# Fingerprint

北京大学 软件与微电子学院
School of Software and Microelectronics, Peking University

*Huiping Sun(孙惠平)*
*sunhp@ss.pku.edu.cn*

Course Overview

# 本次课程内容

## 1 简介 基本原理

- 如何工作
- 概念定义
- 交易构成
- 交易形式

## 2 核心客户端 密钥和地址

- 核心架构
- 地址公私钥
- Base58Check
- 靓号纸钱包

## 3 钱包 交易

- 分类
- 确定性？
- 助记词
- 费用，脚本

## 4 网络 + 区块链

- 节点类型
- 中继网络
- 协议
- 梅克尔树

# 参考资料

CONTRIBUTED PAPER

## A Survey on Web Tracking: Mechanisms, Implications, and Defenses

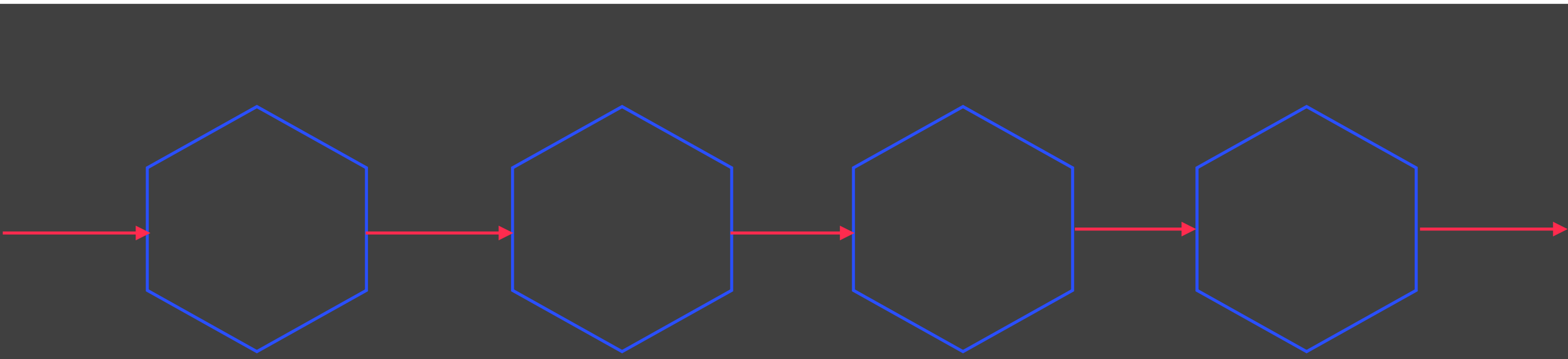
## Browser Fingerprinting: A survey

PIERRE LAPERDRIX, CNRS, Univ Lille, Inria Lille, France
NATALIIA BIELOVA, Inria Sophia Antipolis, France
BENOIT BAUDRY, KTH Royal Institute of Technology, Sweden
GILDAS AVOINE, Univ Rennes, INSA Rennes, CNRS, IRISA, France

2016 IEEE Symposium on Security and Privacy

## Beauty and the Beast: Diverting modern web browsers to build unique browser fingerprints

# Web跟踪



# Web跟踪技术

**基于Session**

- Web认证
- SessionID
- DOM

**基于存储**

- HTTP Cookie
- Flash Cookie
- Flash LCO
- IE userData
- Silverlight
- HTML5
- Web SQL

**基于缓存**

- Web Cache
- DNS Cache
- 操作Cache

**指纹**

- 网路
- 设备
- OS
- 浏览器

**其余方法**

- 电话
- 时间
- SuperCookie

# Web跟踪目的

广告

第三方跟踪

精准定价

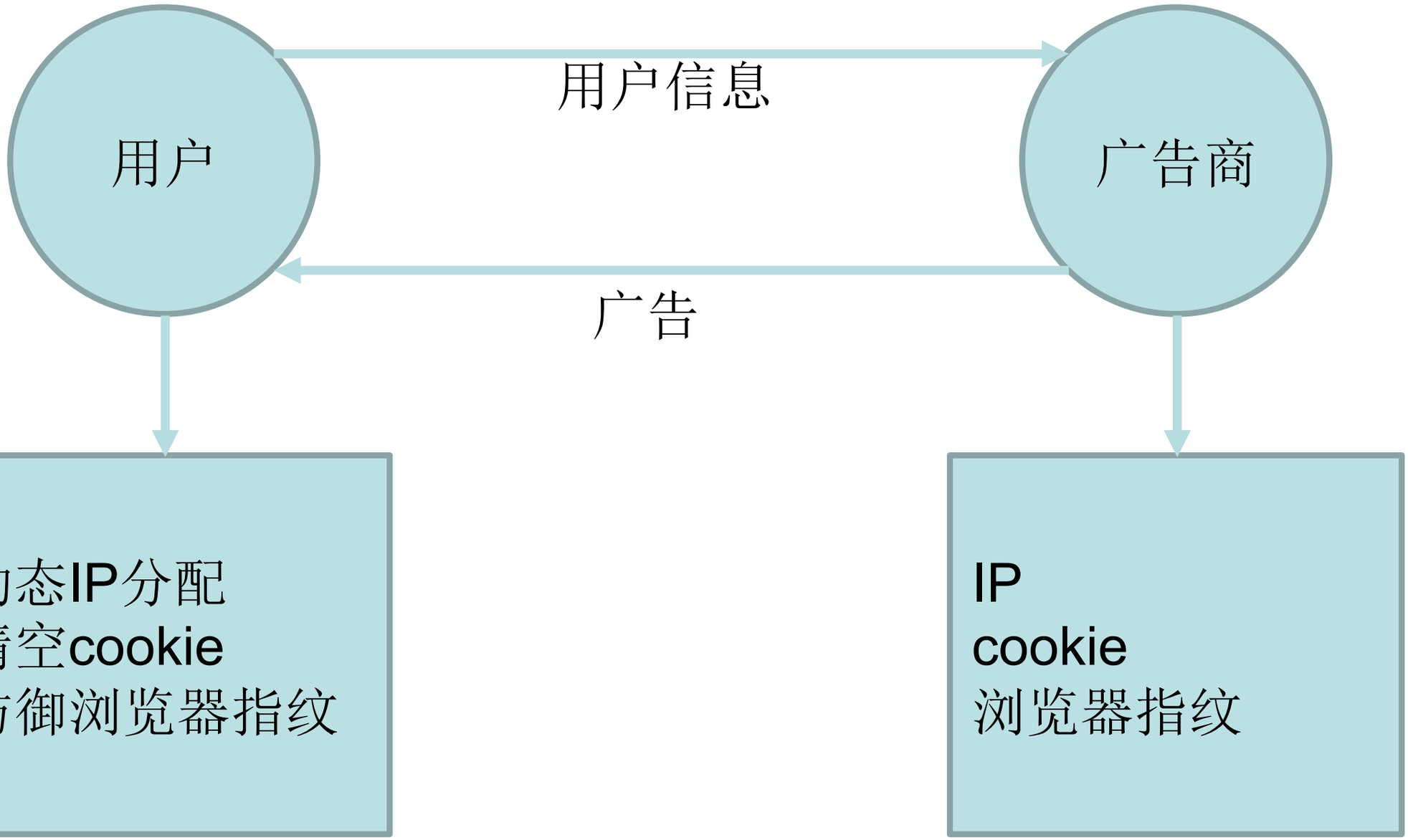
风险控制

个性化服务

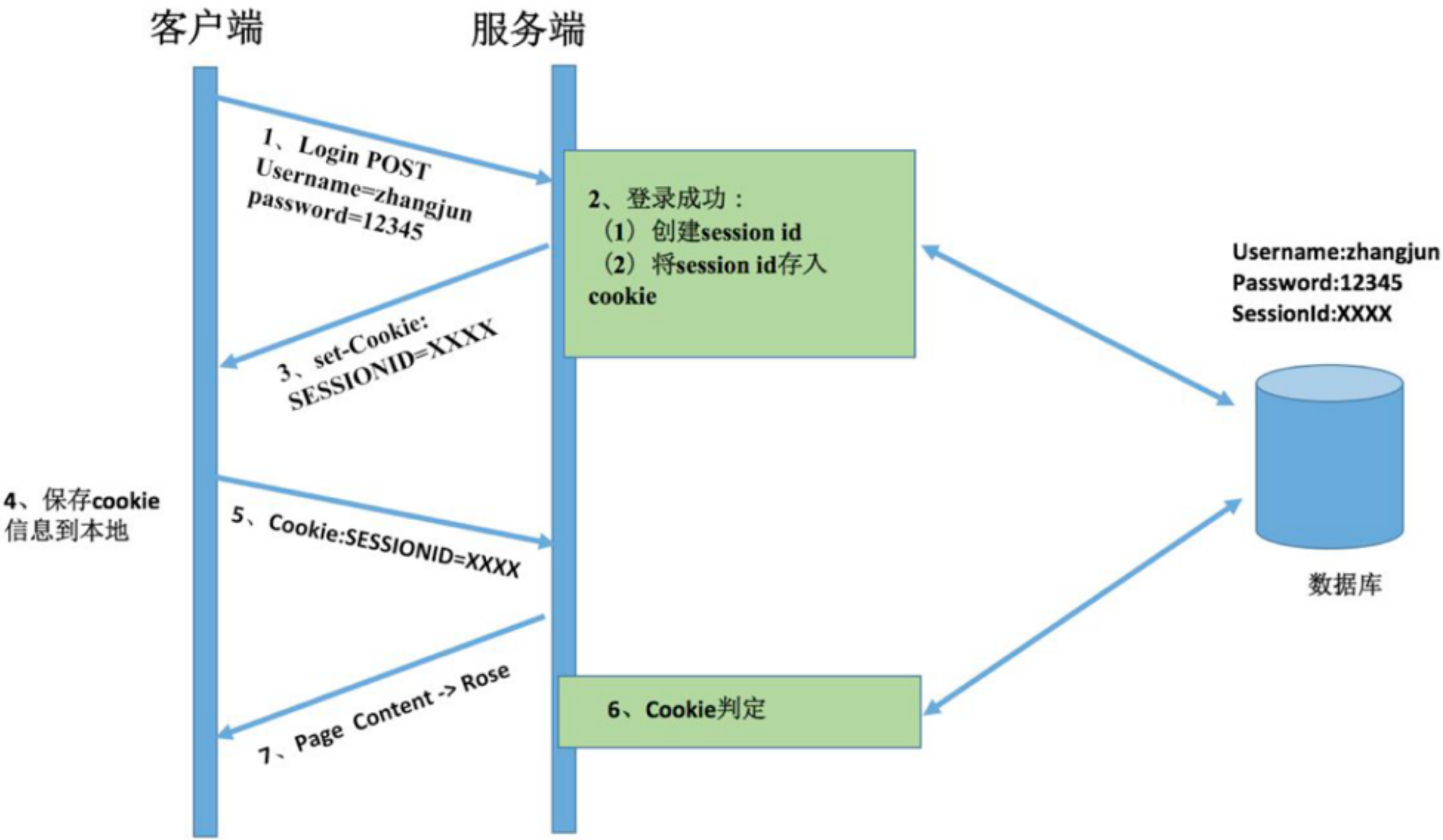
身份欺骗

撸羊毛

# 广告

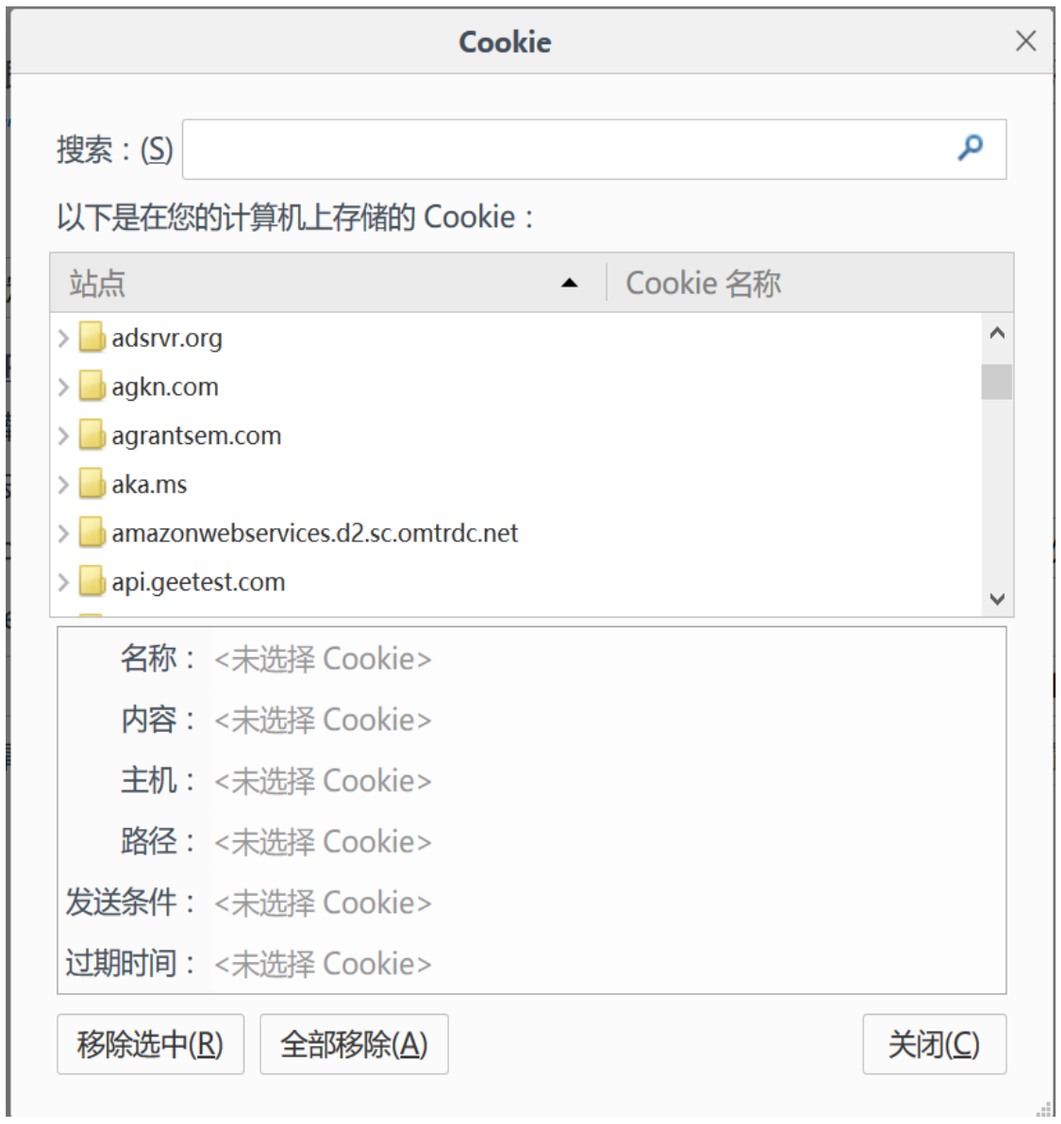


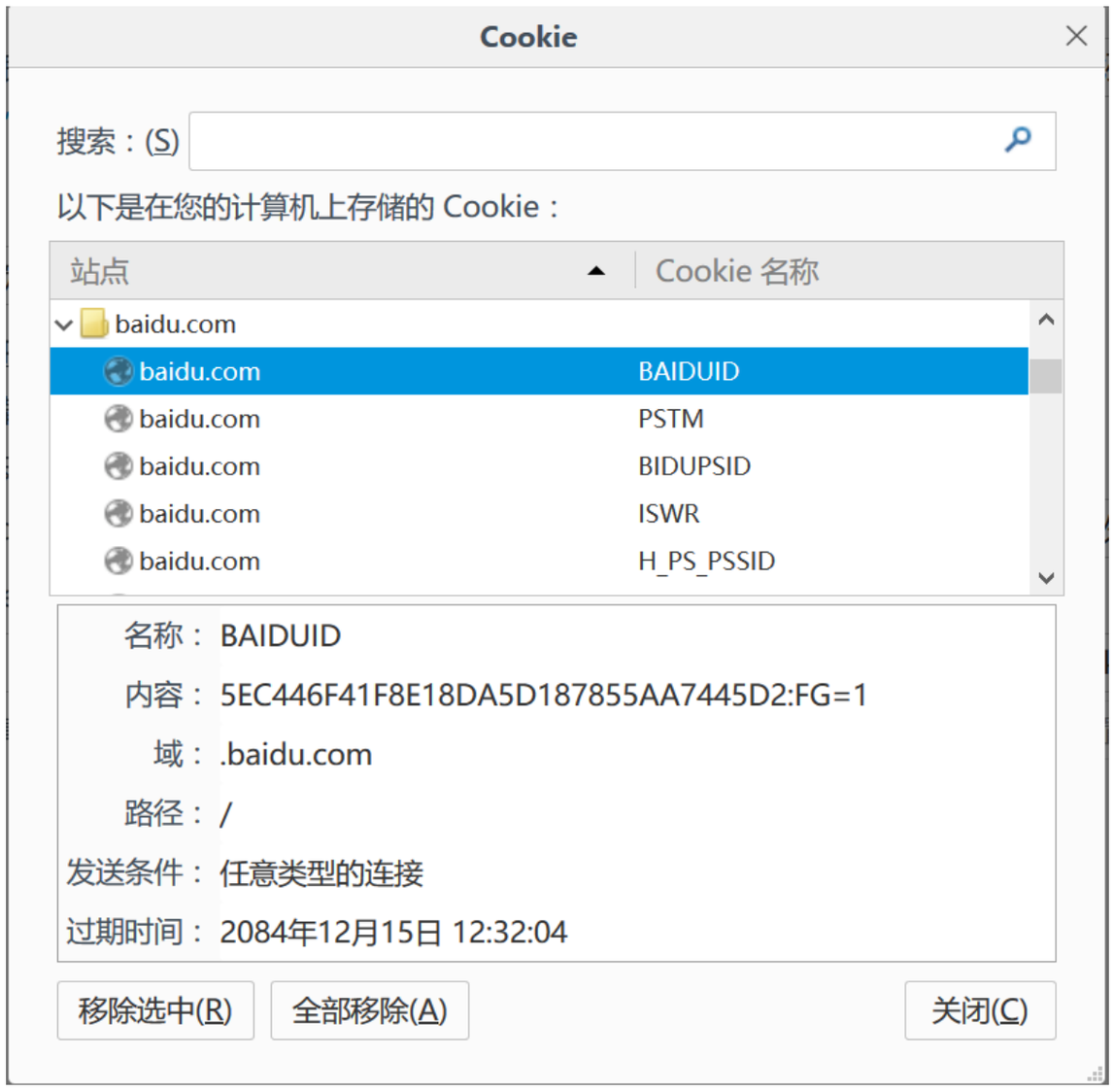
# Cookie


客户端
服务端
1、Login POST
Username=zhangjun
password=12345
2、登录成功：
（1）创建session id
（2）将session id存入
cookie
3、set-Cookie:
SESSIONID=XXXX
Username:zhangjun
Password:12345
SessionId:XXXX
数据库
4、保存cookie
信息到本地
5、Cookie:SESSIONID=XXXX
6、Cookie判定
7、Page Content -> Rose

# Fingerprint

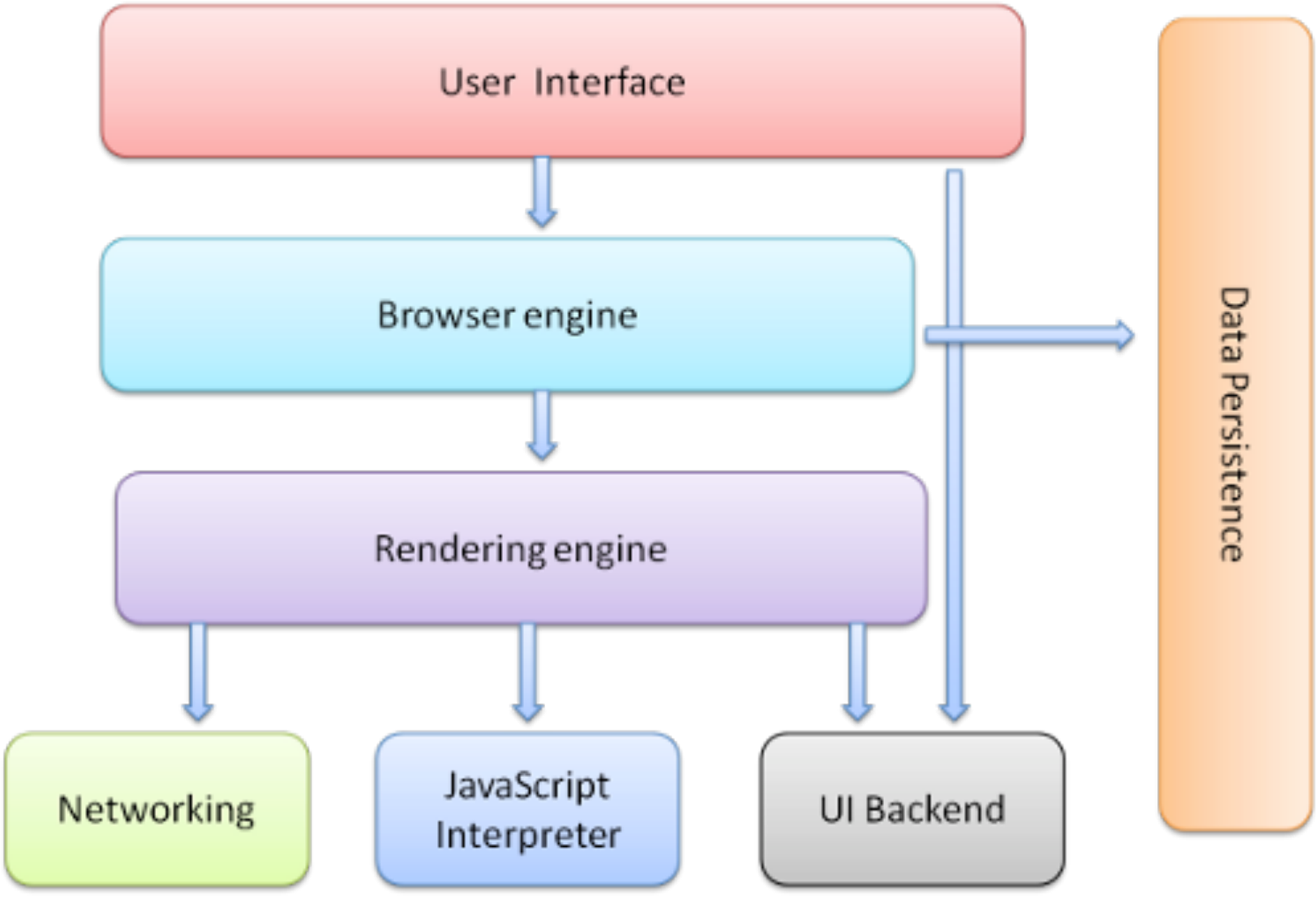
# Cookie



# 浏览器指纹

# 浏览器


User Interface
Browser engine
Rendering engine
Networking
JavaScript Interpreter
UI Backend
Data Persistence

# 浏览器指纹

**浏览器指纹**

一系列和用户设备相关的信息集合，包括硬件设备、操作系统、浏览器、相关配置等

**浏览器指纹技术**

从浏览器指纹收集信息建造设备指纹的过程

浏览器用户产生头文件

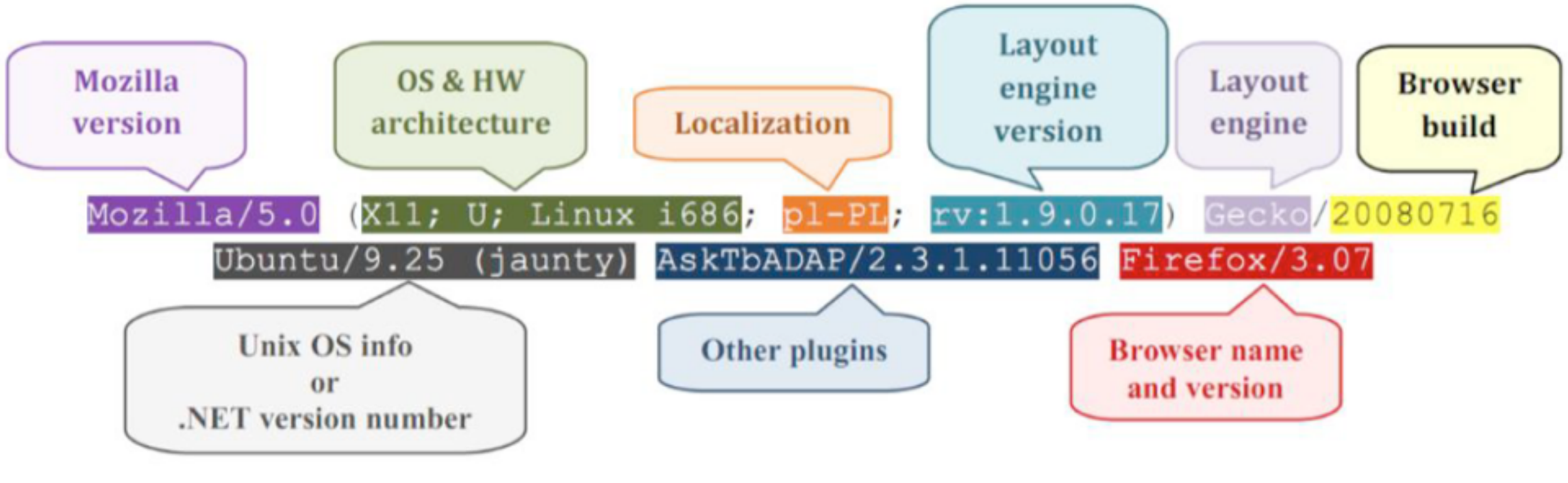
浏览器JS脚本

浏览器API

浏览器指纹新基础

*Fingerprint*

# 浏览器指纹




- 更互动的网络（例如，*JavaScript*库的繁荣，*HTML5*的每周创新）
- 更可用的网络（例如，移动设备的爆炸）
- 更安全的网络（例如，*Flash*正在消失，*NPAPI*插件正在被弃用）
- 更私人的网络（例如，增加立法反对*cookies*，扩展的巨大成功，如*Ghostery*和*AdBlock*）。

# 浏览器指纹分类

| 指纹特征 | 获取方法 | 描述 | 变化情况 |
| --- | --- | --- | --- |
| User Agent | Http、JS | 包含当前系统及浏览器的版本等相关信息 | 浏览器升级等都会对其有影响，容易发生变化 |
| 浏览器插件 | JS | 浏览器安装的插件信息 | 插件的安装、删除、升级都会使其改变，容易发生变化 |
| 字体（fonts） | FLASH、JS | 系统中安装的字体的信息 | 安装和卸载字体会使其改变，一般不会变化 |
| 时区（timezone） | JS | 时区信息 | 极少变化 |
| Cookie | JS | Cookie是否启动 | 变化较小 |
| Local Storage | JS | Local Storage是否启动 | 用户手动开启或关闭时会使其改变，一般很少变化 |
| 可接受信息（Http Accept） | JS | 包含在HTTP的头信息中 | 较少变化 |

Fingerprint

# 浏览器指纹属性

| | | | |
|---|---|---|---|
| 1 | UserAgent | 提供当前浏览器和所在系统的相关信息 | Mozilla/5.0 (Macintosh; Intel Mac OS X 10_10_5) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/56.0.2924.87 Safari/537.36 |
| 2 | 屏幕分辨率（Screen Resolution） | 屏幕的宽度和高度，以像素计 | 1280,800 |
| 3 | 可用屏幕分辨率（Available Screen Resolution） | 屏幕的可用宽度和高度，以像素计。 | 1280,773 |
| 4 | 平台（Platform） | 一个只读的字符串，说明当前的操作系统以及硬件平台的信息。 | MacIntel |
| 5 | 语言（Language） | 当前浏览器的语言 | zh-CN |
| 6 | 时区（Timezone） | 系统设置的地点所在的时区偏移值 | −480 |
| 7 | 色深度（Color Depth） | 缓冲器或者目标设备上的调色板的比特深度 | 24 |
| 8 | 色素率（Pixel Ratio） | 设备物理像素和设备独立像素的比例 | 1.7999999523162842 |
| 9 | 会话存储（Session Storage） | 当前浏览器是否支持Session Storage | 1 |
| 10 | 本地存储（Local Storage） | 当前浏览器是否支持Local Storage | 1 |
| 11 | 索引数据库（Indexed DB） | 当前设备是否支持索引数据库 | 1 |
| 12 | 开放数据库(Open DB) | 当前浏览器是否支持开放数据库 | 1 |
| 13 | CPU 种类（CPU class） | 浏览器所在的计算机系统使用的CPU类型 | Win32 |
| 14 | 禁止追踪（Do Not Track） | 表示用户是否允许网站、广告等的追踪，返回用户设置的值。 | 1 |
| 15 | 插件（Plugins） | 当前浏览器安装的插件的信息 | Widevine Content Decryption Module::Enables Widevine licenses for playback of HTML audio/video content. (version: 1.4.8.962)::application/… |
| 16 | Canvas | 基于HTML的canvas获取的浏览器的指纹信息 | Canvaswinding:yes~canvas fp:… |
| 18 | 广告屏蔽插件（Adblock） | 设备是否安装了广告屏蔽插件 | false |
| 19 | 修改语言（has_lied_language） | 用户是否修改了当前浏览器或者系统语言 | false |
| 20 | 修改分辨率（has_lied_resolution） | 屏幕分辨率是否被修改了 | false |
| 21 | 修改操作系统（has_lied_os） | 用户是否修改了当前操作系统的类型和版本 | false |
| 22 | 修改浏览器（has_lied_browser） | 用户是否修改了当前浏览器的类型和版本 | false |
| 23 | 支持触感（Touch Support） | 当前设备设备是否支持触感，<br>主要针对手持设备 | 0,false,false |

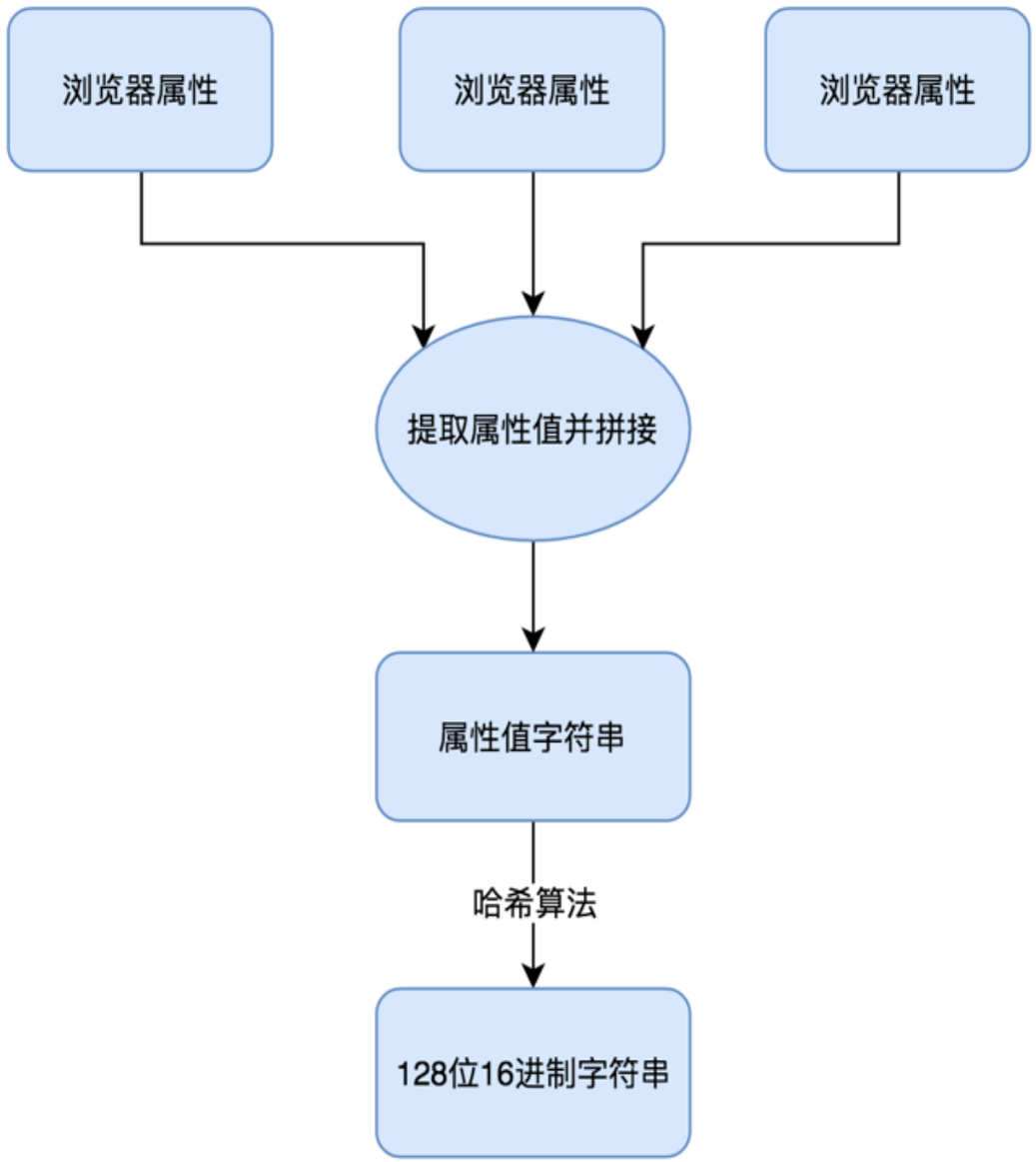
# 浏览器指纹例子

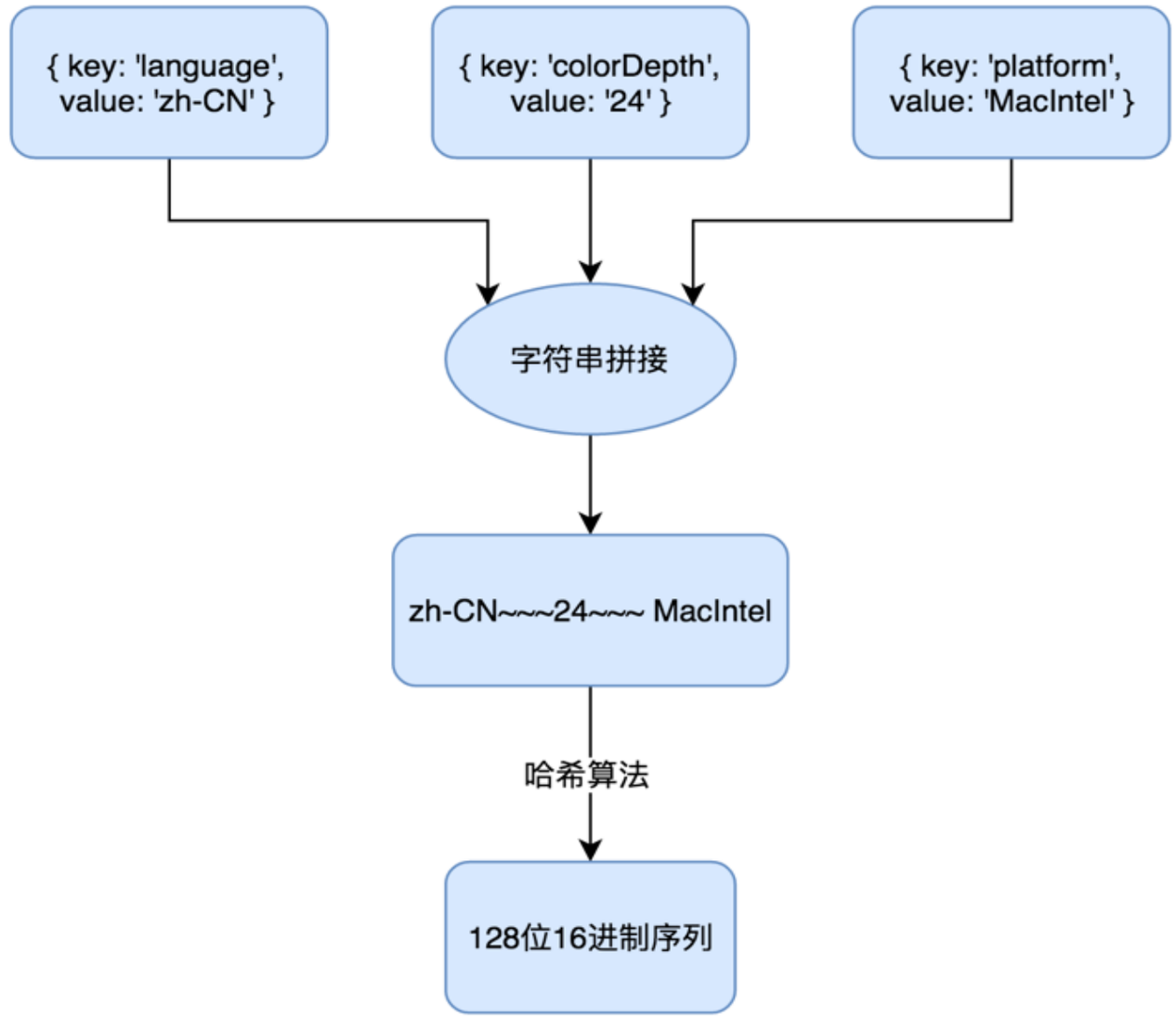
| Attribute | Source | Example |
|---|---|---|
| User agent | HTTP header | Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/64.0.3282.119 Safari/537.36 |
| Accept | HTTP header | text/html,application/xhtml+xml,application/xml;q=0.9,image/webp,image/apng,*/*;q=0.8 |
| Content encoding | HTTP header | gzip, deflate, br |
| Content language | HTTP header | en-US,en;q=0.9 |
| List of plugins | JavaScript | Plugin 1: Chrome PDF Plugin. Plugin 2: Chrome PDF Viewer. Plugin 3: Native Client. Plugin 4: Shockwave Flash... |
| Cookies enabled | JavaScript | yes |
| Use of local/session storage | JavaScript | yes |
| Timezone | JavaScript | -60 (UTC+1) |
| Screen resolution and color depth | JavaScript | 1920x1200x24 |
| List of fonts | Flash or JS | Abyssinica SIL,Aharoni CLM,AR PL UMing CN,AR PL UMing HK,AR PL UMing TW... |
| List of HTTP headers | HTTP headers | Referer X-Forwarded-For Connection Accept Cookie Accept-Language Accept-Encoding User-Agent Host |
| Platform | JavaScript | Linux x86_64 |
| Do Not Track | JavaScript | yes |
| Canvas | JavaScript | Cwm fjordbank glyphs vext quiz, 😃<br>Cwm fjordbank glyphs vext quiz, 😃 |
| WebGL Vendor | JavaScript | NVIDIA Corporation |
| WebGL Renderer | JavaScript | GeForce GTX 650 Ti/PCIe/SSE2 |
| Use of an ad blocker | JavaScript | yes |

# 指纹技术

Fingerprint

# 浏览器指纹提取

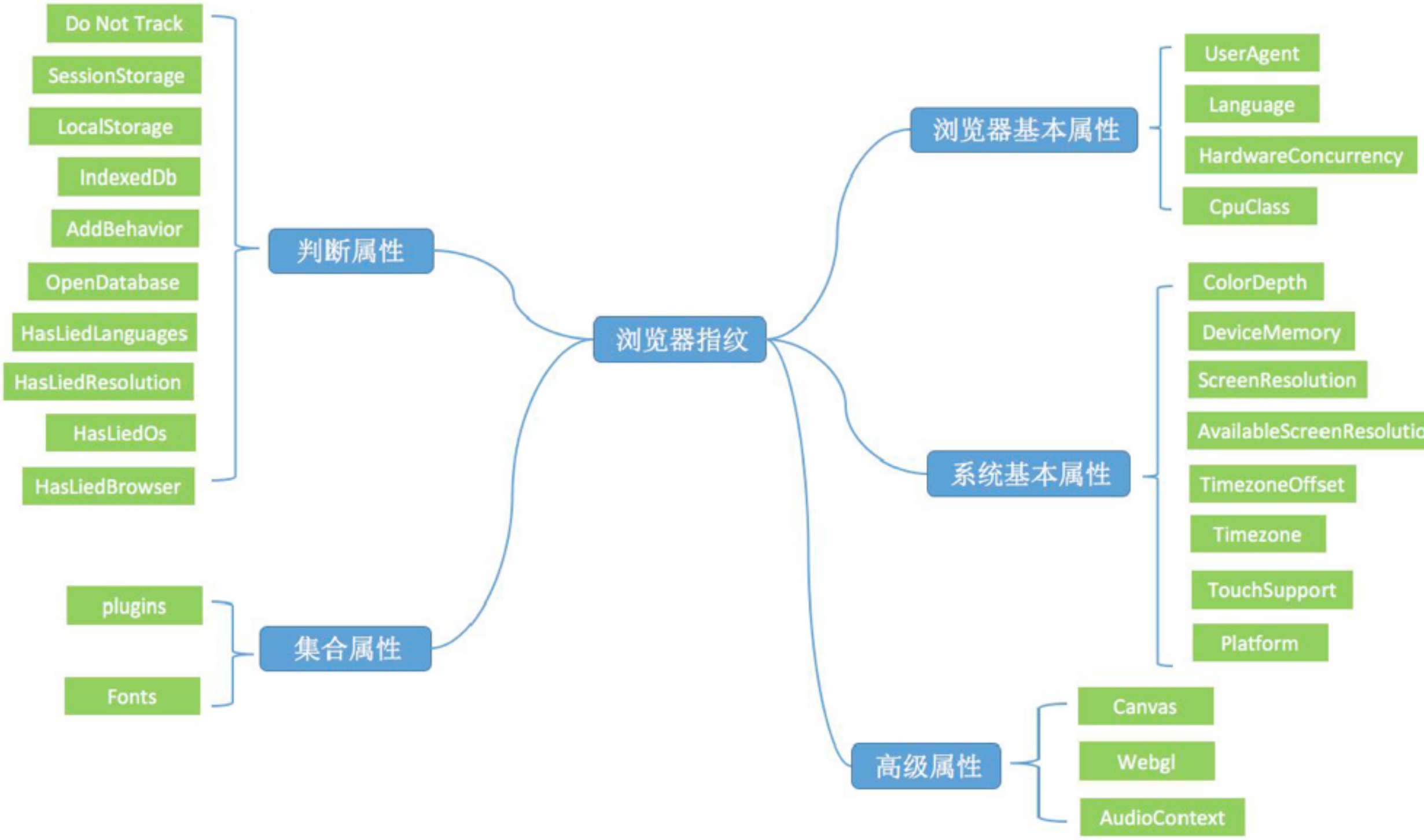


# 浏览器属性分类

| | |
|---|---|
| 可以通过浏览器提供的接口或对象直接获取值的属性 | User Agent、屏幕分辨率、可用屏幕分辨率、平台、语言、时区、色深度、色素率、会话存储、本地存储、索引数据库、开放数据库、CPU种类、触感支持，禁止追踪、插件 |
| 需要通过JavaScript做一些判断，才能确定值的属性 | 是否安装了广告屏蔽插件，用户是否修改了语言、用户是否修改了分辨率、用户是否修改了操作系统、用户是否修改了浏览器 |
| 通过Canvas、WebGL获取到的浏览器属性 | Canvas、WebG |

# 浏览器属性分类


浏览器指纹
判断属性
Do Not Track
SessionStorage
LocalStorage
IndexedDb
AddBehavior
OpenDatabase
HasLiedLanguages
HasLiedResolution
HasLiedOs
HasLiedBrowser
集合属性
plugins
Fonts
浏览器基本属性
UserAgent
Language
HardwareConcurrency
CpuClass
系统基本属性
ColorDepth
DeviceMemory
ScreenResolution
AvailableScreenResolution
TimezoneOffset
Timezone
TouchSupport
Platform
高级属性
Canvas
Webgl
AudioContext

# HTML 5 Canvas

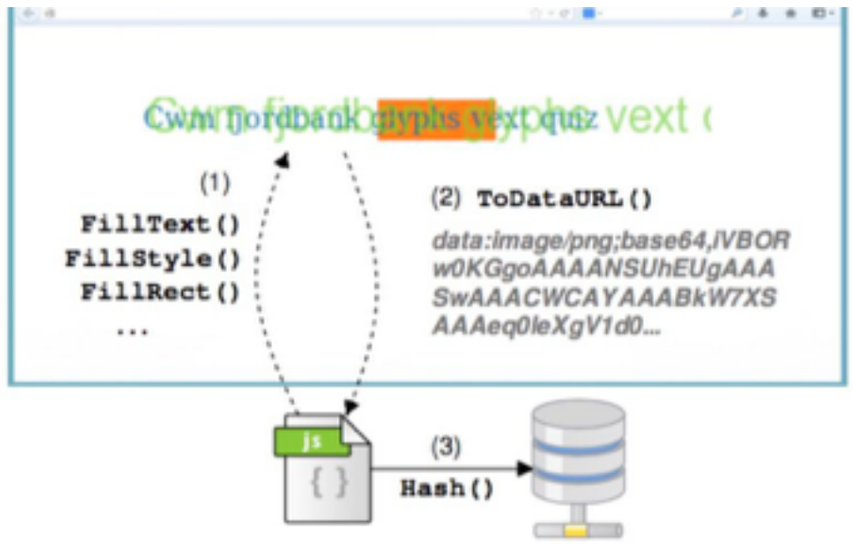
- 当用户访问一个包含*canvas*的指纹脚本，他会被要求绘制一个隐蔽的图形。不同浏览器会有不同的图像处理引擎、导出选项、压缩等级。相同的*JavaScript*代码在不同的平台上执行最终的结果会不同（*OS*，*OS Version*，*Browser*，*Browser Version*，*GPU*）。由于，绘制出的图形会有差别，根据这些差别为用户分配唯一的编号（指纹）。

```
<canvas id="myCanvas"></canvas>

<script type="text/javascript">

var canvas=document.getElementById('myCanvas');
var ctx=canvas.getContext('2d');
ctx.fillStyle='#FF0000';
ctx.fillRect(0,0,80,100);

</script>
```



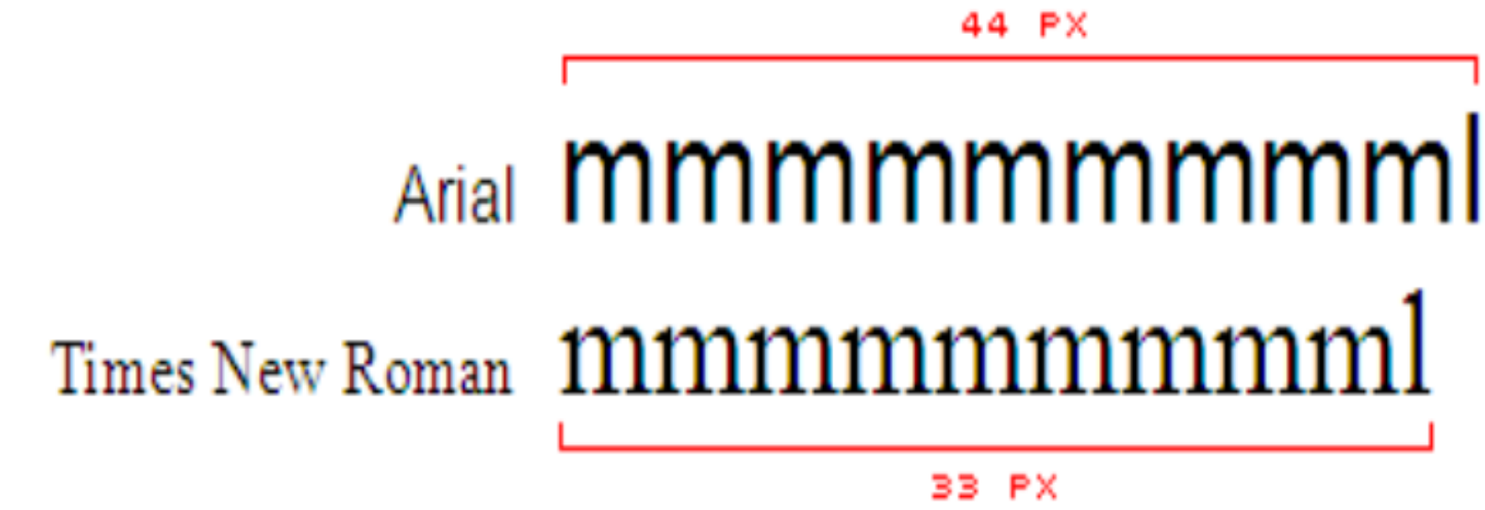

- 持续性；用户透明
- 易获得；*High-Entropy*

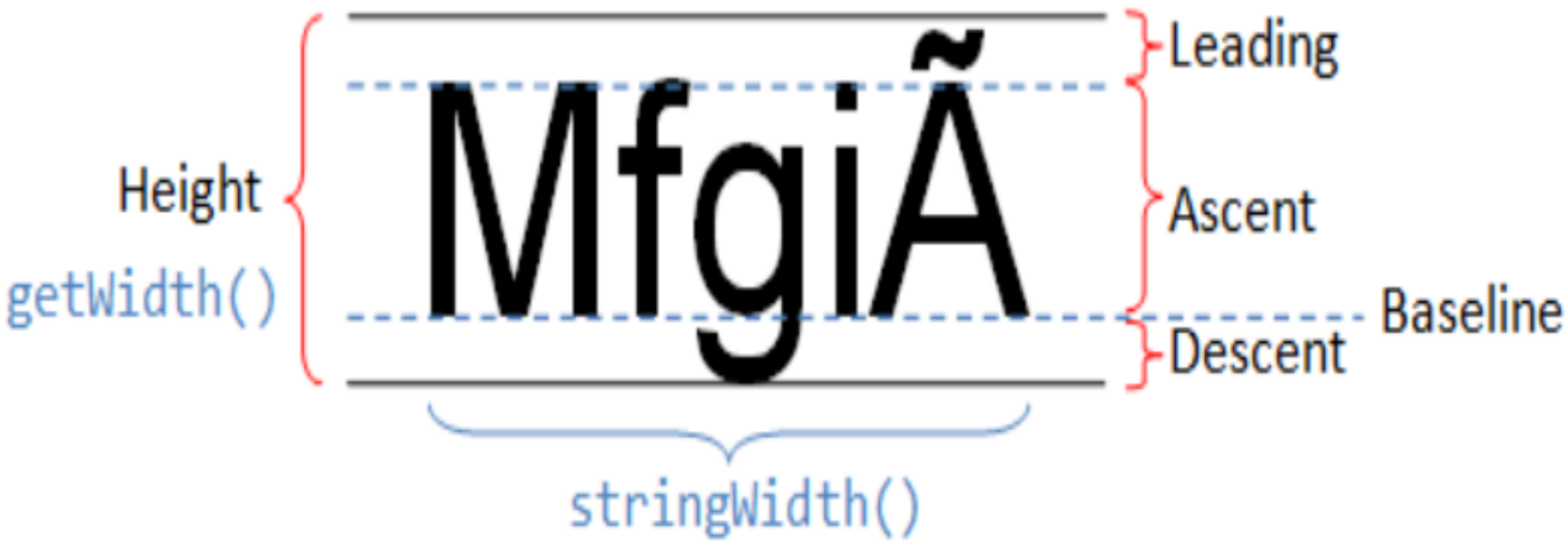
# Font Metrics

- 包含*JavaScript*的*Web*页面，将*code points(*码点*)*插入*DOM*,测量相应字体的维度。将每个码点放入空盒子*(bounding box)*，通过测量盒子的尺寸。

- 变量类型：*browser;browser version;*

- 安装的字体类型；字体的设置；



# Font Metrics

**Table 2.** Code points with the most and least individual entropy.

| rank | individual entropy (bits) | code point | name |
|---|---|---|---|
| #1 | 4.908178 | U+20B9 | INDIAN RUPEE SIGN |
| 2 | 4.798824 | U+20B8 | TENGE SIGN |
| 3 | 4.698577 | U+FBEE | ARABIC LIGATURE YEH WITH HAMZA ABOVE WITH WAW ISOLATED FORM |
| 4 | 4.698577 | U+FBF0 | ARABIC LIGATURE YEH WITH HAMZA ABOVE WITH U ISOLATED FORM |
| 5 | 4.698577 | U+FBF2 | ARABIC LIGATURE YEH WITH HAMZA ABOVE WITH OE ISOLATED FORM |
| 6 | 4.698577 | U+FBF4 | ARABIC LIGATURE YEH WITH HAMZA ABOVE WITH YU ISOLATED FORM |
| 7 | 4.657576 | U+F002 | *Private Use Area* |
| 8 | 4.652798 | U+F001 | *Private Use Area* |
| 9 | 4.646632 | U+FD3D | ARABIC LIGATURE ALEF WITH FATHATAN ISOLATED FORM |
| 10 | 4.640043 | U+FBF8 | ARABIC LIGATURE YEH WITH HAMZA ABOVE WITH E INITIAL FORM |
| 11 | 4.640043 | U+FBFB | ARABIC LIGATURE UIGHUR KIRGHIZ YEH WITH HAMZA ABOVE WITH ALEF MAKSURA INITIAL FORM |
| ⋮ | ⋮ | ⋮ | |
| 125,766 | 2.573742 | U+202A | LEFT-TO-RIGHT EMBEDDING |
| 125,767 | 2.573742 | U+202B | RIGHT-TO-LEFT EMBEDDING |
| 125,768 | 2.573742 | U+202D | LEFT-TO-RIGHT OVERRIDE |
| 125,769 | 2.573742 | U+202E | RIGHT-TO-LEFT OVERRIDE |
| 125,770 | 2.481283 | U+202C | POP DIRECTIONAL FORMATTING |
| 125,771 | 2.462760 | U+000C | FORM FEED (FF) |
| 125,772 | 2.462760 | U+000D | CARRIAGE RETURN (CR) |
| 125,773 | 0.156341 | U+00AD | SOFT HYPHEN |
| 125,774 | 0.000000 | U+0009 | CHARACTER TABULATION |
| 125,775 | 0.000000 | U+000A | LINE FEED (LF) |
| 125,776 | 0.000000 | U+0020 | SPACE |

# 例子

您的浏览器指纹是:

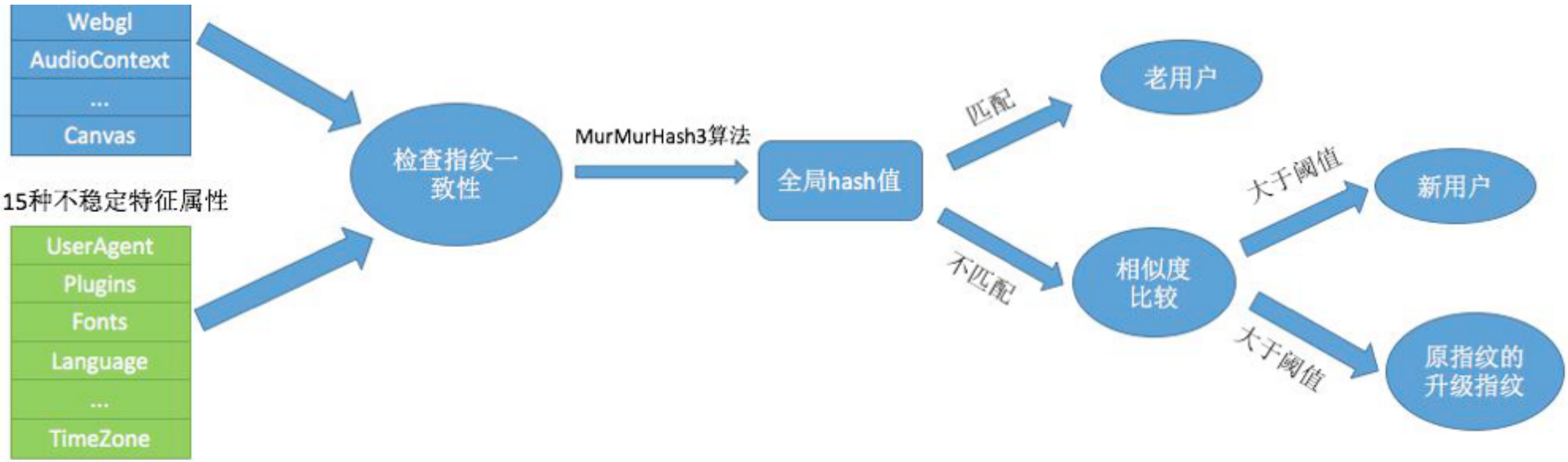
**93a23be01b82930c37707a264def73b9**

这是您第 **1** 次访问。

计算浏览器指纹耗时:**168** ms

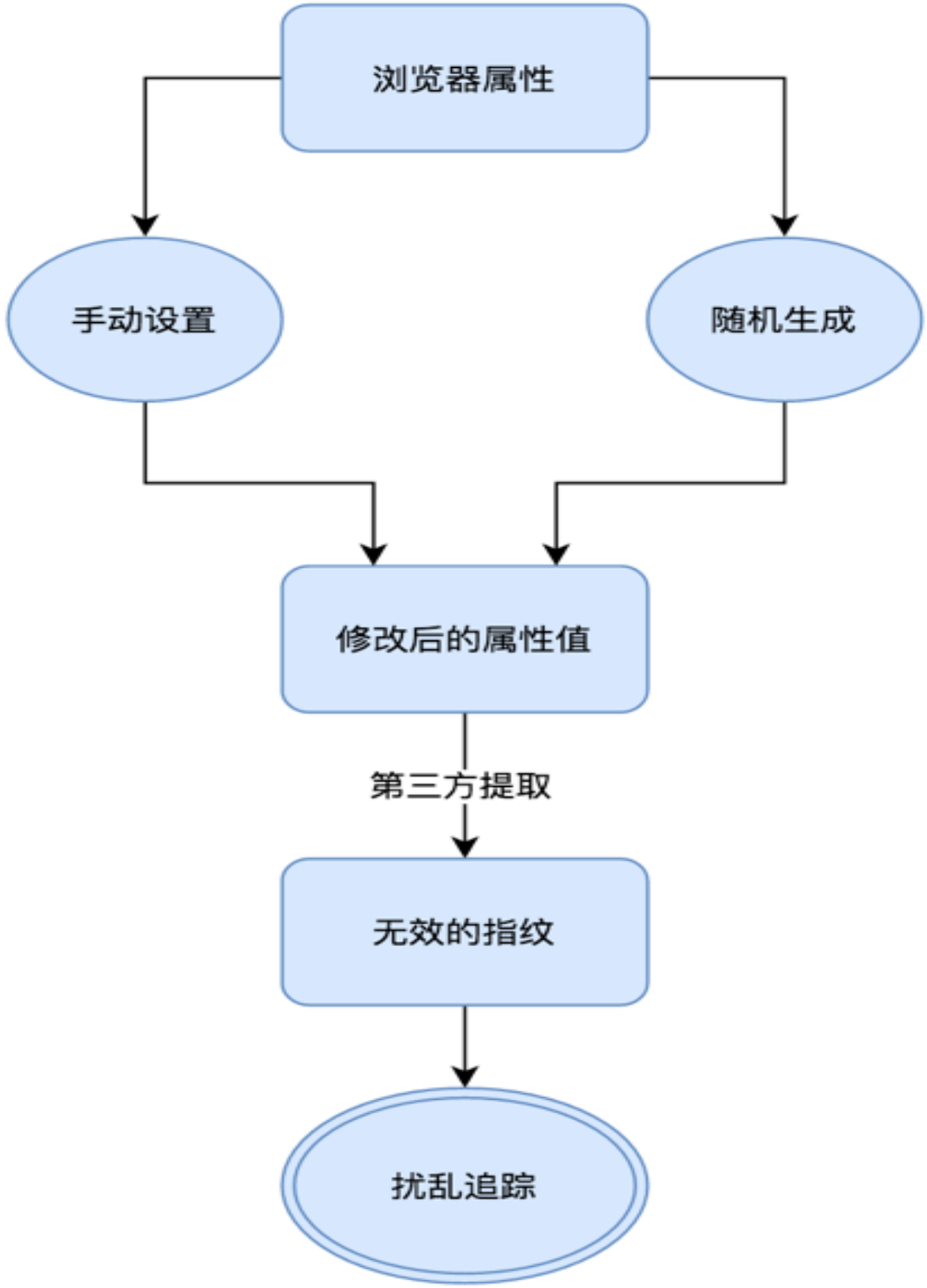
1. **user_agent** : Mozilla/5.0 (Macintosh; Intel Mac OS X 10_12_3) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/57.0.2987.110 Safari/537.36
2. **resolution** : 1440,900
3. **available_resolution** : 1440,873
4. **navigator_platform** : MacIntel
5. **language** : en-US
6. **timezone_offset** : -480
7. **color_depth** : 24
8. **pixel_ratio** : 2
9. **session_storage** : 1
10. **local_storage** : 1
11. **indexed_db** : 1
12. **open_database** : 1
13. **cpu_class** : unknown
14. **do_not_track** : unknown
15. **regular_plugins** : Widevine Content Decryption Module::Enables Widevine licenses for playback of

# 浏览器变化


Webgl
AudioContext
...
Canvas
15种不稳定特征属性
UserAgent
Plugins
Fonts
Language
...
TimeZone
检查指纹一致性
MurMurHash3算法
全局hash值
匹配
老用户
不匹配
相似度比较
大于阈值
新用户
大于阈值
原指纹的升级指纹

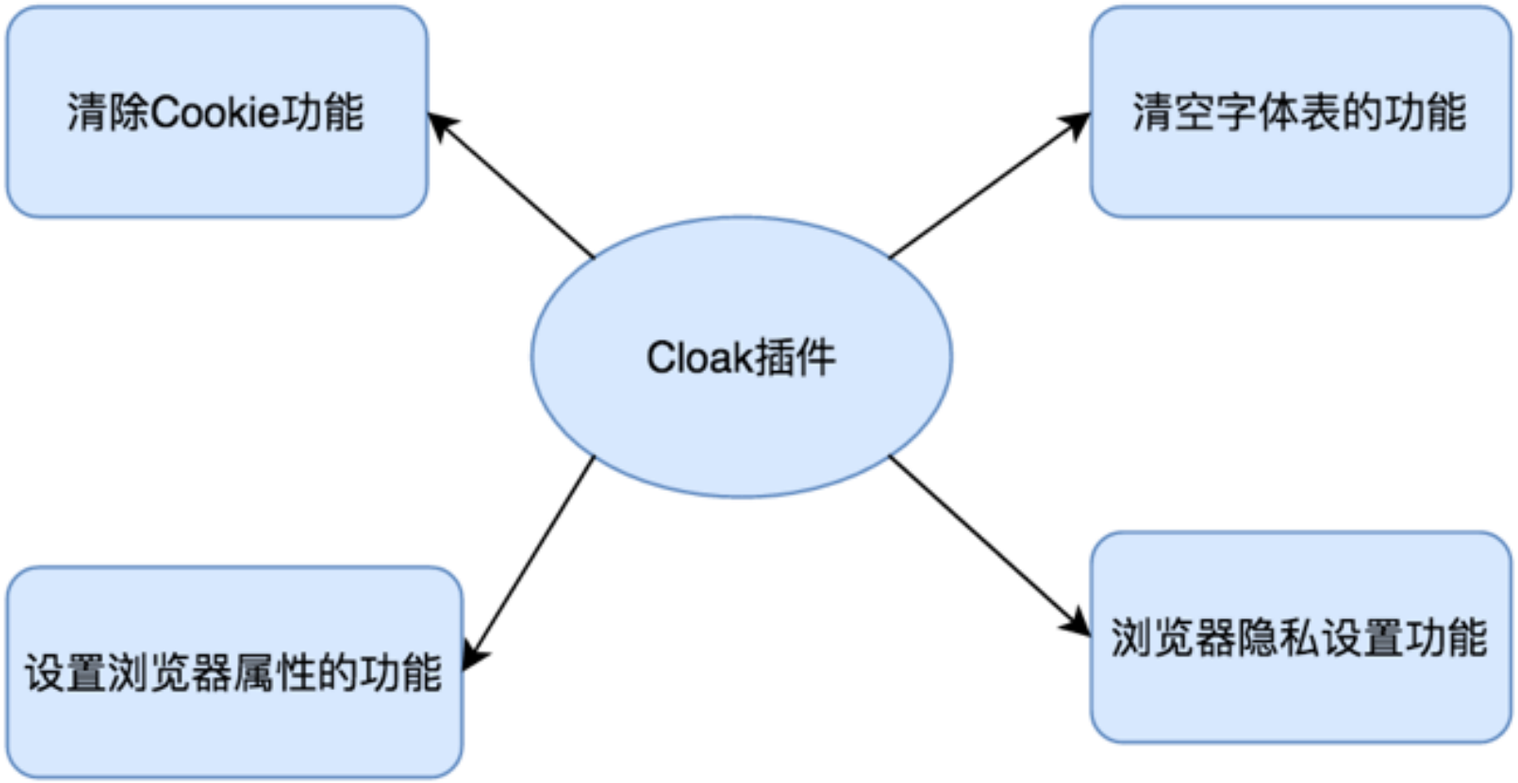
# 浏览器指纹提取防范

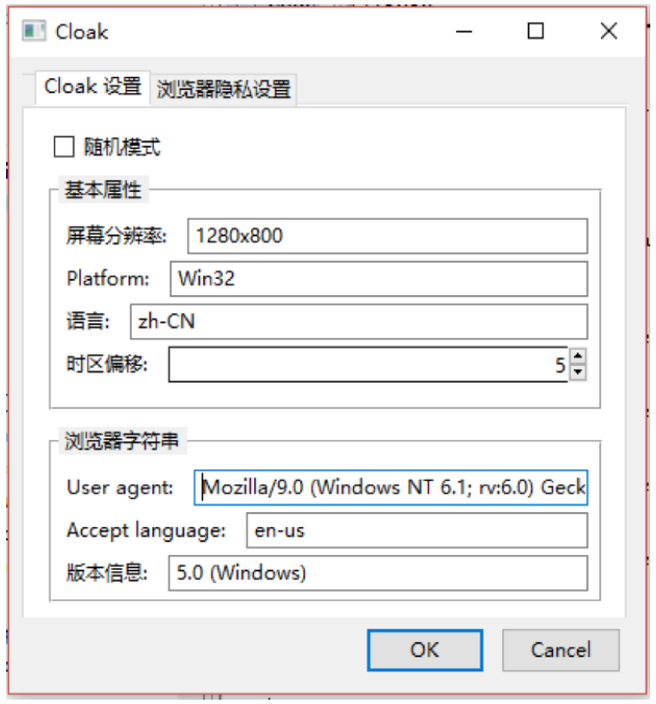
- 手动设置或随机生成某些浏览器属性的值，这样当第三方追踪软件想要提取我们浏览器属性的值时，获取到的就是经过修改或随机生成的值，从而生成无效的浏览器指纹。

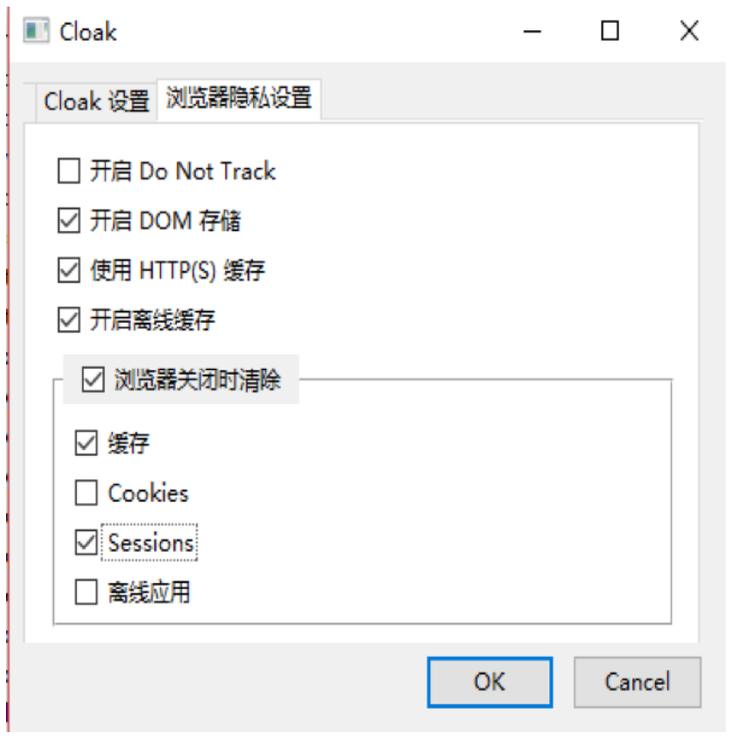


Fingerprint

# 浏览器指纹提取防范系统



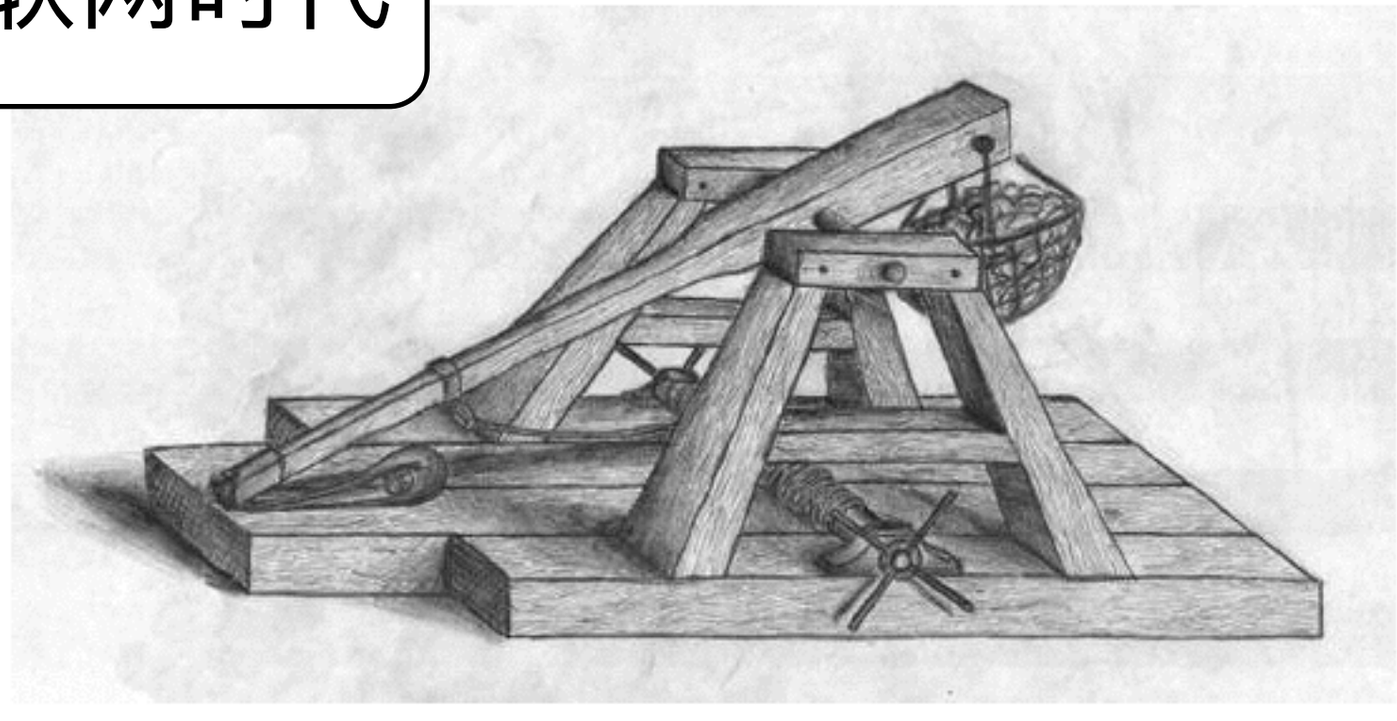



# 物理保护

Fingerprint

# 墙是一种物理保护


互联网时代



# 访问控制

- 对于一个资源和地点的选择性的访问限制
- 物理安全 *vs* 信息安全
- 访问可以是消费、进入、使用、退出等
- 分配存取资源的允许，授权
- 验证：人 *vs* 机器

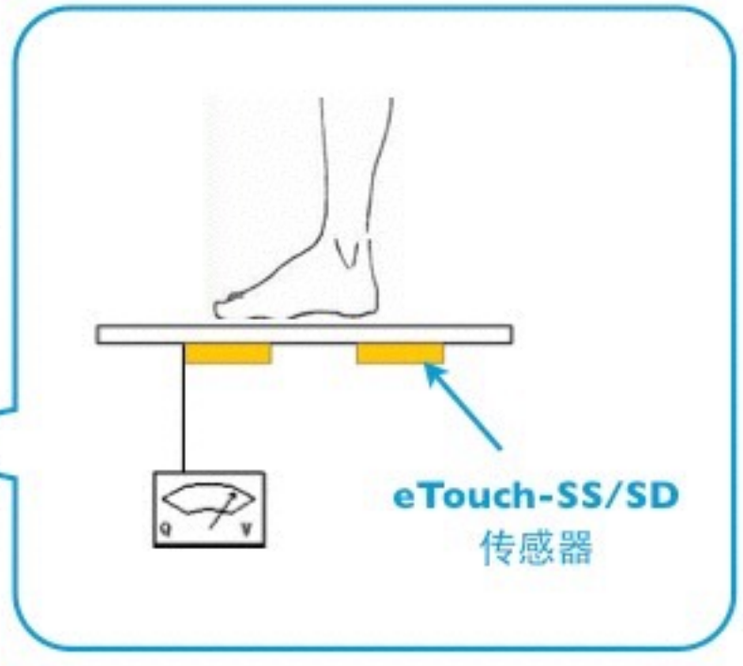
| Who<br>Where<br>When | Key<br>Credential |
|---|---|

- 授权每个员工可以进入自己房间，但不能进入其他人房间，但是保洁员可以访问所有人房间

# 物理防护手段和威胁

- 振动探测
- 红外检测
- 运动监测

- 拒绝服务攻击
- 围墙和安全
- 威慑作用
- 错误警报



# 攻击者能力 vs 手段

- 无目标、无知识、无能力
- 有“专业知识”
- 有目标、有知识
- 内部人员
- 有资助、有支持

- 威慑
- 检测
- 报警
- 延迟
- 响应

# 威慑和障碍

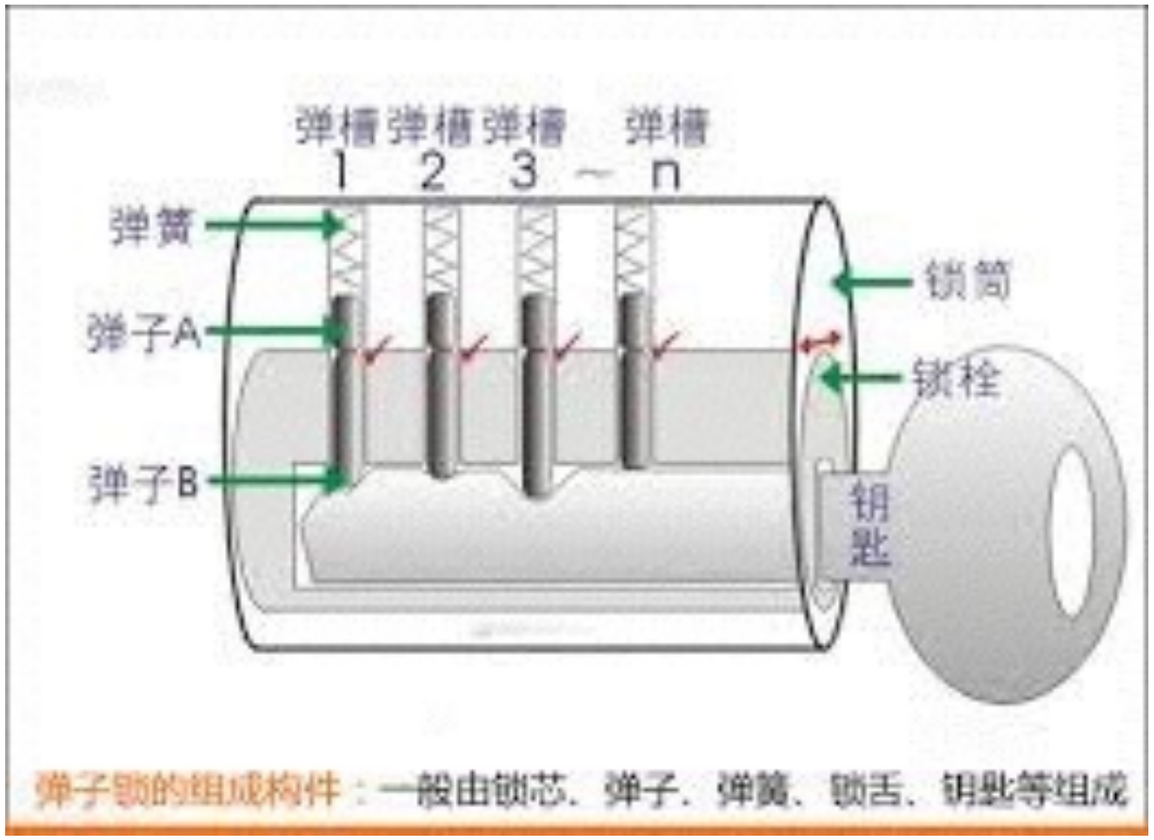
- 默默无闻
- 环境设计
- 保护目标
- 安全边界
- 报警器
- 通信系统

- 好区 vs 坏区
- 可防卫空间
- 破窗理论

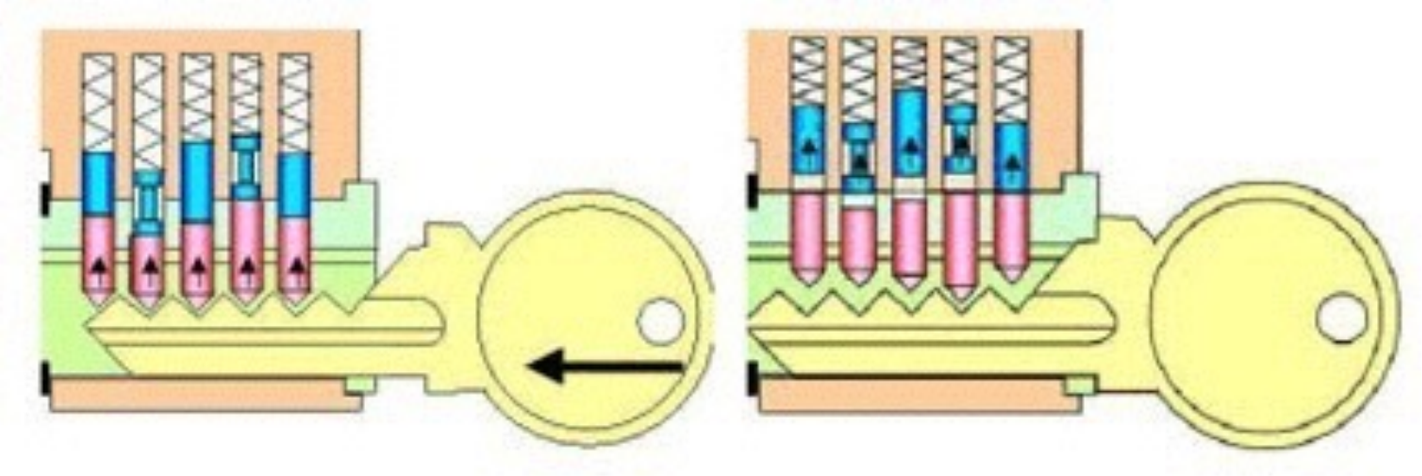
- 银行
- 核电站

- 传感器失灵
- 误报
- 人为破坏

# 存在万能钥匙




撞匙开锁枪



撞钥攻击

主钥攻击

*Fingerprint*

# 电子锁




# 封印

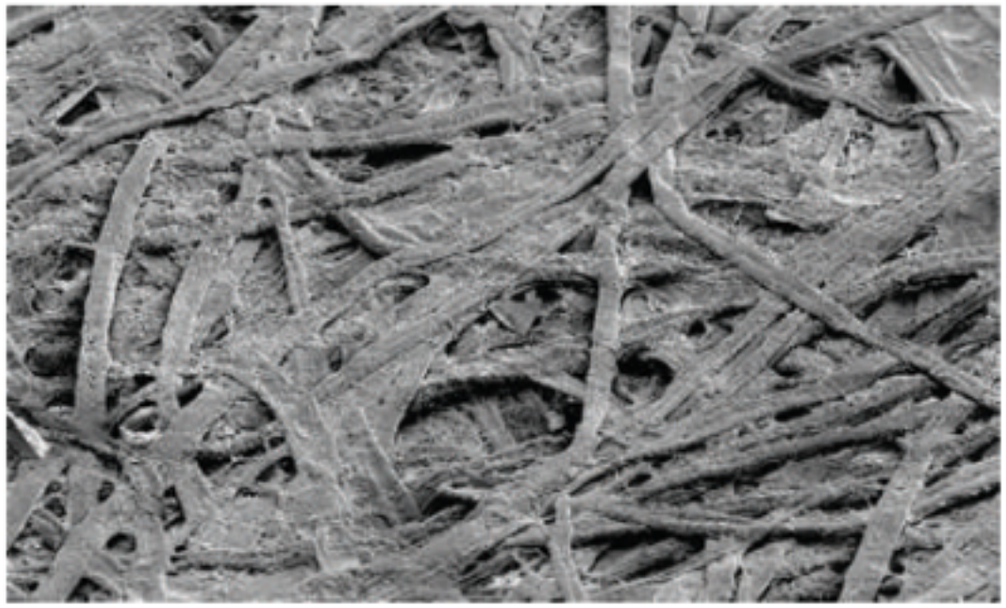

新版
2015年版 第5套
人民币
那些事...
中国人民银行
100
壹佰圆
样币
AA00000000
样币禁止流通
钞票正面
钞票背面
ZHONGGUO RENMIN YINHANG
100
YUAN
2015年
样币禁止流通
多一份金融了解
多一份财富保障
2015年版第5套人民币100元纸币在保持2005年版第五套人民币100元纸币规格、正背面主图案、主色调、“中国人民银行”行名、国徽、盲文和汉语拼音行名、民族文字等不变的前提下，对部分图案做了调整，对整体防伪性能进行了提升。
光变镂空开窗安全线
胶印凹印
胶印对印图案
光彩光变数字
白水印
人像水印
镜变双号码
广发银行
CGB

*Fingerprint*

# 封印

# *Fingerprint*

# 安全打印



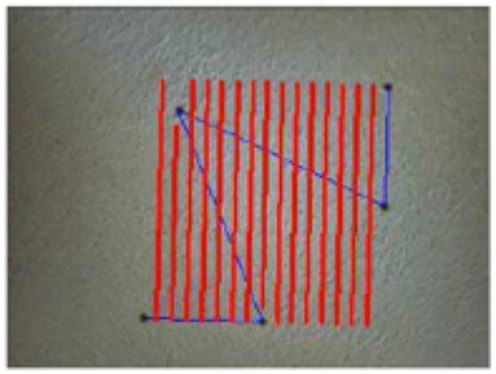
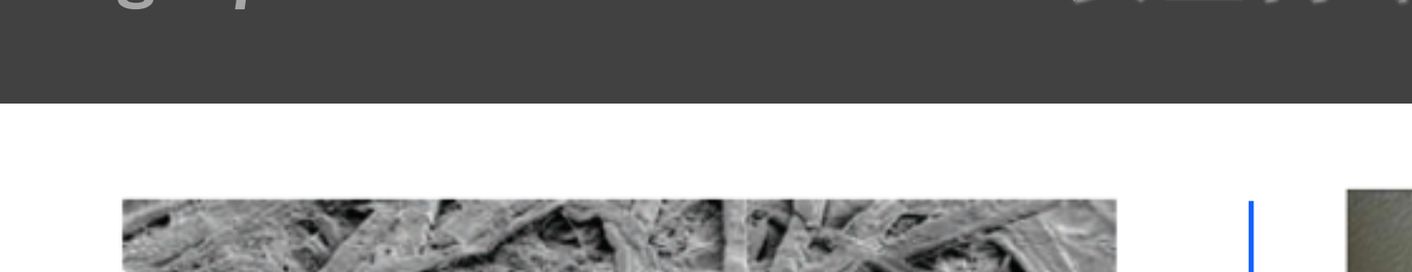
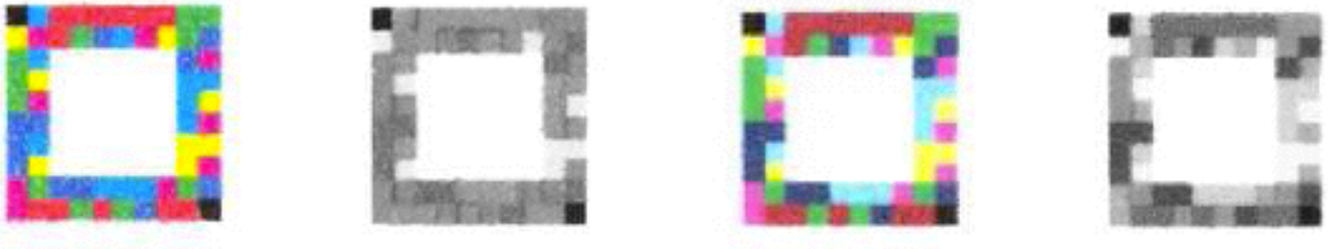
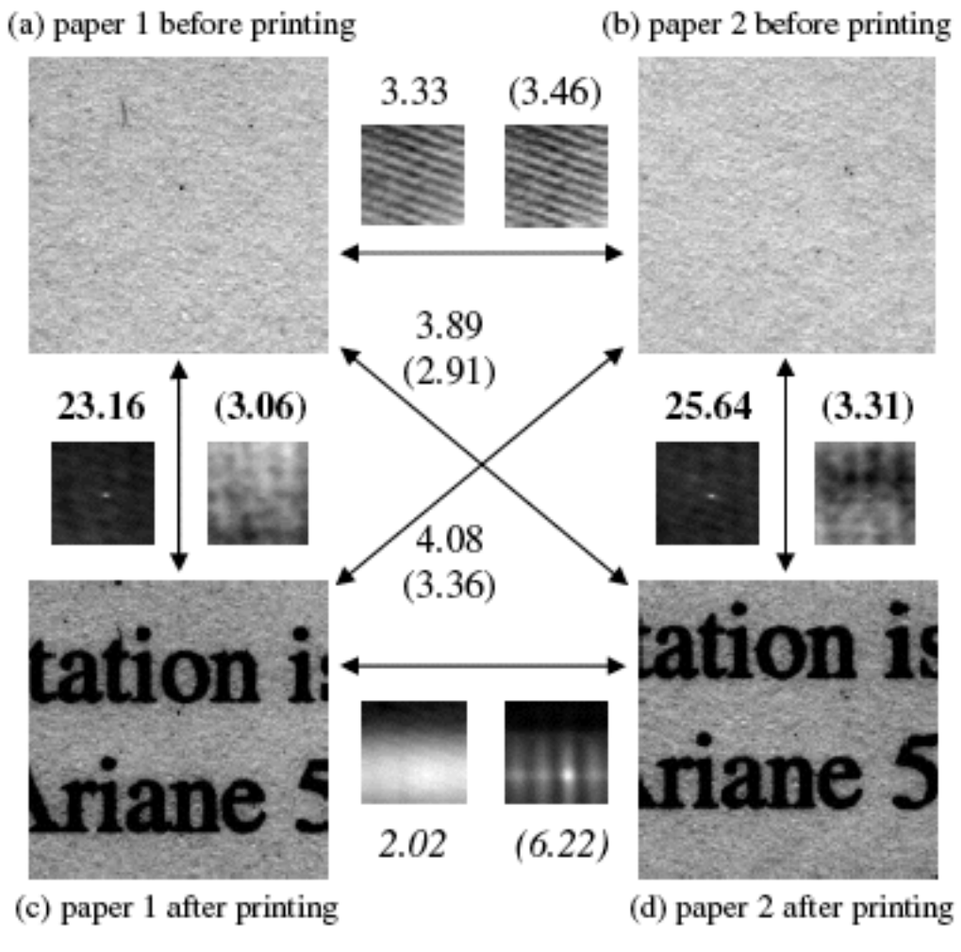


**Figure 3: a) The Fiberfingerprint method; b) Verification device used to extract the substrate individualities (Source: Metois et al. (2002))**







# 攻击安全处理器

- 偷取密钥
- 外壳切割
- 探针攻击
- 剩磁攻击
- 冷冻攻击
- 电磁泄漏
- 接口攻击
- … ….

- 协议分析
- 光学探测
- 冗余监测
- 柠檬市场

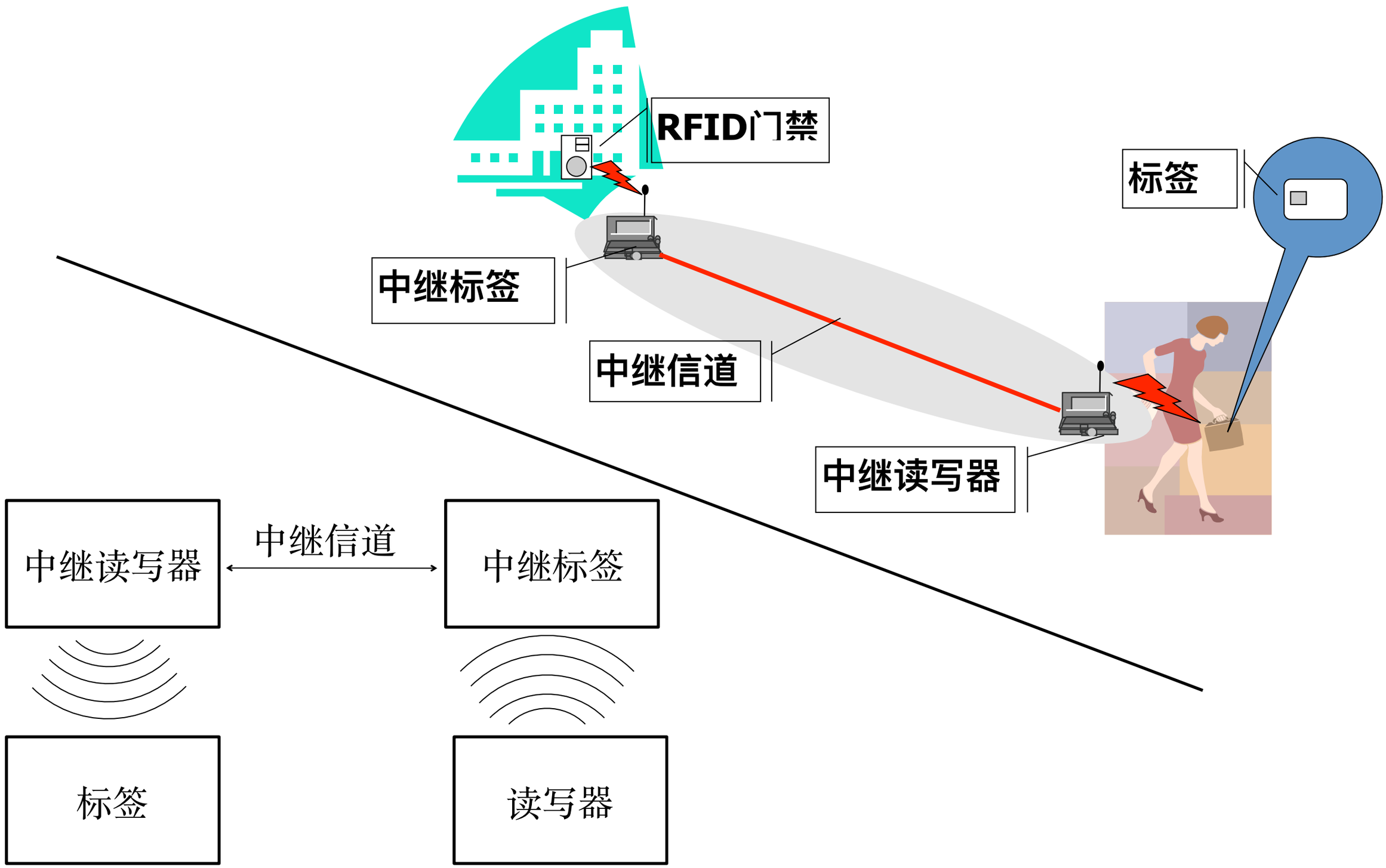
# 物理防篡改 vs 物理攻击


IBM 4758


# 发射安全

- 电话窃听
- 功耗分析
- 电磁泄漏
- 光学旁路
- 声学旁路
- … ….

# 中继攻击



# 课后作业

*Homework*

# 课后作业

## 基于Bitcoin开发一个应用

**1、基于Bitcoin开源库**

**2、实现一个常见的应用**

**3、可以运行，可以演示**

**4、可以改造Bitcoin**

## 12月9日晚上12点前提交给助教

# 谢谢！

*Huiping Sun*

*sunhp@ss.pku.edu.cn*

*https://huipingsun.github.io*