Web Page Classification

Web page classification analyzes the domain, URL, and most importantly, the words and phrases on *every page load* to tally a numerical score in one or more Categories for *that page load*.

The filter Action configuration (Allow/Block/Ignore) for the top-scoring Category is then used to handle that particular page request.

Traffic Visibility Prerequisites

Webpage word and phrase analysis is only possible with full SSL/TLS decryption (`sslbump`), which is the default action for most[1] web requests on TCP ports 80 (HTTP) and 443 (HTTPS).

And, for this to work without browser security errors, all endpoint devices connecting through the DrawBridge must have the **DrawBridge Certificate Authority certificate** installed. See the page **SSL Certs** under the **Devices** module for more information.

[1]**Note:** for security reasons, banking and financial-related websites **are not TLS-decrypted**. It is assumed that these sites are safe from inappropriate content. You can verify a site is Not being TLS-decrypted by clicking the shield or padlock in your browser address bar and viewing the certificate. If the certicate is issued by a public Certificate Authority (and not your DrawBridge), you can know that the DrawBridge is Not intercepting the connection.

Also Note: Certain web traffic (for example some cloud backup services and application traffic) that is not specification-compliant or is otherwise incompatible with content filtering are exempted at a firewall level from the traffic inspection on TCP ports 80 and 443.

Example

Visiting https://www.cabelas.com is most likely to score the most points in the Category `Hunting and Fishing`.

- If the Action assigned to `Hunting and Fishing` is `Allow`, the Cabelas page will load as if nothing happened.
- If the Action assigned to `Hunting and Fishing` is `Block`, a DrawBridge block page is loaded to inform the user that the request was blocked due to filter settings.
- If the Action assigned to `Hunting and Fishing` is `Ignore`, the next-to-top scoring Category action is selected to handle the page load.

> The option to `Ignore` is strongly discouraged except for special situations. If you decide to specify custom Actions for Categories, please only use `Allow` or `Block` to ensure most reliable filtering.

Important Notes

1. About changing default Category Allow/Block settings

The DrawBridge comes with a preset Action for each included (Built-in) Category. When you assign an Action (Allow/Block) to a Category, **you're simply applying a change that gets higher priority than the default setting.**

2. Default Category settings are Business-focused

The default settings for the Built-in Categories are tightly scoped to business-usage needs. Depending on your usage expectations, you may want to set more categories to **Allow** in your *Company Preferences* Access Policy, or in a custom Access Policy.

For more information on Built-In Categories, including how to view default Actions, see **Content Filter: Categories: Built-In Categories**

Further Reading

For more information on Categories and Actions, including how to change the Action for a Category, see page **Overview and Essentials** under the **Content Filter** module.

For more information on Certificates and Certificate Authorities, this Wikipedia article on Public Key Infrastructure may be helpful.

FAQ: Is TLS inspection "bad" or "breaking encryption" or "weakening security"?

In a word, **no** *(if implemented correctly)*

Despite much negative press, blog posts by both Cloudflare and US-CERT acknowledge that legitimate use-cases (and secure methods) of TLS inspection exist.

Some of the concerns raised in the two articles linked above are very valid. However, the DrawBridge filter engine is designed to follow industry best-practices to ensure that it doesn't downgrade security or mask upstream security flaws.

Much of this debate boils down to two things:

1. Intention: Why is the TLS traffic being inspected? (legitimate or malicious?)
2. Privacy: Are the end-users aware of the inspection? (visible/policy or invisible/spycraft?)

For #1: The DrawBridge employs TLS inspection to ensure content filtering properly classifies page content

For #2: Yes: DrawBridge account holders need to purchase the content filter service and need to install a Certificate Authority for the service to work correctly. (It is the responsibility of account holders to inform any user of the service of the content monitoring and inspection.)

This discussion leads to an even deeper question: *Who owns this device*? If you truly own a computer, for example, you should have the authority to decide what Certificate Authorities it will be allowed to trust, and with whom it will communicate. Thankfully, most platforms accomodate adding additional Certificate Authorities, enabling you to know and control the network traffic of your device.

The notable exception is Android, because of an alleged "security" decision by Google. While there were threats they were able to prevent by taking a scorched-earth no-user-CA-trust position[1], this implementation also conveniently prevents auditing of the traffic of third-party apps and bundled Google apps.

[1]Exception: browser apps on Android will trust user-installed Certificate Authorities.

Revision #27
Created 1 November 2022 13:33:57 by Marvin M.
Updated 20 November 2023 17:55:20 by Timothy P.