Anomaly probability calculation with JA3/JA3s hashes

You might need to compute a probability for each event in a result set to help identify anomalies when doing the following:

- Detecting software supply chain attacks

Prerequisites

To succeed in implementing this use case, you need the following dependencies, resources, and information.

- People: Security analyst, threat hunter
- Technologies: Splunk Cloud Platform or Splunk Enterprise
- Data: Deep packet inspection data

Example

You can run a search that uses JA3 and JA3s hashes and probabilities to detect abnormal activity on critical servers, which are often targeted in supply chain attacks. JA3 is an open-source methodology that allows for creating an MD5 hash of specific values found in the SSL/TLS handshake process, and JA3s is a similar methodology for calculating the JA3 hash of a server session.

This search is most effectively run in the following circumstances:

- with an allow list that limits the number of perceived false positives.
- against network connectivity that is not encrypted over SSL/TLS.
- with internal hosts or netblocks that have limited outbound connectivity as a client.
- in networks without SSL/TLS interceptions or inspection.

To optimize the search shown below, you should specify an index and a time range. In this example, Zeek is used to generate JA3 and JA3s data but you can use any other tool which can generate that data.

1. Run the following search:

```
sourcetype="bro:ssl:json" ja3="*" ja3s="*" src_ip IN (192.168.70.0/24) | anomalydetection method=histogram action=annotate pthresh=0.0001 src_ip, ja3, ja3s | stats sparkline max(log_event_prob) AS "Max Prob", min(log_event_prob) AS "Min Prob", values(probable
```
cause) AS "Probable Causes", values(dest_ip) AS "Dest IPs", values(server_name) AS "Server Names",
values(ja3) AS "JA3", values(src_ip) AS "Source IPs" count BY ja3s
| sort "Min Prob" asc

Search explanation

The table provides an explanation of what each part of this search achieves. You can adjust this query based on the specifics of your environment.

<table>
<thead>
<tr>
<th>Splunk Search</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>sourcetype=&quot;bro:ssl:json&quot; ja3=&quot;<em>&quot; ja3s=&quot;</em>&quot; src_ip IN (192.168.70.0/24)</td>
<td>Search Zeek data for JA3 and JA3s hashes within the critical server defined.</td>
</tr>
<tr>
<td></td>
<td>This part of the search uses critical server netblock, 192.168.70.0/24. It's important that you adjust this part of the search to include your own critical servers.</td>
</tr>
<tr>
<td></td>
<td>Uses frequency analysis to detect unlikely(anomalous) events in categorical fields, in this instance, JA3 and JA3s hashes, using a histogram to show distribution.</td>
</tr>
<tr>
<td></td>
<td>You may need to modify the pthresh/probability threshold value depending on your data and how many benign results that you receive.</td>
</tr>
<tr>
<td></td>
<td>Generates a sparkline, or an inline chart that appears within table cells, and renames fields as shown, grouped by JA3s.</td>
</tr>
<tr>
<td></td>
<td>Display the results in a table with columns in the order shown.</td>
</tr>
<tr>
<td></td>
<td>Sort by minimum probability in ascending order.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Result

This search returns abnormal source IP, JA3 and JA3s combinations within your data. In the example below, you can
see results that look suspicious within the top 12 anomalous results.

This search is helpful for identifying malicious abnormal activity, especially where it occurs over a 24 to 48 hour period. Periods longer than this reduce the effectiveness of the query.

If you are searching on a smaller network with a single/24 netblock, malicious activity should be consistently identified in the top results. If you are searching on a larger network with multiple or more extensive netblocks, an allow list could be beneficial to ensure that malicious results are seen within the top results in this search.

Generating a lookup table of stored probabilities can help you run this detection faster.