

A study on web as a boon and bane for today's online world

Goutham S, Soumya shree RK, Vini SP, Mahesh V
Jain University
BCA department

Abstract:

The web or also called as World-Wide-Web (www), is an advanced system of internet servers that is supported by highly formatted documents. These documents are created by a markup language called HTML (hypertext markup language) it supports links to other documents, as well as audio, video, and graphics files. The main issues of web are lack of relevancy, lack of organization and duplicate content. Advantages of web is that its ability to connect billions of people through computer devices. It creates convenience in sharing and receiving information between users and another main advantage of modern internet is its ability for automation.

KEY WORDS: ATTACKS,WEB,THREAT,AFFECT,ONLINE.

1. Introduction:

Worldwide Web (www) came into existence at the end of 1990's by Tim Berners-Lee it leads to all these important marks in the present world. Loads of information in one place which can be accessed by the touch of your fingertips by sitting comfortably your rocking chair, before libraries were the whole and soul of every information or its printed source material information, as we evolved so did our computer from punch cards to integrated chips, these advancements all lead to up gradation now it's all available in the web.

2. Background work:

Knowing all the authenticity of all these information and data people get corrupted which leads to lot of spoof sites which say that it gives everything but is full of malware or unwanted and inappropriate details , many wiki sites where the necessary information is provided voluntarily , mistakes are common and sometimes intentional and purposeful. Even though the internet is at its best to give quality information during an online purchase there are few who (cyber-criminals, hackers, etc.) take advantage of these utilities to snatch personal information and to spread out malware and potentially steal our hard earned cash. These are both the pros and cons of www which is left in our hands to utilize for our good or exploit it to harm us and others. Why the web? Because web can do whatever we want it to be, and we can modify to user's needs. We can shape it. We can mold it. Most importantly, every one each and every one on this earth can be connected this is an ability given to each person, as users and contributors, to improve their lives and communities.

3. Motivation for the above topic:

We usually tend to search all our required information from the internet or the web, but we have all faced some or the other problems at least once, to brief all the problems and give a proper and dignified solution we people are writing this paper, the factors affecting us while using the web motivated all of us to find a solution or a near solution for the above problem. In this paper we have tried our best to recognize most of the web related issues and have also mentioned near possible solution for it.

4. Challenges of web in real life:

4.1 Overload and excess information:

Information Overload is a major problem which occurs due to the increase in the amount of input to the system. Overloaded information has created increasing problems in our daily life including workspace. Here the amount of overloaded information has made the users unable to effectively process and utilize the information. The search in the web has become very slow due to excess information. It decreases the ability of decision making for users. The main causes of information overload are multiple sources, overabundance of information, irrelevant information, and difficulty in managing excess information.

4.2 Efficient information search strategy is required:

Search strategy is a combination of well organised key terms used to search the information. Search precision is improved by using the search strategy while searching in web. Search tools are rapidly increasing on the web nowadays. These are used as a form of artificial intelligence algorithms known as Heuristics, which are designed to help the users in learning, problem solving through self-educating techniques to improve their performance. We can use some special characters like Boolean logic and, or and not. For Phrase searching, it should be enclosed in quotation marks. By using

the Natural language searching we can type a sentence or question as we want it to be, the search tool will only determine the key word in the sentence and gives the relevant output.

4.3 Search is slow majority of the time:

Traffic density is one of the main issues that make the website search slow. Configuration issue in the domain is also a possible for slow search. If the images sizes are un-optimized then it takes long time to load the site. Using multiple sites at the same time increases the time to load. Poor server performance leads to long response time. Lots of traffic in the website may also cause delay in loading of sites. At this time the server can only serves a certain number of people at once, the more people use the site the slower it gets. Text graphics, too many file requests, unnecessary redirects causes loading problems. Out dated CMS, to avoid this problem one should update their software having the latest version helps in faster load times.

4.4 Difficult to filter and prioritize information:

As there are more information available in the web for a single search it is very difficult to prioritize the information as important and not. There will be more information without clear relevance. Users will feel stress when they will be not able to select what they want and what they don't want. Each information will be related to the searched keyword therefore it causes confuse to choose. It will be difficult to filter the information as all will be equally important in their own perspective ways. It leads in poor decision making and reduces attention time. As there will be more information overload it causes a major problem for the information managers to manage, filter and prioritize the information in correct manner.

4.5 Lot of apparently unconnected information:

Whenever we search for something in internet we get many results based on our search which will have both related and unrelated information's. For example:-people create multimedia and duplicate content and share it through online which contains inappropriate grammar usage. In some social sites there will be posts that are non-relevant to many of those who view them. In addition, many traditional media outlets such as newspapers and magazines have websites that continues to maintain the current events in physical form, as general people prefers this the most.

4.6 Large number of users cause internet to overload:

This happens when a website is accessed by too many users. Every website is hosted on a server. Imagine this server to be the shop and the doorway. The website is the product inside the shop and the doorway. The website is the product inside the shop which every customer wants to buy. When the number of people which the server (doorway) can allow in overflows, the server can no longer handle the requests from the user/clients. In such a scenario, the website crashes. Although this might be a very possible scenario, in current web hosting scenarios, most server spaces have the capability of allowing quite a large number of users access to the content concurrently.

4.7 No restricted regulations:

It can be observed that with a large amount of information freely available on the internet, theft of personal information and misuse of this information is in abundance. In this regard from time to time people use some one's information and research materials and pass it off as their own work. Also, spamming, which is the process of sending unwanted or junk email's in bulk, which provide no purpose and consequently hinder the entire system. This in due course is regarded as an illegal activity resulting to frustrate people.

5. TECHNIQUES TO AVOID WEB PROBLEMS :

5.1 Search engine optimization:

SEO is short for search engine optimization or search engine optimizer. Search engine optimization is a methodology of strategies, techniques and tactics used to increase the amount of visitors to a website by obtaining a high-ranking placement in the search results page of a search engine (SERP) -- including Google, Bing, Yahoo and other search engines. SEO helps to ensure that a site is accessible to a search engine and improves the chances that the site will be found by the search engine.

5.2 Contemplate in advance the kind of information we seek:

We as users should have reasonable idea of the kind of information we need to gather. Information will encompasses news about your industry or profession; few notable product and service developments; significant regulations and new legislation; client, customer, or consumer-related information; special applications etc.

5.3 Web filter:

A Web filter is a program that can screen an incoming Web page to determine whether some or all of it should not be displayed to the user. The filter checks the origin or content of a Web page against a set of rules provided by company or person who has installed the Web filter. A Web filter allows an enterprise or individual user to block out pages from Web sites that are likely to include objectionable advertising, pornographic content, spyware, viruses, and other objectionable content.

5.4 Blocking websites:

Blocking or banning is a technical approach intended to restrict access to information or resources. Blocking and unblocking may be implemented by the owners of the computer using certain software, blocking also refers to denying access to a web server based on the ip address of the client machine.

6. Our view on the solutions which can be implemented:

6.1 Self-hosting:

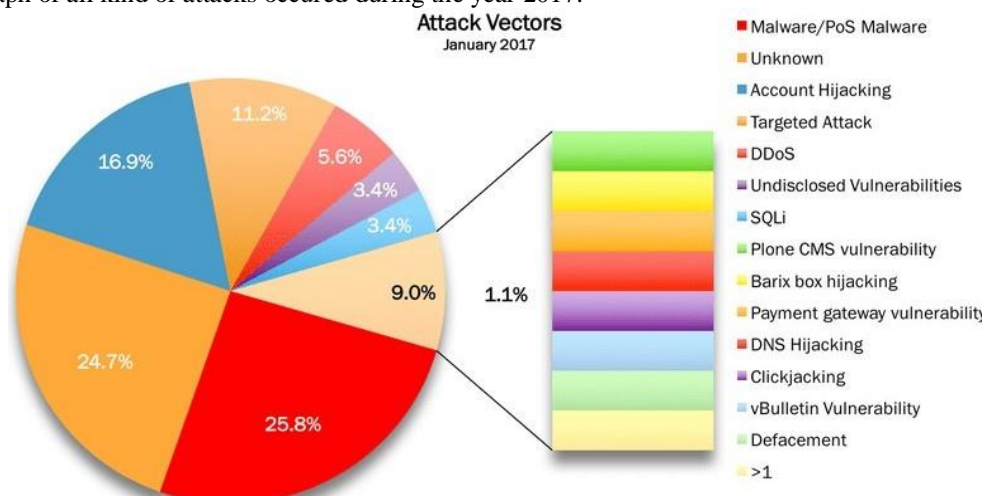
Self-hosting is the best option for web development however this point relates more towards website construction and lesser related to search engine optimization, but one needs to ask “How is it possible to optimize a website effectively, until and unless it has a solid base ? And the base of any website is the server hosting as well as the coding job. It is very difficult to maintain a self-hosted website as lapses in updating maintenance are inevitable. Such lapses take a toll on your SEO. Why should one play around with risky technologies like self-hosting, which are often prone to so many errors and technical glitches.

7. All possible threats on web :

The types of attacks happening on web and its percentage.

8. Attack vector of 2017:

Complete graph of all kind of attacks occurred during the year 2017.



(figure 1.Attack vectors)

8.1 DDoS

Distributed denial of service is a type of denial of service(DOS), Where multiple compromised systems are used to attack a target such as servers. Here an attacker sends a flood of requests or messages to a targeted system which forces it to slow down or even crash the targeted system, thereby denying the service to the legitimate users.

8.2 Account hijacking

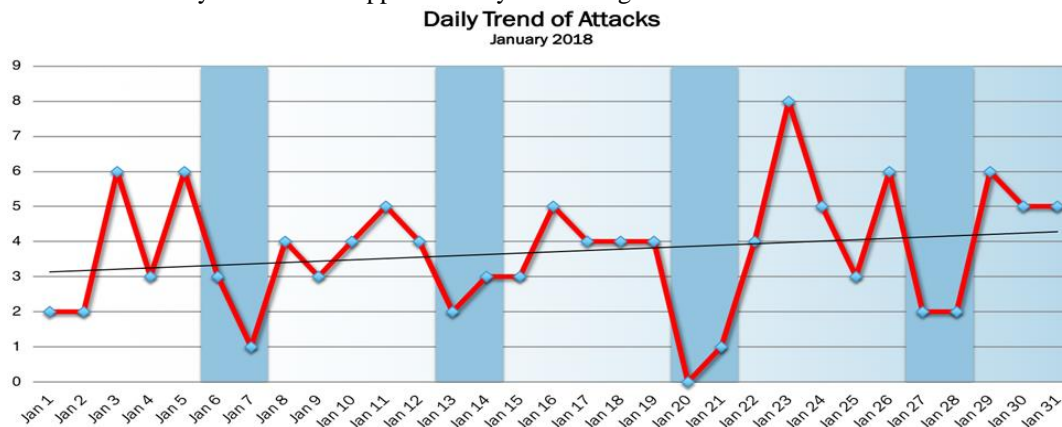
Account hijacking is the process through which an individual’s email account or any other account connected with a devices or services is hijacked by a attacker. The attacker can use the personal informations of the respective account that is been hijacked to perform malicious activity.

8.3 Malware /POS malware

Point of sale malware is a malicious software that is used to steal the customers personal data especially credit card details(financial details) from retail organizations or from business checkout systems . Here the attacker steals the customers informations with the intention of selling the data rather than using it themselves. A pos malware commonly is used by cyber criminals.

9. Daily trend of attacks:

This graph shows how many attacks can happen in a day and through the month.



(figure 2. Daily trends of attacks)

In Jan 23 2018, there was increase in the number of threats up to 8.

Jan 20 2018, there was a decrease in the number of threats that is 0.

10. Conclusion:

“Giving better problems and its solutions on web usage and its background is what we would really like to accomplish in this paper of ours. Making web a completely safe environment for everyone would take a very long time to achieve, henceforth let us all practice a healthy way of using the web. We will all prepare ourselves to tackle all these web related issues and strengthen it up. Our goal might not be easy one to achieve but the things that are worth doing are often not easy”.

References:

- Web services security: challenges and techniques(Author: Anoop Singhal) bologna, It Published in: Eighth IEEE International Workshop on Policies for Distributed Systems and Networks (POLICY'07) bologna, Italy.
- <https://webfoundation.org/about/vision/why-the-web/>
- <http://creative.sulekha.com/>
- Sun, Q.D. Li, "Defending DDoS Attacks Based on the Source and Destination IP Address Database", *Journal of Software*, vol. 18, no. 10, pp. 2613-2623, 2007.
- Y. Xie, S.Z. Yu, "Anomaly Detection Based on Web User 'Browsing Behaviours", *Journal of SoftArt*, vol. 18, no. 4, pp. 967-977, 2007.
- C. Siaterlis, V. Maglaris, "Detecting incoming and outgoing DDoS attacks at the edge using a single set of network characteristics", *Proc. of the 10th IEEE Symp. On Computers and Communications*, pp. 469-475, 2005.