Team Breakdown

Sagar Samtani
- DB management
- Team organizer

Soyeon Kim
- DB management
- Analytics

Ryan Chinn
- Python expert
- Shodan analyst

Gyudong Cho
- Analytics
- Documentation

Hacker Web Group
Shodan Group
Introduction

Cyber Analytics

Hacker Web forums
- **18 hacker forums**
- **Author table**: Author ID, Names, Reputationscore
- **Thread table**: Thread ID, Title, # of View/Posts
- **Posts table**: Post ID, flatContent, postDate, ...
- **Code table**: Code ID, SourceCode, PostID

Shodan
- Search online devices based on **S/W, operating system, IP address**, etc.
- Gathering from Shodan **via Python**
  : scripts with query data / reuse code
Research 1 (Hacker Web): Hypothesis

Are the top five posters from each of the five biggest forums the main contributors of the overall number of posts in a forum and of the source code pertaining to networking related attacks?

Hypothesis: The top five posters of each forum are the main contributors of the overall number of posts in a forum and of the source code pertaining to networking related posts.

- Purpose and Value: Enable security professionals to determine threat vectors and implement reactive and preventative measures.
Research 1 (Hacker Web): Analysis

Determine the five largest forums and their top posting authors

Determine the percentage of postings and source code from these authors

Keyword search
Research 1 (Hacker Web): Findings

- Incorrect hypothesis
- Top 5 authors do not contribute to the majority of source code or number of posts
- Values from analysis do not exceed 50%

<table>
<thead>
<tr>
<th>Forum Name</th>
<th>% of posts</th>
<th>% of source code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antichat</td>
<td>1.812</td>
<td>25</td>
</tr>
<tr>
<td>Cnhonkerarmy</td>
<td>0.897</td>
<td>0</td>
</tr>
<tr>
<td>Arhack</td>
<td>5.285</td>
<td>0.501</td>
</tr>
<tr>
<td>Unpack</td>
<td>4.416</td>
<td>10.727</td>
</tr>
<tr>
<td>Ashiyane</td>
<td>3.264</td>
<td>12.977</td>
</tr>
</tbody>
</table>
Research 2 (Hacker Web): Hypothesis

How does the behavior of posters within forums compare in the larger threads vs the forum as a whole?

Hypothesis: The length of content posted in the top thread of the biggest forums are longer than the average length of content which all of the authors post in the forum

• Purpose and Value: Gain a better understanding into the psyche of the hacker community
Research 2 (Hacker Web): Analysis

Determine the top five forums

Determine the largest threads (in terms of posts) and the length of their posts

Determine length of average post
Research 2 (Hacker Web): Findings

- Incorrect hypothesis
  - The length of top threads is not longer than the length of all posts
Research 3 (Shodan): Hypothesis

How does the number of publicly facing hospital devices change over time and what is the distribution of these devices across organizations?

Hypothesis: The number of publicly facing network devices in the healthcare industry has decreased over time and it is an industry-wide problem.

• Purpose and Value: Provide awareness in the healthcare industry to the current trends and locations of unsecure devices.
Research 3 (Shodan): Analysis

Trend determination

Distribution of devices across organizations

Geographic visualization and distribution
Research 3 (Shodan): Trend analysis

- Incorrect hypothesis
  - Exponentially increasing trend in number of open devices
Research 3 (Shodan): Geographic

- Incorrect hypothesis
- Increasing open devices where population grows at a faster rate
Research 4 (Shodan): Hypothesis

What is the breakdown of the HTTP response codes from hospital related devices, and what does that indicate about their accessibility?

Hypothesis: The majority of healthcare devices will provide a 400 or 500 level HTTP response, denying any access and will be operating on port 443 by using Secure Sockets Layer (SSL) and providing an extra layer of security over the traditional port 80

- Purpose and Value: Will provide insight into truly insecure vs partially secure devices
Research 4 (Shodan): Analysis

Classification of HTTP response codes

Testing and placing of HTTP response codes

Determination of completely secure, inaccessible devices
Research 4 (Shodan): Findings

- Accessibility is determined by type of code.
  - 200/300 level = Accessible, 400/500 level inaccessible
- Incorrect hypothesis
  - 63% of the devices studied are accessible
Research 4: Reasons for Inaccessibility
Future Works

Create a more complete hacker profile

Explore different communities such as carding communities

Test IoT device vulnerability by entering default usernames and passwords
Recap

- Four research questions developed around Hacker Web and Shodan
- All four hypotheses proved incorrect
- When looking for potential threat vectors, look beyond top five
- Hackers act in a different manner in top threads than overall
- The number of open healthcare devices is exponentially growing
- The majority of these healthcare devices are insecure and inaccessible