Cracking Captcha Project Overview

Team Members

| Student Name | Task |
| --- | --- |
| Roberto Guzman | Research and Development |
| Leila Garcia | Research and Development |
| Cameron Jackson | Research and Development |
| Courtney Ingram | Research and Development |

*Objective:* Investigate the current intent, design, and functionality of captcha mechanisms for detecting bots and other automated activities. Research successful exploits and attacks against captcha mechanisms within the last 10 years and what type of counter-captcha research is trending within the last 2 years. Based on the results of the research, develop a proof of concept tool which can identify a captcha request in a web browser, parse the prompt, isolate any images, utilize AI, i.e. computer vision, to identify the correct images, and automatically respond.

Project Assignments

1. [CCA-1] Investigate current intent, design, and functionality of captcha mechanisms.
2. [CCA-2] Investigate the types of systems where captcha mechanisms are deployed and relevant metrics (success/failure, impact on throughput, etc.).
3. [CCA-3] Research how captcha mechanisms are designed and new trends that have arisen.
4. [CCA-4] Research attacks and exploits against captcha mechanisms within the last 10 years, paying special attention to any resources documenting the implementation of countermeasures from the last 2 years.
5. [CCA-5] Research options for parsing the contents of web pages and identifying when a captcha is encountered.
6. [CCA-6] Survey current automation tools for interacting with the browser, e.g. clicking, typing in a box, etc.
7. [CCA-7] Determine how best to extract information from the webpage and classify the type of captcha request.
8. [CCA-8] Research options for using pre-trained computer vision models to review the captcha images and classify them according to the captcha request.
9. [CCA-9] Develop a software solution which ties the web browser interface, computer vision, and automation components together in a single solution.
10. [CCA-10] Create a test website with a captcha implementation for testing the proof of concept solution against.
11. [CCA-11] Generate documentation of any code created/used for the solution and a user’s guide.
12. [CCA-12] Generate documentation on the major hurdles to overcome in developing captcha countermeasures.
13. [CCA-13] Review your proof of concept solution from a defender’s point of view. How would they detect and prevent your solution from functioning correctly?

Deliverables

1. Update slides at each weekly meeting (6-8 total)
2. Research Findings Document
   1. How are captcha designed and implemented?
   2. Where are captcha commonly used and what are the statistics for their success and system/business impact?
   3. How have captcha mechanisms been defeated?
3. Proof Of Concept Solution
   1. Web interface
      1. Identify that a captcha mechanisms has been encountered
      2. Parse the mechanisms to determine what kind/how to interact
      3. Extract and isolated any prompted images
   2. Automation
      1. Based on findings of the computer vision model, provide appropriate keyboard/mouse actions
      2. Scrape the contents of the webpage to support identification of captcha mechanisms and any useful context
   3. Computer Vision
      1. Utilize existing, pre-trained image detection model to classify images from web browser using context from the prompt
      2. Provide automation element necessary information to allow for “clicking” of images or typing of text
4. Final Report
   1. Research survey
   2. Code documentation and User’s guide
   3. FAQ about design process and proof of concept solution