SECURITY'S LEAKS IN SEO SPAMMING

Lilyana Petkova
University of Library and Information technology (UNIBIT), Bulgaria, lilyanapetkova92@gmail.com

Abstract: The full lack of security provides a huge omission in the application or software development life cycle. An example of security flaw is an ineffective search engine optimization (SEO). That can lead to various troubles in an already well developed or during development Internet sites. There has been a rise in SEO spam and malware which purpose is to destroy the application search indexes by adding content, which does not correspond to the idea of the application. Unusual links, pages or comments can confuse the website's visitors and drive the traffic to malicious websites.

A badly secure websites are juicy target for each hacker. In some cases, the obstacle of implementing each type of security protection is just time, but the worst reason is either uninformed or unqualified developers without experience in SEO. The amount of guides and tools to help them is huge which interferes with developers' work by adding more time in investigating it!

The purpose of the current research is to review SEO spams, malware, and the damages they can cause to the Internet applications! Firstly, we describe the concept and definition of SEO spams and malware. Then we are going to review different mechanisms for protecting the Internet applications of those SEO spams and malware. We provide basic information for each protection method as well as some real examples on the SEO spamming threat.

In conclusion, we analyze different damages they can cause as well as different ways of finding those types of attacks by scanning some of the most common sites in Bulgaria. In the analysis, we have used the ALEXA Top 1 Million sites Insight tool. In addition to our new knowledge and skills gained during the research, we have added several more sites developed in our work and on which we have applied those skills.

The concept of our entire study is to show the use of free tools in supporting developers' work progress. Therefore, we are showing some free scanners in order to confirm what we propose in this research!

The need of security integration in the web applications becomes more and more necessary with the evolution of technology. Along with other security's implementations on a programming and on a server's level the ones described in the article bring another layer of security management that mitigates certain types of cyberattacks and vulnerabilities. It does not fully protect our applications but it gives an extra level of security by restricting the attackers' actions.

Those attacks as well as other types of attacks prove that the security of an application must be taken apart and integrated at the beginning of the application development life cycle.

Keywords: SEO, security, spam, optimization, protection, malware

1. INTRODUCTION

Most of the malware attacks against Internet applications are due to the SEO spam and malware. Those cyber threats can significantly damage the website or the entire business. Even though the SEO optimization can be devastating to the website's reputation, many website owners are not aware of its advantages or due to the website's size do not want to spend more time and effort on it. Unfortunately, the attackers do not care of the size of the web site and are always finding a way to target those SEO vulnerable websites.

With the present research, we will demonstrate the high need of SEO optimization and the results from not spending the time to use it properly in the websites.

The present research is divided in two main parts. The first section provides information on the SEO spam and malware structured in several subsections. Each subsection is based on different areas of the SEO world. We start the research by providing basic information on the SEO and the SEO spams for better understanding the purpose of the present article. We give information on how SEO threats affect our web sites; what they can cause and how they can harm us. Finally, we provide several ways of detecting that kind of cyberattacks and what we can do to prevent them.

In the second part of the research, we provide a malware detection statistic on several previously selected with ALEXA web sites on which we based our research. In addition, we include several new sites on which we have applied some of our new knowledge gained over the research. By this analysis, we are trying to provide a status upgrade over the research we are presenting.

2. RESEARCH METHODOLOGIES

In this section, we are giving a definition of what SEO spam is and how it can affect our website! For that purpose, we are giving some basic information on what is SEO spam and how it generally works! Finally, we are exposing some real use cases of that threat and some tips and tricks on how to protect our content from malware.

2.1. SEO SPAM AND MALWARE

To focus our research on the SEO spam and malware first we need to understand the SEO purpose. That way we will find the SEO spam attack target!

Search engines are based on a variety of factors, which helps to determine the rank of a website in the search engine results. One of the most important factors is the number and quality of incoming website's links. Essentially, the more high quality and relevant website's links, the higher its rank will be.

SEO spammers use different techniques to insert links and content on other people's website. A content, which they usually would not include in their websites. That way they are improving their website's search engine rankings. This approach is called spamdexing, which is also known as an attempt to manipulate the search indexes. [8]

Well-known SEO spam techniques are for example the keyword stuffing; doorway pages; link spamming in comments and forums and every technique, which gives the web pages an undue prominence in search results. [4]

A more advanced spamming technique is when a new page is created in the site, serving out its content of choice to any search bot that comes along. In a few extreme cases, installations were discovered that created entirely new sites to do the hacker's bidding. These new sites were located as a subdomain of the actual site, making it appear legitimate to the search engine crawl. [9]

That attack makes the user be redirected to an infected site with a copy of the original site, which even the site owners are not aware of that as they are maintaining their site but the search engines are seeing a completely different site during their crawl. [9]

A good example of spamdexing is the Pharma Spam, which is used by hackers to improve the SEO rank of their pharmaceutical ecommerce websites! Basically they are usually attacking an outdated websites with a plugin that can be exploited by attaching pages, links or files to the website or simply modifying the database! That can gain them a higher SEO rank and gain them more money by advertising. [10]

SEO spam looks like an easy attack to handle, but spammers are experts in hiding their work. They can integrate their malware so well that it can only be displayed from the search engines crawlers. The user and even the website's owner can see only the legitimate content.

According to Sucuri Company, over 50% of the web sites were hacked by SEO spam for 2018, which is almost 8% more than the previous year. [11]

With the following Figure 1 we can see that there is an upward for the second quarter of the year and then we can see a downward in the trends of those two threats. In addition, we can see there is a small upward for the last quarter for the SEO spams. Which makes those two threat always in the run!



Figure 1. Annual trends for SEO spam (in blue) and Malware (in red) [11]

2.2. DETECT AND PROTECT OF SEO SPAM AFFECTION

There are several signs, which can help to indicate whether our website is affected by SEO spam and here are some of them:

Google Webmaster (Search Console) and Google Analytics. Some of the most useful and powerful tools for identifying SEO spam are Google Webmasters and Google Analytics. In many cases, Google will discover that a site has been infected before the owner does. In that case, visitors may get a security warning from their web browsers that our site has been compromised, which is the worst way of discovering that. [12]

Google Search Console allows verifying unusual link or page activity and has penalized the web site. Some of the Google Search Console penalty notifications that might indicate SEO spam include: [8]

- **User-generated spam penalty;**
- **Unnatural links to your website penalty;**
- **Unnatural links from your website penalty;**
- **Hacked website penalty;**

- **Spammy structured markup penalty;**
- **The hidden text or keyword stuffing penalty;**
- **Cloaking or sneaky redirects penalty;**
- **Thin content with low or no added value penalty.**

Google Search Console also informs us of Security Issues affecting the website such as Code Injection, SQL Injection, Cross-Site malware, and server configuration, which are usually related to redirects to a malware-ridden site. [12]

In the Crawl section of the Google Webmaster, we can find a list of not found pages, which may be targets of incorrectly configured SEO spam. [12] [13]

On the other hand, Google Analytics provides traffic monitoring, which in case of increase may be related to SEO spam. The spammers might have started incorporating your website into their link farm, which has given your site a temporary boost. They may have also installed some spam pages on your domain, which sell products to visitors, and they are sending traffic to your website. [9]

Backlinks. A backlink is one of the most used words in the world of the SEO. It presents an incoming link to a webpage. [14] The number of backlinks on a website's rank it higher on all major search engines! [14] As the coin has two sides, the backlinks can be in a huge help, but also can cause a huge problem to the website as being one of the most common SEO spams. For that reason we are going to deeper investigate this topic in a further research.

Unexpected new pages or folders in the website. The appearance of new items in the content as pages, images, posts or even physical files on the host server are a sign of SEO spam attack on the site.

Third-party scanners. There are many third-party security scanners, which can detect if a website is SEO spam affected. As part of our research, we are demonstrating only free security tools! In addition, for this article the one we are using is https://sitecheck.sucuri.net. It gives the necessary information on detecting malware on our website by providing only URL of the site! On the screenshots below, we provide a sample of results from one of the site used in the statistics below.

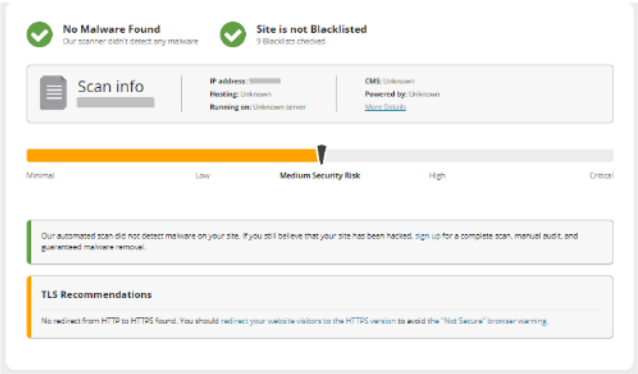


Figure 2. Scanner result – conclusion status

On Figure 2, we can see the results from the scanner with a risk evaluation of the website! On Figure 3 (extend of Figure 2) the scanner provides information if there were malware and security vulnerabilities detected and a blacklist status, which shows if the site was assigned as dangerous by Google or other authorities.

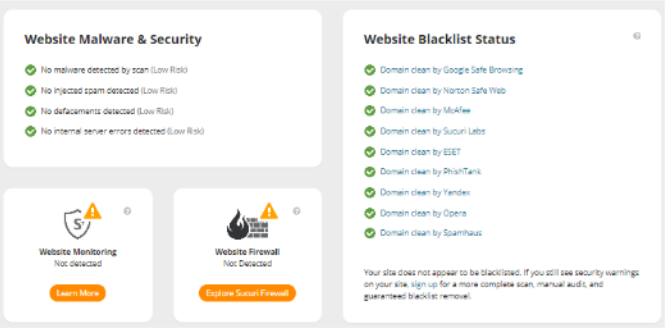


Figure 3. Scanner result – security status

Finally, the last section from the security scanner result provides a list of improvements, which will help in increasing the level of protection on the website (Figure 4).

We cannot fully protect our websites from all incoming attacks, but we can reduce the risk of a huge damage. Here are some good options in protecting our websites: [12]

- **Create a strong password with numbers and special characters;**
- **Create backups of the files and database on a regular basis, but keep in mind that the backup is a snapshot of the current website state, which is not guaranteed that it is a clean copy;**

- **If the website lets, users upload files, talk to the hosting company and ask them how to install antivirus to prevent malware; [6]**
- **Never install plugins or themes from an untrusted source – it can open up to potential unknown threats;**
- **Use of Google Search Console which can give daily basis or on request reports of the status of the website;**
- **Keep plugins, third-party software up-to-date;**

Use of malware scanner to detect SEO spam malware. Those scanners are designed to monitor the security and detect attacks against the website and provide information on how to proceed. The monitoring can be from internal and external point of view. The internal scanning place on the server itself and examines both file structure and site behavior. If a threat is detected, internal monitoring can take corrective action to remove blacklisting from websites. It can also issue alerts and other warnings to site owners and admins. The external from the other hand runs on the client side and works to scan the website from its public-facing components. It is not as powerful as the internal scanning, but since it does not require installation on the server and is not impacted by an infection on the site, it may be remains out of reach to the attack.

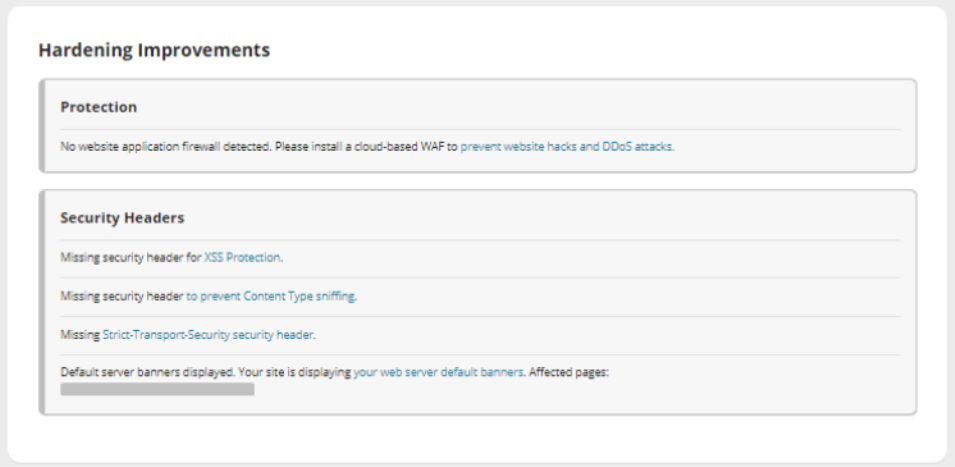


Figure 4. Scanner result – improvements

3. RESULTS

During our research, we are making several security checks on a few selected web applications. To follow the path we have taken we have again worked with the six website (from A to F) selected from ALEXA analytical insight - excerpt of 50 sites based on a selected Country category – Bulgaria. [2] [3]

In addition to our new knowledge gained from the beginning of our research, we have added 3 additional web sites (from G to I) developed by us! In those sites, we have included some protection methods, which we found during the research!

Table 1. Malware detection statistic

Detection	site A	site B	site C	site D	site E	site F	site G	site H	site I
Security Headers	+	-	-	-	-	-	+	+	+
Firewall	+	-	+	-	-	-	-	-	-
Malware	-	+	+	-	+	-	-	-	-
Risk Evaluation	low	critical	medium	low	low	medium	low	low	low

The sites selected for this article were selected on a personal opinion based on the most advertised and the most commonly used website. We are not going to present the name of sites due to law and policies restrictions! [2] [3]

Checking the selected sites with https://sitecheck.sucuri.net we can see that 1 out of 9 websites are with critical risk of being hacked.

The results show that most of the web applications we have selected have lower risk of being compromised. Moreover, considering that we have selected some of the most commonly used website in Bulgaria, that makes us think how much the content we provide to those sites is safe and how much we are exposed to malicious actions.

Moreover, we can see that the protection steps we have taken on the sites we have developed are good enough to increase the security level of the web sites and their content from malware and spams! The only missing thing is the firewall set–up which in the software development teams we participate in was not a part of our responsibilities.

4. CONCLUSION

There is no such thing as a completely secure web application, but experts give us several protection mechanisms, which will reduce the risk of being compromised and targeted. Most of all, the developer must be kept up-to-date to the fast growing Internet innovations. One of the most secure steps in protecting a web site. This can be used as a guideline for each member of the development team!

REFERENCES

Abuwardih, L., Towards Evaluating Web Spam Threats and Countermeasures, (2018), (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 10, 2018
Willian Dimitrov, ICT Security Trends, Sofia, 2017, Avangard, ISBN 978-619-160-766-2
Agrawal, H., (2019) What Are Backlinks in SEO and What Are The Advantages of Backlinks?, February 2019 <https://www.shoutmeloud.com/backlink.html>
Chalk, W., (2019) Cybersecurity in SEO: How website security affects SEO performance, February 2019 <https://www.searchenginewatch.com/2019/02/07/how-website-security-affects-seo/>
Darius, S., (2018) What Is SEO Spam And How Can It Hurt Your WordPress Site, June 2018 <https://blog.threatpress.com/seo-spam-can-hurt-wordpress-site/>
Darius, S., (2018) How To Identify That Your Website Hacked for Pharma Spam, May 2018 <https://blog.threatpress.com/pharma-spam-identify-hacked-website/>
Dimitrov, W., (2017) Software testing, Sofia, 2017, Avangard, ISBN 978-619-160-765-5
Dimitrov, W., (2018) ICT Security Model, Sofia, 2018, Avangard, ISBN 978-619-160-950-5
Petkova, L., (2017) SECURITY STANDARDS in software development,
Petkova, L., (2018) HTTP SECURITY HEADERS,
Petkova, L., (2019) CONTENT SECURITY POLICY VALIDATION,
Quttera, (2018) Malicious SEO Spam Making a Comeback, May 2018 <https://blog.quttera.com/post/malicious-seo-spam-making-comeback/>
SUCURI, Hacked Website Report 2018, 2019 <https://sucuri.net/reports/19-sucuri-2018-hacked-report.pdf>
Wordfence, Recovering Website SEO After a Hack, June 2018 <https://www.wordfence.com/learn/recovering-website-seo-after-a-hack/>