



北京大学软件与微电子学院

极验滑动验证码破解

组员：张嘉豪 王源 龚秋嫦 曹辉

2018-12-3



北京大学



目标：虎嗅网（极验）滑动验证码破解



18/12/10



北京大学



意义

- 一些网站会在正常的账号密码认证之外加一些验证码，以此来明确地区分人机行为，从一定程度上达到反爬的效果，滑动验证码最典型的要属于极验滑动认证了
- 现在极验验证码已经更新到了 **3.0** 版本，截至 **2017** 年 **7** 月全球已有十六万家企业正在使用极验，每天服务响应超过四亿次，广泛应用于直播视频、金融服务、电子商务、游戏娱乐、政府企业等各大类型网站
- 对滑动验证码的破解能够在爬虫自动化等方面提高效率有很大帮助



实现思想

- #步骤一： 点击按钮，弹出没有缺口的图片
- #步骤二： 获取步骤一的图片
- #步骤三： 点击滑动按钮弹出带缺口的图片
- #步骤四： 获取带缺口的图片
- #步骤五： 对比两张图的所有**RGB**像素点，得到不一样像素点的值 x ，即要移动的距离
- #步骤六： 模拟人的习惯,把需要的拖动的总距离分成一段一段小的轨迹
- #步骤七： 按照轨迹拖动，完成验证
- #步骤八： 完成登录



实现工具

- 结合使用python中的selenium模拟滑动实现破解。
- **selenium**是浏览器自动化测试框架，用来模拟鼠标点击拖动过程。



查看图形网页源码

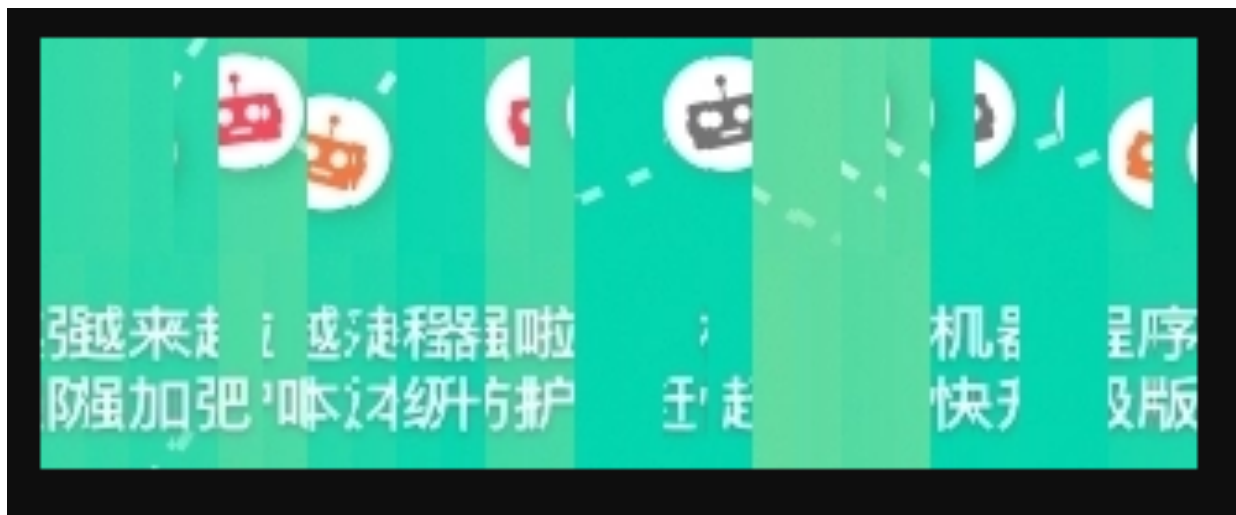
```
▼<div class="gt_cut_fullbg_gt_show">
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -157px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -145px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -265px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -277px -58px;"></div> == $0
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -181px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -169px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -241px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -253px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -109px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -97px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -289px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -301px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -85px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -73px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -25px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -37px -58px;"></div>
  <div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/6bcc3e453/6bcc3e453.webp"); background-position: -37px -58px;"></div>
```

图片+位置





访问图片链接结果





分析过程

首先我们要手动将上述打乱的图片按照正确的顺序拼接起来，这样才能计算图片缺口位置然后获得移动距离

位置坐标里，分别是 x 和 y ， y 轴只有58和0，图片分为上下两部分，div的数量是26块，每一块宽10x高58。整个图片的宽就是260，高116。



分析过程

- 随便找一个特征点，查看元素，看它定位到哪个div元素那里，然后再看后面的位置。
- 选一些位置明显的地方，比如中间，或者边界。

第一个div

```
<div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/0f340b2ba/0f340b2ba.webp"); background-position: -157px -58px;"></div> == $0
```

最后一个div

```
<div class="gt_cut_fullbg_slice" style="background-image: url("https://static.geetest.com/pictures/gt/8cb25b9af/8cb25b9af.webp"); background-position: -205px 0px;"></div> == $0
```





分析过程

```
▼<body style="margin: 0px; background: #0e0e0e;">  
 == $0
```

分析发现，url里的图形（打乱的）比实际大，每个小块+12作为下一个小块的起点。这样的话左右各去掉一个像素，宽度就是10了。而且每个小块是12，26个是312，跟我们看到的拼图大小差不多，按照元素里提供的坐标，取宽度为10的大小即可。



分析过程

- 最终的图片就是把拼图，按照 $x=157$ 、 $y=58$ 、 $w=10$ 、 $h=58$ 截取出来，放在上半部分第一个位置， $x=145$ 、 $y=58$ 、 $w=10$ 、 $h=58$ 截取出来放在上半部分第二个位置，紧挨着第一个，以此类推，拼成一张整图。



计算缺口位置

- 计算两张图片不同位置

ImageChops.difference接口

缺点：有阴影

解决：为两张图片的对比设置容差

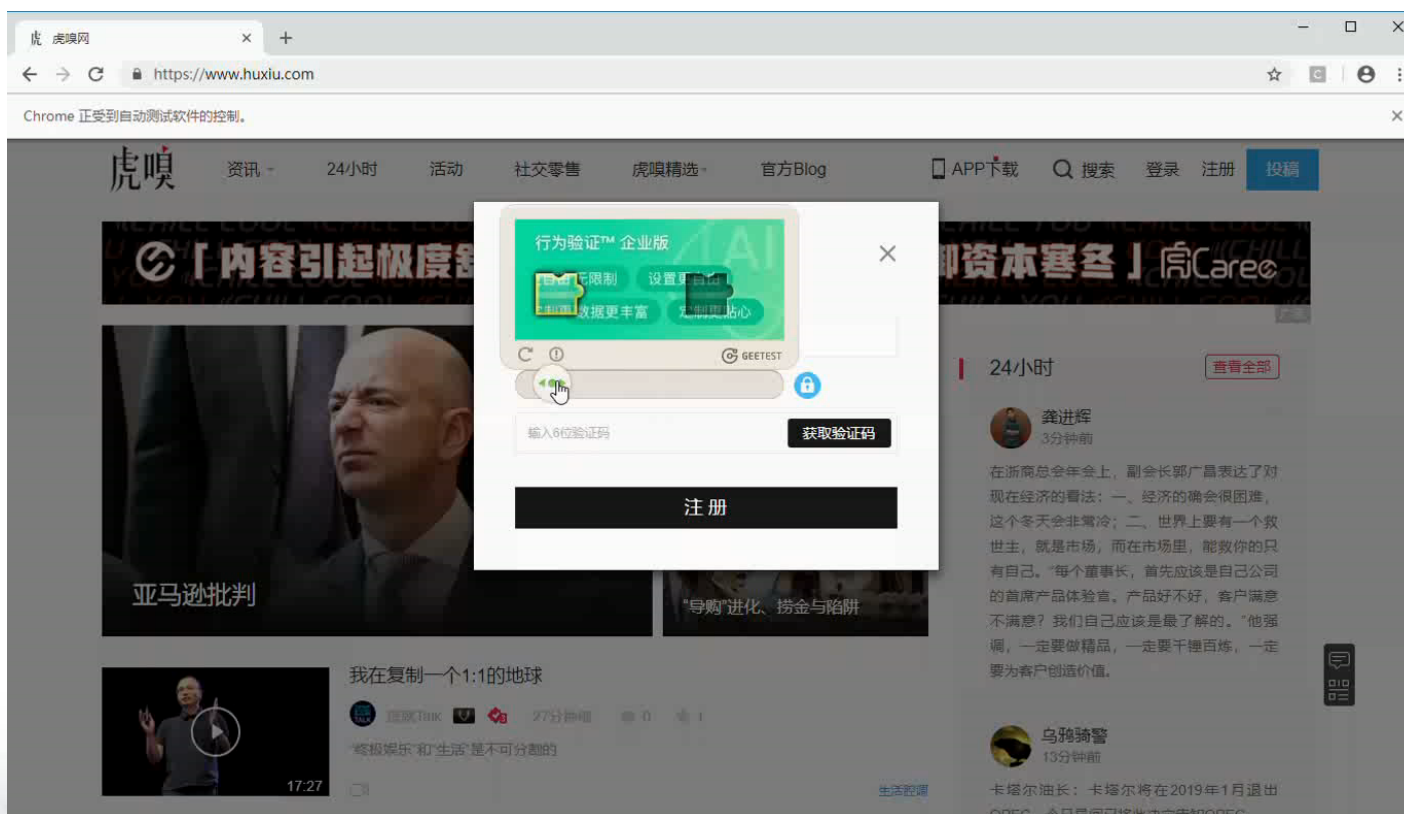


如何确定色差？

获取多张图片，全图和缺陷图，然后使用取色工具，取对应位置的颜色值，确定一个大概范围。



演示



18/12/10



北京大学