



# INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

## Website Blocker Using Python

DR.BALAJI K

**Abstract**— According to the Google Search Statistics, Google processes more than 3.7 billion searches every day, with an average of 40000 searches per second. This illustrates our extreme dependency on the internet for all our daily tasks ranging from complex projects on space science to simplest queries like what is the right way to greet a person in the evening. Our searches are answered through interlinked web pages, known as websites. As of September 2022, there are over 1.98 billion websites available online. Apart from the benefits of having access to huge content and numerous resources on our fingertips, it can also be a great source of distraction and mismanagement of time. Constant online presence of an individual may affect an individual's ability of achieving professional and personal goals. Hence, website blocker is one effective tool that helps an individual take a conscious call on blocking a website. There may be many reasons for blocking a website, which will be discussed in the following sections of the paper. A website blocker, if effectively utilized, can enhance focus, productivity and value of time of an individual and the organization as well. Here, we propose a simple website blocker implementation in Python.

The domain of web mining prominently deals with the implementation of data mining techniques and methodologies to extract the information from the web documents and services. The main objective of web mining is to discover the patterns in the web data by collecting and analysing the data from web sources like web content, server logs and hyperlinks. These patterns will help us to gain an insight into the current trends of the industry and tendencies of the market. The major source of web data is a website. A website is nothing but a collection of interlinked web pages. At present, there are more than 1.98 billion websites which serves information exchange, market expansion, advertising and almost any purpose that we name. Just like a coin has two sides, there are some websites that do not serve a fruitful purpose and in turn be a source of distraction and contain inappropriate contents. Here the website blocker comes into the picture. A website blocker helps in blocking a website with a user's consent for the desired duration. There are several reasons where the necessity of blocking a website arises. Some of them are:

- The irrelevant websites cause unproductivity at workplace.
- Attackers use websites as an easy mode to invade privacy of users.
- Some websites contain malicious content that distributes virus into the client systems.
- There are websites that promote ambiguous content that may raise legal disputes.
- There are several attacks through websites that may pose a threat of theft and hacking to access the sensitive data.

The websites that need to be blocked can be broadly classified into four main categories:

- **Unproductive websites**  
Websites that are irrelevant to a particular context of time and place need to be blocked. For example, the shopping websites should be blocked at offices to minimize distraction among employees. Apart from this, employees also spend their time on social media websites which is a threat for any company's growth.
- **Inappropriate websites**  
Websites containing indecorous and adult contents should be blocked in schools and other educational institutions, as they can be harmful and detrimental in a student's life.
- **Dangerous websites**  
Websites that are designed to look very innocent but are involved in attacking and hacking the system should be blocked as they can spread viruses and invade user privacy.
- **Confidential websites**  
Websites designed for some secret missions, or websites under development should be blocked to prevent theft of sensitive information.

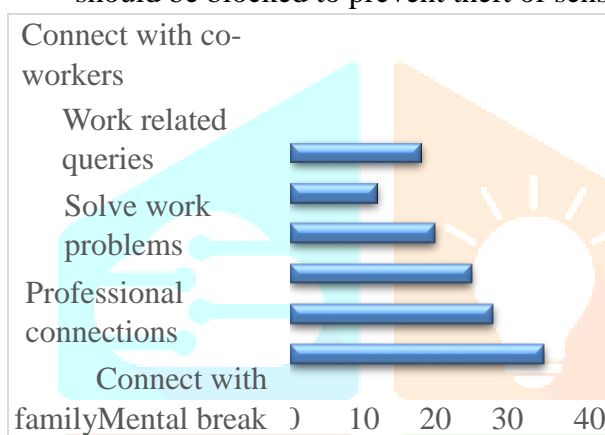


Fig. 1. Online presence of employees during working hours.

The above figure (Fig. 1) shows that 12% of employees use websites productively at the workplace to ask work related queries and 24% of employees use them to build professional connections. Other employees use the websites for surfing through social media which reduces efficient time utilization at the workplace.

## I. LITERATURE REVIEW

A literature survey of over 6 research papers was carried out which gave good insights and conclusions on the topic.

In a research paper written by Veronica Marotta and Alessandro Acquisti, titled "Online Distractions, Website Blockers, and Economic Productivity: A Randomized Field Experiment", an application named "Freedom" has been designed which allows users to create their personalized blocklists which is a list of websites that users want to block while working. The paper concluded with a very important statement that, the productivity increases for individuals who use the tool for being focused.

In another research paper written by Aditya V Shetty, Algeena Carol Dsouza, Aditya Maruti Naik and Reena Lobo, titled "Ad Block and Malicious URL Detection System", a pattern matching mechanism is utilized. We install an extension, feed the URL and validate it. In the validation process, the URL is compared with the blacklisted URLs. If the URL entered is found to be secure, then the user is redirected to the website. If it is found to be in the blacklist, then the mechanism stops the redirection to the website. The main problem is creation of blacklist. With over 1.98 billion websites and several being created every day, it is not quite feasible to be able to include all the malicious websites in the blacklist.

Another survey paper titled, "Internet Blocking in Public Schools" which gives many insightful and thoughtful statements that blocking software does not protect children from exposure to a large volume of material that is harmful to minors within the legal definitions. It also states that the teachers and students above 17 years of age should be exempted, yet suffer the consequences of CIPA implementation.

We then referred to an international study under the title "How do students feel about web filtering?" by Marisa L. Ahmed. In this study, a survey was conducted for 400 students which gave some astonishing facts

and figures. It stated that on average, students spend around five hours per day using display devices. Around 53.6% of students reported of being side-tracked sometimes and further admitted that they were focused only for half of the time. Only 26.60% of students reported of school related work being the biggest usage of their online time. Rest of the students admitted that their major usage of online devices was gaming, entertainment and social media.

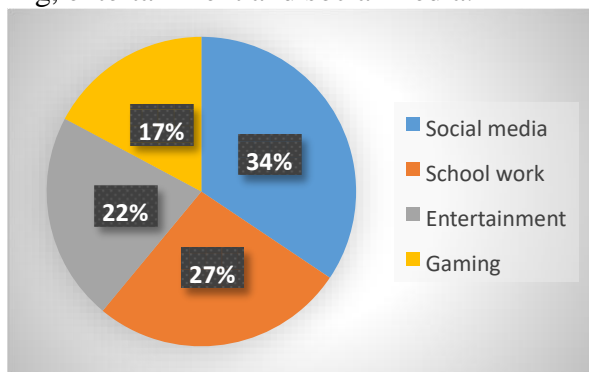


Fig. 2. Websites visited by students for various purposes on daily basis.

The figure above, (i.e., Fig. 2) shows the survey results presence is more associated with distractions and killing of time rather than effective utilization.

From the literature survey, we realized that there are many suggested ways and tools to block the websites and avoid distractions, from the above literature surveys. But there is a potential gap between the users and such tools because the users fear and doubt the integrity and security of these tools as well. They are hesitant to use them as they are afraid of data leaks or data thefts through these applications and websites themselves. One cannot trust a service given for free blindly. Also, one single application or tool may not be an optimal fit in all places. Some are well suited for personal use, some are designed only in organizational point of view and some tools may be potential for blocking only particular type of websites. Hence, we come up with a simple, scalable and a user-friendly implementation of a website blocker.

## II. METHODOLOGY

The simplest way to block a website is by adding the domain name into the Host file of our system. Host files are basically text files that map the IP addresses to the domain names. The functionality is very similar to that of a Domain Name System (DNS). When we try to access to a website on a web browser, we do not mention the IP addresses, instead, we just write the domain names. But to establish a connection, we need an IP address. So the PC checks if the website and its IP address is specified in the Host file. If not, it fetches the IP address from the DNS. The manual blocking is illustrated in the Fig. 3.



Fig. 3. Manual blocking of websites

After understanding the role of Hosts file, which is editable, a website can be easily blocked by providing a mapping of domain name that we want to block with a loopback address. Hosts file contains two columns – one for IP address and another for the respective domain name. So, we add a default IP address 127.0.0.1 mapped with the website that we want to block. The IP address, 127.0.0.1 is the loopback address, which is also known as Localhost in the computer networking. The Localhost is the default name given to the computer on which we are working. This address helps the system in connecting and communicating with itself, which makes it known as loopback address. When we add a

particular website to be blocked along with the Localhost, it is prevented from searching it on Internet and rather searched for it on the local server. But doing so, will not answer the request. connected to or there may be some connectivity issues. Now, we can simply say that the website has been blocked.

#### ALGORITHM

1. Start
2. Enter the URL to be blocked
3. Click "BLOCK"
4. URL added to Hosts file
5. The URL is blocked
6. End

Fig. 4. Algorithm for blocking a website



Fig. 5. Flowchart for blocking mechanism

This simple task can be made even simpler. Going to the Hosts file each time and editing can be tedious, boring and time-consuming in itself, as the path of Hosts file is quite longer one. Hence, we came up with a Python implementation, that would do the same job but with minimal effort and minimum time. First, the user identifies and feeds the sites that are distracting, unproductive and disturbing him/her. The websites that seem to be inappropriate and dangerous can also be blocked right here. There is complete flexibility on the number of websites a user wishes to block, as demonstrated in Fig 4. Also, there might occur instances where, a user may want to reconsider and unblock the blocked site and our model gives that option as well and the mechanism is shown in Fig 5.

#### ALGORITHM

1. Start
2. Enter the URL to be unblocked
3. Click "UNBLOCK"
4. URL truncated from Hosts file
5. The URL is unblocked
6. End

Fig. 6. Algorithm for unblocking a website

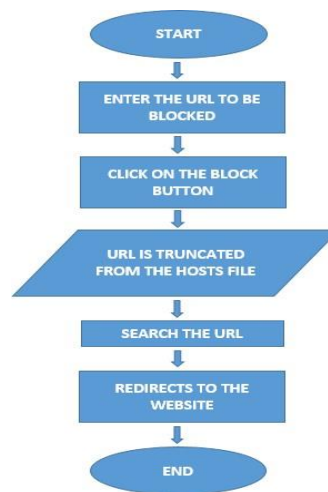


Fig. 7. Flowchart for unblocking mechanism

### III. RESULTS AND DISCUSSIONS

After When we suggested the website blocker to our peers, they found it interesting and very simple. Most of them, however were not aware of blocking the websites through the Hosts file. The simplicity of the blocker where you just have to enter the domain name and block it in a single click made it easy. Also, the idea that unblocking is equally simple, made it usable daily. The user needs to have only a Jupyter notebook or Google Colab, where he can execute the code and easily implement the model. It can help us block few distracting websites like Instagram, Facebook, etc. during study hours and unblock them during our leisure time. The complete control of access to websites will be in our hands. We can redirect the blocked websites to a page that depicts our goal and motivates us or it could be a picture of our role model that would keep us focused.

The implementation of code results in the following outputs:

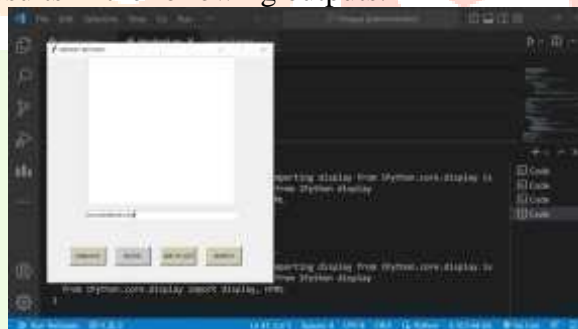


Fig. 8. Output GUI Window

### REFERENCES

- [1] Veronica Marotta and Alessandro Acquisti, "Online Distractions, Website Blockers, and Economic Productivity: A Randomized Field Experiment", Heinz College, Carnegie Mellon University, June, 2017.
- [2] Aditya V Shetty, Aditya Maruti Naik, Algeena Carol Dsouza and Reena Lobo, "Ad Block and Malicious URL Detection System", International Journal of Engineering Research & Technology (IJERT), Mangalore, India, 2019.
- [3] Stefan Kopsell and Ulf Hillig, "How to Achieve Blocking Resistance for Existing Systems Enabling Anonymous Web Surfing", Institute for System Architecture Dresden University of Technology, Dresden, Germany, 2004.
- [4] "Internet Blocking in Public Schools, A Study on Internet Access in Educational Institutions", 2003.
- [5] Marisa L. Ahmed, "How do students feel about web filtering?", AN INTERNATIONAL STUDY, University of California, Berkeley.