ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

Adding Trust to the Social Networking Sites: A Prominent Way of Content Filtering

Aman Sharma[1], Atul Arora[1] , Abhinav Sharma[1] Akshant Gupta[1],Ankit Malik[1]

B.Tech Student, Dept. of Electronics & Computer Engineering, Dronacharya Collage of Engineering, Gurgaon, India [1]

ABSTRACT: This paper reviews about the internet filtering and content filtering techniques of the web pages that helps in removing particular types of encrypted data from that webpage. As soon as the social networking and the internet came into existence the transfer of various types of viruses and unauthorized data which is illegal in law have became a biggest threat to the industry. So, Saving our document from these types of unauthorized viruses and encrypted files is a major issue. Various organizations are working on it to prevent the free flow of these type of viruses from internet to our system and for that reason content filtering has gain importance in making our documentary files secure. This paper includes various types of filters, bases to select a filter and methods for filtering with the conclusion of the content filtering.

KEYWORDS: Filter, Domain, Internet.

I.INTRODUCTION

Today, social networking sites had gained popularity. They have become the greatest source of communication, share a considerable amount of human views, exchange of several types of data like text, videos, etc and even provide employments. Suppose, if somebody wants to search any desired information here come the data mining process which is known to play the most efficient part in it. The search engine selects the input, compare it with all the record saved in the database of the engine. The record in the database is again selected and respective url is displayed. Now generally the information, advertisement or message given on that webpage is shown as such without any change known as general walls. These walls may have that advertisement which is of no use of that user or may have unauthorized photos or encrypted data.

So keeping all this is mind, social networking sites must be equipped with the facility that if something unauthorized is available there on the webpage we have searched for, it will filter the page and remove all its unauthorized contents. This paper reviews about the various techniques and methods of doing it.ML technique[1, 2, 3] and several other methods [4, 5] are also used to study Content based filtering.

II. MODERN TECHNIQUE OF FILTERING

An Internet filter can be something, not necessary required to be only hardware, it can be software well, used to check the flow of undesired data that is present on the internet or travels through the webpages when we click on them. Filters plays a major role in not allowing the free flow of unwanted Advertisement or any vulgar images to show on the web pages. Filters also plays a great role in passing on the free flow of harmful files that can infect our system and lead to any miss-happening. If used in a useful manner these Internet Filters can play crucial role in blocking the access of certain sites, web pages, videos and certain porn cites as well. The greatest achievement in these filters is achieved due to a new emerging technology named as Machine learning (ML) text categorization procedure.[6]. ML technique breaks the content of messages into several parts and place them according to the different categories that may have been created. For example, all the pictures are placed in one category while rest text is placed in another category and so on. On the basses of these categories, only allowed contents are allowed to display and rest of them are cleared off.

In 2002, a warning was issued by the president Hinckley that we are using several techniques to safeguard our home and electronics appliances but we never worried about the security on the internet [7]. Therefore, we must pay some attention to our social networking safety as well. One might be shocked to hear that these social networking sites, website, etc are much easier to crack or hack than our physical Quantities like Electronics gadgets. But nobody pays attention to it. It is highly required to block the flow of harmful contents and that's what the benefits of these filters.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

Many filters provides us the facility of checking the flow of any harmful material and we can used any of them to protect our sites. Several Advanced content filters even provide you security in chatting sites, blocks the unnecessary advertisement on your processed webpage. One might have noticed that there are several Advertisement that may come to see when you are surfing any webpage. These Internet filters provide us the facility to get ride of such unwanted advertisement. Suppose if two users are present in a trusted site in a direct relationship of some type say RT and X then there must be an edge between them which will be connecting them i.e the messages will be directly transferred from one node to another node. [8].All these features of Content filtering keep us safe from the harmful data transfer online and keep us intact on social networking sites.

But these have a major drawback that it can be overcome by providing some certain conditions to it that are specified in its library and hence the filter can be of no use when we talk of the security that it was providing us. But you know at the same time, it's the part and parcel of the game that if something big positive is coming our way then it may have certain negative points also which can be neglected keeping the values of positives.

Due to advancement in the several hacking and account breaking techniques, it is highly recommended to filter the content that one surfing on the internet. So it is highly advised to use you check the capability of your filter that one must use the internet in public areas as to check the maximum unauthorized filtering of data flowing. But if somebody wants to be sure of security then he/she must not use the internet when there is somebody around him/her. This may increase the threat of password leaking and hence destroy security. Also one must use those Operating Systems which have the availability of Internet filters like Windows 7, vista, mac, etc.

In a conference, once President Faust said that As the number of user on any sites increases, user generally shifts from general filters to his/her personal filter to block any untrusted things to flow through. [9].

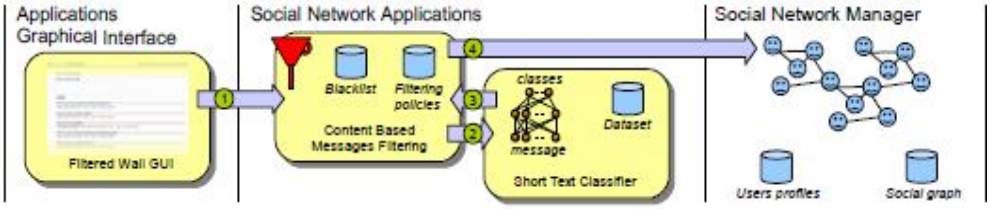

Fig.1 Filtered wall Conceptual Architecture.

III. VARIOUS FILTER AVAILABLE TO APPROACH CONTENT FILTERING

Let us consider a thing that everything that is available on the internet is not secure or it is affected with any untrusted virus which is really a big threat to our system. Thus, keeping the same in our mind, let us make a filter which will safeguard us online, checking everything that we surf, see weather if we are visiting a malware affecting site or a secure site. In case if we are visiting a affected site it will automatically starts it is functioning. For Example if we had a filter that filters any affected mails from our mail account, which automatically removes any infected file or mail from it, if it detects any. Since we are not require to deal with it our self and hence, provide security.

These filters have a very basic way of operation. Before using them, we are just required to set what kind of file we are setting for it as a malware affected file and which one have the right to access through. Keeping all these key points in mind, these filters are grouped into various categories named below-

Using Software based Filters- Software based filters are those filters which are mostly used because of their popularity and ease of usage. These types of Filters are easily available on the web and they can be easily downloaded from there. After they are downloaded, they are installed and then they are all set for usage. Generally we have different types of filters available and they are general purpose i.e they are only use for specific task only. For example we different types of filters available for firewalls, mails extraction and sending. Basic working of these filters is explained as, these filters acts like a kind of interface between your system and internet. Everything you search on the internet if first checked by this filters and then passed on to your system. You are first require to define the set of operation it have to do in beginning. They provide us a good security not even in the case of pornography but also protects us from other threats.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

If you want to compare it with your daily life, it may be considered with the security systems that are installed in banks or high security places. Guest considered as incoming data have to pass on through the security checks and prove its identity so that they can enter through. The same is the functioning of these software filters. If you had the set of instructions that don't allow such types of instructions to pass on then access id denied.

Using Hardware based Filters- These types of filters are also used but most of the people do not know where they actually using them. Generally people think that there is only one type of filter available that is software based filters but that is not what truth is. There is another types of filter which are named based on hardware based. Today people use internet connections at their homes through several means, either they may be using a modem or wireless internet likes wifi etc. These modems or routers have in built data filtering capability. What you require to do is to connect the modem externally to the system, these modems are having a capability to filter the data across the internet. You are not require to install any software further for this functionality. These modems have the access of only trusted and secure sites. So these modems are the best example of hardware filters. These filters are easy to use and provide us the access to the internet with the security that the data is secure or filtered.

Hardware Filters are best understood with the example of a person who is standing just in front of a mailbo who is just checking that whatever mails somebody is actually drops to your mailbox. If these mails were actually belonging to you it will accept otherwise rejects.

Domain Name System- A modern approach- DNS is the modern way for content filtering which provided several benefits of better online security and data filtration.

Another option is to use a DNS service such as OpenDNS to provide filtering. This provides a free option with many of the benefits of having a hardware solution, without having to purchase additional hardware because it will probably work with your existing home networking hardware.

All that is required is to update the primary and secondary DNS entries at the router to point to the opendns servers, and then open an account on open DNS, which allows you to set your filtering options for your home network. Directions for doing so are on the open DNS website.

This option has the advantage of filtering all of the devices in the home; including the computers, cell phones, TVs, game consoles, any device that would connect to your router, either hardwired or wireless. However cell phones connected to the Internet via a cell tower will not be filtered.

Using Internet Service Providers – A Magical Technique.- Well, it might be surprising that even our internet service provider can even provide us the content filtering technique. The reason why content filtering using Internet service providers is magical because one didn't need anything to be downloaded or installed on your system. Please make note that not every service provider gives you the facilities of content filtering but there are only limited one. The main thing in this is that it just present online so there is no worry for you that you are required to download anything or install anything. It is just present between the data you are surfing on the net and your system which checks for the infected file and after checking it passes on to the user. If it is infected, it blocks it and don't allow it to pass through. The key factor that one must be caring about is that the system you are using must be configured with the proxy server.

This can be better understood by the example of a postman which delivers the letters to our home after inspecting them. They are infected, he will not deliver it to us and if they are all ok, means not infected, then it will delivers them at our home.

IV. MAJOR ISSUES CONSIDERED BEFORE SELECTING ANY TYPE OF FILTER

Up to now, we have discussed about the different types of Filters, like hardware, software, DNS, etc but now we will be taking issues that will matter while selecting a special type of filter from all available. The type of filter which you might be selecting will depend upon the following under defined points.

1. Thing which is you trying to save from the untrusted sites?- Now a days, in the modern era, we have seen several devices which uses internet connection and therefore they are required to be protected from the infected sites. One must remember that one is not required to protect older desktops or older versions of software but one is also concerned about the newer technology based devices like palmtops and utrabooks that works on net connection and help us in

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

doing several tasks. One must note that your mobiles does not comes into this category as they uses different net connection for surfing.
if somebody is using the internet connection for several devices, then it must be a broadband connection. It is possible that the devices that one is using may not be functioning well with the hardware one is using so one must be careful about that as well.

2. O.S that one is using in desktop?- Operating System is the interface that act between hardware and software which helps user's functioning on desktop. Which operating system you are using can be identified by rebooting your system. The name that one might see on booting like window xp,7,8 etc tells you what operating system you are using.
First of all, what we require to do is to check what hardware filter one is using. Although it has nothing to do with what operating system one is using. In some operating system, we have couple of hardware filters present or may be more than that. In new versions of operating system like window 8 and mac these hardware filters are already been installed up to the newer version. System's software filters are even better than hardware filters providing better security from several internet issues but they were not present in older versions of operating systems.

3. Using internet at home or also at public places?- This point is although is of less importance but can't be avoided. Is somebody is having desktop then it is highly possible that he/she will be using internet at home only at chance of connecting that desktop away from home to any un-trusted network are pity less. But if somebody is using a laptop then there are fair amount of chances that the person can use internet away from the home so one has to be sure about the security. Therefore, they require a software based filters that they can use to protect their laptops to connect to any fishing site as hardware filters can protect only when one is connected to internet through hardware based internet services.

4. Way by which one is authenticating to the internet?- It might be of less importance and somebody will find it really a unique thing when one say that if somebody is authenticating to the network by using a modem or any WIFI connection then the use of hardware based filters will not be the best option but one must think of using of any software based filter. Also if one is accessing the internet through LAN cable then the hardware filter will be the better option.

V. HOW TO BLOCK A SITE BY USE OF HOST FILE

Sometimes we do not want the users of a particular computer or network to open a particular website. For example we want our children or our friends to stop accessing a website then we needs to block that website on the computer or network. Although web browsers provide facilities to block a website but they do not work always and it is quite easy to undo the settings. So through this guide I will show you how to block websites by using computer's host file. By settings a password for your host file, other users will not be able to unblock the websites back. Here are the steps:
1. Open the 'Command Prompt' and type "notepad C:/Windows/System32/drivers/etc/hosts" and start instead of moving cursor just start typing, the host file will be open in Notepad.
2. Now in the host file find the line that says: "127.0.0.1 localhost,". Bellow this line type the IP address and name of the website that you want to block separating by a space. If you want to block more than one site write the similar details in individual line. For example if you want to block Facebook.com then write the following line:
66.220.158.11 www.facebook.com
3. To find out the IP address of a website open the command prompt and type "ping www.facebook.com" and press Enter key. If a website has more than one IP address then write each one with the website name in separate line.
4. Save the host file after adding the details of all website that you want to block and close the host file and command prompt.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

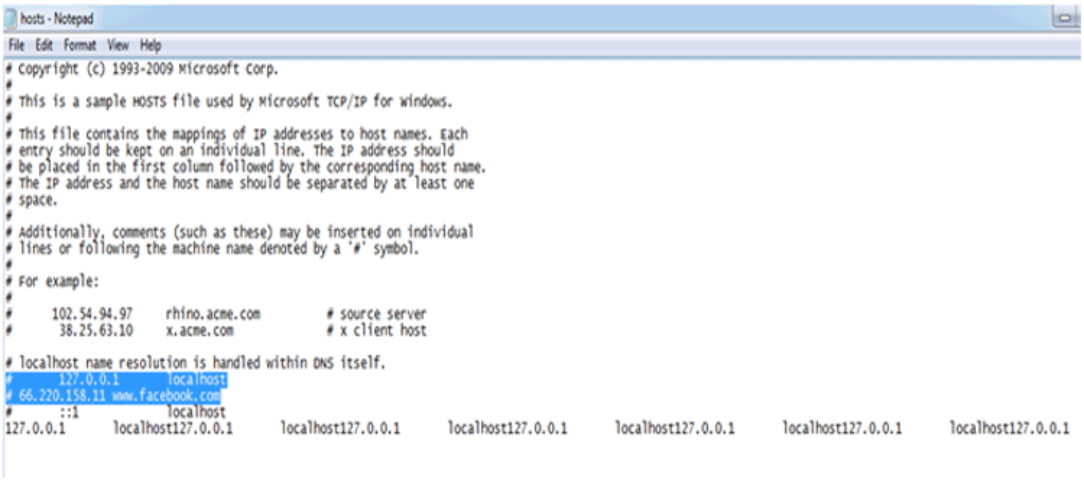

Fig 2. Blocking Facebook on any System.

VI.CONCLUSION

As now we can conclude that from the past few years with the advancement in the field of internet, the issues of online protection had gone up. It is highly required these days to protect your laptop, desktop or nay other device using which you are connected to this virtual world of technology. With the advance ment of several techniques people feel highly secure when they click on any unknown site because they know that if there will be anything fishing behind that then his/her firewall or filters will provide him/her security.

Content filtering is the fast growing field in which in every second any new infected site is produced and on the same second, we generate the way by which the previous junked site can be detected and avoided. Hence, it is ongoing terminology which will keep on fighting with the people's FACEBOOK, twitter, YOUTUBE conflicts for a better online world which will be secure from fishing sites.

REFERENCES

1. Amati, G., Crestani, F.: Probabilistic learning for selective dissemination of information. Information Processing and Management 35(5), 633–654 (1999).
2. Churcharoenkrung, N., Kim, Y.S., Kang, B.H.: Dynamic web content filtering based on user's knowledge. International Conference on Information Technology: Coding and Computing 1, 184–188 (2005).
3. P´erez-Alc´azar, J.d.J., Calder´on-Benavides, M.L., Gonz´alez-Caro, C.N.: Towards an information filtering system in the web integrating collaborative and content based techniques. In: LA-WEB '03: Proceedings of the First Conference on Latin American Web Congress. p.
4. Hanani, U., Shapira, B., Shoval, P.: Information filtering: Overview of issues, research and systems. User Modeling and User-Adapted Interaction 11, 203–259 (2001).
5. Churcharoenkrung, N., Kim, Y.S., Kang, B.H.: Dynamic web content filtering based on user's knowledge. International Conference on Information Technology: Coding and Computing 1, 184–188 (2005).
6. IEEE Computer Society, Washington, DC, USA (2003).
7. Sebastiani, F.: Machine learning in automated text categorization. ACM Computing Surveys 34(1), 1–47 (2002).
8. Overpowering the Goliaths in Our Lives, President Gordon B. Hinckley, Ensign, January 2002.
9. Golbeck, J.A.: Computing and Applying Trust in Web-based Social Networks. Ph.D. thesis, PhD thesis, Graduate School of the University of Maryland, College Park (2005).
10. The Power of Self-Mastery, President James E. Faust, Ensign, May 2000.
11. Pornography, Elder Dallin H. Oaks, Ensign, May 2005.