XCS - Objectionable Content Filtering

Stop Cyber-Bullying in its Tracks - Protect Schools and the Workplace

Technical Brief

WatchGuard® Technologies
Published: January 2011

Cyber-Bullying

Bullies are everywhere – the playground, at school, on sports teams, and today they are also online. During the past six years, cyber-bullying has become more mainstream. Schools, parents and society now need to address this issue proactively. Cyber-bullying opens the door to 24-hour harassment through email, instant messaging, cell phones, gaming consoles, or social networking sites, chat rooms, and other Internet enabled-devices.

> *Stopcyberbullying.org, an expert organization dedicated to Internet safety, security and privacy, defines cyber-bullying as: "a situation when a child, tween or teen is repeatedly 'tormented, threatened, harassed, humiliated, embarrassed or otherwise targeted' by another child or teenager using text messaging, email, instant messaging or any other type of digital technology."*

Why is cyber-bullying different to traditional harassment? Because humiliating rumors, threats and vicious taunts can be viewed by millions and can be devastating to youth and their families.

Statistics from the National Crime Prevention Center include:

- More than 40% of all teenagers with Internet access have reported being bullied online.
- Girls are more likely than boys to be the target of cyber-bullying.
- Only 10% of those bullied told their parents about the incident, and only 18% of the cases were reported to a local or national law enforcement agency.
- Only 15% of parents are "in the know" about their kids' social networking habits, and how these behaviors can lead to cyber-bullying.
- Most common virtual locations for cyber-bullying are chat rooms, social networking web sites, email and instant message systems.
- Social networking sites such as Facebook are growing fast, and so are the cyber-bullying incidents originating from them. Experts believe that they will soon overtake chat rooms as the top source of cyber-bullying problems worldwide.
- 58% of 4th through 8th graders reported having mean or cruel things said to them online; 53% said that they have said mean or hurtful things to others while online; 42% of those studied said that they had been "bullied online".
- Cell phone cameras and digital cameras are a growing problem in the cyber-bullying world. A recent survey found that 10% of 770 young people surveyed were made to feel "threatened, embarrassed or uncomfortable" by a photo taken of them using a cell-phone camera.

Fastest growing Cyber-Bullying tactics are:

- Stealing an individual's name and password to a social networking site, then using their profile to post rumors, gossip or other damaging information.
- Altering photographs using PhotoShop or other photo editing software in order to humiliate the individual.
- Recording conversations without the individual's knowledge or consent, then posting the call online.
- Creating confrontational and mean-spirited online polls about the individual and posting them on different web sites.
- Using web sites and blogs to post hurtful, embarrassing information about another individual.

Although a lot of work can be done in schools, at home, and in the workplace to educate and counsel individuals about cyber-bullying and its effects, it is an issue powered by technology that can be controlled with technology. For example:

- For home-based computers: there are "parental control" softwares that parents can install
- For phones: parents can simply ask the phone company to block features, callers, texting, etc – or for very advanced phones there are parental control apps to install
- For gaming devices: many of the newer games come with some type of parental controls built right in – parents just need to turn them on.

But what about school computers and networks as parents don't have control over Internet-based activities and messaging that are taking place in their children's schools.

Can you afford to ignore these types of activities could be happening on your school's computers or even in the workplace? Are messaging threats from email and web usage at school posing a threat to students, faculty and staff? What are the legal vulnerabilities, wasted network bandwidth, loss of student and faculty productivity? How much time is spent battling viruses, spam, phishing, and malware?

As students increasingly use Internet communications in school, they are exposing themselves, and the school district, to malicious and unacceptable content as well as safety concerns surrounding threatening and dangerous abuse of messaging vehicles for harassment purposes. School and district IT departments are increasingly providing support and network services to growing student bodies and faculty users, offering both in-school and at-home access to email, the Internet, network user folders and coursework via district networks.

With the introduction of recent Federal laws and other regulations set at the County and District level the onus is now on school districts which, under the laws and regulations, have an affirmative obligation to protect the safety and privacy of students and staff. Ultimately, school districts have the responsibility to provide a positive learning environment free of Internet-based threats, harassment and bullying conduct.

Today, cyber-bullying doesn't affect just parents, schools and children, but it has become a problem for IT departments, particularly in Education. With the increased use of social networking sites, such as Facebook, and the chance of slanderous comments being posted, IT departments need ways to prevent cyber-bullying. Most email and web security solutions focus on content coming in from the Internet to protect the internal environment and do not focus on what is being sent out of the network. Some districts believe the solution is to entirely block student access to these messaging tools, however, with the wealth of knowledge available to students on the Internet this could prove to be counterproductive to the learning process. With the advancement in messaging security technology, however, there are ways to fight cyber-bullying before it escalates without the need to completely block student access to email and web protocols via school networks.

WatchGuard XCS Stops Cyber-bullying In Its Tracks

WatchGuard Extensible Content Security (XCS) features the ability to block or flag cyber-bullying, slander and comments related to depression and suicide through traditional email, webmail (such as

Gmail) and Internet sites including Facebook. XCS also offers best-of-breed anti-spam, anti-malware, URL filtering, outbound content control, data loss prevention, and detailed diagnostic tools for email and web traffic.

Eliminate Cyber-bullying Posts to Social Network Sites with XCS Content Control Rules Facebook

Figure 1 shows that an attempted post to Facebook has been blocked by WatchGuard XCS due to the nature of the words used in the post. The user only sees an error message, and would believe that either Facebook has blocked the post, or Facebook is currently down.

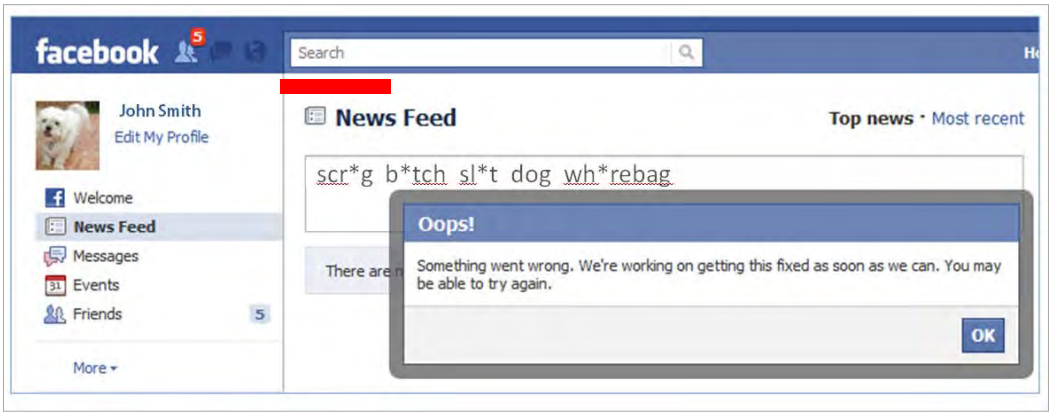


Figure 1. XCS Blocks Facebook Post [i]

Figure 2 shows what the administrator will see on the WatchGuard XCS dashboard. Due to the ***Content Control Rule,*** the administrator knows something was blocked. Also, based on policy, an administrator can trigger an email regarding the breach to be sent. For example, to an HR Manager, a School Principal, or any other individual or group.

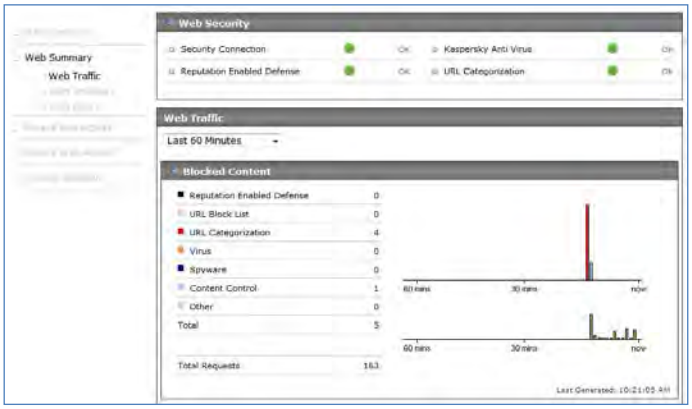


Figure 2. XCS Screen Reporting Blocked Content

By looking at the activity page under Recent Web Activity (Figure 3), the administrator can see that the breach in question was an intended post to www.facebook.com that was blocked by the WatchGuard XCS Content Control Rules.

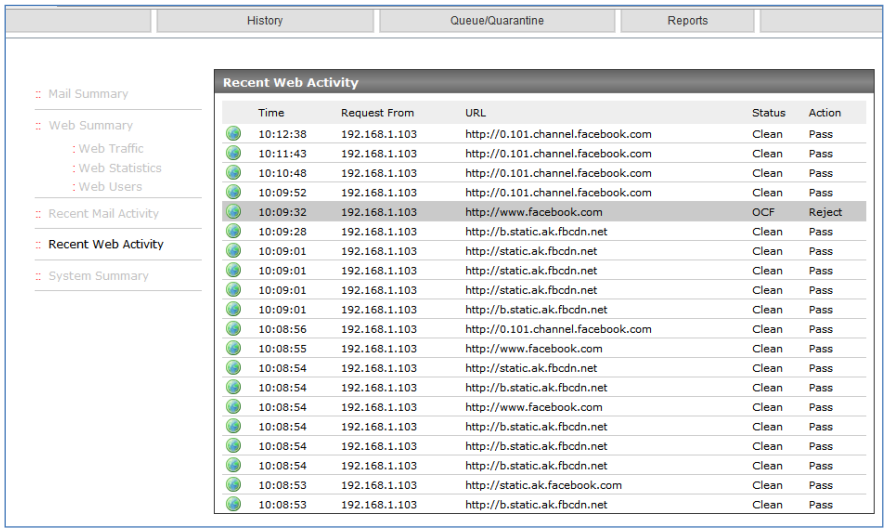


Figure 3. Activity Page

By clicking on the entry, a drill down shows more details about the incident, including the URL.

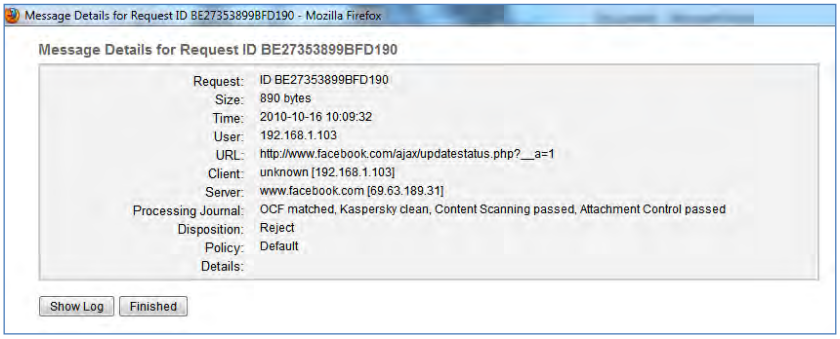


Figure 4. Message Details

By clicking on ***Show Log*** the administrator is able to determine exactly why the posting was blocked.

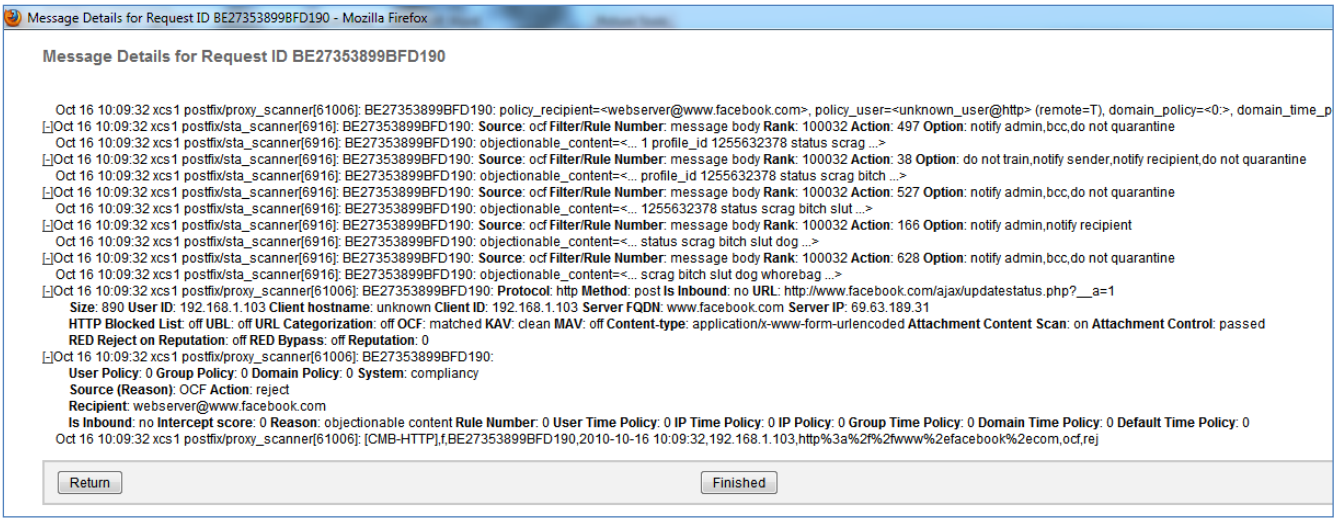


Figure 5. Detailed Message Details

Block Malicious and Slanderous Emails

When sending an email from Gmail (or any other web-based email) that contains inappropriate or disallowed content such as obscenities or content that may be deemed to be considered slanderous or a form of cyber-bullying (example Figure 6), an error pop-up screen would appear for the sender as in Figure 7 below.

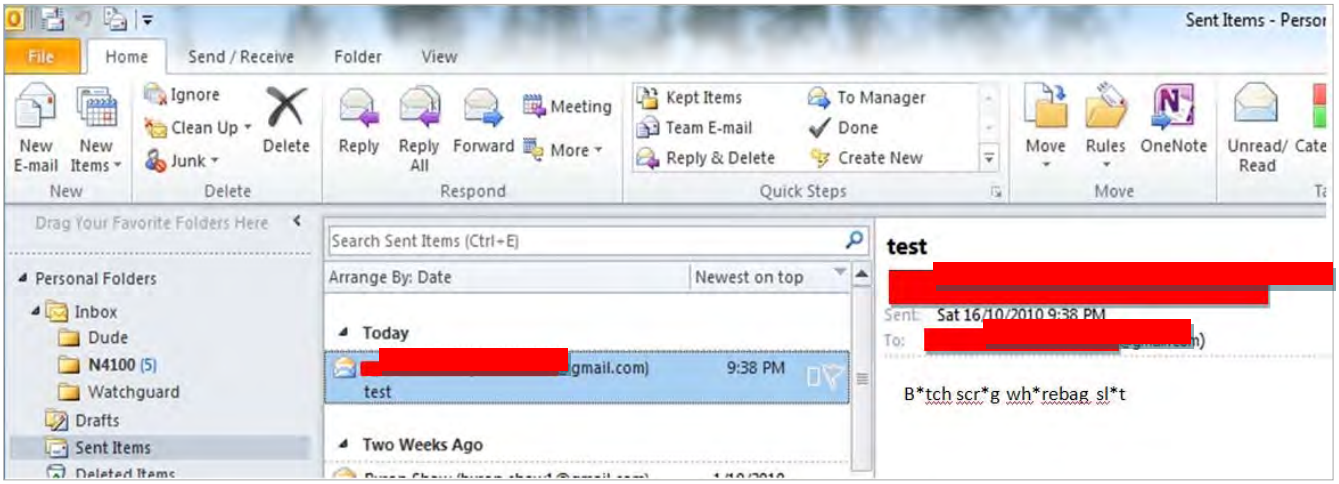


Figure 6. Email with Objectionable Content

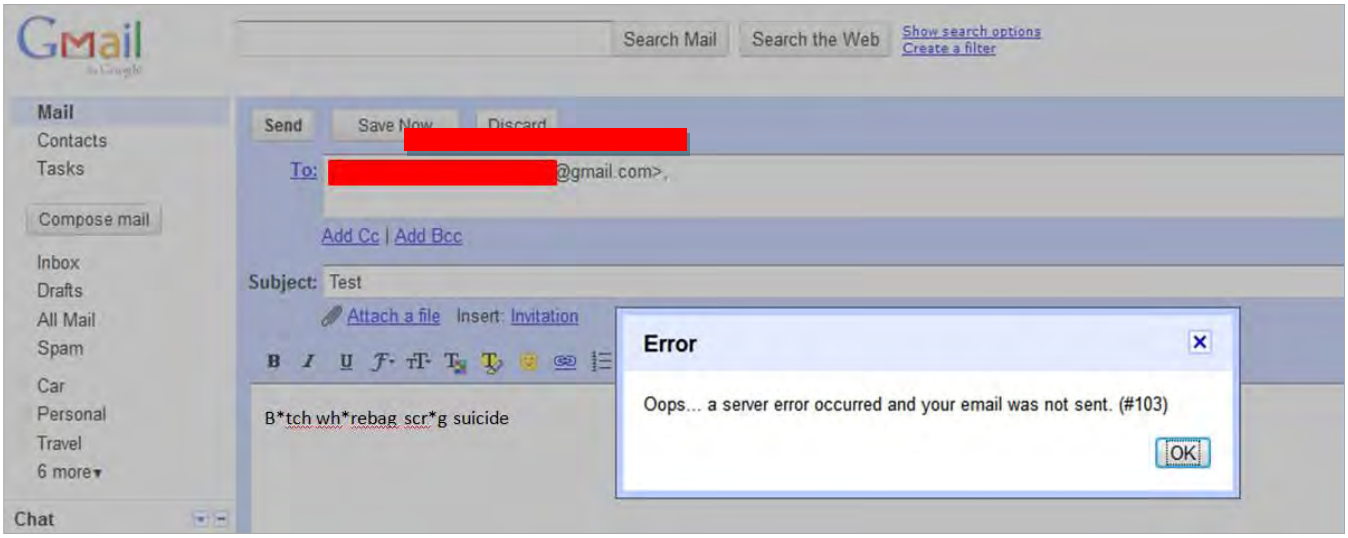
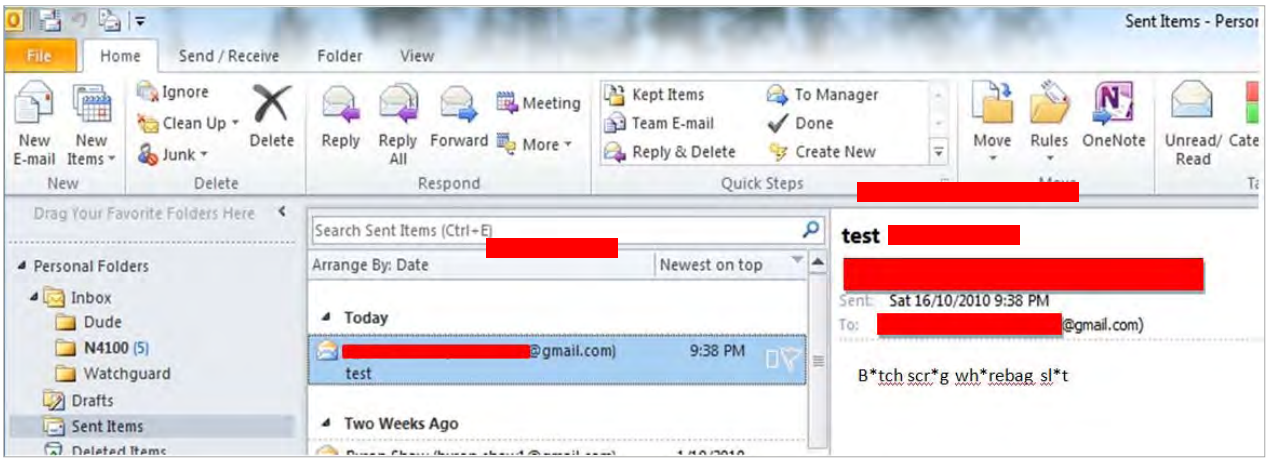
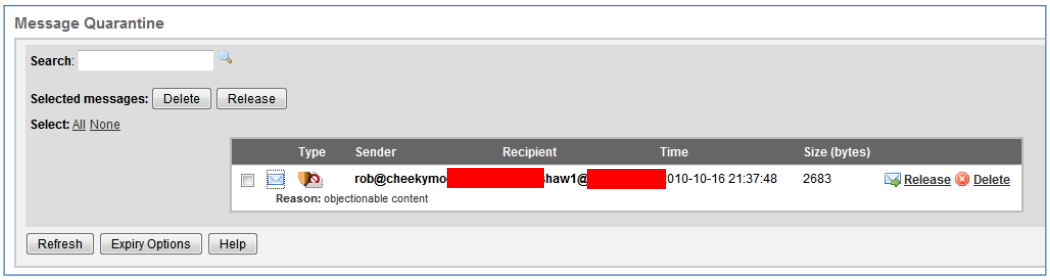


Figure 7. Email Error Message

Traditional emails are blocked or quarantined as they try and leave the Mail Server, such as Exchange or Domino. They will appear to be sent, and remain in the user's Sent Items. WatchGuard XCS can place the email into quarantine or block the email altogether based on policies defined by the administrator. It can also blind-copy or reroute the message as another form of transparent remediation.





Enabling Objectionable Content Filters To Extend Protection Across Email & Web

WatchGuard XCS provides the ability to set Objectionable Content Filters to prevent cyber-bullying. To do this, an administrator would follow the steps below:

1. Ensure the Feature Key includes Objectionable Content Filtering for Email and Objectionable Content Filtering for Web by navigating to ***Administration>> System>> Feature Key***. Figure 8.

   Note: If these features are not enabled, you can request an evaluation key from your local WatchGuard Account Manager, or by raising a Customer Care incident at http://www.watchguard.com/support/contactsupport.asp If you wish to purchase the Web Security Subscription (which is required for Objectionable Content Filtering for Web), then contact your local WatchGuard Reseller.

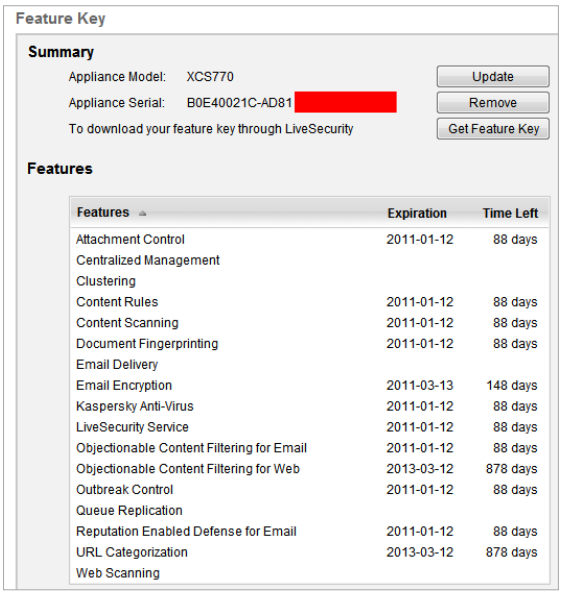


Figure 8. Feature Key Screen

2. In order to inspect both Web and Email traffic, ensure that the interface is set up for HTTP/HTTPS Proxy.
    a. Go to ***Configuration>> Network>> Interfaces***. See Figure 9 below.
    b. On the Interface that will be listening for HTTP requests, ***check the box to enable the HTTP/HTTPS Proxy***.
    c. Click ***Apply*** at the bottom of the screen. Note: To activate, a reboot is required.

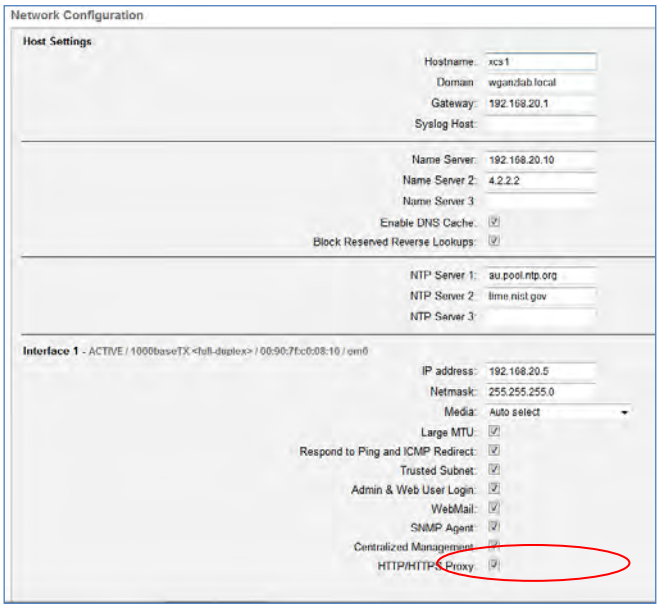


Figure 9. Network Interfaces Screen Hi-Lighting HTTP/HTTPS Proxy

3. Enable the HTTP/HTTPS Proxy at ***Configuration>> Web>> HTTP/HTTPS Proxy***. Ensure the local subnets are added to the ***Allowed Networks*** list. See Figure 10.
4. Click ***Apply***

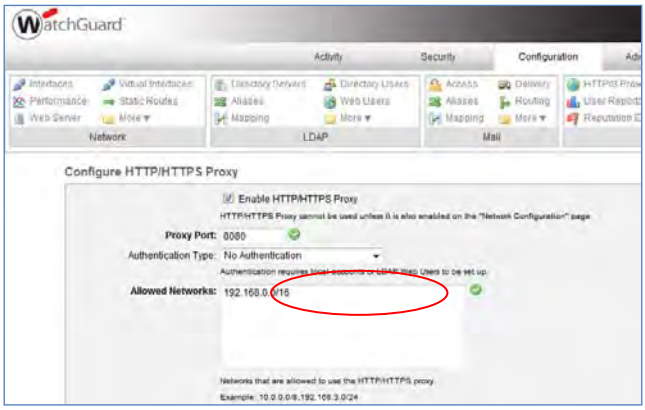


Figure 10. Hi-lighting Subnets

4. Download the latest version of the weighted ***Objectionable Content Filter and Slander Dictionary*** from:
    a. http://www.watchguardtechnologies.com.au/docs/weighted_ocf_slander_dictionary.txt and save it locally. Be warned, the contents are highly offensive, however, you may need to modify this dictionary to suit your organisation.

5. Navigate to ***Security>> Content Control>> More>> Dictionaries and Lists***

6. Click ***Add***

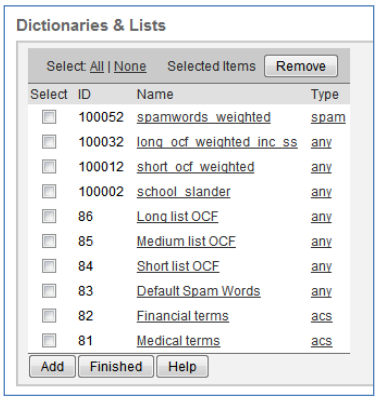


Figure 11. Dictionary & Lists Screen

7. Browse to the file you just downloaded.

8. Click ***Continue***

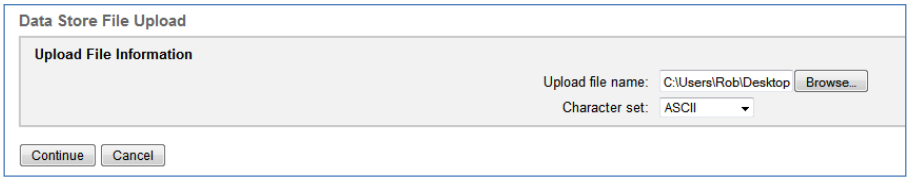
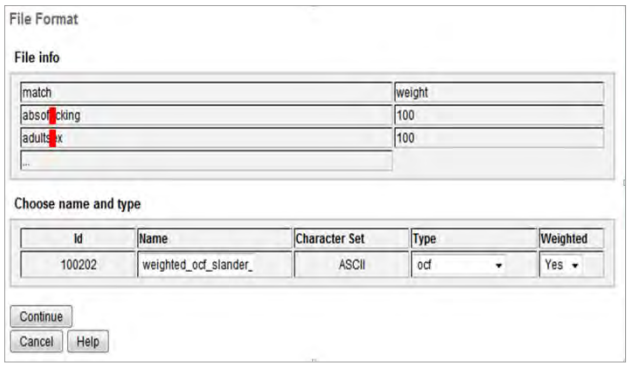
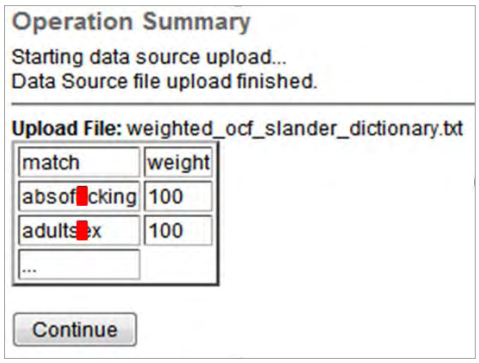


Figure 12. Screen for Uploading File

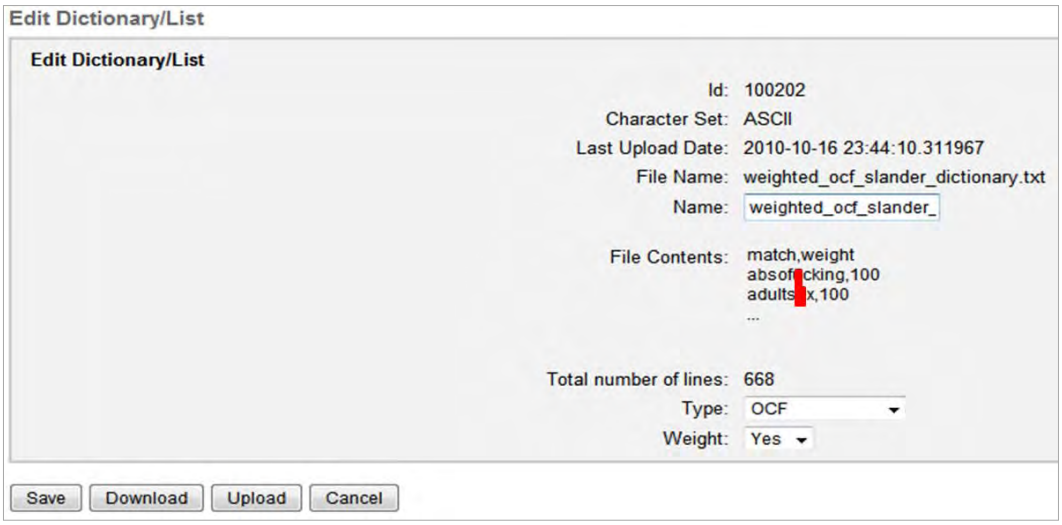
9. Ensure that the ***Type is OCF*** and the ***Weighted drop down*** is ***Yes***

10. ***Click Continue*** and ***Continue again*** (see below)





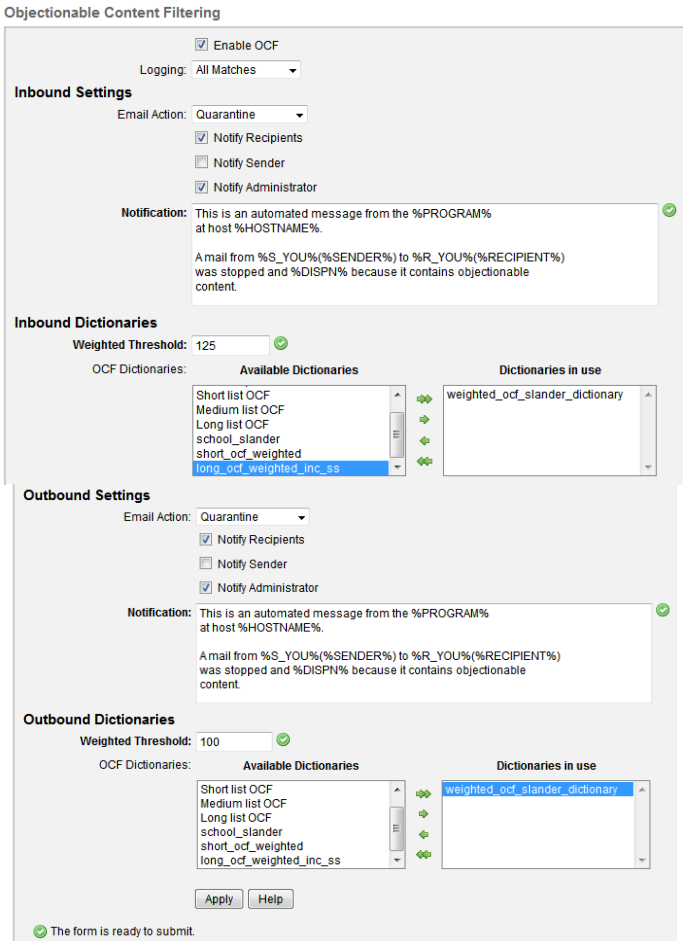
11. Click ***Save***. The new dictionary is now available for policy.



12. Navigate to ***Security>> Content Control>> Objectionable Content***.

13. Enable ***OCF and set as below***. Note the Inbound Threshold is higher, to prevent overblocking blogs, news and other mixed content web sites.

14. Click ***Appy***.



To test, set the proxy of your browser to the IP address of the XCS appliance. You may need to adjust your firewall rules to allow web traffic from the XCS.

Summary

With daily headlines of online bullying and the increasingly alarming impact, including youth suicides, school hostage situations, attacks on schools and students, and other horrendous outcomes, everyone needs to do their part to effect a change and stop cyber-bullying in its tracks.

Although many have labelled cyber-bullying a social or parental issue, cyber-bullying can be controlled and stopped by deploying effective email and web security technologies. Cyber-bullying has certainly brought schools and districts to the forefront of the issue. There are steps schools and universities can take to prevent hostile messages, web posts, and images sent within the school's networks. WatchGuard XCS can help put control in the hands of educators by providing the critical tools required to enforce content controls by filtering what is sent via email and web from the school's network.

WatchGuard XCS is used in schools and universities across the globe to protect students from cyber-related risks and vulnerabilities. Being able to monitor or block offensive or inappropriate words in emails and web posts that may be considered bullying is vital to a school doing their part to stop cyber-bullying from occurring within its networks and make bullies accountable for their actions before a student, or even a faculty member, becomes a victim. With granular content controls across email and web, the school has control of potentially malicious messages so they never leave

the network, and the intended victim – and anyone else for that matter – never sees it. And with the ability to generate alerts on suspected cyber-bullying activities and have the messages sent automatically to principals, teachers or other relevant authorities, schools and districts now have the ability to be proactive and act quickly.

Cyber-bullying may be a growing social problem, but it is a problem that can be solved with technology – count on WatchGuard XCS to stop cyber-bullying from taking place on your networks.

---

[i] Please note: all objectionable content words have been slightly modified or red-lined for censorship purposes. WatchGuard XCS provides the ability for any offensive, slanderous, or otherwise offensive or malicious content to be included in its content control capabilities with objectionable content filtering.

ADDRESS:
505 Fifth Avenue South
Suite 500
Seattle, WA 98104

WEB:
www.watchguard.com

NORTH AMERICA SALES:
+1.800.734.9905

INTERNATIONAL SALES:
+1.206.613.0895

ABOUT WATCHGUARD

Since 1996, WatchGuard Technologies has provided reliable, easy to manage security appliances to hundreds of thousands of businesses worldwide. WatchGuard's award-winning extensible threat management (XTM) network security solutions combine firewall, VPN, and security services. The extensible content security (XCS) appliances offer content security across email and web, as well as data loss prevention. Both product lines help you meet regulatory compliance requirements including PCI DSS, HIPAA, SOX and GLBA. More than 15,000 partners represent WatchGuard in 120 countries. WatchGuard is headquartered in Seattle, Washington, with offices in North America, Latin America, Europe, and Asia Pacific. For more information, please visit www.watchguard.com.