Araña: Discovering and Characterizing Password Guessing Attacks in Practice

Mazharul Islam, Marina Sanusi Bohuk, Paul Chung, Thomas Ristenpart, Rahul Chatterjee
Online password guessing attacks are damaging

123456
1234567
password1
Password
qwerty
987654321
0000000
1111111

Attacker

Authentication Server

Guess1
Fail
Guess2
Fail
Guess3
Success
Online password guessing attacks are damaging.

According to the 2021 Data Breach Investigations Report, 89% of web application hacking attempts come in the form of credential abuse through stolen credentials or brute-force attacks.

One Stolen Password Took Down The Colonial Pipeline — Is Your Business Next?
Few prior works characterized guessing attacks

Detection
• Schechter et al. (Euro S&P ’16)
• Freeman et al. (NDSS ’17)
• Herley et al. (NDSS ’19)
• Bohuk et al. (USENIX ’22)

How do attackers execute password guessing attacks in practice?
How to detect password guessing attacks in practice?
Araña: Detecting Password Guessing Attacks

• Framework for detecting password guessing attacks
• Discover 25 new attack clusters
• Characterized attack clusters

Araña found 1,157 of new compromised accounts
Challenges to detect password guessing attacks

- Large scale of real-world login dataset
- Lack of ground truth on real word logins
- Unknown attack strategies

Araña uses Filter, Cluster, Analysis

Benign filters
Unsupervised clustering
Diverse feature types
Araña uses *Filter-Cluster-Analysis* Approach

Login L-sets

Login requests grouped by their IP addresses and date
Araña uses **Filter**-Cluster-Analysis Approach

- Login L-sets
- Potentially Malicious
  - Login requests grouped by their IP addresses and date
  - IP addresses exhibiting outright benign behavior

**Benign filters**
- High success rate
- Org’s Private IP addresses
- Successful 2FA
- ...
Araña uses Filter-Cluster-Analysis Approach

Login L-sets

Login requests grouped by their IP addresses and date

Potentially Malicious

IP addresses exhibiting outright benign behavior

Malicious Filter

L-sets exhibiting similar behavior are clustered together

L-set clusters

Benign Filter

Agglomerative Clustering

Features

- Client user agent, IP subnet, ISP
- Password features
- Volumetric features
- Timing, success rates etc.
Araña uses Filter-Cluster-\textit{Analysis} Approach

- **Login L-sets**: Login requests grouped by their IP addresses and date
- **Potentially Malicious**: IP addresses exhibiting outright benign behavior
- **L-set clusters**: L-sets exhibiting similar behavior in the same cluster
- **Attack clusters**: Manually analyze high volume clusters

\textbf{Sampling criteria}:
High precision detection of attack clusters
Araña evaluated on real world login dataset

Login L-sets: 453,361

Potentially Malicious: 8,125

L-set clusters: 1,006

Attack clusters: 29

Gossamer: Securely Measuring Password-based Logins

Marina Sanusi Bohuk, Cornell University
Mazharul Islam, University of Wisconsin–Madison
Suleman Ahmad, Cloudflare
Michael Swift, University of Wisconsin–Madison
Thomas Ristenpart, Cornell Tech
Rahul Chatterjee, University of Wisconsin–Madison

USENIX Security ’22

34 million logins for 7 months at 2 universities
Araña detected thousands of compromised accounts

<table>
<thead>
<tr>
<th>Number of</th>
<th>Univ 1</th>
<th>Univ 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>attack clusters</td>
<td>9</td>
<td>20</td>
<td>29</td>
</tr>
<tr>
<td>IP addresses</td>
<td>756</td>
<td>1,668</td>
<td>2,424</td>
</tr>
<tr>
<td>logins</td>
<td>75,884</td>
<td>287,16</td>
<td>363,051</td>
</tr>
<tr>
<td>users attacked</td>
<td>11646</td>
<td>152278</td>
<td>163924</td>
</tr>
<tr>
<td>compromised users</td>
<td>41</td>
<td>1,116</td>
<td>1,157</td>
</tr>
</tbody>
</table>

Identified multiple attack strategies
Attack cluster behaviors and strategies

Filtered usernames

Username, pw pair

Breach dataset

Weakest link of basic auth

IP addresses

Multiple days

Bursty

Araña can improve existing defenses
Araña: A new framework for detecting password guessing attacks in practice

- Filter-Cluster-Analysis approach
- Identified several attack clusters (potentially attack campaigns)
- Help learn attack strategies to improve existing techniques

https://github.com/islamazhar/Arana-Public

@Mazharul_13  mazharul@cs.wisc.edu

Thanks!