A Website Classification Database: An Ideal Source of Threat Data

Posted on November 5, 2019

A simple rule applies in today's infosec environment: organizations must consider the effectiveness of their threat data sources. In fact, this should be a primary concern, especially if they wish to get the most out of their threat intelligence platform. But not all companies know which sources of threat intelligence can benefit them.

TIPs are critical for many enterprises. A TIP is capable of collecting and managing threat data coming from multiple external sources. It lets companies correlate the information with one another to come up with insightful findings. This process allows them to identify which threats they need to prioritize. A TIP can also reduce risks by answering who is responsible for and what comprises an attack.

So, just like any data analytics software, what organizations get from a TIP largely depends on what goes into it. After all, TIPs consume threat intelligence, which is why the sources they get them from should be topnotch.

But how can one distinguish good threat intelligence sources from bad ones? Let's take a look at the following four criteria.

Backed by a Qualified and Trusted Third-Party Provider

Many organizations lack the capability to gather, organize, and analyze threat data on their own. That is why they make use of TIPs to do all of these activities for them.

A reputable third-party provider should be behind your threat intelligence source. The provider should vet its data sources for accuracy because if it doesn't, the results they give can overwhelm analysts with false positives.

Should Provide Information on Active Campaigns

Most companies today already have a good understanding and information on exploits, vulnerabilities, malware, and other kinds of threats. What they may lack are insights into active campaigns. This information will tell them who is behind the attack, what vectors the perpetrators used, where the attack comes from, when ?? began, and how to address the threat. The best kind of insight is one that is relevant to the organization's structure and business context.

Should Provide Relevant Insights to Users

Most types of threat intelligence provide an overview of the risks and business impacts related to threats. But such insights are only useful if they are relevant to the organization. For instance, threat data related to technologies that it doesn't use is meaningless. As such, threat intelligence sources should match a user's systems, processes, and assets.

Utilizes an Algorithmic Approach

Algorithm research has come a long way in giving products the capacity to explore and analyze a vast amount of data. Algorithmic approaches not only allow users to collect threat data from multiple sources at a rapid rate, but also provide for automatic and near-real-time analysis. As such, TIPs that employ artificial intelligence (AI) and machine learning (ML), for instance, are highly recommended.

WhoisXML API Offers an Excellent and Relevant Source of

Threat Intelligence

Websites are an essential part of any threat actor's arsenal. Cyber attackers can't launch attacks without them. Therefore, it is only natural that those working against them should study website data.

WhoisXML API provides a well-structured database that contains extensive information on websites. Its product uses a machine learning (ML) engine with natural language processing capabilities. It can retrieve website content as well as metatags while assigning categories to domains for website classification purposes. Its Web crawlers parse millions of pages regularly to obtain active domain name contact information. This data can be used for many security functions but is particularly handy in identifying malicious domains. It also provides more detailed insights for threat investigations.

At present, our web categorization service classifies domains into 25 different categories. Categories that show up include arts, beauty, home, people and society, sports, shopping, recreation, news and more. These categories get updates on a regular basis, and users can always request new ones as the need arises. Besides these, you can also find meta-information, social media accounts, and emails related to the websites.

All of the information in the database is well-parsed and normalized to a standard format. Users can get them either parsed or raw, depending on their needs. The downloads can come as data dumps or as comma-separated values (CSV) files. As such, the database can be easily integrated into existing business processes and systems. Since the data sets are standardized, users can quickly correlate them without the need to translate or reformat. Easy integration is, after all, highly essential, especially with the growing complexity of today's technologies.

Finding the right data source for a TIP is essential so organizations can use their threat intelligence solutions to the fullest. After all, what use is a TIP that cannot keep up with today's cyber threats?

A reliable threat intelligence repository, like our **website classification** database, can help. With it,

cybersecurity teams can ensure their web threat insights are legitimate and up-to-date.

Would you like to know more about how our database can benefit you? Send us a message today for more information.