POSE A THREAT

HOW PERCEPTUAL ANALYSIS HELPS BUG HUNTERS

PRESENTED BY ROB RAGAN AND OSCAR SALAZAR
LET'S SEE SOME HANDS

WHO HERE IS

A PENTESTER?
QUESTION
WHO HAS HAD THEIR AWS CREDENTIALS STOLEN?
WHAT IS ELASTICBEANSTALK
5 AVENUES OF ATTACK

- Weak Passwords
- Insecure Applications
- Missing Patches
- Single Misconfiguration
- Sensitive Info Leaks
What did we observe?

What's in the list of targets?

88,022 total targets
30 minute total labor time
880,220 total requests
52,586 responses & screenshots
4,007 filtered for review
321 confirmed exposures
10 total paths
AWS CREDENTIALS: 64
DATABASE CREDENTIALS: 44
PAYMENT PROCESSING KEYS: 22
APP KEYS: 36
SMTP CREDENTIALS: 120
REDIS CREDENTIALS: 178
SLACK TOKENS: 01
SOCIAL MEDIA TOKENS: 37
ELASTIC SEARCH: 13
SOURCE CODE LEAK
CRITICAL EXPOSURE

/.git/config

Source Code Contains

- Encryption Keys
- DB Passwords
- API Tokens
- Internal Systems
- Application Vulnerabilities
CONTENT DISCOVERY 10 YEARS AGO
EVLVING OLD WAYS

OWASP DirBuster (2008)

Multi-threaded Crawling

Wordlist Path Brute-forcing
Directory
File
CONTENT DISCOVERY FAILS
EVOLVING OLD WAYS

/thisdefinitelydoesnotexist

Use “Fail Cases”
- Soft 404 Checking
- DOM Difference Analysis
- Content Length
CONTENT DISCOVERY FAILS
PAINFUL + TIME CONSUMING

/.docker/variables.env

Manual Review
Load Times
SSL Errors
Human Errors
LIKE FINDING A

NEEDLE IN A

HAYSTACK
PROCESS PROBLEMS
CONTENTS DISCOVERY 2008 - 2018

INCONSISTENCY – Actions performed pertaining to content discovery during assessments are inconsistent despite the value proposition. It’s not feasible to run large-scale comprehensive dictionaries on every target in a small time window using traditional techniques.

DIVERGENCE – Information storage is disparate as new targets are discovered throughout an assessment. Penetration testers often work off of separate datasets and ensuring all targets were reviewed the same way is problematic.

EFFICIENCY – Results validation is time consuming when false positives are too numerous. Too much noise and not enough signal. Dictionaries that are used are often outdated or inefficiently utilized.
CONTENT DISCOVERY + SCREENSHOTS

Evolving Old Ways

Pre-render Applications
Approximately 100 per minute
Tag for further review

Goal
Find Content to Attack
Sensitive Information
HUMANS ARE NOT THE SOLUTION
WHAT IF THERE WERE A WAY
TO GET THE
SAME RESULTS...
404 CHECKING & SIGNATURES...
QUESTION

...AT

SCALE?
EXPLORING
SOLUTIONS FOR HUMANS
GOAL

MAKE IT EASY FOR HUMAN

TO GET EYES ON INTERESTING SCREENSHOTS FROM URL PATH BRUTE-FORCE
HOW PERCEPTUAL ANALYSIS HELPS BUG HUNTERS

AUTOMATED TRIAGE

DISCOVER
Discover hostnames through OSINT and subdomain enumeration

OPEN SERVICES
Confirm open ports through continuous probes on a daily basis

REQUEST PATHS
Brute-force paths to discover unknown content

DETECT EXPOSURES
Display screenshots by visual similarity and disregard similar matches

IDENTIFY OUTLIERS
Perform perceptual analysis and OCR screenshot then save to meta data

CAPTURE SCREENSHOT
OWASP Amass (2018)

Open Source Intelligence (OSINT)

- Scrape web pages with DNS dragnet data
- Aggregate Passive DNS API data
- Crawling internet archives
- Recursive brute-forcing subdomains
- Permutations/alternative character substitutions
- Reverse DNS lookups
- Querying ASNs and netblocks
OWASP Amass (2018)

Open Source Intelligence (OSINT)

- Scrape web pages with DNS dragnet data
- Aggregate Passive DNS API data
- Crawling internet archives
- Recursive brute-forcing subdomains
- Permutations/alternative character substitutions
- Reverse DNS lookups
- Querying ASNs and netblocks

TARGET DISCOVERY 2018
DOZENS OF SOURCES
TECHNOLOGY STACK

REKOGNITION
ELSATICSEARCH
AWS LAMBDA
HEADLESS CHROME
S3 STORAGE
PERCEPTUAL ANALYSIS

WHAT IS IT?

APPROACH

- Reduce size. Shrink the image to a constant in order to ignore size differences
- Reduce color. Convert to grayscale to reduce colors
- Compute difference. Observe difference between adjacent pixels to identify gradient direction and store difference in bits.
- Assign bits. Each bit represents if the left pixel is brighter than the right pixel.

STEPS

- Shrink to 9x8 or 72 pixels
- Convert to 72 colors
- Compute 9 pixel differences or 8 differences per row which becomes 64 bits
- Assign 1 if \( P[x] < P[x+1] \) else 0

\[ \text{Image} = \begin{array}{c}
\text{Pixel}
\end{array} \Rightarrow \text{Hex} = \text{3a6c6565498da525} \]
FUTURE ENHANCEMENTS

OCR OF SCREENSHOTS
- To perform inclusive AND exclusive filtering based on text

ML
- To perform unsupervised categorization of screenshots and supervised training for valid consumer facing

COLLABORATIVE FILTERING
- To manage human analysis in real-time & centralize results
CONCLUSION

THANK YOU FOR YOUR TIME