特征的系统
  - 有效获得先验知识。先验知识指分析者知道的有关隐写者所采取算法与参数等的信息，它能帮助分析者更好地提取隐写分析特征并构造特征识别系统

DCT马氏特征





空域高阶小波特征

纲要

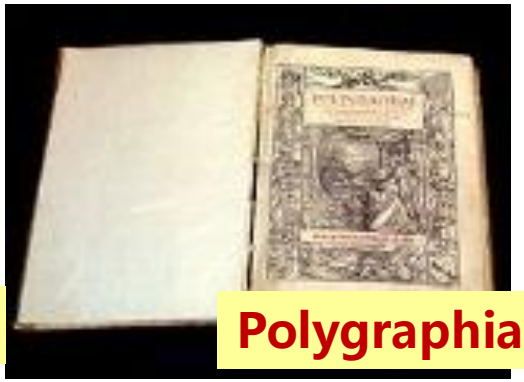
1. 从密码到信息隐藏与隐写
2. 隐写的发展与应用
3. 隐写安全指标
4. 文献阅读推荐

2-1 古代和近代的隐写通信（1）

- 隐写（Steganography）一词源自希腊语词根“στεγαν-ς”和“γραφ-ειυ”，意思是密写，这说明古代隐写是以文本隐藏为主的
- Herodotus(公元486-425)《Histories》：公元前440年的希腊，奴隶的头发被剃光，在头皮上面刺消息以策划反对波斯人的起义
- Trithemius(1462-1516)《Steganographice》：是信息隐藏乃至信息安全领域最早的专著，提出了在拉丁文、德文、意大利文和法文中隐藏文本

1499
1606

Steganographice



1518

Polygraphia

2-2 古代和近代的隐写通信（2）

- Schott（1608-1666）在《Schola Steganographica》：乐谱中隐藏消息；离合诗（acrostic），即用诗歌的特定位置表达隐藏的文字
- 离合诗的类似应用在我国古代记载也较多

Schola Steganographica

离合诗

卢花潭上有扁舟，俊杰黄昏独自游。

义到尽头原是命，反弓逃难必无忧。

施耐庵（1296-1370）《水浒传》第61回

- 在20世纪的两次世界大战和以后的间谍活动中，隐形墨水受到了广泛的应用（化学方法）

2-3 古代和近代的隐写通信（3）

- 北宋年间，曾公亮（999—1078）与丁度（990—1053）合著的《武经总要》反映了北宋军队对军令的伪装方法（算法）：
  - 原文：今约军中之事，略有四十余条，以一字为暗号；以旧诗四十字，不得令字重，每字依次配一条，与大将各收一本。如有报覆事，据字于寻常书状或文牒中书之
  - 大意：将全部40条军令编号并汇成码本，以40字诗对应位置上的文字代表相应编号；通信中，代表某编号的文字被隐藏在普通文书中，接收方知道它的位置，可以通过查找该字在40字诗中的位置获得编号，通过码本获得军令

按现在观点，综合了基于密码本的加密和文本同义词替换/语言学的信息隐藏（Linguistics Steganography）

2-4 应用情况——情报对抗

dark READING
Protect The Business Enable Access
Welcome Guest. | Log In | Register | Membersh
ATTACKS / BREACHES | VULNER
SECURITY MANAGEMENT | STORAGE

Tech Center: Ir

E-mail this page | Print this page | BOOKMARK

Busted Alleged Russian Spies Used Steganography To Conceal Communications

'Deep-cover' Russian intelligence agents hid electronic messages behind computer images

Jun 29, 2010 | 06:46 PM | 1 Comments

By Kelly Jackson Higgins

case that smacks of a Cold War spy novel, the FBI has arrested 11 pected Russian spies who for years had blended into day-to-day American life in the suburbs and cities. Aside from hiding their true identities and posing as legitimate American citizens, the suspects also

俄国间谍
隐写报道

USA TODAY
Home
News
Money
Sports
Life
Tech
Main Categories
Tech briefs
Web Guide
Tech Investor
Product reviews
More Tech

Tech

E-mail this story | Subscribe to the newspaper | Sign-up for e-mail news

02/05/2001 - Updated 05:17 PM ET

Terror groups hide behind Web encryption

By Jack Kelley, USA TODAY

WASHINGTON — Hidden in the X-rated pictures on several pornographic Web sites and the posted comments on sports chat rooms may lie the encrypted blueprints of the next terrorist attack against the United States or its allies. It

AP

U.S. officials say Osama bin Laden s posting instructions for terrorist ctivities on sports chat rooms, ornographic bulletin boards and ther Web sites.

ars technica
MAIN MENU MY STORIES: 1 FORUMS SUBSCRIBE JOBS

MINISTRY OF INNOVATION / BUSINESS OF TECH

Steganography: how al-Qaeda hid secret documents in a porn video

Digital steganography hides files in plain sight, concealed in image and media files.

by Sean Gallagher - May 2 2012, 8:02pm +0800

CYBERWAR IT PRIVACY 64

Photo illustration by Aurich Lawson

When a suspected al-Qaeda member was arrested in Berlin in May of 2011, he was found with a memory card with a password-protected folder—and the files within it were hidden. But, as the German newspaper *Die Zeit* reports, computer forensics experts from the German Federal Criminal Police (BKA) claim to have eventually uncovered its contents—what appeared to be a pornographic video called "KickAss."

Within that video, they discovered 141 separate text files, containing what officials claim are documents detailing al-Qaeda operations and plans for future operations—among them, three entitled "Future Works," "Lessons Learned," and "Report on Operations."

So just how does one store a terrorist's home study library in a pirated porn video file? In this case the files had been hidden (unencrypted) within the video file through a well-known approach for concealing messages in plain sight: steganography.

基地组织
隐写报道

SKLOIS
信息安全国家重点实验室

2-5 应用情况——犯罪组织

哥伦比亚贩毒头目J. C. R. Abadia使用图像隐写逃避网络追踪和审计，用于记录交易信息

贩毒组织隐写图像

图像送到美国检查

2-6 应用情况——恶意代码

Topic: Malware

Android malware makes steganography

Summary: Malware makers are turning to quite sophisticated tric of rogue applications.

By Adrian Kingsley-Hughes for Hardware 2.0 | January 30, 2012 -- Upd

用于隐藏代码（PNG图像隐写）
www.f-secure.com/weblog/archives/00002305.html

steganography to hide the control parameters for rogue code.

First, what is steganography? It's the technique of hiding messages case, an icon file.

F-Secure first suspected that Android malware was making use of st researchers came across this line of code:

```
localObject2 = ((ByteArrayOutputStream)localObject2).toByte
int k = paramInt + (-4 + new String(localObject2).indexOf("
if (k < 0)
  throw new IOException("Chank tEXt not found in png");
```

Image credit: F-Secure

Further digging revealed more code, and it soon became clear that t referenced here was the icon file bundled with the rogue application

```
| invoke-virtual (v5), [method]android.app.Activity.getAssets  #---proto:No
| move-result-object v0
| const-string v2, "icon.png"   image file name
| invoke-virtual (v0, v2), [method]android.content.res.AssetManager.open  #
| move-result-object v1
| iget-object v0, v5, [field]com.termate.MainActivity.d  #---type:Lcom/term
| const/4 v2, 1   method that checks for the tEXt chunk
| invoke-virtual (v0, v1, v2), [method]com.termate.a.a  #---proto:('descript
| if-eqz v1, off:0xff79
```

Image credit: F-Secure

ZDNet Join | Log In | Privacy | Cookies

virus BULLETIN Fighting malware and spam

Extend your in with a subscri

News | Resources | Magazine | VB100 | VBSpam | Conferenc

Home » News » Latest news » Alureon trojan uses steganography to receive co... cur

Alureon trojan uses steganography to receive commands

Messages hidden inside images create extra layer of redundancy.

archers at *Microsoft* have discovered a new variant of the 'Alureon' trojan that uses steganography to itself invincible against the takedown of botherders' domains.

用于隐蔽通信（位图图像隐写）
www.virusbtn.com/news/2011/09_26.xml

Steganography, sometimes referred to as 'hiding in plain sight', is the art and science of writing messages in such a way that no one but the intended recipient would even suspect that a message is present. Images are often used for this purpose: the sender uses an existing image and modifies the least significant bit(s) of the colour components of each pixel to contain the message. The difference between the old and the new image will be barely noticeable, but the intended recipient can easily extract the message from it.

Alureon (which also goes by the name of TDSS or TDL) is an oft-researched malware family that uses a number of advanced techniques to avoid detection and increase redundancy. Steganography is the latest such technique: the malware is capable of downloading innocent-looking images from free hosting sites. These images contain an updated configuration file and thus provide an extra layer of redundancy against the domains used by the malware becoming unavailable.

With malware researchers and law enforcement agencies becoming increasingly successful in taking down malicious domains and command and control centres used by botherders, the latter are constantly looking for new ways to control their bots. The use of steganography, as well as for instance the use of DNS TXT records by the Morto worm, show that malware researchers should keep their eyes wide open and may find control commands to be hidden in places where they might least expect them.

More on the *Microsoft Malware Protection Center* blog here, while here is a blog post by *Symantec* that shows another way in which crooks could use steganography.

26 September 2011

2-7 隐写的信息安全威胁（可下载1，2016）

1	Blindside	BMP	共享	www.blindside.co.uk
2	BMP Secrets	BMP	共享	www.pworlds.com/products/i_secrets.html
3	Camouflage	Windows格式文件	共享	www.camouflagesoftware.co.uk
4	StegMark	JPG,GIF,TIF, PNG,MIDI,WAV, AVI,MPEG	商业	www.datamark-tech.com
5	dc-Steganograph	PCX	共享	members.tripod.com/~Nikola_Injac /stegano/
6	Digital Picture Envelope	BMP	共享	www.know.comp.kyutech.ac.jp/BPCSe /Dpenve/DPENVe-home.html
7	Empty Pic	GIF	共享	www.crtelco.com/~robertw/
8	EZStego	GIF	共享	www.stego.com
9	F5	JPG	共享	wwwrn.inf.tu-dresden.de/~westfeld /f5.html

2-8 隐写的信息安全威胁（可下载2，2016）

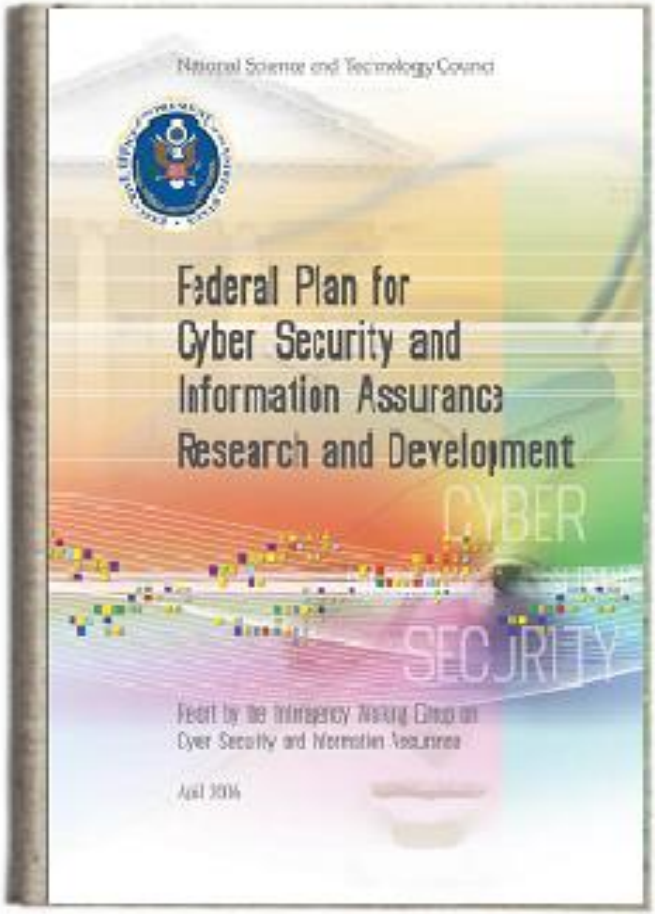
10	FFEncode	文本文件	共享	www.rugeley.demon.co.uk /security/ffencode.zip
11	Gif-It-Up	GIF	共享	crypto.radiusnet.net/archive /steganography/gif-it-up/
12	Gifshuffle	GIF	共享	www.darkside.com.au/gifshuffle/
13	Hide and Seek	BMP	共享	ftp://ftp.hacktic.nl/pub/crypto/ incoming/hideseek95.zip
14	Hide In Picture	BMP	共享	www.brasil.terravista.pt/Jenipabu /2571/e_hip.htm
15	Invisible Secrets	JPG, BMP, PNG	共享	www.innovatools.com/software /isecrets/
16	Invisible Secrets 3	JPG, PNG, BMP, HTML, WAV	付费	www.neobytesolutions.com/invsecr /index.htm
17	JP Hide and Seek	JPG	共享	linux01.gwdg.de/~alatham/stego.html
18	JSteg	JPG	共享	zooid.org/~paul/crypto/jsteg/

2-9 隐写的信息安全威胁（可下载3，2016）

19	MP3Stego	MP3	共享	www.cl.cam.ac.uk/~fapp2/steganography/mp3stego/
20	Steghide	BMP, WAV	共享	www.crosswinds.net/~shetzl/steghide/index.html
21	StegoWav	WAV	共享	www.geocities.com/SiliconValley/9210/stegowav.zip
22	Stella	GIF,BMP,JPG	共享	www.stella-steganography.de/
23	S-Tools	BMP,GIF,WAV,软盘	共享	ftp://ftp.ntua.gr/pub/crypt/mirrors/idea.sec.dsi.unimi.it/code/stools4.zip
24	TextHide	文本文件	商业	www.texthide.com
25	wbStego	BMP,TXT,HTML,XML,PDF	付费	wbstego.wbailer.com
26	轻松图片加密	BMP	共享	www.onlinedown.net/soft/30471.htm
27	渗透 3.0	FLASH SWF, BMP, JPG	共享	www.onlinedown.net/soft/4069.htm
28	BMP文件隐藏加密器	BMP	共享	www.onlinedown.net/soft/26335.htm

2-10 隐写的信息安全威胁（谁最关注）

- 2006年由美国科学技术委员会颁布的“网络安全与信息保障联邦研发计划” 中指出：The affordability and widespread availability of these tools makes steganography an enabling technology for U.S. adversaries. The threat posed by steganography has been documented in numerous intelligence reports.



https://www.nitrd.gov

2-11 隐写的信息安全威胁（谁最关注）

Jessica Fridrich is a research professor at the Center for Intelligent Systems at the State University of New York, Binghamton. her M.S. degree in om Czech Technical Czech Republic, and cience in 1995 from New York in Bing- earch interests are in hy and steganalysis, authentication and tamper detection, and forensic analysis of digital images. In the last six years, Fridrich's research has been steadily supported by the US Air Force in the form of 13 research grants total worth over US$1.3 million, generating five US and international patents.

Manuscript received February 28, 1998; revised October 23, 1998. This work was performed through collaborative participation in the Advanced Telecommunication/Information Distribution Research Program (ATIRP) Consortium sponsored by the U.S. Army Research Laboratory under Cooperative Agreement DAAL01-96-2-0002. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Naohisa Ohta.

L. M. Marvel and C. T. Retter are with the U.S. Army Research Laboratory, Aberdeen Proving Ground, MD 21005 USA (e-mail: marvel@arl.mil).

Organizers of IH2011 can be reached at:
organizers@ihconference.org

Lecture Notes in Computer Science
LNCS LNAI LNBI

Consider visiting IEEE ICASSP 2011, May 22-27 after IH.

ICASSP

- Neil F. Johnson, Booz Allen Hamilton and JJTC, U
- Stefan Katzenbeisser, TU Darmstadt, Germany
- Darko Kirovski, Microsoft Research, USA
- John McHugh, University of North Carolina, USA ar
- Ira S. Moskowitz, Naval Research Laboratory, USA
- Ahmad-Reza Sadeghi, Ruhr-Universität Bochum, Ge
- Rei Safavi-Naini, University of Calgary, Canada
- Phil Sallee, Booz Allen Hamilton, USA
- Berry Schoenmakers, TU Eindhoven, The Netherlar
- Kaushal Solanki, Mayachitra Inc., USA
- Kenneth Sullivan, Mayachitra Inc., USA
- Paul Syverson, Naval Research Laboratory, USA

学术界与政府、军队背景的研究机构均展开了相关研究

SKLOIS 信息安全国家重点实验室

2-12 隐写被动防范（国内外隐写防范产品）

除了前述的研究投资，商业产品也开始出现

- 国外：WetStone的Stego Suite：Stego Hunter、Stego Watch、Stego Analyst、Stego Break；Backbone Security的StegAlyzerSS（隐写分析）与StegAlyzerRTS（网关）；共享：StegDetect、Stegospy； 美国政府和军队；病毒扫描软件
- 国内：我国科研机构研发了部分分析软件，当前也开始注重网络环境下的分析

WetStone研制的Stego Suite隐写检查软件

Only Bit Defender and Symantec alerted on the JPEGS. Bit Defender found Trojan.HideFrog.A in all three (they are images of a frog)

Symantec alerted as follows:
NT1.JPG W32.Looksky!gen
NT2.JPG Trojan.Desktophijack.B
NT3.JPG Trojan.Jupillites

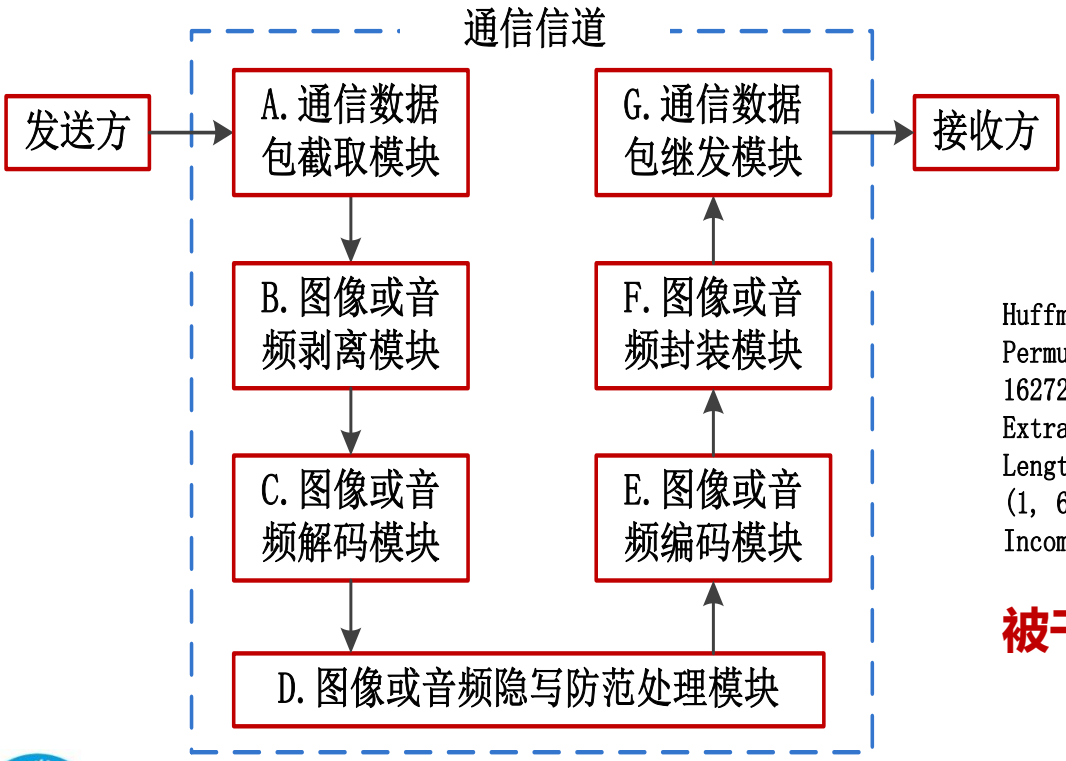
病毒检查中的隐写分析

www.pcreview.co.uk/forums/steganography-sample-malware-t2979331.html

SKLOIS 信息安全国家重点实验室

2-13 隐写主动防范——媒体隐写干扰

- 对隐写的主动防范（专利ZL201010230477.6）。基于图像和音频再编码处理，以极低的代价使隐写通信失效




```
Huffman decoding starts
Permutation starts
1627200 indices shuffled
Extraction starts
Length of embedded file: 996317 bytes
(1, 67108863, -6) code used
Incomplete file: only 0 of 996317 bytes extracted
```

被干扰后，F5隐写软件的出错输出

2-14 隐写主动防范——网络通信数据隐写干扰

对网络数据包的重组发送，获得“抗逃避能力”

时序型网络隐写依然可穿透

纲要

1. 从密码到信息隐藏与隐写
2. 隐写的发展与应用
3. 隐写安全指标
4. 文献阅读推荐

3-1 基于分布偏差的隐写安全指标（1）

- Cachin基于信息论模型提出了隐写的理论安全性描述
- 由于隐写安全性主要是指隐写后媒体特征变化的隐蔽性，在理论上可以认为，载体分布的变化程度可以描述这类隐蔽性
- 载体与隐写后载体的分布分别是 $P_C(x)$ 与 $P_S(x)$，则用以下KL偏差（Kullback–Leibler Divergence）描述分布之间的差别

$$D(P_C||P_S)=\sum_{x\in X}P_C(x)\log_2\frac{P_C(x)}{P_S(x)}=\sum_{x\in X}P_C(x)\left(\log_2(P_C(x))-\log_2(P_S(x))\right)$$

- 当 $D(P_C||P_S)\le\varepsilon$，称隐写方法针对被动攻击者（Passive Adversary）是 $\varepsilon-$安全的（$\varepsilon-$secure）；当 $D(P_C||P_S)=0$，称隐写方法是完美安全的（Perfectly Secure）

- 由于载体维度极高，即使采用一个很大数据集，获得可靠载体分布也是困难的，因此，载体KL偏差是难以计算的

3-2 基于特征投影偏差的隐写安全指标（2）

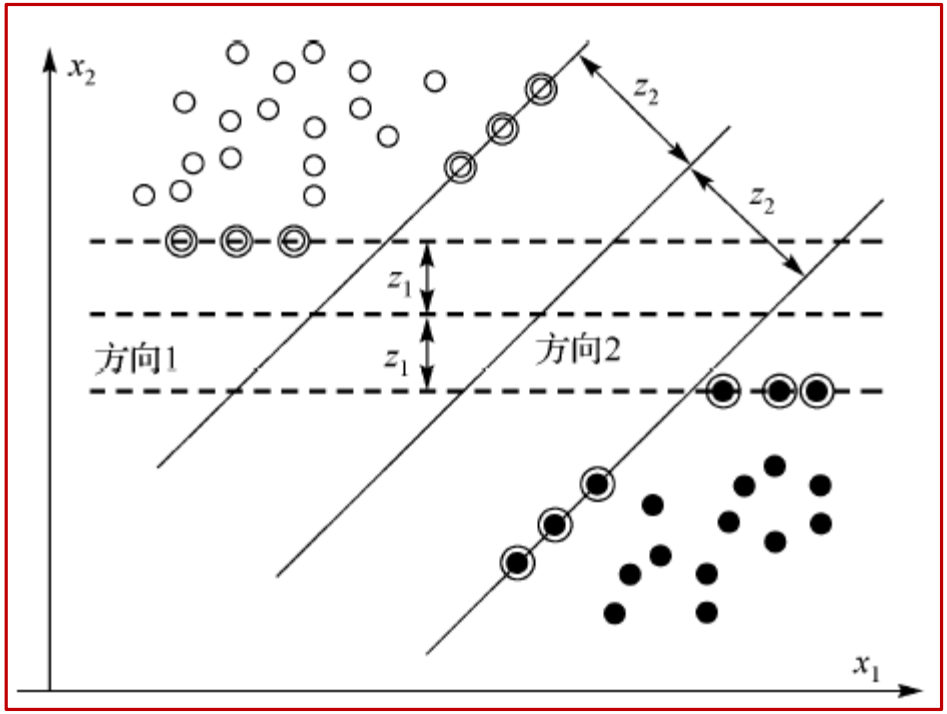
- 为了得到一个可计算的隐写安全指标值，T. Pevny等人提出基于选定的一组**特征偏差**衡量隐写安全
- 采用最大平均偏差（MMD，Maximum Mean Discrepancy）来衡量载体特征与隐文特征在分布上的差异。设F为一组函数集合，代表投影与变换；$x\sim p$、$y\sim q$ 为**特征**及其分布，则MMD的定义是：

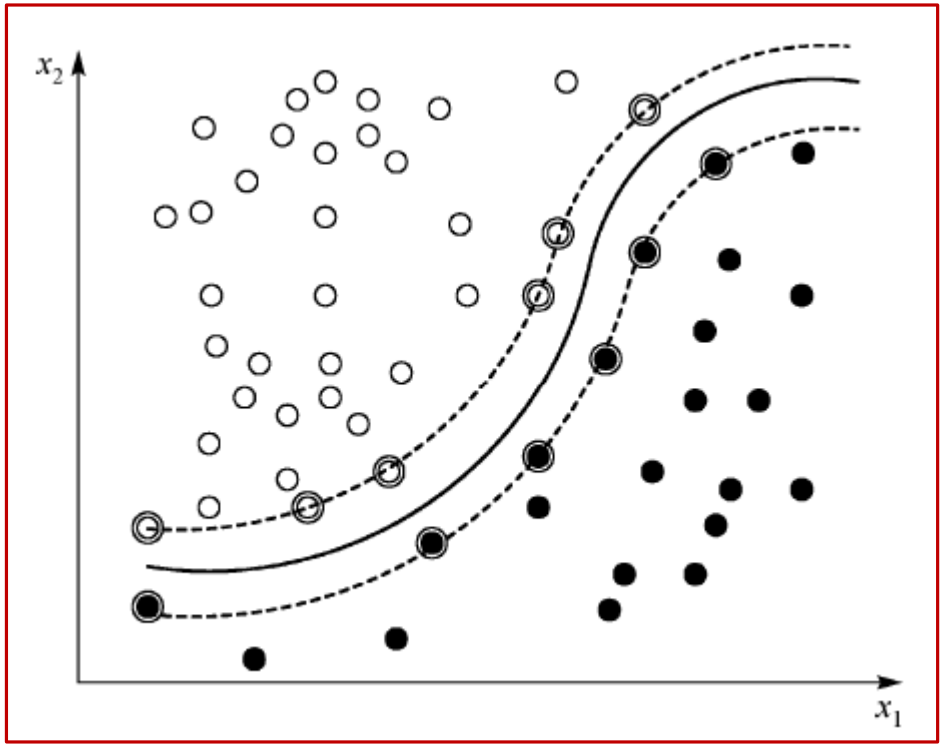
$$\mathrm{MMD}(\mathsf{F},p,q)=\sup_{f\in\mathsf{F}}\left(\mathrm{E}_{x\sim p}\left(f(\boldsymbol{x})\right)-\mathrm{E}_{y\sim q}\left(f(\boldsymbol{y})\right)\right)$$

$$=\sup_{f\in\mathsf{F}}\left(\frac{1}{D}\sum_{i=1}^{D}f(\boldsymbol{x}_i)-\frac{1}{D}\sum_{i=1}^{D}f(\boldsymbol{y}_i)\right)$$，假设出现概率相同

- MMD考察的是，在特定投影或变换下，阴阳样本集合上一组特征2组值(阴阳各一组)的最大平均偏差

选择特征投影偏差的原因





3-3 基于抗隐写分析性能的指标（工程实用）

- **正确率**（Accuracy Rate）。是隐写分析的主要技术指标，一般认为真阳性率与真阴性率同等重要，则可表示为

$$\text{正确率} = 1 - \frac{\text{漏检率} + \text{虚警率}}{2} = \frac{\text{真阳性率} + \text{真阴性率}}{2}$$

- 由于隐写者的敌手是隐写分析者，因此，基于抗隐写分析性能也能衡量隐写的安全。显然，抗隐写分析性能可以直接由隐写分析的正确率来表示
- 在有关信息隐藏安全性的描述上，尤其在有关隐写安全的描述上，虽然人们认为非授权方是不知密钥的，但往往存在假设Kerckhoffs准则与不假设该准则的两种情况
  - 验证模型下的安全性（遵照Kerckhoffs准则）：假设敌手知道，在此情况下安全则更满足需求；但是不一定与实际情况相符
  - 盲检测下的安全性（不遵照或者部分遵照Kerckhoffs准则）

4 文献阅读推荐

- 教材第1章

- 推荐参考书中相关内容
  - J. Fridrich（美）著， 张涛、奚玲、张彦、许漫坤 译，数字媒体中的隐写术——原理、算法和应用，北京：国防工业出版社，2014年4月

谢谢！

中国科学院信息工程研究所
INSITUTE OF INFORMATION ENGNEERING CAS

SKLOIS
信息安全国家重点实验室