

Website Fingerprinting and Distributed Proxy Detection System

A CIPAFilter White Paper

In education today, many websites have become hot button issues with administration, faculty, and parents. Websites like Myspace, Facebook and YouTube have become synonymous with technical, social, and political problems. More and more often we've been hearing, "What we need is a filter that always blocks Myspace. I can't afford to have a single problem with Myspace; it could cost me my job." The demand is there, but serious technical, social, and economic phenomena have come together to make most blacklisting systems on the market today nearly obsolete.

First, a quick rundown on what has happened. Most school network administrators are familiar with anonymous web proxies. These are websites such as "unblockmyspace.com" which exist for the sole purpose of helping students bypass blacklist based filters. These sites spring up like weeds, get blacklisted, and die just as quickly only to resurface days later under a new name. **The reason is simple. There is big money to be made in helping children bypass Internet filters.** These websites replace the ads that Myspace would normally run, or they run their own ads as banners across the top of each page. They focus on generating revenue from products that interest children and teens like ringtones and emoticons. Many also peddle pornography and propagate spyware to the computers of the victims who choose to use them, but if these computers are your responsibility, that victim is you!

At CIPAFilter, we've known about the flaws in the blacklisting paradigm for a long time. We foresaw these problems back in 2000, when the first CIPAFilters were being designed. We created a pornography blocking system that was immune to these anonymous proxy servers, and now we've created a system that solves the problem of these hot button websites as well.

What can match the energy of millions of students across the country when it comes to finding and using new proxies? Only the students themselves, of course!

Website Fingerprinting

CIPAFilter has implemented our new anti-proxy solution as a two-prong system. The first technique we call fingerprinting. Using unique strings of syntax from a particular website, a trained analyst can create a fingerprint that can be used to detect that website regardless of how it's being rendered. For the students to view the website, certain portions of the code of that site must remain intact on the way to the student's workstation. We can isolate and create pattern matching systems to detect these fragments of code. **With a CIPAFilter configured for high security¹, students will not be able to reach Myspace, Facebook, or any other fingerprinted website, from your network.**

Our propriety fingerprinting system is an outgrowth of the ground breaking work we did in high speed pattern detection nearly 8 years ago for our non-blacklist based, zero-false-positive content filtering system.

Distributed Proxy Detection System

Our first technique detects these fingerprints in real time, immediately blocking the student from reaching the fingerprinted website (e.g. Myspace), but our second technique takes this one step further. If Myspace is on the blacklist, and a user has just accessed it, they must have done so through a proxy website. We immediately add the newly detected proxy site they were using to your local blacklist.

Additionally, the hostname of that proxy website will then be transmitted back to CIPAFilter's corporate offices with the URL for confirmation and a digest hash for security. These newly discovered blacklist entries will then be confirmed and distributed twice a day to all CIPAFilter customers. **This means that when any student uses a proxy server to access a fingerprinted website in the morning, the proxy server they used will be blacklisted nation-wide by the afternoon.**

Conclusion

This two-pronged approach of completely eliminating "hot button" websites with fingerprinting and aggressive automatic detection and distribution of web proxy servers will give CIPAFilter customers the edge, ensuring that all students attempting to circumvent the filter, are stopping proxy servers with a global "trap" that is immediate and allows all CIPAFilter units to work together to put an end to anonymous proxies.

1. A CIPAFilter configured for high security is a unit configured as a drop firewall to block proxy software like tor, and not allowing SSL website access for students, or only allowing SSL access to a preselected group of websites.