Notorious Ways of Exploitation Search Engine

Sohesh Doshi[1] Meet Gadhiya[2] Sanket B Suthar[3]
[1,2,3]Department of Information Technology
[1,2,3]Charusat University, Changa, 388421, India

Abstract— As time goes the use of internet & technology increases and become the part of the human in routine life. All people continuously use to share their information either it's private or public via various source on the internet. This information directly and indirectly interact with Google database searching the keyword on Google search engine give precise information but in vary large scale ,but use of Google dorks help us to narrow this scale and retrieve specific result. This research paper's objective is mainly focus on how hackers use Google dorks to steal information on the database.

Key words: Google Search Engine, Exploitation Search Engine

I. INTRODUCTION

In Our Daily Lives, we are connected to the internet. 32 billion people are online on the internet. In 1995 it was less than 1% people who used the internet. In 2016 there was more than 3 billion user. As per the current situation we do analysis of internet users by regions worldwide since year 2000 to January 2017.

Year	Internet User	Non-Internet User	World Population	%
2000	414,794,957	5,711,827,164	6,126,622,121	6.8
2005	1,030,101,289	5,489,534,561	6,519,635,850	15.8
2010	2,023,202,974	4,906,522,069	6,929,725,043	29.2
2015	3,185,996,155	4,163,475,944	7,349,472,099	43.4
2017*	3,575,870,700	3,911,038,788	7,486,909,488	47.76

Table 1: Internet users

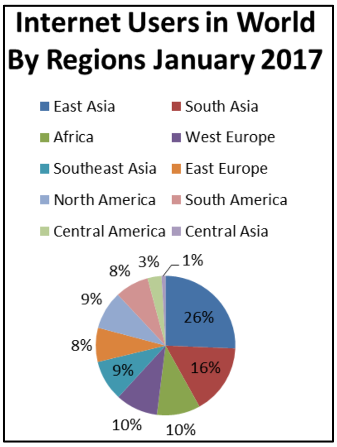


Fig. 1: Internet Users in World by Regions January 2017

Major Actives Performed by people on the internet are sent & Read Email, a search engine to find information, search for map driving direction, Look for info on a hobby or interest etc. 85% People are Used Search engine 6,586,013,574 searches a day worldwide Google can process 40,000 query per a second, 3.5 billion per a day and 1.2 trillion searches per a year [6].

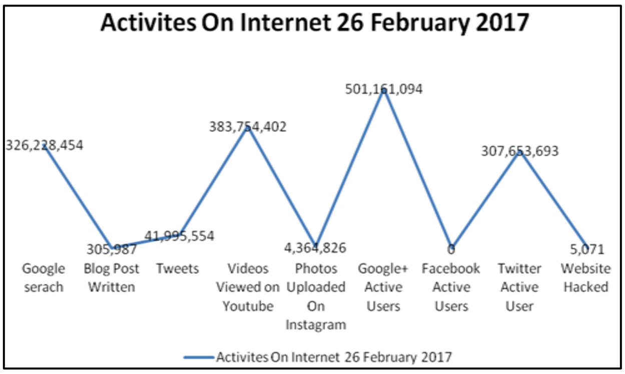


Fig. 2: Activities on Internet by 26 February 2017

Steve Mansfield-Devine says that instead of using penetration testing software like NMAP, Nessus etc. use Google and your first stop should be something much more basic: Google [5].

II. Review of Literature

A. Google Basic

When you search in a Google search engine it gives appropriate search compare to others Search engine.

Google used some special kind of algorithm while other search engine doesn't. Google used some auto med program called "spider" or "crawl". Google also used some trademarked algorithm called "Page Rank" which assigns each Web page a relevancy score [4].

This Algorithm depends on a few factors:-

- Frequency and location of the keywords within a web page: If some of the keywords have a low frequency on a web page then it assigns law rank to the web page.
- How Long Webpage has Existed: People created web page daily goggle assign that web page how they important.
- The number of other Web pages that link to the page in question: Google shows that how many other web pages link to that site.

The third rule are easy than others let's consider example we search keyword "Techno" in Google. Google show that pages on top whose rank is important or more than others. Because Google looks link to the web pages as a vote. The Information we search on internet first content discovered then indexing and analysed than content and store in a huge database when some query fetch gives a response or show relevant pages.

III. Proposed Work

A. What Is Google Dorks?

Google Dorks is that kind of a person who gives secure, sensitive, and particular information from all the internet sources. We can use Google Dorks for faster, better, and powerful search. If your website has some vulnerability than Google provides all information to the hacker using this dork. The concept of Google dorks create by Johnny long 2002 and uncovered the vulnerable system and disclosed sensitive information. Google hacking techniques were the focus of a book released by Johnny Long in 2005, called Google Hacking for Penetration Testers, Volume 1 Don't Underestimate power of Google dorks most powerful search in world.

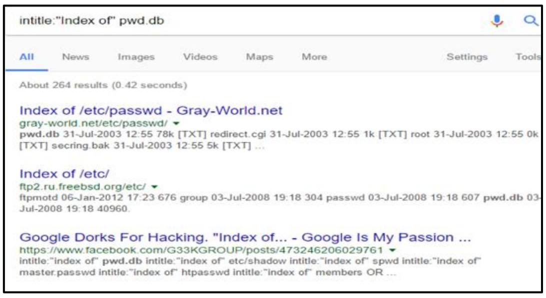
B. Types of Vulnerability Google Dorks Can Revel

- Footholds
- Web server detection
- Files contains username
- Sensitive directory
- Vulnerable files
- Files contain password
- Vulnerable server
- Sensitive online shopping info
- Error message
- Juicy files
- Login portal
- Various online device
- Network files or more

C. List of dorks with example

In this section we are going to test total 7 dorks and do analysis the results retrieved from Google search engine among thousands of dorks. Here by using "intitle" we retrieves the "page title", by using "filetype" we retrieves the particular file from the website, by using "cache" we retrieves particular cache of the website, by using "intext" we retrieves "website text", by using "link" we retrieves links of particular website that are connected with other website, by using "phonebook" we retrieves the phone number of particular person, by using "related" we retrieves the website which are related with each other [2], by using "site" we can able to search for particular site or domain

1) In title Dork



intitle:"Index of" pwd.db

About 264 results (0.42 seconds)

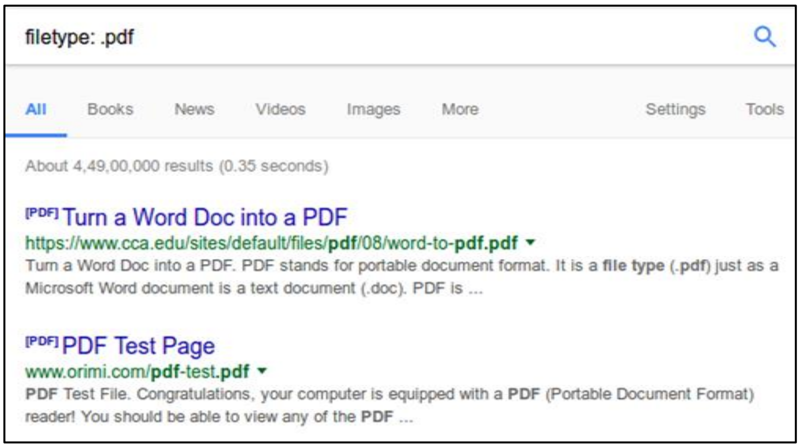
Index of /etc/passwd - Gray-World.net
gray-world.net/etc/passwd/
pwd.db 31-Jul-2003 12:55 78k [TXT] redirect.cgi 31-Jul-2003 12:55 1k [TXT] root 31-Jul-2003 12:55 0k [TXT] secring.bak 31-Jul-2003 12:55 5k [TXT] ...

Index of /etc/
ftp2.ru.freebsd.org/etc/
ftpmotd 06-Jan-2012 17:23 676 group 03-Jul-2008 19:18 304 passwd 03-Jul-2008 19:18 607 pwd.db 03-Jul-2008 19:18 40960.

Google Dorks For Hacking. "Index of... - Google Is My Passion ...
https://www.facebook.com/G33KGROUP/posts/473246206029761
intitle:"index of" pwd.db intitle:"index of" etc/shadow intitle:"index of" spwd intitle:"index of" master.passwd intitle:"index of" htpasswd intitle:"index of" members OR ...

Fig. 3: Searching database password files with the help of "intitle dork"

2) File type Dork



filetype: .pdf

About 4,49,00,000 results (0.35 seconds)

[PDF] Turn a Word Doc into a PDF
https://www.cca.edu/sites/default/files/**pdf**/08/word-to-**pdf.pdf**
Turn a Word Doc into a PDF. PDF stands for portable document format. It is a **file type** (.**pdf**) just as a Microsoft Word document is a text document (.doc). PDF is ...

[PDF] PDF Test Page
www.orimi.com/**pdf-test.pdf**
PDF Test File. Congratulations, your computer is equipped with a **PDF** (Portable Document Format) reader! You should be able to view any of the **PDF** ...

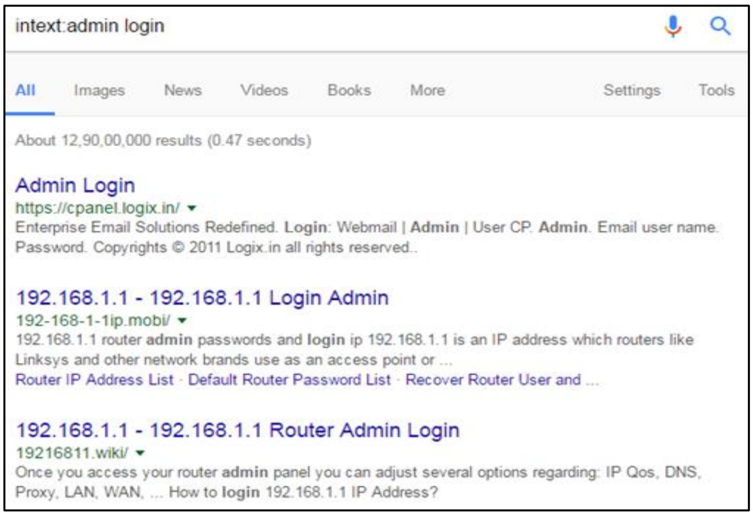
Fig. 4: Searching file formats using “filetype dork”

3) Cache Dork

– Cache: www.timesofindia.com

This keyword gives us particular cache of the website for specific date and time. If we want to read the content of particular website before the current time, it can be possible by entering “cache” keyword.

4) In text Dork



intext:admin login

About 12,90,00,000 results (0.47 seconds)

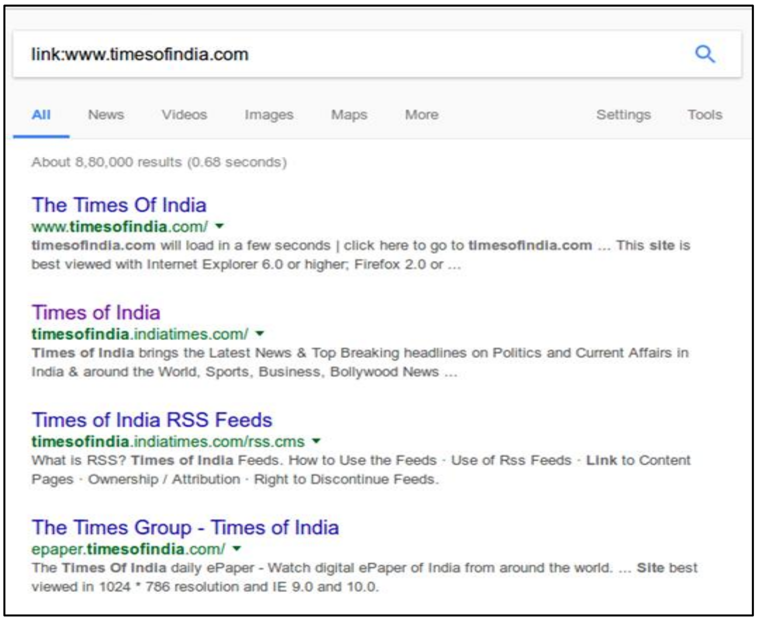
Admin Login
https://cpanel.logix.in/
Enterprise Email Solutions Redefined. **Login**: Webmail | **Admin** | User CP. **Admin**. Email user name. Password. Copyrights © 2011 Logix.in all rights reserved.

192.168.1.1 - 192.168.1.1 Login Admin
192-168-1-1ip.mobi/
192.168.1.1 router **admin** passwords and **login** ip 192.168.1.1 is an IP address which routers like Linksys and other network brands use as an access point or ...
Router IP Address List · Default Router Password List · Recover Router User and ...

192.168.1.1 - 192.168.1.1 Router Admin Login
19216811.wiki/
Once you access your router **admin** panel you can adjust several options regarding: IP Qos, DNS, Proxy, LAN, WAN, ... How to **login** 192.168.1.1 IP Address?

Fig. 5: Searching admin login page using “intext dork” [1]

5) Link Dork



link:www.timesofindia.com

About 8,80,000 results (0.68 seconds)

The Times Of India
www.**timesofindia**.com/
timesofindia.com will load in a few seconds | click here to go to **timesofindia.com** ... This **site** is best viewed with Internet Explorer 6.0 or higher; Firefox 2.0 or ...

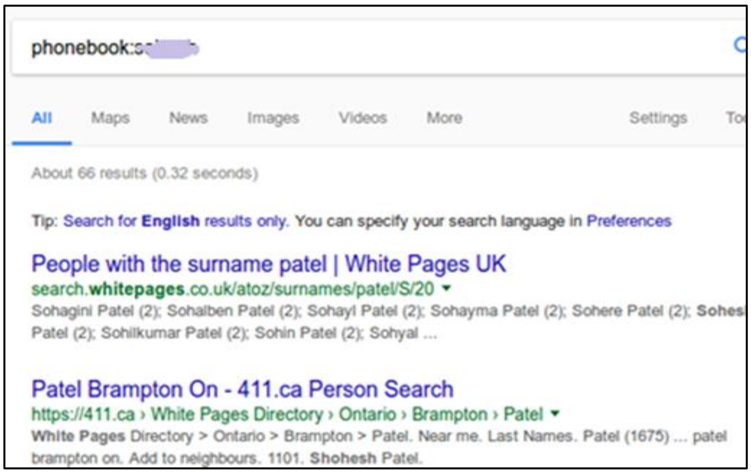
Times of India
timesofindia.indiatimes.com/
Times of India brings the Latest News & Top Breaking headlines on Politics and Current Affairs in India & around the World, Sports, Business, Bollywood News ...

Times of India RSS Feeds
timesofindia.indiatimes.com/rss.cms
What is RSS? **Times of India** Feeds. How to Use the Feeds · Use of Rss Feeds · **Link** to Content Pages · Ownership / Attribution · Right to Discontinue Feeds.

The Times Group - Times of India
epaper.**timesofindia**.com/
The **Times Of India** daily ePaper - Watch digital ePaper of India from around the world. ... **Site** best viewed in 1024 * 786 resolution and IE 9.0 and 10.0.

Fig. 6: Searching all the links for particular website using "link dork"

6) Phonebook Dork



phonebook:s[redacted]

About 66 results (0.32 seconds)

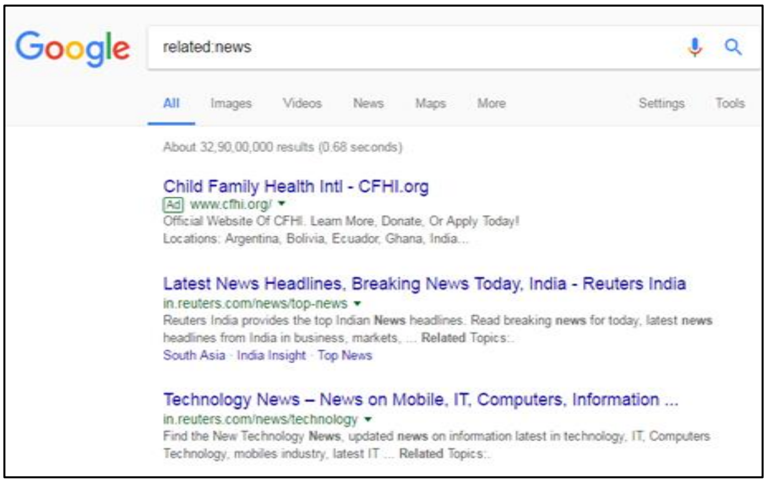
Tip: Search for **English** results only. You can specify your search language in Preferences

People with the surname patel | White Pages UK
search.**whitepages**.co.uk/atoz/surnames/patel/S/20
Sohagini Patel (2); Sohalben Patel (2); Sohayl Patel (2); Sohayma Patel (2); Sohere Patel (2); **Sohes**
Patel (2); Sohilkumar Patel (2); Sohin Patel (2); Sohyal ...

Patel Brampton On - 411.ca Person Search
https://411.ca › White Pages Directory › Ontario › Brampton › Patel
White Pages Directory > Ontario > Brampton > Patel. Near me. Last Names. Patel (1675) ... patel
brampton on. Add to neighbours. 1101. **Shohesh** Patel.

Fig. 7: Searching phone number of particular person using "phonebook dork"

7) Related Dork



related:news

About 32,90,00,000 results (0.68 seconds)

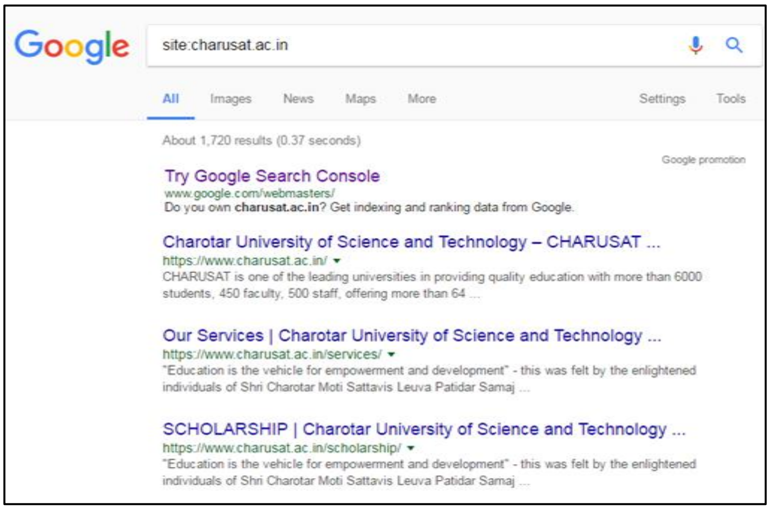
Child Family Health Intl - CFHI.org
Ad www.cfhi.org/
Official Website Of CFHI. Learn More, Donate, Or Apply Today!
Locations: Argentina, Bolivia, Ecuador, Ghana, India...

Latest News Headlines, Breaking News Today, India - Reuters India
in.reuters.com/news/top-news
Reuters India provides the top Indian **News** headlines. Read breaking **news** for today, latest **news**
headlines from India in business, markets, ... **Related** Topics:.
South Asia · India Insight · Top News

Technology News – News on Mobile, IT, Computers, Information ...
in.reuters.com/news/technology
Find the New Technology **News**, updated **news** on information latest in technology, IT, Computers
Technology, mobiles industry, latest IT ... **Related** Topics:.

Fig. 8: Related website list which contain news keyword

8) Site Dork



site:charusat.ac.in

About 1,720 results (0.37 seconds)

Google promotion

Try Google Search Console
www.google.com/webmasters/
Do you own **charusat.ac.in**? Get indexing and ranking data from Google.

Charotar University of Science and Technology – CHARUSAT ...
https://www.charusat.ac.in/
CHARUSAT is one of the leading universities in providing quality education with more than 6000
students, 450 faculty, 500 staff, offering more than 64 ...

Our Services | Charotar University of Science and Technology ...
https://www.charusat.ac.in/services/
"Education is the vehicle for empowerment and development" - this was felt by the enlightened
individuals of Shri Charotar Moti Sattavis Leuva Patidar Samaj ...

SCHOLARSHIP | Charotar University of Science and Technology ...
https://www.charusat.ac.in/scholarship/
"Education is the vehicle for empowerment and development" - this was felt by the enlightened
individuals of Shri Charotar Moti Sattavis Leuva Patidar Samaj ...

Fig. 9: List of all site which are include with "charusat.ac.in" using "site dork”

D. Real Time Implementation

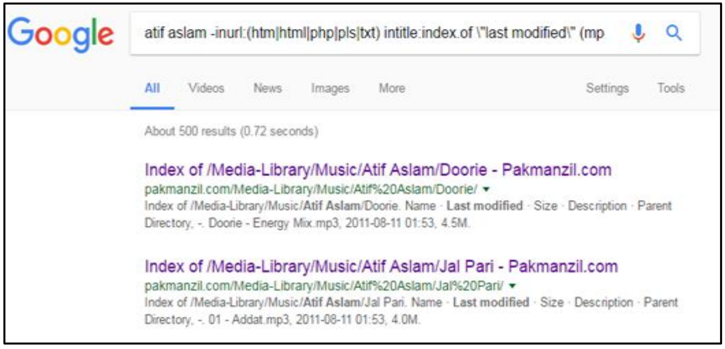
As we have seen individual dork examples in previous section for retrieving documents, images, personal files, phone numbers, website admin related links etc., now we can combine these dorks and create some queries and we get talismanic results.

Here below we executed some queries by applying some combinations of dorks:

1) Multimedia File Query

multimedia name -inurl:(htm|html|php|pls|txt) intitle:index.of \"last modified\" (mp4|wma|aac|avi)"

In this query the multimedia name can be like movie name, music artist name etc. This query “inurl dork” searches html, htm, php, pls, text content and “intitle dork” searches for index for particular website that contains mp4, wma, aac, avi files. User can also modify as per their requirements.



atif aslam -inurl:(htm|html|php|pls|txt) intitle:index.of \"last modified\" (mp

About 500 results (0.72 seconds)

Index of /Media-Library/Music/Atif Aslam/Doorie - Pakmanzil.com
pakmanzil.com/Media-Library/Music/Atif%20Aslam/Doorie/
Index of /Media-Library/Music/Atif Aslam/Doorie. Name · Last modified · Size · Description · Parent Directory, -. Doorie - Energy Mix.mp3, 2011-08-11 01:53, 4.5M.

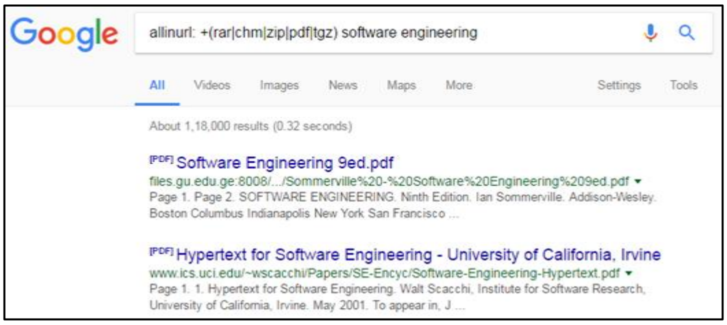
Index of /Media-Library/Music/Atif Aslam/Jal Pari - Pakmanzil.com
pakmanzil.com/Media-Library/Music/Atif%20Aslam/Jal%20Pari/
Index of /Media-Library/Music/Atif Aslam/Jal Pari. Name · Last modified · Size · Description · Parent Directory, -. 01 - Addat.mp3, 2011-08-11 01:53, 4.0M.

Fig. 10: Multimedia file query

2) Find Book Using This Query

– allinurl: +(rar|chm|zip|pdf|tgz) Book name

By using this query we can find books from parent directory of particular website. In above query “allinurl dork” can search in parent directory that modified last time which gives us file in rar, chm, zip, pdf and tgz format. By modifying this query user can get file in any format.



allinurl: +(rar|chm|zip|pdf|tgz) software engineering

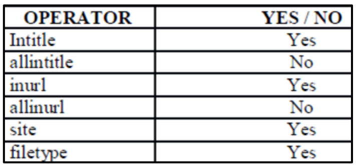
About 1,18,000 results (0.32 seconds)

[PDF] Software Engineering 9ed.pdf
files.gu.edu.ge:8008/.../Sommerville%20-%20Software%20Engineering%209ed.pdf
Page 1. Page 2. SOFTWARE ENGINEERING. Ninth Edition. Ian Sommerville. Addison-Wesley. Boston Columbus Indianapolis New York San Francisco ...

[PDF] Hypertext for Software Engineering - University of California, Irvine
www.ics.uci.edu/~wscacchi/Papers/SE-Encyc/Software-Engineering-Hypertext.pdf
Page 1. 1. Hypertext for Software Engineering. Walt Scacchi, Institute for Software Research, University of California, Irvine. May 2001. To appear in, J ...

Fig. 11: Find Book using This Query

E. List of dorks that can be mixed or not with other operator



OPERATOR	YES / NO
Intitle	Yes
allintitle	No
inurl	Yes
allinurl	No
site	Yes
filetype	Yes

Fig. 12: Dorks that mixed with other operators [7]

IV. CONCLUSION

We have already outlined some of the dorks and the combination of the dorks which are used to download secret or private documents from the index of the website and by using this we get talismanic results. Google dorks have their own pros and cons, end user able to hack a website and web developer cannot protect their information. The hackers can able to get database related information which is very serious issue.

REFERENCES

[1] Priyanka VK. “Detection of SQL Injection Attack and Various Prevention Strategies”. International Journal of Engineering and Advanced Technology (IJEAT) vol. 2013, pp. 457-60.

[2] Wilhoit, Kyle. "Who’s Really Attacking Your ICS Equipment?." Trend Micro (2013).

[3] Lancor, Lisa, and Robert Workman. "Using Google hacking to enhance defence strategies." In ACM SIGCSE Bulletin, vol. 39, no. 1, pp. 491-495. ACM, 2007.

[4] Invernizzi, Luca, and Paolo Milani Comparetti. "Evilseed: A guided approach to finding malicious web pages." In Security and Privacy (SP), 2012 IEEE Symposium on, pp. 428-442. IEEE, 2012.

[5] Mansfield-Devine, Steve. "Google hacking 101." Network Security 2009, no. 3 (2009): 4-6.

[6] Zhang, Jialong, Jayant Notani, and Guofei Gu. "Characterizing Google hacking: a first large-scale quantitative study." In International Conference on Security and Privacy in Communication Systems, pp. 602-622. Springer International Publishing, 2014.

[7] Vaibhav Sharma, Dr. Yashpal Singh International Journal of Enhanced Research in Management & Computer Applications, Vol. 5 Issue 6, June-2016.