

# *Information Security Engineering*

2017.11.08

## CAPTCHA



*Huiping Sun(孙惠平)*  
[sunhp@ss.pku.edu.cn](mailto:sunhp@ss.pku.edu.cn)

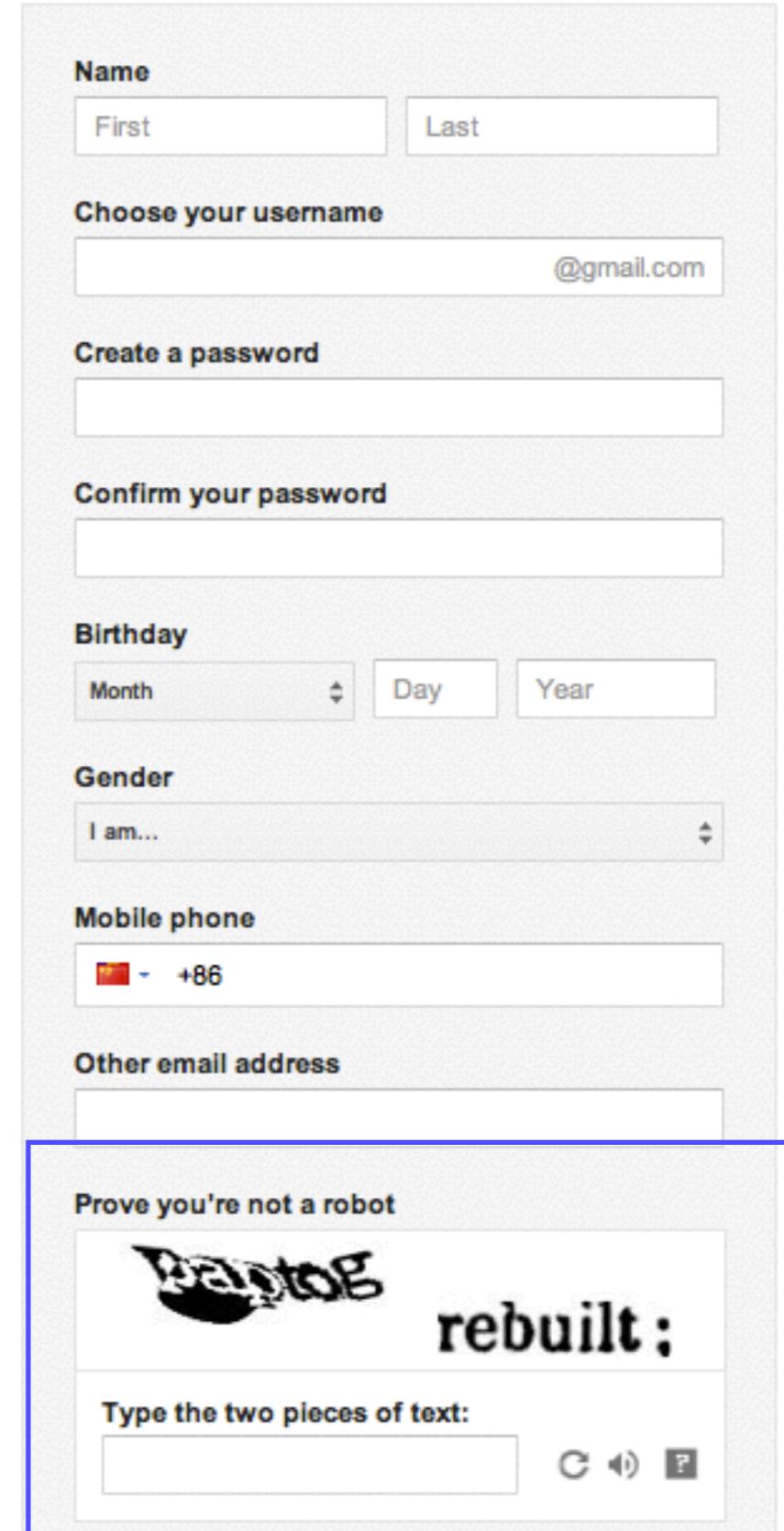
# 什么是CAPTCHA

## ● CAPTCHA

Completely  
Automated  
Public  
Turning test to tell  
Computers and  
Humans  
Apart

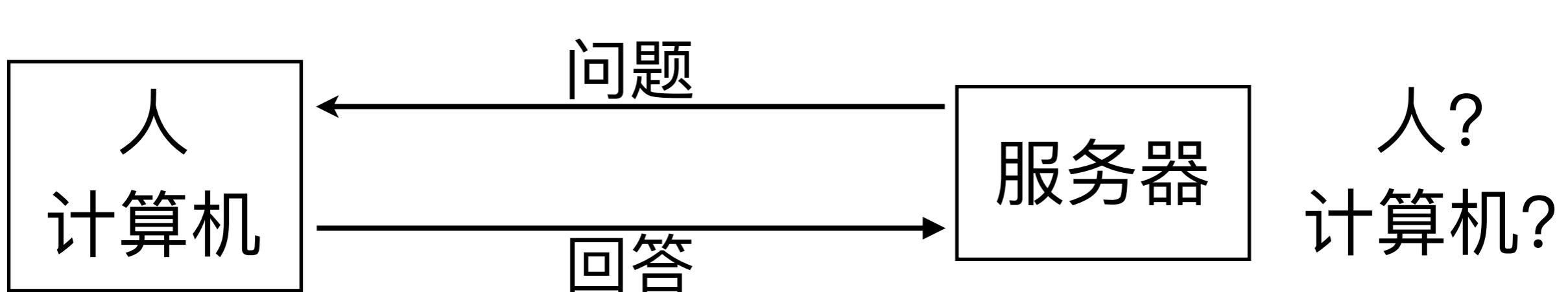
<http://www.captcha.net/>

<http://en.wikipedia.org/wiki/CAPTCHA>

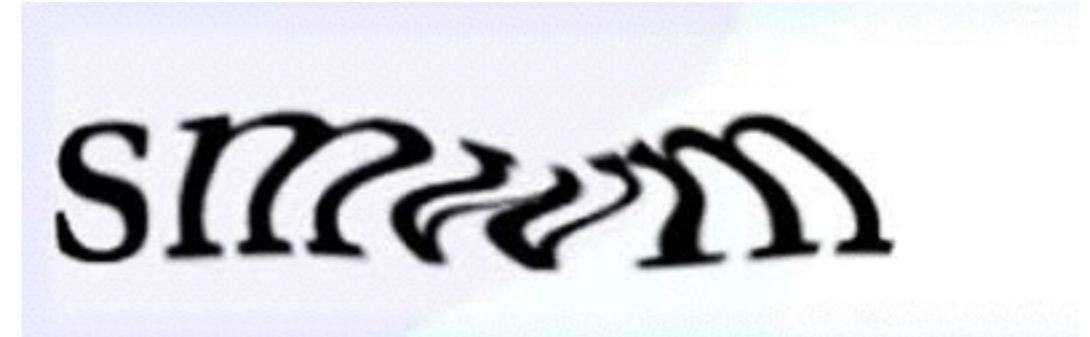


# CAPTCHA定义

- A program that generates and grades tests that are human solvable but beyond the capabilities of current computer programs
- 全自动区分计算机和人类的图灵测试
- 验证码，一种区分用户是计算机和人的公共全自动程序



- 1996, Moni Naor
- 1997年, AltaVista
- Yahoo邮件是第一个大规模采用CAPTCHA的网站
- 2000年, von Ahn和Blum设计出第一个抵御OCR的CAPTCHA
- 2007年, reCAPTCHA
- 2007年, NoCAPTCHA
- ... ...



---

HIPs:  
Human Interactive Proof

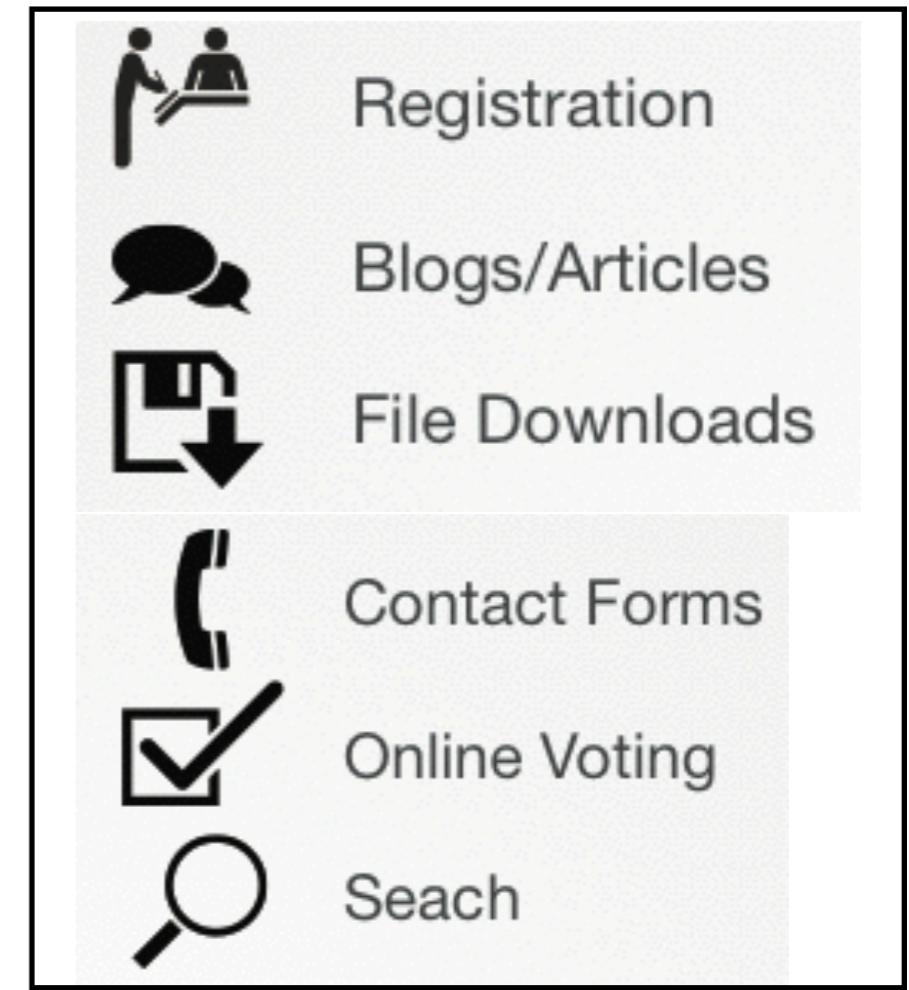
---

OCR:  
Optical Character Recognition

- 判断人还是机器

- \* 保护网站用户注册， 提供免费服务给用户而不是bots
- \* 保护搜索引擎， 防止评论spam
- \* 确保在线调查更可信， 防止虚假投票
- \* 防止spam制作者很容易的获得email地址
- \* 防止针对口令系统的字典攻击
- \* 抵御bots

SPAM



# CAPTCHA vs AI

- 设计CAPTCHA

- \* Email

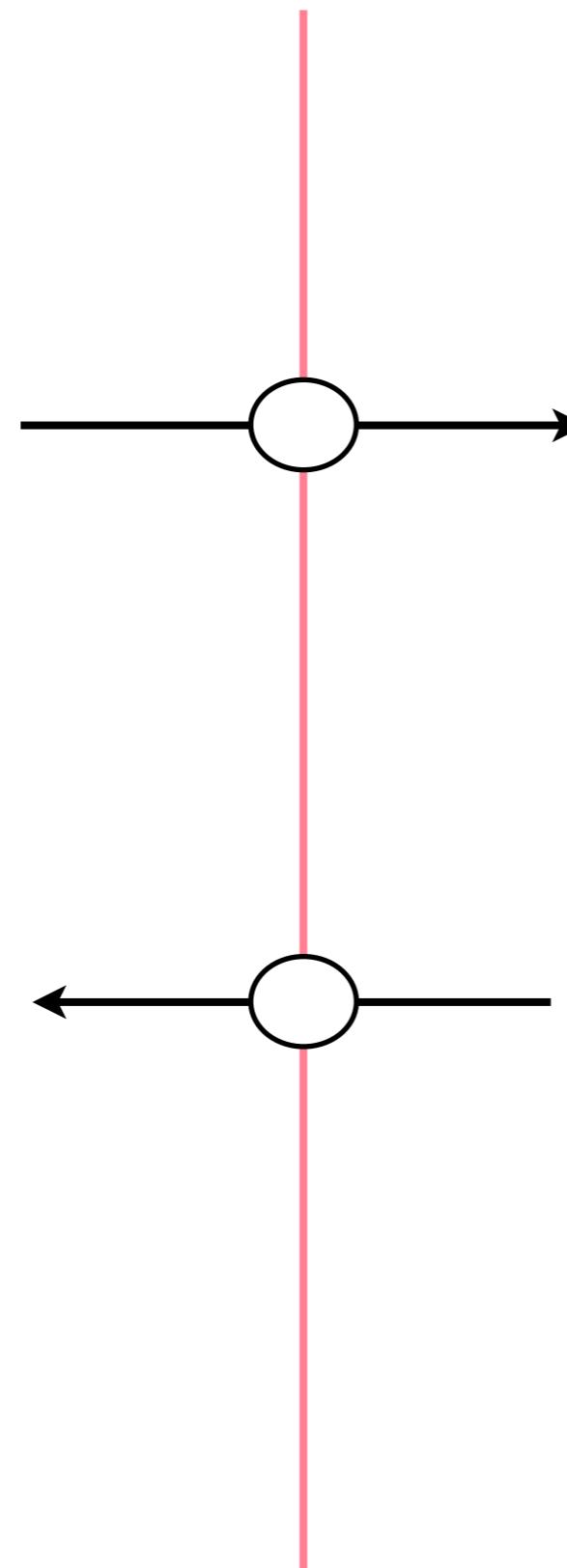
- \* BBS

- \* Blog

- \* SNS

- \* Security

- \* ...



- 攻击CAPTCHA

- \* 垃圾信息生产者

- \* 僵尸网络掌握者

- \* 打码等其余攻击者

- \* 自动化测试人员

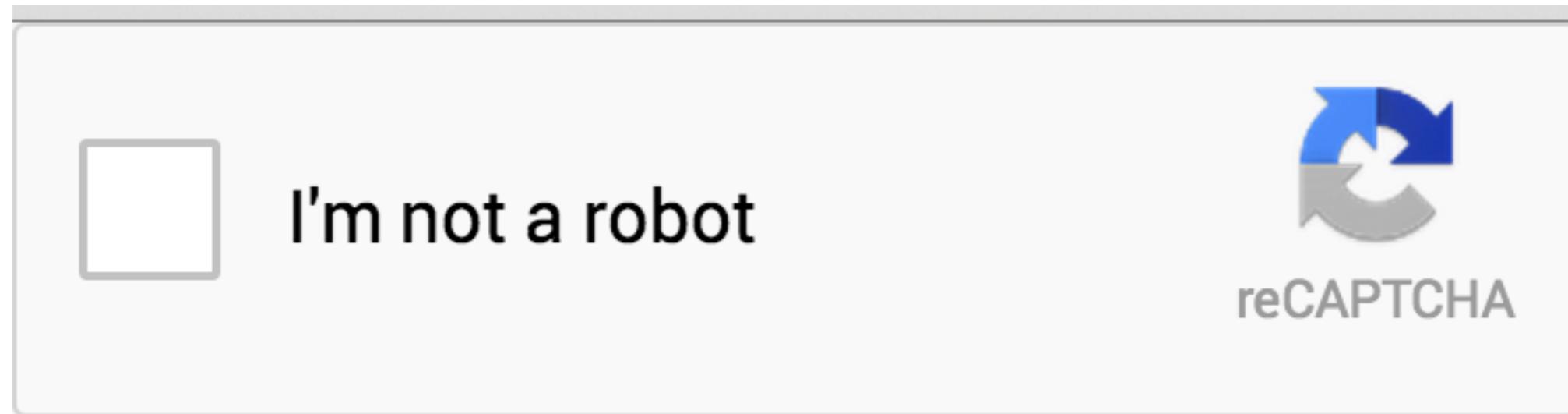
- \* ...

AI难题

机器学习

# Introduction

# NoCAPTCHA



A standard image-based CAPTCHA interface. It includes a checkbox labeled "I'm not a robot" and the reCAPTCHA logo. Below this, the text "Select all images below that match this one:" is displayed. A grid of 12 small images is shown, with the top-left one being a turkey. At the bottom are three control icons (refresh, volume, info) and a "Verify" button.

A standard image-based CAPTCHA interface. It includes a checkbox labeled "I'm not a robot" and the reCAPTCHA logo. Below this, the text "Select all images below that match this one:" is displayed. A grid of 12 small images is shown, with the top-left one being a cat. At the bottom are three control icons (refresh, volume, info) and a "Verify" button.

A text-based CAPTCHA interface. It includes a checkbox labeled "I'm not a robot" and the reCAPTCHA logo. Below this is a text input field with the placeholder "Type the text". A blurred image of a house number "250" is displayed above the input field. At the bottom are three control icons (refresh, volume, info) and a "Verify" button.

文本CAPTCHA

## 文本 vs 图片

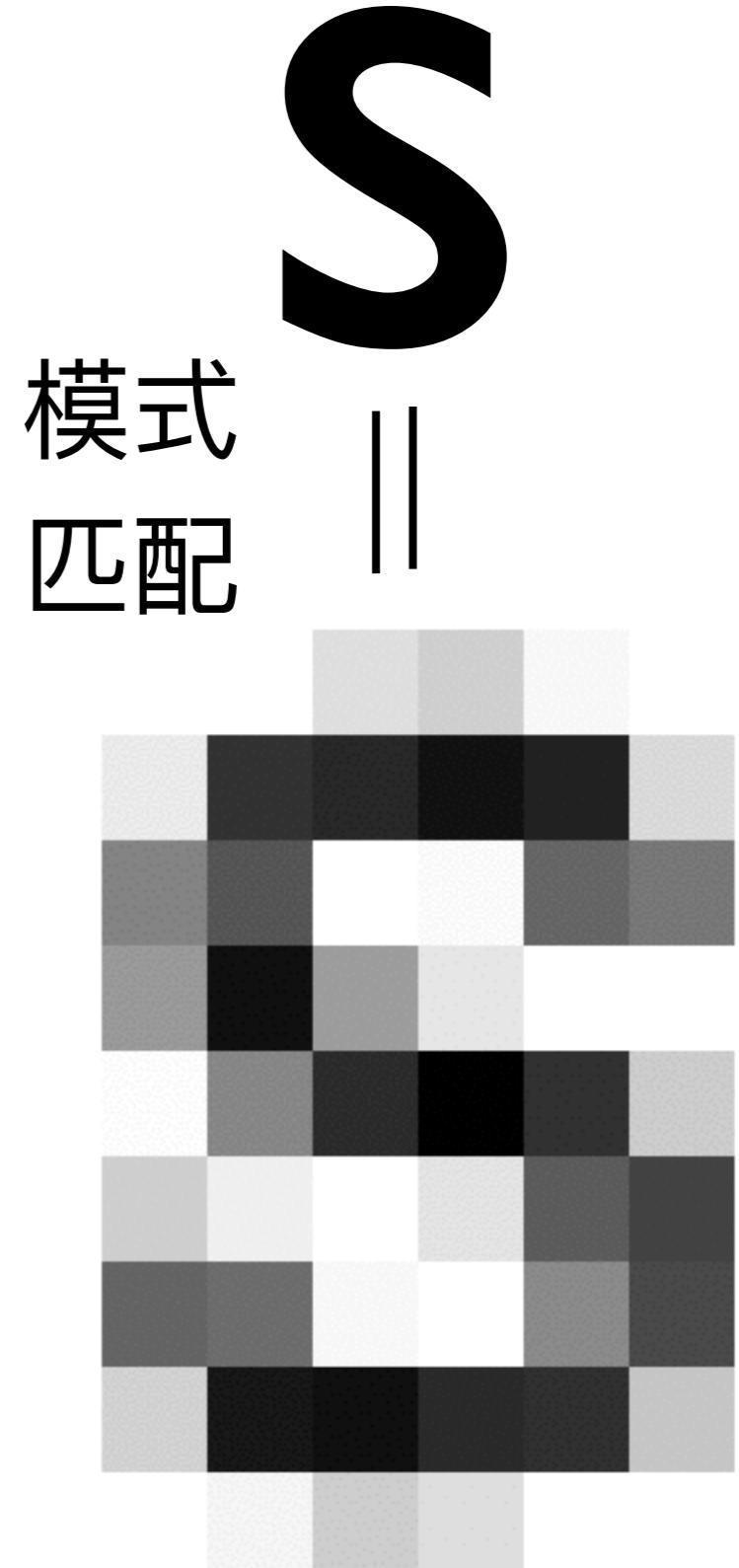
- 请进入如下图片的文字

sunhp@ss.pku.edu.cn

- 文字如下：

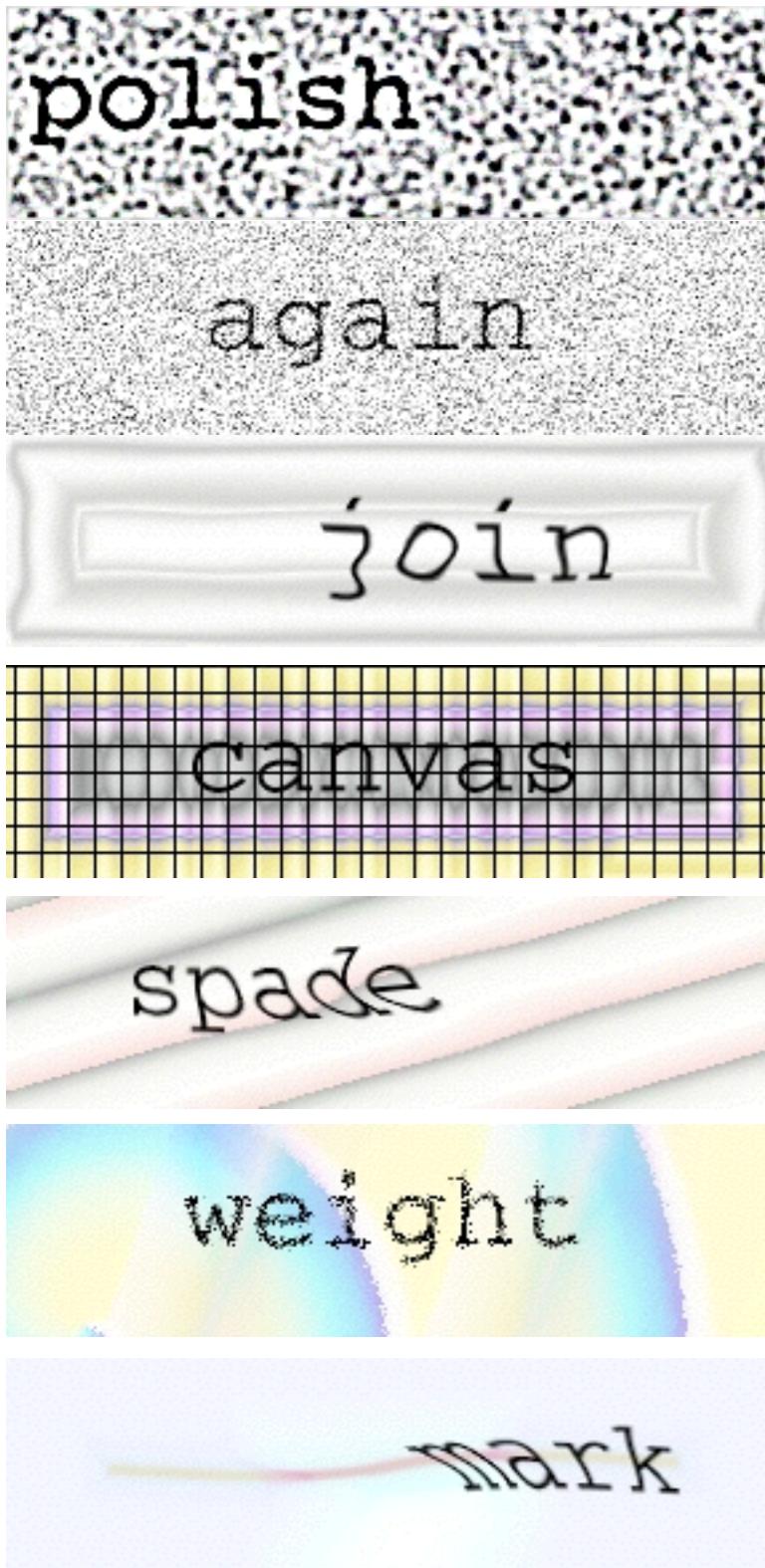
sunhp@ss.pku.edu.cn

- 像素点
- 图片：二维的像素点
- 计算机不能识别图片中包括什么文本
- 计算机能识别图片中哪些像素相同不相同



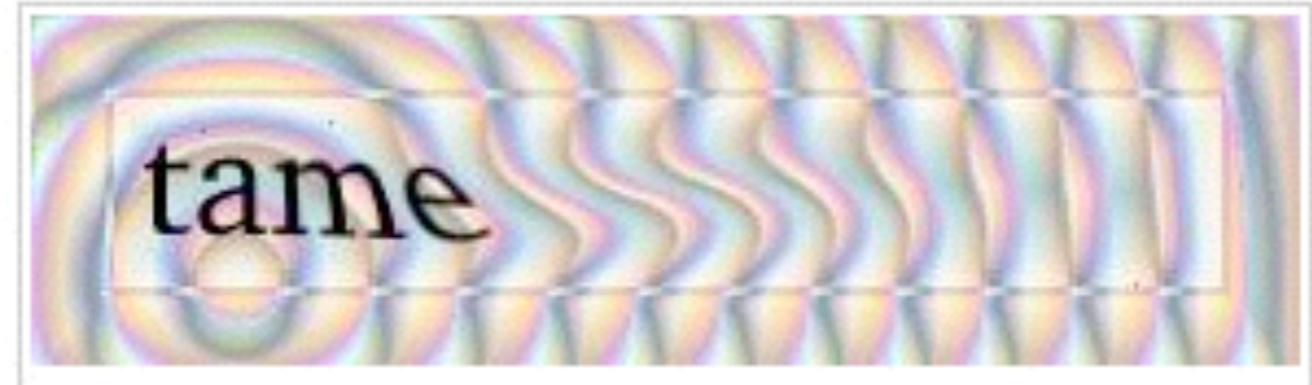
# Text CAPTCHA

## GIMPY / EZ-GIMPY

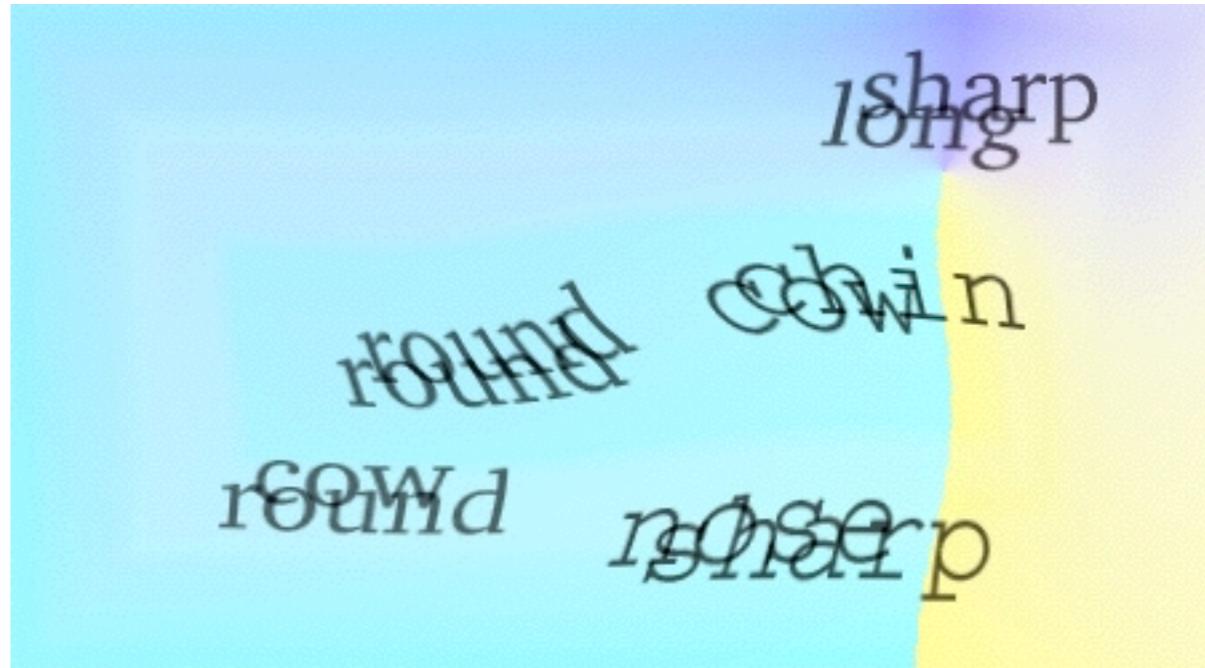


### EZ-GIMPY

Enter the word as it is shown in the box below.



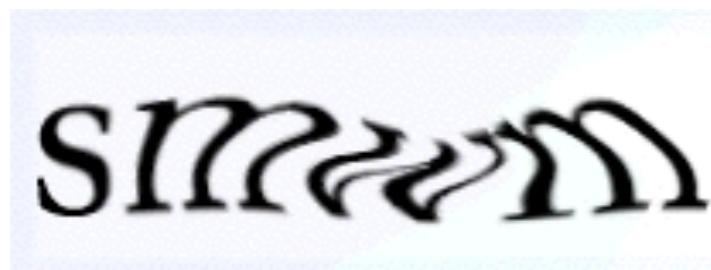
### GIMPY



<http://www.cs.sfu.ca/~mori/research/gimpy/>

## 文本CAPTCHA原理

Yahoo's EZ-Gimpy



假设：人能读  
扭曲的文本，  
但计算机不能

一般情况下要求  
用户输入扭曲图  
像中的文字

fellowing

finding

增加一条曲线使得  
图像分割更困难

将符号彼此拥挤在一起，但是人也比较难于识别

Eurotrip

klopsh

3m53

vqkz

# Text CAPTCHA

## 常见文本CAPTCHA



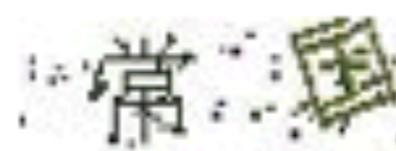
谷歌的验证码



QQ的验证码



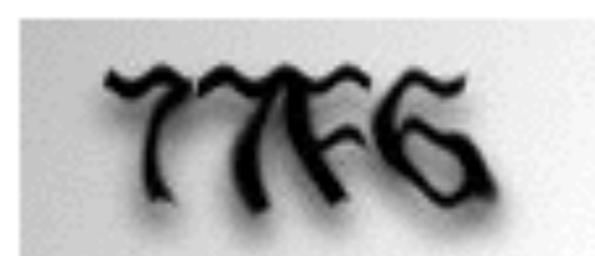
新浪的验证码



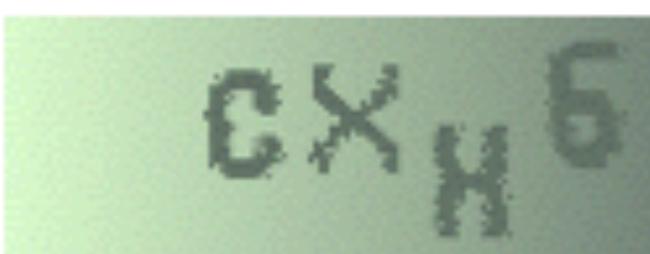
网易的验证码



人人网的验证码



天涯的验证码



Discuz的验证码



北大软微论坛

# Text CAPTCHA

# 人为攻击CAPTCHA

任务名称	每码价格	更新周期	结算状态	赚友推荐	操作
银河答题（超简单）	28¥	1天	05月07日金币已发	165	继续任务
阳光打码（日赚30）	15¥	1天	05月07日金币已发	1435	立即参与
打码免（日赚50元）	17¥	1天	05月07日金币已发	2625	继续任务
知码打码（强力推荐） <b>活动</b>	24¥	1天	05月07日金币已发	3311	立即参与
点图打码（夜班力荐）	28¥	1天	05月07日金币已发	26	立即参与
图像打码（积分结算）	15¥	1天	05月07日金币已发	371	立即参与
极速打码（工号加前缀）	16¥	1天	05月07日金币已发	476	立即参与
超速打码（奖励超高） <b>活动</b>	16¥	1天	05月07日金币已发	1036	立即参与
发财打码（积分结算）	14¥	1天	05月07日金币已发	237	立即参与
大众打码（积分结算）	10¥	1天	05月07日金币已发	399	立即参与

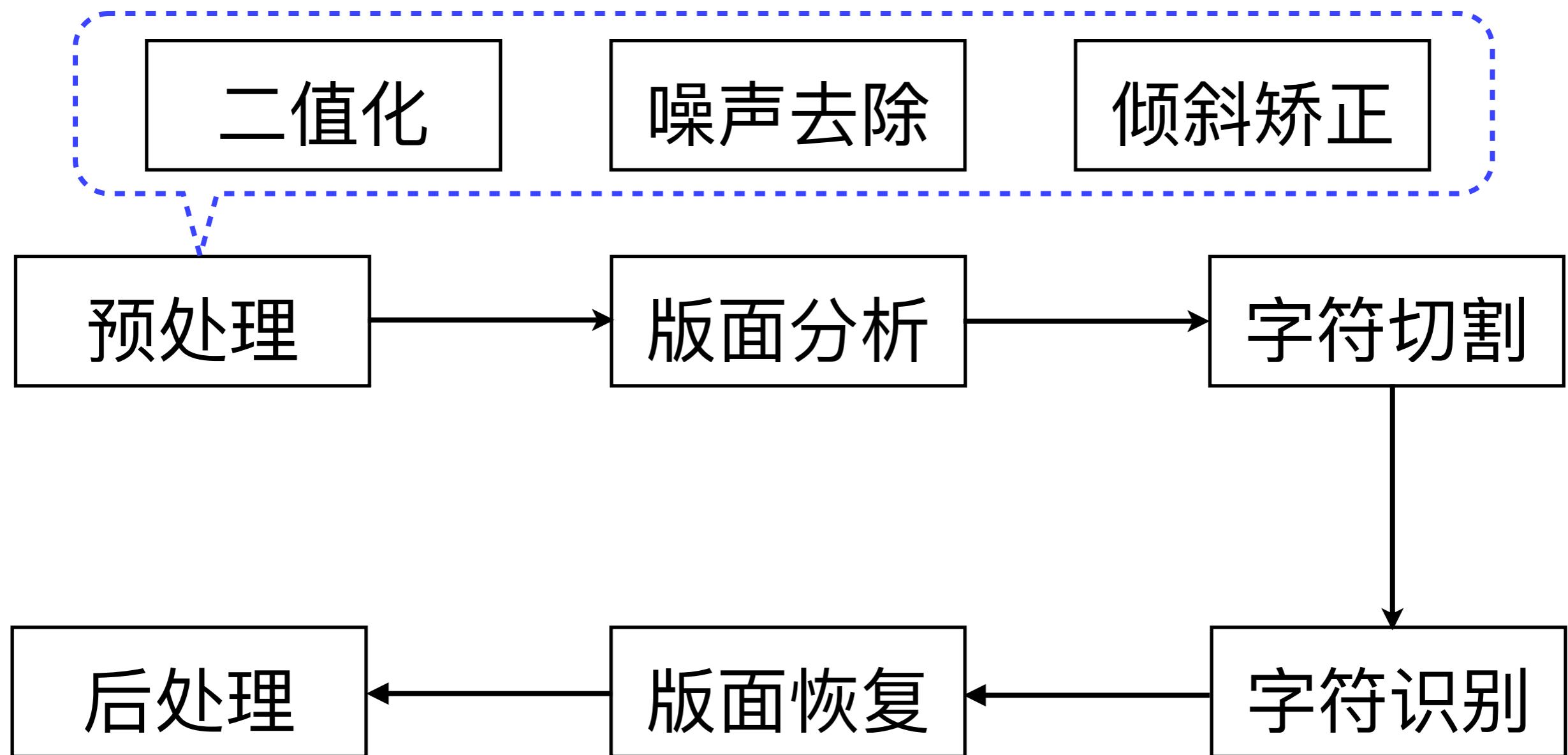


Bypass  
CAPTCHA

5k CAPTCHAs  
10k CAPTCHAs  
25k CAPTCHAs  
50k CAPTCHAs  
100k CAPTCHAs

Only US\$6.95  
Only US\$13.90  
Only US\$34.75  
Only US\$69.50  
Only US\$139.00

- OCR: Optical Character Recognition, 光学字符识别



# Text CAPTCHA

漏洞



当当

字符分离



京东

字符简单粘连



字符特征粘连

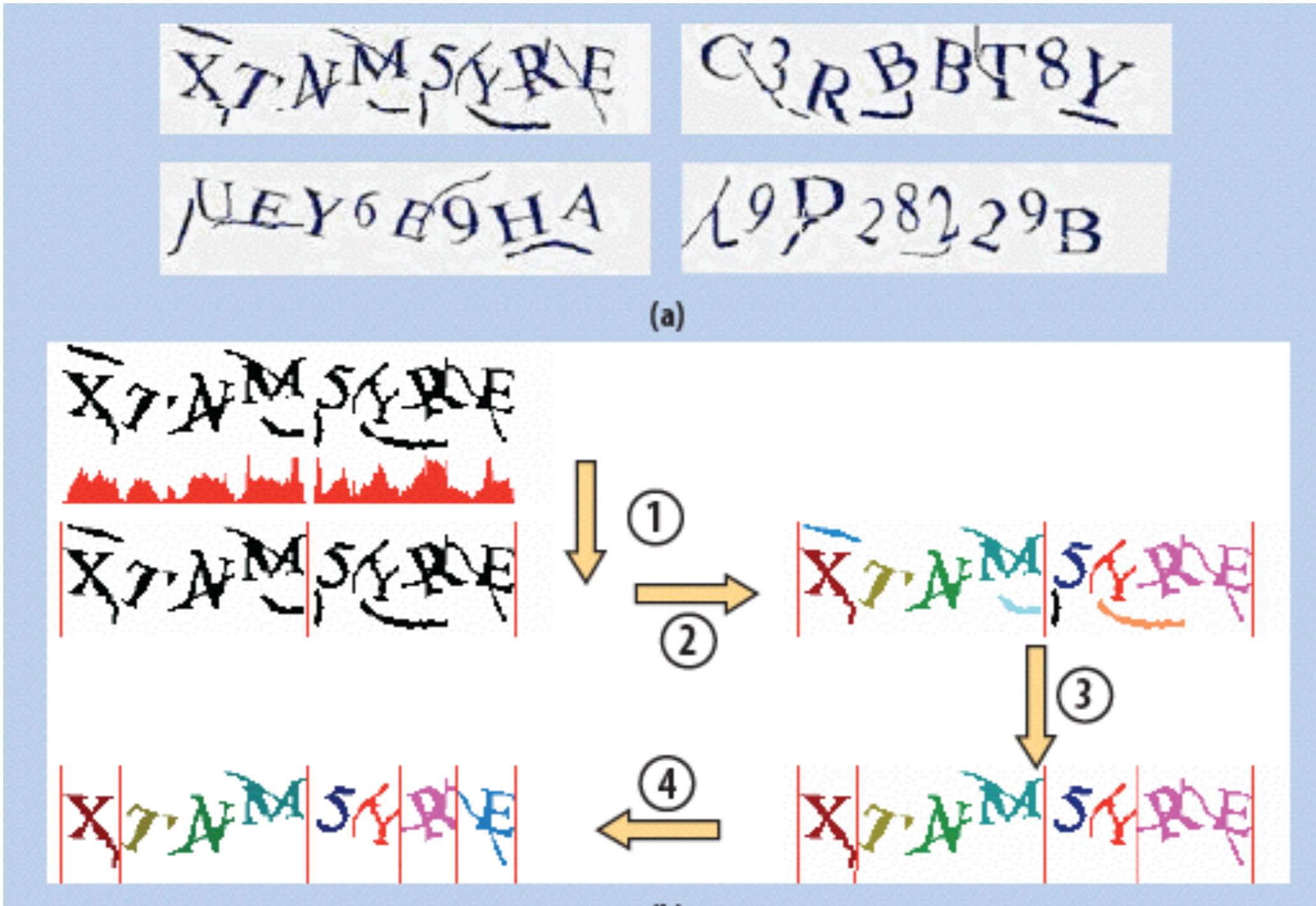
- 像素点个数
- 长宽比
- 像素点密度
- 像素点重心

攻击门限      0.01% or 1%

挑战空间

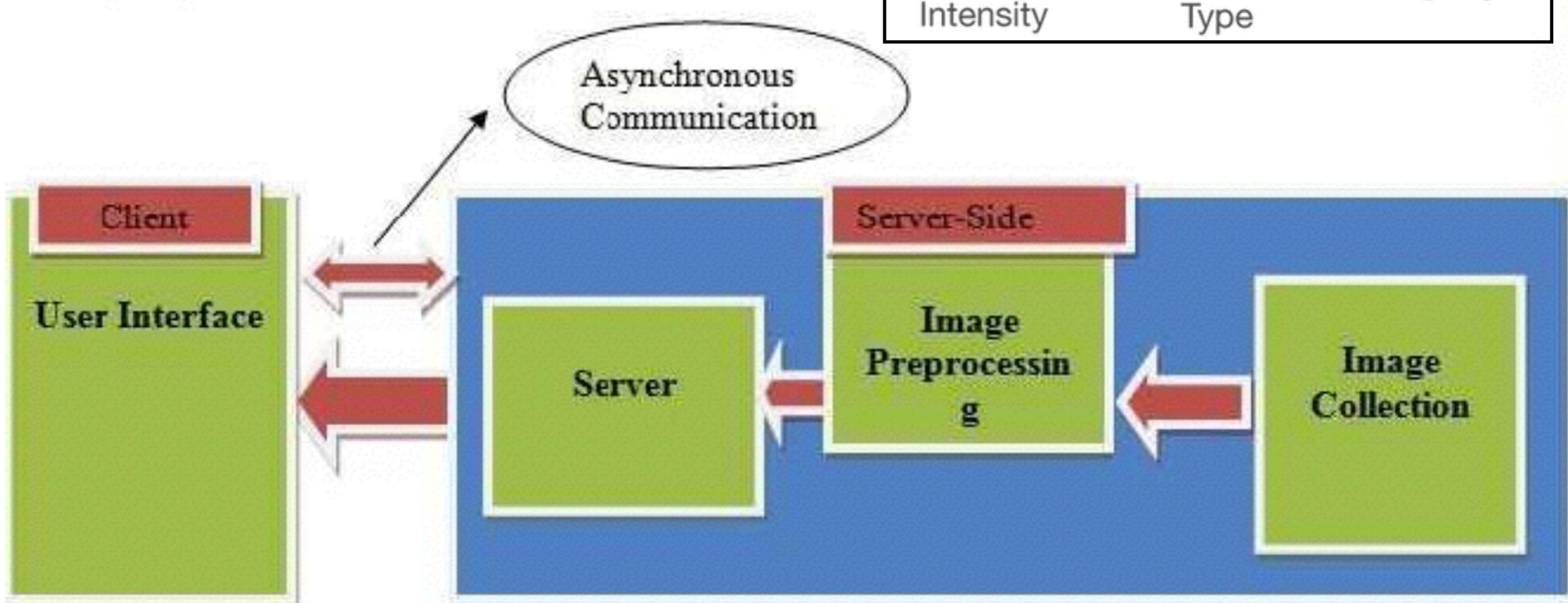
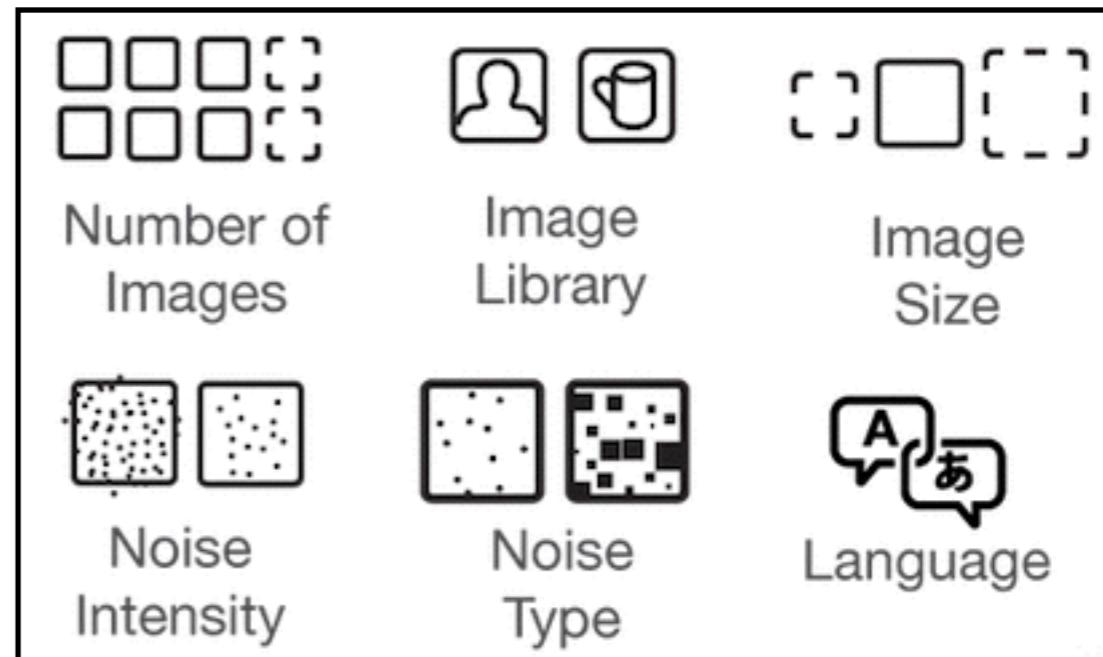
# Text CAPTCHA

## 攻击例子



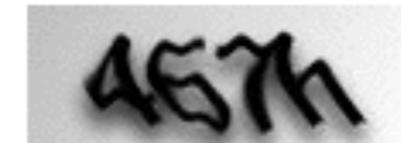
## 设计框架

字符：长度、大小、语  
言、集合、颜色、背景、  
干扰、扭曲、...



# Text CAPTCHA

## 常用技术



a. 百度使用的验证码



b. 开心网使用的验证码

b. 谷歌使用的验证码

粘连

扭曲

干扰线 其余



a. QQ使用的验证码



b. 新浪网使用的验证码

- 不定字符
- 不定字体和大小
- 背景干扰

AUD, IR, MIOR

## Other CAPTCHA

### 声音CAPTCHA



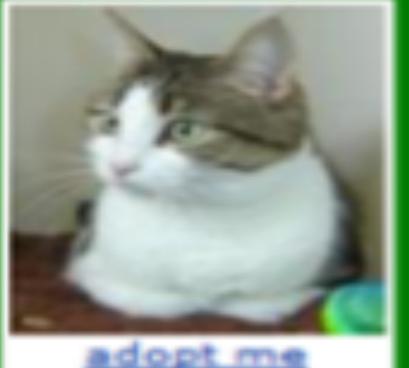
# Other CAPTCHA

## 分类CAPTCHA

Please click on all the images that show cats:

[adopt me](#) [adopt me](#) [adopt me](#) [adopt me](#)

[Score Test](#) **You're a human!**

## Other CAPTCHA

### 分类CAPTCHA

Select ALL the images containing one or more **buildings**

[Powered by Picatcha](#)

**Refresh**

**Submit**

# Other CAPTCHA

## 逻辑CAPTCHA



# Other CAPTCHA

## 拼图CAPTCHA

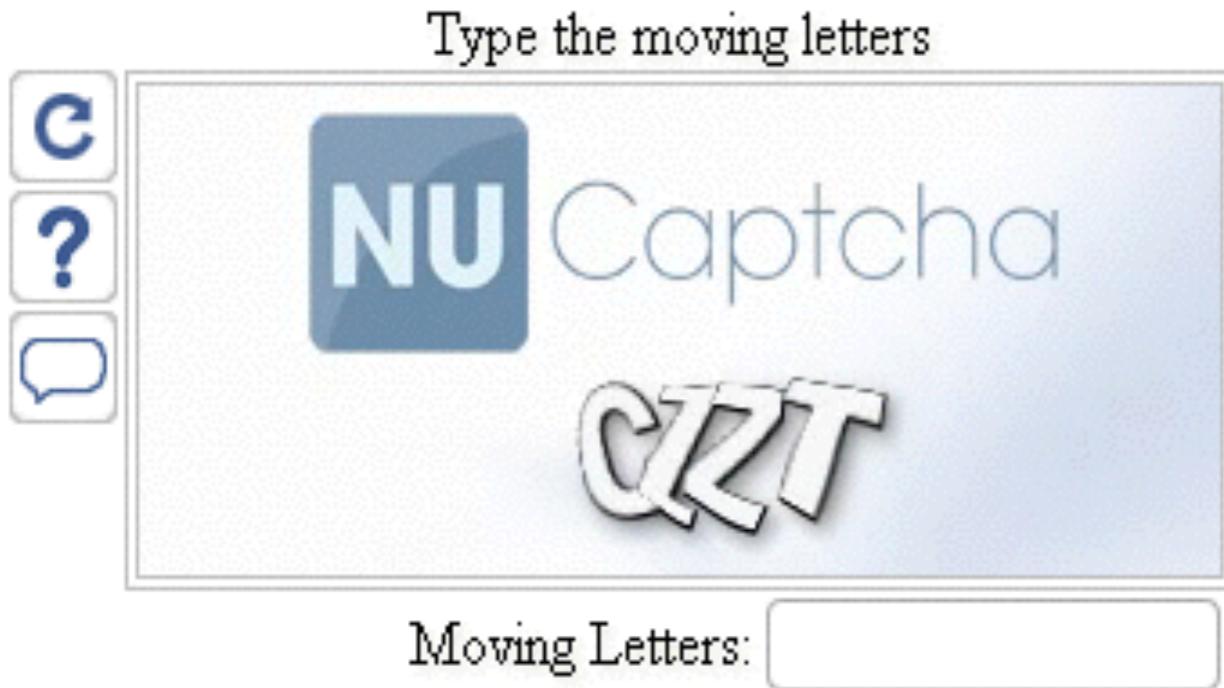


基于拼图



# Other CAPTCHA

## NuCAPCTHA



NU Captcha

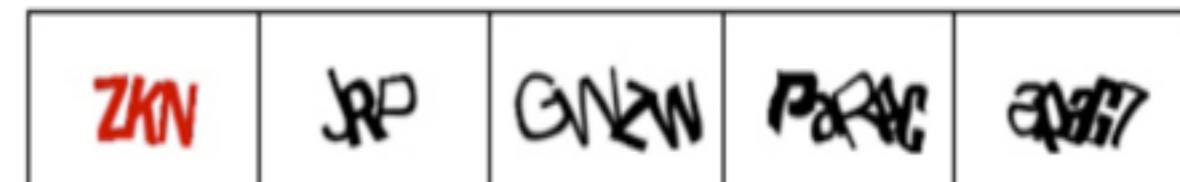
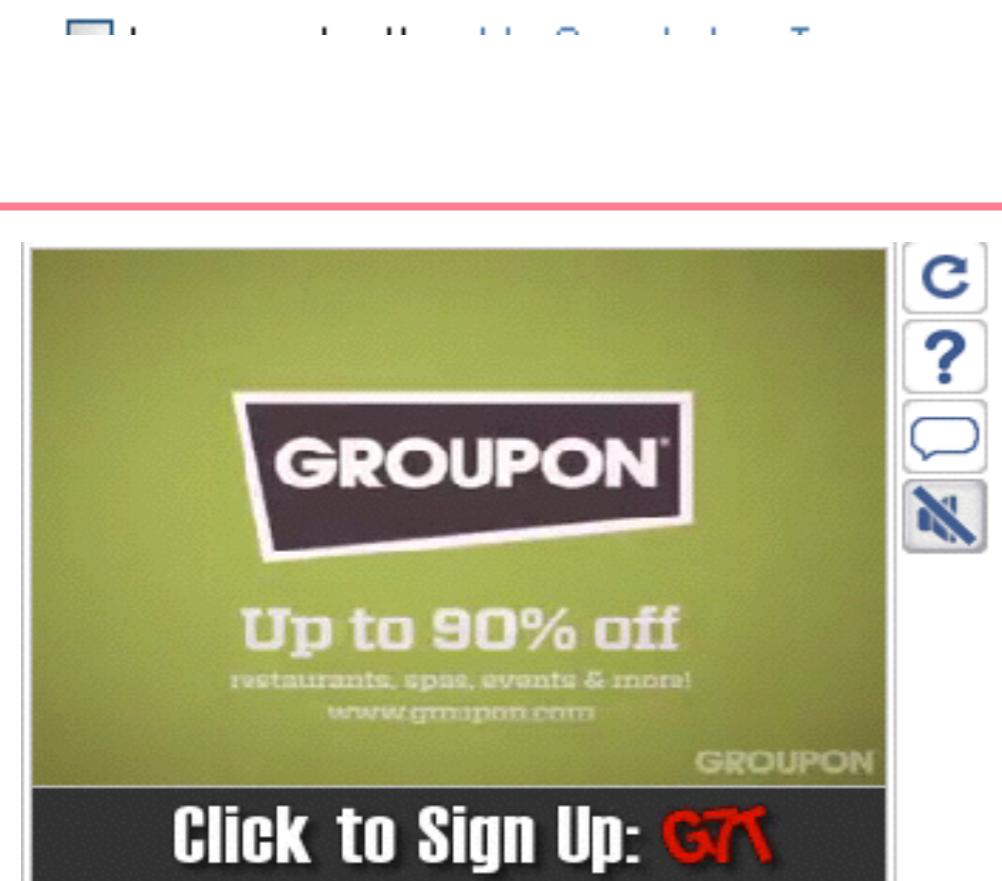


Figure 1: Sample screenshot non animated security settings for NuCaptcha

用户行为分析系统



Moving Letters:

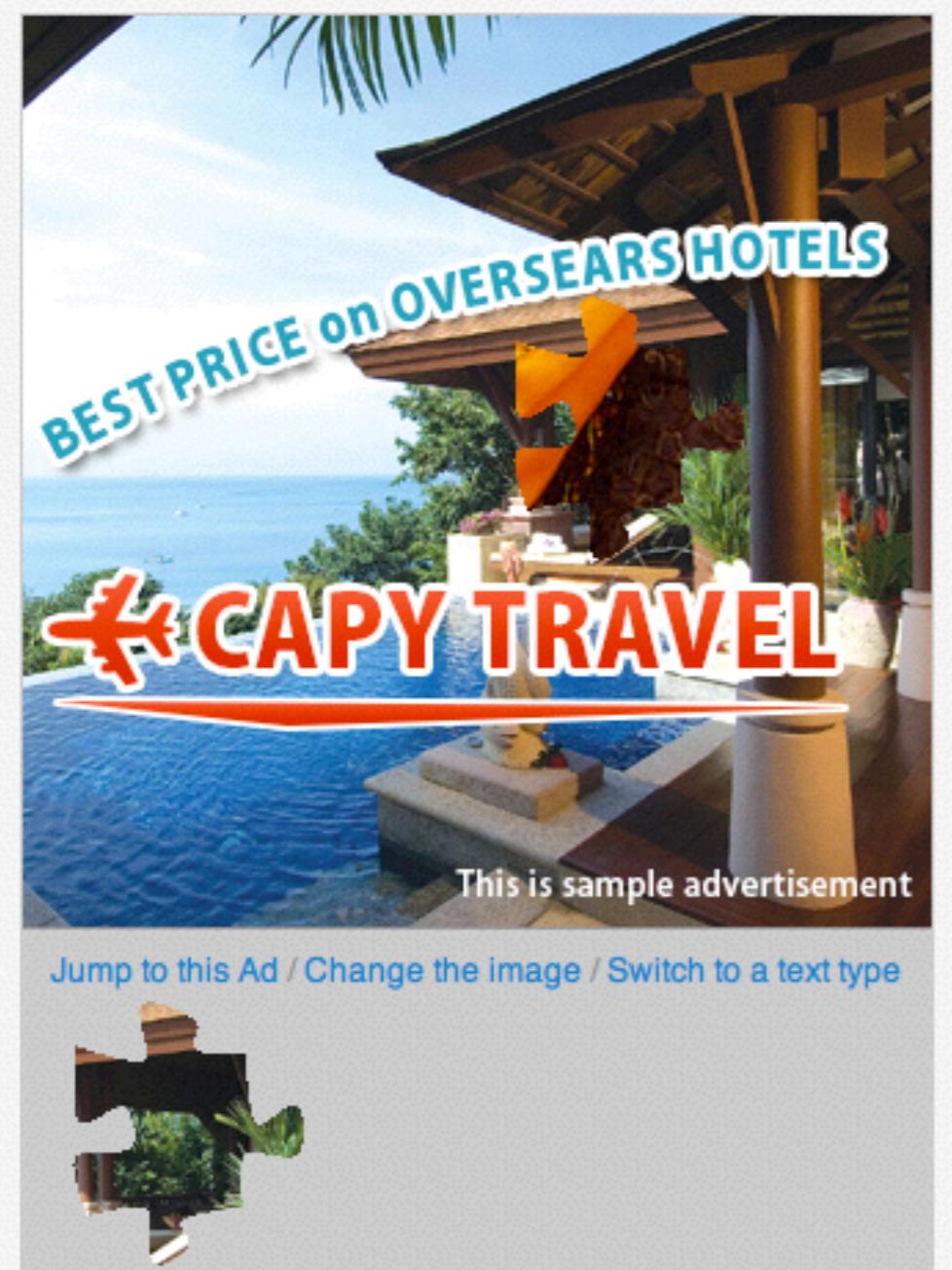


powered by mucaptcha

Submit

# Other CAPTCHA

CAPY

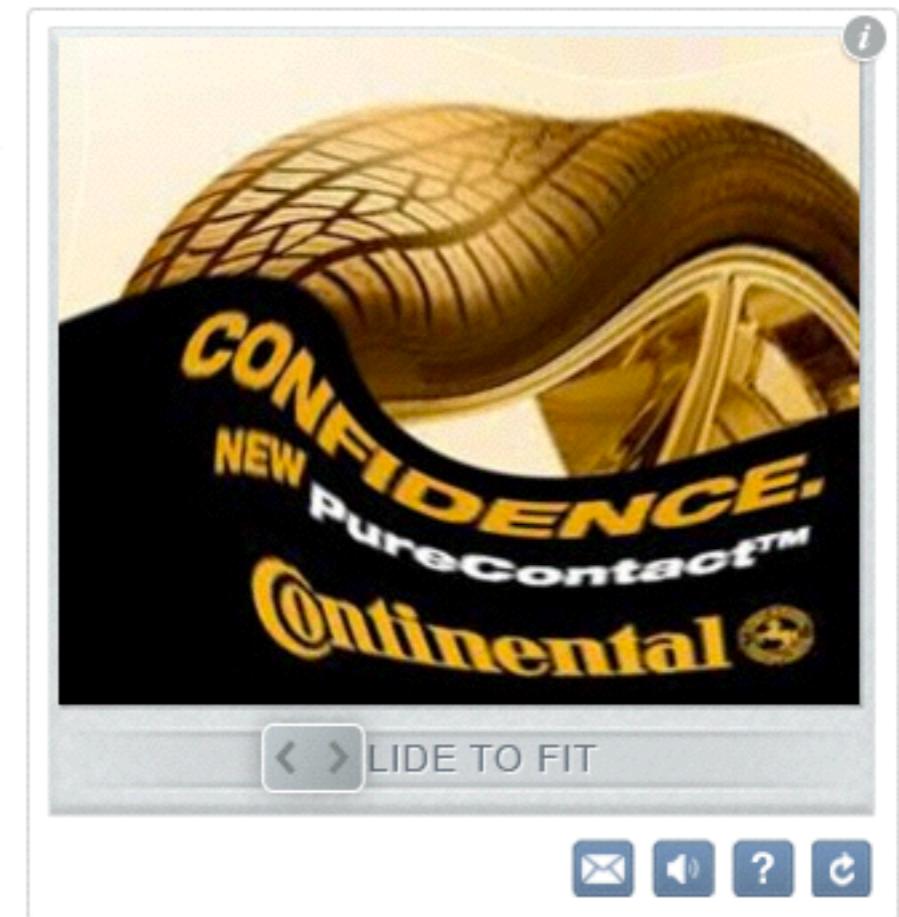
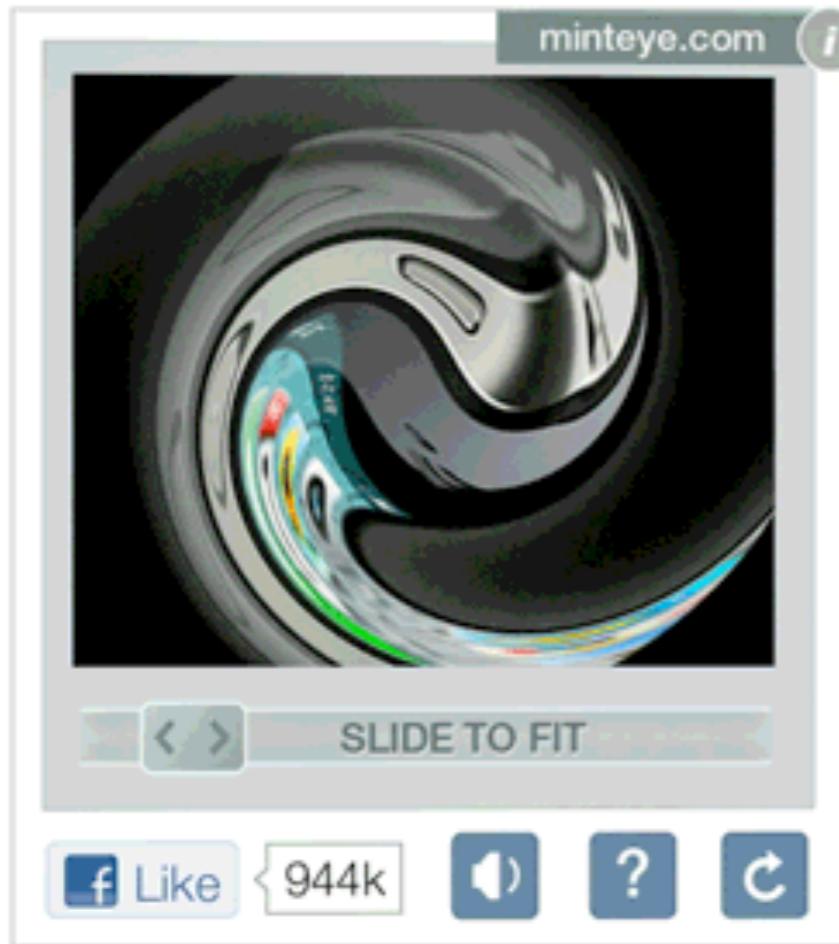


Confirm

Send

# Other CAPTCHA

## Mineye

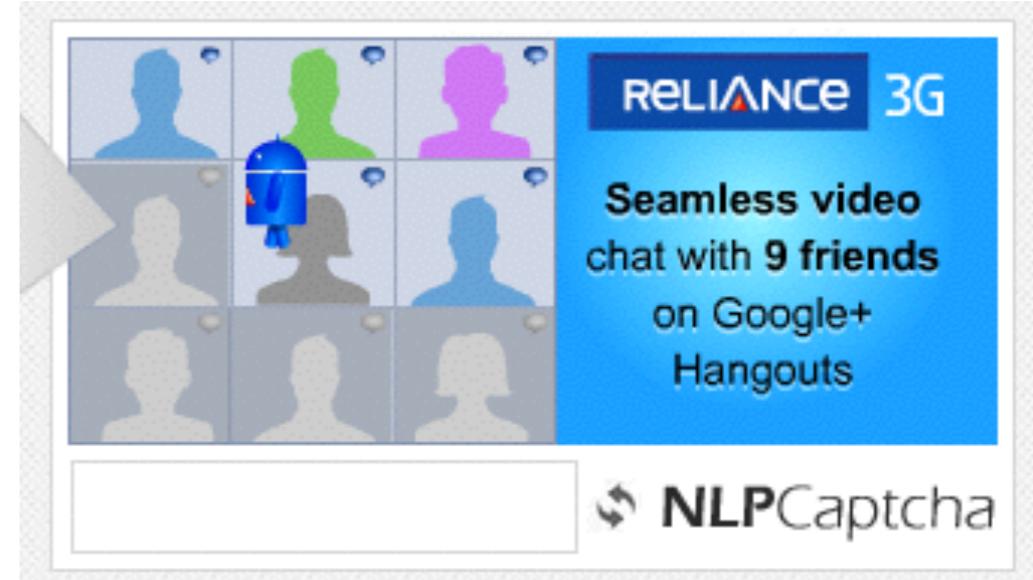
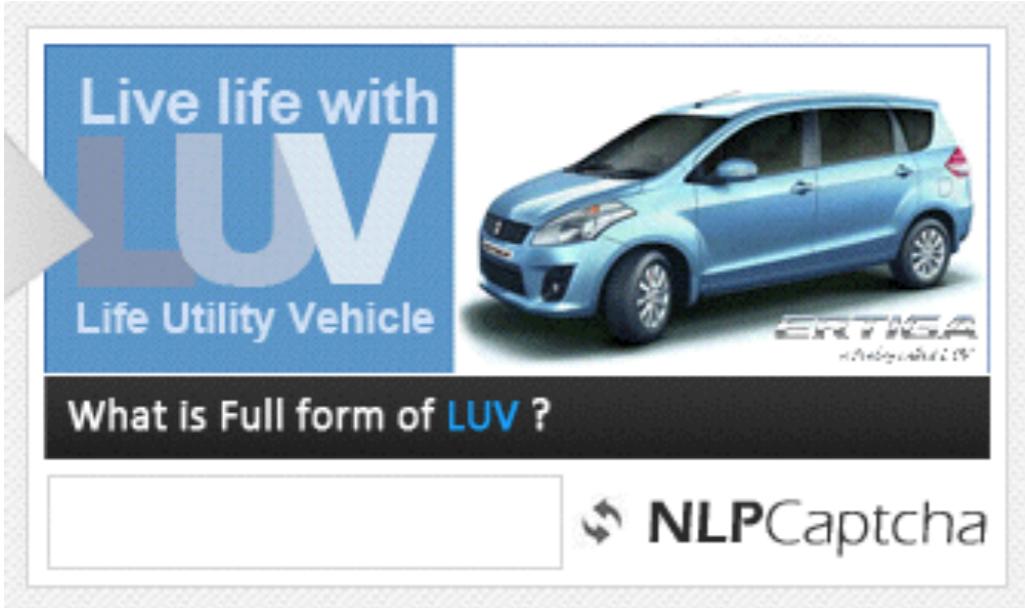


Submit

CAPTCHA  
广告

# Other CAPTCHA

## NLPCAPTCHA



**NLP**Captcha  
Digital Engagement Solutions



# Other CAPTCHA

## AdCAPTCHA



Type the word(s) above

By clicking on 'Create my account now' below, you confirm that you accept the Terms of Service.

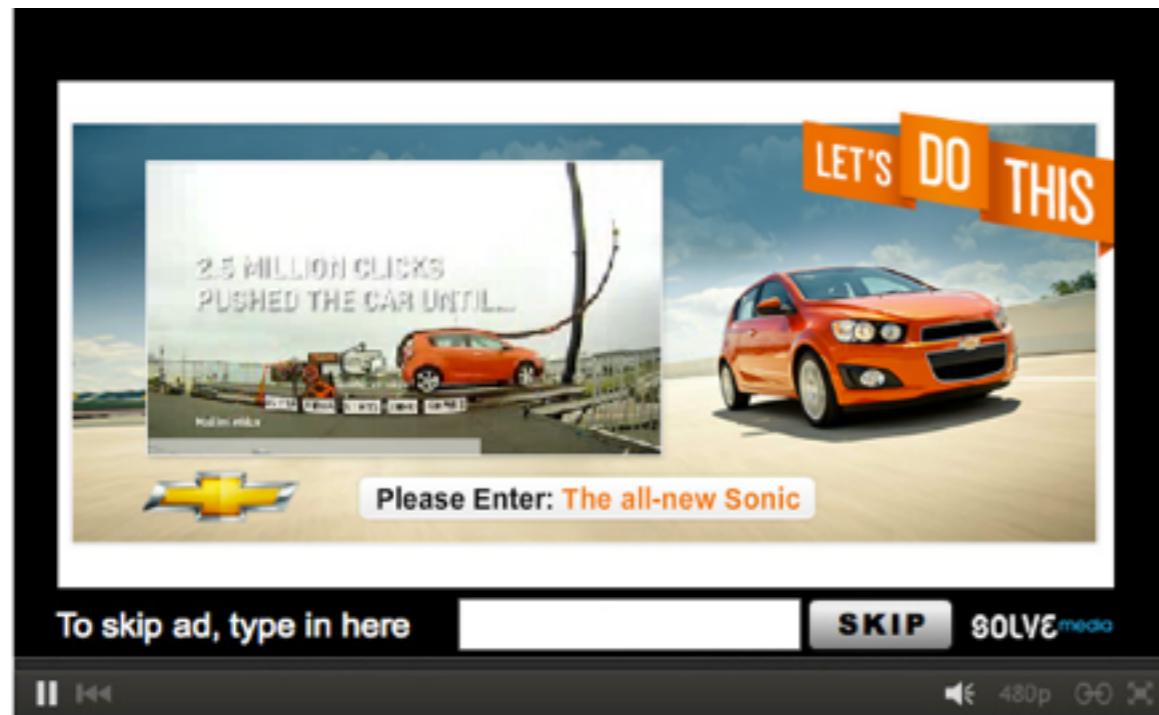
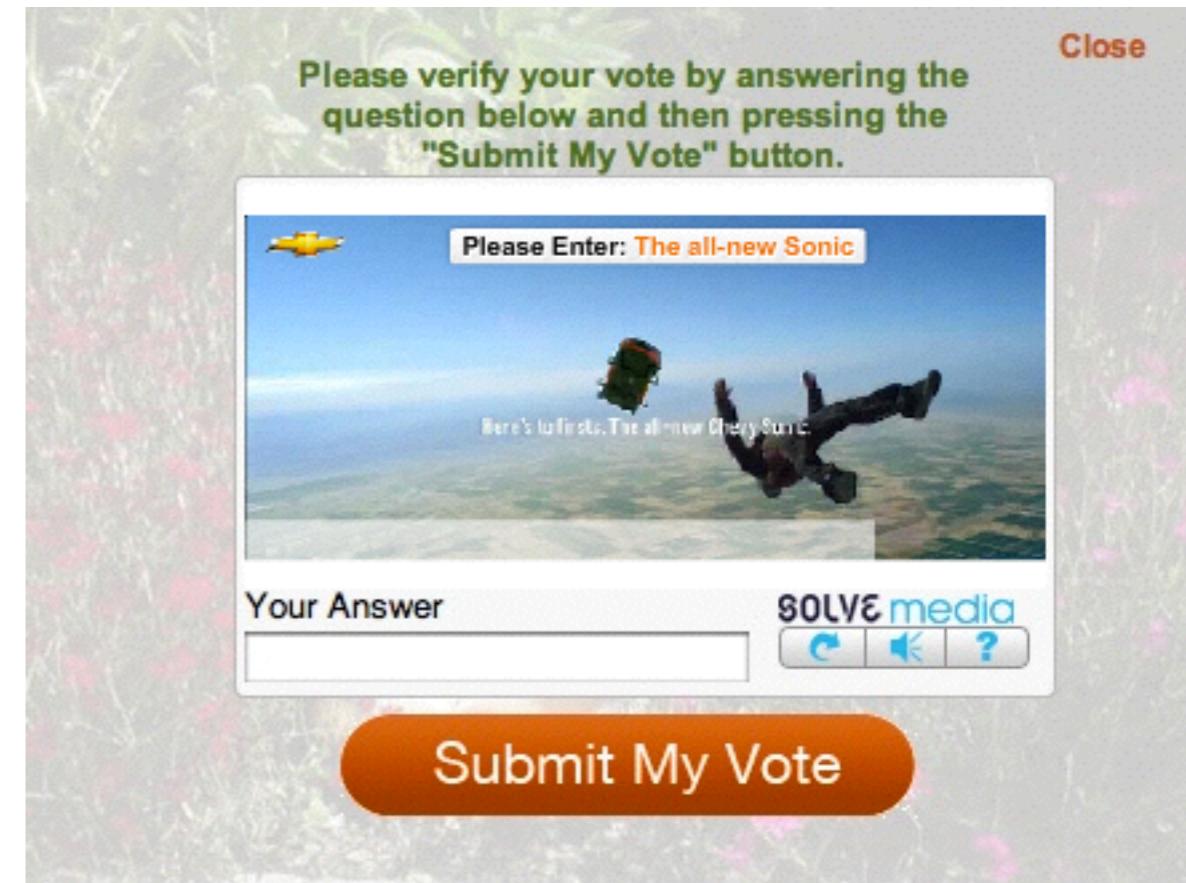
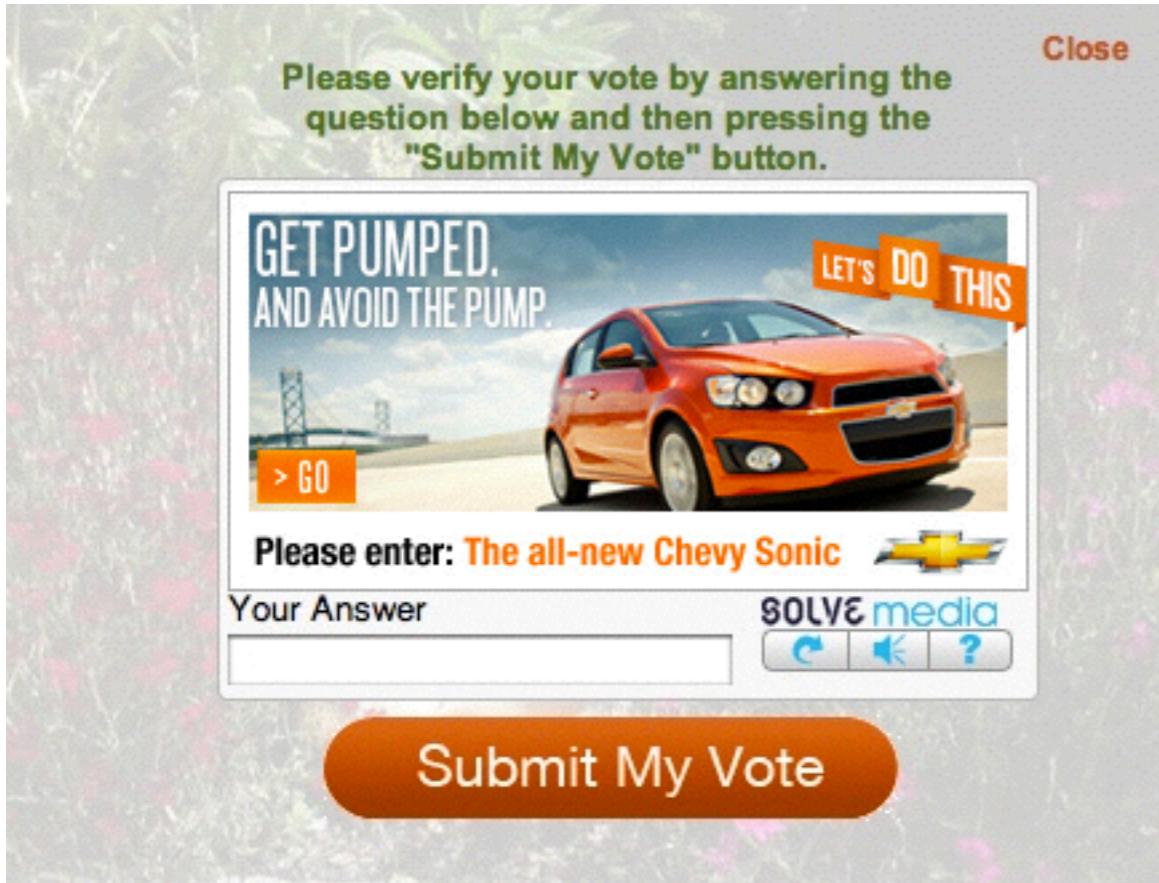
your message

Type the word(s) above

Try it: Submit

# Other CAPTCHA

## SolveMedia



# Other CAPTCHA

## DoubleRecall



The visitor of a website reads your message.

To access the content, type-in the highlighted words.



**Apple iPod** @apple

Movies, TV shows, games, and music. Now playing on an iPod near you.

Now playing iPod |

Get content

[What's this?](#)

DOUBLERECALL

A graphic with a possibility to link to your site.

To access the content he must type-in the words **highlighted** in green.

<http://doublerecall.com/>

## Other CAPTCHA

## YoCAPTCHA



yo! CAPTCHA

YourBrand.com  
**25% OFF**

Get 25% discount on every transaction at [yourbrand.com](http://yourbrand.com)

Please fill the blank based on **the ad** above.

YourBrand is offering  %  
discount on every transaction till February 2012.

?

# Other CAPTCHA

## 印象码

请输入验证码

**NALA** | 化妆品商城 中国最大的化妆品购物网站! [www.nala.com.cn](http://www.nala.com.cn)

可货到付款 30天退货 100%正品

 购物0风险

请输入图中的红色的数字 [点击查看>](#)

换一张 

印象码

**提交**



请输入验证码

平安助您获取  
**无抵押贷款**  
最快一天到账



在线申请

请输入上图中 黑色 的文字

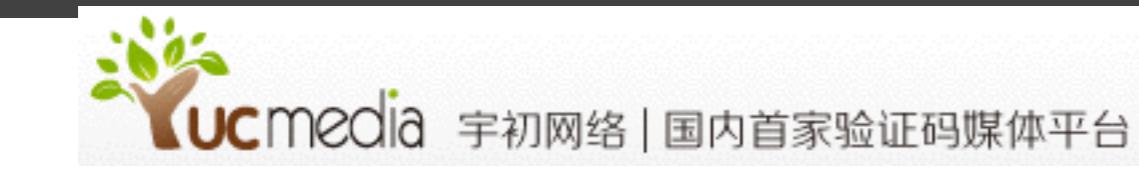
印象码

**提交**

**印象码**  
YinXiangMa.com 中国第一家验证码广告云服务平台

# Other CAPTCHA

## YucMedia



**SONY**  
make.believe

Xperia™ MT27i 摩炫新势力

参与活动

约15mm 浮空触屏技术

请输入图中红色数字



回复状态把握梦想

脉动 CUBE 梦立方

脉动梦立方震撼来袭  
开启**百万**级梦想

百万梦想之旅 抢夺神秘包裹 赢取梦想基金

请输入上图中橙色汉字



**SONY**  
make.believe

Xperia™ MT27i

参与活动

时尚 **NFC** 智触卡

SONY MT27i与你一起智触未来！

请输入图中蓝色字符



随时脉动回来

3999

梦想基金等你拿

开启梦想之旅

请输入上图中橙色数字



# CAPTCHA 和Smartphones

# Smartphones CAPTCHA

## Noise CAPTCHA

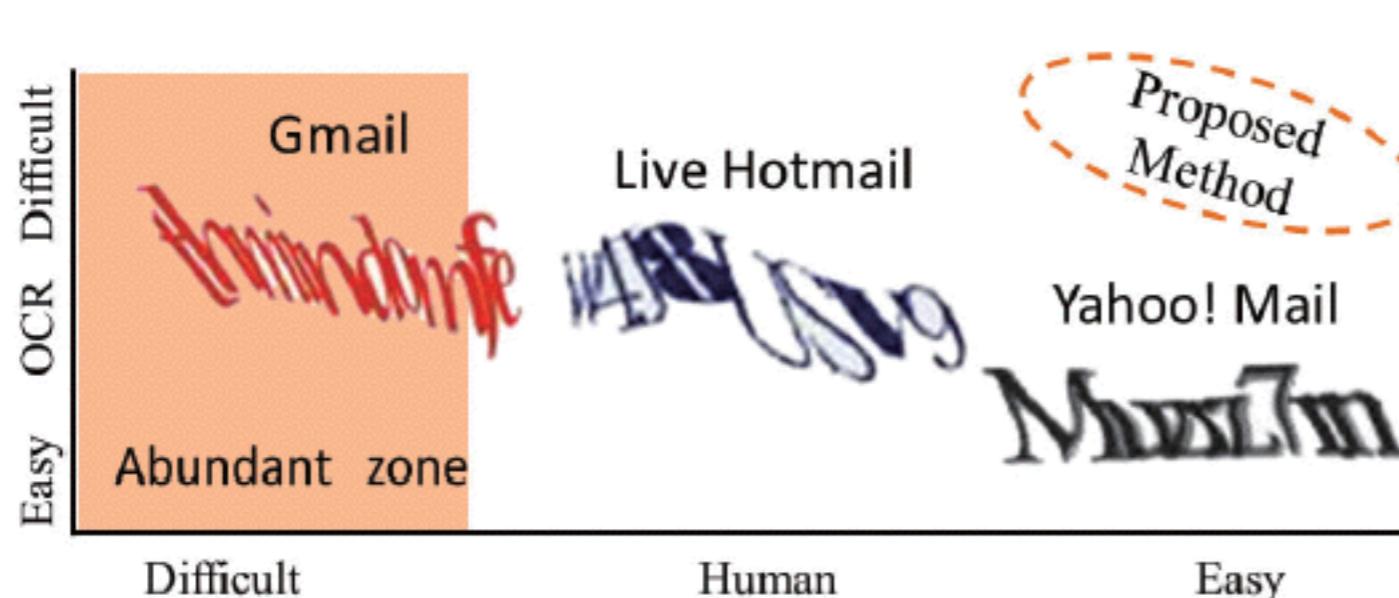


Fig. 2. Gimp images of Google, Live Hotmail, Yahoo mail

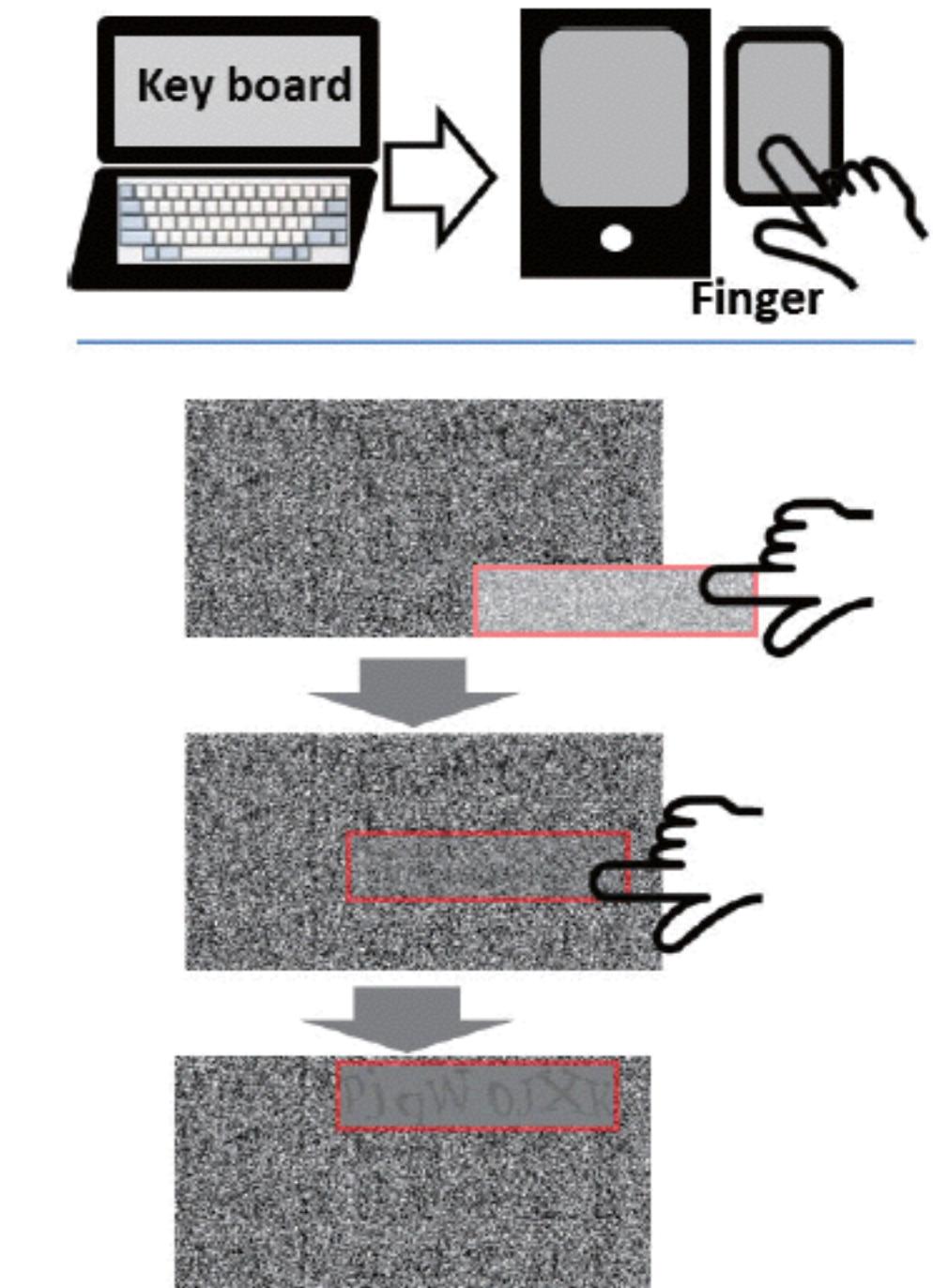
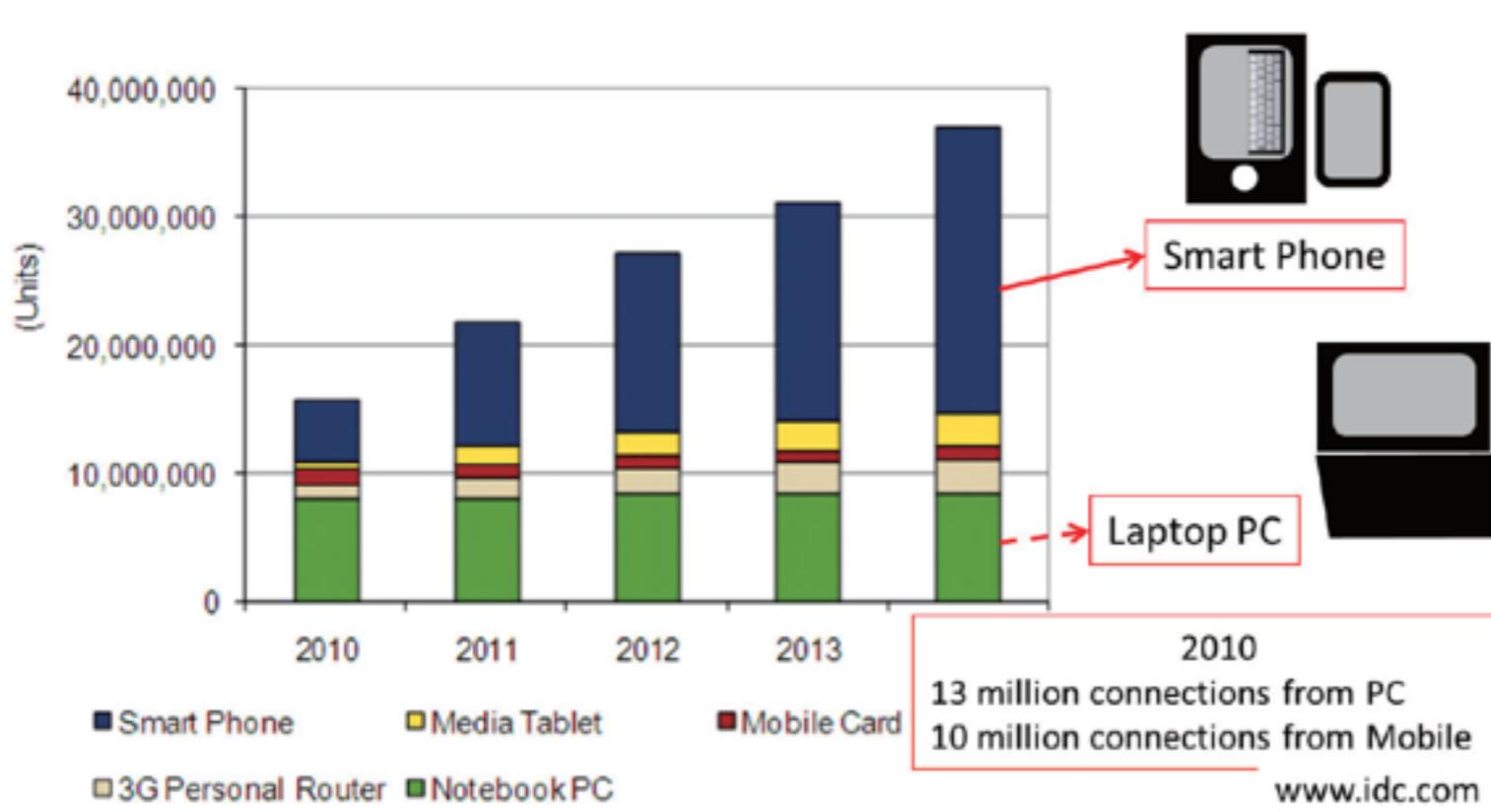
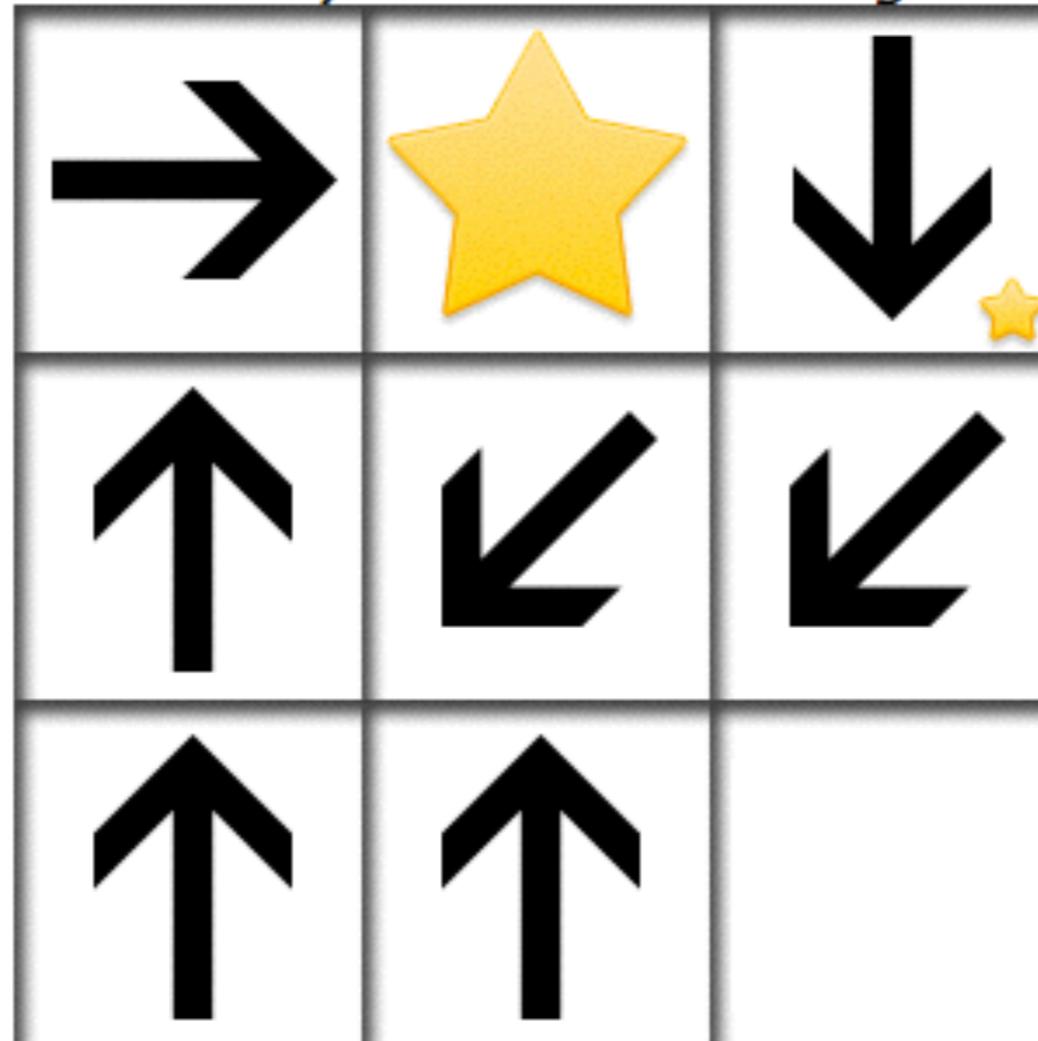


Fig. 5. New CAPTCHA Technique Using Noise Images

# Smartphones CAPTCHA

## Touch CAPTCHA

Click/Tap through the Maze:  
*Start at tiny star and end at big star*



Steps taken: 0/8 [Start Over](#)

[Test](#)

# Smartphones CAPTCHA

其余



(a) Clickable captchas [36]



(b) What's UP captcha [59]



(c) Captcha zoo [81]

Enter the correct characters.

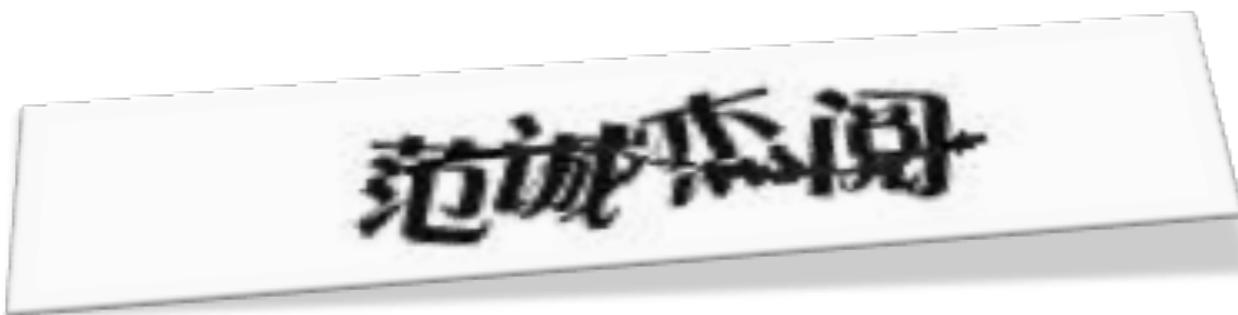


(d) Lin's et al. captcha for mobile [81]

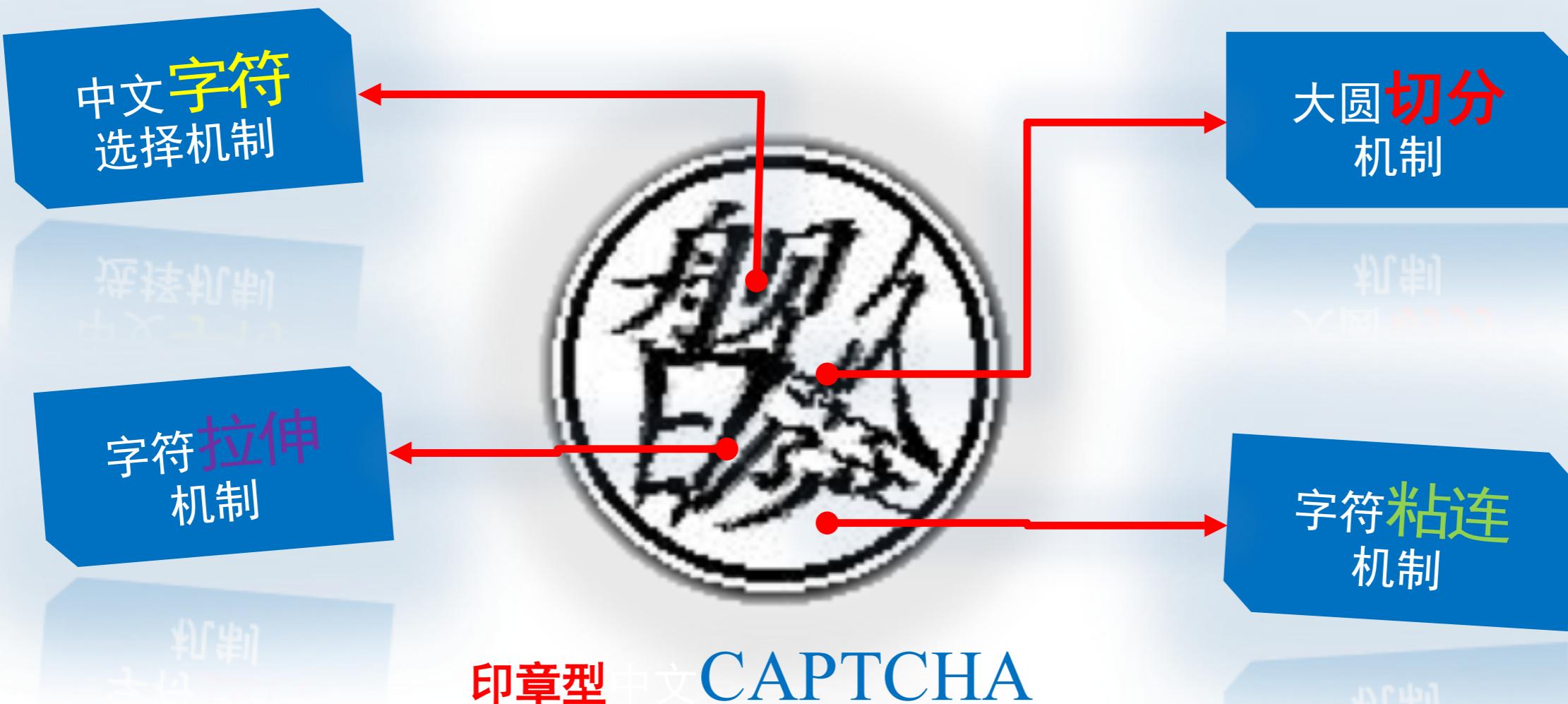
STAMP  
CAPTCHA

Stamp  
CAPTCHA

中文CAPTCHA

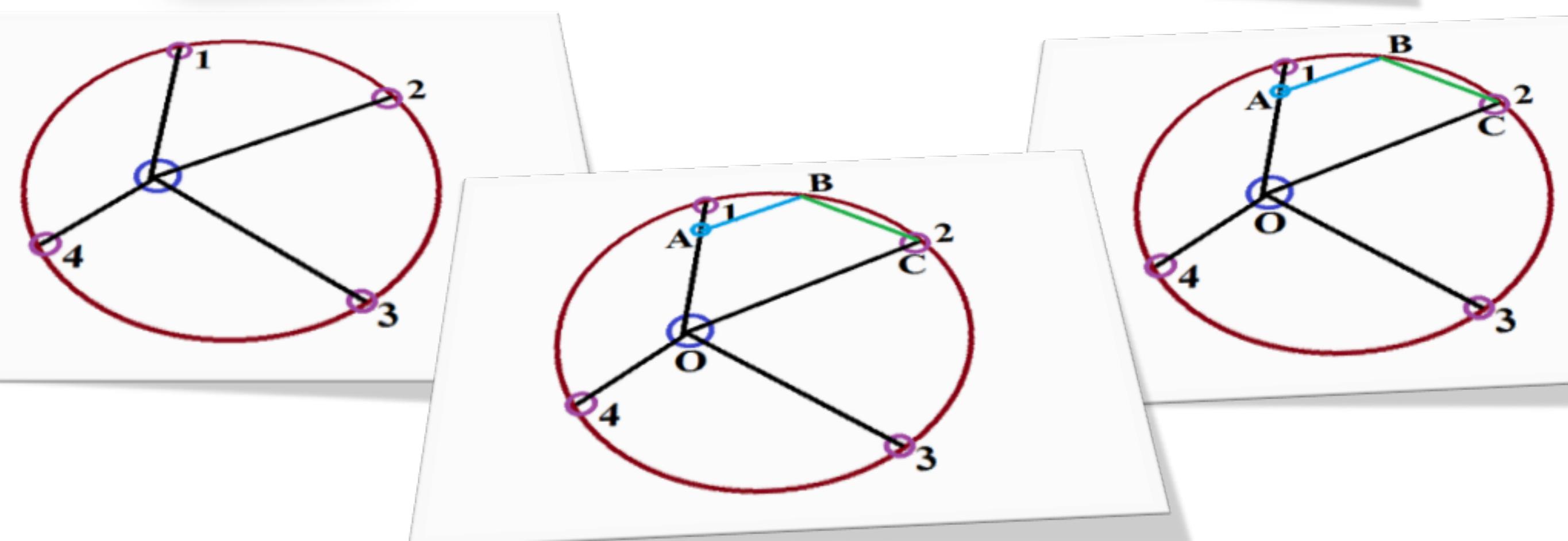
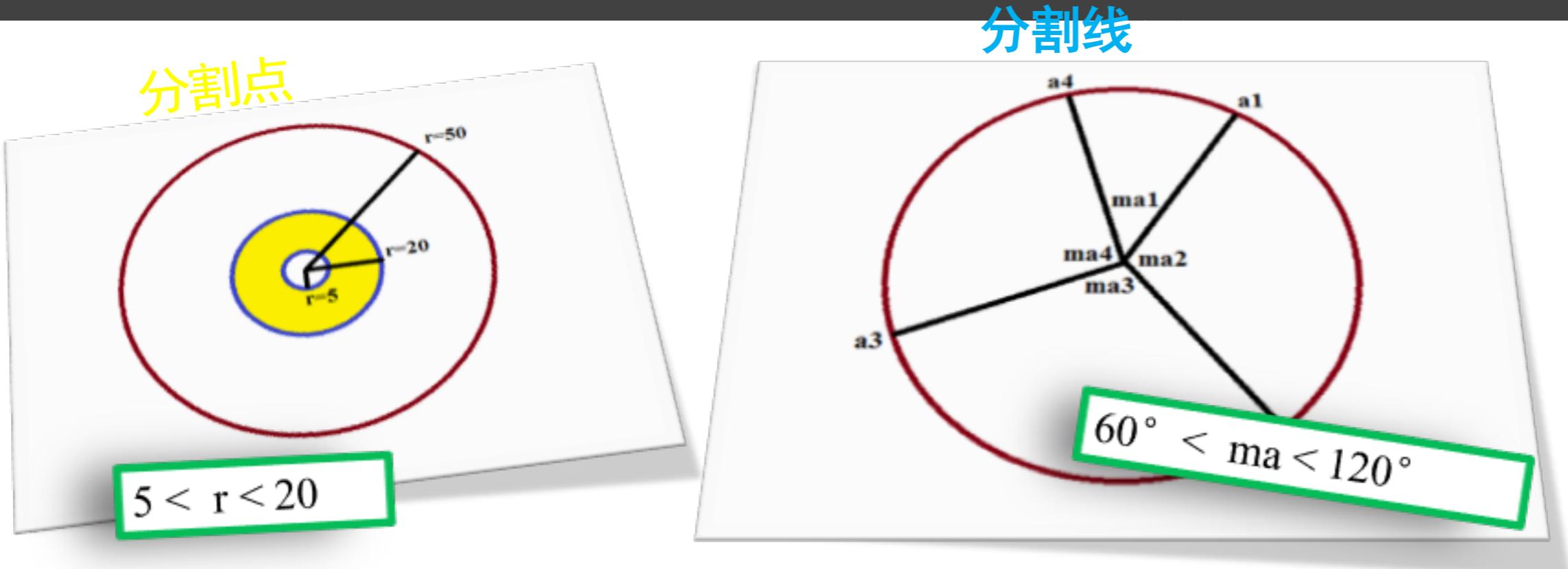


# 设计思想



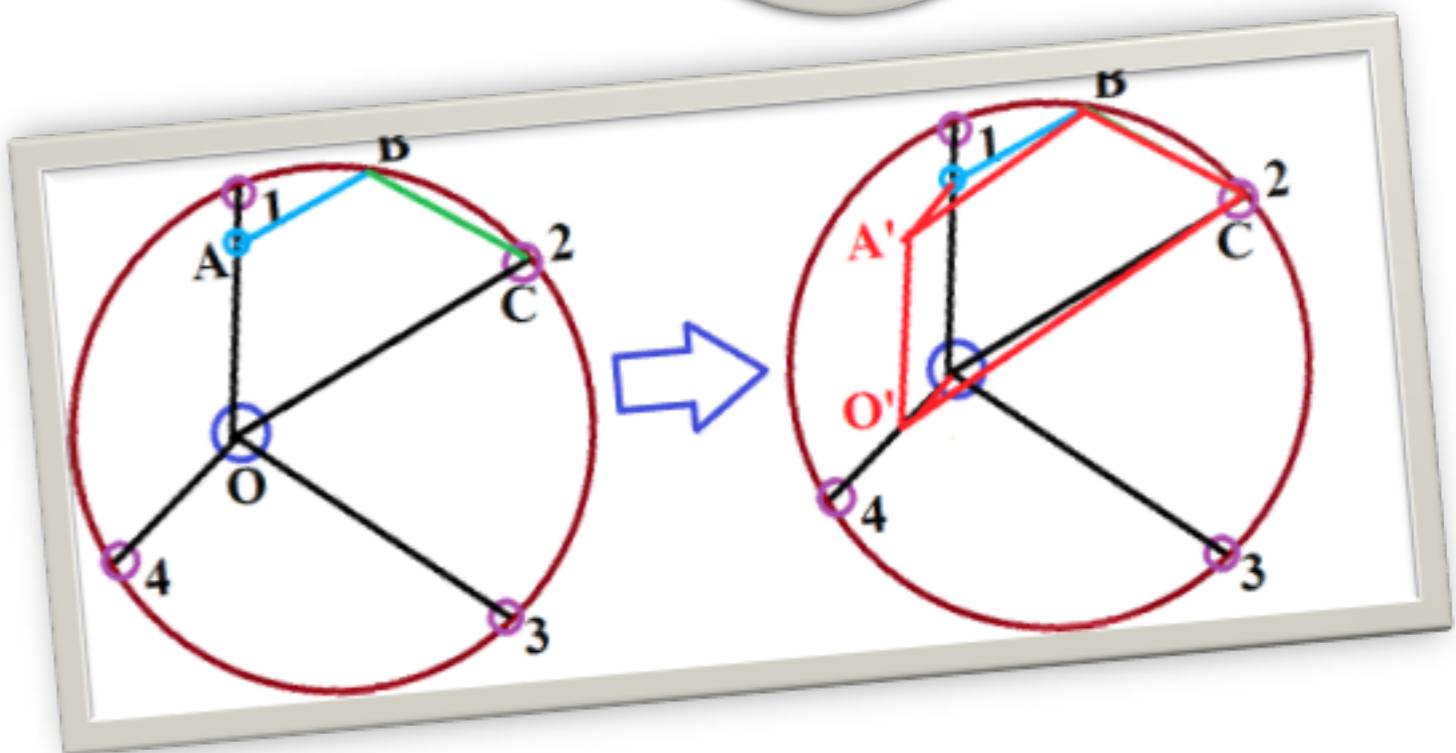
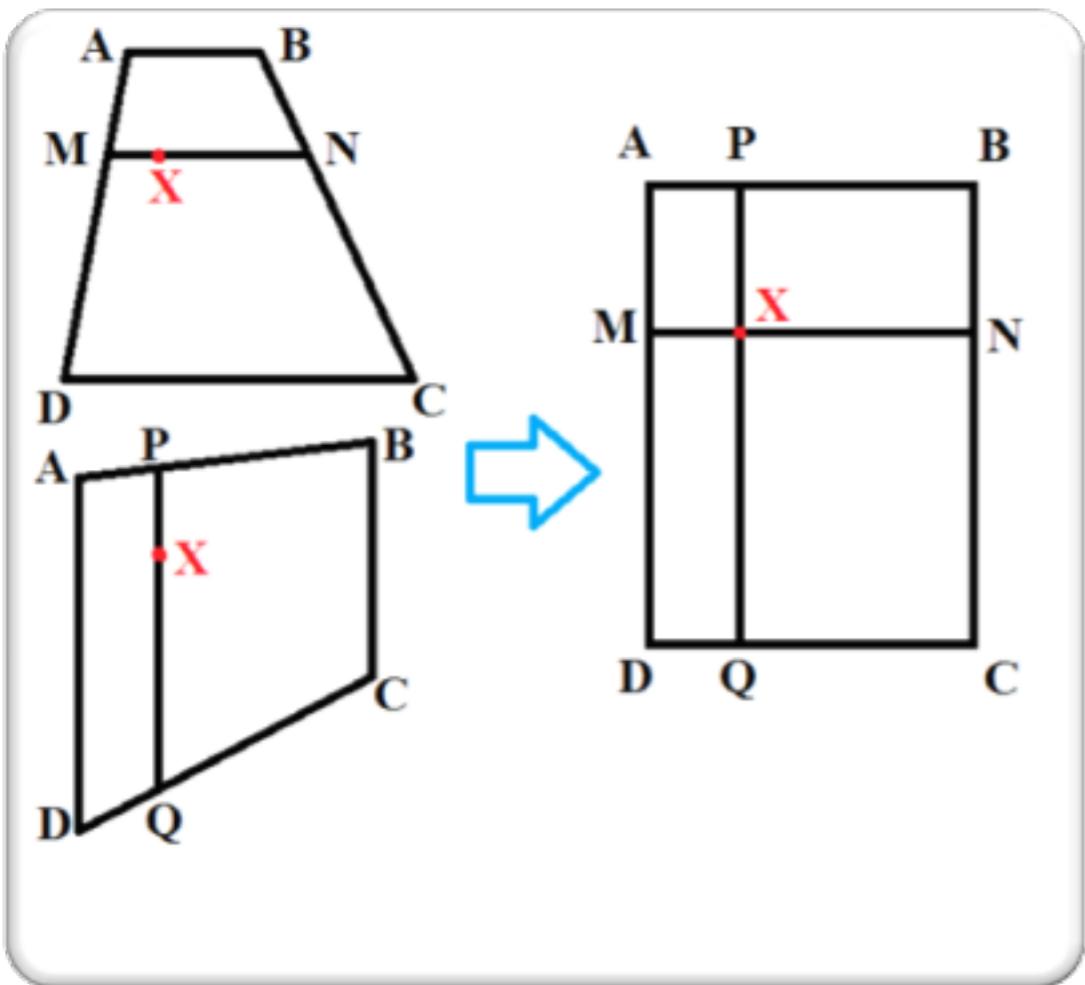
# Stamp CAPTCHA

## 主要机制



# Stamp CAPTCHA

## 主要机制



谢谢！

孙惠平

[sunhp@ss.pku.edu.cn](mailto:sunhp@ss.pku.edu.cn)