INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 4, Issue 4)
Available online at: www.ijariit.com

Threats involved with internet advertisements and attack on botnet network

Prakash Kumavat
kumawat730@gmail.com
Veermata Jijabai Technological Institute, Mumbai, Maharashtra

Nikhil B. Khandare
nikcoep@gmail.com
Veermata Jijabai Technological Institute, Mumbai, Maharashtra

ABSTRACT

After the research on online advertisement industries, the global committees have identified various threats and risks to consumer's privacy and security which are hidden by the consumer. Various malicious software (malware) attacks take place through online advertisements without any click or interaction by user with advertisements contents. The scope of this research is to identify such threats and provide ideas to counter attack on botnet network to prevent privacy and financial loss.

Keywords: *Internet Advertisement, Botnet, C&C Server, IRC Server, Domain Generation algorithm, Steganography, Active attack, Passive attack, Ad Network, Trojan*

1. INTRODUCTION

Consumers can be the major victim of malware attacks even if they don't participate in any mainstream activity other than just visiting the website. The architecture of online advertisement is too complex for consumers to understand the malware linked with the ads network and determine if the ad network or host website has prevented the ad from such threats. Although companies also suffer reputational and economical losses due to such attacks. Many of the times Ad-hosting websites are also not aware of what advertisements will be running on their websites.

2. CONSUMER EXPOSURE TO MALWARE

Two major concern which are as follows:

A. Malwares in online advertisements often does not require any mouse click or interaction for user.

B. Malwares are usually placed on most popular sites, most popular contents and even on most reputable brands.

3. SYSTEM MODEL OF ONLINE ADVERTISEMENTS

Considering the whole process as an individual system we can try to find out the limitations in it. It consists of various components (process steps) which are as follows:

a. User sends request for a webpage.
b. Publisher sends webpage contents.
c. User is directed to webpage and Ad.
d. Ad Server sends a script.
e. User's browser executes the script and request for the Ad.
f. Ad Server receives request and send exact matching Ad.
g. User's browser receives Ad and shows it to user. [1]

3.1 Advertisers

Advertisers are the companies who are interested to have their product or services advertisements on online websites to collaborate more business.

3.2 Ad Network

They are the organizations who facilitates ads on appropriate websites. They have specialized ad servers who are responsible for managing ads on websites.

3.3 Data Collector or Data Brokers

These people collect user's information to sell it to Ad Networks to optimize and improve advertisement experience for user.

3.4 Internet Service Provider(ISP)

ISP provides internet services to user and Ad Networks. It plays very important role in communication through internet.

3.5 Publisher

These target websites which are used as a platform for advertisements.

3.6 Botnets

Botnets are collection of software bots which automatically executes various harmful scripts and responsible for exploiting web browser vulnerability. They are consisting of virus, Trojan horse and worms. They also steal user's private information and generate sources or loop holes in the user's system. Many times, these bots edit the link of the ads and make user redirect to a different webpage where he supposed to be. A bot Master controls all botnets remotely and using botnets for ad fraud is becoming very popular. Bots creating fake versions of other websites for phishing, a technique called "domain spoofing". [1][3]

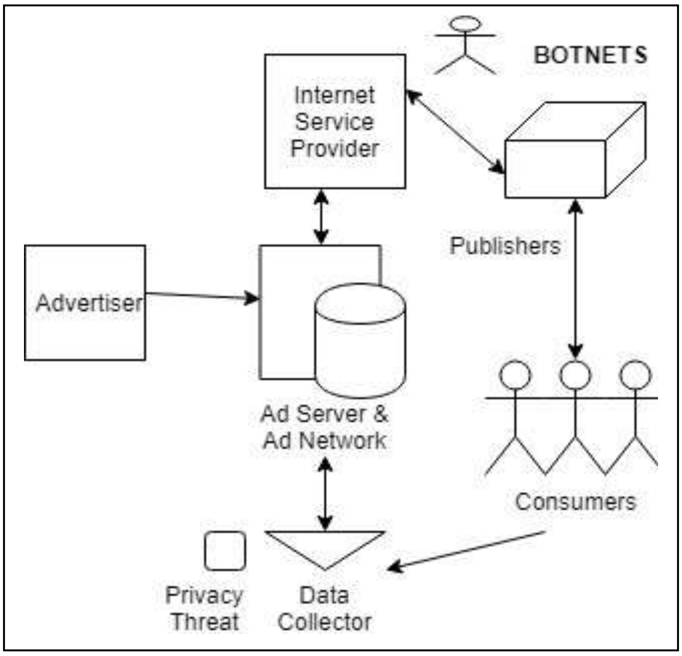

Fig. 1: Internet advertisement system model

3.7 Data Collectors or Brokers

These are people who collects data about the user's personal information, browsing history and interests in variety of products and categories from various sources for selling it to another third party (online sellers, ad networks, publishers etc.) without the concern of consumer. The aim of data collectors and brokers are to provide information to such companies to compile data about users and then target online advertisements on individual. There is lack of transparency about the data collected from such brokers or collectors for making online market convenient. It hits the law of privacy and securities because the data are collected without the consumer's concern.

3.8 Internet Service Providers(ISP)

The common role which we all know about ISP's are to provide internet access. However, recently ISP have started taking some additional tasks. They need to track and gather detailed records about the user's activities and provide these details to law enforcement organizations. With the increase in internet uses and internet as a platform or medium for business and advertisements, the risk and threats have also been increased. Many cyber criminals and cyber terrorist are using this platform for their vulnerable activities. Hence, ISP's responsibilities also have been increased to provide a safe channel and safe environment to the users along with the access to internet. According to new laws, ISP's should detect malware infected machines of their subscribers and take needed actions to avoid problems. Initiative from ISP's can actual make botnets life complex and restrict them for entering user's machine. However, ISP's required enough funding for such initiative. [1]

3.9 Threats and Ad Frauds

Due to increase in revenue of internet advertisement market, many threats are also attracted towards this. And because of such Ad Frauds, internet advertisement market is losing huge revenue. According to a report from CNBC news Nov 2017, the online advertisement scam has been exposed, that could be costing businesses, primarily in the US, almost $1.3 million a day. [1]

4. COUNTER MEASURES

4.1 One simple way to improve the security is to serve content and ads over HTTPS instead of HTTP. Because of the poor implementation of certification based authentication HTTP is not as secured as HTTPS is.

4.2 To protect revenue from stealers, cooperate and collaborate with ISP's and eliminate major risks(Botnets). They can also provide funding or economical incentives to ISP's so that they can fight back to botnets and provide protection to internet advertisement channel. Since ISP's have high privileges to fight back with Botnets.

4.3 Analysis of cookies to prevent the leak of personal information and helping user to prevent their data from attackers

5. STEGANOGRAPHY

Steganography is the art and science of hiding information behind objects. The objects could be the representative of any digital entity such as text document, image, audio, video files etc. The Steganography is ancient technique used to hide information or object behind another legitimate object. The Steganography is being used widely to hide objects behind the advertisement contents. Consumer or audience are unaware of the threat behind the legitimate objects. The hidden objects can be harmful enough for consumer to give financial loss or computational damage or life threatening (because of cyber terrorism). User is unaware of content which is being downloaded along with the content he/she is wanting to download. [2]

For example: Sometimes when user wants to download any file and click on download button then some other file is downloaded instead of that specific file. Many of the times such files are .exe file which prompt user to install it for fast downloading or unlimited access to more content which user is interested in. Since user trust the content provider, he/she install the application and then try to download the contents but here behind that application there is some malicious files associated with it which are triggered once user install the application and many times there is no action required by user. Sometimes the malicious content is associated with the legitimate file and user were not aware of it. Later user face various challenges and problems. [1]

The Steganography can be classified in different categories which are as follows:

5.1 Language Based Steganography: It is the process of hiding information within objects in some non-obvious way. This includes making non-obvious changes in the language or content to achieve the goals of steganography. This can be further categorized into semagrams and open codes.

5.2 Semagrams: It uses symbols and signs to hide information behind objects.

Visual Semgrams: A visual semagrams uses commonly used or innocent or legitimate objects to hide information to transfer messages.

5.3 Text Semagram: It modiefies the appearance of the carrier text.

5.4 Open Codes: It is the process of hiding messages in legitimate carrier messages that are not obvious. The carrier messages are called overt communication while the messages hidden behind the carrier messages are called covert communication. This is subcategorized into Jargon codes and Covered cipher.

5.5 Jargon codes: It uses a kind of language or communication terminology that is understandable only by specific group of people and not by others. It also uses pre-arranged phrases to convey a meaning.

5.6 Covered cipher: It hides the message openly in the carrier message so that it can be recovered by anyone who knows the secret for how it was concealed.

5.7 Grille cipher: A grille cipher employs a template that is used to cover the carrier message; the words that appear in the openings of the template are the hidden message.

5.8 Null cipher: A null cipher hides the message according to some prearranged set of rules such follow a specific number letter or read only 3rd letter of the word.

5.9 Technical Steganography: Technical Steganography uses scientific methodology to hide the message content.

Technical Steganography follows 3 major protocols which are as follows:

i. **Pure Steganography:** It entirely depends upon the secrecy of steganography algorithm. It does not include any external key for protection.

ii. **Single Secret Key Steganography:** It uses the same key for embedding the data into objects and extracting it.

iii. **Public Steganography:** In this the sender uses individual private key to embed the data into the objects and the recipient can extract the actual data hidden in the object by public key. This case ensures the legitimacy of data that it came from the right source. To protect the confidentiality the method can also be followed vice versa.[2]

Cyber-terrorist groups using steganography to bypass the security layers. Data hiding by the members of terrorist organizations is revealed on many occasions, but for sure it can be said that the number of cases where the data transmission covered using steganography methods is not registered by security services is much larger. Terrorist organization usually finds the object that are very commonly transferred and usually large in size such as audio, video files which is a difficult to scan and determine the hidden content by just looking at size. In previous studies and criminal cases, it has been revealed that criminal and hackers keep their data behind pornographic contents and found sending secret files by combining it into songs or movie files to avoid being capture.

6. STEGANALYSIS

Steganalysis is the study of detecting messages hidden using steganography; this is analogous to cryptanalysis applied to cryptography. Perhaps, the objects can be extracted or either destroyed to prevent losses. No matter how advanced the steganography process is, still after the completion of process it produces “Stego-Objects” which are somehow always different from the original content. They somehow make changes in the actual architecture of the original content. The main aim of Steganalysis is to identify the pattern differences to identify the objects. [2]

Steganalysis can be conducted using three main techniques. Visual Detection, Statistical analysis and Structural analysis.

i. Visual Technique

Visual analysis includes observation of file or objects with open naked eyes or using computer assistance to find the hidden content behind the visible objects or file. When I say computer assistance it means using computer advance software to represent the image in bit-planes which helps in analyzing file in more depth. [2]

ii. Statistical Analysis

Statistical Analysis helps in detecting tiny alternations in file’s statistical behavior due to incorporating Steganography. It includes various analysis such as Histogram analysis. This is more advance technique but it is more complex and time consuming as well.[2]

iii. Structural Analysis

This one is based on the analysis of the actual content or characteristics of the stego-objects. In this different kind of comparisons is done with stego-objects and the original objects by considering some factors such as file size, checksum, difference in date/time, content modifying etc. [2]

.

7. BOTNETS: THREATS TO ADVERTISEMENT NETWORK

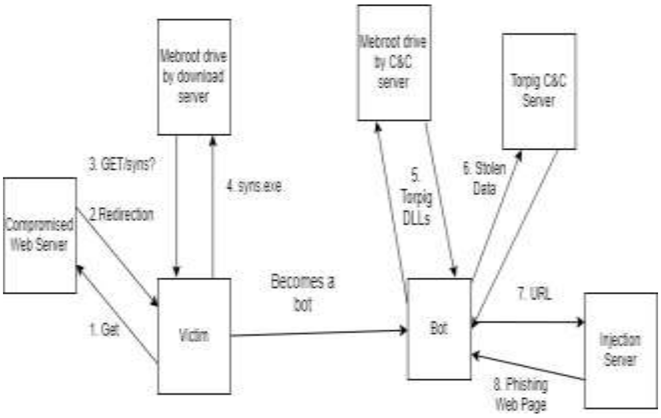

Fig. 2: Botnet attack on victim’s computer system

Botnets are one of the most harmful components in a communication network. Since they cover very large number of hosts and node groups to perform various types of attacks. They target stable systems and get the control over to make them bots for executing operations. Such bots can be controlled and managed from far distance.

Botnets uses advanced techniques to capture and stead data from various victims over the internet. It can cause major financial and privacy loss to the victim. Botnets relies on very complex network infrastructure.

Victim’s system consider malware as their Mebroot rootkit, which controls the whole system by changing the system’s MBR (Master Boot Record). The Attackers targets weak systems to make them botnets to create multiple hosts for communication and making their network stronger. Attackers target advertisement content and modify them from legitimate content to harmful scripts which make browser request for JavaScript and execute it on the victim’s browser. This script takes actions against the browser’s components such as ActiveX controls, browser plugins and add-ons. By this, attackers get the control over the browser activity and data it holds. Such as login credentials, cookies, payment information, request-response from server etc. [4]

Mebroot does not contain any harmful content or perform any malicious activity by itself. But it has a capability to activate other modules or DLL’s. The malicious contents are kept on System32 directory so that it can be used when user restart or reboot the machine without contacting the online server. The rootkit stores data that's required to survive reboots in physical sectors instead of files. This means that the data, including the real payload, is not visible or in any way accessible to normal applications. Therefore, the rootkit does not have to hook the normal set of interfaces to keep them hidden. The MBR is the rootkit's launch point. Therefore, it doesn't need to make any registry changes or to modify any existing startup executables in order to launch itself. This means that the only hooks it needs to make are used to hide and protect the modified MBR. [6]

The Mebroot contacts server for sending and receiving the information. The communication occurs on HTTP networks with a suitable custom algorithm. The Mebroot get the malicious content from the communicating server and injects them into various application such as mail clients, browsers, FTP clients etc.

The botnets observe the communication and data transferred through those infected programs. By this it can have access to important information such as bank details, login credentials. Botnets gets access to such information easily which can cause major loss to the victim. The communication channel uses HTTP which is protected by simple XORing with 8-byte key and base 64 encoding. This security mechanism has already been broken by security researchers in 2008 and automation tool for decrypting this which are available on the source from www.secureworks.com/ research/tools/untorpig/. Don Jackson’s Untorpig is one of them.

Here the botnet which we are talking about is also known as Torpig or Sinowal or Mebroot or Arserin. The main purpose of this botnet is to steal sensitive information such as Bank account and credit card information. The discussion focusses on the experiment discussed by the reference research paper.

8. BOTNETS COMMUNICATION

Botnets must maintain communication with their servers. To identify and detect their server over the internet, they use IP Address or DNS Names. Servers uses advance techniques to protect themselves from getting detected by anti-malware tools security mechanisms by IP fast-flux techniques. In this mechanism the botnets connect to one domain name which is mapped with multiple IP address. Hence, the server gets the capability to change its IP Addresses frequently. However, IP fast flux uses one domain name, hence if the domain name is failed or blocked, the botnets won’t able to connect to their server. Hence to resolve these issue botnets uses Domain flux mechanism which uses DGA (Domain Generation Algorithm) to maintain the list of domains. The domain resolves to an IP address to connect with the corresponding server. This algorithm can also be used to connect with driven by download server and ad Network etc. In DGA algorithm it generates periodic domain name it uses for a specific period or until the domain name is not blocked by security mechanism. [6]

8.1 Daily Domain Generation Algorithm

The below Domain Generation algorithm uses random key generation to generate domain and later uses it for specific period like a week or for a day. It simply adds the TLD (Top-Level-Domain) at the end of domain name. [5]

The domain name generation algorithm generates DNS for server in a sequence which is predictable by attacker and malware. Hence, botnets can simply select some selective domain names and loop over to reach the destination server. Once the domain is generated simply add the TLD(Top-Level-Domain) to it and one DNS sample is ready. [5]

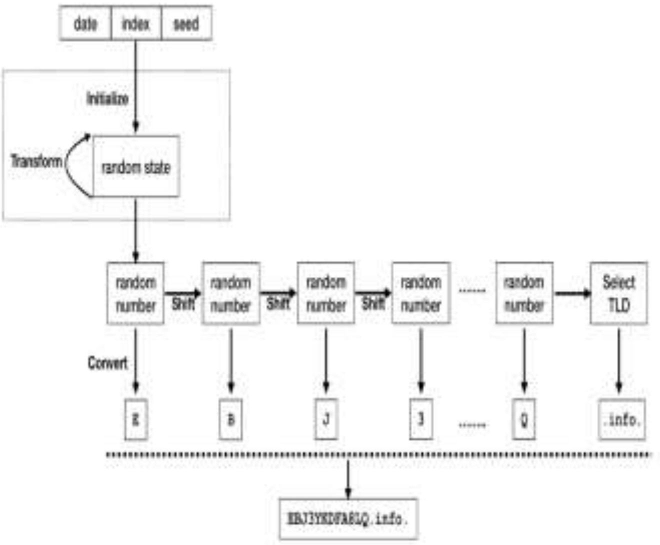

Fig. 3: Architecture of Domain Generation Algorithm

8.2 Controlling the Botnet

The major aspect to take the control over botnets is to understand how they communicate with their botnet master or C&C server. To detect or identify the botnet we need to keep tract on their activity. The botnets perform various activities which are malicious in nature. The behavior of botnets includes botnet-based click fraud, scanning of file and data frequently and the application or program not being registered under any anti-virus or anti-malware or search application, botnet-based spamming and botnet-based DoS attacks etc. The botnet uses IP-Addresses or DNS Names to communicate in peer-to-peer network. There are some other techniques such as IP-fast flux which is being used by some botmasters. IP-fast flux provides domain name which corresponds to multiple IP-Address Set which are used until the next round of domain name generation algorithm does not provide any other destination address for communication. This cycle is repeated in specified or dynamic time intervals.

The botmaster is main lead to manage all botnets. It must make sure of few things which are as follows.

1. Bots should not generate too many future domains.
2. The domains which are registered should not be registered by any other server network.
3. It should notify the bots about changing the domain name or domain blocked.

In the referred research it was found that many research and experiment has been done to find the botnets communication server (through which botnets receive all notification and to whom they send the information collected from victim) and hijack it or change the path of communication to become the server to receive the information. But there are some limitations in the proposed phenomena which is discussed further.

1. The botnets keep updating the domain name after specific period hence difficult to track the future domains.
2. Botnets communication with server can only be routed until our server knows the future domains and then control over botnets.
3. The domain captured by our server can be blocked by the domain registrar due to malicious content activity and abuse report.
4. When malicious server doesn’t receive any information from the botnets then they may switch to any other victim which is not known to us.

5. The data received is confidential and important hence, leak of that data may cause major problems to victims.
6. Holding this private information of victim also causes to break the privacy protocol or privacy law irrespective of our intension.
7. The sever can be closed by security mechanism at any point if the malicious activities are caught or observed.
8. The data is travelling through less protected channel, hence open for other attackers to get some benefit out of it before it reaches to our server.

8.3 IRC Communication

Botnets uses IRC communication as channel to communicate to the botnet master or botnet server for transfer of information or instructions. The standard IRC protocol uses TCP 6667 port for communication. Hence this channel port can be used to observe and detect the botnet instruction commands. The attackers keep making changes in their communication address or IP address. [8]

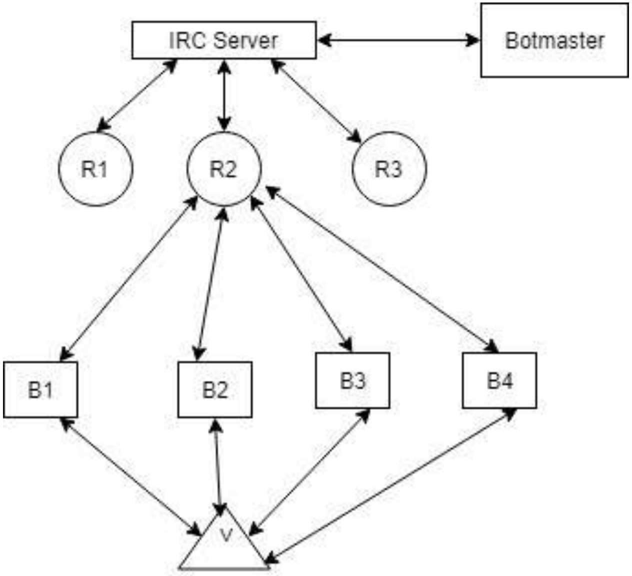

Fig. 4: Abstract diagram represents communication between bots and botmaster through IRC server

In the above diagram the V represents the Victim, B1, B2, B3 and B4 are botnet systems and R1, R2 and R3 represents router which transmit the communication packets. Routers play major role in communication between the botnets and botmaster. The instructions or command from botmaster to botnets are transferred through IP packets and may have fake IP header. But they travel routed through the routers. If we record these routers activity then the source path can be identified. The technique is described in more detail in reference given. [8]

9. THE PROPOSED SOLUTION

The proposed solution is to focus on botnets activities and taking control over it. The idea is to take the control over these botnets, understand their behavior and observe their activity. Once we get enough information regarding the botnets, we can hijack the activities of botnets and then make our own botnets with configuration we observed. The domain details which we have captured can be used here to communicate to the botnet server. Since the botnet we have setup will communicate with botnets server.

The data sent by our botnet will be not genuine. It is being created using the permutation and combination standards or the original data is being modified in such a way that it will still maintain the standards of data. It replaces the content with similar matching random generated content. Some standards of data are discussed below.

1. Password should contain at least 1 uppercase, 1 lowercase, 1 numerical and 1 special digit character etc.
2. If the data is related with credit/debit card details, then it will follow generation of random digits for each number, which will help from various threats and risk.

The generated content can be sent to botnet server in two ways:

Passive Attack

The file is sent directly to botnet server. Since the data is generated with random data generation algorithm , the original content is modified by replacing some content by adding or deleting the data in it. Hence, we won't bother about the protection of data. Now the botnet server will filter the data to extract useful information such as login credentials, bank details etc. So, when attacker use that data to get access into any other's system by using login credentials or bank details to steal the money by programs or manually, they will fail. Since the data is not genuine but they don't know about it so when they try multiple attempts on any website or system, the system's security mechanism will be triggered due to number of attempts and security protocol breached, hence the intentions of attacker fails.

Active Attack

Another approach is to add tracking programs and trojan horse programs with the file by using the steganography techniques to hide the actual object behind the legitimate objects. Attacker is unware of the data being received and when the attacker tries to use those information received at their end, the stego-objects trigger the content associated with the file and find their way in other malicious programs or application in their system. Now this object can server multiple purpose such as we can identify the attacker or track the source of attacker or can sit idle to observe all the activities being performed on the other legitimate data receive by botnet server from other sources and make it unreliable or adding random content in it. By this it can prevent the data of multiple victims.

The attackers don't have much security mechanism at their end since they have lot of such malicious content which can be trapped into the anti-malware programs. Attacker's server generally uses less secure channel for communication which is unable to detect malicious content in files. For E.g. the C&C server or botnet server uses HTTP protocol for communication rather than HTTPS which is more secure and reliable protocol for communication. Hence, it opens a path to exploit their system by active attack technique.

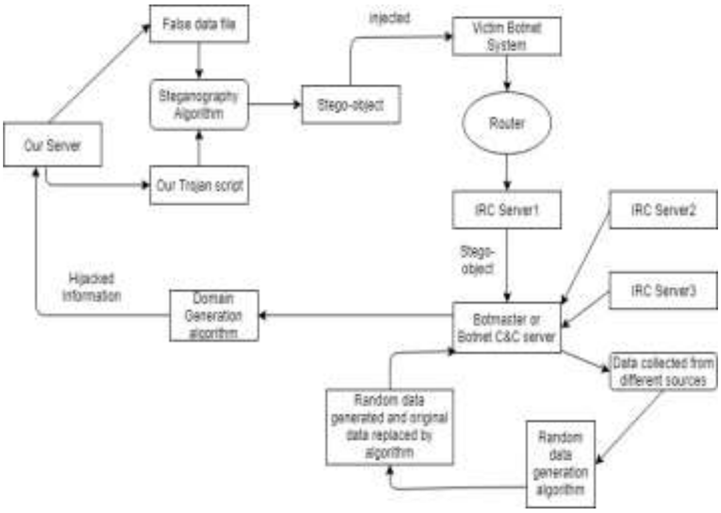
False data file
injected
Victim Botnet System
Steganography Algorithm
Stego-object
Our Server
Router
Our Trojan script
IRC Server1
IRC Server2
Stego-object
IRC Server3
Hijacked information
Domain Generation algorithm
Botmaster or Botnet C&C server
Data collected from different sources
Random data generated and original data replaced by algorithm
Random data generation algorithm

Fig. 5: Active attack approch on botnet network

In the above diagram, the approach is discussed where we can inject our script in false data file using steganography algorithm to generate stego-object. These stego-object is then injected to victim's system which is infected by botnet system. Since the file maintain confidential information standard, so it will be scanned and captured by botnet and send through router to IRC Server. Botmaster or Botnet C&C server gets the stego-object but are not aware of the content hidden behind the file. Since the channel of communication used by botnet server is not secure, hence it makes easy for the stego-objects to reach destination server without being detected. Once it reaches the destination server, it triggers its malicious script to track the activities and analyze data. It sends back the important information related to activities to our server and uses random data generation algorithm to replace the content collected from different sources. The intension is to protect the data collected from different sources to be used for un-ethical purpose. The following technique or idea can be customized according to the choice of algorithms and accuracy. The results can vary depending upon the efficiency of experiment.

10. CONCLUSION

The paper discusses about the better understanding of botnet-based attacks and threats involved with internet or online advertisements. Botnets have capability to exploit the network without the need of any human assistance. The much-needed solution for such problems is that it cannot be limited to prevention but must be extended to attack back. I have proposed several prevention techniques and techniques or ideas to attack back on botnets sources. The changing behavior of botnets brings a new shape to the problem, researchers has to anticipate for future techniques or strategies of botnet makers and design effective response technique. The research discusses about ideas which can be implemented in future to take down the intensions of botnets. The discussion can be extended to more techniques and implementation of the above discussed ideas that are under the scope of future discussion.

11. REFERENCES

[1] Written Testimony of Craig D. Spiezle before the Senate Committee on Homeland Security & Government Affairs Permanent Subcommittee on Investigations, May 15, 2014; see also Caitlin Condon, StopBadware steps down as leader of the Ads Integrity Alliance, STOP BADWARE BLOG (Jan. 20, 2014), J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68-73.

[2] M. Bogdanoski, A. Risteski and S. Pejoski, "Steganalysis — A way forward against cyber terrorism," 2012 20th Telecommunications Forum (TELFOR), Belgrade, 2012, pp. 681-684.

[3] T. Holz et al., "Measuring and Detecting Fast-Flux Service Networks," Proc. 16th Network and Distributed System Security Symp., Internet Soc., 2008;

[4] B. Stone-Gross, M. Cova, B. Gilbert, R. Kemmerer, C. Kruegel and G. Vigna, "Analysis of a Botnet Takeover," in *IEEE Security & Privacy*, vol. 9, no. 1, pp. 64-72, Jan.-Feb. 2011.doi: 10.1109/MSP.2010.144

[5] A Death Match of Domain Generation Algorithm. https://blogs.akamai.com/2018/01/a-death-match-of-domain-generation-algorithms.html

[6] Barford P., Yegneswaran V. (2007) An Inside Look at Botnets. In: Christodorescu M., Jha S., Maughan D., Song D., Wang C. (eds) Malware Detection. Advances in Information Security, vol 27. Springer, Boston, MA

[7] Hindawi Publishing Corporation.EURASIP Journal on Wireless Communications and Networking. Volume 2009, Article ID 692654,11pages.doi:10.1155/2009/69

[8] Z. Chi and Z. Zhao, "Detecting and Blocking Malicious Traffic Caused by IRC Protocol Based Botnets," 2007 IFIP International Conference on Network and Parallel Computing Workshops (NPC 2007), Liaoning, 2007, pp. 485-489.doi: 10.1109/NPC.2007.77

[9] Botnet Detection Countering the Largest Security Threat Library of Congress Control Number:
ISBN-13: 978-0-387-68766-7 eISBN-13: 978-0-387-68768-1